A FORMAL MODEL AND MIXED-INTEGER PROGRAM FOR AREA AGGREGATION IN MAP GENERALIZATION

Jan-Henrik Haunert

Institut für Kartographie und Geoinformatik
Leibniz Universität Hannover
Appelstraße 9a, 30167 Hannover, Germany
jan.haunert@ikg.uni-hannover.de

KEY WORDS: GIS, Generalization, Land Cover, Abstraction, Automation, Modelling

ABSTRACT:

This paper presents a model and an optimization method for a problem that appears when reducing the scale of a topographic database. Such a database commonly contains areas of different land cover classes that define a planar subdivision. When reducing its scale, some areas become too small and need to be aggregated. In order to produce contiguous aggregates that are not smaller than a user-defined threshold, it is necessary to change the classes of some areas. As generalization intends to preserve the characteristic features of the map, we aim to change classes as little as possible. A second objective is to create simple, compact shapes. Based on a previous work that neglected this second objective, we define a more general problem in this paper that reflects both aims of generalization. The problem was proven to be NP-hard, meaning that it is unlikely to find an efficient solution. Therefore, we propose a mixed-integer program (MIP) and heuristics, which enable the production of near-optimal results. The paper concludes with the presentation of some results we obtained using our method.

1 INTRODUCTION

In topographic databases information about land use or land cover is commonly represented by areas that are assigned to different classes, such as settlement, water, or different kinds of vegetation. The areas in such a database collectively define a subdivision of the plane, i.e., overlaps and gaps are not allowed. Generalizing this kind of map requires algorithms for different problems (Bader and Weibel (1997)). A challenging task is the aggregation of areas, which aims to satisfy size thresholds for the target scale. In an earlier paper we proposed a method for this generalization problem based on mixed-integer programming - a technique for combinatorial optimization (Haunert and Wolff (2006)). This method ensures different kinds of constraints coming from the specifications of the data sets and produces solutions with minimum change of land cover classes. The results were promising, but it was observed that the resulting geometries were not compact. Figure 1 (left) shows an example of a map at the original scale and the result which was obtained according to the defined objective (right). The settlement in the result (red) contains a narrow isthmus that was created to satisfy the area constraint while expending a minimum cost for class changes. In order to avoid such complex shapes, we define compactness as additional objective in this paper. The possibilities for the application of different compactness measures are discussed.

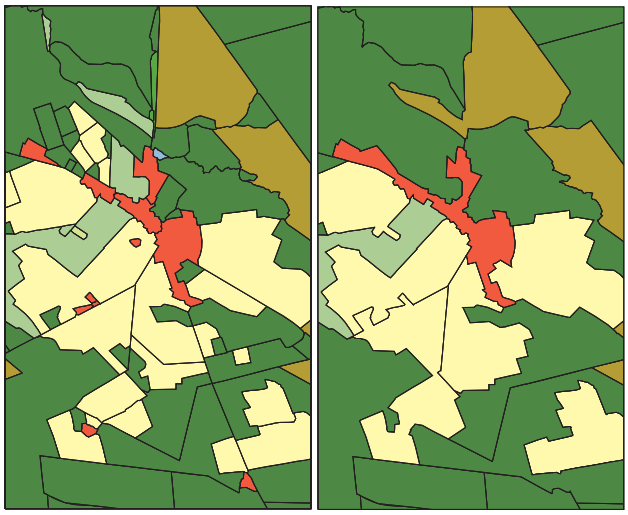


Figure 1: An example from the input data set at scale 1:50.000 (left) and a result for the scale 1:250.000 when minimizing changes of classes.

When designing optimization problems always two things need to be taken into account: The adequacy of the optimization objective and the possibility to solve the problem. Because of this, we concentrate on compactness measures that can be expressed by linear expressions. Regrettably, with this requirement, it is not possible to express size-invariant compactness measures. We discuss this deficit and its effects in detail. To cope with this, we add further requirements to the problem.

The paper is structured as follows: In Section 1.1 we discuss related work. Section 2 gives a formal problem definition and discusses the possibilities and difficulties to model compactness. In Section 3 we define a problem with additional requirements that allows to better express the cartographer's aim of generalization while getting along with the defined measures. In Section 4, we present our new mixed-integer program, results and an outline of an approach for the processing of large data sets. Finally we give a conclusion.

1.1 Related Work

The problem of area aggregation in map generalization has extensively been analyzed by researchers (Timpf (1998); van Smaalen (2003)). However, from an algorithmic point of view little success has been made in tackling its combinatorial nature. Different researchers have proposed iterative methods for the area aggregation problem. The following algorithm is described by van Oosterom (1995):

In each iteration the feature with lowest importance is selected. The selected feature is merged with a neighbor, which is chosen

according to a collapse function, and the next iteration is processed. The iteration can be terminated, if all areas satisfy the minimal dimension that is required for the target scale.

Many proposed algorithms are specializations of this general method. Jaakkola (1997) uses the method within a more comprehensive generalization framework for raster based land cover maps. Podrenek (2002) discusses preferences for merges, which reflects the collapse function. Generally, semantic similarity of classes, boundary lengths and area sizes are considered as criteria that need to be incorporated into the collapse function. The main problem with these iterative approaches is that consequences for future actions are not taken into account, when greedily selecting a neighbor. Therefore, we present a global approach in this paper.

Though there has not been any global optimization approach to area aggregation in map generalization, there exists a multiplicity of related problems that have been investigated by researchers. Especially, in the field of operations research, optimization methods for districting and aggregation problems have been developed. A typical application is the definition of sales districts presented by Zoltners and Sinha (1983). Their solution to find optimal districts is based on mathematical programming. As it is aimed to minimize distances between customers and stores, compactness is also aimed in their approach. We discuss the applied measure in Section 2.2.1. Other researchers have applied metaheuristics such as simulated annealing (Bergey et al. (2003)). The major disadvantage of these methods is the requirement for the definition of several tuning parameters, which are not inherent to the aggregation problem. Because of this, we concentrate on mathematical programming.

2 AGGREGATION PROBLEM

2.1 General Problem Statement

In this section, we first give a formal problem definition, which models the requirements and objectives of area aggregation in map generalization, and then explain this definition in detail. We simply refer to this problem as “Area Aggregation”.

Given

- a planar graph $G(V, E)$ with node weights $w : V \to \mathbb{R}^+$ and a coloring of nodes $\gamma : V \to \Gamma$, where Γ is the set of all colors, i.e. land cover classes,
- a function $\theta : \Gamma \to \mathbb{R}^+$, defining minimal allowed weights for colors,
- a function $d : \Gamma^2 \to \mathbb{R}_0^+$, expressing a distance between colors,
- a function $c : 2^V \times \Gamma \to \mathbb{R}_0^+$, defining the non-compactness of an aggregate,
- and a scalar weight factor $s \in [0, 1]$,

define a new coloring $\gamma' : V \to \Gamma$ of nodes and find a partition $P = \{V_1, V_2, \ldots, V_p\}$ of V, such that

- for each node set $V_i \in P$
  - the graph induced by V_i is connected,
  - all nodes in V_i receive the same new color $\gamma_i' \in \Gamma$, i.e., $\gamma'(v) = \gamma_i' \quad \text{for all} \quad v \in V_i$,
  - there is at least one node $v \in V_i$ with unchanged color, i.e., $\gamma'(v) = \gamma(v)$,
  - and V_i has total weight at least $\theta(\gamma_i')$,
- and the cost

  $s \cdot \sum_{v \in V} w(v) \cdot d(\gamma(v), \gamma'(v)) + (1 - s) \cdot \sum_{V_i \in P} c(V_i, \gamma_i')$

  is minimized.

The graph G is the dual graph of the planar subdivision. It contains a node for each shape and an edge between two nodes if the corresponding shapes share a common boundary. Node weights represent the sizes of areas. Land cover classes are represented by colors.

The defined requirements for connectivity and weight feasibility come from the specifications of data sets. Such specifications have been introduced as data standards by mapping authorities. To model minimal allowed area sizes that are defined for different land cover classes in the target scale, the weight threshold θ is defined as a function of color. Additionally, the requirement for a node with unchanged color in each part is introduced, to avoid that new classes pop up in the generalized map. Throughout this paper, such a node, which defines the color of an aggregate, will be referred to as *center*. Note that in this definition each node is a potential center. Figure 2 shows an instance of the problem and a solution, which is feasible according to the defined requirements. The partition P defines the shapes for the target scale, which can be obtained by geometrical union of shapes that correspond to the nodes contained in each element $V_i \in P$.

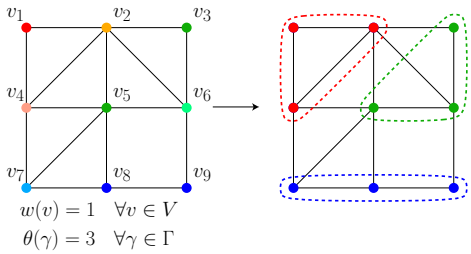


Figure 2: An instance of the aggregation problem (left) and a solution with $P = \{\{v_1, v_2, v_4\}, \{v_3, v_5, v_6\}, \{v_7, v_8, v_9\}\}$ (right).

The objective function expresses the cartographer’s preferences for different feasible solutions. Two different objectives can be identified: Firstly, it is aimed to change the original classes as little as possible. Secondly, compact shapes are preferred. To model these two aims, the two functions d and c are introduced, which are combined in a weighted sum. These functions need to be explained in detail.

The function d defines costs that are charged to change an area of unit size from one color into another. The values of this function could be given explicitly by a quadratic matrix with $|\Gamma| \times |\Gamma|$ elements. Generally, this matrix is not symmetric: Objects of rare land cover classes are often considered more important than others. It is unwanted to loose these objects; because of this, one will rather change a frequent class into a rare class than vice versa.

The function c defines a penalty being charged for the non-compactness of an aggregate, i.e., an area in the target scale that is defined by a subset of nodes and their new color (2^V refers to the power set of V, i.e., the set of all subsets of V). We assume that c attains high values for complex shapes, but we will simply

use the term compactness measure in the following. It is clear that explicitly expressing the values of this function is prohibitive due to limited time and storage space. However, no assumptions are made for this function here, in order to allow for different variations of the problem.

The complexity of this problem was investigated for the special case that compactness is neglected, i.e., for $s = 1$. Even for this case, the described problem was proven to be NP-hard, which means that it is very unlikely to find an efficient algorithm that reaches the optimal solution (Haunert and Wolff (2006)). This is still true, if $|\Gamma| = 2$, i.e., if the map only contains areas of two different classes. The NP-hardness justifies that we will later turn to mixed-integer programming and to heuristics.

2.2 Definition of Compactness

In the last section, the objective for compactness of shapes was expressed in a very general sense by the function c. This definition is specified here. Many different compactness measures have been proposed by researchers for the analysis of shapes. A detailed discussion is given by Maceachren (1985). The presented measures could be expressed as objectives to be optimized for the area aggregation problem.

However, before deciding for a specific model, it needs to be pointed out that this definition will influence the solvability of the problem. For this reason, we only consider those measures here that can be expressed by means of linear expressions, including integer and fractional variables. Additionally, we would like to bound the number of variables by a polynomial of low degree, e.g., quadratic in the input size. With this in mind, we discuss the possibilities to model compactness.

2.2.1 A Measure based on Distances to a Center A simple measure of compactness is defined by Zoltners and Sinha (1983). The aggregate defined by the set $V' \in 2^V$ contains a node $u \in V'$ whose corresponding shape defines the *geometrical center* of the aggregate by its centroid. The aggregate is considered to be compact, if the centroids of all other nodes in V' are close to this. To respect different sizes of areas, a penalty is charged for each node, which is equal to the product of the node's weight and the distance of its centroid from the geometrical center.

To apply this measure here, we claim that the geometrical center and the center according to Section 2.1 are defined by the same node. This definition slightly eases the problem, because less variables are needed to express the model. Nevertheless, it is reasonable, since it is undesired that nodes with unchanged color only appear at the margin of an aggregate. In other words, it is preferred, that nodes "gather around" a center of unchanged color. If there are several nodes with this color, then, among these, the center is defined by the node for which the overall penalty is minimal.

To formalize this measure, let $\delta : V^2 \to \mathbb{R}_0^+$ be the Euclidean distance between centroids of the shapes corresponding to two nodes. With this, we define the measure $c_1 : 2^V \times \Gamma \to \mathbb{R}_0^+$ as

$$c_1(V', \gamma') = \min \left\{ \sum_{v \in V'} w(v) \cdot \delta(v, u) \mid u \in V' \wedge \gamma(u) = \gamma' \right\} . \quad (1)$$

The function c_1 attains high values for complex shapes. Certainly, this measure only coarsely reflects the geometrical characteristic of a shape, since shapes are approximated by centroids. Because of this, the aggregate's perimeter is introduced as a second measure.

2.2.2 Measuring Compactness by the Perimeter of a Region The previously discussed iterative approaches to area aggregation in map generalization usually consider the length of boundaries as criterion when choosing a neighbor for merging (van Oosterom (1995)). To formalize the perimeter of an aggregate, let $\lambda : E \to \mathbb{R}^+$ be the length of the common boundary between two areas. Now, the perimeter $c_2 : 2^V \to \mathbb{R}_0^+$ becomes

$$c_2(V') = \sum_{e \in E'} \lambda(e) , \quad (2)$$

with E' being the set of edges incident to one node in V', i.e.,

$$E' = \left\{ \{u, v\} \in E \mid \left| \{u, v\} \cap V' \right| = 1 \right\} .$$

Similar to c_1, the compactness measure c_2 attains high values for complex shapes, which supposably have greater perimeters.

2.2.3 Discussion of Proposed Measures Both measures can result in side-effects, when being applied as global objectives. These need to be discussed. The measures c_1 and c_2 can result in two different biases:

1. When minimizing $\sum_{V_i \in P} c_1(V_i, \gamma_i')$, solutions with many small aggregates are preferred compared to solutions with few large aggregates. This is simply because average distances to centers are shorter for smaller aggregates.
2. When minimizing $\sum_{V_i \in P} c_2(V_i)$, solutions with few large aggregates are preferred compared to solutions with many small aggregates. In fact, when neglecting the objective for minimal color change, the globally optimal result would contain only one single aggregate, since in this case the total boundary length of the resulting partition would be minimal.

Both effects are due to the fact that the measures are not size invariant. In order to avoid these effects, the functions c_1 and c_2 could be normalized. However, we cannot satisfy the earlier claimed possibility for modeling the objective by means of linear expressions when using size invariant compactness measures.

It is important to note that aggregates will not become too small when applying c_1, since the size of each aggregate is bounded from below by the threshold θ. However, when applying the measure c_2 we run the risk of creating unintentionally large aggregates. To avoid this danger, we add additional hard requirements to the problem statement from Section 2.1. A detailed explanation of this method is given in Section 3.

3 AN APPROACH BASED ON PREDEFINED CENTERS

To avoid the creation of too large aggregates one could define an upper bound for the weights of the elements in the partition P, or a lower bound for the number of elements in P. Both definitions are probably too global and do not take local differences in the data set into account. Because of this, we chose another approach. It is based on a set of nodes that are predefined as centers. We give a general outline of this approach, formalize the modified problem and explain the definition of centers in detail.

3.1 Outline of Approach

In our previous paper (Haunert and Wolff (2006)), we proposed a heuristic that allowed for the elimination of certain variables. The idea was to fix relatively large areas as centers of aggregates. This

resulted in solutions with slightly higher values for the objective function, i.e., the total change of colors increased approximately by 10%. In the same way we subjectively perceived a decrease of quality.

Our first experiments with the proposed compactness measures, however, revealed that the defined objective function does not suffice to model our aim: Without fixed centers, the aggregates did not become compact enough when giving low weights to c_2 and the aggregates became unintentionally large for higher weights. The reason for this effect was explained in Section 2.2.3. Nevertheless, by fixing centers we did obtain nice and compact results. Since each aggregate can contain at most one center, the expansion of aggregates is limited. We assume that this model sufficiently reflects the aims of area aggregation in map generalization if the set of fixed centers is reasonably defined. Because of this, we include the previously defined heuristic in the problem statement.

3.2 Modified Problem Statement

The modified problem is defined as generalization of the problem Area Aggregation from Section 2.1. We refer to this problem as "Area Aggregation With Predefined Centers". In addition to an instance of Area Aggregation we require a set of predefined centers $C \subseteq V$ as input and define the constraints that for each node set $V_i \in P$

- at most one center is contained, i.e, $|V_i \cap C| \leq 1$ and
- if V_i contains a center $v \in C$, then all nodes $u \in V_i$ receive the color of the center, i.e., $\gamma'(u) = \gamma'_i = \gamma(v)$.

With the concept of predefined centres, the measure c_1 is generalized by the function $c_3 : 2^V \times \Gamma \rightarrow \mathbb{R}_0^+$:

$$c_3(V', \gamma') = \begin{cases} \sum_{v \in V'} w(v) \cdot \delta(v, u) & \text{if a node } u \text{ is in } V' \cap C\,, \\ c_1(V', \gamma') & \text{else, i.e., if } V' \cap C = \emptyset\,. \end{cases} \quad (3)$$

This simply means that if there is a predefined center in V', then the corresponding centroid defines the geometrical center of the aggregate, which is used to measure the compactness. Note that for $C = \emptyset$ the problem is the same as the original problem. Because of this, it is also NP-hard.

3.3 Definition of Centers

To define the set C, we recall that the application of c_1 as global objective is rather unproblematic, i.e., the influence on the size of aggregates is limited due to strict lower bounds for their weight. However, it was argued that the geometrical compactness is only coarsely reflected. We therefore propose a two-steps approach:

1. We solve the problem first, expressing compactness solely by the distances to a center, i.e., $c := c_1$. Presumably, such a solution is close to the result wanted by a cartographer.
2. Based on the resulting partition P, we define the set C to contain one node for each element in P. For this, we chose the center according to the measure c_1. With this definition, the problem can be solved a second time, this time applying the following combination of measures:

$$c := s' \cdot c_3 + (1 - s') \cdot c_2, \quad s' \in [0, 1] \quad (4)$$

The scalar weight factor s' is introduced to define a compromise of the two objectives c_3 and c_2. We discuss a solution of the problem in the next section.

4 A MIP FOR AREA AGGREGATION WITH PREDEFINED CENTERS

Different possibilities exist to model the aggregation problem as MIP. A difficult task is to express the connectivity of aggregates by means of variables and linear constraints. Williams (2002) and Shirabe (2005) have found different solutions for the problem of ensuring connectivity when selecting a subset of nodes from a graph. These approaches can be adopted in a straight forward way to model the aggregation problem, leading to a quadratic number of variables and constraints. Both methods have been implemented and tested using the software ILOG CPLEX 9.100 on a Linux server with 4 GB RAM and a 2.2 GHz AMD-CPU. In conclusion, the obtained running time was prohibitive – the largest instance that could be solved contained only 30 nodes. An improvement was made using a new MIP based on a single commodity flow model that requires only a linear number of variables and constraints (Haunert and Wolff, 2006). Still, without heuristics, it was not possible to process more than 50 areas.

Because of these experiences and the absence of existing approximation algorithms, heuristics need to be applied. Therefore we now define a more restrictive requirement for the connectivity of aggregates. This leads to an alternative MIP formulation. A similar approach was used by Zoltners and Sinha (1983) for the problem of optimally defining sales territories.

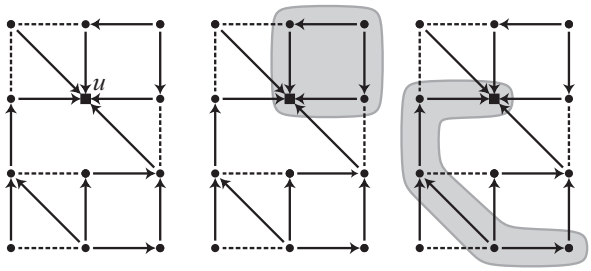
4.1 Connectivity based on Precedence Relationship

Evidently, in order to end up with connected parts, a node v can only be assigned to a distinct center u, if at least one of its neighbors is also assigned to u. However, it is important to note that this does not suffice. Consider two adjacent nodes being assigned to the same center: Both nodes will mutually satisfy their requirements without ensuring the connectivity to others, i.e., the problem with the neighbor relationship is that it contains cycles. To cope with this, we introduce the stricter, acyclic *precedence relationship*.

Given a graph $G(V, E)$ with edge lengths $\alpha : E \rightarrow \mathbb{R}^+$, a center $u \in V$ and a node $v \in V\,, u \neq v$, we define the set of predecessors of v with respect to center u as

$$\mathrm{Pred}_u(v) := \{w \in V \mid \mathrm{D}(u\,, w) < \mathrm{D}(u\,, v) \wedge \{v, w\} \in E\}\,, \quad (5)$$

with $\mathrm{D}(i\,, j)$ being the length of the shortest path in G from i to j using edge lengths α. Different possibilities exist for defining the edge lengths α. The definition which is applied here is based on the minimal size of a potential aggregate containing u and v. This definition is discussed in our earlier paper. An example for the precedence relationship with the setting of equal edge length is illustrated in Figure 3(a). Arcs are drawn from each node $v \in V$ to its predecessors $\mathrm{Pred}_u(v)$. The resulting directed graph is acyclic and the center u is the only terminal. For some edges of the adjacency graph, both incident nodes have the same distance to the center. These edges are displayed as dashed lines. However, when using non-uniform edge lengths, these cases are rare exceptions.



(a) Precedence relationship Pred_u.

(b) A feasible aggregate according to Pred_u.

(c) A connected aggregate that is not feasible according to Pred_u.

Figure 3: Precedence relationship with respect to center u (displayed as square) and feasibility of aggregates in the presented MIP.

The definition of the precedence relationship can be used to define a simple requirement that ensures connectivity of aggregates: A node may only be assigned to a distinct center if at least one of its predecessors with respect to this center is also assigned to it. The constraint clearly forbids disconnected aggregates since the center can always be reached from an assigned node via predecessors, i.e., without leaving the aggregate. Figure 3(b) shows an example which satisfies the requirement. However, by defining this requirement several connected aggregates will be excluded. An example for this is displayed in Figure 3(c). The aggregate does not contain any predecessor of the node located in the bottom right corner. Zoltners and Sinha legitimate the restriction of their model with their preference for compact sales districts: The non-feasible connected parts likely are non-compact. So, they can probably be excluded without loosing good solutions. As we also aim compactness, their model is a reasonable approach. However, since the optimal solution might be missed, we refer to it as heuristic. Comparisons with results that were attained with a model allowing for connectivity in a general sense, i.e., with our flow model, have shown that this heuristic only marginally affects the result. The attained results for the objective function were at most 5% worse than the optimum. The processing time, however was greatly reduced. As mentioned in the introduction, our tests were made without consideration of compactness. It is likely that the results are even closer to the optimum when defining compactness as additional objective.

4.2 MIP Formulation

In this section, we present a MIP hat models the requirements and objectives of the problems from Sections 2.1 and 3.2. We first introduce the formulation as a whole and then explain it in detail.

$$x_{uv} \in \{0,1\}\,, \quad \text{with } x_{uv} = 1 \text{ if node } v \in V \text{ belongs to center } u \in V\,.$$

$$y_{ue} \in [0,1]\,, \quad \text{with } y_{ue} = 0 \text{ if at least one incident node of } e \in E \text{ does not belong to center } u \in V\,.$$

Minimize

$$\begin{aligned} & s \cdot \sum_{u\in V}\sum_{v\in V} w(v)\cdot x_{uv}\cdot d(\gamma(v),\gamma(u)) \\ +(1-s)\cdot s' \cdot & \sum_{u\in V}\sum_{v\in V} w(v)\cdot x_{uv}\cdot \delta(v,u) \\ -(1-s)\cdot(1-s')\cdot & \sum_{u\in V}\sum_{e\in E} 2\cdot\lambda(e)\cdot y_{ue} \end{aligned} \tag{6}$$

subject to

$$\sum_{u\in V} x_{uv} = 1 \qquad \forall v \in V\,, \tag{7}$$

$$\sum_{v\in V} w(v)\cdot x_{uv} \geq \theta(\gamma(u))\cdot x_{uu} \qquad \forall u \in V\,, \tag{8}$$

$$\sum_{w\in \mathrm{Pred}_u(v)} x_{uw} \geq x_{uv} \qquad \forall u, v \in V : u \neq v\,, \tag{9}$$

$$\left.\begin{array}{rcl} y_{ue} & \leq & x_{uv} \\ y_{ue} & \leq & x_{uw} \end{array}\right\} \forall u \in V, \quad e = \{v,w\} \in E\,. \tag{10}$$

The binary variables x_{uv} define the solution of the problem: All nodes u with $x_{uu} = 1$ constitute the set of centers that define the color and geometrical centers of aggregates. To assign a node v to a center u, x_{uv} needs to be set to 1. The cost for the color change that is charged for this assignment is defined by the first term in Equation 6. The second term defines the cost for compactness according to Equation 1. The third term defines a benefit for each edge $e = \{v, w\}$ that is totally contained in one aggregate, i.e., there is a center $u \in V$ with $x_{uv} = 1$ and $x_{uw} = 1$. Auxiliary variables y_{ue} are defined to express this case. Giving a benefit for interior edges has the same result as charging a cost for the perimeter, as the objective was defined in Equation 2. The factor 2 is needed, as each edge belongs to the boundaries of two areas.

We now describe our set of constraints. Constraint 7 expresses that each node must be assigned to exactly one center. Constraint 8 does not have any effect for $x_{uu} = 0$, i.e., if u is not selected as center. For $x_{uu} = 1$ it ensures that the aggregate with center u is weight feasible, i.e., the threshold for the target scale is satisfied. Constraint 9 ensures connectivity according to the precedence relationship, as defined in Section 4.1: Node v can only be assigned to center u if there is also a predecessor w which is assigned to u. Finally, Constraint 10 is used to couple the variables x_{uv} and those of type y_{ue}: If one of the incident nodes of e, i.e., v or w, is not assigned to center u, then y_{ue} is forced to be 0 and no benefit will be given. Otherwise, the constraint defines that $y_{ue} \leq 1$. Since a benefit proportional to y_{ue} is given, y_{ue} will always take the value of its upper bound. Because of this y_{ue} will be 1 for edges included in aggregates. Thus we do not have to make y_{ue} explicitly a 0-1 variable, which usually speeds up MIP solvers.

Our MIP models both, the problem from Section 2.1 and the modified problem with predefined centers. In the second case, it is possible to simply define $x_{uu} = 1$ for all u in C. Additional variables can be fixed after this. To define a MIP without the restricting precedence relationship, one can simply replace Constraint 9 by formulations that have been presented by Williams (2002) and Shirabe (2005). These, however, require additional auxiliary variables. In our previous paper, we presented two additional heuristics that can be applied to speed up the processing. The first is to set $x_{uu} = 0$ for nodes u with very small weights, i.e., to exclude them from the set of potential centers. The second heuristic is to set $x_{uv} = 0$, if the distance between u and v is large. These heuristics have been formally defined and discussed in detail. We present results of this method and the addition of criteria for compactness in the next section.

4.3 Results

We used the presented formulation to express the problem as a MIP and solved it by application of standard branch-and-cut methods. The performance was similar to the MIP without the application of the compactness objective, which has extensively

been tested in our previous paper. With the presented heuristics it is possible to solve instances with 400 areas in modest time, i.e., less than one hour. Figure 4 shows the same sample as Figure 1, but this time the proposed measure for compactness was applied in combination with the objective for minimum class changes. The resulting aggregates are clearly more compact. As in the first example, the red settlement was saved by sacrificing smaller neighbors, but instead of building a narrow bridge to smaller areas of the same color, a neighbor on the right side was included, leading to a simpler shape. However, the resulting map certainly does not constitute a finished product. For example, one would need to apply a line simplification algorithm to remove further details. Nevertheless, it is perceived that the formulated optimization problem sufficiently models the aims of aggregation.

4.4 Processing Large Data Sets

In the presented form, the method is still not suitable for cartographic production, as the expense of time is too high. To process large datasets we have developed a heuristic approach (Haunert (2007)). The idea is to predefine each node $v \in V$ with $w(v) \geq \theta(\gamma(v))$ as center, i.e., those areas in the original scale that are sufficiently large for the target scale. Let G' be the sub graph of G that is induced by all other nodes, then the aggregation problem with the compactness measures from Section 2.2 can be solved independently for each connected component of G'. This fact allows to decompose the problem into smaller instances. However, for our data set, the resulting instances are still too large to be processed. We have solved this problem by definition of intermediate size thresholds, such that the number of predefined centers increases until the problem instances are manageable, i.e., do not contain more nodes than a user-specified number k. Via these intermediate scales, the target scale can be reached in several steps. We have shown that our method generalizes the existing iterative method of van Oosterom (1995), i.e., for $k := 1$ both methods are the same. However, for a complete map sheet of a topographic map, our method with $k := 200$ resulted in 20% less class change, 2% less cost for non-compact shapes and 8% less total cost.

5 CONCLUSION

We have proposed a new method for the aggregation of areas in a planar subdivision that takes compactness and class similarity into account and enables the application of mixed-integer programming. With this restriction, we could model compactness only by adding requirements that avoid the creation of too large aggregates. To define these requirements, we developed a two-step approach. First we apply a coarse measure of compactness for the definition of centers and second we create a high-quality map by applying a more sophisticated measure. The obtained results showed that this approach satisfies the aims of aggregation in map generalization. Due to the NP-hardness of the problem, heuristics needed to be introduced to solve instances of interesting size. We gave an outline of a heuristic approach that decomposes the problem into manageable instances.

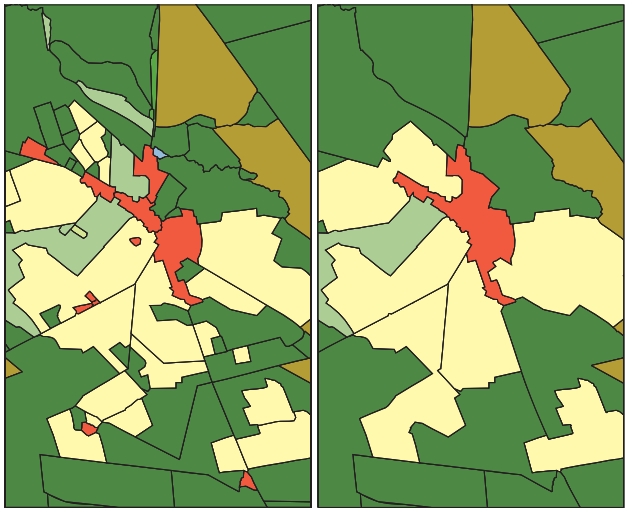


Figure 4: An example from the original map at scale 1:50.000 (left) and a result of the proposed method (right). The sample contained 200 areas in the original map. These were grouped into 32 aggregates. The processing took 9s.

References

Bader, M. and Weibel, R., 1997. Detecting and resolving size and proximity conflicts in the generalization of polygonal maps. In: Proc. 18th International Cartographic Conference, Stockholm, Sweden, pp. 1525–1532.

Bergey, P. K., Ragsdale, C. T. and Hoskote, M., 2003. A simulated annealing genetic algorithm for the electrical power districting problem. Annals of Operations Research 121, pp. 33–55.

Haunert, J.-H., 2007. Efficient area aggregation by combination of different techniques. In: Proc. of 10th ICA Workshop on Generalisation and Multiple Representation, 2–3 August 2007, Moscow, Russia.

Haunert, J.-H. and Wolff, A., 2006. Generalization of land cover maps by mixed integer programming. In: Proc. 14th Int. ACM Sympos. Advances in Geographic Information Systems (ACM-GIS '06), ACM Press, New York, pp. 75–82.

Jaakkola, O., 1997. Quality and Automatic Generalization of Land Cover Data. PhD thesis, Department of Geography, University of Helsinki.

Maceachren, A. M., 1985. Compactness of geographic shape: Comparison and evaluation of measures. Geografiska Annaler. Series B, Human Geography 67(1), pp. 53–67.

Podrenek, M., 2002. Aufbau des DLM50 aus dem Basis-DLM und Ableitung der DTK50 – Lösungsansatz in Niedersachsen. Karthographische Schriften Band 6, Kartographie als Baustein moderner Kommunikation, pp. 126–130.

Shirabe, T., 2005. A model of contiguity for spatial unit allocation. Geographical Analysis 37, pp. 2–16.

Timpf, S., 1998. Hierarchical structures in map series. PhD thesis, Technical University Vienna, Austria.

van Oosterom, P., 1995. The gap-tree, an approach to 'on-the-fly' map generalization of an area partitioning. In: J.-C. Müller, J.-P. Lagrange and R. Weibel (eds), GIS and Generalization - Methodology and Practice, GISDATA, Taylor & Francis, London.

van Smaalen, J. W. N., 2003. Automated Aggregation of Geographic Objects. PhD thesis, Wageningen University, The Netherlands.

Williams, J. C., 2002. A zero-one programming model for contiguous land acquisition. Geographical Analysis 34(4), pp. 330–349.

Zoltners, A. A. and Sinha, P., 1983. Sales territory alignment: A review and model. Management Science 29(11), pp. 1237–1256.