

Network-Based Dissolution

René van Bevern¹, Robert Brederick¹, Jiehua Chen¹, Vincent Froese¹,
Rolf Niedermeier¹, and Gerhard J. Woeginger²

¹ Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany

² Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands

Abstract. We introduce a graph-theoretic dissolution model that applies to a number of redistribution scenarios such as gerrymandering in political districting or work balancing in an online situation. The central aspect of our model is the deletion of certain vertices and the redistribution of their loads to neighboring vertices in a perfectly balanced way.

We investigate how the underlying graph structure, the pre-knowledge of which vertices should be deleted, and the relation between old and new vertex loads influence the computational complexity of the underlying graph problems. Our results establish a clear borderline between tractable and intractable cases.

1 Introduction

Motivated by applications in areas like political redistricting, economization, and distributed systems, we introduce a class of graph modification problems that we call *network-based dissolution*. We are given an undirected graph where each vertex carries a load consisting of discrete entities (e.g. voters, tasks, data). These loads are *balanced*: all vertices carry the same load. Now a certain number of vertices has to be *dissolved*, that is, they are to be deleted from the graph and their loads are to be redistributed among their neighbors such that afterwards all loads are balanced again.

Indeed, our dissolution problem comes in two flavors called DISSOLUTION and BIASED DISSOLUTION. DISSOLUTION is the basic version, as described in the preceding paragraph. BIASED DISSOLUTION is a variant that is motivated by gerrymandering in the context of political districting. It is centered around a bipartisan scenario with two types A and B of discrete entities. The goal is to find a redistribution that maximizes the number of vertices in which the A-entities form a majority. See Section 2 for a formal definition of these models.

Our main focus lies on analyzing the computational complexity of network-based dissolution problems, and in getting a good understanding of polynomial-time solvable and NP-hard cases.

Three application scenarios. We discuss three example scenarios for dissolution applications in some detail. The first and third example relate to BIASED DISSOLUTION, while the second example is closer to DISSOLUTION.

To appear in *Proceedings of the 39th International Symposium on Mathematical Foundations of Computer Science (MFCS '14)*, Budapest Hungary, August 2014.
© Springer.

Our first example comes from political districting, the process of setting electoral districts. Let us consider a situation with two political parties (A and B) and an electorate of voters that each support either A or B. The electorate is currently divided into n districts, each consisting of precisely s individual voters. A district is won by the party that receives the majority of votes in this district. The local government performs an electoral reform that reduces the number of districts, and the local governor (from party A) is in charge of the redistricting process. His goal is of course to let party A win as many districts as possible while dissolving some districts and moving their voters to adjacent districts. All resulting new districts should have equal sizes s_{new} (where $s_{\text{new}} > s$). In the *network-based dissolution* model, the districts and their neighborhoods are represented by an undirected graph: vertices represent districts and edges indicate that two districts are adjacent.

Our second example concerns economization in a fairly general form. Let us consider a company with n employees, each producing s units of a desirable good during an eight-hour working day; for concreteness, let us say that each employee proves s theorems per working day. Now, due to the increasing support of automatic theorem provers, each employee is able to prove s_{new} theorems per day ($s_{\text{new}} > s$). Hence, without lowering the total number of proved theorems per day, some employees may be moved to a special task force for improving automatic theorem provers: this will secure the company's future competitiveness in proving theorems, without decreasing the overall theorem output. By company regulations, all theorem proving employees have to be treated equally and should have identical workloads. In the *network-based dissolution* model, employees correspond to vertices. Employees in the special task force are dissolved and disappear from the scene of action; their workload is to be taken over by neighboring employees who are comparable in qualification and research interests.

Our third and last example concerns storage updates in parallel or distributed systems. Let us consider a distributed storage array consisting of n storage nodes, each having a capacity of s storage units, of which some space is free. As the prices on cheap hard disk space are rapidly decreasing, the operators want to upgrade the storage capacity of some nodes and to deactivate other nodes for saving energy and cost. As their distributed storage concept takes full advantage only in case all nodes have equal capacity, they want to upgrade all (non-deactivated) nodes to the same capacity s_{new} and move capacities from deactivated nodes to non-deactivated neighboring nodes. In the resulting system, every non-deactivated node should only use half of its storage capacity.

Related work. We are not aware of any previous work on our network-based dissolution problem. Our main inspiration comes from the area of political districting, and in particular from gerrymandering [8, 11, 12] and from supervised regionalization methods [5]. Of course, graph-theoretic models have been employed before for political districting; for instance Mehrota et al. [10] draws a connection to graph partitioning, and Duque [4] and Maravalle and Simeone [9] use graphs to model geographic information in the regionalization problem. These models are tailored towards very specific applications and are mainly used for the

purpose of developing efficient heuristic algorithms, often relying on mathematical programming techniques. The computational hardness of districting problems has been known for many years [1].

Remark on nomenclature. For the ease of presentation, throughout the paper we will adopt a political districting point of view on network-based dissolution: the words districts and vertices are used interchangeably, and the entities in districts are referred to as voters or supporters.

Contributions and organization of this paper. We propose two simple models DISSOLUTION and BIASED DISSOLUTION for network-based dissolution (Section 2). In the main body of the paper, we provide a variety of computational tractability and intractability results for both models. Furthermore, we investigate how the structure of the underlying graphs or an in-advance fixing of vertices to be dissolved influence the computational complexity (mainly in terms of polynomial-time solvability versus NP-hard cases).

- In Section 3, with network flow techniques we show that BIASED DISSOLUTION is polynomial-time solvable if the set of districts to be dissolved and the set of districts to be won are both specified as part of the input. The general version is NP-hard for every fixed $s \geq 3$.
- Section 4 presents a complexity dichotomy for DISSOLUTION and BIASED DISSOLUTION with respect to the old district size s and the increase Δ_s in district size (= difference between new and old district size). DISSOLUTION is polynomial-time solvable for $s = \Delta_s$, and BIASED DISSOLUTION is polynomial-time solvable for $s = \Delta_s = 1$; all other cases are NP-hard.
- Section 5 analyzes the complexity of DISSOLUTION and BIASED DISSOLUTION for various specially structured graphs, including planar graphs (NP-hard), cliques (polynomial-time solvable), and graphs of bounded treewidth (linear-time solvable if s and Δ_s are constant).

Due to the lack of space many proofs are only contained in the full version of the paper, which is available on arXiv (arXiv:1402.2664 [cs.DM]).

2 Formal setting

Let $G = (V, E)$ be an undirected graph representing n districts. Let $s, \Delta_s \in \mathbb{N}^+$ be the *district size* and *district size increase*, respectively. For a subset $V' \subseteq V$ of districts, let $Z(V', G) = \{(x, y) \mid x \in V' \wedge y \in V(G) \setminus V' \wedge \{x, y\} \in E(G)\}$ be the set of pairs of districts in V' and their neighbors that are not in V' . The central notion for our studies is that of a *dissolution*, which basically describes a valid movement of voters from dissolved districts into remaining districts. The formal definition is the following:

Definition 1 (Dissolution). Let $G = (V, E)$ be an undirected graph and let $D \subset V$ be a subset of *districts to dissolve* and $z: Z(D, G) \rightarrow \{0, \dots, s\}$ be a function that describes how many voters shall be *moved* from one district to its non-dissolved neighbors. Then, (D, z) is called an (s, Δ_s) -*dissolution* for G if

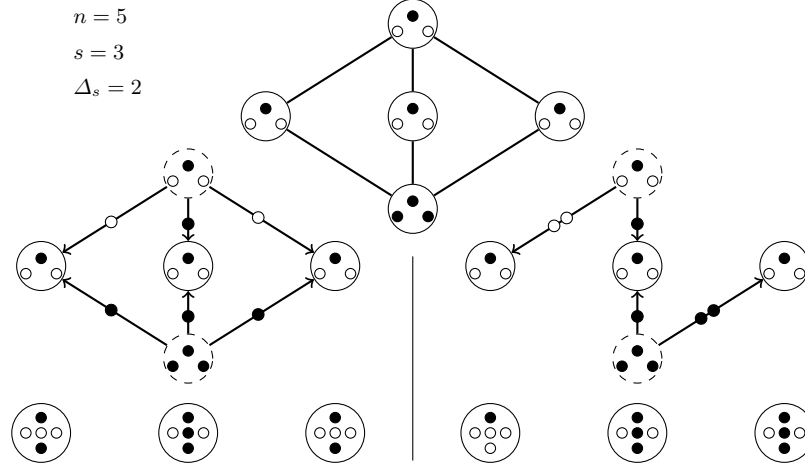


Fig. 1. An illustration of a 1-biased (3,2)-dissolution (left) and a 2-biased (3,2)-dissolution (right). Black circles represent A-supporters while white circles represent B-supporters. The graph on the top shows a neighborhood graph of five districts, each district consisting of three voters. The task is to dissolve two districts such that each remaining district contains five voters. The graphs in the middle show two possible realizations of dissolutions. The graphs on the bottom show the two corresponding outcomes. The arrows point from the districts to be dissolved to the “goal districts” and the black/white circle labels on the arrows indicate which kind of voters are moved along the arrows.

a) no voter remains in any dissolved district:

$$\forall v' \in D : \sum_{(v',v) \in Z(D,G)} z(v',v) = s, \text{ and}$$

b) the size of all remaining (non-dissolved) districts increases by Δ_s :

$$\forall v \in V \setminus D : \sum_{(v',v) \in Z(D,G)} z(v',v) = \Delta_s.$$

Throughout this work, we use $s_{\text{new}} := s + \Delta_s$ to denote the new district size, $d := |D| = |V| \cdot \Delta_s / s_{\text{new}}$ to denote the number of dissolved districts, and $r := |V| - d$ to denote the number of remaining, non-dissolved districts.

We write *dissolution* instead of (s, Δ_s) -dissolution when s and Δ_s are clear from the context. By definition, a dissolution only ensures that the numbers of voters moving between districts fulfill the given constraints on the district sizes, that is, the size of each remaining district increases by Δ_s . Figure 1 gives an example illustrating two possible (3,2)-dissolutions for a 5-vertex graph.

Motivated from social choice application scenarios, we additionally assume that each voter supports one of two parties A and B. We then seek a dissolution

such that the number of remaining districts won by party A is maximized. Here, a district is won by the party that is supported by a strict majority of the voters inside the district. This yields the notion of a *biased dissolution*, which is defined as follows:

Definition 2 (Biased dissolution). Let G be an undirected graph and let $\alpha: V(G) \rightarrow \{0, \dots, s\}$ be an *A-supporter distribution*, where $\alpha(v)$ denotes the number of A-supporters in district $v \in V$. Let (D, z) be an (s, Δ_s) -dissolution for G . Let $r_\alpha \in \mathbb{N}$ be the minimum number of districts that party A shall win after the dissolution and $z_\alpha: Z(D, G) \rightarrow \{0, \dots, s\}$ be an *A-supporter movement*, where $z_\alpha(v', v)$ denotes the number of A-supporters moving from district v' to district v . Finally, let $R_\alpha \subseteq V(G) \setminus D$ be a size- r_α subset of districts. Then, $(D, z, z_\alpha, R_\alpha)$ is called an r_α -*biased* (s, Δ_s) -*dissolution* for (G, α) if and only if

- c) a district cannot receive more A-supporters from a dissolved district than the total number of voters it receives from that district:

$$\forall (v', v) \in Z(D, G) : z_\alpha(v', v) \leq z(v', v),$$

- d) no A-supporters remain in any dissolved district:

$$\forall v' \in D : \sum_{(v', v) \in Z(D, G)} z_\alpha(v', v) = \alpha(v'), \text{ and}$$

- e) each district in R_α has a strict majority of A-supporters:

$$\forall v \in R_\alpha : \alpha(v) + \sum_{(v', v) \in Z(D, G)} z_\alpha(v', v) > \frac{s + \Delta_s}{2}.$$

We also say that a district *wins* if it has a strict majority of A-supporters and *loses* otherwise.

Figure 1 shows two biased dissolutions: one with $r_\alpha = 1$ and the other one with $r_\alpha = 2$. We are now ready to formally state the definitions of the two dissolution problems that we discuss in this work:

DISSOLUTION

Input: An undirected graph $G = (V, E)$ and positive integers s and Δ_s .

Question: Is there an (s, Δ_s) -dissolution for G ?

BIASED DISSOLUTION

Input: An undirected graph $G = (V, E)$, positive integers s, Δ_s, r_α , and an A-supporter distribution $\alpha: V \rightarrow \{0, \dots, s\}$.

Question: Is there an r_α -biased (s, Δ_s) -dissolution for (G, α) ?

Note that DISSOLUTION is equivalent to BIASED DISSOLUTION with $r_\alpha = 0$. As we will see later, both DISSOLUTION and BIASED DISSOLUTION are NP-hard in general. In this work, we additionally look into special cases of our dissolution problems and investigate where the causes of intractability lie.

3 Complexity for partially known dissolutions

In this section, we discuss some relevant special cases of our (in general) NP-hard dissolution problems. These include situations where the districts to be dissolved or to win are fixed in advance. We see that BIASED DISSOLUTION is only polynomial-time solvable if both are fixed, and NP-hard otherwise.

Sometimes, the districts to be dissolved and the districts to win are already determined beforehand. For this case, we show that BIASED DISSOLUTION can be modeled as a network flow problem which can be solved in polynomial time.

Theorem 1. *Let $I = (G = (V, E), s, \Delta_s, r_\alpha, \alpha)$ be a BIASED DISSOLUTION instance, and let $D, R_\alpha \subset V$ be two disjoint subsets of districts. The problem of deciding whether (G, α) admits an r_α -biased (s, Δ_s) -dissolution in which D is the set of dissolved districts and in which all districts in R_α are won can be reduced in linear time to a maximum flow problem with $2|V| + 2$ nodes, $2|V| + 3|E|$ arcs, and maximum arc capacity $\max(s, \Delta_s)$.*

With the above flow network construction we can design a polynomial-time algorithm for BIASED DISSOLUTION when the number of districts is a constant.

Corollary 1. *Any instance $((V, E), s, \Delta_s, \alpha)$ of BIASED DISSOLUTION can be solved in time $O(3^{|V|} \cdot (\max(s, \Delta_s) \cdot |V| \cdot |E| + |V|^3))$.*

On the contrary, we obtain NP-hardness for BIASED DISSOLUTION once one of the two sets D and R_α is unknown. For the case that only the set D of dissolved districts is given beforehand, the remaining task is to decide how many A-supporters are moved to a certain non-dissolved district. However, we will see in Section 4.2 that in the hardness construction for Theorem 2 it is already determined which districts are to be dissolved. Furthermore, DISSOLUTION is the special case of BIASED DISSOLUTION with $r_\alpha = 0$ (which implies $R_\alpha = \emptyset$) and DISSOLUTION is NP-hard for the case of $s \neq \Delta_s$ (Theorem 2).

4 Complexity dichotomy with respect to district sizes

In this section, we study the computational complexity of DISSOLUTION and BIASED DISSOLUTION with respect to the ratio of the two integers: old district size s and district size increase Δ_s . We start by showing some useful structural observations for dissolutions in Section 4.1 before we come to the results for DISSOLUTION in Section 4.2 and for BIASED DISSOLUTION in Section 4.3.

4.1 Structural properties

Using the flow construction from Theorem 1, we can show the equivalence of (s, Δ_s) -dissolutions and star partitions for the cases where s is any multiple of Δ_s .

Lemma 1. *There exists a $(t \cdot \Delta_s, \Delta_s)$ -dissolution for an undirected graph G if and only if G has a t -star partition.*

We observe a symmetry concerning the district size s and the district size increase Δ_s in the sense that exchanging their values yields an equivalent instance of DISSOLUTION. Intuitively, the idea behind the following lemma is that the roles of dissolved and non-dissolved districts in a given (s, Δ_s) -dissolution can in fact be exchanged by “reversing” the movement of voters to obtain a (Δ_s, s) -dissolution.

Lemma 2. *There exists an (s, Δ_s) -dissolution for an undirected graph G if and only if there exists a (Δ_s, s) -dissolution for G .*

4.2 Complexity dichotomy for Dissolution

In this subsection, we show a P vs. NP dichotomy of DISSOLUTION with respect to the district size s and the size increase Δ_s . Using Lemma 1, we can show that finding an (s, s) -dissolution essentially corresponds to finding a perfect matching and can thus be done in polynomial time. If $s \neq \Delta_s$, then DISSOLUTION becomes NP-hard. We use Bézout’s identity to encode the NP-complete EXACT COVER BY t -SETS problem into our dissolution problem.

Theorem 2. *If $s = \Delta_s$, DISSOLUTION is solvable in $O(n^\omega)$ time (where ω is the matrix multiplication exponent); otherwise the problem is NP-complete.*

Proof. Let $I = (G, s, \Delta_s)$ be a DISSOLUTION instance with $\Delta_s = s$. Set $t := s/\Delta_s = 1$. Lemma 1 implies that I is a yes-instance if and only if G has a t -star partition. A t -star partition with $t = 1$ is indeed a perfect matching, which can be computed in $O(n^\omega)$ time, where ω is the smallest exponent such that matrix multiplication can be computed in $O(n^\omega)$ time. Currently, the smallest known upper bound of ω is 2.3727 [13].

For the case $s \neq \Delta_s$, we show that DISSOLUTION is NP-complete if $s > \Delta_s$. Due to Lemma 2, this also transfers to the cases where $s < \Delta_s$. First, given a DISSOLUTION instance (G, s, Δ_s) and a function $z : Z(D, G) \rightarrow \{0, \dots, s\}$ where $D \subset V(G)$, one can check in polynomial time whether (D, z) is an (s, Δ_s) -dissolution. Thus, DISSOLUTION is in NP.

To show the NP-hardness result, we give a reduction from the NP-complete EXACT COVER BY t -SETS problem [6] for $t := (s + \Delta_s)/g > 2$, where $g := \gcd(s, \Delta_s) \leq \Delta_s$ is the greatest common divisor of s and Δ_s . Given a finite set X and a collection \mathcal{C} of subsets of X of size t , EXACT COVER BY t -SETS asks whether there is a subcollection $\mathcal{C}' \subseteq \mathcal{C}$ that partitions X , that is, each element of X is contained in exactly one subset in \mathcal{C}' .

Let (X, \mathcal{C}) be an EXACT COVER BY t -SETS instance. We construct a DISSOLUTION instance (G, s, Δ_s) with a neighborhood graph $G = (V, E)$ defined as follows: For each element $u \in X$, add a clique C_u of properly chosen size q to G and let v_u denote an arbitrary fixed vertex in C_u . For each subset $S \in \mathcal{C}$, add a clique C_S of properly chosen size $r \geq t$ to G and connect each v_u for $u \in S$ to a unique vertex in C_S . Figure 2 shows an example of the constructed neighborhood graph for $t = 3$.

Next, we explain how to choose the values of q and r . We set $q := x_q + y_q$, where $x_q \geq 0$ and $y_q \geq 0$ are integers satisfying $x_q s - y_q \Delta_s = g$. Such integers

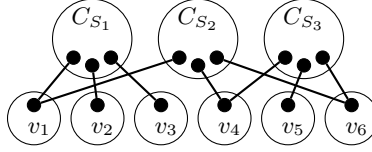


Fig. 2. The constructed instance for $t = 3$.

exist by Bézout's identity. The intuition behind is as follows: Dissolving x_q districts in C_u and moving the voters to y_q districts in C_u creates an overflow of exactly g voters that have to move out of C_u . Notice that the only way to move voters into or out of C_u is via district v_u . Moreover, in any dissolution, exactly x_q districts in C_u are dissolved because dissolving more districts leads to an overflow of at least $g + s + \Delta_s > s$ voters, which is more than v_u can move, whereas dissolving less districts yields a demand of at least $s + \Delta_s - g > \Delta_s$ voters, which is more than v_u can receive. Thus, v_u must be dissolved since there is an overflow of g voters to move out of C_u and this can only be done via district v_u .

The value of $r \geq t$ is chosen in such a way that, for each $S \in \mathcal{C}$ and each $u \in S$, it is possible to move g voters from v_u to C_S (recall that v_u must be dissolved). In other words, we require C_S to be able to receive in total $t \cdot g = s + \Delta_s$ voters in at least t non-dissolved districts. Thus, we set $r := x_r + y_r$, where $x_r \geq 0$ and $y_r \geq t$ are integers satisfying $x_r s - y_r \Delta_s = -(s + \Delta_s)$. Again, since $-(s + \Delta_s)$ is divisible by g , such integers exist by our preliminary discussion. It is thus possible to dissolve x_r districts in C_S moving the voters to the remaining y_r districts in C_S such that we end up with a demand of $s + \Delta_s$ voters in C_S . Note that the only other possibility is to dissolve $x_r + 1$ districts in C_S in order to end up with a demand of zero voters. In this case, no voters of any other districts connected to C_S can move to C_S . By the construction of C_u above, it is clear that it is also not possible to move any voters out of C_S because no v_u can receive voters in any dissolution. Thus, for any dissolution, it holds that either all or none of the v_u connected to some C_S move g voters to C_S .

The proof of correctness is as follows. Suppose (X, \mathcal{C}) is a yes-instance, that is, there exists a partition $\mathcal{C}' \subseteq \mathcal{C}$ of X . We can thus dissolve x_q districts in each C_u (including v_u) and move the voters such that all y_q non-dissolved districts receive exactly Δ_s voters. This is always possible since C_u is a clique. If we do so, then, by construction, g voters have to move out of each v_u . Since \mathcal{C}' partitions X , each $u \in X$ is contained in exactly one subset $S \in \mathcal{C}'$. We can thus move the g voters from each v_u to C_S . Now, for each $S \in \mathcal{C}'$, we dissolve any x_r districts that are not adjacent to any v_u and for the subsets in $\mathcal{C} \setminus \mathcal{C}'$, we simply dissolve $x_r + 1$ arbitrary districts in the corresponding cliques. By the above discussion of the construction, we know that this in fact yields an (s, Δ_s) -dissolution. Hence, (G, s, Δ_s) is a yes-instance.

Now assume that there exists an (s, Δ_s) -dissolution for (G, s, Δ_s) . As we have already seen in the above discussion, any (s, Δ_s) -dissolution generates an overflow of g voters in each C_u that has to be moved over v_u to some district

in C_S . Furthermore, each C_S either receives g voters from all its adjacent v_u or no voters at all. Therefore, the subsets S corresponding to cliques C_S that receive $t \cdot g$ voters form a partition of X , showing that (X, \mathcal{C}) is a yes-instance. \square

4.3 Complexity of Biased Dissolution

Since DISSOLUTION is a special case of BIASED DISSOLUTION, the NP-hardness results for $s \neq \Delta_s$ transfer to BIASED DISSOLUTION. It remains to see whether BIASED DISSOLUTION remains polynomial-time solvable when $s = \Delta_s$. Interestingly, this is true for $s = \Delta_s = 1$.

We introduce a notion called “edge set” for a given dissolution (D, z) of a given graph G . Let $E_z \subseteq E(G)$ contain all edges $\{x, y\}$ with $(x, y) \in Z(D, G)$ and $z(x, y) > 0$. Then, we call E_z the *edge set used by the dissolution* (D, z) .

The following lemma shows that finding an r_α -biased $(1, 1)$ -dissolution essentially corresponds to finding a maximum-weight perfect matching.

Lemma 3. *Let $(G = (V, E), s = 1, \Delta_s = 1, r_\alpha, \alpha)$ be a BIASED DISSOLUTION instance. There exists an r_α -biased $(1, 1)$ -dissolution for (G, α) if and only if there exists a perfect matching of weight at least r_α in (G, w) with $w(\{x, y\}) := 1$ if $\alpha(x) = \alpha(y) = 1$ and $w(\{x, y\}) := 0$ otherwise.*

As shown in the proof of Theorem 2, the edge set used by a $(1, 1)$ -dissolution is a perfect matching. By appropriately setting α and r_α we can enforce that the edge set used by any r_α -biased $(2, 2)$ -dissolution only induces cycles of lengths divisible by four. We end up with a restricted two-factor problem which was already studied in the literature [7] and can be used to show NP-hardness (Theorem 3).

Lemma 4. *Let $G = (V, E)$ be an undirected graph with $4q$ vertices ($q \in \mathbb{N}$). Then G has a two-factor E' whose cycle lengths are all multiples of four, if and only if (G, α) admits a q -biased $(2, 2)$ -dissolution where $\alpha(v) = 1$ for all $v \in V$.*

Theorem 3. BIASED DISSOLUTION on graphs $G = (V, E)$ can be solved in $O(|V|(|E| + |V| \log |V|))$ time if $s = \Delta_s = 1$; otherwise it is NP-complete for any constant value $s = \Delta_s \geq 2$.

5 Complexity on special graph classes

In a companion paper [3], we have shown that computing star partitions—and hence by Lemma 1 also DISSOLUTION—remains NP-hard even on subcubic grid graphs and split graphs. In this section, we discuss the complexity of BIASED DISSOLUTION on special graph classes.

An interesting special case of BIASED DISSOLUTION occurs if voters can move from any district to any other district, that is, the neighborhood graph is a clique. Then, the existence of an (s, Δ_s) -dissolution depends only on the number $|V|$ of districts, the district size s , and the size increase Δ_s . Clearly, a DISSOLUTION instance is a yes-instance if and only if $d := |V| \cdot \Delta_s / (s + \Delta_s)$ is an

integer. We show that BIASED DISSOLUTION can likewise be solved in polynomial time if the neighborhood graph is a clique. Yuster [14, Theorem 2.3] showed that the H -FACTOR problem is solvable in linear time on graphs of bounded treewidth when the size of H is constant. This includes the case of finding x -star partitions, that is, $(x, 1)$ -dissolutions resp. $(1, x)$ -dissolutions when x is constant. We can show that the more general problem BIASED DISSOLUTION is solvable in linear time on graphs of bounded treewidth when s and Δ_s are constants. By a polynomial-time reduction from the NP-hard PERFECT PLANAR H -MATCHING problem [2], we get NP-hardness for DISSOLUTION on planar graphs.

Theorem 4.

- (1) BIASED DISSOLUTION is solvable in $O(|V|^2)$ time on cliques.
- (2) BIASED DISSOLUTION is solvable in linear time on graphs of bounded treewidth when s and Δ_s are constant.
- (3) DISSOLUTION on planar graphs is NP-complete for all $s \neq \Delta_s$ such that Δ_s divides s or s divides Δ_s . It is polynomial-time solvable for $s = \Delta_s$.

Proof (Sketch for (1)). In fact, we show how to solve the optimization version of BIASED DISSOLUTION, where we maximize the number r_α of winning districts. Intuitively, it appears to be a reasonable approach to dissolve districts pursuing the following two objectives: Any losing district should contain as few A-supporters as possible and any winning district should contain exactly the amount that is required to have a majority. Dissolving districts this way minimizes the number of “wasted” A-supporters. We now show that this greedy strategy is indeed optimal.

Let $G = (V, \binom{V}{2})$ be a clique, let α be an A-supporter distribution over V , and let s and Δ_s be the district size and the district size increase. With G being complete, we are free to move voters from any dissolved district to any non-dissolved district. Let $\mu := \lfloor (s + \Delta_s)/2 \rfloor + 1$ be the minimum number of A-supporters required to win a district. Thus, a district with less than $(\mu - \Delta_s)$ A-supporters can never win. Denote by $\mathcal{L} := \{v \in V \mid \alpha(v) < \mu - \Delta_s\}$ the set of *non-winnable* districts.

Our first claim corresponds to the first objective above, that is, the losing districts should contain a minimal number of A-supporters.

Claim 1 *Let $v, w \in V$ be two districts with $\alpha(v) \leq \alpha(w)$. If there exists an r_α -biased dissolution where v is winning and w is losing, then there also exists an r_α -biased dissolution where v is losing and w is winning.*

The next claim basically corresponds to the second objective, in the sense that districts with a large number of A-supporters (possibly more than the required) should be dissolved in order to distribute the voters more efficiently.

Claim 2 *Let $v, w \in V$ be two districts with $\alpha(v) \leq \alpha(w)$. Assume that there exists an r_α -biased dissolution where r_α is optimal. If v is dissolved, then the following holds: (i) If w is losing, then there also exists an r_α -biased dissolution where w is dissolved and v is losing. (ii) If w is winning and v is winnable, that is, $v \notin \mathcal{L}$, then there exists an r_α -biased dissolution where w is dissolved and v is winning.*

Using the two claims above, we now show how to compute an optimal biased dissolution. In order to find a biased dissolution with the maximum number of winning districts, we seek a dissolution which loses a minimum number of remaining districts. Thus, for each $\ell \in \{0, \dots, r\}$, we check whether it is possible to dissolve d districts such that at most ℓ of the remaining r districts lose. To this end, assume that the districts v_1, \dots, v_n are ordered by increasing number of A-supporters, that is, $\alpha(v_1) \leq \alpha(v_2) \leq \dots \leq \alpha(v_n)$ and let $V_\ell := \{v_1, \dots, v_\ell\}$. Now, if there exists an $(r - \ell)$ -biased dissolution, then there also exists an $(r - \ell)$ -biased dissolution where the losing districts are exactly V_ℓ . This follows by repeated application of the exchange arguments of Claim 1 and Claim 2(i). Hence, given ℓ , we have to check whether there is a set $D \subseteq V \setminus V_\ell$ of d districts that can be dissolved in such a way that all non-dissolved districts in $V \setminus (V_\ell \cup D)$ win and the districts in V_ℓ lose.

First, note that in order to achieve this, all districts in $\mathcal{L} \setminus V_\ell$ have to be dissolved because they cannot win in any way. Clearly, if $|\mathcal{L} \setminus V_\ell| > d$, then it is simply not possible to lose only ℓ districts and we can immediately go to the next iteration with $\ell := \ell + 1$. Therefore, we assume that $|\mathcal{L} \setminus V_\ell| \leq d$ and let $d' := d - |\mathcal{L} \setminus V_\ell|$ be the number of additional districts to dissolve in $V \setminus (\mathcal{L} \cup V_\ell)$. By Claim 2(ii), it follows that we can assume that the d' districts with the maximum number of A-supporters are dissolved, that is, $V^{d'} := \{v_{n-d'+1}, \dots, v_n\}$. Thus, we set $D := \mathcal{L} \setminus V_\ell \cup V^{d'}$ and check whether there are enough A-supporters in D to let all $r - \ell$ remaining districts in $V \setminus (V_\ell \cup D)$ win.

Sorting the districts by the number of A-supporters (as preprocessing) requires $O(n \log n)$ arithmetic operations. For up to n values of ℓ , to check whether the remaining districts in $V \setminus (V_\ell \cup D)$ can win requires $O(n)$ arithmetic operations each. Thus, assuming constant-time arithmetics, we end up with $O(n^2)$ time. \square

6 Conclusion

We initiated a graph-theoretic combinatorial approach to concrete redistribution problems occurring in various application domains. Obviously, the two basic problems DISSOLUTION and BIASED DISSOLUTION concern highly simplified situations and will not be able to model all interesting aspects of redistribution scenarios. For instance, our constraint that before and after the dissolution all vertex loads are perfectly balanced may be too restrictive for many applications. All in all, we consider our simple (and yet fairly realistic) models as a first step into a fruitful research direction that might yield a stronger linking of graph-theoretic concepts with districting methods and other application scenarios.

We end with a few specific challenges for future research. We have left open whether the P vs. NP dichotomy for general graphs fully carries over to the planar case: it might be possible that planar graphs allow for some further tractable cases with respect to the relation between old and new district sizes. Moreover, with redistricting applications in mind it might be of interest to study special cases of planar graphs (such as grid-like structures) in quest of finding polynomial-time solvable special cases of network-based dissolution problems. Having identified

several NP-hard special cases of DISSOLUTION and BIASED DISSOLUTION, it is a natural endeavor to investigate their polynomial-time approximability and their parameterized complexity; in the latter case one also needs to identify fruitful parameterizations.

Acknowledgments. René van Bevern was supported by the DFG, project DAPA (NI 369/12), Robert Bredereck by the DFG, project PAWS (NI 369/10), Jiehua Chen by the Studienstiftung des Deutschen Volkes, Vincent Froese by the DFG, project DAMM (NI 369/13), and Gerhard J. Woeginger while visiting TU Berlin by the Alexander von Humboldt Foundation, Bonn, Germany.

Bibliography

- [1] M. Altman. *Districting Principles and Democratic Representation*. PhD thesis, California Institute of Technology, 1998.
- [2] F. Berman, D. Johnson, T. Leighton, P. W. Shor, and L. Snyder. Generalized planar matching. *Journal of Algorithms*, 11(2):153–184, 1990.
- [3] R. van Bevern, R. Bredereck, L. Bulteau, J. Chen, V. Froese, R. Niedermeier, and G. J. Woeginger. Star partitions of perfect graphs. In *Proc. of 41st ICALP*, volume 8572 of *LNCS*, pages 174–186, 2014. Full version available as arXiv:1402.2589 [cs.DM].
- [4] J. C. Duque. *Design of Homogeneous Territorial Units: A Methodological Proposal and Applications*. PhD thesis, University of Barcelona, 2004.
- [5] J. C. Duque, R. Ramos, and J. Surinach. Supervised regionalization methods: A survey. *International Regional Science Review*, 30(3):195–220, 2007.
- [6] M. R. Garey and D. S. Johnson. *Computers and Intractability*. W. H. Freeman, 1979.
- [7] P. Hell, D. G. Kirkpatrick, J. Kratochvíl, and I. Kríz. On restricted two-factors. *SIAM Journal on Discrete Mathematics*, 1(4):472–484, 1988.
- [8] Z. Landau and F. Su. Fair division and redistricting. *Social Choice and Welfare*, 32(3):479–492, 2009.
- [9] M. Maravalle and B. Simeone. A spanning tree heuristic for regional clustering. *Communications in Statistics—Theory and Methods*, 24(3):625–639, 1995.
- [10] A. Mehrota, E. L. Johnson, and G. L. Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, 1998.
- [11] C. Puppe and A. Tasnádi. A computational approach to unbiased districting. *Mathematical and Computer Modelling*, 48:1455–1460, 2008.
- [12] C. Puppe and A. Tasnádi. Optimal redistricting under geographical constraints: Why “pack and crack” does not work. *Economics Letters*, 105(1):93–96, 2009.
- [13] V. Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. In *Proc. 44th STOC*, pages 887–898. ACM, 2012.
- [14] R. Yuster. Combinatorial and computational aspects of graph packing and graph decomposition. *Computer Science Review*, 1(1):12–26, 2007.