

AN OVERVIEW OF THE DUTCH APPROACH TO AUTOMATIC GENERALIZATION

Vincent van Altena, Jan Bakermans, Ben Bruns, Hedwig Kupers, Ron Nijhuis, Marc Post

After years of discussion on feasibility of fully automatic generalization of cartographic maps, the journey has come to an end. The Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster) has succeeded in fulfilling this challenge. The fully automatic generalization process of Kadaster is an international breakthrough in the production of geo-information.

Introduction

In the Netherlands national mapping is regulated by law. It is Kadaster's statutory task to maintain a number of registrations. One of these registrations, the Basis Registratie Topografie (Key register Topography, BRT) consists of digital topographic data sets at different map scales (1:10k, base data set. and derived / generalized data sets at scale 1:50k, 1:100k, 1:250k, 1:500k and 1:1.000k). These data sets are available as open data. Governmental organizations are obliged to use the available BRT data sets for the exchange of geographical information.

The introduction of the BRT as open data placed a major challenge on the acquisition and production of the data sets itself. Austerity measures resulted in a declining budget, while the demand for increased actuality was laid down in legal obligations. These obligations require an actuality of less than two years for the whole range of BRT product family. At that point only the base data set. (1:10k) met this requirement.

Other issues were complicating the situation. The traditional production method of the derivation and generalization of mid- and small scale data sets and map series were at the end of their life cycle. Utilization of CAD software, for interactive generalization with the purpose to produce derived data sets in a spaghetti structure, was thought to be not appropriate for the increased demands for actuality and object oriented data.

These developments, combined with the state of art of generalization tools within ArcGis 10 (and up), urged Kadaster in 2010 to start a feasibility study on a 100% automatic generalization of the 1:50k map from a 1:10k data set. The results of the feasibility study were very promising and led to the development and implementation of a new production line for the 1:50k map, based on fully automated generalization and map automation (Stoter et al. 2011).

Dutch approach / How

There are two distinguishing aspects of the Dutch automatic generalization approach: one is the composition of the development team. Each member was selected for partaking as developer because of an out-of-the-box mentality and because of a persevering attitude. A cartographic background was not essential.

The second very important characteristic is the iterative agile approach. The process was started with a minimal set of requirements defining the scope and bandwidth of the development:

- the new generalization process should run 100% automated
- the "new map" does not have to be an exact imitation of the "old map"
- the automatic process should connect directly to the base data production flow to achieve an identical actuality (less than two years).
- the BRT product family must be achieved cost-effectively.
- during the development, active user participation by evaluating results is necessary.

The task to develop a new 50k map generalization and automation process involved the translation of assessments in cartographers mind into a form of artificial intelligence. A cartographer is able to base his judgments and decisions about the representation of topographical reality on a smaller scale on an evaluation of multiple criteria simultaneously. To equate this with serial computing technology was a huge challenge. One of the issues was the interaction between different objects. If for instance the symbolization of a road has to be exaggerated, this affects adjacent buildings which should be displaced accordingly. Eventually this resulted in the requirement to generalize all topographic objects in one pass.

Work flow

Unfortunately, the generalization work flow is too complicated to be executed for the whole of the Netherlands at once. Therefore, the base data has been separated in subsets which are divided by real topographic, linear objects (see Figure 1). These objects must remain untouched during the whole process to be able to stitch the subsets together to one topologically correct data set.

The primary road network (highways and primary roads) expanded with some other linear topographic features near the coast suited this purpose. The resulting partitions (about 500) are each processed entirely within an hour. As a side effect, partitioning opens the door to parallel processing of multiple areas and decreases processing time enormously.



Figure 2: Partitions - geographical areas as processing units

There are however some topographical features, such as railroads, high tension lines and administrative borders, that cannot be divided as they span multiple partitions. These objects are processed at once after all individual partitions are successfully generalized.

Generalization results in the individual partitions are propagated where a *nation wide* object is affected.

All objects are integrally preprocessed before generalization starts. Enhancement and enrichment of the source data improved achieved results substantially. Detected errors are corrected in our dataset to keep the production process going, but are also reported to the

basemap production team for a substantial solution in the long term.

All subsequent processing takes place at two levels: at local partitions or nationwide. Both branches are modular and consist of three steps: model generalization, geometric generalization (displacement) and graphic generalization (cartographic conflict resolution). For a detailed description, see (Stoter et al. 2013).

The advantage of this modular approach is the potential to replicate the models, adapt parameter values or even substitute parts of the process and produce other map series (i.e. a 100K map series) See Figure 2 below

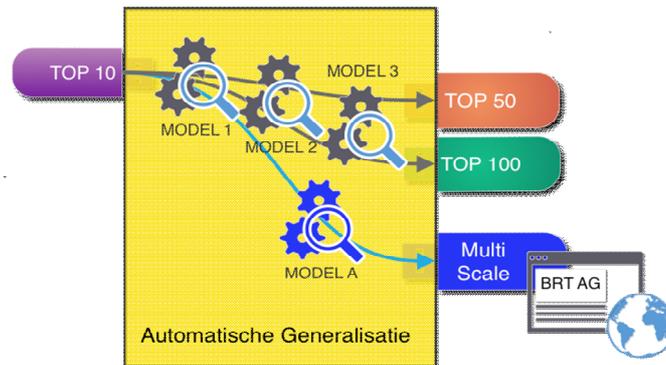


Figure 4: the modular approach

Web services

The modular approach is very flexible and provides other interesting possibilities. The preprocessed data is used as input for an automatically derived and generalized base map. A multi-scale web service was designed using Scalemaster methodology (Brewer et al. 2007) and the technical development was implemented in ArcGIS. The resulting 15 layer base map has to be used by governmental organizations.

User feedback

During the development of the automatic generalization process close ties were kept with the key users of the 50k map. After each development iterations the newly acquired results were discussed with the key users which gave us valuable feedback for the next development iteration. Originally the 50k map is a military product which must comply with NATO standards. Hence, one of the most important customers of this topographic product is the Royal Dutch Army.

In course of the aforementioned feedback cycle a field test was conducted with an automatically generalized 50k map in June 2013. An infantry division of the Training Command performed a two week field training in the south of The Netherlands. In these two weeks the soldiers used the map intensively and were very satisfied with the results. In particular the increased actuality was an improvement eagerly looked forward to. After this successful test Kadaster decided to release the first version of the automatically generalized maps in September 2013.

Results and benefits

Both Kadaster and users benefit from the implementation of a fully automated generalization production workflow.

The automation of manual cartographic labor led to a significant cost reduction. Using the new procedure enabled Kadaster to produce the map series at one fourth of the original budget.

Processing time has been reduced tremendously and in effect this increased the update frequency of derived map series from a six- to a two year cycle. This is in full sync with the basedata acquisition. Nowadays the source dataset and derived map series are released simultaneously with the same actuality five times a year.

Map content itself is optimized by applying generalization algorithms consistently for the whole of the county and by improving the base dataset and algorithms after each development and production iteration.

Ultimately, users are the real beneficiaries of this innovation: getting actual maps more frequent at lower costs.

References

- Brewer, Cynthia A., Barbara P. Buttenfield, Charlie Frye, and Jessica Acosta. 2007. "Scalemaster: multi-scale mapmaking from multiple database resolutions and for multiple map purposes." In Moscow.
http://www.personal.psu.edu/cab38/ScaleMaster/ICC2007_paper_Brewer_Buttenfield_Frye_Acosta_May31_07.pdf.
- Stoter, Jantien, Ron Nijhuis, Marc Post, Vincent van Altena, Jan Bulder, Ben Bruns, and John van Smaalen. 2011. "Feasibility Study on an Automated Generalisation Production Line." In Paris, France.
- Stoter, Jantien, Marc Post, Vincent van Altena, Ron Nijhuis, and Ben Bruns. 2013. "Fully Automated Generalization of a 1:50k Map from 1:10k Data." *Cartography and Geographic Information Science* (August 22): 1–13.
doi:10.1080/15230406.2013.824637.