Political Districting for Elections to the German Bundestag: An Optimization-Based Multi-Stage Heuristic Respecting Administrative Boundaries

Sebastian Goderbauer

1 Electoral Districts in Elections to the German Bundestag

In general, the election to the German federal parliament, the Bundestag, takes place every four years. The 299 electoral districts play an important role in those elections. In fact, the voters of each district elect one representative into parliament ensuring that each part of the country is represented. These elected representatives make up half of the members of the Bundestag. The allocation of electoral districts needs regular updates due to an ever-changing population distribution and is subject to a variety of legal requirements as listed in the following.

In order to comply with the principle of electoral equality as anchored in the German constitution, the differences in population between the districts have to be preferably small. The law defines a tolerance limit, saying that the amount of deviation from the average district population should not exceed 15%. Moreover, an amount of deviation beyond 25% is illegal. Every district should be a contiguous area and it is prefered that its allocation aligns with existing administrative boundaries. In addition to that, the law demands that the districts strictly comply with the borders of the German federal states. The law specifies that the Sainte-Laguë method [7] has to be used to distribute the 299 districts among the 16 states. The electoral districts ought to be visually compact counteracting the suspicion of applying Gerrymandering [6, 9, 11]. In the context of setting electoral districts, Gerrymandering is a practice that attempts to create an advantage or disadvantage for a certain political party or candidate by manipulating district boundaries.

In this contribution, which is an extended abstract of the author's master's thesis, the problem of dividing a country into electoral districts is defined as a multi-criteria graph partition problem. To solve this regularly current practical problem, an optimization-based multi-stage heuristic is introduced and successfully applied

Sebastian Goderbauer
RWTH Aachen University, Operations Research, Kackertstrasse 7, 52072 Aachen, e-mail: goderbauer@or.rwth-aachen.de

to population data of the latest German census [12]. The computed results show that the presented algorithm allocates electoral districts, which are not only in accordance with the law, but also fulfill the tolerances mentioned in the law more closely than the current districting.

2 The Political Districting Problem

The Political Districting Problem is defined on the basis of a so-called population graph. In a population graph

$$G = (V, E)$$

a node $i \in V$ represents an geographical area, e.g., the area of a municipality, and is weighted with its population p_i. An undirected edge $(i, j) \in E$ with nodes $i, j \in V$ exists, iff the corresponding areas share a border. The Political Districting Problem is an optimization problem in which a node-weighted population graph has to be partitioned into a given number of connected, weight-restricted subgraphs.

More precisely, let S be the set of all 16 German states. Given the total number of electoral districts $d \in \mathbb{N}$, which has to be set, the number of districts $d(s) \in \mathbb{N}$ for each state $s \in S$ is computable with the mentioned Sainte-Laguë method [7]. Of course, $\sum_{s \in S} d(s) = d$ holds. Furthermore let $\varnothing_p := \frac{1}{d} \sum_{i \in V} p_i$ be the average population of an electoral district. Finally, a partition

$$D_k \subseteq V,\, k = 1, \dots, d \quad \text{with} \quad D_l \cap D_m = \emptyset,\, l \neq m \quad \text{and} \quad \cup_k D_k = V$$

is called a feasible solution (districting) for the Political Districting Problem, if the following holds:

$$\forall 1 \leq k \leq d \;\; \forall i, j \in D_k : i \text{ and } j \text{ are in same state}, \tag{1}$$

$$\forall s \in S : |\{D_k : \text{state } s \text{ contains district } D_k\}| = d(s), \tag{2}$$

$$\forall 1 \leq k \leq d : G[D_k] \text{ connected}, \tag{3}$$

$$\forall 1 \leq k \leq d : 0.75\varnothing_p \leq \sum_{i \in D_k} p_i \leq 1.25\varnothing_p. \tag{4}$$

To complete the definition and as a result of analyzing the legal requirements, the multi-criteria objective is as follows:

$$\mathbf{max}\, |\{D_k : 1 \leq k \leq d \text{ and } 0.85\varnothing_p \leq \sum_{i \in D_k} p_i \leq 1.15\varnothing_p\}|, \tag{5}$$

$$\mathbf{min}\;\; \text{amount of deviations between district population} \sum_{i \in D_k} p_i \text{ and } \varnothing_p, \tag{6}$$

$$\mathbf{max} \text{ match between district and existing administrative boundaries}, \tag{7}$$

$$\mathbf{max} \text{ geographical and visual compactness of the districts}. \tag{8}$$

Fig. 1 The population graph of North Rhine-Westphalia (NRW) consists of 396 nodes and 1 084 edges at borough level. All in all, there are 11 339 boroughs in Germany, thus NRW's population graph is one of the smaller ones.

To obtain the population graph of Germany and thus the required graphs for each German state, data of the latest German census [12] were combined with geoinformation [8] (cf. Fig. 1).

The Political Districting Problem includes incomparable and conflicting objective criteria. Analyzing the complexity of the subproblems, in which only one objective is considered, Altman [1] concludes the complexity of the Political Districting Problem.

Theorem 1. *The Political Districting Problem is NP-hard.*

To gain a profound understanding of the complexity, it is possible to analyze the graph partition problems on which the Political Districting Problem is based [2, 4, 5]. It is possible to compute feasible districtings in polynomial or even linear time on paths and special trees. Suitable partitions with minimal differences in population between the components can be found in polynomial time as well. Considering general trees a feasible districting is computable in polynomial time, but the problem gets NP-hard by adding the objective of minimizing the differences in population.

3 Optimization-Based Multi-Stage Heuristic

Since the 1960s the Political Districting Problem has been discussed and approached by many authors in operations research and social science. In 1961 Vickrey [9] provided a multi-kernel growth procedure. Although his proposal was rather informal and rudimentary it marked the start of a large variety of work on this topic.

Hess et al. [3] considered the problem of setting electoral districts as a modified facility location problem in 1965. The technique of column generation was applied by Mehrotra et al. [6] in 1998. In 2009 Yamada [10] formulated the problem as a spanning forest problem and presented a local search based heuristic.

Most heuristics and exact methods in the literature implement the requirements and objectives of the German Political Districting Problem only partially, e.g., support for matching district boundaries with existing administrative bounderies is disregarded. Beyond that, the numbers of nodes and edges in the population graphs of most German states outnumber all graph orders and sizes considered in the literature. The optimization-based heuristic described hereafter was developed to overcome these shortcomings. The multi-stage algorithm uses the existing hierarchical administrative divisions in Germany (cf. Fig. 2) and iteratively divides the Political Districting Problem into smaller subproblems. This has two advantages: The goal of aligning electoral district boundaries with existing administrative bounderies is realizable in an adequate way and in addition to that, the graphs of the subproblems will be of manageable size.

1st stage: states As mentioned above, it is required by law to align the districts with the boundaries of the German states. Therefore, a union of electoral districts for the individual states is a solution of the Political Districting Problem for Germany.

2nd stage: governmental districts The four most highly populated states are composed of so called governmental districts. For those states, the number of electoral districts of a state is distributed over the governmental districts by reapplying the Sainte-Laguë method [7]. With regard to the final solution, this is mostly a good and valid choice. If it is not valid, two neighbouring governmental districts are merged and seen as one in the application of the Sainte-Laguë method [7].

3rd stage: rural districs, urban districts Subsequently, the population graphs on rural and urban district level of a state (or governmental district) are considered. Intuitively, each node represents a rural or urban district. On this population graph

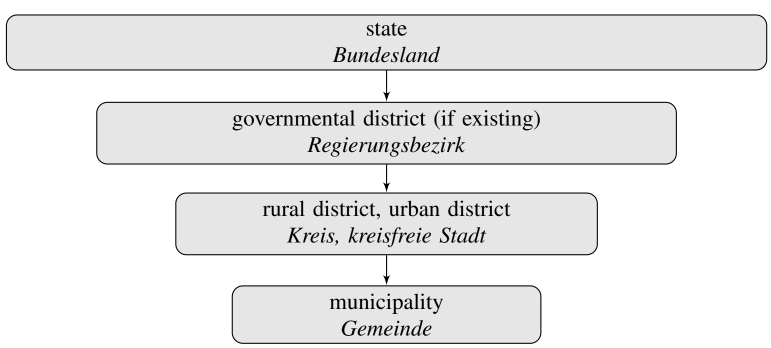
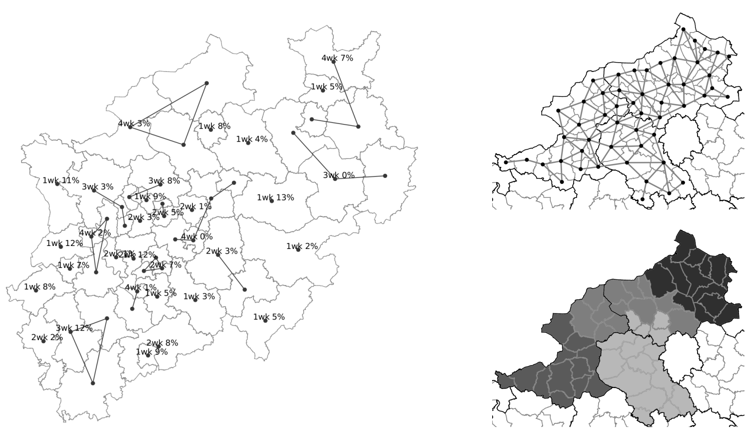


Fig. 2 The multi-stage algorithm is based on the structure of the hierarchical administrative divisions of Germany.

a modified set partitioning problem is solved: A graph partition into connected subgraphs is computed and a number of electoral districts is assigned to each subgraph. The sum of those numbers has to be equal to the number of districts assigned to that state (or governmental district) (cf. Fig. 3a). As the major part of the objective in this set partitioning problem the resulting average differences in district population are minimized. Components and thus subproblems with exactly one electoral district are solved, because the electoral district is already set.

4^{th} stage: municipalities Subproblems which are still open after the third stage are solved at the municipality level. For setting the remaining electoral districts almost all algorithms from the literature can be used, because the problem size is at this point mostly manageable. In this work a simple heuristic was implemented (cf. Fig. 3b).

The districtings computed by the optimization-base multi-stage heuristic are useable for the elections to the German Bundestag in general, because latest population data was used and the heuristic respects all legal requirements. It is worth to note that the developed algorithm supports matches between existing administrative boundaries and electoral district boundaries. This aspect is hard to implement in a compact formulation of the problem. Beside this, the computed districts follow the mentioned objectives transferred from the law more closely than the current districting applied in elections to the Bundestag in 2013.



a) b)

Fig. 3 a) A solution of the set partitioning problem at the third stage for North Rhine-Westphalia. b) The population graph on municipality level of the rural districts Borken, Coesfeld, and Steinfurt (*top*) and the allocation of four electoral districts after applying the heuristic (*bottom*).

4 In Search of an Optimal Number of German Electoral Districts

Currently, Germany is divided into 299 districts for elections to the Bundestag. The question arises if that is a well chosen number of electoral districts. In the following, one option of approaching this issue is outlined.

The Sainte-Laguë method [7] distributes the districts between the states on the basis of the state's population. However, the population of a state is usually not an integer multiple of the average electoral district population, thus differences in population between the districts are unavoidable. A question is which realizable number of German electoral districts causes the lowest maximal amount of average deviation in district population in a state. Considering the distribution of 299 electoral districts with the Sainte-Laguë method [7], the state Bremen has the highest average deviation of 17.2%. This value is greater than the tolerance limit of 15% specified in the law. However, when Germany is divided into 242 districts, the maximal average deviation is reduced to a mere 5.2% (Bremen again). This value is the minimum in the range between 1 and 376 distributed districts. It can therefore be concluded, that 242 districts would observe the rules of the law more closely than the current choice of 299 districts. In other words, 242 districts embody the German population distribution between the 16 states better than 299. Interestingly, when the number of districts is increased to 319 no legal districting can be found as the maximum deviation limit of 25% would always be exceeded in Bremen.

References

1. Altman, M.: Is Automation the Answer? The Computational Complexity of Automated Redistricting. Rutgers Computer and Law Technology Journal. **23(1)**, 81–142 (1997)
2. De Simone, C., Lucertini, M., Pallottino, S., Simeone, B.: Fair dissections of spiders, worms, and caterpillars. Networks. **20(3)**, 323–344 (1990)
3. Hess, S.W., Weaver, J.B., Siegfeldt, H.J., Whelan, J.N., Zitlau, P.A.: Nonpartisan Political Redistricting by Computer. Operations Research. **13(6)**, 998–1006 (1965).
4. Ito, T., Zhou, X., Nishizeki, T.: Partitioning a graph of bounded tree-width to connected subgraphs of almost uniform size. Journal of Discrete Algorithms. **4(1)**, 142–154 (2006)
5. Lucertini, M., Perl, Y., Simeone, B.: Most uniform path partitioning and its use in image processing. Discrete Applied Mathematics. **42(2-3)**, 227–256 (1993)
6. Mehrotra, A., Johnson, E.L., Nemhauser, G.L.: An Optimization Based Heuristic for Political Districting. Management Science. **44(8)**, 1100–1114 (1998)
7. Sainte-Laguë, A.: La représentation proportionnelle et la méthode des moindres carrés. Annales scientifiques de l'École Normale Supérieure. **27**, 529–542 (1910)
8. Shapefile der Verwaltungsgrenzen. Bundesamt für Kartographie und Geodäsie. (2013) - http://www.zensus2011.de, http://www.bkg.bund.de. Cited 30 Jul 2014.
9. Vickrey, W.S.: On the Prevention of Gerrymandering. Pol. Sci. Quar. **76(1)**, 105–110 (1961)
10. Yamada, T.: A Mini-max Spanning Forest Approach to the Political Districting Problem. Intern. J. Syst. Sci. **40(5)**, 471–477 (2009)
11. Young, H.P.: Measuring the Compactness of Legislative Districts. Legislative Studies Quarterly. **13(1)**, 105–115 (1988)
12. Zensus 2011: Bevölkerung - Ergebnisse des Zensus am 9. Mai 2011. Statistisches Bundesamt, Wiesbaden. (2014) - http://www.zensus2011.de. Cited 30 Jul 2014.