PROGRESSIVE REPRESENTATION AND PROGRESSIVE GENERALIZATION OF STREET NETWORK VECTOR DATA

J. Tian^{ab} *

^aSchool of Resource and Environment Science, Wuhan University, 129 Luoyu Road, 430079Wuhan, China -(yutaka-2010@163.com)

^bKey Laboratory of Geographic Information System, Ministry of Education, Wuhan University, 129 Luoyu Road, 430079Wuhan, China

KEY WORDS: Cartography, Generalization, Representation, Model, Data Structures, Algorithms

ABSTRACT:

Street network vector data is represented as Regular Closed Network. The combinatorial structure of the Regular Closed Network is modelled through abstract cell complex. Two abstraction operators and two refinement updates are cited to define the street elimination operator and street refinement operator. Two alternative progressive representation models for street network vector data are proposed. Topological consistency is implicit in the progressive representation models. The ordered generalization tree structure can be seen as a form of data enrichment. The definition, construction and branch-pruning of this structure are introduced. The progressive generalization algorithm based on this structure is proposed. This algorithm brings together the map generalization knowledge and the specific data structure, in such a way that the generalization process can be implemented intelligently.

1. INTRODUCTION

In recent years, the progressive transmission of vector data has become a hot issue. Bertolotto et al (2001) first presented the concept model of progressive transmission of vector map data based on a set of generalization operators. Buttenfiled (2002) implemented an algorithm based on modified strip tree structure and RDP simplification algorithm for line transmission. Ai et al (2004) presented a Changes Accumulation Model for the progressive transmission of polygons. GAP-face tree and GAPedge forest structure are presented in van Oosterom (2005). Yang et al (2005, 2007) presented a multi-resolution vector data model for rapid and efficient transmission based on repeated elimination of point algorithm (Visvalingam et al 1993). It is clear from the research reported above that the multi-resolution representation model and the corresponding generalization method for constructing the representation are two key issues involved in the progressive transmission of vector map data.

Abstract cell complexes (Kovalevsk 1989) have been used to model the combinatorial structure of vector maps by several researchers (Puppo 1995, Bertolotto 1998, Bertolotto et al 2001, Viaña et al 2006). There exist seven functions, called generalization operators or abstraction operators and seven corresponding inverse functions called inverse of generalization operators or refinement updates (Bertolotto 1998, Bertolotto 2001, Viaña et al 2006). Topological consistency is implicit in multiple representation models on the basis of such operators (Bertolotto 1998).

Several researchers have presented one type of multi-resolution model, which consider the spatial representation from one resolution to another as execution of some operators with a specific sequence (Puppo 1995, Bertolotto 1998, Bertolotto et al 2001, Viaña et al 2006).Several researchers have presented another type of multi-resolution model, which consider the spatial representation from one resolution to another resolution as an accumulation of the set of changes with a specific sequence (Ai et al 2004, Yang et al 2007). A sequence of operators produces a sequence of changes. Thus, these two type models are intrinsically the same.

Street network generalization is an important generalization operation, and may be a prerequisite to other generalization operations such as building generalization. It is a complex task which requires a good understanding of the geometrical, topological and semantic aspects of the street network. Generalization of street networks can be based on different principles: (1) Filtering-selection method based on space syntax. Mackaness (1995) applied "line of sight" pattern analysis algorithm to urban road networks. Jiang et al (2004a) proposed a novel generalization model for selecting characteristic streets using graph principles where vertices represented named streets and links represented street intersections. (2) Importancemodelling method. Thomson et al (1995) used a graph-theoretic approach based on the concept of minimum spanning trees to select important street segment. Richardson et al (1996) presented a technique for determining and quantifying the functional importance of road segments according to a set of source and destination points that reflect the user's context. Morrist et al (1997) used the multi-agent model to calculate the importance of streets based on amount of 'street use'. (3) strokeordering method. Thomson et al (1999) showed how the 'good continuation' grouping principle could serve as the basis for analyzing a road network into a set of liner elements (strokes). They ordered the strokes to reflect their relative importance in the network. Edwards et al (2000) used the strokes as one tool and the areal dual of the network as a second structure to perform the generalization. Zhang (2004) presented a new stroke ordering method based on connection analysis. Hu et al (2007) presented the definition of mesh density to support stroke ordering. (4) Intelligent method to support the generalization process. Peng et al (1996) introduced an

^{*} Tian Jing. e-mail: yutaka-2010@163.com; phone: 86-13628636229

intelligent approach to the generalization of urban road network based on dynamic decision tree structure. Tian et al (2007) did similar work follow Peng's idea by improve the tree construction process and branch-clipping rules. Jiang et al (2004b) proposed an approach to selection of important streets from a network based on the technique of SOM. Li et al (2004) acquired the knowledge of road elimination through the analysis on Hong Kong topographic map series. ⁽⁵⁾Information theoretic approach. BjØrke (2003) presented a new generalization algorithm that computes how many roads can be presented in a map by maximizing the amount of useful information.

"Progressive Representation" was introduced by Hoppe (1998) for the triangulated meshes. It has been pointed out by Betolotto et al (2001) that the progressive mesh scheme proposed by Hoppe (1998) presented a fundamental milestone for progressive transmission. Similar to the idea of "Continuous Generalization" (Sester et al 2004), "Progressive Generalization" is a generalization strategy, which generates a set of consecutive generalization results. Every generalization result is feasible and the current generalization result is the basis of next generalization step. Progressive representation of vector map data can be achieved by progressive generalization. This paper introduces two alternative progressive representation models for street network. Topological consistency is implicit in the progressive representation models. A new generalization algorithm based on the ordered generalization tree structure to support the progressive representation model is proposed. This

algorithm brings together the map generalization knowledge and the data structure, in such a way that the generalization process can be implemented intelligently. It guarantees that every generalization result is feasible.

The rest of this paper is structured as follows. Section 2 introduces two alternative progressive representation models for street network based on the theory of plane Euclidean graph and abstract complex cell. Section 3 presents a new generalization algorithm based on the ordered generalization tree structure to support the progressive representation model. Finally, section 4 concludes the paper and offers subjects for future work.

2. PROGRESSIVE REPRESENATION OF STREET NETWORK

2.1 Street Network

A street network is composed of a number of entire streets, blocks and street intersections. The entire street satisfies specific continuity rules and usually consists of one or multiple street segments. The principle of continuity proposed by Jiang et al (2004a) is the street name: two different arcs of the original street network are assigned the same street identity if they shared the same street name. Porta et al (2006) presented Intersection Continuity Negotiation model based on the preference to go straight at intersections. Streets can be divided into the major street and the minor street according to Wang et al (1993). In this work named-street approach is used in priority to define the entire street, namely different street segment are assigned the same street identity if they shared the same street name. When street name are unavailable, the approach proposed by Thomson et al (1999) is applied with one improvement, namely major street segment and minor street segment can not group together. The block is the building area

surrounded by entire streets or street segments. (street and entire street will be used interchangeably hereafter in the paper). **2.2 Regular Closed Network**

A Regular Closed Network (*RCN*) is induced by a Plane Euclidean Graph (De Floriani et al 1993) with some restrictions.

Definition 1: A *RCN* is a triple (*V*, *E*, *F*), where:

V, called the set of vertices, is a set of points in the plane. $\forall v \in V$, degree(v) ≥ 2 . The boundary of a vertex is empty.

E, called the set of edges, is a set of polygonal chains having their endpoints in *V*, and any two edges of *E* never intersect except at their endpoints. A polygonal chain of *E* is defined as a sequences of points $e = (v, p_1, \dots, p_k, w)$, where $v, w \in V$ and $p_1, \dots, p_k \notin V$. The boundary of an edge is formed by vertices of *V* which are bounding it.

F, called the set of faces, is a set of polygons bounded by edges of *E*. The boundary of a face is formed by the vertices of *V* and edges of *E* which are bounding it. There exists one infinite face. Figure 1 is a typical example of the *SN*. f0 is the infinite face. Street network vector data can be represented as Regular Closed Network. v represents the street intersection, *e* represents the street or street segment and *f* represents the block.



Figure 1. A typical RCN.

2.3 Abstract Cell Complex

An Abstract Cell Complex (*ACC*), defined by Kovalevsk (1989), can capture the combinatorial structure of the *RCN*.

Definition 2: An *ACC* is a triple (C, \prec, \dim) , where:

C, called the set of cells, is a finite set.

 \prec , called the bounding relation, is a strict partial ordering on the elements of C.

dim: $C \rightarrow N$, called the dimension function, is such that

$$\gamma \prec \gamma' \Longrightarrow \dim(\gamma) < \dim(\gamma'), \ \forall \gamma, \gamma' \in C \text{ and } \gamma \neq \gamma'.$$

Proposition 1: RCN = (V, E, F) is a Regular Closed Network.

Let: (1)
$$C = V \bigcup E \bigcup F$$
;

(2) \prec is defined as follows:

 $\gamma \prec \gamma' \Leftrightarrow \gamma$ belongs to the boundary of γ'

(3) dim function is defined as follows:

$$\dim(\gamma) = \begin{cases} 0 & \text{if } \gamma \in V \\ 1 & \text{if } \gamma \in E \\ 2 & \text{if } \gamma \in F \end{cases}$$

Then, the (C, \prec, \dim) is an ACC.

2.4 Operators

Street Elimination Operator (*SEO*) is defined by composing two abstraction operators and Street Refinement Update (*SRU*) is defined by composing two refinement updates.

The two abstraction operators (Bertolotto 1998, Bertolotto 2001, Viaña et al 2006) are described in the following:

(1) Edge-merge: fusing of two edges sharing a vertex into one edge. $em : \{(v, e, e'), \prec\} \rightarrow \{e''\}$, with $v \prec e$ and $v \prec e'$.

(2) Face-merge: fusing of two faces sharing an bounding edge into one face. $fm: \{(e, f, f'), \prec\} \rightarrow \{f''\}$, with $e \prec f$ and $e \prec f'$.

The two refinement updates (Bertolotto 1998, Bertolotto 2001, Viaña et al 2006)are described in the following:

(1) Edge-split: splitting of an edge into two edges sharing an vertex. $es: \{e''\} \rightarrow \{(v, e, e')\}$, with $v \prec e$ and $v \prec e'$.

(2) Face-split: splitting of a face into two faces sharing an bounding edge. $fs: \{f''\} \rightarrow \{(e, f, f')\}$, with $e \prec f$ and $e \prec f'$.

Definition 3: A Street Elimination Operator (*SEO*) is composed of a set of Edge-merge and a set of Face-merge with a specific sequence. *SEO* eliminates a street segment or an entire street. A Street Refinement Update (*SRU*) is composed of a set of Edgesplit and a set of Face-split a specific sequence. *SRU* reconstructs a street segment or an entire street.

Figure 2 shows an example of *SEO* and *SRU*, suppose e27 is an entire street. *SEO*: fusing *f*6 and *f*8 into *f*12 by *fm*, then fusion e24 and e25 into e30 and fusion e21 and e22 into e31 by *em*. *SRU*: splitting e30 into e24 and e25 at v15 and splitting e31 into e21 and e22 at v18 by *es*, then splitting *f*12into *f*6 and *f*8at e27 by *fs*.



Figure 2. An example of SEO and SRU.

2.5 Progressive Representation Model

The change between two representations is defined as one entire street. The initial street network SN0 can be simplified into a coarser street network SNi by applying a sequence of m successive SEO and i street are eliminated.

Definition 4: two alternative progressive representation models of the street network can be defined as:

(1) $SN_0 = (SN_i, (SRU_1, SRU_2 \dots SRU_m), +)$

Where: SNi is the *ith* coarser representation of SN_0 , $(SRU_1, SRU_2 \dots SRU_m)$ is a set of SRU, + represents the execution sequence of street refinement updates.

(2) $SN_0 = (SNi, (street_1, street_2 \dots street_i), +)$

Where: SNi is the *ith* coarser representation of SN_0 , (*street*1, *street*2.....*streeti*) is a set of eliminated street, + represents the eliminating sequence of streets.

3. PROGRESSIVE GENERALIZATION OF STREET NETWORK

Three experience rules are defined to guide the generalization process:

Rule1: Street network boundaries must not be broken.

Rule2: The entire street will be eliminated if one of the segments of it is eliminated.

Rule3: When a small block needs to be merged, consider its smallest neighbour block first.

3.1 Ordered Generalization Tree Structure

Auxiliary data structures which are useful to represent contextual topological relation may provide additional support to generalization. They are essential for the automation of generalization process, as they give specific information on data organization which allows a better understanding of geographical meaning and contextual analysis. Peng et al (1996) introduced a dynamic decision tree structure to support urban road network. This paper introduces an ordered generalization tree structure. The notion of structure includes the definition, construction and branch-pruning. The ordered generalization tree structure can be seen as a form of data enrichment.

3.1.1 Definition: Definition 5: An Ordered Generalization Tree OGT is a quaternion (N, R, B, L), where:

N, called the set of nodes, represents a set of blocks of which the area is smaller than a given value.

R, called root, is a specific node representing a block.

B, called the set of branches, represents a set of common streets of the two blocks which are adjacent to each other.

L, called the set of leaves, represents a set of specific blocks. There exist three types of leaves. I-type leaf: a leaf links to its parent by preserved street or street network boundary. II-type leaf: a leaf links to its parent by a street segment. III-type leaf: represent a block of which the area is larger than or equal to a given value. If a leaf is both I-type and II-type leaf, it is I-type leaf. If a leaf is both II-type and III-type leaf, it is I-type leaf.

3.1.2 Construction: The construction of the OGT is to represent the local structure of street network. Table 1 shows the construction algorithm. The input is the root R and the output is the OGT of R. $N_{current}$ is the node current considered.

Algorithm: Construction_OGT
Input: root R
Output : OGT of R
$N_{current} \leftarrow R$
Find all neighbours of $N_{current}$ by the use of topology. The
result is block_set.
If $(N_{current} \text{ is not the root})$
Delete the parent of N _{current} from block_set.
End if
Sort the block set ascending according to the area of block.
For each block in block set
Construct the branch linked to $N_{current}$, check the block's
condition and set it for nodes and leaves.
End for
If (all the blocks in block_set have been treated)
If $(N_{current} is the root)$
Exit
End if
Else
If (The son N_{son} of $N_{current}$ is a node)
Construction_OGT (N _{son})
End if
End if

Table 1. OGT construction algorithm

In figure 1, assume that blocks f5, f6 and f8 are smaller than a given value and Area (f6) <Area (f5) < Area (f8). Figure 3 shows the OGT of f6. f8 and f5 are nodes. f0 is a I-type leaf. f2, f3, f7, f9 are II-type leaf.



Figure 3. OGT of f6

3.1.3 Branch-pruning: The branch-pruning of the OGT is to determine whether the branch is deleted or not. The recursive tree searching order was adopted. The input is the root R of the OGT and the output is the branch deleted.

Algorithm: Branch-pruning
Input: R of OGT
Output: deleted_branch B
$N_{current} \leftarrow R$
While (not all the sons of N _{current} are leaves)
Move to the first node son N _{son}
$N_{current} \leftarrow N_{son}$
End while
If (all the sons of $N_{current}$ are leaves)
Backtrack to the parent N_{parent} of $N_{current}$
If $(N_{parent} == NULL)$
If (exist II-type leaf in the sons of $N_{current}$)
$B \leftarrow Delete$ the branch linked between $N_{current}$ and
the last II-type leaf
Exit
Else if (exist III-type leaf in the sons of $N_{current}$)
$B \leftarrow Delete$ the branch linked between $N_{current}$ and the
first III-type leaf
Exit
End if
Else
$B \leftarrow Delete$ the branch linked between $N_{current}$ and its
parent
Exit
End if
End if

Table 2. OGT branch-pruning algorithm

In figure 3, branch *e27* is deleted according to the algorithm. The searching order is $f6 \rightarrow f8 \rightarrow f7 \rightarrow f8$.

3.2 Generalization Algorithm

The importance of block is calculated as: $Importance(b_i) = Area(b_i) / \sum Length(b_i)$, where: $Area(b_i)$ is area of the block b_i and $\sum Length(b_i)$ is the total length of non-preserved streets or street segments surrounding the blocks b_i . This parameter is similar to the concept "mesh density" (Hu et al 2007) and reflects the local density of the street network. Table 3 shows the basic algorithm. The input is block area threshold and the output is street eliminated. *Calculate_Importance* calculates the importance of the block. Update() is a function that updates a new state of the street network.

Algorithm: Progressive_Generalization_Street Network
Input: area_threshold
Output: street_list
Improvement $\leftarrow Ture$
Do
<i>If</i> (Not Exist block_area < area_threshold)
Improvement \leftarrow False
Else
block_set ← Search_Block (area_threshold)
End if
<i>For</i> each block in block_set
Calculate_Importance (block)
End for

OGT←Construction_OGT (min_importance_block) deleted_branch ←Branch-pruning (R of OGT) Eliminate the corresponding street or street segment If (deleted_branch is one segment of the street) Delete other segment of the street End if street_list ←Delete_Street Update() While(Improvement)

Table 3. Progressive generalization algorithm

Figure 4 presents some of the results in the experiment. The data used in this experiment comes from the Compilation specifications for 1:25000 1:50000 topographic maps (GB 12343-90, China), consisting of 32 streets and 40 blocks. The area threshold is set to be 5000 units. Figure 4a describes the original street network, while Figure 4b-d presents three generalized maps corresponding to 5th, 8th and final step. Table 4 shows the corresponding data of generalization process.



Figure 4. Experiment results

Step	Total Length	Street Amount	Eliminate StreetID	Block Amount	Minimum Area	Minimum Importance
1	7293.88	32	30	40	2574.99	12.281080
2	7231.42	31	25	39	2748.14	12.302701
3	7137.72	30	6	38	2748.14	12.684876
4	7053.42	29	8	37	2748.14	12.990365
5	6994.72	28	9	36	3107.18	13.078569
6	6914.57	27	13	35	3292.52	13.448095
7	6750.51	26	4	33	3292.52	14.019975
8	6465.30	25	7	30	3950.19	14.873889
9	6372.50	24	23	29	3950.19	15.184131
10	6257.03	23	27	28	3950.19	15.239936
11	6182.89	22	15	27	4052.48	15.341656
12	5924.20	21	5	24	5385.27	18.277200

Table 4. Corresponding data of generalization process

The proposed progressive generalization algorithm is controlled by a few parameters, produces predictable and repeatable results once the threshold is set to a certain value. It guarantees that every generalization result is feasible. Generalization of the street network can be progressively and dynamically achieved through the algorithm based on the *OGT* structure. The algorithm has to be progressive as it can eliminate the street one by one. It has to be dynamic in that elimination of a street will create a new status, the importance of blocks is recalculated, and then a new *OGT* will be constructed.

The progressive representation can be represented as follows: $SN_0 = SN_{12} + Street_{30} + Street_{25} + \dots + Street_{15} + Street_{55}$

4. CONCLUSIONS AND FUTURE WORK

This paper makes two major contributions. First, it introduces two alternative progressive representation models of street network. Second, it presents a new generalization algorithm based on the ordered generalization tree structure to support the progressive representation models.

There are a number of avenues for future work, including:

- Extending the generalization algorithm to deal with the dangling streets.
- Development of an explicit link of various inputs of generalization algorithms.
- Investigation of other generalization algorithms to support the progressive representation model.

ACKNOWLEDGEMENTS

This research is financed by from China National Science Foundation under grant No 40571133.

REFERENCES

Bertolotto, M., Egenhofer, M. J., 2001. Progressive Transmission of Vector Map Data over the World Wide Web. *GeoInformatica*, 5(4), pp. 345-373.

Buttenfield, B. P., 2002. Transmitting Vector Geospatial Data across the Internet. In: Egenhofer M J and Mark D M (Eds), GIScience 2002. Berling: Springer-verlag, Lecture Notes in Computer Science, Vol. 2748, pp. 51-64.

Ai, T, Li, Z., Liu, Y., 2004. Progressive Transmission of Vector Data Based on Changes Accumulation Model. In: 11th International Symposium on Spatial Data Handling, Leicester, UK, pp. 85-96.

Van Oosterom, P., 2005. Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAPface Tree and GAP-edge Forest. *Cartography and Geographic Information Science*, 32(4), pp. 331-346.

Yang, B. S., Purves, R., Weibel, R., 2007. Efficient transmission of vector data over the Internet. *International Journal of Geographical Information Science*, 21(1-2), pp. 215-237.

Visvalingam, M., Whyatt, J. D., 1993. Line generalization by repeated elimination of points. *The Cartographic Journal*, 30(1), pp.46-51.

Kovalevsk, V. A., 1989. Finite topology as applied to image analysis. Computer Vision, *Graphics and Image Processing*, Vol. 46, pp.141-161.

Puppo, E., Dettori, G., 1995. Towards a formal model for multiresolution spatial maps. In: Egenhofer M J, Herring J R (Eds), Advances in Spatial Database, Lecture Notes in Computer Science, Springer, Berlin Heidelberg New York, Vol. 951, pp. 152-169.

Bertolotto, M., 1998. Geometric modeling of spatial entities at multiple levels of resolution. ph.D. Thesis. Department of Computer Science, University of Genova.

Viaña, R., Magillo, P., Puppo, E., Ramos, P., 2006. Multi-VMap: A Multi-Scale Model for Vector Maps. *GeoInformatica*, 10(3), pp. 359-394.

Mackaness, W., 1995. Analysis of Urban Road Networks to Support Cartographic Generalization. *Cartography and Geographic Information Systems*, 22(4), pp.306-316.

Jiang, B., Claramunt, C., 2004a. A Structural Approach to the Model Generalization of urban Street Network. *GeoInformatica*, 8(2), pp.157-173.

Richardson, D. E., Thomson, R. C., 1996. Integrating Thematic, Geometric, and Topologic Information in the Generalization of Road Networks. *Cartographica*, 33(1), pp. 75-83.

Morisset, B., Ruas, A., Simulation and agent modeling for road selection in generalization, In: 18th International Cartographic Conference, Stockholm, Sweden, 1997, pp. 1376-1380.

Thomson, R. C., Richardson, D. E., 1999. The "Good Continuation" Principle of Perceptual Organization Applied to the Generalization of Road Networks. In: 19th International Cartographic Conference, Ottawa, Canada, pp.1215-1223.

Edwardes, A. J., Mackaness, W., 2000. Intelligent generalization of urban road networks. In: GIS Research UK Conference, York, UK, pp.81-85

Zhang, Q. N., 2004. Road network generalization based on connection analysis. In: 11th International Symposium on Spatial Data Handling, Leicester, UK, pp.343-353

Hu, Y. G., Chen, J., Li, Z. L., Zhao, R. L., 2007. Selective Omission of Road Features Based on Mesh Density for Digital Map Generalization. *Acta Geodaetic et Cartographica Sinica*, 36(3), pp. 351-357.

Peng, W., Muller, J. C., 1996. A Dynamic Decision Tree Structure Supporting Urban Road Network Automated Generalization. *The Cartographic Journal*, 33(1), pp.5-10.

Tian, J., Guo, Q. S., 2007. A progressive approach to the generalization of an urban street network. In: 15th International Conference on GEOINFORMATICS, Nanjing, China, Vol 6751.

Jiang, B., Harrie, L., 2004b. Selection of Streets from a Network Using Self-Organizing Maps. *Transactions in GIS*, 8(3), pp. 335-350.

Li, Z. L., Choi, Y. H., 2002. Topographic Map Generalization: Association of Road Elimination with Thematic Attributes. *The Cartographic Journal*, 39(2), pp. 153-166.

BjØrke , J. T., 2003. Generalization of Road Network for Mobile Map Service: An Information Theoretic Approach. In: 21st International Cartographic Conference, Durban, South Africa, pp.127-135.

Hoppe, H., 1996. Progressive meshes. In: SIGGRAPH96, pp. 99-108.

Sester, M., Brenner, C., 2004. Continuous Generalization for Visualization on Small Mobile Devices. In: 11th International Symposium on Spatial Data Handling. Leicester, UK, pp. 355-368.

De Floriani, L., Marzano, P., Puppo, E., 1993. Spatial queries and data models. In: Frank A U, Campari I(Eds), Spatial Information Theory-A theoretical basis for GIS, Lecture Notes in Computer Science, Springer, Berlin Heidelberg New York, , Vol. 716, pp. 113-138.

Porta, S., Crucitti, P., Latora, V., 2006. The network analysis of urban streets: A dual approach. *Physica A*, 369, pp.853-866.

Wang, J. Y. et al, 1993. *Cartographic Generalization of General Map*, Surveying and Mapping Press, Beijing.