Vulnerability Analysis of Weighted Graphs: a Measure of the Vulnerability of Spatial Networks by Using Betweenness Centrality

Zhe Zhang[1], Hannes Seppänen[2,]Virrantaus Kirsi[3]
[1,2,3]Aalto University, Espoo, Finland
[1] Corresponding author: Zhe Zhang, e-mail: zhe.zhang@aalto.fi

Abstracts

The infrastructure of critical networks is important for human life. Vulnerability analysis of critical networks has often been discussed in the literature because of the increasing numbers of natural and man-made hazards and disasters. Betweenness is one of the topological centrality measure methods in the field of graph analysis. It has been used to identify the critical location for an un-weighted spatial network, where it is assumed that the length of each graph segment is the same and the shortest path for each pair of locations (vertex) is the path which contains the fewest locations (vertices). In reality, the shortest path between two locations in a spatial network should have shortest length; therefore betweenness analysis should be applied to a weighted graph, where weight refers to the length of each graph segment. In this article, a betweenness centrality measure was applied to three weighted graphs (transportation, drinking water and natural gas networks) to measure the importance of a graph segment according to its topological location. This research work is part of the "Critical Infrastructure Protection (Matine)" project. The goal of the project is to analyse the vulnerability of the critical networks of the city of Tampere for crisis management purposes.

Keywords: critical networks, Betweenness centrality, graph analysis, crisis management

1. Introduction

Hazards, disasters, and concerns such as the increasing amount of terrorism cause insecurity in society. Crisis management with respect to urban structures has become one of the core tasks of governmental authorities. Critical networks such as transportation networks, electricity networks etc. play an important role in the entire crisis management process. Therefore, the vulnerability analysis of critical networks is often discussed in the government's preparedness planning and mitigation tasks. This research work is part of the "Critical Infrastructure Protection (Matine)" project. The project was funded by the Scientific Defence Council of Finland. Several organisations, such as Aalto University, Tampere University of Technology, the Finnish Defence Force Institute, and the City of Tampere, were involved in the project.

In Finland, the Pirkanmaa location is very central and it includes 22 municipalities. The Prikanmaa Region had 491,000 inhabitants at the end of the year 2011, and the population of the central city of Tampere is 217,603 (Population Register Centre, 2013). Prikanmaa, and especially the city of Tampere, is on the route of the major highways and railways. For instance, Highway No. 3 goes from north to south, Highway No. goes to the south east direction and Highway No. 12 follows an east-west direction. In addition, the city of Tampere and its neighbouring municipalities are the focus of the operations of the entire country's main activities in the field of commercial and industrial enterprises, as well as military units. Therefore, Tampere is an important logistical city for Finland. The goal of the project is to analyse the vulnerability of the critical networks (transportation networks, water pipe networks, and natural gas networks) of the Tampere area for crisis management purposes.

Demšar et al. presented a mathematical method for modelling the vulnerability of the elements of a network, which can be used for the identification of critical locations in a spatial network. It combines line graph modelling with connectivity analysis and topological measures. Graph centrality measure was used in this research work. The method was tested with the street network of the Helsinki Metropolitan Area in Finland. The vertices of the line graph that correspond to critical locations have one or more of the following three properties: they are cut vertices and they have a high betweenness value or a low clustering coefficient. Betweenness attributes were used as one of the parameters to compute the vulnerability of the street network. The vulnerability value estimation method produced results of low accuracy because the betweenness value was computed to an un-weighted graph in which it was assumed that the length of each graph segment is the same and the shortest path for each pair of locations (vertex) is the path which contains the fewest locations (vertex). In reality, the shortest path between two locations in a spatial network should contain the shortest length; therefore betweenness analysis should be applied to weighted graphs where weight refers to the length of each graph segment (Demšar et al., 2008).

In this article, a betweenness centrality measure was applied to three weighted graphs (transportation, drinking water, and natural gas networks) to measure the importance of each graph segment according to its topological location for crisis management. The results were visualised with the help of GIS tools. The article is organized as follows: section 2 gives an overview of the methodology research such as mathematical terms and definitions. The methods which used in this research are presented in section 3. The testing of the methods and results is presented in section 4. Section 5 presents some conclusions and discussions.

2. Methodology

The concept of vulnerability has many definitions according to many authors. The meaning of the term may depend on the context. In this article, the term vulnerability is applied to urban infrastructure such as spatial networks to evaluate the importance of a graph segment with respect to a crisis according to its topological location.

Vulnerability studies of spatial networks have attracted a lot of attention recently. Berdica defines the vulnerability of the transportation system as susceptibility to incidents that can result in considerable reductions in the serviceability of a road network. A similar definition was introduced by Nicholson and Du (1994). D'Este and Taylor (2003) argue that "vulnerability" should only be concerned with consequences. They state that in a network, a node is vulnerable "if loss (or substantial degradation) of a small number of links significantly diminishes the accessibility of the node, as measured by a standard index of accessibility". Wakabayashi and Iida (1992) and Bell and Iida (1997) think the vulnerability analysis of complex spatial networks is very closely related to minimal paths and cuts. The minimal routes refer to the minimum number of successive links needed to connect a pair of nodes. For instance, the network will function well if any one of the routes functions. A minimum cut set is the minimum number of links needed to disconnect a pair of nodes. If any one of the cut sets fails, the networks fail.

The idea of network centrality was first introduced by Bavelas (1948) for social networks. He suggested that when a particular person in a group is strategically located on the shortest communication path connecting pairs of others, that person is in a central position. A social network can be represented as an undirected, non-valued graph G, consisting of a set of N nodes (or vertices) and a set of K edges (or lines) connecting pairs of nodes. Degree centrality is based on the idea that the important nodes are those with the largest number of ties to other nodes in the graph. The degree centrality of node *i* is defined as below (Nieminen, 1974):

$$C_i^D = \frac{k_i}{N-1} = \frac{\sum_{j\in G} a_{ij}}{N-1} \tag{1}$$

where k_i is the degree of node *i*.

The closeness centrality of node *i* is based on the concept of minimum distance or geodesic d_{ij}, the minimum number of edges traversed to get from *i* to *j*. It is defined (Scott, 2000) (Boccaletti, 2006) (Sabidussi, 1966) as:

$$C_i^C = (L_i)^{-1} = \frac{N-1}{\sum_{j\in G} d_{ij}} \tag{2}$$

where L_i is the average distance from *i* to all the other nodes.

Betweenness centrality can be defined on a vertex or an edge, and it measures how many times a vertex or an edge is found on the shortest path between other vertices in comparison with the total number of shortest paths. For graph G:=(V,E) with n vertices, the Betweenness C_B(v) for vertex v is:

$$C_B(v) = \sum_{\substack{s,t\in V \\ s,t\notin v}} \frac{\vartheta_{st}(v)}{\vartheta_{st}} \tag{3}$$

where $\vartheta_{st}(v)$ is the number of shortest paths passing across v and ϑ_{st} is the total number of shortest paths (Barthélemy ,2004).

The centrality measure was used in spatial network vulnerability analysis and it ranks the relative importance of the vertices and edges in a graph. Wang et al. (2010) defined new measures of centrality for a power grid structure that are based on its functionality. Asta et al. (2006) analysed the vulnerability of an air transportation network to attacks and random failures by using the centrality measure. In their work an air transportation system was modelled as a network in which the vertices are commercial airports and the edges are non-stop passenger flights. The weight of the edges refers to the capacities (number of people) of the corresponding flight. They first characterised the relevant topological and weighted centrality measures and then used these quantities as selection criteria for the vulnerability analysis.

3. Methods

Dataset

The national road and street database (Digiroad) was used for the betweenness analysis of the Tampere road network. Digiroad is a national database which contains precise and accurate data on the location of all roads and streets in Finland, as well as their most important physical features (Juslén, 2010). The Digiroad data set was downloaded from the Paituli geospatial data service in Shape format. The Tampere road network data set consists of 26,321 road segments. All the road segments were grouped into seven classes. They are highways, main roads, regional roads, private roads, important private roads, other private roads, and light traffic roads. In this project, major roads such as highways, main roads, regional roads, and private roads were selected. Each major road was cut into many small road segments, making a total of 5215 road segments. The Gasum natural gas and drinking water pipe data set were originally delivered in CAD format and converted to ArcGIS Shape format. The Gasum natural gas network data consist of 465 line segments and the drinking water pipe data consist of 46,194 line segments.

Computation process

Figure 1 illustrates the process of computing the betweenness value for the critical networks in Tampere. At the beginning, each network was translated into a line graph. Given a graph G, the line graph L (G) takes the edges of G as its vertices, i.e. V (L (G)) = E (G). Two vertices e and f in the line graph are connected if and only if the respective edges e and f are adjacent in G. The line graph is sometimes called an edge dual of its original graph (Jungnickel, 2005). In a line graph, each line segment has the following attributes attached to it: identification number; X- and Y-coordinates of the starting and ending points of each line segment, and the length of the line segment. ArcGIS was used to add the starting and ending points and X- and Y-coordinates to each line segment. The length of each line segment can be calculated by means of the formula

$$length = \sqrt{(Xcoordinate_{end} - Xcoordinate_{start})^2 + (Ycoordinate_{end} - Ycoordinate_{start})^2} \quad (4)$$

The Java programming language was used to read the coordinates of the starting and ending points of each line segment and create an adjacency matrix. In an adjacency matrix, two line segments are adjacent to each other if they share the same starting or ending point. A graph was built on the basis of the adjacency matrix, and it contains information on each line segment and its neighbour. Dijkstra's (1959) shortest path algorithm was used to calculate the shortest path between a pair of vertices and this was used to compute the betweenness value for each graph segment. ArcGIS was used to join the road network betweenness value result (Excel) to the original data set (Shape) on the basis of the unique identification number of each line segment, and the results were visualised on the map.

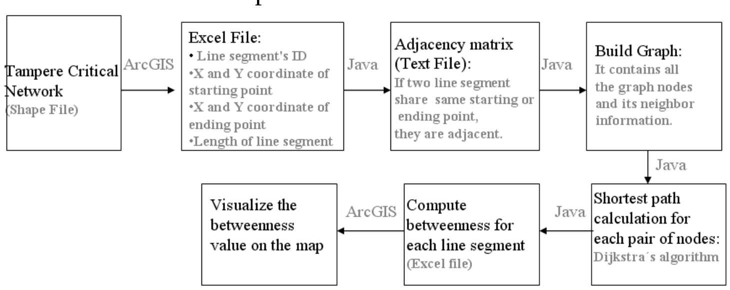

Figure 1 Betweenness value computation process for the Tampere spatial network.

4. Results

Tampere transportation network

The betweenness value map for the major roads in Tampere is illustrated in Figure 2. This map consists of two connected graphs. The southern part of the graph refers to the area of the city of Tampere. The northern part of the graph represents the Teisko area. In the Teisko area, Terälahdentie received the highest betweenness value. Siltasavontie, Kaitavedentie, Teiskonkirkkotie, and part of Viitapohjantie, Kapeentie, and Velaatantie received the second highest betweenness value. Terälahdentie and Siltasavontie received very high betweenness values because if these two roads are cut it will cause the road network in the Teisko area to become disconnected. Therefore, these two roads are vulnerable because of their topological location. In the area of the city of Tampere, Tampereen Valtatie received the highest betweenness value because it connects two parts of the city across the lake. Highways No. 12 and 9 received the second highest betweenness value.

Tampere natural gas and drinking water pipe network

In order to see where the most vulnerable area is on the basis of the topological structure of the networks, these two networks were overlaid on top of each other. In this article, part of the natural gas and drinking water pipe network data which overlapped each other was selected for betweenness analysis. Figure 3 illustrates the results of the analysis. For instance, area A is much more vulnerable than area B because area A has more natural gas and drinking water pipes than area B and the networks seem to have a higher betweenness value than area B.

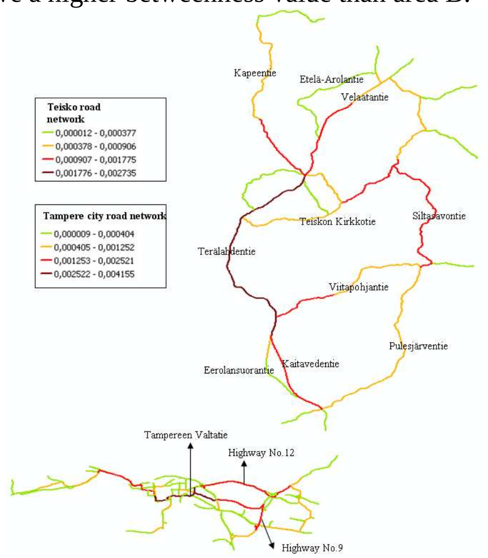

Figure 2 Results of the betweenness analysis of the Tampere transportation network.

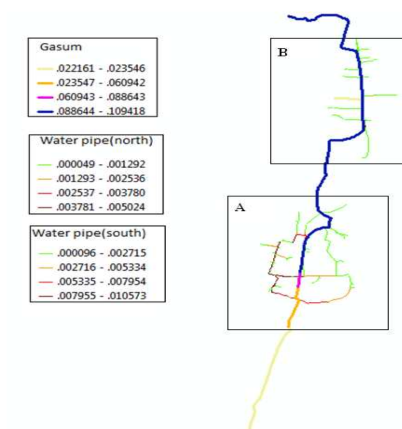

Figure 3 Results of the betweenness analysis of the Tampere natural gas and drinking water pipe networks.

5. Conclusions

In this article, the betweenness centrality measure was used to compute the vulnerability of spatial networks in the Pirkanmaa region. The results were visualised with the help of GIS tools. A road which has a higher betweenness value is more vulnerable because there are more "shortest paths" passing along this road than other roads and people may use this road more often for their journeys. The data quality issue was analysed during the betweenness computation process. The Digiroad data seem to be of reasonably high quality in terms of geometry, but missing features can be identified. For instance, the road network in the area of the city of Tampere was not connected with the Teisla road network because of a missing feature in the data set. In the Tampere and Teisko Digiroad data, each road was cut into many small road segments. The data consist of a total of 5215 segments. This increases the complexity of the computation; therefore, a generalisation tool will be needed in the future. This problem can also cause some error in the results because the road segments which belong to the same road can receive different betweenness values in the calculation of the shortest path. The data sets for the natural gas and water networks in Tampere were outdated and of poor quality. They contain many duplicated arcs and lines, which makes the computation more difficult. Therefore, the data were first cleaned by using GIS tools and only part of the data was selected for use in this work.

References

Asta, L., Barrat, A., Barthélemy M., and Vespignani, A. (2006) “Vulnerability of weighted networks”, *Journal of Statistical Mechanics: Theory and Experiment*, 2006.

Bavelas, A. (1948) “A mathematical model for group structure”, *Human Organization*, 7, 16-30.

Barthélemy, M. (2004) “Betweenness centrality in large complex networks”, *The European Physical Journal B-Condensed Matter and Complex Systems*, 38, 162-168

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. (2006) “Complex networks: structure and dynamics”, *Physics report*, 424, 175-308.

Bell, M.G.H. and Iida, Y. (1997), *Transportation network analysis.* Wiley (John) &Son, Sussex England

Berdica, K., (2002) “An introduction to road vulnerability: what has been done, is done and should be done”, *Transport Policy*, 9, 117-127.

Demšar, U., Špatenková, O., and Virrantaus, K. (2008) “Identifying critical location in a spatial network with graph theory”, *Transactions in GIS*, 12, 61-68.

D’Este, G.M. and Taylor, M.A.P. (2003) “Network vulnerability: an approach to reliability analysis at the level of national strategic transport networks”, *Proceedings of the 1st International Symposium on Transportation Network Reliability (INSTR),* Bell and Iida (Eds.), Elsevier, Oxford, 23-24.

Dijkstra, E.W. (1959) “A note on two problems in connexion with graphs”, *Numerische Mathematik,1,* 269-271.

Juslén, J. (2010) “Digiroad – a national road and street database for promoting transport telematics in Finland”, Available from web: http://www.digiroad.fi/en_GB/ (cited 19.03.2013)

Jungnickel, D. (2005) *Graphs, Networks and Algorithms*, Springer-Verlag, Berlin.

Nicholson, A. and Du, Z. (1994) “Improving transportation system reliability: a framework”. *In: Proceedings of 17th ARRB conference*, 17, 1-17.

Nieminen, J. (1974) “On the centrality in a graph”. *Scandinavian Journal of Psychology*, 15, 332-336.

Population Register Centre, 2013.Kuntien asukasluvut aakkosjärjestyksessä. Available from web: http://vrk.fi/default.aspx?docid=6865&site=3&id=0 (cited 19.03.2013)

Scott J. (2000) *Social network analysis: A Handbook*, Sage, London.

Sabidussi G. (1966) “The centrality index of a graph”, *Psychometrika*, 31, 581-603.

Wakabayashi, H. and Iida, Y. (1992) “Upper and lower bounds of terminal reliability in road networks: an efficient method with Boolean algebra”, *Journal of Natural Disaster Science*, 14, 29-44.

Wang, Z., Scaglione, Z., and Thomas, R.,(2010) “Electrical centrality measures for electric power grid vulnerability analysis”, *49th IEEE conference on decision control*, 5792-5797.