

Joint generalization of city points and road network for small-scale mapping

*T. E. Samsonov^{1,2}, A. M. Krivosheina¹

¹Lomonosov Moscow State University, Faculty of Geography, 1 Leninskiye Gory, Moscow, Russia, 119234

*Email: tsamsonov@geogr.msu.ru

²Delaunay Laboratory of Discrete and Computational Geometry, Yaroslavl State University, 150000, Sovetskaya Str. 14, Yaroslavl, Russian Federation

1. Introduction

Multiscale cartographic web services such as Google Maps, Microsoft Bing and OpenStreetMap are produced using large amounts of spatial data that is initially prepared in large and intermediate scales. Since these services are published in multiscale environment a problem of data generalization for small scales arises. Usually this task is solved by selection of objects by their attributes. The smaller scale is, the more important object classes are selected. However, this strategy is only suitable for territories with mostly uniform distribution of settlements, for example, Western Europe. Large and climatically diverse regions, such as Russia, Australia, and Northern Africa are notable for their extremely non-uniform pattern of topographic elements. Attribute selection does not work in such cases, since in sparsely populated areas we should depict even some of the smallest cities and low-category roads to show that the territory is lived-in. At the same time the ratio of population densities between densely and sparsely populated territories should be retained as far as possible.

Another important aspect of small-scale mapping is that road network generalization should correspond strongly to location of generalized city points. Every selected road should start and finish at the city point or another road. If two cities are connected in the source map, this connection should be retained after generalization. A brief survey of web services mentioned above shows that the rules of small-scale generalization are often violated. This leads to inadequate representation of the territory with lots of road dangles and only the largest cities retained.

This study aims to develop an algorithm for joint generalization of settlement points and road network with regulation of settlements distribution density over the territory.

2. Background

Small-scale representation of settlements and roads is usually implemented using point set and network generalization algorithms. Li (2007) singles out two types of algorithms for point set generalization: selective omission and structural simplification. The first group includes settlement-spacing, distribution-coefficient, gravity-modeling, set-segmentation, quadrant-reduction (Langran and Poicker 1986) and circle-growth (van Kreveldt et al. 1995) algorithms. The second group of algorithms simplifies the set of points by removing some of them on the basis of various characteristics. One of the most popular strategies is to use Voronoi-based generalization, which ranks points according to their Voronoi areas and thematic importance and iteratively removes points with smallest weights (Ai and Liu 2002, Yan and Weibel 2008). Qian et al (2006) introduced the polarization transformation approach.

For simplification of geographical networks graph theory is usually used (Mackaness and

Beard 1993). Using various measures the importance of individual edges and nodes can be determined (Li and Choi 2002). Thomson and Brooks (2007) generalized road networks using “strokes” with the assumption that the most probable edges for traveling are those with minimal travel cost.

Analysis of the generalization techniques reveals the following problems:

- Most of the algorithms for point generalization are *global*, so they do not allow adjustment of generalization parameters to the characteristics of local pattern and density of objects. For example, Ai and Liu (2002) algorithm preserves point distribution pattern well. However it does not consider cartographic requirements, which compel generalization thresholds to tighten in densely populated regions to free the space for symbols and labels. At the same time generalization thresholds should be relaxed in sparsely populated regions to show the presence of habitation. Recent progress in hydrography generalization with regional parameterization (Buttenfield Stanislawski and Brewer 2011) shows that this technique can be successfully applied to build more cartographically oriented generalization process.
- The task of joint generalization of point and network features and its cartographic application is not developed enough. An experience in agent-based generalization (Touya and Duchene 2011) that combines multiple layers and criteria together shows that collaborative generalization strategy gives the most satisfying results. However it is applied primarily to the large-scale maps, while the small-scale generalization is different because of increasing competition between objects and attention to patterns and clusters instead of individual objects.

3. Methodology and results

We solve the task of joint generalization in two steps. At first, the city points are generalized, and then the road network is adjusted to generalized set of points.

The problem of point density regulation is treated by clustering the points into regions of various densities and then generalizing each region independently with individual percentage of omitted points. We based Voronoi polygons removal on three factors:

- Voronoi polygon area A with weight w_a .
- City population P with weight w_p .
- City administrative status S with weight w_s .

Both population and administrative statuses were classified into 5 categories from 1 to 5 and the weight W_i of every point is calculated as:

$$W_i = w_a A_i \cdot w_p P_i \cdot w_s S_i$$

Using the weights w_a , w_p and w_s we can control the influence of each factor on the importance of the point.

The main difference from previous Voronoi-based techniques is that we:

- do not generalize all the points in one set, but do this operation by subsets of points with different clustering tolerance;
- do not freeze adjacent neighbours of currently deleted point, but increment their neighbour index N_i — the number of deleted neighbours — and use it along with weight in point sorting; this simplifies algorithm and derives additional useful information about generalization cores (see further).

The outline of the algorithm is as follows:

1. Define several levels L_j of clustering tolerance. For example these levels can be 25, 50, 100 and 200 kilometres. The polygonal cluster is a group of points in which the distances from every point to its two nearest neighbours is less than defined threshold T_j .

2. Iterate through L_j and delineate clusters using triangulation of the points and merge the adjacent triangles in which two edges are shorter than threshold T_j . At each level of clustering a set of polygons is produced.
3. Add the largest clustering level, which covers *all* the points.
4. From every polygon subtract all polygons of smaller clustering tolerance.
5. Buffer the resulting polygons by the distance $d=0.5T_j$.
6. Define the percentage R_j of the points to be removed for every clustering level.
7. For every point define the smallest clustering level to which it belongs.
8. For every point calculate W_i according to clustering level; set N_i equal to zero.
9. Iterate through clustering levels.
 - a. Iterate through cluster polygons at current level and omit the points using the following strategy:
 - i. Select the points belonging to current polygon and current clustering level.
 - ii. Derive Voronoy diagram for these points and clip it by current polygon.
 - iii. Sort the points by N_i (first) and W_i (second) fields in increasing order.
 - iv. Iterate until the percentage is achieved.
 - Select the first point in the list and remove it.
 - Increment N_i of adjacent Voronoy neighbours by one.
 - Rebuild and clip the Voronoy diagram.
 - Recalculate W_i for adjacent Voronoy neighbours.
 - Sort the points by N_i (first) and W_i (second) in increasing order.

Our algorithm was programmed using Model Builder for ArcGIS Desktop 10. Experiments showed that it is flexible, allowing to find an optimal balance in preserving the point pattern, retaining points in sparsely populated areas and rarefying extremely dense regions for labeling and symbol placement (Figure 1). The balance can be controlled by selection of different number of clustering levels L_j , different thresholds T_j and point removal percentages R_j for each level. Resulting values of neighbour index N_i (which is incremented each time the neighbour of the point is deleted) show generalization cores. The points with high value of order are located in intensively generalized regions.

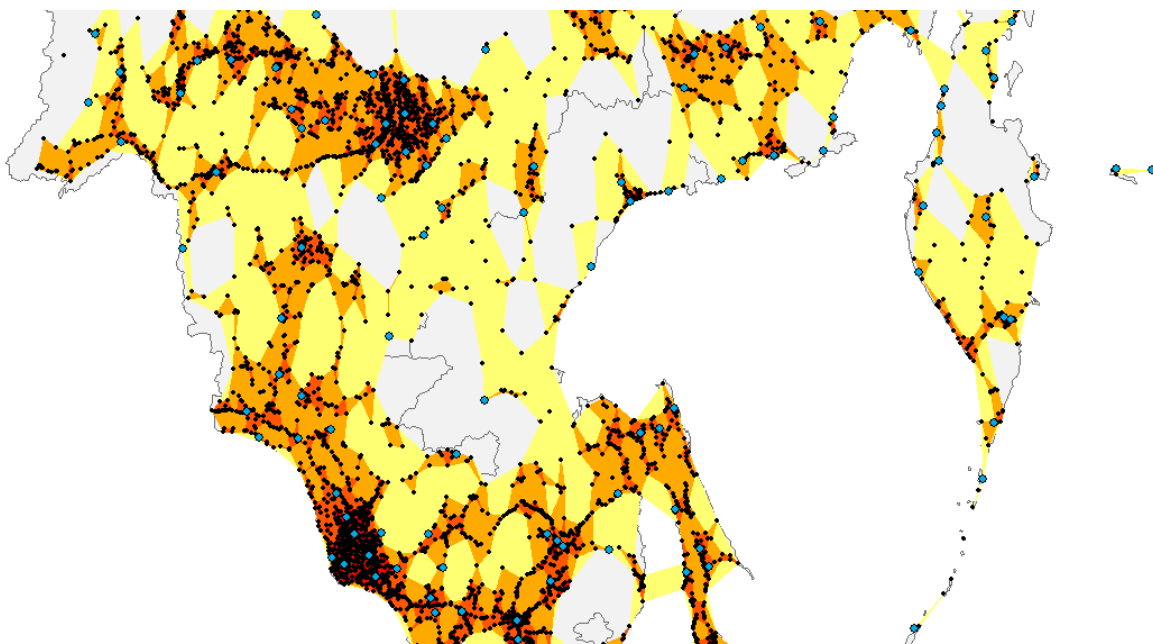


Figure 1. The result of city points generalization using polygonal clusters. Selected points are shown in blue and clusters in yellow-red gradations.

The next stage is to generalize road network. Since we have a set of generalized city points they can be used in generalization process. Here we applied a following strategy:

1. Build the graph using the source road data. Introduce 3 levels of hierarchy in the network to make the highways preferable for routing.
2. Derive the routes between every pair of generalized points. In fact the number of pairs can be limited to a few without altering the result.
3. Select edges, which are included in one or more routes.
4. Reconstruct road network using selected edges.
5. Simplify roads geometrically.

This algorithm was also implemented using ArcGIS Model Builder and ArcGIS Network Analyst Extension. For geometric simplification of roads we used bend simplify method (Wang and Muller 1998). Example of application of algorithm is presented on Figure 2.

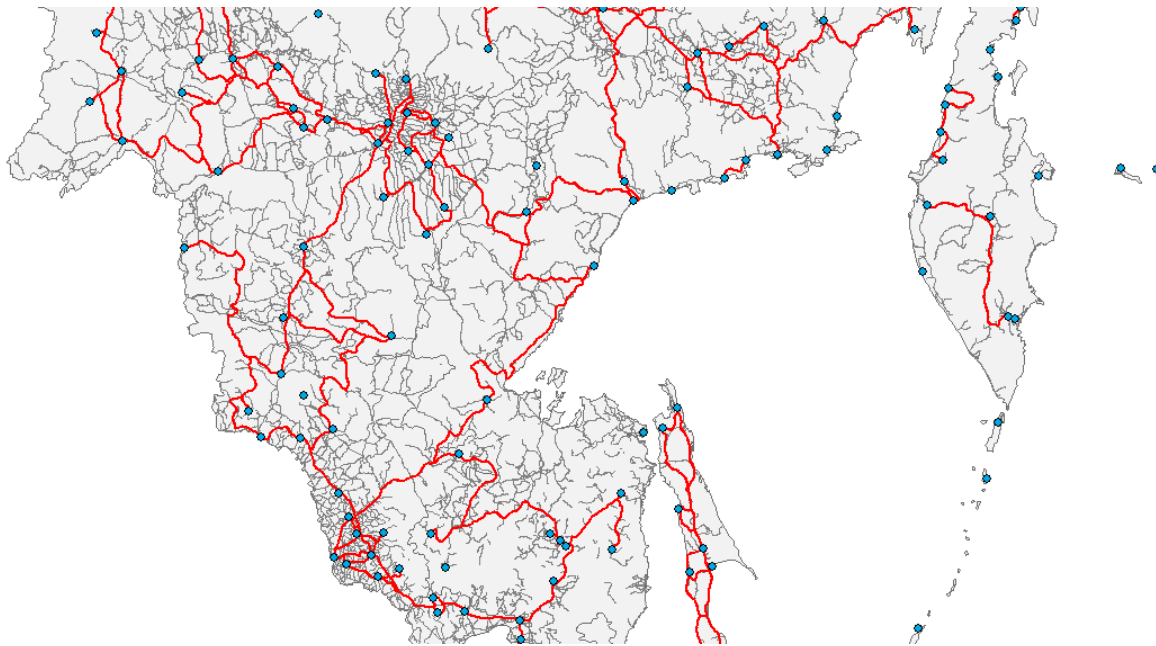


Figure 2. Result of road network generalization using selected cities. Some points are located on isolated parts of the network and some are not reached by the roads in source dataset.

Our approach was applied to city points and road generalization from 1:1 000 000 scale to 1:10 000 000 for Far East region of Russia which is well known for its non-uniform settlement distribution (Figure 3). The number of city points was reduced from 4703 to 184 and the number of road segments was reduced from 8414 to 455. For comparison of the results we prepared a map derived by attribute selection, which is the common practice in modern web services (Figure 4).

Our results show high degree of reconciliation between city and road layers that is close to manual generalization. However there are many ways in which algorithm can be improved. The first is to extend the variety of factors that influence the importance of objects. The next is to embed road network pattern recognition and detection of nodal points that comprise the distinctive features of this pattern. Another suggestion is to implement collaborative algorithm that generalizes points and roads together instead of generalizing them in two steps. In current version of algorithm the point generalization step is fully independent from road generalization. We plan to improve the methodology and estimate its applicability in multiscale basemapping.



Figure 3. Result of joint cities and road network generalization using new approach



Figure 4. The result of generalization using selection by attributes (city population and administrative status, road class).

Acknowledgements

This work is supported by the Russian Government project 11.G34.31.0053. Data+ Ltd provided the data for this experiment.

References

- Ai T and Liu Y, 2002, A method of point cluster simplification with spatial distribution properties preserved, *Acta Geodaetica et Cartographica Sinica*, 25(1), 35–41.
- Buttenfield BP, Stanislawski LV and Brewer CA, 2011, Adapting Generalization Tools to Physiographic Diversity for the USGS National Hydrography Dataset. *Cartography and GIS* 38(3): 289-301.
- Langran CE. and Poiker TK, 1986, Integration of name selection and name placement, *Proceedings of 2nd International Symposium on Spatial Data Handling*, pp. 50–64.
- Li Z, 2007, *Algorithmic Foundation of Multi-Scale Spatial Representation*. CRC Press, 2007, 310 c.
- Li Z and Choi YH, 2002, Topographic map generalization: association of road elimination with thematic attributes. *The Cartographic Journal*, 39(2) 153-166
- Mackaness WA and Beard MK, 1993, Use of the graph theory to support map generalization. *Cartography and Geographic Information Systems*, 20(4), 210-221
- Qian H, Meng L, Wu F and Wang J, 2006, The generalization of point clusters and its quality assessment based on a polarization approach. *Mapping and Image Science*, 4, pp. 55-63
- Thomson R and Brooks R. 2007, Generalization of Geographical Network. In: W.A. Mackaness, A. Ruas and L.T. Sarjakoski (Editors), *Generalization of Geographic Information: Cartographic Modelling and Applications*, pp 255-267.
- Touya G, Duchêne C, 2011, Collagen : collaboration between automatic cartographic generalisation processes, *Advances in Cartography and GIScience Vol.1, Selection from ICC'11, Paris*, pp 541-558
- van Kreveld M, van Oostrum R and Snoeyink, J. (1995) Efficient settlement selection for interactive display, in *Proceedings of Auto Carto 12, Bethesda, MD*, pp. 287–296.
- Wang Z and Muller JC, 1998. Line Generalization Based on Analysis of Shape, *Cartography and Geographic Information Systems*, 25, pp 3-15.
- Yan H and Weibel R, 2008, An Algorithm for Point Cluster Generalization Based on the Voronoi Diagram. *Computers & Geosciences*, 34(8): 939-954.