Boundary Effects on the Intelligibility and Predictability of Spatial Systems

Ref 086

Hoon-Tae Park
University College London, The Bartlett School of Graduate Studies, London, United Kingdom
hoon-tae.park@ucl.ac.uk

Keywords
boundary; intelligibility; understandability; predictability; intensity

Abstract

The boundary determination problem is one of the most fundamental and difficult issues in spatial analysis. This is also true of space syntax where syntactic values are meaningful only with reference to the system boundary we have chosen to draw for analysis.

This paper investigates how our choice of a boundary can affect our understanding of a system's spatial structure and functionality. It will be argued that to draw a boundary is not only a priori condition for syntactic analysis but also a posteriori product which should be carefully designed based on our understanding of a syntactic structure it induces.

The four standard test areas in London - Barnsbury, Calthorpe, South Kensington and Brompton - are re-employed for this purpose. We apply various boundary conditions, both in topological and metric spaces, to these areas, and compare how sensitively the predictability of human movement patterns by global syntactic values change under each condition. Although many researches in the space syntax community have untilised the same datasets already, few views the boundary condition as an independent variable.

We then recall one of the key propositions in space syntax that the degree to which movement rate is predictable from a spatial system is a function of its intelligibility. This proposition is tested against the empirical findings, which will lead to a critical discussion as to how and why the conventional measure of intelligibility falls short of holding the proposition to be valid.

Our aim in this regard is to develop a novel morphological index that may work, in place of the measure of intelligibility, more effectively for the proposition. That is, we will find that the new measure, called 'understandability' in this paper, is a better indicator of the spatial system's predictability. Based on this it is suggested that such kind of measure has the potential to serve as a useful guide and reference for designers as well as researchers at their earlier stage of practices.

1. Introduction: The Boundary Determination Problem

The boundary determination problem is one of the most deep and difficult issues in spatial analysis. This is also true of space syntax where syntactic values are meaningful only with reference to a system boundary we choose for analysis. To draw a boundary in this respect is an unavoidable starting point of analysis, but the important question is: How is such a starting point itself given? How can syntactic values, which are deemed objective and thus undisputable, be grounded on a subjective choice that is open to debate?

In space syntax, we can identify at least three different practices carried out to cope with the problem:

The first is to apply *local analysis*, in which only neighbouring spaces encountered within some fixed radius from each space are taken into account (Penn et al. 1998). But one may argue that this restrictive method, although hugely successful in many empirical studies, only translates the boundary determination problem into the problem of determining a radius (see, for instance, Joutsiniemi, 2005). Moreover, Park (2007) has pointed out that local properties, such as local integration, have a strong tendency to reduce to connectivity due to the presence of scaling laws. This means that a genuine distinction can be made only in terms of the immediate local (e.g. connectivity) and the global (e.g. global integration), resulting the status of local analysis to be rather unstable.

The second undertakes the problem more directly by posing a series of heuristic rules to follow in defining a *contextual area*. It states that, if we are interested in studying a particular location, the contextual area must i) have our location of interest in the geometric centre of that contextual system; ii) approximate a circle as closely as possible; iii) be large enough to cover the 'catchment area of the catchment area' of whatever function we aim to study (Hillier et al. 1993). But it does not give an explicit answer to the question of how large the contextual area should be to be 'large enough.' Also, it is sometimes awkward to reconcile the concept of imposing a geometric circle in Euclidean space to study underlying topological structure.

The third may seem value neutral as it suggests to reserve one's decision by employing a *whole city map* wherever possible. But it can be demonstrated that network effects, so crucial in syntactic analysis, tend to dissipate as a system gets larger (see the main text for more details). If this is the case, that is, if the modification of parts cannot induce any significant differences at the level of a whole, how could it serve as a useful tool for design practices? It seems that a whole city map tends to produce a result that may be pleasing, through its simplicity and clarity, to the panoptic eyes of analysts, while not very instructive in the process of designing, for instance, a new pedestrian road at the local scale.

This paper does never intend to solve the problem completely. Rather it begins with the assumption that the subjective or better creative role of 'Cartographers' cannot be altogether dispelled from the process of drawing boundaries. From this the argument will follow that to draw a boundary is not only a priori condition for analysis but also a posteriori product which should be carefully designed based on our understanding of syntactic structure it induces. In other words, we need to see the boundary determination problem as a kind of *design problem*. To see the problem in this way and to present a common ground on which we can discuss it further will be the main aim of this paper.

2. Boundary Effects on Intelligibility and Predictability

Boundary effects can be considered in two aspects: the *spatial extensity* and *internal structure* of a bounded system. On the one hand, changing boundaries should result in systems of different size (i.e. the number of axial lines). These results are trivial though. We can never expect to resolve the boundary determination problem in terms of size alone. On the other, changing boundaries may induce systems of different internal structure and consequently different levels of predictability. These differences can be revealed by means of various syntactic measures or their association with functional properties. Our first task is then to distinguish these latter structural effects clearly from mere size effects.

For this purpose we have re-employed the four standard test areas in London - Barnsbury, Calthorpe, South Kensington and Brompton[1]. We apply various boundary conditions, in both metric and topological space, to these areas, and compare how sensitively the intelligibility (i.e. the correlation coefficient between connectivity and integration) and the predictability of movement from global syntactic values (i.e. r-squared value) change under each condition. Figure 1 shows the example of Barnsbury area, in which metric radius is set to vary from 1000 to 5000m and topological radius from 2 to 5 step depths. Although many researches in the space syntax community have utilised the same datasets already, few view the boundary condition as an independent variable in this way. Note also that all the measures used in the analysis have been Box-Cox transformed to ensure the normality condition[2].

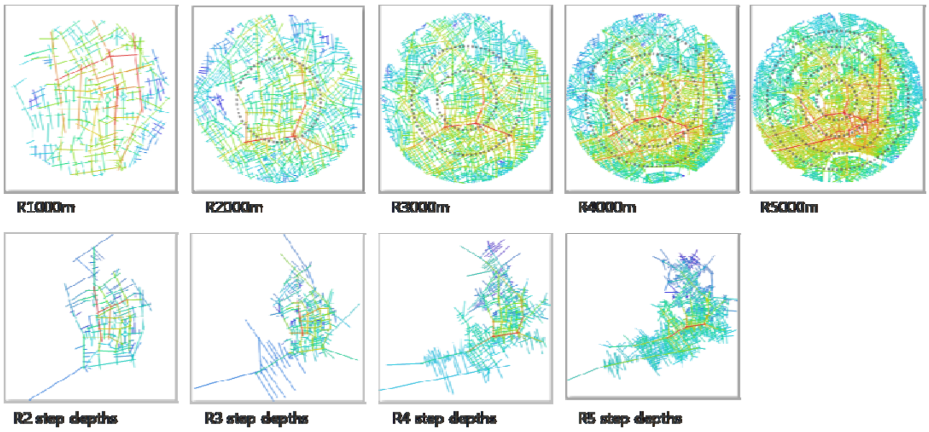

Figure 1
The changing boundary conditions in both metric and topological spaces – Barnsbury example

Areas	Radius	N	Intlligb (integ)	Intlligb (choice)	Pedestrian			Vehicle		
					Conn	ITG	ITS	Conn	ITG	ITS
Barnsbury	1 km	282	**0.672**	**0.824**	0.732	**0.559**	**0.667**	0.695	0.615	0.643
	2 km	1040	0.517	0.780	0.733	0.532	0.606	0.695	**0.629**	**0.672**
	3 km	2315	0.503	0.760	0.733	0.456	0.557	0.695	0.571	0.626
	4 km	4126	0.418	0.746	0.733	0.434	0.515	0.695	0.548	0.582
	5 km	6057	0.381	0.730	0.733	0.429	0.497	0.695	0.540	0.560
Calthorpe	1 km	368	**0.716**	**0.792**	0.529	0.584	0.763	0.465	0.577	0.592
	2 km	1343	0.541	0.755	0.529	0.759	**0.783**	0.465	**0.592**	**0.597**
	3 km	2825	0.513	0.749	0.529	**0.779**	0.761	0.465	0.546	0.568
	4 km	4441	0.471	0.741	0.529	0.765	0.753	0.465	0.517	0.543
	5 km	6355	0.433	0.729	0.529	0.753	0.756	0.465	0.480	0.533
South Kensington	1 km	375	**0.638**	0.750	0.493	**0.532**	**0.538**	0.773	0.700	0.729
	2 km	1019	0.586	**0.767**	0.468	0.442	0.509	0.766	**0.727**	0.743
	3 km	2075	0.474	0.733	0.468	0.399	0.444	0.766	0.693	**0.757**
	4 km	3384	0.414	0.721	0.468	0.372	0.388	0.766	0.631	0.663
	5 km	5117	0.392	0.713	0.468	0.354	0.348	0.766	0.648	0.651
Brompton	1 km	374	**0.694**	**0.806**	0.307	0.441	0.366	0.472	**0.582**	0.503
	2 km	1006	0.510	0.741	0.343	**0.556**	0.537	0.475	0.548	0.576
	3 km	2007	0.486	0.740	0.343	0.541	**0.588**	0.475	0.502	**0.588**
	4 km	3594	0.447	0.723	0.343	0.536	0.557	0.475	0.496	0.531
	5 km	5417	0.433	0.712	0.343	0.528	0.542	0.475	0.485	0.508
Average	1 km	349.8	**0.680**	**0.793**	0.515	0.529	0.584	0.601	0.619	0.617
	2 km	1102.0	0.539	0.761	0.518	**0.572**	**0.609**	0.600	**0.624**	**0.647**
	3 km	2305.5	0.494	0.746	0.518	0.543	0.587	0.600	0.578	0.635
	4 km	3886.3	0.438	0.733	0.518	0.527	0.553	0.600	0.548	0.580
	5 km	5736.5	0.410	0.721	0.518	0.516	0.536	0.600	0.538	0.563

Table 1
Metric boundary effects
(Note: N = number of axial lines; Conn = connectivity; ITG = global integration; ITS = global intensity; Intlligb (integ) = the correlation between connectivity and integration; Intlligb (choice) = the correlation between connectivity and choice)

Table 1 summarises the results. First we notice that the predictability of both pedestrian and vehicular movements from connectivity is not affected by the boundary conditions, while that from global measures - integration and intensity - are. Although the differences in r-squared value may seem insignificant numerically, an important point to note here is that there seems to be a certain radius for each area at which the predictability reaches the maximum. This may indicate the presence of size-independent boundary effects on internal structure, given that the size of each area cannot but increase monotonously with radius.

To examine the differences in internal structure directly, we have measured the intelligibility for each area. The reason why we are interested in intelligibility is that it is a whole-map value that may reflect the structural simplicity of a system - 'The degree to which the whole can be read from the parts.' This leads to the idea that, in the case where two boundaries are given, we should be able to choose one that induces a more intelligible structure. Just as we draw a boundary following a circle in the geometric space, there is no a priori reason to prefer a biased and complicated structure in a topological sense. If a system has the simplest and thus perfectly intelligible structure, its predictability of movement patterns should not differ, in principle, whether it be based on connectivity or integration.

However, as shown in Table 1, the most critical problem with the intelligibility measure is that it seems entirely dependent on the size of an area only, independently of its internal structure and predictability. That is, intelligibility only decreases as an area gets larger, while predictability does not. Figueiredo and Amorim (2007) have suggested that intelligibility can be also conceived in terms of the correlation between connectivity and choice. But, although much less sensitive, this choice-based intelligibility can be still regressed to the size of an area. In the next section, we will give more detailed analysis to the idea of intelligibility and why it fails to capture the differences of internal structure.

Table 2 then shows the results in the case of topological boundaries and we notice that the arguments can go exactly the same way as in the case of metric boundaries. In average, the maximum value of predictability can be found at the topological radius of 3 step depths, which roughly corresponds to the metric radius of 2000m. This evidences once again that changing boundaries may induce different internal structures, whether we apply metric or topological boundaries. These differences can be reflected indirectly or post-dictively by the different levels of predictability, but cannot be detected directly or predictively by the different level of intelligibility.

Areas	Radius	N	Intlligb (integ)	Intlligb (choice)	Pedestrian			Vehicle		
					Conn	ITG	ITS	Conn	ITG	ITS
Barnsbury	rad 2	188	0.708	**0.861**	0.719	0.546	**0.661**	0.689	0.470	0.635
	rad 3	323	**0.712**	0.812	0.733	**0.585**	0.623	0.695	0.533	**0.669**
	rad 4	646	0.559	0.801	0.733	0.557	0.557	0.695	**0.562**	0.653
	rad 5	1313	0.488	0.756	0.733	0.508	0.457	0.695	0.538	0.596
Calthorpe	rad 2	225	**0.727**	**0.837**	0.541	0.541	0.638	0.435	0.493	0.495
	rad 3	532	0.543	0.834	0.617	0.736	**0.748**	0.495	**0.583**	**0.586**
	rad 4	1237	0.476	0.789	0.528	0.758	0.642	0.465	0.515	0.479
	rad 5	2404	0.472	0.781	0.528	**0.763**	0.627	0.465	0.482	0.411
South Kensington	rad 2	321	**0.567**	**0.834**	0.377	0.460	0.428	0.744	0.689	0.698
	rad 3	748	0.525	0.792	0.523	**0.477**	**0.446**	0.818	**0.755**	**0.714**
	rad 4	1414	0.494	0.763	0.468	0.383	0.413	0.766	0.687	0.703
	rad 5	2221	0.489	0.761	0.468	0.320	0.411	0.766	0.625	**0.714**
Brompton	rad 2	428	**0.623**	**0.836**	0.347	0.448	0.532	0.483	0.511	0.549
	rad 3	746	0.570	0.771	0.343	0.548	0.547	0.475	**0.562**	0.559
	rad 4	1268	0.458	0.741	0.343	**0.550**	0.556	0.475	0.536	0.559
	rad 5	2051	0.441	0.760	0.343	0.517	**0.580**	0.475	0.491	**0.566**
Average	rad 2	290.5	**0.656**	**0.842**	0.496	0.499	0.565	0.588	0.540	0.594
	rad 3	587.3	0.587	0.802	0.555	**0.586**	**0.591**	0.620	**0.608**	**0.632**
	rad 4	1141.3	0.497	0.773	0.519	0.562	0.542	0.600	0.575	0.598
	rad 5	1997.3	0.472	0.764	0.519	0.527	0.519	0.600	0.534	0.572

Table 2
Topological boundary effects

3. The relationship between intelligibility and predictability

The relationship between intelligibility and predictability is important as it has been suggested as a key theoretical component of space syntax. In the earlier development stage of space syntax, Hillier et al. (1987) have hypothesised that *the degree to which movement is predictable from integration is a function of the intelligibility of each area* (see also Hillier et al. 1993). They have also suggested that "the technique of optimising the reference area to give spatial parameters which best predict the pattern of movement may permit the identification of something like natural areas and sub-areas in a continuous fabric." Combining this post-dictive technique with the predictive hypothesis, we can naturally draw a working proposition that a boundary can be drawn in such a way as to *maximise the intelligibility of an area it encloses*. In order to reason out the mechanism underlying this proposition, we need to introduce what they call "the model of measurement" (Figure 2).

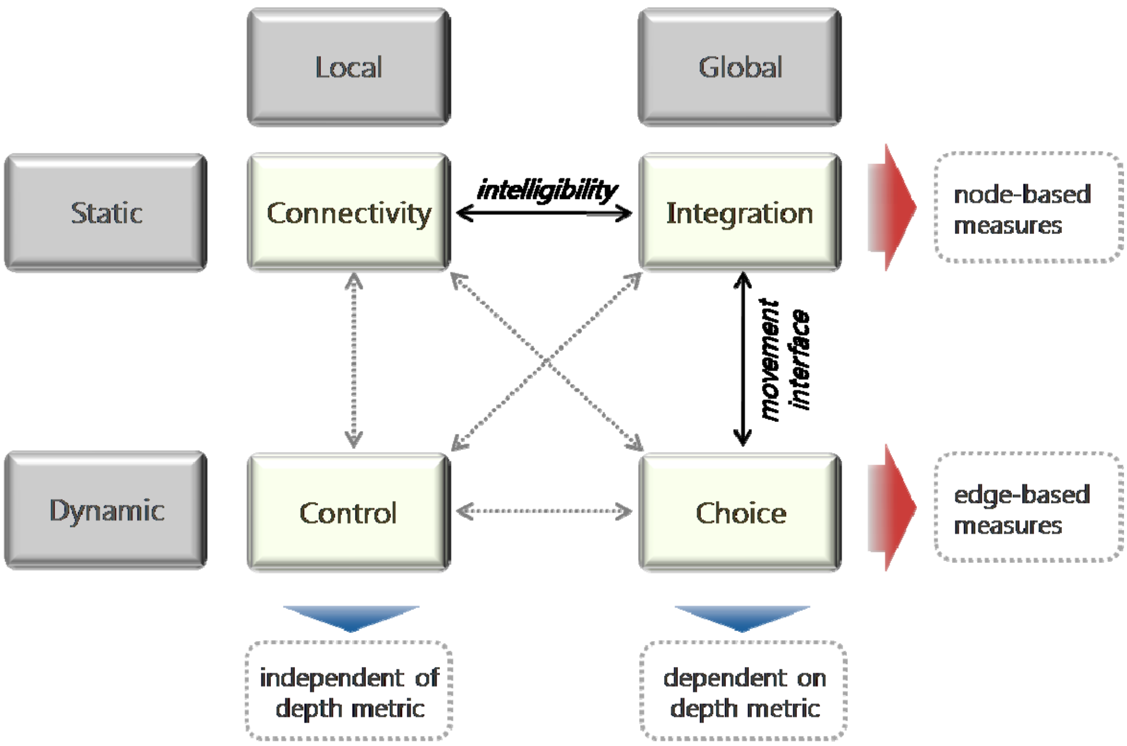

Figure 2
Model of Measurement

Firstly, they distinguish static measures as those about a fixed spatial system from *dynamic* measures as those about a set of mobile individuals superimposed on the system. The global static measure (integration) corresponds then to the movement pattern of 'strangers' who rely entirely on the reading of a spatial layout, and therefore, the commonly observed good correlation between integration and movement is taken to explain that there are always a good number of relative strangers in the system. By contrast, the global dynamic measure (choice) is a better indicator of movement for 'inhabitants' with better spatial knowledge than for strangers. So, as they reason, the higher correlation between integration and choice may indicate a stronger "movement interface" between inhabitants and strangers.

Yet at the same time, if the structure of a spatial system is so simple and thus highly intelligible, this will nullify the distinction between inhabitants and strangers simply because there is not much spatial knowledge to be retrieved. The movement patterns of both inhabitants and strangers would converge to each other consequently, creating a strong movement interface represented by a high correlation between integration and choice. Put together, these conjectures lead to a conclusion that the predictability of movement from integration (or choice) will become more powerful, through a strong movement interface, as a system becomes more intelligible. In the ideal case, the three measures of connectivity, integration and choice would be seen as one and the same thing.

The properties of intelligibility can be aptly described in terms of the notion of *assortativity by connectivity*. We recall that a graph is called assortative (or dissortative), if high-connectivity nodes tend to be associated with nodes with high (or low) connectivity; and neutral if there is no such tendency (Newman 2002; Pastor-Satorras et al. 2001). Assortativity by connectivity reflects well-defined core-periphery structure, in which integration, assortative by its spatial nature, decreases with connectivity from core to periphery (see Figure 3). By contrast, this correlation will break down in a dissortative structure which consists of the hierarchy of distinctive components. This means that the idea of intelligibility as the correlation between connectivity and integration is in fact equivalent to the idea of assortativity by connectivity.

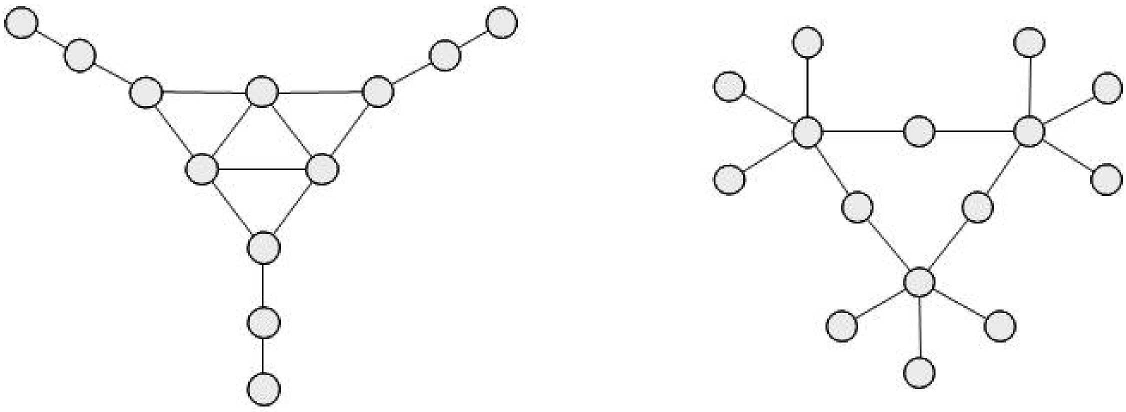

Figure 3
Simple graph examples of assortativity (left) and dissortativity (right) by connectivity

However, Goh et al. (2003) have shown that if a graph is dissortative by connectivity and thus has hierarchical structure, connectivity and choice should behave in a very similar way, so that higher connectivity will guarantee higher choice. This time, then, the idea of intelligibility as the correlation between connectivity and choice, which Figueiredo and Amorim have seen more appropriate for urban street networks, becomes equivalent to the idea of dissortativity by connectivity. This relationship will break down for assortative graphs, since the neighbours of high-connectivity nodes also have high connectivity and thus they do not necessarily lie in the shortest path between a certain pair of nodes at the same time. And this will greatly decrease the choice of those high-connectivity nodes in favour of the formation of integration-core.

So it seems now clear that we are eventually left with two partial notions of intelligibility set in opposite directions: 'assortative-intelligibility' and 'dissortative-intelligibility.' In this respect, to say that a graph is intelligible may not mean much unless the sense in which it is so is specified at the same time. But in any sense, what is more important here is to note that there is no way to actualise both notions of intelligibility at the same time. In other words, the high level of movement interface would be impossible for a highly intelligible graph. If they appear somehow correlated, it is because they both are decidedly dependent on size. And if it is the case that they both decrease with size, the only reason we can think of is that a graph they refer to does not remain assortative or dissortative, but becomes neutral as it grows[3].

What does that a graph is neutral by connectivity explain about its structure? In the first place, it literally means the absence of both assortativity and dissortativity (or the absence of intelligibility in any sense). But it may just reflect the incapability of the formal measures, not the deficiency of the 'real' intelligibility of systems we observe. For instance, consider a city whose 'foreground network' of arterial streets is assortative, while whose 'background network' of tributary streets is dissortative

(Hillier et al. 2007). This dual form is in fact very common and tends to help us to navigate in an intelligible way; but cannot be expressed sufficiently if averaged through such a crude measure as assortativity or intelligibility. If the property (rather than measure) of assortativity itself is concerned, this homogeneity at the whole level seems highly artificial so that it is quite unlikely to find it for naturally grown cities. Xulvi-Brunet and Sokolov (2005)'s experiment, for instance, has shown that the mean distance travelled by randomly moving agents can be minimised when a system as a whole is neutral by connectivity. This may imply that neutrality results from a very complex way in which a city evolves to optimise 'natural movements' in it; that there is perhaps not less, but more in the absence of assortativity. How can we then express this 'more' in a positive way?

4. A New Measure of Understandability

Based on the review in the last section, we now turn to develop a novel graph-level measure that may overcome the shortcomings of the measure of intelligibility. We assume that this measure should be, in order to achieve this objective, defined i) independently of the idea of assortativity; ii) independently of size effects and iii) independently of connectivity at the local level. The third condition is necessary as any graph-level measure based on connectivity tends to reduce to the measure of assortativity or its derivatives. As such, the new measure, which we will call *understandability*, diverges from the original concept of intelligibility that the whole can be read from the parts.

The fundamental concept underlying the graph-level measure of understandability is *structural similarity* among distance degree sequences (or j-graphs). And it will be implemented through the node-level measure of intensity. Intensity is a composite measure of mean depth and informational entropy of a distance degree sequence, defined as H/λ where H and λ denote entropy and mean depth respectively (see Park 2005 for more details). It is a global static measure that is more sensitive to the 'shape' of distance degree sequences than integration, and thanks to this property, gives a systematically better correlation with observed movement patterns than integration (see Table 1 again). For instance, the more positively a distance degree sequence is skewed, so that most movements occur within a close range to the reference node, the higher intensity it will have (Figure 4 Left). From this point of view, we may understand that intensity measures a kind of *movement efficiency*, to which integration is comparably unresponsive.

Yet, intensity also has a tendency to reduce to integration *under a certain structural condition*: if H in the numerator part is expressed as some concave function of λ, intensity can be transformed into a simple integration-like function of λ. This indicates in essence that a structure captured by intensity is very simple and thus can be described by much compressed information. But, what does that entropy is expressible as a function of mean depth mean? It means that a set of distance degree sequences comprising a graph belong to the identical family of distribution and thus they are all structurally similar. For a typical example, a Boltzmann distribution with mean value parameter λ, delimiting the upper bound of entropy maximisation, has the entropy function given by $Hmax = \lambda \ln \lambda - (\lambda - 1) \ln (\lambda - 1)$ (Figure 4 Right)[4]. By comparison, the entropy function of a Poisson distribution with mean value parameter λ is known to have an approximate form of $H = 0.5 \ln \lambda + 0.5 \ln(2pe)$ for large λ (Evans et al. 1988). In both cases, intensity in which entropy is cancelled becomes a monotonously decreasing function of λ only, such that it cannot be differentiated from integration.

In this paper, we develop this idea further by introducing the measure of normalised intensity. Normalised intensity is simply a ratio of intensity to its possible maximum given λ, that is, $(H/\lambda)/(Hmax/\lambda) = H/Hmax$, which must lie in the range of 0 and 1. Here, $Hmax$ can be achieved when a distance degree sequence follows a Boltzmann distribution and it corresponds to an equilibrium state in which nodes are completely independent from each other (somewhat like ideal gas particles). As a system gets larger and larger, it can be shown that entropy tends indeed to increase toward its maximum, resulting in the normalised intensity value of 1. Yet, while entropy remains away from its maximum, there will be always a difference represented by $Hmax - H$, resulting in normalised intensity value less than 1. In information theory, this quantity is often conceptualised as the amount of redundant information, that is, information that can be developed further from the current state toward the equilibrium. Hence, we can say that entropy H measures the amount of information that has already been developed for the current state. Normalised intensity indicates just this kind of actualised information against the virtual background of redundancy.

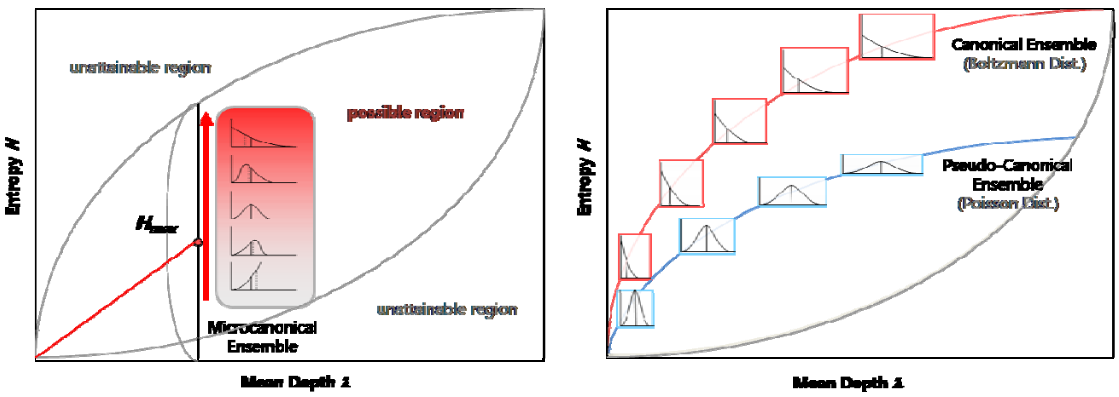

Figure 4

The phase space defined by mean depth λ (horizontal axis) and informational entropy H (vertical axis). Individual distance degree sequences can be represented by points in the phase space, while a whole graph by a set of points. Intensity is then simply the gradient of the line connecting the origin (i.e. the homogeneous sequence) and each point. The possible region delimits the upper and lower bounds of entropy, given mean depth.

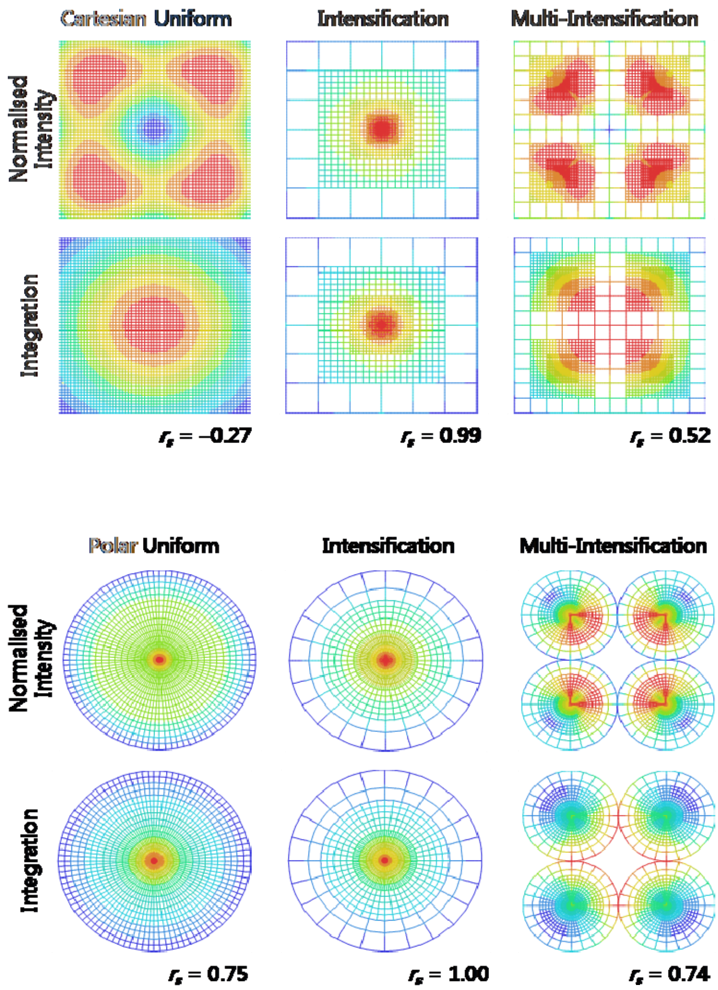

Figure 5

Hypothetical Examples of Cartesian and Polar Grids

We will then define the measure of understandability as *Spearman's rank correlation coefficient between normalised intensity and integration*. Because rank correlation is free from the linearity condition, it may work for any monotonic relationship between mean depth and entropy. For instance, if all distance degree sequences observed follow Poisson distribution, understandability will get the maximum value of 1. But if they follow Boltzmann distribution, understandability will be zero since normalised intensity becomes all uniform. In the case when entropy is described by a convex function of mean depth, the value of understandability will be even negative (a sort of anti-understandability) as normalised intensity increases with mean depth. We also note that a highly understandable structure may be also highly intelligible in the original sense, while understandability may vary independently of connectivity[5]. See the examples below, in which all axial lines are segmented to make connectivity more or less uniform.

The hypothetical examples above show that the maximum value of understandability can be found for a grid system with a single intensive centre. If a grid system consists of multiple centres with local intensification, we may increase its understandability by dividing it into separate systems with their own centres. On the other hand, a Cartesian uniform grid is anti-understandable with negative correlation. This is perhaps related to the fact that there is no objectively meaningful way to divide a Cartesian uniform grid as it will repeat itself indefinitely. We also note that the intensity-image of a Cartesian uniform grid, although it is global and based on topological depths, appear quite similar to that which can be found through the application of metric mean depth (Hillier et al. 2007). This all is possible because intensity in the first place is very sensitive to the local (but more than immediate local) condition of systems and thus capable of inducing a more concrete image than integration. This in turn implies that understandability may reflect our cognitive process in which an intensity-image (content) is abstracted into an integration-image (form).

5. An Application of Understandability

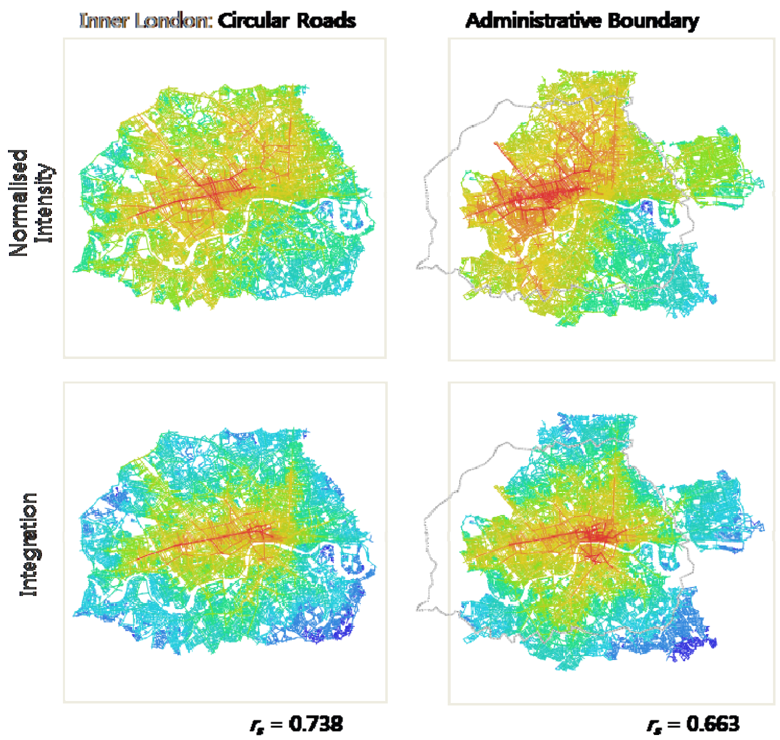

Figure 6

The understandability of the two whole axial maps of Inner London

First, we analyse the two whole axial map of Inner London (Figure 6). One is drawn roughly according to the North and South Circular Roads and the other is to the administrative boundary of Inner London (Source: the Office for National Statistics, 2001). The average values of normalised intensity are 0.81 (N = 17,320) and 0.83 (N = 19,655) respectively. A comparison of understandability suggests that the former (0.74) is based on a better choice of boundary than the latter (0.66). Indeed, the spatial distributions of normalised intensity and integration are not very different in the former, while, in the latter, integration core has been shifted to the east with the administrative boundary. Note also that the distribution of normalised intensity seems to be less affected by a choice of boundary than that of integration (e.g. the area around Oxford street remains to be the intensity core in both cases).

Table 3 below compares predictability for the four test areas embedded in those two whole maps of Inner London. For each area and each mode of traffic, it is the more understandable area within the Circular Roads that gives a better predictability of movement pattern from integration (indicated by red colour). But, we should ensure whether this is not due to size effects (see below), for not only understndability but also both intelligibility values are higher for this smaller area. Apart from this question, it is noteworthy that normalised intensity is the best predictor of movement in 11 out of 16 cases. This suggests that we had better rely on normalised intensity to get a more true-to-life global image when a whole city map is employed.

	N	Intlligb (integ)	Intlligb (choice)	Unders.	Areas	Pedestrian			Vehicle		
						ITG	ITS	N.ITS	ITG	ITS	N.ITS
Circular Roads	17320	0.297	0.703	0.738	Barnbury	0.451	**0.477**	0.385	0.492	**0.530**	0.459
					Calthorpe	0.694	**0.737**	0.730	0.448	**0.483**	0.475
					Sth.Ken.	0.239	0.284	**0.423**	0.531	0.581	**0.678**
					Brompton	0.468	0.506	**0.580**	0.431	0.468	**0.540**
Admin.	19655	0.257	0.695	0.663	Barnbury	0.434	0.459	**0.505**	0.505	0.522	**0.534**
					Calthorpe	0.716	**0.756**	0.700	0.357	0.416	**0.473**
					Sth.Ken	0.238	0.279	**0.411**	0.526	0.571	**0.675**
					Brompton	0.460	0.497	**0.574**	0.428	0.464	**0.545**

Table 3

Predictability of the test areas embedded in the concatenated areas

However, the predictability of movement from normalised intensity diminishes significantly if the test areas are separated to be embedded in their own contextual areas. Table 4 summarises average predictability for the four test areas as their contextual boundaries change both metrically and topologically. We find that the predictability from normalised intensity is worst in most cases. Note also that topological boundaries are no better than metric boundaries in predicting movement patterns (indicated by red colour), while there seems to be no significant difference for the predictability patterns of pedestrian and vehicular movements. For instance, if it is connectivity that gives the best predictability for pedestrian movement, then it is also true for vehicular movement. To sum up, the predictability of movement patterns can be improved generally by embedding the test areas in their own contextual areas, but the predictability pattern from individual syntactic variables becomes much untidier than when they are embedded in the whole maps.

The fact that the best predictor can be any syntactic variables but normalised intensity when individual contextual boundaries are applied suggests that the consistent relationship between understandability and predictability may be lost. To find out if this is the case, we examine the variation of the understandability of each test area as its contextual boundary changes and its relationship with the variation of predictability. Figure 7 illustrates the example of South Kensington area in the metric boundary condition, and Figure 8 summarises the results for all the other test areas.

Area	Boundary	Pedestrian				Vehicle			
		Conn	ITG	ITS	N.ITS	Conn	ITG	ITS	N.ITS
Barnsbury	Metric	**0.733**	0.482	0.568	0.549	**0.695**	0.581	0.617	0.478
Calthorpe		0.529	0.728	**0.763**	0.585	0.465	0.542	**0.567**	0.428
S Kengsinton		**0.473**	0.420	0.445	0.410	**0.768**	0.680	0.709	0.646
Brompton		0.336	**0.520**	0.518	0.292	0.474	0.522	**0.541**	0.324
Barnsbury	Topo	**0.729**	0.549	0.574	0.252	**0.693**	0.526	0.638	0.366
Calthorpe		0.553	**0.699**	0.664	0.340	0.465	**0.518**	0.493	0.243
S Kengsinton		**0.459**	0.410	0.425	0.271	**0.774**	0.689	0.707	0.429
Brompton		0.348	0.516	**0.554**	0.262	0.476	0.525	**0.558**	0.243

Table 4
Average predictability of the test areas embedded in their own contextual areas

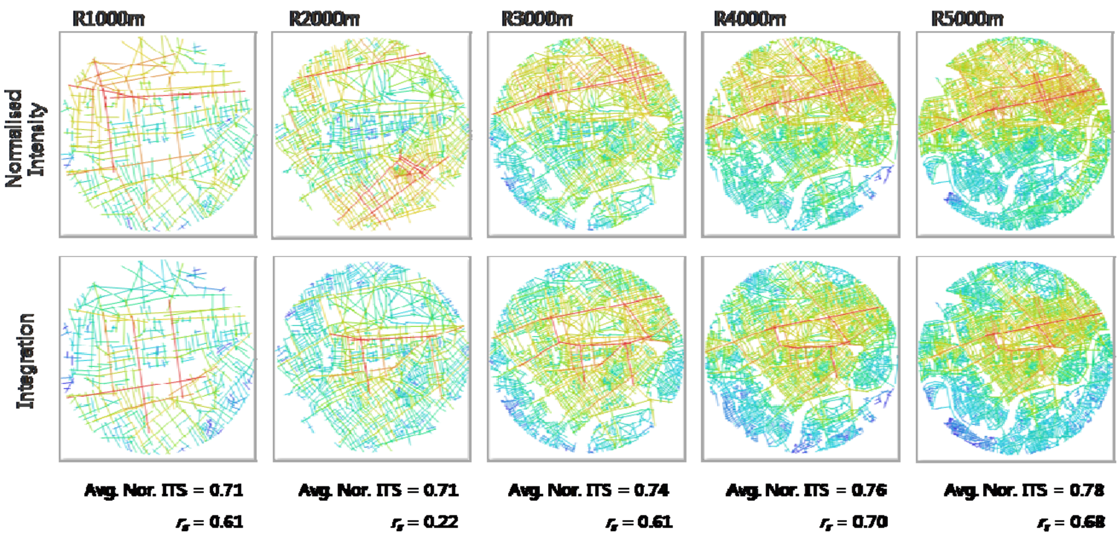

Figure 7
The variation of the understandability with metric boundaries – South Kensington area

Firstly, it can be noted that the average value of nomarlised intensity tends to increase with size and this is found to be the case for all the other test areas. It seems to suggest that, for an infinitely large system, nomarlised intensity will reach to its maximum value of 1 and all network effects will dissipate to make it completely not understandable. Second, the variation of understandability seems independent of size effects. Although there is some positive correlation in the case of topological boundaries, this is statistically insignificant considering the sample size. This helps to strengthen the results we have found for the case of whole axial maps. However, as we have expected, the relationship between understandability and predictability collapses completely when individual contextual boundaries are applied. It is even embarrassing to find negative correlations for some of the cases. Yet at the same time, we report that intelligibility can be no better alternative to understandability (r-squared value varies from 0.001 to 0.12). After all, the consistent relationship between syntactic structure and predictability still remains beyond the scope of our understanding.

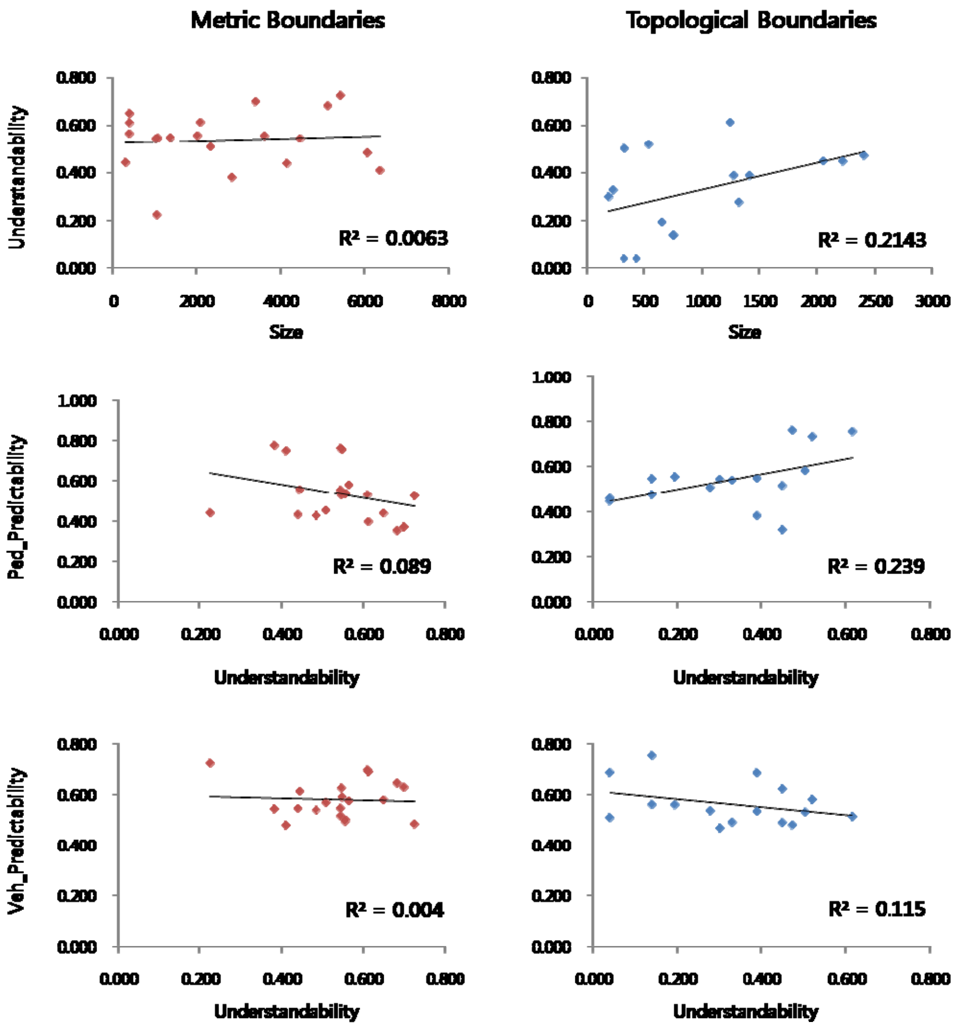

Figure 8

The size-independence of understandability (the upper two) and the relationship between understandability and the predictability of both pedestrian and vehicle movements from integration

6. Discussions

This paper has begun with the assumption that the boundary determination problem should be regarded as a kind of design problem, in that its possible resolution lies primarily in understanding better the internal structure of a system which a subjective choice of boundary induces. In a way to objectify the boundary drawing process it has introduced a new experimental measure of understandability in place of intelligibility and presented a working proposition that a boundary can be drawn in such a way as to maximise the understandability of a system it encloses. This proposition has been tested against the relationship of understandability and predictability from integration, with a view to verifying the classic hypothesis of space syntax that the degree to which movement is predictable depends on the structural simplicity of a system.

The measure of understandability has been defined as a correlation between the two global static measures of 'normalised intensity (the concrete)' and 'integration (the abstract)', thus diverging from the original scheme of 'the local' versus 'the global' in intelligibility. In the case of whole city maps, we can verify the proposition as the more understandable system gives the more predictable movement patterns from integration. But this breaks down altogether as we apply individual contextual areas to the test areas. In the former, it is normalised intensity that is consistently the best predictor of movement, and its strong correlation with integration may thus ensure the relationship between understandability and predictability. But in the latter, the best predictor of movement pattern can be any syntactic variable but normalised intensity, and thus, understandability becomes completely unrelated to predictability. We have also found that the application of topological boundaries does no better than that of metric boundaries in terms both of understandability and predictability.

There are many other graph-level indices than intelligibility or understandability, such as scale-freeness, assortativity, centralisation, hierarchicalisation and so on. But we do not understand very clearly their relationships yet. In general, many commonly used graph-level indices are notoriously ill-behaved (Anderson et al. 1999). But nevertheless, to have a better understanding of those indices is one of the main syntactic issues, and this is particularly so when we need to derive a whole-map value for some real applications. The boundary determination problem is one such an application and the measure of understandability is one of many possible alternatives that may serve the same purpose. It is the main concern of future researches to provide the measure with a more solid theoretical basis and to verify it against a larger set of observations.

Acknowledgements

This research was supported by a grant (07High Tech A01) from High tech Urban Development Program funded by Ministry of Land, Transportation and Maritime Affairs of Korean government.

Notes

1 The datasets are publicly available from the UCL Eprints (http://eprints.ucl.ac.uk/1232/).
2 The Box-Cox transformation is a procedure to normalise data and thus to improve the correlation between variables (see http://en.wikipedia.org/wiki/Power_transform).
3 Park (2007) has shown that 'normalised control' can be utilised to probe the presence of assortativity. For neutral structure, normalised control should not have any meaningful relationship with connectivity.
4 This entropy function can be approximated by ln λ + 1 for large λ, and as such has been derived by applying the entropy maximisation principle to a discrete infinite base with a fixed mean.
5 For Poisson distributions, connectivity is proportional to λ/eλ, which decreases monotonously with λ. This will induce not only perfect intelligibility but also perfect understandability.

References

Anderson B S, Butts C and Carley K. 1999. The interaction of size and density with graph-level indices, *Social Network* 21: 239-267.
Evans R J, Boersman J, Blachman N M and Jagers A A. 1988. The entropy of a Poisson distribution: problem 87-6, *SIAM Review* 30(2): 314-317.
Figueiredo L and Amorim L. 2007. Decoding the urban grid: or why cities are neither trees nor perfect grids, *Proceedings of the 6th International Symposium in Space Syntax*, Istanbul Technical University.
Goh K I, Oh E, Kahng B and Kim D. 2003. Betweenness centrality correlation in social networks, *Physical Review E.* 67(1): 017101.
Hillier B, Turner A, Yang T and Park H T. 2007. Metric and topo-geometric properties of urban street networks: some convergence, divergence, and new results, *Proceedings of the 6th International Symposium in Space Syntax*. Istanbul Technical University.
Hillier B, Burdett R, Peponis J and Penn A. 1987. Creating life: or, does architecture determine anything? *Architecture and Behaviour*, 3(3):233-250.
Hillier B, Penn A, Hanson J, Grajewski T, and Xu J. 1993. Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Planning and Design*. 20 (1): 29-66.
Joutsiniemi A. 2005. Generic accessibility challenges axial maps: case Helsinki, *Proceedings of the 5th International Symposium in Space Syntax*. Delf University of Technology.
Newman M E J. 2002. Assortative Mixing in Networks, *Physical Review Letter*. 89: 208701.
Park H T. 2005. Before integration: a critical review of integration measure in space syntax, *Proceedings of the 5th International Symposium in Space Syntax*. Delf University of Technology.

Park H T. 2007. The structural similarity of neighbourhoods in urban street networks: a case of London, *Proceedings of the 6th International Symposium in Space Syntax*. Istanbul Technical University.

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

Penn A, Hillier B, Banister D and Xu J. 1998. Configurational modelling of urban movement networks. *Environment and Planning B: Planning and Design*, 25:59-84.

Xulvi-Brunet R and Sokolov I. 2005. Changing Correlations in Networks: Assortativity and Dissortativity, *ACTA Physica Polonica B*, 36(5).