A Structural Approach to Model Generalisation of an Urban Street Network

B. Jiang [(1)] and C. Claramunt [(2)]
[(1)] Division of Geomatics, Institutionen för Teknik
University of Gävle, SE-801 76 Gävle, Sweden
Email: bin.jiang@hig.se

[(2)] Naval Academy Research Institute
BP 600, 29240, Brest Naval, France
Email: claramunt@ecole-navale.fr

Abstract. This paper proposes a novel generalisation model for selecting characteristic streets in an urban street network. This model retains the central structure of a street network, it relies on a structural representation of a street network using graph principles where vertices represent named streets and links represent street intersections. Based on this representation, so-called connectivity graph, local and global measures are introduced to qualify the status of each individual vertex within the graph. Two streets selection algorithms based on these structural measures are introduced and implemented. The proposed approach is validated with a case study applied to a middle-sized Swedish city.

Keywords : model generalisation, structural analysis, space syntax, graph modelling

Introduction

Cartographic generalisation is a constraint-based process used by cartographers to reduce the complexity in a map in a scale reduction process. It involves intensive human knowledge obtained through professional cartographic expertise and practise. Automatic generalisation has long been a research effort by both scientific researchers and cartographic practitioners (Buttenfield and McMaster 1991, Muller et al. 1995, AGENT 1998). In particular the idea of one single master database used to automatically derive maps at different scales has been a dilemma faced by many national mapping agencies. Amongst many application domains, cartographic generalisation is for example used to reduce the complexity of an urban street network in a scale reduction process while retaining its general structure. Although generalising an urban street network is often a cartographical task, this can be also considered as an operation where the objective is rather to understand the structure and organisation of the city. This is an important aim of many urban studies that focus on the understanding of urban structures and configurations. In particular, space syntax (Hillier and Hanson 1984) has developed as an important quantitative way to analyse and understand the complexity of urban street networks using a graph-theoretic method. These principles support the belief that spatial layout or structure has great impact on human social activities. The application of space syntax covers many urban studies such as modelling pedestrian movement, vehicle flows, crime mapping, and human wayfinding process in complex built environments (Hillier 1996). Many empirical studies have demonstrated the interest of the space syntax for modelling and understanding of urban patterns and structures (Hillier 1997, Holanda 1999, Jiang *et al.* 1999, Peponis et al. 2001).

This paper proposes a model generalisation of an urban street network whose objective is to retain the functional structure of the city. First our approach, which is based on a computational application of graph modelling principles, uses vertices to represent named streets and edges to represent street intersections, so a form of derived graph instead of a street network modelled as a graph. Integrating named streets (e.g. Kennedy avenue, 45th avenue) as a basic modelling unit gives a form of functional representation of the city that complements the structural view of the urban street network given by the graph-based approach (let us remark that this approach applies to cities where streets are labelled using either names or identifiers). This functional component comes from the observed fact that named streets often denote a logical flow unit, or commercial environment that is often perceived as a whole by people acting in the city. Secondly, and from this structural representation, two filtering algorithms are introduced and implemented. Our model considers both local and global measures to represent the structural property of each vertex in the graph, namely connectivity and average path length. Eventually, selection of characteristics streets is achieved through two filtering algorithms based on these structural measures. These algorithms are flexible in the sense that they support different levels of generalisation. The proposed approach is validated with a case study applied to a middle-sized Swedish city.

The remainder of this paper is organised as follows. Section 2 briefly introduces basic concepts of graph theory. Section 3 introduces a structural representation of a street network and the related structural measures. Section 4 develops the principles of the selection algorithms and illustrates their application to a case study. Section 5 discusses some related work. Finally section 6 draws some conclusions.

Graph theory principles

In order to develop a structural representation of a street network, let's introduce some basic graph concepts. For a more complete introduction to graph theory, readers can refer for example to (Gross and Yellen 1999). A *graph G* consists of a finite set of vertices (or nodes) V and a finite set of edges (or links) E (note that we use vertices and nodes, and edges and links interchangeably). A graph is often denoted as $G(V,E)$ where V is the set of vertices, $V = \{v_1, v_2, ..., v_n\}$, and E is the set of edges, $E = \{v_i v_j\}$. For example, figure 1 shows a graph $G_1(V_1, E_1)$ with the set of vertices $V_1 = \{a, b, c, d, e, f, h, j, k\}$ and the set of edges $E_1 = \{\overline{ab}, \overline{ac}, \overline{ad}, \overline{cf}, \overline{ch}, \overline{de}, \overline{df}, \overline{dh}, \overline{ej}, \overline{ek}, \overline{fh}\}$. Let us remark that this simplified graph example is unweighted and undirected, and it is also connected as there is no isolated vertex.

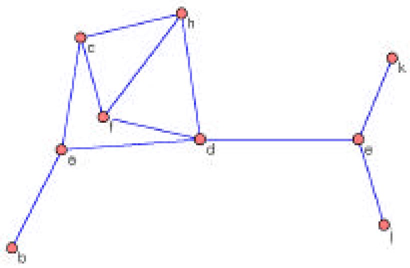


Figure 1: A simplified example of graph

We say that a graph H is a subgraph of a graph G if the vertices of H give a subset of the vertices of G. Conversely, if H is a subgraph of G, we say that G is a supergraph of H. For a vertex subset U of a given graph G, a subgraph whose vertices belong to U is said to be *induced* on the vertex subset U. Any two adjacent vertices v_i, v_j of G (i.e., $v_i, v_j \in E$) are said to be neighbours. The *neighbourhood* of a vertex v_i of a graph G, denoted $N_G(v_i)$, is the subgraph induced by the set of vertices consisting of v_i and all its neighbours, i.e., $N_G(v_i) = \{v_j \mid v_i v_j \in E, i \neq j\}$.

For computational purposes we represent a connected, undirected and unweighted (i.e. all links with a unit distance) graph by an adjacency matrix *R(G)*:

$$R(G) = \lfloor r_{ij} \rfloor_{n \times n} \text{ where } r_{ij} = \begin{cases} 1 & if\ v_i v_i \in E \\ 0 & otherwise \end{cases}$$

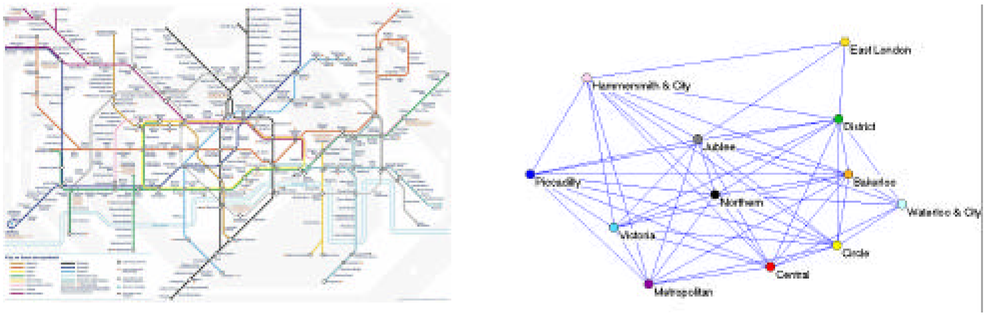
It should be noted that for an undirected graph G, this adjacency matrix *R(G)* is symmetric, i.e. $\forall r_{ij} \Rightarrow r_{ij} = r_{ji}$. Also all diagonal elements of *R(G)* are equal to zero so are not needed. Thus the lower or upper triangular matrix of *R(G)* is sufficient for a complete description of the graph G.

An approach to the structural representation of street networks

Given an urban system, the underlying street network can be considered as a structuring element for many other cartographic objects (e.g. built environment, electricity and gas networks) and socio-economical activities in the city. So reducing the complexity of an urban street network has many application interests. A street network has its own intrinsic logical and spatial structure that must be represented and retained in applying a scale reduction process.

We represent a street network using some basic graph theoretic principles; named streets (note that a named street is not a street segment but the entire named street considered as a basic modelling unit) are represented as nodes and street intersections as links of a graph. One can remark that any graph derived using such an approach is connected, i.e. one can reach any vertex of the graph from any vertex.

To illustrate this approach, let's consider the example shown in figure 2. To the left of the figure is the London underground map; while to the right is the corresponding connectivity graph. The derivation of the connectivity graph is based on the following rules: the underground lines give the nodes of the connectivity graph, two nodes of the derived graph (i.e. two underground lines) are connected if there is at least one direct link between these two nodes (i.e. a connection between these two underground lines). It should be noted that multiple intersections between pairs of connected underground lines are not represented in the connectivity graph.



a b

Figure 2: London underground map (a) and its connectivity graph (b)

We introduce two measures for the description of node status within a connectivity graph. *Connectivity* of a vertex v_i, denoted $s(v_i)$, is the number of vertices directly linked to this vertex, so a local measure. For a given graph *G*, the *connectivity* satisfies the following condition: $\sum_{i=1}^{n} s(v_i) = 2m$, where *m* is the total number of edges, and *n* is the total number of vertices of the graph *G*.

The *average path length* of a given vertex v_i, denoted $L(v_i)$, considers not only those directly connected vertices, but also those within a few steps, so a form of global measure when the number of steps considered is high. Given two vertices v_i v_j $\in$ *V*, let $d_{\min}(i, j)$ be the shortest distance between these two vertices. The *average path length* of a given vertex v_i is given by (*n* being the total number of vertices of the graph *G*):

$$L(v_i) = \frac{1}{n}\sum_{j=1}^{n} d_{\min}(i, j)$$, where n is the total number of vertices of the graph *G*.

Table 1: Two measures for the nodes of graph G_1

Node ID	Connectivity	Average path length
A	3	1,6667
B	1	2,4444
C	3	2,0000
D	4	1,3333
E	3	1,6667
F	3	1,7778
H	3	1,7778
J	1	2,4444
K	1	2,4444

The above two measures (*connectivity* and *average path length*) present respectively some local and global properties of each considered vertex within its connectivity graph. For illustration purpose, table 1 lists the two calculated measures for the graph G_1 shown in figure 1. We can remark that less connected nodes are less important from a structural point of than those well connected at the local level. From a global perspective, the average path length measures how each node connects to every other in the connectivity graph. This gives a sense to what extent any vertex is integrated or segregated to every other within a connectivity graph. The lower the value of that measure is, the more integrated the node is. This property can be illustrated in figure 3, where all nodes are arranged in terms of how far (shortest distance) every other node is from the two nodes: *a* and *d* respectively. We can observe that node *d* is better integrated to every other than node *a* is. Conversely we can remark that node *a* is relatively "far from" every other, while node *d* is "close to" every other. The connectivity structure of the graph is important in deriving a series of subgraphs which retain the main structure of initial graph. For instance, the node *d* should have a higher probability than *a* to be kept during the processing of the reduction scale algorithm, as it is better integrated to every other node at the global level, and also better connected to other nodes at the local level.

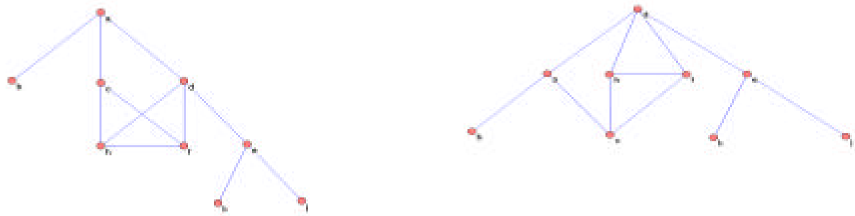


Figure 3: Respective integration/connection of nodes a and d

The above example illustrates how *connectivity* gives a sense on nodes' integration with immediate neighbors (local level), while the *average path length* reflects the way each node is integrated to its k-neighbors (global level). Overall a relevant structural approach to model generalization of urban street network should keep well-integrated nodes (or in other words to eliminate less integrated nodes). Logically we can remark that well-connected and -integrated streets tend to be more important from a structural point of view than those less connected and integrated.

Structural generalisation of an urban street network

We propose a generalization process based on the derivation of a series of subgraphs from the initial connectivity graph. In a related work we have illustrated the fact that street connectivity conforms to a power law distribution, that is, most named streets have low connectivity values while a few have high connectivity values (Jiang *et al*. 2001). This provides an interesting clue for the definition of a reduction scale algorithm based on those connectivity values at the local level (such an algorithm will be very much selective for average connectivity values). By contrast applying a filtering algorithm on the average path length reflects some global properties of the network.

The following sub-sections introduce these approaches: firstly with a connectivity-based selective algorithm, secondly with a selective algorithm proposed for average path length selection, finally with the application of a recursive algorithm that considers the hierarchical nature of a street network. The respective properties and advantages of these algorithms are discussed.

Connectivity-based generalisation

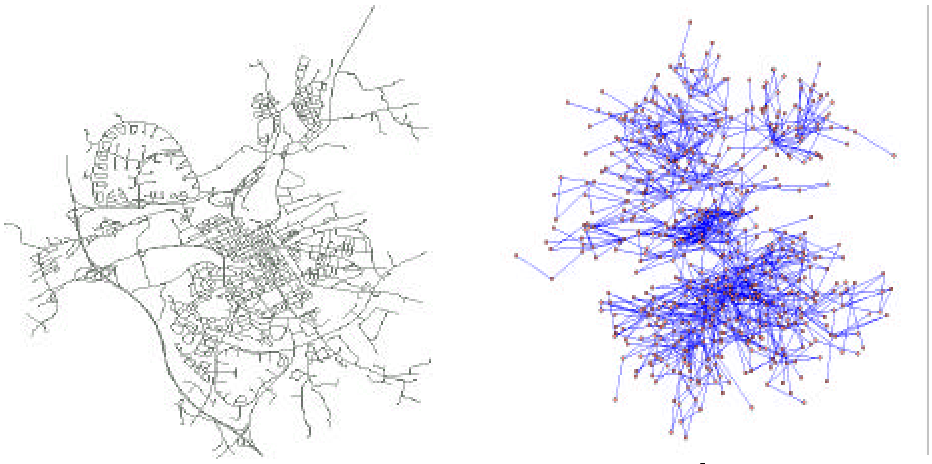
Let's use a Gävle city network for example to illustrate the different principles used for the selection of streets in a reduction scale process. This network involves 565 named streets, so 565 nodes in the connectivity graph (figure 4 a and b). It is composed of street central lines topologically interconnected, i.e. no isolated streets. A script determines how each given street intersects to every other, and then creates a connectivity graph. Informally the algorithm can be read as follows: for each street, check if it intersects another street, if yes, $r_{ij} = 1$, otherwise $r_{ij} = 0$. This algorithm is as follows:

```
Algorithm create-connectivity-graph
// V is a set of streets
// assume that street ID range from 0 to n-1
// R is the output matrix of the connectivity graph
Begin
```

$R \leftarrow [0...n-1][0...n-1]$ // *an empty matrix with all elements as zero*

$V' \leftarrow V$ // *target set of vertices*

```
for every v ∈ V do
  for every w ∈ V do
    if INTERSECTION(v, w) then
      R[v][w] ← 1
    else
      R[v][w] ← 0
    end if
  end for
end for
end create-connectivity-graph
```



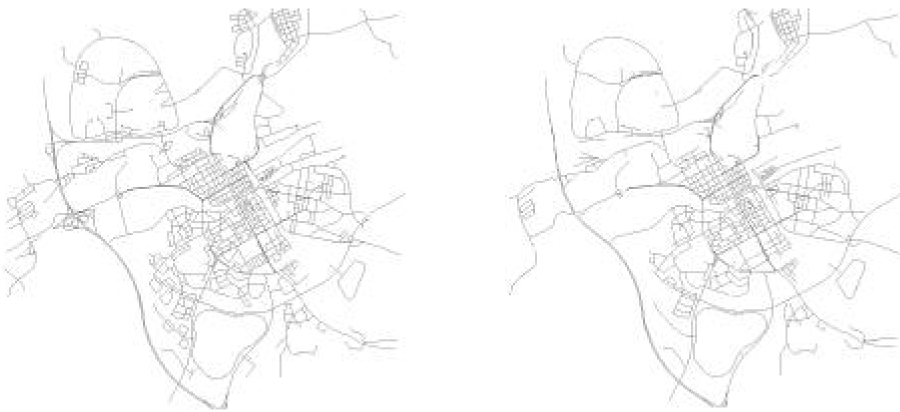
a b

Figure 4: Gävle street network (a), and its connectivity graph (b)

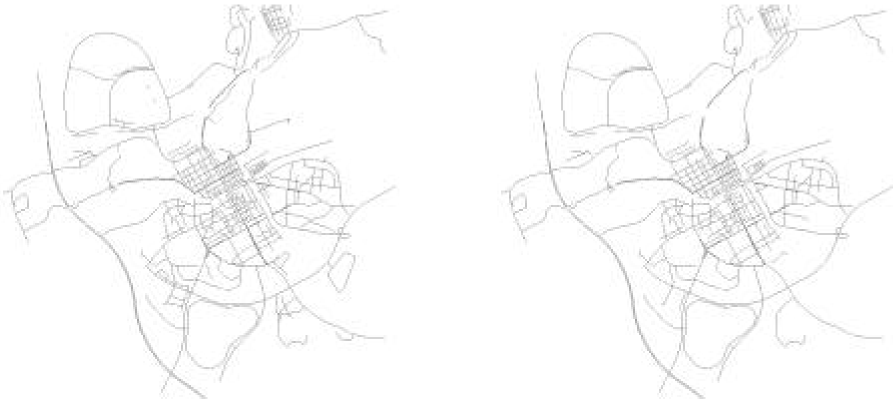
Based on this connectivity graph, or more specifically on the connectivity measure, we can start to select or eliminate streets for generalisation purposes. As mentioned in the previous section, well-connected streets tend to more important then less connected. Therefore the first rule for the selection is defined as follows:

"If a street connectivity is greater than a given threshold, then keep it; otherwise eliminate it."

For illustration purpose, figure 5 shows a series of generalised maps with threshold values respectively equal to 1, 2, 3 and 4.



(a) (b)



(c) (d)

Figure 5: Streets generalisation with connectivity values 1 (a), 2 (b), 3 (c) and 4 (d)

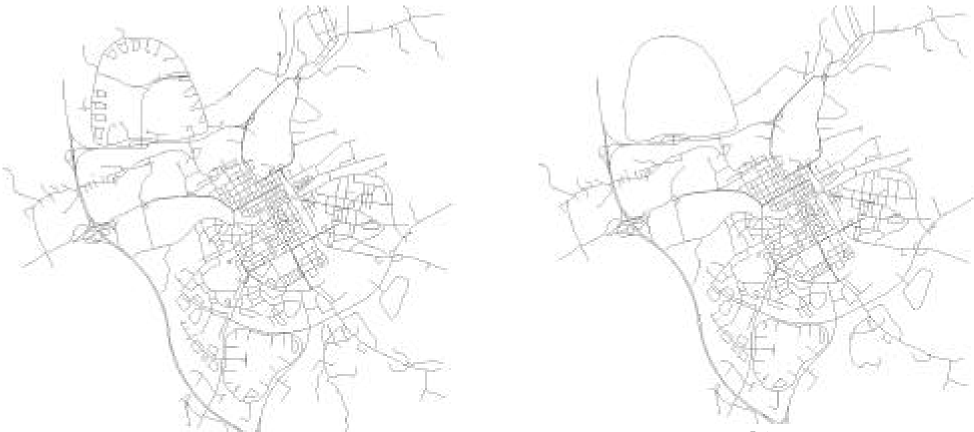
Average path length-based generalisation

Connectivity considers local streets "directly" connected, that is streets within a range of one step. On the other hand, average path length considers streets within k steps, and this reflects how a given street is integrated to every other within an urban street network. So an average path length-based algorithm selects those well-integrated streets. The rule for the selection of these streets is defined as follows:

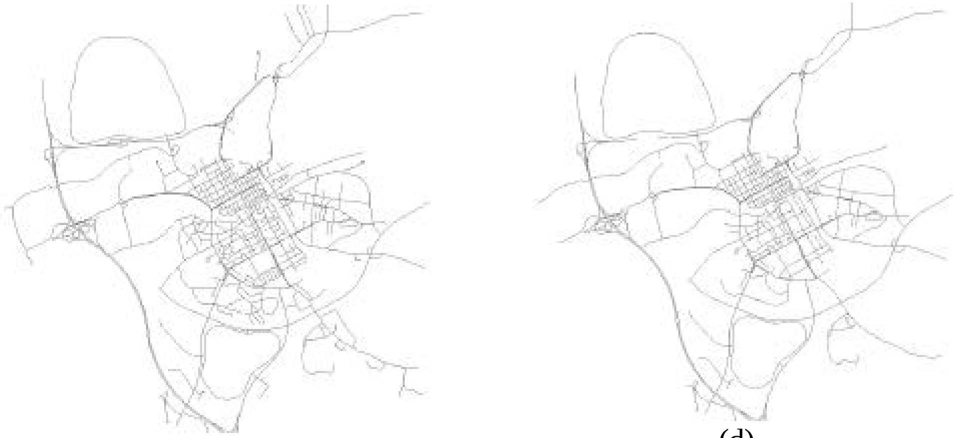
"If the average path length of a street is less than a given threshold, then keep it; otherwise eliminate it."

Figure 6 illustrates a series of generalised maps with the threshold values of average path length equal to 6.5, 6.25, 5.75, and 5.5 (respectively figures (a) (b) (c) and (d)).

It should be noted that in both figure 5 and 6, thresholds are defined for illustration purpose. End users can choose appropriate thresholds according to their particular objectives in applying such a reduction algorithm (this might be an exploratory and interactive process).



(a) (b)



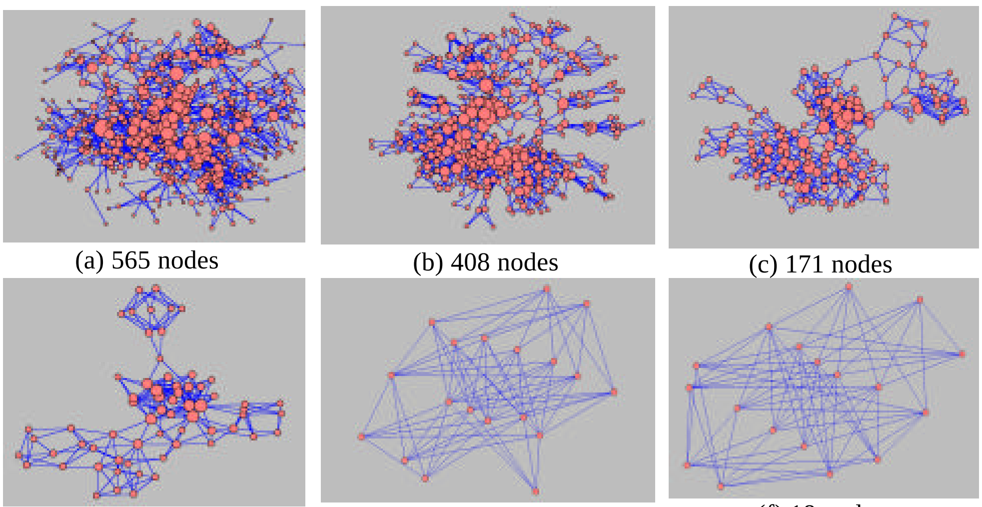
(c) (d)

Figure 6: Streets generalisation with threshold values of average path length equal to 6.5 (a), 6.25 (b), 5.75 (c) and 5.5 (d)

Hierarchy-based generalisation

Streets are hierarchically organised in terms of connectivity and average path length measures. To illustrate this, let's consider the connectivity graph whose nodes are represented by different size in terms of the magnitude of connectivity of each node (i.e. individual street mapped from the street network) (figure 7a). These figures display well-connected streets using larger node sizes, and less connected streets using smaller node sizes. These patterns illustrate the fact that these nodes are arranged at different levels of a hierarchy. So another way of generalising a street network can be based on such a hierarchical property. We introduce a recursive algorithm to reduce the number of nodes based on that hierarchical structure.

The proposed algorithm follows a recursive process. For example, if one set a connectivity threshold of 2, then in the generalised graph the minimum connectivity of nodes is equal to two as shown in figure 7(b) (so no nodes with a connectivity value of 1 are left). Similarly, one can then generalise this resulting graph with a connectivity threshold being equal to 3 (171 nodes left as in figure 7c), and then 4 (65 nodes left as in figure 7d), and then 5 (19 nodes left as in figure 7e) and then 6 (18 nodes left as in figure 7f). It should be noted that from this stage, an additional iteration of the algorithm gives no nodes left at all, because of the recursive nature of this approach. We can also remark that this algorithm retains the central structure of the city while the importance of outlying streets is diminished.



(a) 565 nodes (b) 408 nodes (c) 171 nodes

(d) 65 nodes (e) 19 nodes (f) 18 nodes

Figure 7: A hierarchy-based connectivity streets generalisation

Let us map the schematic graph in figure 7f into the street network. Figure 8 shows the most generalised network (represented as thicker lines) derived from the 18 resulting nodes of the scale reducing process. A cross-check of the roles of those resulting streets in the city of Gävle shows that these streets constitute the central structuring part of the city, and are most accessible in terms of transportation and commercial activities allocation. For example, the fours streets quoted in figure 8 namely Nygatan, Drottningatan, Kungsgatan and Rådmansgatan are most important commercial streets. Important landmarks such as central station, theatre, city hall, and central shopping mall are also located within this generalised street network.

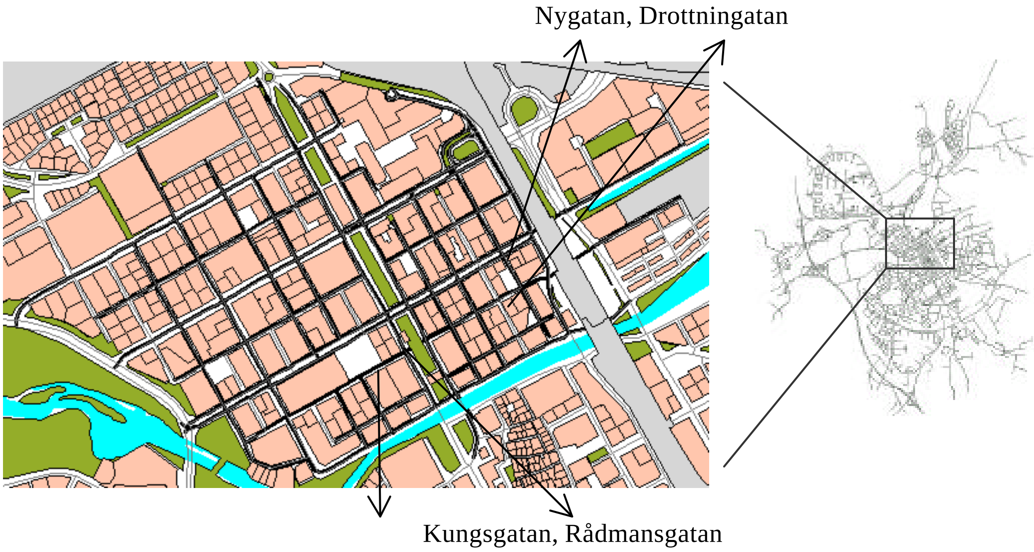


Figure 8: A final generalised map with 18 major streets in the centre of Gâvle

Discussion

Let us briefly discuss related work in the domains of cartographical generalisation and space syntax studies. In the cartography and GIS research community, one can make a distinction between two types of generalisation, namely model and graphic generalisation (Muller et al. 1995). Model generalisation is mainly oriented to data filtering in a scale reduction process, while graphic generalisation is more concerned with graphic representation or visualisation at the visual output level (Weibel 1995). These two generalisation approaches are closely related, often model generalisation being a pre-process of graphic generalisation. Combining the two permits not only to consider geometric simplification but also to integrate spatial structure factors in the generalisation process. Street network generalisation constitutes an important research challenge in cartographic generalisation, as it has also an influence on other cartographic object generalisation. There have been many research efforts since the appearance of the seminal Douglas-Peucker algorithm for line simplification (Douglas and Peucker 1973).

Recently, graph-based approaches have been investigated for linear object generalisation like street and hydrological networks. Mackaness and Beard (1993) discussed the potential of graph theory principles for derivation of information at the topological level to support generalisation of linear objects. They applied weighted graph, directed graph, and minimum spanning trees in the description of street and drainage networks, and derived some preliminary rules for generalisation process. Thomson and Richardson (1995) used the concept of minimum spanning tree in road network generalisation. More recently Kreveld and Peschier (1998) proposed a three-step approach to road network generalisation by considering basic geometric, topological and semantic requirements in line simplification.

Mackaness (1995) applied space syntax principles and demonstrated how they can be used to derive hierarchies of urban road networks. His study shows that street segment inter-connections (note herein street segments rather than named streets) and space syntax parameters can be used to illustrate the structure of an urban street network. Although no implementations of these principles have been achieved so far, this proposal illustrates the potential of space syntax for the structural analysis of an urban street network. Space syntax has also developed many computational solutions to the analysis of an urban street network. Space syntax studies often derive some local or global properties of a given urban network. For example, a long straight street tends to have many streets interconnected, thus it has a high connectivity; similarly, the same street tends to be well integrated to every other street with a shorter average path length. However, and to the best of our knowledge, space syntax principles have not been applied so far to generalise an urban street network while retaining its main structure.

We can remark that our approach to the structural representation bears some similarity to Richardson's (2000) approach based on human's spontaneously perceptual organisation (or grouping) on linear objects. She used a term 'stroke' to define the elementary units of a network based on movement continuity. Each stroke is actually a basic modelling unit, a similar concept to the one we adopt by considering named street as vertices of the connectivity graph. However, this approach that considers a cognitive-based graphic representation is different in essence from the structural and model-based generalisation we propose in this paper, and it is also not directly computable. For instance, and at the exception of regular and orthogonal networks, street networks may not appear to have such immediate and structuring visual properties.

Conclusion

An urban street network is a structuring component of the city so defining and implementing filtering algorithms that keep the main and central structure of an urban street network is of much interest for many urban applications and studies. This paper proposes a model generalisation approach for the selection of characteristic streets in an urban street network. It is based on the application of selective and flexible algorithms that considers both local and global structuring properties of named streets that correspond to basic functional elements in the city. The proposed algorithms are flexible as thresholds are user-defined and controlled giving thus an interactive solution to the application of such algorithms to a street network.

The case study presented in the paper shows how the structure of a street network is retained with subsequent filtering of streets. Our model and approach extend Mackaness (1995) proposal by an implementation of two model generalisation algorithms. This model generalisation can also be treated as a prerequisite for further generalisation processes such as line simplification or building blocks generalisation. Further work concerns the integration of weighted graph with the connectivity graph by considering a semantics classification of streets and the application of the method to other urban contexts.

References

AGENT (1998), *Constraint Analysis*, ESPRIT report, Department of Geography, University of Zurich.

Buttenfield B. P. and McMaster R. B. (1991), *Map Generalisation: Making Rules for Knowledge Representation*, Longman Scientific & Technical.

Douglas D. and Peucker T. (1973), Algorithms for the Reduction of the Number of Points Required to Represent a Digital Line or its Caricature, *The Canadian Cartographer*, Vol. 10, pp. 112 – 22.

Gross J. and Yellen J. (1999), *Graph Theory and Its Application*, CRC Press: London.

Hillier B. (1996), *Space is the Machine: A Configurational Theory of Architecture*, Cambridge University Press, Cambridge, UK.

Hillier B. and Hanson J. (1984), *The Social Logic of Space*, Cambridge University Press: Cambridge.

Hillier, B. (editor, 1997), *Proceedings, First International Symposium on Space Syntax*, University College London, London, 16-18 April, 1997.

Holanda F. (editor, 1999), *Proceedings, Second International Symposium on Space Syntax*, Universidade de Brasilia, Brasilia, 29 March-2 April 1999.

Jiang, B., Claramunt, C. and Batty, M. (1999), Geometric Accessibility and Geographic Information: extending desktop GIS to space syntax, *Computer Environment and Urban Systems*, Elsevier Science (Pub.), 23 (2), pp. 127-146.

Jiang B. Claramunt C. and Vasek V. (2001), Small World Patterns in Urban Street Networks, *Working paper at the University of Gävle*.

Kreveld M. V. and Peschier J. (1998), On the Automated Generalization of Road Network Maps, *GeoComputation'98*, available at http://divcom.otago.ac.nz/SIRC/webpages/Conferences/GeoComp/GeoComp98/geocomp98.htm

Mackaness W. A. and Beard M. K. (1993), Use of Graph Theory to Support Map Generalisation, *Cartography and Geographic Information Systems*, Vol. 20, pp. 210 – 221.

Mackaness W. A. (1995), Analysis of Urban Road Networks to Support Cartographic Generalization, *Cartography and Geographic Information Systems*, Vol. 22, pp. 306 – 316.

Muller J. C., Lagrange J. P. and Weibel R. (1995, eds.), *GIS and Generalization: Methodology and Practice*, Taylor and Francis: London.

Peponis J., Wineman J. and Bafna S. (editors, 2001), Proceedings, Third International Symposium on Space Syntax, Georgia Institute of Technology Atlanta, May 7-11, 2001.

Richardson D. (2000), Generalization of Road Networks, available at http://www.ccrs.nrcan.gc.ca/ccrs/tekrd/rd/apps/map/current/genrne.html

Thomson R. C. and Richardson D. E. (1995), A Graph Theory Approach to Road Network Generalisation, in: *Proceeding of the 17th International Cartographic Conference*, pp. 1871 – 1880.

Weibel R. (1995), Three Essential Building Blocks for Automated Generalisation, in: Muller J. C., Lagrange J. P. and Weibel R. (eds.), *GIS and Generalization: Methodology and Practice*, Taylor and Francis: London, pp. 56 - 69.