A Generic Approach to Simplification of Geodata for Mobile Applications

Theodor Foerster[1], Jantien Stoter[1], Barend Köbben[1] and Peter van Oosterom[2]

[1] International Institute for Geo-Information Science and Earth Observation (ITC)
P.O. Box 6, 7500 AA Enschede, the Netherlands
{foerster,stoter,kobben}@itc.nl
[2] Delft University of Technology, OTB, Section GIS-technology
P.O. Box 5030, 2600 GA Delft, the Netherlands
oosterom@tudelft.nl

Simplification of geodata is promising in the context of data reduction for mobile applications. However, choosing the best simplification algorithm depends on the correct parameterization of the algorithm with respect to the specific data. This is a difficult task for any mobile application provider. This paper proposes a generic approach to simplification, which gives the mobile application provider a means to choose the appropriate algorithm including the most optimal parameter value according to his/her demands. This demand-driven notion adds a new aspect to the research about simplification. The paper proposes a design of an architecture for the determination of such simplification algorithms and demonstrates satisfying results based on a study applied to Dutch topographic road data from TOP10NL.

Keywords: ratio-based simplification, data reduction, mobile applications, generalization

INTRODUCTION

Geodata are used in a broad variety of mobile applications for different domains such as transport, tourism, security and disaster management. All mobile applications have characteristics in common that limit the amount of data that can be portrayed: limited bandwidth of the network, limited power supply of the mobile devices and limited display capabilities of the mobile devices. This requires that the amount of transferred geodata has to be reduced as much as possible, while taking the readability of the portrayed geodata into account. The ways to reduce geodata are a) to define the specific thematic selection or filter criteria, b) to compress the data by applying common zipping algorithms, c) to clip the data by cropping with the geometry of the desired spatial extent and d) to delete unimportant or unperceivable aspects of the data. For the last method, techniques from the field of generalization (Weibel and Dutton, 1999) can be applied, as some of them simplify the geometries and thereby reduce the number of data elements (vertices) that have to be transferred while preserving the major geometric characteristics. Thus these so called simplification algorithms of generalization are not designed to merge or delete any features. This is an important aspect, as the client can still access all features of the complete dataset and request desired details if needed.

However, incorporating generalization functionality appropriately into applications is very complex. Besides choosing the correct simplification algorithm it is also challenging to determine the correct parameter values in order to produce the desired result. This is due to the different and interrelated effects of parameter values and the different parameter semantics of the different algorithms in general but also due to the heterogeneity and the characteristics of the dataset itself. Therefore it is necessary to give the mobile application provider a reliable mechanism at hand, which always produces sufficiently reduced datasets, while retaining their essentials. As it is always possible to measure the effect of data reduction by calculating the ratio of vertices before and after the simplification, we propose a relative measurement,

which defines the degree of simplification in relation to the original dataset. This ratio-based approach adds a new aspect to simplification, as it also introduces a demand-driven and generic means to simplification. This ratio-based simplification is demand-driven, as the data reduction ratio is the major concern of the mobile application provider. The proposed approach is generic, because it provides a shared basis to configure different simplification algorithms in the same way. This shared basis allows easily comparison of several simplification algorithms and their simplified results.

This ratio-based approach enables the determination of the desired degree of data reduction (i.e. simplification ratio) and to compare the simplified results, which are all based on the same simplification ratio, but applied on different algorithms. Providing these results as maps to the mobile application provider will give him/her a valid means to select the appropriate algorithm. Therefore a prototype has been designed in which different simplification scenarios (algorithms, ratios and their effects in the form of maps) are presented to the mobile application provider.

This paper will first give an overview about simplification and describe its relevance in the context of mobile applications. Then the ratio parameter is described in more detail. After that the paper presents some results (based on Dutch road data) and demonstrates the positive effect of our approach, which enables the comparison of different data reduction scenarios (what algorithms and what values to choose) for mobile applications based on the most relevant criteria (visual effects and effective data reduction). The paper ends with a conclusion.

INTRODUCTION TO SIMPLIFICATION

The theory of generalization describes the concept of operators and algorithms (Weibel and Dutton, 1999). Operators such as simplification describe the effect on features and collections of features in a generic fashion. These operators are the basis for algorithms, which implement their generic notion. As there are more algorithms for an operator, the applicability of such algorithms is hard to determine by a non-expert (e.g. the mobile application provider). This is the reason, why we propose a generic view on such an operator. Simplification[1] is performed on lines but also on polygons. Applying simplification to a line, consisting of start and end point and an ordered set of intermediate points affects the line by eliminating the unimportant (i.e. selecting important) intermediate points, which do not hold a desired simplification tolerance. It is important to note, that simplification is strictly about selection/elimination of intermediate points (vertices), not about moving them (Saalfeld 1999).

Research has yielded a lot of algorithms (Li, 2006), which apply different strategies to eliminate points. This leads to parameters with different semantics, which cause different effects on the data. So there is no generic basis to use these algorithms homogeneously. For instance the Douglas Peucker (DP) algorithm (Douglas and Peucker, 1973) uses a distance measurement to test each single point. Instead the Visvalingam and Whyatt (VW) algorithm (Visvalingam and Whyatt, 1993) applies an area measurement. Thus these algorithms require different parameter types and cause different effects on the data, as it was recently again demonstrated visually by a web application of Harrower and Bloch (2006). The DP algorithm for instance is spikier than the VW. But the VW fails in some data situations to produce good cartographic results because it does not take the shape of the measured area in account. These two

---

[1] We are aware of the fact, that the paper deals with the special case of line simplification, but as the terms simplification and line simplification are used interchangeably in the literature (McMaster and Shea, 1992), we will use the more general term simplification.

algorithms are taken as examples for our study to demonstrate the applicability of our proposed approach. We are aware of the fact, that there are different modifications to these algorithms, which solve topological errors (Saalfeld, 1999) and include more sophisticated measurements (Zhou and Jones, 2004), but even then the problem of heterogeneity of parameters still holds.

Simplification in the Context of Mobile Applications

Most of the recent projects (Sarjakoski and Sarjakoski, 2005; Burghardt et al., 2004) about mobile applications with a link to generalization did not address the need of data reduction. Even the application of data reduction algorithms for generalization has been doubted by some researchers (Li 1993), because it can affect the cartographic representation of the data significantly.

Besides the already mentioned benefits of simplification in the context of mobile application is that the data to be processed on the client side is reduced to a reasonable amount. The reduced amount of data improves the performance (due to less communication time) and decreases the power consumption of the mobile client while running the desired mobile application. In general we claim that using appropriate simplification algorithms reduces data reasonably without considerably modifying the visual result (as also shown in the section about the results using DP algorithm). However we are aware of the fact that the quality of the visual result depends on the target scale and the data characteristics.

We admit that simplification might result in drawbacks concerning cartographic representation, as it deletes aspects of the data, which might be relevant for appropriate cartographic representation. However, as we apply simplification only to reduce the transferred data, we do not suggest displaying the reduced data directly on the client side without further modification. Instead we recommend applying further generalization operators such as displacement or enhancement for display purposes afterwards.

DESIGN OF A RATIO-BASED SIMPLIFICATION

The proposed ratio-based simplification approach is called generic, because it provides a valid basis to perform and evaluate simplification with the same configuration applied to different simplification algorithms. Additionally it can measure the effectiveness of any algorithm. Our approach reflects the demand of data reduction, as this demand is valid for most cases, in which simplification is required. In order to meet this popular demand directly, we introduce the simple measurement of *simplification ratio*. This ratio can easily be converted automatically to the corresponding parameter value of the algorithm (by comparing the amount of data of the original with the result), which makes it quite easy to apply it on the actual algorithms. This was at least the experience we got while implementing our approach prototypically.

The simplification ratio is defined as the ratio between the number of vertices (i.e. number of intermediate points plus start and end point) in the dataset before and after simplification. It is important to note that this definition is purely focusing on the data reduction aspect of simplification and is applied to the complete dataset. We are aware of the fact, that there is also the possibility to define it based on cartographic aspects.

RESULTS

The test environment is based on the generalization web service introduced in Foerster and Stoter (2006), which features a client plug-in for JUMP[2]. This framework is available under open source license and

---

[2] Java Unified Mapping Platform website: http://www.jump-project.org

hosted at the 52º North open source initiative[3]. The client can trigger remote processes on any generalization web service by XML messaging and can visualize their results. For test purposes we deployed the DP algorithm (based on JTS[4]) and the VW algorithm (own implementation with classic effective area) in the generalization web service.

The test dataset for this scenario is taken from the Dutch topographic data product TOP10NL (Bakker and Kolk, 2003), which is applicable to a scale range between approximately 1:10000 and 1:25000. We applied our approach on a TOP10NL road dataset, which covers an area of 10 x 6 kilometers size and consists of 6500 line features, which are constructed out of more than 36000 points (i.e. intermediate points and end and start points).

We tested simplification ratios, as defined in the previous section, from 0.0 (no simplification) to 1.0 (full simplification) in steps of 0.1 to the data using both simplification algorithms. The individual simplification tolerance values for each algorithm are calculated from the requested ratio. The requested simplification ratio (x-axis) and the effective data reduction ratio (right y-axis, denoted in percent) are drawn for both algorithms in charts. The effective data reduction ratio describes the ratio of the request and response message size as measured on the simplification service. Additionally the charts contain on the left y-axis the calculated tolerance values of each algorithm to produce the result according to the requested simplification ratio.

It has to be noted, that the effective data reduction also includes the size of the XML messaging overhead. However this overhead is always of constant size and its size relatively decreases, as the overall size of the dataset increases. So the overhead can be ignored in case of larger datasets.

The chart in Figure 1 illustrates the effect of the VW algorithm on the road data and shows a quite linear decrease of data reduction. Note that though the requested simplification ratio is only 0.1 (or 0.2, or 0.3), more reduction is obtained in these cases (indicating the estimation of the tolerance values could be further improved). The data reduction cannot be lower than a little bit below 60 percent. This effect is caused by the nature of simplification, which is to not delete any line features, thus the line features at this threshold only consist of start and end points. This is also true for the case of the DP algorithm. So, the charts after 0.6 are not meaningful anymore.

---

[3] 52º North website: http://www.52north.org
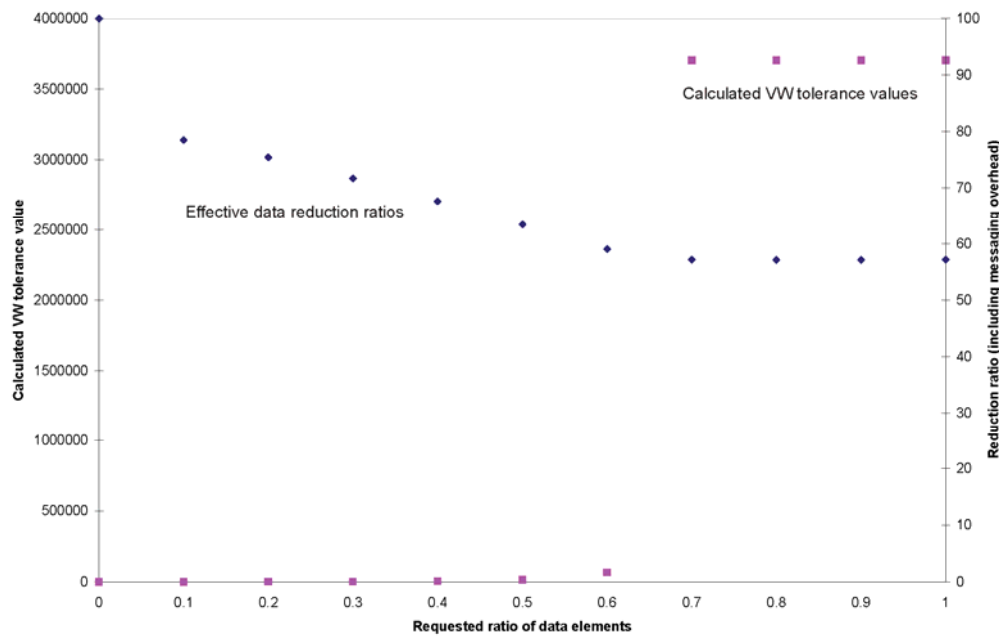[4] Java Topology Suite website: http://www.vividsolutions.com/jts/jtshome.htm

Figure 1: Chart demonstrating the effect of different simplification ratios (from no simplification (0.0) to full simplification (1.0)) to the effective data reduction and the tolerance values using the VW algorithm.

Figure 2 depicts the produced results of the DP algorithm. The values for the effective data reduction ratio vary a bit in comparison to the first chart. The reason for this effect is the nature of the VW algorithm, which assigns in some cases the same value of the effective area to more points during the update of the effective area. Both charts are not very meaningful without looking at the cartographic representation. They show both the difficulty to determine the actual tolerance value for the simplification algorithm because they require different tolerance value ranges and they also have a different course over the increasing simplification rate. Both charts show a big jump, which is due to a certain number of inhomogeneous segments.
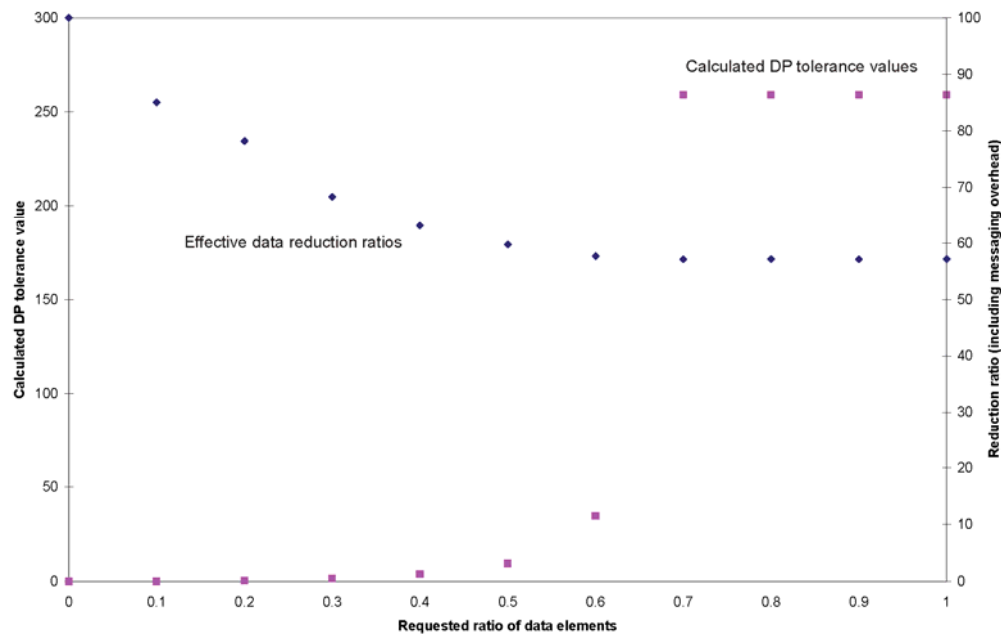
Figure 2: Chart demonstrating the effect of different simplification ratios (from no simplification (0.0) to full simplification (1.0)) to the effective data reduction and the tolerance values using the DP algorithm.

However as the simplified data have to be displayed finally on the mobile device, they have to be evaluated under cartographic aspects as well. This is actually a subjective process and is not the focus of this paper. We want to point out, that our approach eases this cartographic evaluation as it adds transparency as it enables to identify the corresponding counterparts for comparison. These counterparts are produced by different algorithms based on the same simplification ratio. So now the results can be compared on the same basis. This in the end can improve the cartographic evaluation as it adds a certain mechanism of objectiveness to it.

Such a cartographic evaluation could be applied on maps as depicted in Figure 3. In this case of evaluation it is important to choose and compare significant features, which give a good indication about the cartographic applicability of a chosen simplification ratio. Prominent features in this context are short roads, which represent curves. It becomes clear that the effect of simplification is significant as such roads will be simplified roughly with even slightly increasing simplification ratios. As suggested before the DP algorithm produces better results. However a mobile application provider would not have this kind of knowledge and also in the case of accessing web-based generalization algorithms a visual check would be still necessary as well.
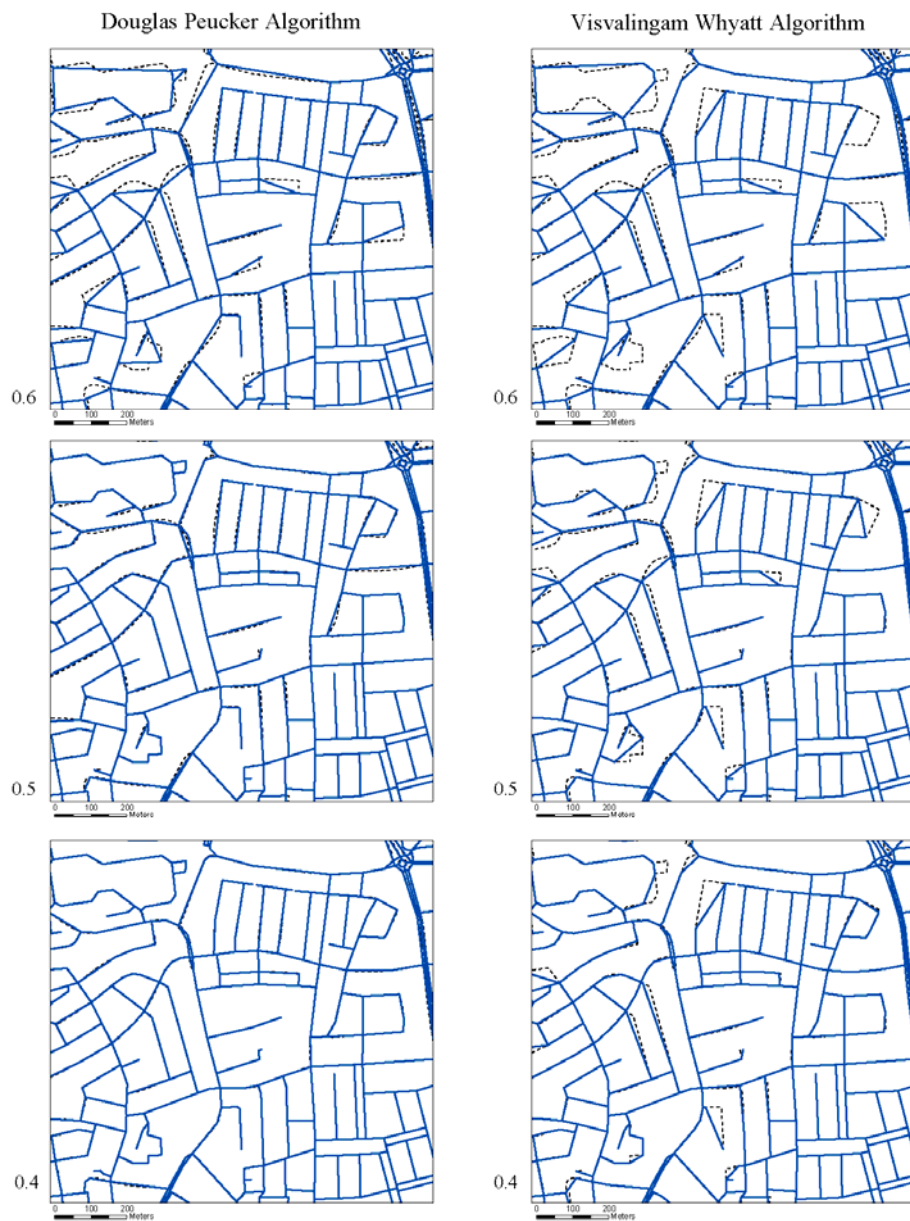
Figure 3: Both algorithms applied to a simplification ratio of 0.6 (top), 0.5 (middle) and 0.4 (bottom) - dashed line symbolizes the original line; a straight line symbolizes the generalized one.

DISCUSSION & CONCLUSION

This paper opens up a new discussion about a ratio-based web service interface for generalization, the deployment of such an interface in a service-oriented architecture and the benefits of ratio-based simplification in the context of progressive transfer (van Oosterom et al., 2006). Additionally the paper addresses the need for an intelligent framework to automatically provide the most suitable data reduction ratio according to the data provider's requirements.

Finally the proposed ratio-based approach introduces a new aspect to simplification. It covers demands of the mobile application provider seeking for appropriate data reduction mechanisms. Ratio-based simplification combined with cartographic evaluation of the simplified results support the mobile application provider determining the appropriate simplification algorithm and its parameter values.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

Bakker, N.; Kolk, B.: A new object-oriented topographical database in GML. In: ICC 2003. Durban, South Africa, 2003.

Burghardt, Dirk; Edwardes, Alistair ; Mannes, Juerg: An architecture for automatic generalisation of mobile maps. In: Gartner, G. (Ed.): 2nd Symposion on Location based service and telecartography. Vienna, Austria, 2004.

Douglas, David H.; Peucker, Thomas K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. In: The Canadian Cartographer 10 (1973), No. 2, P. 112–122.

Foerster, Theodor; Stoter, Jantien: Establishing an OGC Web Processing Service for generalization processes. In: ICA workshop on Generalization and Multiple Representation, 2006.

Harrower, Mark; Bloch, Matthew: Map-Shaper.org: A Map Generalization Web Service. In: IEEE Computer Graphics and Applications (2006), July/August 2006, P. 22–27.

ISO/TC 211: Geographic information - Simple feature access - Part 1: Common architecture / International Organization for Standardization.2004 (ISO 19125-1). – ISO Standard.

Li, Z.: Some observations on the issue of line generalisation. In: Cartographic Journal 30 (1993), No. 1, P. 68–71.

Li, Z.: Algorithmic foundation of multi - scale spatial representation, CRC, 2006.

McMaster, Robert: Automated line generalization. In: Cartographica 24 (1987), No. 2, P. 74–111.

McMaster, Robert B.; Shea, K. S.: Generalization in Digital Cartography. American Association of Geographers, 1992.

van Oosterom, P.; de Vries, M; and Meijers, M: Vario-scale data server in a web service context ICA Workshop, June 25th 2006, map generalisation and multiple representation, Vancouver, Washington USA.

Saalfeld, Alan: Topologically Consistent Line Simplification with the Douglas-Peucker Algorithm. In: Cartography and Geographic Information Science 26 (1999), No. 1, P. 7–18.

Sarjakoski, Tapani; Sarjakoski, Tiina: The GiMoDig public final report. 2005. – Project report.

Visvalingam, M.; Whyatt, J. D.: Line generalization by repeated elimination of points. In: Cartographic Journal 30 (1993), No. 1, P. 46–51.

Weibel, R.; Dutton, G.: Generalising spatial data and dealing with multiple representations. In: Longley, P. (Ed.) ; Goodchild, M. (Ed.) ; Maguire, D. (Ed.) ; Rhind, D. (Ed.): Geographic Information Systems - Principles and Technical Issues Vol. 1. 2nd Edition. John Wiley & Sons, 1999, P. 125–155.

Zhou, Sheng; Jones, Christopher B.: Shape-aware Line Generalisation with Weighted Effective Area. In: Fisher, Peter F. (Ed.): 11th International Symposium on Spatial Data Handling. Leicester, UK : Springer-Verlag, 23rd - 25th August 2004 2004, P. 369–380.