Small World Modeling for Complex Geographic Environments

Bin Jiang

Abstract. This paper aims to provide some insights into geographic environments based on our studies using various small world models. We model a geographic environment as a network of interacting objects - not only spaces, places and locations, but also vehicles and pedestrians acting on it. We demonstrate how geographic environments might be represented as a form of networks and be illustrated as small worlds. Furthermore we try to shed light on the implications of small world properties from various application perspectives.

1. Introduction

Conventional geometric-based networks have various limitations when modeling geographic environments. Essentially they are geometric representations, and thus interrelationships between locations or objects are not well defined and represented. Therefore it is difficult for the model to deal with issues such as dynamics, growth and evolution. Now researchers tend to view geographic environments as a complex network, and treat them as such with a range of features such as nonlinearity, interdependence and emergence. For instance, a topological based network representation is suggested as an alternative model for the study of the evolution of street networks (Jiang and Claramunt 2004). In modeling a built environment and terrain surface, the concept of a visibility graph is introduced to show the interrelationship of individual locations in terms of visual accessibility (Turner et al. 2001, Jiang and Claramunt 2002). These efforts represent a new wave of studies of geographic environments using a topological-based network view following the main stream of study on complex networks.

We model geographic environments as a network of interacting objects – not only spaces, places and locations, but also vehicles and pedestrians acting on it. This view goes beyond the conventional geometric-based network, in the sense of dealing with neighboring, adjacency and relationships, and it provides an alternative representation of geographic environments with respect to geographic modeling. For instance, agent-based and cellular automata modeling can be based on the topological based representation (O'Sullivan 2001). This topological model has been considered in a variety of disciplines in the study of complex networks involving the Internet, cells, scientific collaboration, and social networks, to mention a few examples (see Strogatz 2001 for an overview). The nature of geographic environments, as a complexity system, can be illustrated via the underlying networks. We can examine whether or not there is a hidden order behind a geographic environment, on the one hand, and study how global properties are emerged via the interaction of individuals modeled as the vertices of the underlying network, on the other. It helps understand that complexity systems are more than sum of individual components, and this extends the reductionism towards interrelations of the individual components. This has been a major development stream in the current study of complexity theory (Holland 1995).

Recent advances in the study of complex networks have been tremendous. The availability of various datasets of real-world networks and powerful computers has made possible a series of empirical studies since the seminal work by Watts and Strogatz (1998). The concept of small world networks has featured in many scientific journals and conferences, and has been becoming an increasingly interesting topic being investigated as an emerging science (Barabasi 2002, Watts 2003).

The small world network is a network with a kind of hidden order between regular and random networks in analogy with the small world phenomenon observed in social systems (Milgram 1967). The small world phenomenon states that in a large social system, e.g. a population of a country, the distance between any two randomly chosen persons, e.g. yourself and the president of your country, is just about six persons away, so called “six degrees of separation”. Whether the number really is six remains a matter for debate (Kleinfeld 2002), but most real world networks are indeed small enough, as evident in many real systems mentioned above. A real world system with which the underlying network demonstrates the small world phenomenon is likely to be an efficient and stable system in terms of the diffusion of a variety of phenomena ranging from information to terrorist networks and AIDS epidemics. Although originally observed as a property of social systems, the small world phenomenon has been examined in a variety of natural systems with which basic units are interconnected as a complicated network.

This paper aims to investigate how small world models can be used for modeling and understanding geographic environments from a structural point of view. We start with an introduction to small world networks, followed by Watts and Strogatz’s small world model and several other models that expand on the small world concept in various ways. We then examine how small world properties are demonstrated with geographic environments and further discuss the implications of the small world properties from various application perspectives, in particular linking to the issue of search efficiency with a geographic environment, and of designing a geographic environment with high search efficiency. This chapter concludes with a summary and outlook for future work.

2. Small world networks and models

2.1. Small world networks

To explain what a small world network is, let us consider a visibility graph in terms of how each point location is visible to every other within, for example, a downtown area. We impose in a grid fashion a set of point locations between buildings and examine how these point locations, represented as nodes, are visible to each other. If there is no obstacle between two point locations (nodes), then there would be a link between the two nodes. In this way, we would see that those visible locations from a given location are likely to be visible as well. In other words, the visibility structure is highly clustered. This is the first characteristic of small world network, and it is measured by the *clustering coefficient*. If visible locations from a given location are not visible to each other at all, the cluster coefficient equals 0. On the other hand, if visible locations from a given location are all visible to each other, then the clustering coefficient equals 1. The ratio of actual links over all possible links among a set of visible locations (for n visible locations, all possible links is $n(n-1)/2$) from a given location is defined as clustering coefficient. The average of all locations' clustering coefficient is that of an entire network.

There is another important property that characterizes a small world network. In most cases, two locations are not directly visible, but they may be visually accessible via another location. In this case, these two locations are visually separated by a distance of two. In some cases, two locations may be separated by a distance of three or four etc. The distance is actually the notion of *shortest path length*. The average of the shortest path length between all pairs of locations for a visibility graph is called the *characteristic path length*, which shows the separation between any two randomly chosen locations.

In summary a small world network is the network with a high clustering coefficient and a low characteristic path length. It resembles a regular graph in the sense of high clustering and resembles a random graph in the sense of short path length. The two properties determine an efficient structure for a small world network in terms of information propagation. In the following subsections, we will briefly introduce various models that illustrate a number of features of the real world networks.

2.2. Rewiring simulation (W-S model)

A regular graph has a high clustering coefficient and a long characteristic path length. On the other hand, a random graph has a low clustering coefficient and a short characteristic path length. These properties of both regular and random graphs prompt exploration of small world behaviour. Watts and Strogatz (1998) designed a simulation through rewiring a few links of a regular graph, i.e. to introduce a few random links to replace the original neighbouring links (Figure 1). They define a parameter to control the level of randomness. When the parameter equals to 0, it represents purely ordered graph, not random at all; when the parameter equals to 1, it represents purely random graph. Imagining a regular graph represented as in Figure 1, the simulation starts rewiring a few links, which means that fix one end of the links and randomly rewire to another node in the ring. Through carefully controlling the parameter, a range of graphs with different levels of randomness is created. Between the two extreme graphs, it is found that there is a range where networks maintain both high clustering coefficient as the regular graph and a low path length as the random graph. The networks that combine the two properties are the small world networks.

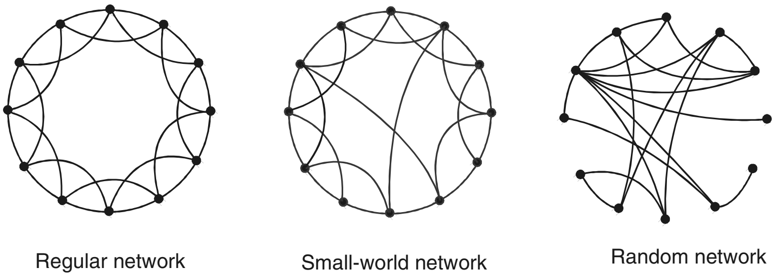

Fig. 1. Regular, small world and random networks

2.3. Efficient behaviour of small world networks (M-L model)

Marchiori and Latora (2000) investigated various small world networks and suggested a generalized concept of connectivity length toward a better understanding of small world behaviour. The concept is defined by the harmonic mean (rather than the arithmetic mean) of the shortest distances among all pairs of nodes within a graph. The model also relaxed several constraints set by the initial small world model by Watts and Strogatz (1998) and extended it to weighted and unconnected graphs. What is important for the model is that it brings an efficiency view into the small world networks, i.e. a small world network is an efficient network with both local and global efficiency in terms of information propagation.

2.4. Scale free property of small world networks (B-A model)

In studying the small world property of World Wide Web, Barabasi and Albert (1999) noted that the degree distribution of individual web pages is extremely uneven. Some pages are extremely well connected, while most pages have nearly the same low level of connection. This distribution is called scale-free, as it is different from normal distribution, which has ”a characteristic *scale* in its node connectivity, embodied by the average node and fixed by the peak of the degree distribution” (Barabasi 2002, p. 70). The scale-free property can be used to explain the growth mechanism of real-world networks, i.e. the rich get richer and preferential attachment in Barabasi’s term. Most dynamically evolved networks investigated by various researchers demonstrate the scale-free property. Note that a network with scale-free property is a small world, but not vice versa.

2.5. Directed search (W-D-N model)

The W-S model proves existence of short chains of intermediaries with many real world networks. However, how to find a short chain by individuals constitutes another challenge. This is actually the concept of directed search rather than broadcast search in analogue with Breadth First Search (BFS), which is impossible in practice (Watts 2003). With W-S model, each node is chosen at a uniform randomness for rewiring, while Kleinberg (2000) realized that in reality people use different senses of distance such as geographic distance, professionals, and race to decide to whom they want to make an acquaintance. Based on this observation and insight, Kleinberg constructed a two-dimensional cellular space in which each cell is connected by four immediate neighbours plus a random link. He found through the study that some networks that meet some particular condition are searchable (see Kleinberg 2000 for details). Watts et al. (2002) took a step further and integrated all what they called social dimensions (or identities) into search strategies. Eventually Watts and his colleagues concluded that most social networks are searchable because that various identities are involved in determining a next target of a short chain by individuals. It sets a significant difference from Kleinberg's finding where the condition is hard to meet for a network to be searchable.

3. Geographic environments as small worlds

3.1. Small world properties of geographic environments

Small world properties can be considered as a perfect combination of regularity and randomness, and they seem available in geographic contexts. On the one hand, spatial processes are far from a random process, as investigated for instance by spatial autocorrelation (Cliff and Ord 1973); on the other, common sense appears to suggest that geographic environments and spatial processes are not regular or ordered at all. The first law of geography elegantly states that "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). The law appears to suggest that with geographic environments everything has many links to neighboring things (regularity), but a few distant links between the distant things (randomness). The regularity can also be said to be a sort of high clustering. Taking a transport network for example, the block-by-block street network constitutes a sort of regular network with a high degree of clustering.

Regarding randomness, various transport systems, such as bus and underground networks, imposed on the top of a street network provide some randomness or shortcuts. This is the very small world mixed with the nature of regularity and randomness used in Kleinberg's model. Sui (2004) made a similar remark in a forum on Tobler's first law of geography with AAG annuals. Batty (2001) has also commented how cities might be treated as small worlds, in particular, how new technologies have shrunk the cities with new transport means such as underground and high speed trains for global cities. It should be noted that near things should not be understood only in the sense of Euclidean geometry. For instance, two distant locations could be near things if they are visually accessible, or two GIS people are near people as they are in the same field, no matter how far they are in physical space. It is worth noting that the small-world problem was studied in the '70s and '80s in a geographic context. One of the earliest works is about the small-world problem in a spatial context (Stoneham 1977), although some issues like integration and segregation can now be better measured by clustering coefficient, a basic measure in the W-S model.

Geographic environments can be represented as topological networks in which vertices represent geographic entities and edges represent possible links or relationships between entities. For example, from a topological perspective, we represent a street network as a topological network based on a "named street"-oriented view (Jiang and Claramunt 2004), i.e. all the named streets are represented as nodes, and street intersections as links of a graph (Figure 2). Compared to the conventional geometric view, this representation provides a complementary view to street networks for modeling purposes, as it is defined at a higher level of abstraction. It can be used to study morphological structure and evolution of urban street networks. Apart from street network, a built environment and a terrain surface can be represented as a visibility graph as briefly introduced in section 2.1.

With the topological representations or street topologies, we have found that small world properties appear with urban street networks. Thus streets are highly clustered on the one hand, and are separated by a short chain of intermediate streets on the other. In a similar way, we examined visibility graphs with a built environment and a terrain surface (Figure 3). All the studies show that these geographic environments are indeed small worlds (Jiang 2005, Jiang 2004). In the context of the visibility graph, it means that two randomly chosen locations within an area are visually separated by a short chain of intermediate locations, no matter how many point locations one imposes on the spaces.

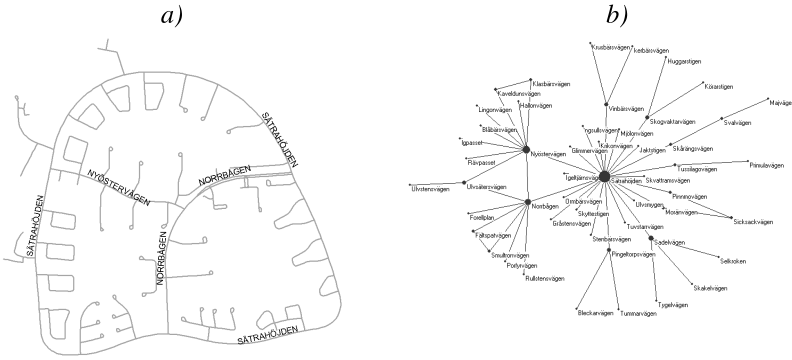

Fig. 2. A small street network (a) and its connectivity graph (b)

(Note: every node in (b) is labelled by the corresponding street name, and the size of nodes shows the degree of connectivity of individual streets)

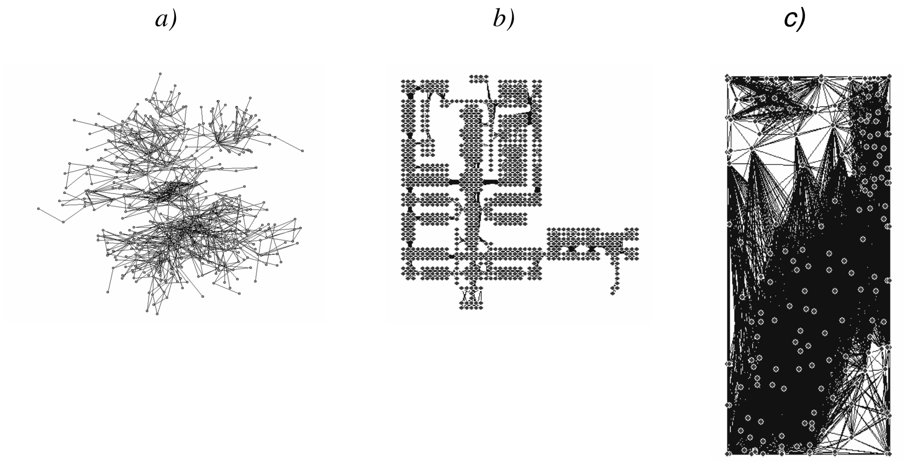

Fig. 3. Street-based topology of Gävle city (a), visibility graphs with the Tate Gallery (now called Tate Britain) (b) and a terrain surface (c)

3.2. Efficiency distribution for geographic environments

The exhibition of small world properties supports the idea that urban street net works are efficient at both local and global levels in terms of traffic flow. In a similar fashion, the Tate Britain and the terrain surface are efficient for search and navigation. It means that locations are visually reachable in an efficient manner with a built environment due to a high clustering coefficient at a local level and a low path length at a global level. This provides theoretic evidence whether or not a geographic environment in general is efficient. However, the efficiency view is with respect to an entire system and it can be extended to individual levels.

It is interesting to assess how these measures differentiate among individual geographic objects, locations within a geographic environment. Thus spatial effect for the small world properties of a geographic environment can be examined. For instance, using the M-L model, the efficiency for each street and each point location can be computed to illustrate its distribution. As an example, Figure 4 shows the distribution of both local and global efficiency for the Tate Britain. We can note that those locations with large dots have a higher efficiency.

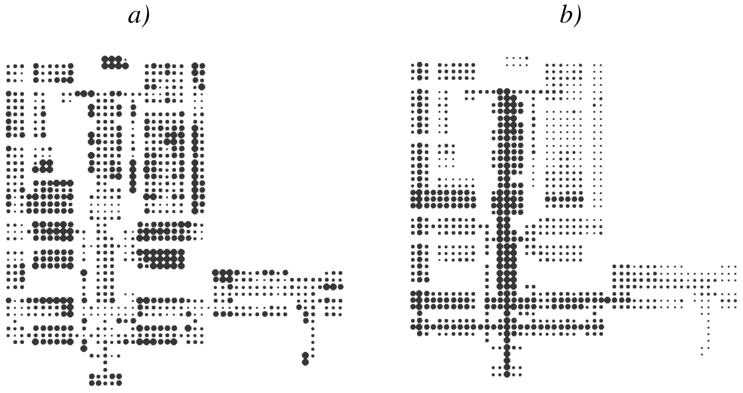

Fig. 4. Distribution of local (a) and global (b) efficiency for the individual locations with the Tate Britain

The distribution of efficiency provides new insights into geographic environments from both analysis and design perspectives. From an analytical perspective, we can examine the questions as follows: which parts are more efficient than others? When would a street network be evolved into most efficient? From a design perspective, it would be interesting to study how to create an environment with both local and global efficiencies, which are often desired. Additionally we can introduce weights into the graph representations. For example, in the case of street network, we did not represent multiple intersections between a pair of streets, which could be defined as a weight for a link, i.e. the more intersections, the stronger ties between a pair of streets. In the similar way, visibility link can be weighted in terms of distance, i.e. the shorter the distance between two point locations, the stronger the visibility link.

3.3. Re-examination of scale-free property

True scale-free property is only applicable for infinitely large networks, but in practice, most networks have a cutoff with their power law distribution (Watts 2003). This leads us to re-examine the finding we made in our recent work (Jiang and Claramunt 2004). Taking the case of Gävle street network for example, its log-log plot is indeed not strictly linear as shown in Figure 4a. However, a linear tendency seems exist if you make an appropriate cutoff, in particular when compared to the log-log plot for visibility degree (Figure 4b) that shows a stronger exponential distribution. We could characterize street networks as a broad-scale network, i.e. connectivity distribution has a power law regime followed by a cutoff, according to Amaral et al. (2000). The distribution difference shown in Figure 4a and 4b suggests the fact that scale free property is indeed a signature of dynamic networks, since street networks are evolved dynamically while visibility graphs are static in essence. It is important to note that it is these well-connected streets within a network, or hubs as they are often called, that keep a network small. For example, some streets are connected up to 29 other streets in the Gävle network and some locations are visible up to 135 other locations among a total of 960 locations in the Tate Gallery. These streets and locations constitute the hubs for the respective networks.

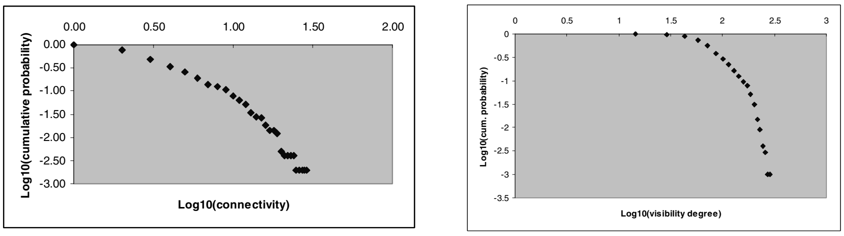
a)
b)

Fig. 5. Log-log plot of street connectivity versus cumulative probability (a) and visibility degree versus cumulative probability (b)

Based on the remark that the scale-free property is the patent signature of self-organization in complex systems (Barabasi 2002), we tend to suggest that street networks or street topologies are self-organized because of the scale-free property with street topologies - another signature that cities are self-organizing systems as elaborated by Portugali (2000). In this connection, the evolution of street network is much similar to that of the Internet and Web. Two important laws governed the development of the Internet and Web (Barabasi 2002) seem to be applicable to the street topologies as well, that is, growth - earlier streets have more time to be linked by other streets, and preferential attachment - preferences are given to the already well connected streets. The two laws together generate the scale-free property, or a power law of the street topologies.

3.4. Directed search in geographic environments

Search and navigation in geographic environments have been a research issue for a long time in the fields of behavioral geography and environmental psychology. How to design an artificial environment that is easy to navigate has also been a basic requirement in architectural and urban design. The small world models in general and the W-D-N model in particular provide some insight into various issues involving search and navigation in a geographic context. The visual separation within a geographic environment is short. However, short visual separation does not imply that an individual can easily find a short path that leads to a specific target. This is the issue of *directed search*. Directed search is a decentralized search strategy, in contrast to broadcast search in analogue with BFS algorithm.

Put differently, individuals use only local information about the network to determine the next link and eventually a short chain is found with the collective effort of individuals. Although it is defined in the context of social networks, we believe that people conduct directed searches from time to time in a geographic context, in particular when one comes to a new place with frequent search within an airport, an unfamiliar city and even a building complex.

The W-D-N model assumes that individuals are hierarchically organized within a social network, and it relies on the height of the hierarchy to reach a target. On the other hand, Kleinberg's model assumes a constant degree of the nodes, and finds that geographic distance is an important factor to be considered in the directed search. These two models have special implications to geographic environments, as both distance and hierarchical organization are natures of a geographic environment. Both physical worlds and cognitive maps are organized hierarchically (Portugali 1996). For example, places or buildings are organized into neighborhoods, neighborhoods into districts, and districts into cities. The same hierarchical structure is available in human internal representation of geographic environments, so called cognitive maps.

A geographic environment for human navigation can be constructed as a search network in which nodes are individual locations and edges are links between the locations. The links could be in different ways like a visual link (visibility), transport link (a bus line), or a road link. Therefore visibility is one of the factors that impact on search and navigation. Several other factors such as available transportation means, maps, and guided information given by someone are also involved in search decision. Within the network, landmarks, defined as having key characteristics to be easily recognizable and memorable in an environment, act as hubs within a search network. Sorrows and Hirtle (1999) suggested three types of landmarks that involve visual, cognitive and structural landmarks. The final structural landmarks can be considered as those point locations with the highest visibility degree. Although cognitive and visual landmarks are not so visually accessible compared to structural landmarks, they are often treated as hubs and are easily reminded because of its distinction in meaning. We can say that because of various landmarks of an environment it becomes searchable.

4. Discussion and conclusion

A network of interconnected and interacting objects provides an alternative representation and modeling approach to geographic environments. The representation is complementary to the conventional geometric representation in the sense of representing relationship of interconnected things. Therefore from a modeling perspective, it facilitates detection of the hidden structure of small world, i.e. most geographic environments are neither ordered nor random, but somewhere in between. Geographic environments seem to show high efficiency at both local and global levels. It provides new theoretical evidence that most geographic environments are efficient in search and navigation, as studied in behavior geography and environmental physiology.

This chapter took an exploratory approach to small world modeling for a better understanding of geographic environments. The ideas discussed here need some further research and experiments. The demonstration of small world properties with various geographic environments is a first step towards the understanding of geographic environments. Our finding of small world properties in the city context appears to support Alexander's famous contention that "a city is not a tree" (Alexander 1965). Indeed, the street topologies and visibility graphs with the Tate Britain tend to suggest the kind of semilattice structure rather than a tree. Some ongoing studies with small world modeling in a geographic context have taken a step forward. For instance, the idea of efficiency has been applied to the study of information and knowledge diffusion in the context of interactive learning (Morone and Taylor 2004). The small world structure and representations have been implemented in various agent-based modeling software platforms (Dibble and Feldman 2004).

A challenging issue from a small world perspective is how to design a geographic environment that is easy for search and navigation, or alternatively how to make a geographic environment closer to a small world. As a small world network shows both local and global efficiency, this should be an ideal design target. A simple principle appears to be: neither regular nor random but somewhere in between. With this principle, we still need to measure the degree of in-between in a quantitative manner. For instance, with the above-mentioned case studies, typical path length is a bit longer than that of equivalent random graphs. In the case of street topology, the average path length for the three levels of detail is 5.3, while the average path length for the equivalent random graphs is 4.1. It implies that such geographic environments could be better designed towards a more random one, i.e. to be a more efficient system at a global level. This can help in the course of design adjust a design into a right direction: to be more random or more regular. Another alternative approach could be using a randomness measure to show the level of randomness and regularity. These ideas need further experimental studies.

We have not spent much time in this chapter on how mobile vehicles and people are networked and interacting with each other and to geographic environments, therefore it is worthwhile to add some speculations on the issue. Nowadays, the world and people are more wired than ever before because of pervasive use of cell phones, tracking and positioning devices. This is particularly true in the city context. In the cities, more and more vehicles are equipped with GPS receivers and tracking systems, so it is convenient for online logistics and fleet management. On the other hand, pedestrians or people in general are interconnected with cell phones and mobile devices. This is the very concept of SwarmCity (Mitchell 1999) refers to. In many aspects, geographic environments are similar to the Internet space and many other real world networks. This similarity opens up possibilities and continues to give new inspirations to geographic modeling from the network point of view.

Acknowledgements

The author would like to thank the participants of the ESLab's international workshop who provided valuable inputs and comments to an earlier version of chapter when it was presented there. Thanks are also due to Daniel Z. Sui and Juval Portugali who suggested some valuable references.

References

Alexander, C. (1965), A city is not a tree, *Architectural Forum*, 122 (1+2), pp. 58 – 62.

Amaral, L.A.N., Scala, A., Barthelemy, M. and Stanley, H.E. (2000), Classes of behaviour of small-world networks, *Proceedings of the National Academy of Sciences*, 97, 11149 – 11152.

Barabasi, A.L. (2002), *Linked: The New Science of Networks*, Perseus Publishing: Cambridge, Massachusetts.

Barabasi, A.L. and Albert, R. (1999). Emergence of scaling in random networks, *Science*, 286, pp. 509 – 512.

Batty, M. (2001), Editorial: cities as small worlds, *Environment and Planning B: Planning and Design*, Pion Ltd, 28, pp. 637 – 638.

Cliff, A.D. and Ord, J. K. (1973), *Spatial autocorrelation*, London: Pion.

Dibble, C. and Feldman, P.G. (2004), The GeoGraph 3D computational laboratory: network and terrain landscape for RePast, *Journal of artificial societies and social simulation*, 7(1), available at http://jasss.soc.surrey.ac.uk/7/1/7.html (accessed 2004-04-05).

Holland, J.H. (1995) *Hidden Order: How Adaptation Builds Complexity*, Addison-Wesley Publishing Company, Reading, MA.

Jiang, B. and Claramunt, C. (2004), Topological analysis of urban street networks, *Environment and Planning B: Planning and Design*, Pion Ltd, 31, pp. 151 – 162.

Jiang, B. and Claramunt, C. (2002). Integration of space syntax into GIS: new perspectives for urban morphology, *Transactions in GIS*, Blackwell Publishers Ltd, 6(3), pp. 295 – 309.

Jiang, B. (2005), A structural perspective on visibility patterns with a topographic surface, To appear in *Transactions in GIS*, Blackwell Publishers, 9, pp. xx – xx.

Jiang, B. (2004), Small world properties of visibility graph with a built environment, *University of Gävle Working Paper*, 15 pages.

Kleinberg, J. (2000), Navigation in a small world, *Nature*, 406, pp. 845.

Kleinfeld, J.S. (2002), The small world problem, *Society*, 39(2), pp. 61 – 66.

Marchiori, M. and Latora, V. (2000), Harmony in the small-world, *Physica A*, 285, 539 – 546.

Milgram, S. (1967), The small world Problem, *Psychology Today*, Vol. 2, pp. 60 – 67.

Mitchell, W.J. (1999), *e-topia: Urban Life, Jim--But Not As We Know It*, Cambridge: The MIT Press.

Morone, P. and Taylor, R. (2004), Small world dynamics and the process of knowledge diffusion: the case of the metropolitan area of greater Santiago De Chile, *Journal of artificial societies and social simulation*, 2(2), available at http://jasss.soc.surrey.ac.uk/JASSS.html

O'Sullivan, D. (2001), Graph-cellular automata: a generalised discrete urban and regional model. *Environment and Planning B: Planning & Design* **28**(5), pp. 687 – 705.

Portugali, J. (1996), *The construction of cognitive maps*, Kluwer Academic Publishers, Dordrecht.

Portugali, J. (2000), *Self-organization and the city*, Springer, Berlin.

Sorrows, M. and Hirtle, S. (1999), The nature of landmarks for real and electronic spaces, in C. Freksa and D. Mark, editors, *COSIT 99: Spatial information theory: cognition and computational foundations of geographic information science*, Springer, pp. 37 – 50.

Stoneham, A.K.M. (1977), The small-world problem in a spatial context, *Environment and Planning A*, **9**, pp. 185 – 195.

Strogatz S. (2001), Exploring complex networks, *Nature*, 410, pp. 268 – 276.

Sui, D.Z., (2004), Tobler's first law of geography: A big idea for a small world? *Annals of the Association of American Geographers* 94(2): pp. 269 – 277.

Tobler, W.R. (1970), A computer movie simulating urban growth in the Detroit region, *Economic geography*, 46. pp. 234 – 240.

Turner, A., Doxa, M., O'Sullivan, D. and Penn, A. (2001), From isovists to visibility graphs: a methodology for the analysis of architectural space, *Environment and Planning B: Planning and Design*, Pion Ltd., pp. 103 – 121.

Watts, D.J. and Strogatz, S.H. (1998), Collective dynamics of 'small-world' networks, *Nature*, 393, 440 – 442.

Watts, D.J., Dodds, P.S. and Newman, M.E.J. (2002), Identity and search in social networks, *Science*, 296, pp. 1302 – 1305.

Watts, D.J. (2003), *Six degrees – the science of a connected age*, W. W. Norton & Company: New York, London.