A Refined Complexity Analysis of Fair Districting over Graphs

Extended Abstract

Niclas Boehmer
TU Berlin
Berlin, Germany
niclas.boehmer@tu-berlin.de

Tomohiro Koana
TU Berlin
Berlin, Germany
tomohiro.koana@tu-berlin.de

Rolf Niedermeier
TU Berlin
Berlin, Germany
rolf.niedermeier@tu-berlin.de

ABSTRACT

We study the NP-hard Fair Connected Districting problem: Partition a vertex-colored graph into k connected components (subsequently referred to as districts) so that in every district the most frequent color occurs at most a given number of times more often than the second most frequent color. Fair Connected Districting is motivated by various real-world scenarios, such as district-based elections, where agents of different types, which are one-to-one represented by nodes in a network, have to be partitioned into disjoint districts. We conduct a fine-grained analysis of the (parameterized) computational complexity of Fair Connected Districting: We study its parameterized complexity with respect to various graph parameters, including treewidth, and problem-specific parameters, including the numbers of colors and districts, and its complexity on graphs from different classes (such as paths, stars, and trees).

KEYWORDS

Parameterized Algorithmics; Electoral Districting; Vertex-Colored Graphs

ACM Reference Format:
Niclas Boehmer, Tomohiro Koana, and Rolf Niedermeier. 2022. A Refined Complexity Analysis of Fair Districting over Graphs: Extended Abstract. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

1 INTRODUCTION

Stoica et al. [18] recently introduced graph-based problems on fair (re)districting, employing "margin of victory" as the measure of fair representation. They performed theoretical and empirical studies; the latter clearly supporting the practical relevance of these problems. In our paper, we focus on the theoretical aspects, significantly extending their findings in this direction.

Dividing agents into groups is a ubiquitous task. Electoral districting is one of the prime examples: All voters are assigned to political districts, in which their own representatives are chosen. Another example emerges in education; in many countries, children are assigned to schools based on their residency. In such scenarios, the agents (in the settings above, voters or school children) are often placed on a (social or geographical) network.

In districting, there are various objectives. What we study here can be interpreted as a "benevolent" counterpart of gerrymandering, which is well-studied in voting theory. For gerrymandering, every voter is characterized by their projected vote in the upcoming election. The goal is then to find a partition of the voters into connected districts such that some designated alternative gains the majority in as many districts as possible. Following Stoica et al. [18], we consider an opposite objective. That is, we assume that some central authority wishes to partition the agents, which are of different types, into connected districts that are *fair*, where a district is deemed fair if the margin of victory in the district is smaller than a given bound. The *margin of victory* of a district is the minimum number of agents whose deletion results in a tie between the two most frequent types in the district: In electoral districting where agents' types represents their projected vote, a low margin of victory may foster competition among politicians, thereby motivating elected officials to do a great job. When partitioning children into school districts, types may model sociodemographic attributes such as race and gender, and a low margin of victory could be beneficial to prevent the existence of schools where one trait is in a clear majority and which may thus be only associated with this single trait (see Stoica et al. [18] for a more extensive discussion).

In our work, we search for tractable special cases of fair districting over graphs focusing on Fair Connected Districting:

Fair Connected Districting (FCD)
Input: An undirected graph $G = (V, E)$, a set C of colors, a function $\mathrm{col} : V \to C$ assigning each vertex one color from C, a number $k \leq |V|$ of districts, a maximum margin of victory ℓ, and two integers $s_{\max} \geq s_{\min} \geq 1$.
Question: Does there exist a partition of the vertices into k districts $(V_1, \ldots, V_k)$ such that, for all $i \in [k]$, V_i is ℓ-fair, $|V_i| \in [s_{\min}, s_{\max}]$, and the graph G induced by the vertices from V_i is connected?

A set of vertices V' is *ℓ-fair* if the difference between the occurrences of the most and second-most frequent color in V' is at most ℓ.

2 CONTRIBUTION

It is easy to see that FCD generalizes the NP-hard Perfectly Balanced Connected Partition problem [4, 6], which asks for a partition of an undirected graph into two connected components of the same size. This motivates a parameterized complexity analysis and the study of restrictions of the underlying graph in order to identify tractable special cases. We investigate the influence of problem-specific parameters (the number $|C|$ of colors, the number k of districts, and the margin of victory ℓ) and the structure of the underlying graph on the computational complexity of FCD.

We show that FCD is NP-hard even if $|C| = k = 2$ and $\ell = 0$ but polynomial-time solvable on paths, cycles, stars, and caterpillars (for stars, our algorithm even runs in linear time).[1] Subsequently,

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

[1] While in real-world applications these simple graphs may occur not often, they are the building blocks of more complex graphs. This also motivated a study of these graph classes for gerrymandering over graphs [2, 9, 11]. Moreover, initially, it is not at all clear that FCD is polynomial-time solvable on these graphs, as, for instance, gerrymandering

we extend our polynomial-time algorithms for paths and cycles to a polynomial-time algorithm for all graphs with a constant max leaf number (mln), which are basically graphs that consist of a constant number of paths and cycles (where the two endpoints of each path and one point from each cycle can be arbitrarily connected).

Remarkably, in our most involved hardness reduction, we show that FCD already becomes NP-hard and even W[1]-hard with respect to $|C|+k$ on trees. However, when the number of colors or the number of districts is constant, FCD on trees becomes polynomial-time solvable. In fact, we show that these results hold for some tree-like graphs as well. Herein, the tree-likeness of a graph is measured by one of three parameters, namely, the treewidth (tw), the feedback edge number (fen), and the feedback vertex number (fvn). More precisely, as our most involved algorithmic results, we establish polynomial-time solvability of FCD when the number of colors and the treewidth are constant. We achieve this with a dynamic programming approach on the tree decomposition of the given graph empowered by some structural observations on FCD. Moreover, we observe that there is a simple polynomial-time algorithm on graphs with a constant feedback edge number when there are a constant number of districts. On the other hand, we prove that FCD is NP-hard for two districts even on graphs with fvn = 1 (and tw = 2). Lastly, we show that FCD is polynomial-time solvable on graphs with a constant vertex cover number (vcn) and fixed-parameter tractable with respect to the vertex cover number and the number of colors. A summary of our parameterized results can be found in Figure 1. Notably, all our hardness results also hold without size constraints.

In our studies, we identify several sharp complexity dichotomies. For instance, FCD is polynomial-time solvable on trees with diameter at most three but NP-hard and W[1]-hard with respect to $|C|+k$ on trees with diameter four. Similarly, FCD is NP-hard and W[1]-hard with respect to $|C|+k$ on graphs with pathwidth at least two but polynomial-time solvable on pathwidth-one graphs.

To summarize, we show that FCD without size constraints is NP-hard even in very restricted settings, e.g., on trees or if $|C| = k = 2$ and $\ell = 0$. To make the problem tractable, one possibility is to significantly restrict the input graph, e.g., to consist of a constant number of paths and cycles, or to combine structural parameters of the given graph with the number $|C|$ of colors or the number k of districts [2]. For small $|C|$ and k, the tractability of FCD extends to certain tree-like graphs and graphs with a small vertex cover number. In contrast to the parameters $|C|$ and k, which have a strong influence on the complexity of FCD, the bound ℓ on the margin of victory has only little impact as all hardness results already hold for $\ell = 0$ and all our algorithmic results hold for arbitrary ℓ. The full version of our paper is available on arXiv [3].

3 RELATED WORK

Stoica et al. [18] introduced Fair Connected Regrouping, which is a generalization of our FCD problem. Fair Connected Regrouping differs from FCD in that, in Fair Connected Regrouping, one is additionally given a function that specifies for each vertex to which district it can belong. They proved that Fair Connected Regrouping is NP-hard even for only two colors and two districts. Moreover, Stoica et al. [18] considered special cases of Fair Connected Regrouping: Fair Regrouping (omitting connectivity constraints) and Fair Regrouping_X (further omitting any restriction to which districts vertices can belong). They proved that Fair Regrouping is NP-hard for three colors but in XP with respect to the number of districts, and that Fair Regrouping_X is in XP with respect to the number of colors.

FCD is relevant in district-based elections. Several papers have studied how to assign voters to constituencies so as to "fairly" reflect the political choices of voters [1, 12, 13, 16, 17]. Well-studied in this context is gerrymandering, which can be regarded as a "malicious" counterpart to our problem. In gerrymandering, the task is to partition a set of voters into districts obeying certain conditions such that a designated alternative wins in as many districts as possible. Initially, gerrymandering has been predominantly studied from the perspective of social and political science [8, 10, 15] but more recently also different variants of gerrymandering (over graphs) have been considered from an algorithmic perspective [2, 5, 7, 9, 11, 14]. In particular, as done here for FCD, Bentert et al. [2], Gupta et al. [9], and Ito et al. [11] analyzed the complexity of gerrymandering on paths, cycles, and trees and studied the influence of the number of candidates/colors and the number of districts.

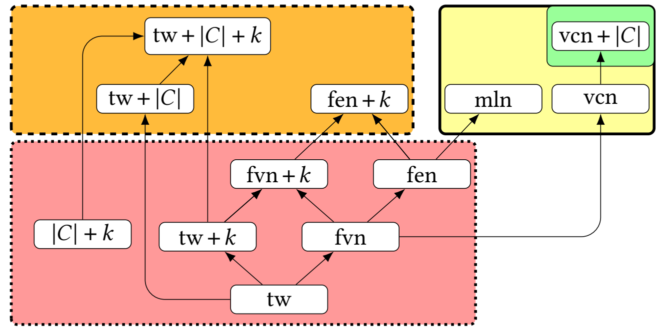

Figure 1: Overview of our parameterized complexity results. Each box represents one parameterization of FCD. An arc from parameter p to another parameter p' indicates that p is upper-bounded by some function of p'. For parameters in the red area (dotted), we prove that FCD is NP-hard even if the parameter is a constant. For parameters in the orange area (dashed), we prove W[1]-hardness and present an XP-algorithm. For parameters in the yellow area (solid thick), we have an XP-algorithm but W[1]-hardness is unknown. The green area (solid) indicates fixed-parameter tractability.

ACKNOWLEDGMENTS

NB was supported by the DFG projects MaMu (NI 369/19) and ComSoc-MPMS (NI 369/22). TK was supported by the DFG projects FPTinP (NI 369/18) and DiPa (NI 369/21). This work was started at the research retreat of the TU Berlin Algorithmics and Computational Complexity research group held in Zinnowitz (Usedom) in September 2020.

over graphs is NP-hard even on paths [2, Theorem 1]. This also proves that FCD is sometimes easier than the corresponding gerrymandering over graphs problem.

[2]Notably, in most applications, the number of colors and districts should be relatively small. E.g., in their experiments, Stoica et al. [18] partitioned 50,000 voters into 10 voting districts and 41,834 schoolchildren into 61 school districts with $|C| = 7$.

REFERENCES

[1] Yoram Bachrach, Omer Lev, Yoad Lewenberg, and Yair Zick. 2016. Misrepresentation in District Voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI '16)*. AAAI Press, 81–87.

[2] Matthias Bentert, Tomohiro Koana, and Rolf Niedermeier. 2021. The Complexity of Gerrymandering over Graphs: Paths and Trees. In *Proceedings of the 47th International Workshop on Graph-Theoretic Concepts in Computer Science (WG '21)*. Springer, 195–206.

[3] Niclas Boehmer, Tomohiro Koana, and Rolf Niedermeier. 2021. A Refined Complexity Analysis of Fair Districting over Graphs. *CoRR* abs/2102.11864 (2021). arXiv:2102.11864 https://arxiv.org/abs/2102.11864

[4] Janka Chlebíková. 1996. Approximating the Maximally Balanced Connected Partition Problem in Graphs. *Inf. Process. Lett.* 60, 5 (1996), 223–230.

[5] Amittai Cohen-Zemach, Yoad Lewenberg, and Jeffrey S. Rosenschein. 2018. Gerrymandering Over Graphs. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '18)*. IFAAMAS, 274–282.

[6] Martin E. Dyer and Alan M. Frieze. 1985. On the complexity of partitioning graphs into connected subgraphs. *Discrete Appl. Math.* 10, 2 (1985), 139–153.

[7] Eduard Eiben, Fedor V. Fomin, Fahad Panolan, and Kirill Simonov. 2020. Manipulating Districts to Win Elections: Fine-Grained Complexity. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI '20)*. AAAI Press, 1902–1909.

[8] Robert S Erikson. 1972. Malapportionment, gerrymandering, and party fortunes in congressional elections. *Am. Political Sci. Rev.* 66, 4 (1972), 1234–1245.

[9] Sushmita Gupta, Pallavi Jain, Fahad Panolan, Sanjukta Roy, and Saket Saurabh. 2021. Gerrymandering on Graphs: Computational Complexity and Parameterized Algorithms. In *Proceedings of the 14th International Symposium on Algorithmic Game Theory (SAGT '21)*. Springer, 140–155.

[10] Samuel Issacharoff. 2002. Gerrymandering and political cartels. *Harv. Law Rev.* 116 (2002), 593–648.

[11] Takehiro Ito, Naoyuki Kamiyama, Yusuke Kobayashi, and Yoshio Okamoto. 2021. Algorithms for gerrymandering over graphs. *Theor. Comput. Sci.* 868 (2021), 30–45.

[12] Zeph Landau, Oneil Reid, and Ilona Yershov. 2009. A fair division solution to the problem of redistricting. *Soc. Choice Welf.* 32, 3 (2009), 479–492.

[13] Zeph Landau and Francis Edward Su. 2013. Fair Division and Redistricting. *AMS Special Sessions on The Mathematics of Decisions, Elections, and Games* (2013), 17–36.

[14] Yoad Lewenberg, Omer Lev, and Jeffrey S. Rosenschein. 2017. Divide and Conquer: Using Geographic Manipulation to Win District-Based Elections. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems (AAMAS '17)*. ACM, 624–632.

[15] David Lublin. 1999. *The Paradox of Representation: Racial Gerrymandering and Minority Interests in Congress.* Princeton University Press.

[16] Adway Mitra. 2020. Electoral David vs Goliath: How does the Spatial Concentration of Electors affect District-based Elections? *CoRR* abs/2006.11865 (2020). arXiv:2006.11865 https://arxiv.org/abs/2006.11865

[17] Clemens Puppe and Attila Tasnádi. 2008. A computational approach to unbiased districting. *Math. Comput. Model.* 48, 9-10 (2008), 1455–1460.

[18] Ana-Andreea Stoica, Abhijnan Chakraborty, Palash Dey, and Krishna P. Gummadi. 2020. Minimizing Margin of Victory for Fair Political and Educational Districting. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '20)*. IFAAMAS, 1305–1313.