A GRAPH THEORY APPROACH TO ROAD NETWORK GENERALIZATION

Robert C Thomson & Dianne E Richardson
Canada Centre for Remote Sensing
MSD, 4th Floor, 588 Booth Street
Ottawa, Ontario K1A 0Y7, Canada
fax: (613) 947 1383

Abstract

The development of techniques for the automatic integration of remotely sensed data into a GIS environment is one focus of research at Canada Centre for Remote Sensing. One aspect of this effort is the automatic creation of databases of geographic data which can be reconfigured on demand to yield the information relevant to a given context, e.g. for map compilation. A theoretical framework is provided by a model developed at CCRS for automated spatial and thematic generalization, and techniques are being developed for automatic structuring, classification and coding of unstructured road network data into a suitable form supporting generalizations.

An appropriate classification of road data must take into account features such as surface type and number of lanes, and also functional aspects such as a road's relative importance in linking a given set of locations. Graph representations offer a convenient means of handling the topological and associated information describing a road network, and the use of graph theory in supporting network analysis and generalization is briefly reviewed. Graph theoretic techniques, such as the shortest path between network nodes and spanning trees, are then shown to provide a solution to the important problem of deriving measures of the functional relevance of network road segments, given a context defined in terms of a set of points of interest. Thus, for a given context, a set of rankings reflecting the importance of the segments is created which can serve as the basis for the attenuation of the network to any required degree for use in map density reduction and generalization. Spanning trees can be used in addition to maintain connectivity between destination points during attenuation.

Results from a prototype implementation of the network analysis system are presented. Preliminary tests indicate the effectiveness of the analyses: graph theoretic methods allow the efficient extraction and handling of the pertinent topological properties of a network, and as such naturally support generalization which aims to find the essential or representative characteristics of a data set for a given context.

1 Introduction

To date, most system-based developments in generalization provide techniques that emulate human cartographers. Consequently, although generalization techniques have been developed in systems environments, the execution of the routines usually reflects a cartographic perspective in that the final representation has been the priority rather than a consistent and logical treatment of the information. Often this implies the application of algorithms to individual features taken out of context. Typically, the processes invoked in the systems environment have usually consisted of graphically based techniques such as simplification, smoothing, displacement and symbolization, following the paradigm for manually derived generalization. Additional techniques have been developed to address selection and elimination, classification and aggregation. These latter forms of generalization, which result in a new representation, are often derived using attribute queries. According to the literature, little has been achieved which takes into consideration the context of the spatial representation in a systems environment. Certainly, theoretical models have been developed [3;15;16]; successful implementations, however, are not reported.

One of the difficulties in building an effective generalization system is the predominance of the conventionally based generalization paradigm, exemplified by numerous attempts at replicating rules derived from manual environments in a GIS or CIS. What is lacking is a refocussing of our perspective on what generalization achieves. By readdressing generalization from an information perspective, the focus in generalization models shifts from representational issues to the information and the database, and consequently developments should shift away from the cartographic perspective towards an information driven approach. This, however, places demands on operational mapping organizations for database building, data classification and subsequent coding.

Only recently have techniques been established to automatically classify, code, and build databases [19;20]. Since one impediment to generalization has been establishing a logical database, one approach at CCRS has been to identify techniques which allow users to automatically create the database without having to interactively code the features. Three different datasets are being used for experimentation and include hydrography, transportation networks, and built-up areas: features which represent a fundamental reference for most types of small scale maps, including topographic and thematic. The objective here is to load unclassified categories of features, automatically derive a series of classifications, automatically build a database, and subsequently generalize the data based on the information available and the user requirements, for any smaller scale of representation.

Three levels of modelling are involved. The first is the user level, which is important in identifying requirements and specifications for the final expected outcome. In each of the datatypes the concept of the user view is that it should allow users considerable flexibility in using the system, provide logical approaches in reducing data density, render consistently applied generalization techniques for particular scales of representation, and be information based.

The next level of development includes a generalization model applicable to the three datatypes, and a data model with which the generalization can operate to comply with anticipated user requirements. It should be pointed out, however, that establishing a system design based on anticipated user requirements restricts the versatility of the prototype development. In the generalization model this has been accommodated by establishing a generalization system referred to as GENSYSTEM that allows outcome based on anticipated requirements, yet also allows attenuation and amplification of any generalization to render greater flexibility than the prescribed designated scales. The data model component has been designed to support both the generalization operations and the user requirements. The data model for all three datatypes is based in the spatial domain on geometry and topology, and in the thematic domain on object oriented classification and aggregation hierarchies [18].

The lowest level of development is that of the logical data structures; this is the level at which spatial and attribute information are implemented in the systems environment. Ensuring that the data used are in a suitable format is an important aspect for developing automatic database building. At this level, topology must be checked for closure before the automatic classification, coding and database building can take place.

2 Objective

The objective of this work is to develop techniques for automatic structuring, classification and coding of transportation network data to support spatial and thematic generalization and density reduction. To realize this objective, the automatic structuring process must establish hierarchical structures, such as classification and aggregation hierarchies for the generalization processes to be successful. To create the hierarchical levels graph theoretic methods have been applied. Functional classifications have been derived based on access to and communication between destination points. These rely on derivations of shortest path spanning trees. These techniques provide the foundation within the spatial and attribute domain required for invoking contextually based generalization. The focus in these procedures is to render results that are consistent, logical, and contextually have relevance to the user.

3 Background

The basic topological structure of a network is captured by a graph: a set of vertices or nodes and a set of edges or arcs, where each edge connects two vertices. The valency of a vertex is the number of edges incident. A path is a sequence of edges, with consecutive edges sharing an end vertex. A graph is connected if a path can be found linking any pair of its vertices. A cycle is a path which leads from a vertex to the same vertex. A tree is a connected graph with no cycles. A spanning tree for a vertex set is a tree of edges in the graph in which a path can be found linking any pair of vertices. Clearly graphs can represent the topology of road networks in a natural way, with vertices corresponding to road junctions, intersections, dead-ends and locations of interest on the roads such as settlements, and edges corresponding to road segments between such points. The representation can be enhanced by associating cost values with each edge to represent road length or travel time.

Graph theory can be used in the generalization of networks in two different ways: to derive quantitative measures of topological or metrical properties of arcs and nodes [8;12], or to identify and represent important topological information which is needed for the effective application of generalization procedures [14]. The graph theoretic approach to quantifying the relative importance of the individual segments in river and stream networks (trees) is well known, notably the ordering methods of Horton [11], Strahler [25] and Shreve [24]. Horton ordering has been shown to be a suitable basis for the objective generalization of river networks where density reduction through feature elimination is required [22]; these ordering techniques have been used to support automated generalization of river networks [18].

Haggett [10] suggested a method of transferring the techniques for stream ordering to transport networks. Given a set of vertices of special interest, termed 'poles', the network is first partitioned into a set of 'polar networks' centred on the poles in a process analogous to finding Dirichlet regions in the plane. Within each polar network any cycle present is broken, producing a tree of paths rooted on each pole to which standard stream ordering techniques can be applied. Although this method appears to produce arc weights suitable for supporting the generalization of roads in a network which only has one dominant pole, the technique of network partitioning is not a generally acceptable solution to ranking the road segments in a network since partitioning splits major highways connecting poles. In contrast to such an approach, a method is required which recognizes and weights the major routes highly.

4 Method

Given a set of nodes of special interest in a road network, with associated arc costs giving distance or travel time information, and in the absence of other data, the road segments which are most likely to be used in travelling between those nodes can be reasonably assumed to be those that make up the shortest paths between the nodes. Further, by traversing each of these paths in turn, and by keeping a tally of the number of times each network arc is used, a set of weighting values can be built up for the arcs which directly reflects their relevance to this given node set. This observation leads to a process of deriving arc weights which can serve to support the generalization of road networks in different contexts defined by sets of nodes of special interest.

The complexity of the cat's cradle of paths which would be generated in this procedure can be tamed by recognizing that it can be decomposed into a set of tree structures: each set of shortest paths originating at one particular node and connecting it to the other nodes of interest constitutes a spanning tree for the node set called a shortest path spanning tree (SPST). The set of routes of interest can therefore be viewed as a stack of SPSTs. Further, the calculation of SPSTs in networks is a well known task with established algorithms.

The set of weights which will appropriately reflect the arc relevances in the given context can be built up from weights calculated for each SPST. Tracing each shortest path in the SPST from root to destination and noting each arc traversal provides the required weighting values for the tree arcs. This process is equivalent to calculating the Shreve magnitudes of the tree arcs [24].

When the nodes in the set of interest are not of equal interest, which could be modelled by associating different weighting values with each node, it would be desirable to to use this information in the arc weight calculations to give a more appropriate set of values. A simple way of doing this is to multiply the arc (Shreve) magnitudes found for each SPST by the weight associated with the root node of the tree before adding to the tallies for the arcs.

In the absence of any sort of thematic information the arc costs would be arc length values, calculated from the associated geometric data. When thematic information such as road surface type or number of road lanes is also available this can be used in conjunction with the length information to produce arc costs which give an indication of relative travel times and whose use in SPST calculation are more likely to give an accurate reflection of the relative use of arcs and so provide improved arc weights to support generalization. In this case all forms of network information - thematic, geometric and topological - have combined to produce a single measure. (Additional costs could be associated with the nodes encountered on a path, or on some measure of arc sinuosity [4] to refine the model further.)

The fact that spanning trees have been established for the node set of interest can also be exploited as a means of maintaining connectivity amongst the nodes during generalization and density reduction of the network: by ensuring the retention of the edges of a spanning tree during generalization no vertex becomes isolated from the rest of the network, and so the map conveys the notion of connectivity between features. The minimum cost spanning tree (MCST) has been proposed for this purpose [14]; the suggestion here is that one of the SPSTs already calculated should be used. The SPST rooted at the node of greatest relative importance - i.e. the most probable major source of traffic - seems the optimal choice, regardless of its cost: this SPST is likely to reflect the actual routes favoured between the major centre and the destinations and so it represents a very sensible set of connecting roads to retain. The option of using a SPST to maintain connectivity rather than the MCST clearly also has the advantage that it requires no additional calculation. The arcs of the spanning tree can be protected from deletion simply by setting their weights equal to the maximum value present after accumulation of the SPST data.

5 Implementation

The practical derivation of this sort of information requires the construction of a representation which combines thematic, spatial and topological information about the road network, and provides appropriate linkages between them. Procedures to test the ideas presented here have been implemented in POP-11, a flexible environment for experimentation and prototyping [2].

As described above, there is a natural correpondence between a road network and a graph structure. In modelling the network, however, to support the necessary associations between thematic, geometric and topological information, the basic arc-node structure needs some augmentation. In these implementations a distinction has been made between non-structural nodes (of valency 2) and structural nodes (of valency 1 or 3 or higher). Non-structural nodes may serve either to represent a point of interest, such as a settlement, or to mark a point where a thematic attribute of a road changes. Roads are modelled as sequences of 'sections' which are uniform in their thematic attributes, such as number of lanes or road surface type, and have no internal points of interest. Sequences of sections connecting at non-structural nodes make up 'arcs'. Arcs terminate at structural nodes (which may also be points of interest or attribute change).

Arcs, sections and nodes are given unique identifiers, which serve as keys to database records of associated geometric and thematic data. The arc-node and arc-section relationships are also stored in relational form, defining the topology and allowing the easy extraction of an abstract graph representation of the network for analysis. The thematic and geometric road information is associated with the sections using relational tables: this association allows calculation of section costs from which arc costs are calculated. In recording paths in the network it is only necessary to store at each node the identity of the preceding node in the path or the identity of the previous arc traversed, and paths can be recovered by a recursive process. Linkages are maintained between graph and network database so that, for example, derived attributes can be stored with appropriate associations, or geometric information found to allow cartographic representations to be built.

5.1 Examples

Figure 1 shows an example road network extracted from a 1:250,000 map sheet of Vancouver Island. Fourteen network vertices have been selected as representing hypothetical sites of special interest, and assigned weights to reflect their relative importance: these weights are indicated by the size of the symbols used to mark the sites. A set of relative 'navigability' values for the arcs in figure 1 were estimated heuristically using the section thematic information on number of lanes and road surface type: the values calculated are reflected in the shading used, with the 'faster' roads shown darker.

Figure 2 shows one SPST for the set of marked nodes, rooted at one of the nodes of maximum importance rating. In this case arc costs represented road distance alone, without regard to the thematic attributes. Shreve magnitudes were calculated for this tree and the relative values are indicated by the grey levels. The magnitudes in this case range from 1, for 'fingertip' arcs of the SPST, to 8; arcs outside the SPST have zero weight.

A series of SPSTs for the marked set of nodes were calculated, rooted on each node in turn. At first only arc length was used as the arc cost. The Shreve magnitudes were calculated for each SPST found, giving a weight value for each network arc. These weights were accumulated for the different SPSTs, giving a total weight value for each arc. In each case the value associated with the root node, representing that node's relative importance, was used to multiply the weights provided by the Shreve ordering. In figure 3(a) the total accumulated weights which result for each arc are indicated by relative greyscale. This procedure can be seen to have picked out the major routes between the designated points of interest and has produced a ranking of the network arcs according to their functional relevance in this context.

Figure 3(b) shows the result of using the SPST rooted at one of the more important nodes to ensure connectivity by boosting its arc values to the maximum weight value found in figure 3(a).

Figure 4(a) may be compared with figure 3(a). Here the process of SPST calculation has used arc costs which take into account both road length and road navigability, and so represent estimated relative travel times. This change can be seen to have produced a different set of overall arc weightings; in particular, in the lower right of the figure, a longer but faster connecting path now carries higher weighting than its shorter but slower alternative.

Figure 4(b) is based on 4(a): here the SPST rooted at one of the more important nodes has again been used to ensure connectivity by enhancing its arc values.

5.2 Further development

The above examples represent preliminary tests of this approach. Although results are already very promising there are other variations which merit examination. One area for experimentation is how weights are assigned to the arcs of the SPSTs. An interesting strategy to be investigated for assigning a path weight is to follow accessibility potential theory and take a weight which is proportional to the product of the weights (importance ratings) of the nodes being connected and inversely proportional to some power of the path length [9].

Problems of excessive computation times must also be anticipated when dealing with such graph-search algorithms, and simplifying strategies considered. The algorithm used for these test implementations is due to Dijkstra [7]. The basic form of this algorithm calculates the shortest paths from each vertex in a network to every other, and so may represent a lot of unnecessary computation when the set of nodes of interest is small in comparison with the total number of nodes in the network. In such cases it may be more efficient to use an algorithm such as the A* algorithm [17, ch.3] which is more suited to finding single shortest paths between pairs of network nodes, although such non-exhaustive methods can bring the risk of reaching non-optimal solutions.

The task of finding shortest paths in networks also arises, for example, in location-allocation analyses [6] and transportation studies, and suggestions for improving the performance times of path search methods can be found in those areas of the literature. The use of an algorithm which uses path search outward from both nodes being connected has been suggested [21]. Such an approach, where the optimal path from A to B is found by combining optimal paths from A to C and from C to B, could also be adapted to find sub-optimal paths, and so investigate alternative routings in the network. This technique has been described [13] in the context of planning problems, but has also been long known in speech template matching studies [23]. The use of sub-optimal paths is also of interest since it would be expected to produce non-zero weightings for more of the network arcs than the current method, and so support a smoother rate of density reduction.

Of interest also are hierarchical path searching methods which could be considered if there were adequate thematic information present to categorize the roads into a hierarchy of 'navigability' classes [5;21]. Further, there have been investigations into the use of parallel algorithms to muffle the combinatorial explosion of path-finding [1]. Clearly such developments in location-allocation studies and network analysis are of direct relevance to the implementation of the generalization methods presented here.

6 Conclusions

This paper has described current work in the development of automatic structuring, classification and coding of road network data to support spatial and thematic generalization and density reduction. In achieving this the need to previously code arcs or line segments has been eliminated. The processes discussed result in automatic database derivations based on points of interest expressed by the user. It is an entirely flexible process which can provide output consistent with topographic representations, and yet which also provides an environment which can be altered for contextual representations.

Graph representations offer a natural and convenient means of handling the topological and associated information, and the use of graph theory in supporting network analysis and generalization was briefly reviewed. Graph theoretic techniques were then shown to provide a solution to the important problem of deriving measures of the functional relevance of network arcs given a context defined in terms of a set of points of interest, possibly with associated values reflecting their relative importance. The derived arc weights give a partial ordering of the arcs which can be used as a basis for the attenuation of the network for use in map density reduction and generalization. This method also illustrates another role of graph theory in generalization since the shortest path spanning trees whose calculation underlies the arc weight derivation can be used in addition to maintain connectivity between destination points during attenuation.

Results of prototype implementations show that this method of road segment classification can successfully integrate thematic, geometric and topological information into a single measure appropriate for the generalization of the network. Further developments of the work, and its relation to other fields have been discussed.

7 References

[1] Armstrong, M. P. & P. J. Densham, 1992. "Domain decomposition for parallel processing of spatial problems", *Computers, Environment, and Urban Systems*, **16**(6), pp. 497-513.

[2] Barrett, R., A. Ramsay & A. Sloman, 1985. *POP-11, a Practical Language for Artificial Intelligence*, Chichester, UK: Ellis Horwood.

[3] Brassel, K. E. & R. Weibel, 1988. "A review and conceptual framework of automated map generalization", *International Journal of Geographical Information Systems*, **2**(3), pp. 229-244.

[4] Buttenfield, B. P., 1991. "A rule for describing line feature geometry", in B. P. Buttenfield & R. B. McMaster (eds.), *Map Generalization*, pp. 150-171, Harlow, UK: Longman.

[5] Car, A. & A. U. Frank, 1994. "General principles of hierarchical spatial reasoning - the case of wayfinding", *Proceedings of the 6th International Symposium on Spatial Data Handling* (5-9 Sept., Edinburgh), pp. 646-664.

[6] Densham, P. J. & M. P. Armstrong, 1993. "Supporting visual interactive locational analysis using multiple abstracted topological structures", *Proceedings of Auto Carto 11* (30 Oct - 1 Nov 93, Minneapolis, MN), pp. 12-22.

[7] Dijkstra, E. W., 1959. "A note on two problems in connexion with graphs", *Numerische Mathematik*, **1**, pp. 269-271.

[8] Dramowicz, K., 1994. "Application of graph theory to network evaluation: theoretical and practical issues", *Proc. 6th Canadian Conf. on GIS* (Ottawa, 6-10 June), **2**, pp. 1657-1668.

[9] Geertman, S. C. M. & J. R. Ritsema van Eck, 1995. "GIS and models of accessibility potential: an application in planning", *Int. Journal of Geographical Information Systems*, **9**, pp. 67-80.

[10] Haggett, P., 1967. "Network models in geography", in R.J.Chorley & P.Haggett (eds), *Models in Geography*, pp. 609-668, London: Methuen.

[11] Horton, R. E., 1945. "Erosional development of streams and their drainage basins", *Bulletin of the Geological Society of America*, **56**, pp. 275-370.

[12] James, G. A., A. D. Cliff, P. Haggett & J. K. Ord, 1970. "Some discrete distributions for graphs with applications to regional transport networks", *Geografiska Annaler*, **52B**, pp.14-21.

[13] Lombard, K. & R. L. Church, 1993. "The gateway shortest path problem: generating alternative routes for a corridor location problem", *Geographical Systems*, **1**, pp. 25-45.

[14] Mackaness, W. A. & M. K. Beard, 1993. "Use of graph theory to support map generalization", *Cartography and Geographic Information Systems*, **20**(4), pp. 210-221.

[15] McMaster, R. B. & K. S. Shea, 1988. "Cartographic generalization in a digital environment: a framework for implementation in a geographic information system", *Proceedings GIS/LIS '88* (San Antonio, TX), **1**, pp. 240-249.

[16] Nickerson, B. G. & H. Freeman, 1986. "Development of a rule-based system for automatic map generalization", *Proceedings of the 2nd International Symposium on Spatial Data Handling* (July, Seattle), pp. 537-556.

[17] Rich, E., 1983. *Artificial Intelligence*, New York: McGraw-Hill.

[18] Richardson, D. E., 1993. *Automatic Spatial and Thematic Generalization using a Context Transformation Model*, Doctoral Dissertation, Wageningen Agricultural University, NL.

[19] Richardson, D. E., 1994. "Generalization of spatial and thematic data using inheritance and classification and aggregation hierarchies", *Proceedings of the 6th International Symposium on Spatial Data Handling* (5-9 Sept., Edinburgh), **2**, pp. 957-972.

[20] Richardson, D. E., 1995. "Generalization of spatial and thematic data using inheritance and classification and aggregation hierarchies", *Proc. GIS AM/FM ASIA'95* (August, Bangkok).

[21] Ritsema van Eck, J. R. & T. de Jong, 1990. "Adapting datastructures and algorithms for faster transport network computations", *Proceedings of the 4th International Symposium on Spatial Data Handling* (23-27 July, Zurich), pp. 295-304.

[22] Rusak Mazur, E. & H. W. Castner, 1990. "Horton's ordering scheme and the generalisation of river networks", *The Cartographic Journal*, **27**, pp. 104-112.

[23] Russell, M. J., R. K. Moore & M. J. Tomlinson, 1986. "Dynamic programming and statistical modelling in automatic speech recognition", *Journal of Operational Research*, **37**(1), pp. 21-30.

[24] Shreve, R. L., 1966. "Statistical law of stream numbers", *Journal of Geology*, **74**, pp. 17-37.

[25] Strahler, A. N., 1957. "Quantitative analysis of watershed geomorphology", *Trans. of the American Geophysical Union*, **38**, pp. 913-920.

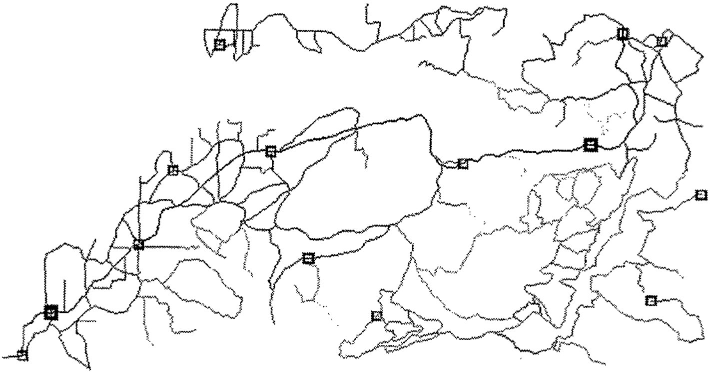

Figure 1: Example road network.
Derived navigability ratings are shown by greyscale and nodes of special interest are highlighted

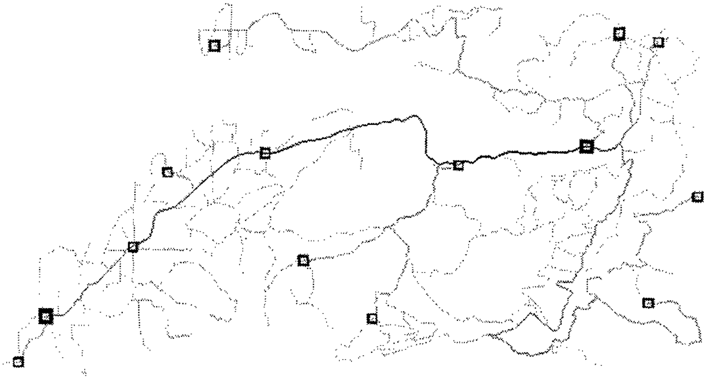

Figure 2: One shortest path spanning tree for selected nodeset.
SPST is calculated using length as arc cost; Shreve magnitudes of arcs are shown by greyscale

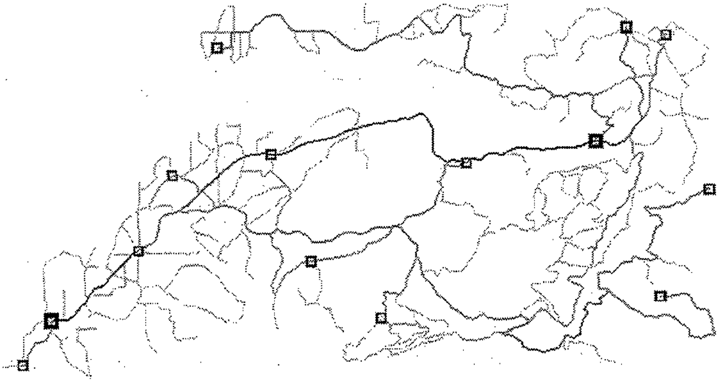

Figure 3(a): Arc weights for network (path finding uses length as cost).

Arc weights are derived by combining the SPSTs rooted at each node in the nodeset, with length used as arc cost and arc weights multiplied by root node weighting factor before summation

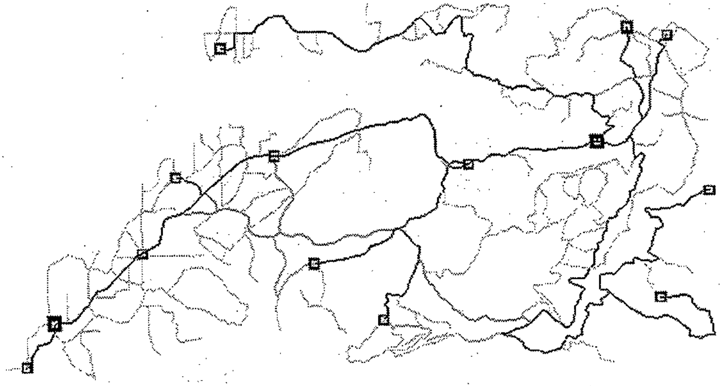

Figure 3(b): Arc weights with single spanning tree enhanced.

These weights are those of figure 3(a) with the SPST rooted at a major node enhanced to retain connectivity during generalization

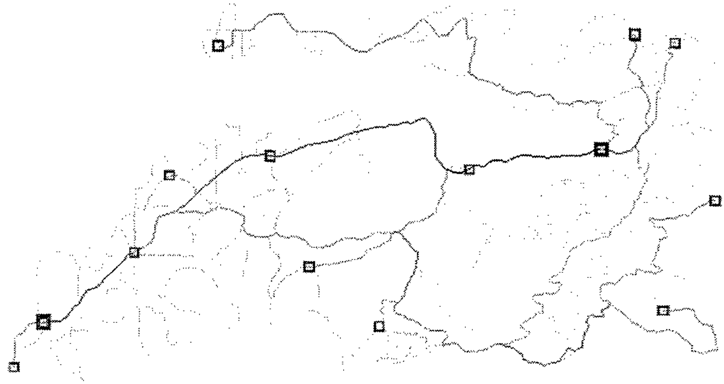

Figure 4(a): Arc weights for road network (path finding uses travel time as cost).

Arc weights are derived by combining the SPSTs rooted at each node in the nodeset, with estimated travel time used as arc cost and arc weights multiplied by root node weighting factor before summation

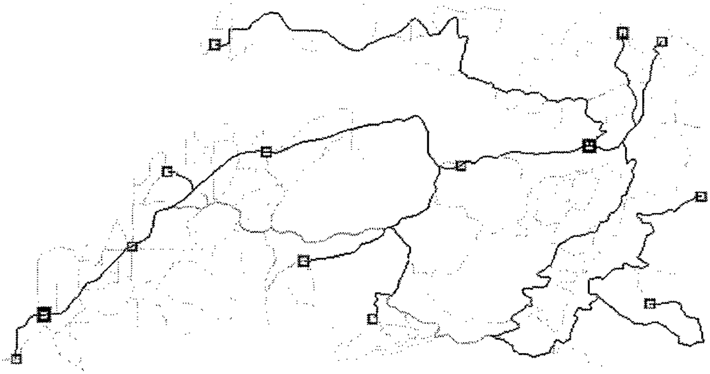

Figure 4(b): Arc weights with single spanning tree enhanced.

These weights are those of figure 4(a) with the SPST rooted at a major node enhanced to retain connectivity during generalization