# Map Generalization from Scale of 1:500 000 to 1:2 500 000

Li Limeng, National Geomatics Center of China, Beijing, 100044, P. R. China
Wang Lixin, National Geomatics Center of China, Beijing, 100044, P. R. China

**ABSTRACT:** Generalization is for making smaller scale maps from larger scale maps. It involves various techniques to select proper features and simplify the selected liner features in order to suit the final map scale. This paper introduces the method and procedure of map generalization used to produce the Map of China at 1:2.5 million from maps at 1:500 thousand with the assistance of the computer. To meet the requirements of the scale of 1:2.5 million, proper amount of point features and liner features are selected, then the selected liner features are simplified. The ArcInfo generalization tools are used to reduce the difficulties of manual works. Because of the defects of automatic map generalization, it is also necessary to employ the human interactive editing.

## INTRODUCTION:

The 1:500 thousand Digital Map of China is digitized in 1998, later on it is revised with latest information. It is mainly used for building thematic Information Database with the GIS technology and making thematic maps. It is in ArcInfo coverage format. It consists of 12 layers (Coverage), including political and Oceanic Layer, Populated place Layer, Road layer, Railroad Layer, Drainage Layer, etc. The data volume is about 100 Megabytes.

The information in 1:500 thousand map is too much for the wall mounted National Thematic Maps of China and come Information Systems. For this and economical reasons, a Digital map of smaller scale is preferred. A 1:2.5 million Digital Map of China is requested to be produced to meet the qualification. The computer aided generalization method is engaged in the production. The following are the reasons to do it this way other than digitizing from 1:2.5 million paper maps. First, the 1:500 thousand digital map is newer than the 1:2.5 million paper map. Second, the consistency of location and attribute of features between 1:500 thousand and 1:2.5 million Digital maps can be easily kept. Third, the cost may be reduced.

The 1:2.5 million Digital Maps of China have 10 layers, including political and Oceanic Layer, Populated place Layer, Road layer, Railroad Layer, Drainage Layer, etc. All the features including Polygon features, liner features and point features are generalized from 1:500 thousand Digital Maps. The generalized data set volume is about 20 Megabytes.

It takes four engineers around one month to finish all the work. The working environment includes four Pentium II/III NT workstations in a network and ArcInfo software in each workstation.

## PROCEDURE:

Different methods are used for point, liner and polygon features.

### 1. Point Features

Generalizing point features are easier than generalizing other features. Collecting proper points is what should be done. This operation is called typification. In this generalizing project, the point shape populated place is the only point shape coverage.

First, the attributes of the points are used for selecting. All the points of county level or higher are selected. Second, the remaining points are selected manually according to the density of the points; the 1:2.5 million paper maps are used as references. At last, 12447 point shape populated places are collected from 19004 points in the 1:500 thousand Digital Map.

## 2. Liner Features

Roads, Railroads Liner Rivers and sea features (islands, corals and reefs) are the layers engaging liner feature generalization. Only some of features are needed in the smaller scale maps. Therefore the first step in generalizing liner features is typification which reduces feature density and simplifies the structural pattern without destroying the overall impression of a feature group at the smaller scale. The attribute value of each feature is used for collecting lines. In the case of road, national and provincial roads are collected.

Line simplification is the second step, the selected liner features are simplified by remove the extra vertexes in the lines. The ArcInfo GENERALIZE command is used. It is used twice, each time with different simplification operator. In the first time, with POINTREMOVE operator the lines are simplified using the Douglas-Peucker's algorithm. In the second time, with BENDSIMPLIFY operator extraneous bends are removed. When using GENERALIZE command, the weed tolerance must be determined. According to the accurate demand of the final map, 0.0025 degree (about 0.1mm on the scaled map unit) is used as weed tolerance when using GENERALIZE command with POINTREMOVE operator. With BENDSIMPLIFY operator, the weed tolerance is chosen after test of a portion of coastline, 0.02 degree (about 0.9mm on the scaled map unit) is accepted.

It is based on a test to use the GENERALIZE command twice. In this way the redundant points along lines can be efficiently removed and the changes of the shape of lines are limited. The following table shows the comparison of each method for Road Layer.

| Coverage | Number of Arcs | Number of Vertexes after GENERALIZE | | | |
| --- | --- | --- | --- | --- | --- |
| | | Source | POINTREMOVE Only | BENDSIMPLIFY Only | POINTREMOVE & BENDSIMPLIFY |
| Roads | 16306 | 600469 | 189664 | 420592 | 184151 |

After the GENERALIZE command, the section and route feature classes are remained. The following table shows the number of vertexes for each liner layer after the GENERALIZE command.

| Coverage | Number of Arcs | Number of Vertexes | | |
| --- | --- | --- | --- | --- |
| | | Source | After POINTREMOVE | After BENDSIMPLIFY |
| Roads | 16306 | 600469 | 189664 | 184151 |
| Railroads | 854 | 67172 | 18023 | 17879 |
| Liner Rivers | 146101 | 3071643 | 886337 | 840822 |
| Sea Features | 7611 | 394428 | 104641 | 94812 |

## 3. Polygon Features

The Political and Oceanic layer, Polygon shape Populated Place layer and Drainage layer are generated from polygon shape layers.

In ArcInfo, the lines and polygons can stay in one coverage. The polygons are closed with lines, and there can be some lines which don't make any polygon. For example, in the drainage layer, the drainage network consist of polygons (polygon shape river and lake) and lines (liner river). The line generalization is one must step in the polygon feature generalization.

According to the different characteristics of each polygon shape layer, the area or perimeter of feature attribute is used in selecting. The polygons with areas bigger or perimeters longer than a pre-defined value are selected. The value is calculated according to the scale of the final product; in this case, it is 1:2.5 million. Individual lines (not the border of polygon) are also selected. Then the selected polygons and lines are simplified by the GENERALIZE command (see liner features generalization above). The table below shows the number of vertexes for each polygon layer after the GENERALIZE command.

| Coverage | Number of Arcs | Number of Vertexes | | |
|---|---|---|---|---|
| | | Source | After POINTREMOVE | After BENDSIMPLIFY |
| **Political and Oceanic** | 8575 | 708166 | 189495 | 179163 |
| **Drainage** | 17928 | 1094131 | 301312 | 231851 |

After the GENERALIZE command, the CLEAN command is used to restore the polygon and region topology.

The Polygon Shape Populated Place Layer is a special coverage in generalization, because many polygons representing the street blocks of a city in the larger scale map should be merged into one polygon or a few polygons in the smaller scale map to represent the city. In ArcInfo 8.0.2, there is a new command –AREAAGGREGATE to do this kind of work. This command involves the use of GRID, an extension software. Therefore, the GRID license is needed. The AREAAGGREGATE command first converts the input polygons to grid (raster), and then uses GRID functions EXPAND, SHRINK, and so on, to group features within the specified distance of each other. The result is then converted back to vector with proper construction of new boundaries. After this command, a one-to-many relation table that linked the aggregated preliminary regions to their source polygons can be used to derive attributes for the output features. Unfortunately, because there is a limitation for the number of lines and columns in generating the GRID, for whole China, the cell size is too big to get an accurate result.

The BUFFER command can also be used to merge the street blocks. There are two steps, First, use BUFFER command to extend the street blocks to make one polygon. Second, use BUFFER command with negative buffer distance to shrink the polygon to the correct size. The problems of this method are as follow. First, the attributes of the polygon should be edited (copying label points from the original coverage). Second, this method can not deal with the polygons with holes properly. The result will omit the holes.

In simplifying the Polygon Shape Populated Place Layer from 1:500 thousand map to 1:2.5 million, because only 34 cities were selected as polygons, the on-screen digitizing method is used. In the digitizing process, the original polygons of street blocks are displayed in the background. At last the attributes of cities are transferred.

## DATA EDITING:

After the point, line and polygon features are simplified, manual editing is required. The reasons are as follow.

First, when applying GENERALIZE command, topological errors are produced, including line-crossing, line-overlapping, and zero-length lines; For polygon coverage, errors of no label or multiple labels can also be produced; To deal with this problem, manual works are necessary.

Second, there is new information for Political Boundaries, national roads and railroads, the data should be revised manually.

In the editing process, the Drainage Layer and Liner River editing is difficult. For Liner River Layer, as there is no rank item in the attribute table, manual work is the only way to collect lines to form the typified Liner River Layer. For Drainage Layer, there are a lot of topological errors to be corrected manually.

## PROBLEM:

The main problem is that the GENERALIZE command may create label errors for polygon coverages as well as topological errors for polygon and/or line coverages. To correct these errors, a lot of manual work has to be done.

In the upcoming Workstation ArcInfo 8.1 release, an option, NOERRORCHECK/ERRORCHECK, will be added to the GENERALIZE command. This option specifies whether to check for these errors or not. If ERRORCHECK is specified, the GENERALIZE command will find and avoid errors generated by the line simplification. If any

topological errors are found, the arcs involved will be re-generalized using a reduced tolerance. Then the result will be checked for topological errors again. This process iterates until no errors are found. With this option, the difficulty and time of manual work would be greatly reduced.

## CONCLUSION:

The 1:2.5 million digital map was successfully done after working for one month. Computer aided generalization is an easy way of making a digital map when larger scale maps are available. It has the advantage of reducing time, man-power and cost. Without the help of the computer and software, the generalization work would never be finished in a short time.
Workstation ArcInfo is a reliable platform in generalization work. With the upcoming new features of its generalization tools, the ArcInfo software will be more powerful.
It is better for the geo-data to have a rank item. It can be used in automatic typification to collect needed points and lines to the output.

## REFERENCE:

Environmental Systems Research Institute, INC., 2000, Map Generalization in GIS: Practical Solutions with Workstation ArcInfo Software

Environmental Systems Research Institute, INC., 1994, ARC Commands

Environmental Systems Research Institute, INC., 1994, Editing Coverages and Tables in ARCEDIT

National Geomatics Center of China, 1998, Technical Paper of 1:500 Thousand Digital Map Database

National Geomatics Center of China, 2000, Data Description of 1:2.5 Million Digital Map Database

Environmental Systems Research Institute, INC., 1993, Data Dictionary of Digital Chart of the World