Centrality measures and vulnerability of spatial networks

Urška Demšar
National Centre for Geocomputation,
National University of Ireland, Maynooth, Ireland
urska.demsar@nuim.ie

Olga Špatenková
Helsinki University of Technology,
Espoo, Finland
olga.spatenkova@hut.fi

Kirsi Virrantaus
Helsinki University of Technology,
Espoo, Finland
kirsi.virrantaus@hut.fi

ABSTRACT

Effective management of infrastructural networks in the case of a crisis requires a prior analysis of the vulnerability of spatial networks and identification of critical locations where an interdiction would cause most damage and disruption. This paper presents a preliminary study into how a graph theoretic structural analysis could be used for this purpose. Centrality measures are combined with a dual graph modelling approach in order to identify critical locations in a spatial network. The results of a case study on a street network of a small area in the city of Helsinki indicate that 'betweenness' is the most promising centrality measure for this purpose. Other measures and properties of graphs are under consideration for eventually developing a risk model not only for one but for a group of co-located spatial networks.

Keywords

Network analysis, centrality measures, graph theory, spatial networks.

INTRODUCTION

One of the main tasks of the governments is to ensure the safety of their people and the security of the vital functions of the society. Natural catastrophes and disasters as well as terrorism can cause acute crises in places which are not traditionally thought of as areas of potential disruption (YETT, 2006). The most vulnerable and also vital structures in a modern society are various spatial networks: electricity, energy, water and transportation networks. What happens to a city if part of the electricity network is down, or if no water is available or if transportation is not functioning? These are threatening questions which need to be answered in advance and not when the crisis is already happening.

There has been much interest in examining vulnerabilities and risk in critical network infrastructures. Vulnerability or survivability of a network depends on its connectivity and flow properties. Recent approaches are therefore mainly based on analysis and optimisation of network flows and investigation of topological measures that indicate scale-free and small-world properties of the networks (Carlier, Li and Lutton, 1997; Latora and Marchiori, 2004; Grubesic and Murray, 2006; Guida and Maria, 2007; Murray, Matisziw and Grubesic, 2007). These studies focus on global networks at a continental or a country scale, such as telecommunication networks, Internet, electrical power systems, airport networks and other global transportation networks. US Internet as an example of a spatial information network has received a particularly thorough investigation of its structure and vulnerability (Wheeler and O'Kelly, 1999; Gorman and Malecki, 2000; O'Kelly and Grubesic, 2002; Grubesic, O'Kelly and Murray, 2003; Gorman and Kulkarni, 2004; Gorman, Schintler, Kulkarni and Stough, 2004). Not much focus has been given to smaller local networks, such as infrastructural networks in a city.

A common topic in the vulnerability analysis of networks is interdiction, where network elements are intentionally or otherwise disabled, which disrupts the flow through the network (Murray et al., 2007). One of the common problems is identification of critical infrastructure elements where interdiction would cause most damage to the network. This problem is important for global as well as for local networks.

This paper presents an approach towards identification of critical elements in spatial networks using a dual graph modelling approach (meaning that the line graph is used for analysis instead of the original network) and centrality measures: 'degree', 'closeness' and 'betweenness'. Linear geographic phenomena are digitally represented as graphs (Worboys and Duckham, 2004) and using graph theory for the analysis of these phenomena is therefore a logical choice. Much has been in done in this area, in particular in transportation (Miller and Shaw, 2001), but most of the GIS applications only use traditional methods. Well-known examples include finding the shortest path or the nearest facility, calculating the flow capabilities of network elements or analysing connectivity of a network with the sole aim to correct the data if the network turns out to be unconnected. Standard GIS packages do not provide tools beyond the few standard ones and geographers are rarely acquainted with mathematically more advanced topics, such as probabilistic graph theory or structural graph theory. Using the dual modelling together with structural measures from social network analysis combines two such advanced concepts and therefore presents a novel approach towards the analysis of spatial networks.

The approach has been tested in a small pilot case study on a local spatial network: the street network of an area in the city of Helsinki. A more extensive investigation including other graph theoretic measures is currently under way, the goal of which is to develop a risk model first for one single local spatial network and ultimately for a group of all co-located infrastructural networks in a city. Even though the model is developed using local networks, it could be extended for global network modelling or even used for other applications, where spatial networks are used as a spatial metaphor for non-spatial phenomena.

METHODOLOGY

A critical location in the infrastructure can be defined as an object whose removal or destruction changes the structure of the infrastructural network in terms of flow and connectedness. If the infrastructure objects at such locations are removed or damaged, this disconnects large areas of the network from each other and causes the flow between two areas to be relocated through a longer detour path. An example of such locations in cities are bridges through which several infrastructural networks are drawn – a road, a railway track, a subway track, electrical and gas lines, etc. might all be located on a single bridge. If such a bridge is removed or damaged, the areas on each side of the bridge become disconnected and the flow between these areas is disrupted.

The method presented in this paper attempts to use a dual modelling approach from graph theory combined with centralities for identification of critical locations. This means that the centrality measures (Freeman 1979) are calculated on a line graph of the street network. Similar dual modelling approaches have been used to describe the structural properties of urban space in space syntax (Hillier and Hanson 1984). Space syntax describes the urban space using a connectivity graph, which is derived from a selected space representation in a similar manner as the line graph from a graph. Several space representations are possible: convex spaces (Hillier and Hanson 1984), axial lines (Hillier and Hanson 1984), isovists (Batty and Rana 2002) and named streets in a street network (Jiang and Claramunt, 2004a, 2004b; Rosvall, Trusina, Minnhagen and Sneppen, 2005; Crucitti, Latora and Porta, 2006). In all these approaches, the importance of spaces is described by calculating centralities and topological measures for their corresponding vertices in the connectivity graph. The approaches that analyse street networks all use larger units (i.e. named streets) as their basic spatial elements. While this works if the interest is in structural analysis of urban space, it is not appropriate for risk modelling of a local spatial network. Therefore our approach takes the original vertices and edges of the street network as the basic spatial elements in the calculation of the line graph.

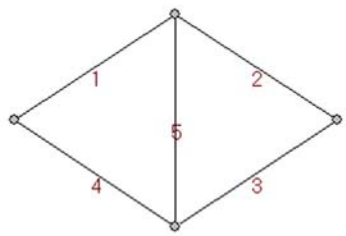
Dual graph modelling and centrality measures

This section introduces basic definitions from graph theory (Jungnickel 2005), used in our method.

A graph G=G(V,E) is a pair of two finite sets V and E, where $V \cap E = \emptyset$ and E is a subset of $V \times V$, that is, a set of two-element subsets of V. The elements of V are called vertices, the elements of E edges. An edge e = {u, v} from the set E connects vertices u and v and is usually denoted as e = uv. Two vertices x, y of G are adjacent or neighbours if xy is an edge of G. Two edges of G are adjacent if they share a common vertex. The usual way to draw a graph is to represent each vertex by a dot and each edge by a line connecting two dots.

A line graph L(G) of a graph G is defined as follows: the edges E of G are the vertices of L(G). If e and f are two edges of graph G (and so the vertices in its line graph L(G)), there exists an edge between them in L(G) if and only if e and f are adjacent in G. An example of a graph G and its line graph L(G) are shown in figure 1. The line graph is sometimes also called an edge-dual of its original graph and is used when edge properties need to be translated into vertex properties - this approach is called dual graph modelling.

a)



b)

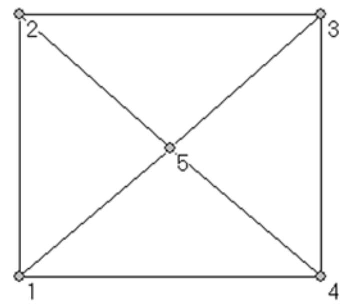


Figure 1. a) Graph G and b) its line graph L(G). The labelling of the edges of the graph corresponds to the labelling of the vertices in the line graph.

The structure of the graph can be described by centrality measures, which are calculated for each vertex and represent the structural importance of the vertex – vertices with high centrality have a larger impact on other vertices. Three commonly used centrality measures are degree, closeness and betweenness (Freeman 1979). The degree of a vertex v, usually denoted as d(v), is the number of its neighbours. It describes the local structural importance of the vertex. The closeness, c(v), is the shortest distance from v to all other vertices (distance is measured as the number of edges in the shortest path between v and some other vertex). The third measure is betweenness, b(v), which is defined as the proportion of the shortest paths between every pair of vertices that pass through the given vertex v towards all the shortest paths. The vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not. Centrality measures were first used for analysis of social networks (Freeman 1979), where vertices represent persons or institutions and edges relationships between them. In social networks, degree and closeness describe the reachability of a person in the network, while betweenness describes the extent to which the person is needed as a link in the chains of contacts that facilitate the spread of information in the network (de Nooy, Mrvar and Batagelj, 2005).

The method

The approach presented in this paper calculates centrality measures of the line graph of the network in order to identify critical infrastructure locations. First, all attribute and geometry information is removed from the spatial network, and its graph is created using solely topological information. A line graph is then derived from this graph. The reason for introducing this step is that critical locations correspond to edges in the original network. Since the centrality measures can only be calculated for vertices and not for edges, it therefore makes sense to first translate edge properties into vertex properties by generating a line graph.

The next step is to calculate centrality measures for the vertices of the line graph. These values are inserted back into the original spatial network as new attributes of the respective links. These new attributes have only been used for visual presentation but in the future these and other topological measures could be used for calculating a risk measure for each link in the original spatial network.

The procedure is illustrated in the next section by a small case study, where the centrality measures were calculated for a small sub-network of the street network of Helsinki.

Figure 2. Location of the test area in Helsinki street network with the two bridges that represent critical locations (indicated in red).

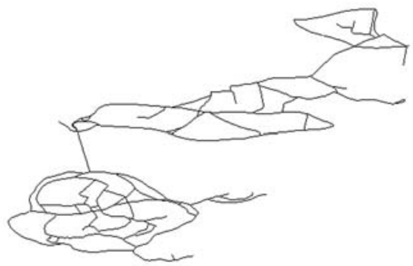
CASE STUDY

The proposed approach was tested on a real example, consisting of a street network of a small area located east of the centre of Helsinki. The chosen area consists of two small islands where Helsinki ZOO is located. Even though the area is not of high importance strategically, the two bridges in this area represent good examples of critical locations (indicated in red in figure 2). The topological structure of the street network in this area was therefore considered to be a good area for testing our methodology. The selected street network consists of 142 road segments.

The original data came from Digiroad – a National Road and Street Database of Finland (Finnish Road Administration, 2007) and came as a shapefile. The network was exported into ASCII format – the only information needed were the numbers of from-node and to-node for each road segment, which were used to build a graph representing the network, shown in figure 3. This graph contains only topological information - all geometrical and attribute information has been removed from the data. The edges in this graph correspond to the road segments in the original network.

The line graph of the network graph (figure 4) was derived using a Java programme written specifically for this purpose. This graph has 142 vertices that correspond to the edges of the graph in figure 3b. The vertices of the line graph in figure 4 are drawn in roughly the same positions as the edges that they correspond to in the graph in figure 3b. These locations have nothing to do with the geometry of the original network, but are calculated using the energy-preserving graph-drawing algorithms implemented in the network analysis software package Pajek (Batagelj and Mrvar 2006), which was used for graph visualisation and calculation of centrality measures in this study.

a) b)



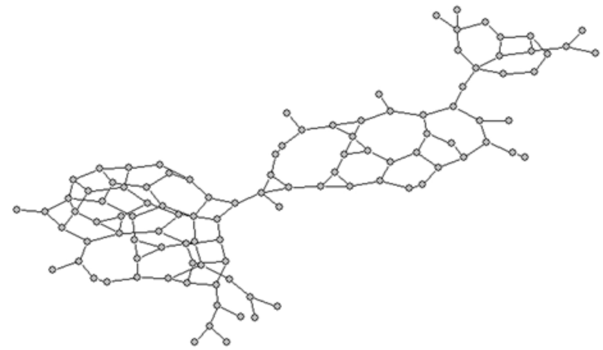


Figure 3. a) The original road network and b) its graph, where only topological information has been retained.

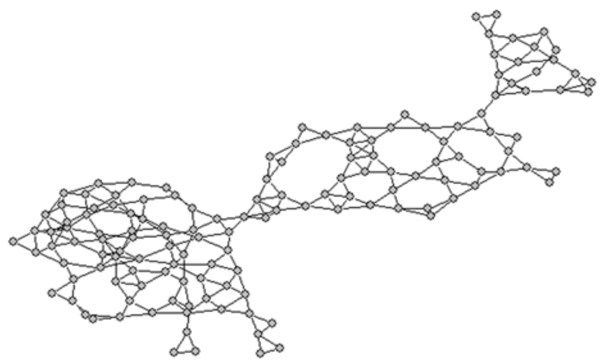


Figure 4. The line graph.

In the next step the centrality measures were calculated for the line graph. The results are shown in figure 5, where the size of the vertices indicates values of degree, closeness and betweenness respectively.

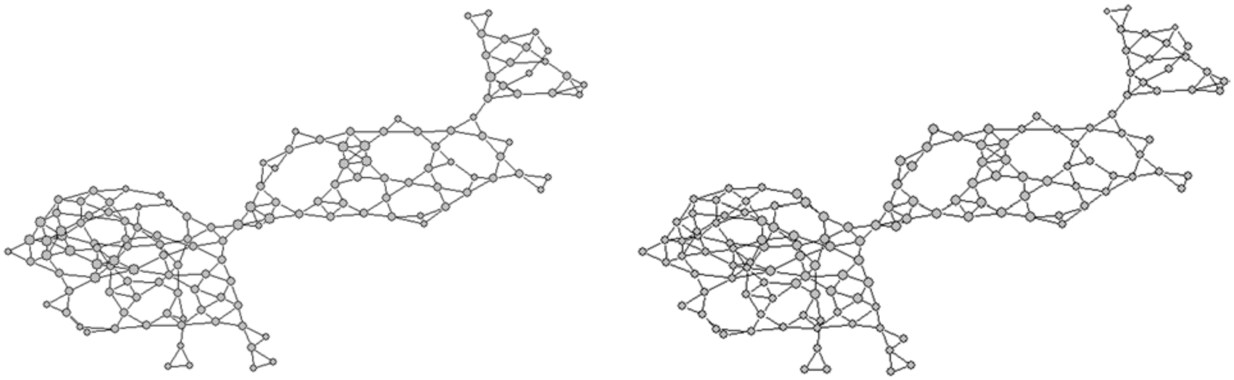
All three centrality measures were combined with the original network in the shapefile and visualised as maps in GIS software. The three maps are shown in figure 6.

RESULTS

The goal of this experiment was to investigate if any of the three centrality measures can be used for identification of critical locations in a spatial network. Each of the resulting maps in figure 6 was therefore examined to see what the measures indicate for the two bridges that represent critical locations in the test network.

Since the degree is a local measure that only considers the immediate neighbourhood of a vertex, it identifies links in critical locations only if these links have many directly adjacent links in the network. In the test network, this is the case with one of the two bridges – the southern bridge has a high degree and is drawn in red on the map in figure 6a. However, the northern bridge in the network consists of more than one link and none of the two bridge links has a high degree that would made the area stand out from the rest of the network. Moreover, there are several other areas with red links (links with a high degree) that are not located near the two critical locations, such as for example a small red sub-network in the southern island.

a) b)



c)

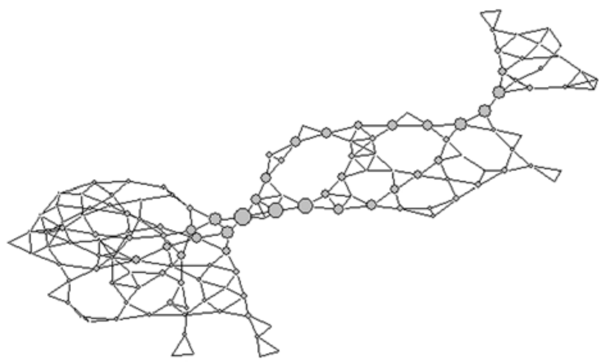


Figure 5. Centrality measures calculated at the vertices of the line graph: a) degree b) closeness, c) betweenness. The size of vertices indicates the value of each measure for that particular vertex.

Closeness describes how well a vertex is integrated in the network: a vertex with high closeness is more connected to all other vertices in the sense that the shortest paths from this vertex to all other vertices are short comparing to vertices with lower closeness. While this property can be used to identify local centres in the network, it is not very relevant for critical locations. The result in figure 6b confirms this: high closeness (links in red and orange) defines a central area around the connection between the northern and the southern island, but the two critical bridges are not identified.

Betweenness is the centrality measure that is linked to the flow in the network. Removing a vertex with high betweenness from the graph breaks off many shortest paths between other vertices, which disrupts the flow in the graph or redirects it through a longer detour. This is one of the properties that a critical location should have. Betweenness could therefore be used for identification of critical locations – perhaps not directly, but in combination with some other measure, since other links (not only those located in critical places) might also have high betweenness values. The map of betweenness in figure 6c confirms this: both critical bridges have high betweenness (the southern one is red, which indicates the highest value and the northern one is orange, which also indicates a high value), but there are other links in the neighbourhood of the bridges that also have high betweenness values.

a) b)

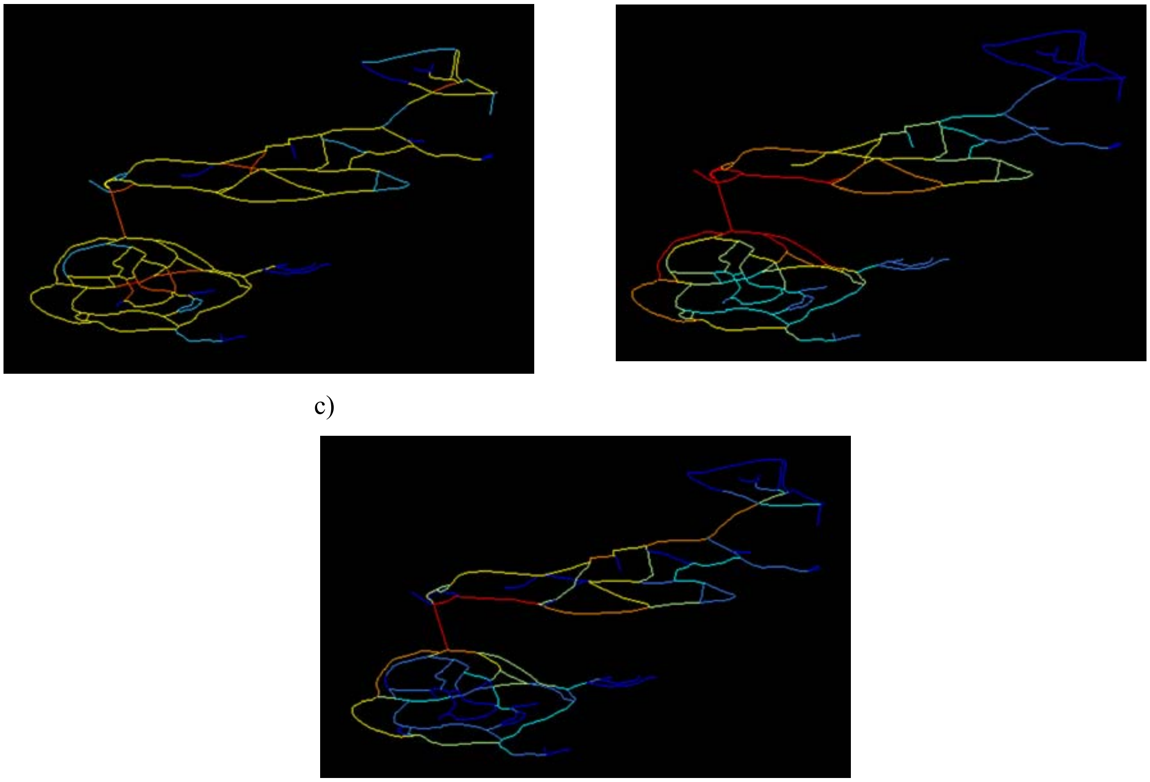


Figure 6. Centrality measures mapped on the original network: a) Degree b) Closeness, c) Betweenness. In all three cases red colour indicates high values, yellow medium values and blue low values of each measure.

CONCLUSIONS & FURTHER RESEARCH

This paper presents an approach towards identifying critical locations in a spatial network by dual graph modelling and calculation of centrality measures. Of all three compared centrality measures, degree, closeness and betweenness, it is the latter one that is the most promising for identification of critical locations, possibly in combination with some other topological or connectivity property of the network elements.

The approach presented in this paper is the first step in the investigation of how graph theoretic structural analysis could be used for modelling of vulnerability and survivability of local spatial networks. Other graph properties, such as biconnected components (Jungnickel 2005), clustering coefficient and the average shortest path (Watts and Strogatz, 1998) are being considered for inclusion in the risk modelling.

Another research direction is to develop a method that would identify critical locations using topological information from several co-located networks (for example, street, energy and information networks), not just one. Eventually, the analytically calculated risk measure could also be combined with a set of critical locations, identified by experience and attribute information. For this, other infrastructural information could be taken into consideration, such as locations of vital and vulnerable points of the society (for example places where a lot of people are located at the same time, such as shopping centres, railway stations, hospitals, and important objects, such as central governmental buildings, historical buildings, shelters, etc.).

Since the method uses only topology to reveal knowledge about the functionality of the network, it could be used more broadly than just for analysis of spatial networks. It could be considered as a spatial data mining approach towards extracting information from a structured but unlabeled dataset and used for tasks such as classification of unlabeled data, cluster analysis and identification of link types. Another advantage is that since the geometry does not play a role in the method and is only used for visualisation purposes, the method could be used for knowledge extraction in any application where a network is seen as a spatial metaphor for the (not necessarily spatial) phenomenon under investigation.

ACKNOWLEDGMENTS

The authors would like to thank Jukka Krisp from Helsinki University of Technology for fruitful discussions during the development stage of this method and for all the help with preparation of the spatial network datasets used in this study.

REFERENCES

1. Batagelj, V. and Mrvar, A. (2006) Pajek - Program for Analysis and Visualization of Large Networks, version 1.17., 11 Nov 2006, available at: http://vlado.fmf.uni-lj.si/pub/networks/pajek/ (last accessed 14 Dec 2006).
2. Batty, M. and Rana, S. (2002) Reformulating space syntax: the automatic definition and generation of axial lines and axial maps. Centre for Advanced Spatial Analysis, University College, London. http://www.casa.ucl.ac.uk/working_papers/paper58.pdf (last accessed 13 Dec 2006).
3. Carlier, J., Li, Y. and Lutton, J.-L. (1997) Reliability evaluation of large telecommunication networks. *Discrete Applied Mathematics*, 76, 61-80.
4. Crucitti, P., Latora, V. and Porta, S. (2006) Centrality measures in spatial networks of urban streets. *Physical Review E*, 73, 036125.
5. Finnish Road Administration (2007), *Digiroad – A National Road and Street Database*. http://www.digiroad.fi/en_GB/
6. Freeman, L.C. (1979) Centrality in social networks: Conceptual clarification. *Social Networks*, 1, 215–39.
7. Gorman, S.P. and Malecki, E.J. (2000) The Networks of the Internet: An Analysis of Provider Networks in the USA. Telecommunications Policy, 24(2), 113-134.
8. Gorman, S.P. and Kulkarni, R. (2004) Spatial small worlds: New geographic patterns for an information economy. Environment & Planning B, 31, 273-296.
9. Gorman, S. P., Schintler, L., Kulkarni, R. and Stough, R. (2004) The Revenge of Distance: Vulnerability Analysis of Critical Information Infrastructure. *Journal of Contingencies and Crisis Management*, 12(2), 48-63.
10. Grubesic, T. H., O'Kelly, M.E. and Murray, A.T. (2003) A geographic perspective on commercial Internet survivability. *Telematics and Informatics*, 20 51-69.
11. Grubesic, T. H. and Murray, A. T. (2006) Vital Nodes, Interconnected Infrastructures and the Geographies of Network Survivability. *Annals of the Association of American Geographers*, 96, 64-83.
12. Guida, M. and Maria, F. (2007) Topology of the Italian airport network: A scale-free small-world network with a fractal structure? *Chaos, Solitons and Fractals*, 31, 527-536.
13. Hillier, B. and Hanson, J. (1984) *The social logic of space*. Cambridge University Press.
14. Jiang, B. and Claramunt, C. (2004a) A Structural Approach to the Model Generalisation of an Urban Street Network. *Geoinformatica*, 8, 2, 157-171.
15. Jiang, B. and Claramunt, C. (2004b) Topological analysis of urban street networks. *Environment and Planning B: Planning and Design*, 31, 151-162.
16. Jungnickel D. (2005) *Graphs, Networks and Algorithms*. 2nd edition. Springer-Verlag.
17. Latora, V. and Marchiori, M. (2004) How the science of complex networks can help develop strategies against terrorism. *Chaos, Solitons and Fractals*, 20, 69-75.
18. Miller, H. J. and Shaw, S.-L. (2001) *Geographic Information Systems for Transportation*. Oxford University Press.

19. Murray, A. T., Matisziw, T. C. and Grubesic T. H. (2007) Critical network infrastructure analysis: interdiction and system flow. *Journal of Geographical Systems*, online pre-publication.

20. de Nooy, W., Mrvar, A. and Batagelj, V. (2005) *Exploratory Social Network Analysis with Pajek*. Cambridge University Press.

21. O'Kelly, M.E. and Grubesic, T.H. (2002) Backbone topology, access, and the commercial internet. *Environment and Planning B*, 29, 533-552.

22. Rosvall, M., Trusina, A., Minnhagen, P. and Sneppen, K. (2005) Networks and Cities: An Information Perspective. *Physical Review Letters*, 94, 028701.

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

24. Wheeler, D. and O'Kelly, M.E. (1999) Network topology and city accessibility of the commercial internet. *The Professional Geographer*, 51(3), 327-339.

25. Worboys, M. and Duckham, M. (2004) *GIS – A Computing Perspective*. CRC Press.

26. YETT (2006) The Strategy for Securing the Functions Vital to Society. Governmental resolution. Available in English: http://www.defmin.fi/index.phtml?l=en&s=335 (last accessed 19 Jan 2007).