Gerrymandering under Uncertain Preferences (Student Abstract)

Benjamin Kelly

Rensselaer Polytechnic Institute
110 8th Street
Troy, New York 12180
benjaminkelly@me.com

Abstract

Gerrymandering is the manipulating of redistricting for political gain. While many attempts to formalize and model gerrymandering have been made, the assumption of known voter preference, or *perfect information*, limits the applicability of these works to model real world scenarios. To more accurately reason about gerrymandering we investigate how to adapt existing models of the problem to work with *imperfect information*. In our work, we formalize a definition of the gerrymandering problem under probabilistic voter preferences, reason about its complexity compared to the deterministic version, and propose a greedy algorithm to approximate the problem in polynomial time under certain conditions.

Introduction

Gerrymandering refers to manipulating district assignments such that one candidate or political party gains some kind of advantage. This advantage manifests as a substantial difference in the proportional of seats won by a candidate and the proportion of votes won. Gerrymandering is a commonly used tactic in many countries around the world including the United States where a recent example includes North Carolina, where a state court ruled the district map unconstitutional under the North Carolina State constitution in 2019 (Wines 2019). Some court cases have restricted gerrymandering, but it still remains widely used and so developing models with which we are able to reason about the effect and potential of gerrymandering remains critically important. While existing work has been done into formalizing a definition of gerrymandering, much of the work makes the assumption of constant and known voter preferences. This assumption allows models to be simpler and reflects how many people will vote, but we believe it also limits the applicability of these existing models to real world scenarios.

In particular, we use a definition of gerrymandering as a decision problem from (Cohen-Zemach, Lewenberg, and Rosenschein 2018), A_{GM}, and extend it to include voters with non-constant preferences. A_{GM}, shown to be NP-Complete, was originally defined to be whether a graph of voters could be partitioned to create connected components (districts) such that a given candidate would win a specific number of districts in the election.

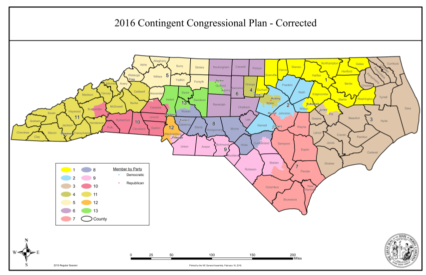

Figure 1: North Carolina district map ruled unconstitutional in 2019. Sourced from the North Carolina General Assembly (Joint Select Committee on Congressional Redistricting 2016).

Problem Definition

We define *PROB-GERRY* as follows: given an undirected graph of voters, each with a probability distribution over the linear orders of candidates, can we partition the graph to create a set of connected components corresponding to a valid district assignment? A valid district assignment must ensure that a target candidate has at least a given probability to win a specified number of districts, as well at most another probability to lose a specified number of districts under a given voting rule. Lastly, there is a restriction placed on the ratio of the size from the largest district to the smallest district.

We include two bounds, a lower bound on number of districts won, and an upper bound on number of districts lost. Adding both of these conditions ensures that the candidate has sufficient odds to both win by a given margin, but also never to lose by a given margin as well. A potentially gerrymandered map and district assignment is given for both the deterministic and non-deterministic version of the problem are given in 2. Other assignments of the probabilistic map give the same chance for green to win 3+ districts, but also increase the chance green loses 4+ districts.

Problem Complexity

Theorem 1 *PROB-GERRY is NP-Hard*

To prove this, we show that there exists a reduction from A_{GM} to *PROB-GERRY* by creating an instance of *PROB-GERRY* where voters are given constant preferences. Since

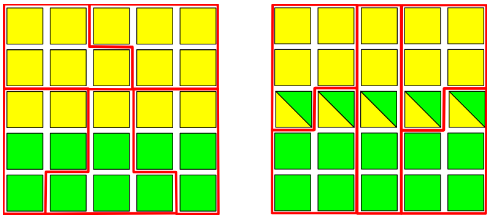

Figure 2: An example of a potential assignment for a deterministic map (left) and probabilistic map (right)

A_{GM} is NP-Hard, *PROB-GERRY* must also be NP-Hard.

Theorem 2 *PROB-GERRY is NP-Complete for constant candidate number and weights bounded by* $poly(n)$.

Using the result of the previous theorem, we know that *PROB-GERRY* is NP-Hard. To show that *PROB-GERRY* is NP-Complete under these conditions, we provide a polynomial time verifier for *PROB-GERRY* using the algorithm described in (Hazon et al. 2012) for calculating winning probabilities for candidates with probabilistic voters. This verifier shows that *PROB-GERRY* is in NP, and we know that it is NP-Hard, so under these conditions it is NP-Complete.

Greedy Algorithm

Additionally, since even the version of *PROB-GERRY* with bounded candidate number and voter weights is NP-Complete, we offer a greedy algorithm for approximately solutions to *PROB-GERRY* in polynomial time under bounded candidate number and voter weights. The algorithm is derived as an extension of the algorithm presented in (Cohen-Zemach, Lewenberg, and Rosenschein 2018) using the dynamic programming approach described in (Hazon et al. 2012) to calculate probabilities of candidates winning the election. An abbreviated description is provided in Algorithm 1. Let k be the number of total districts, l be the number of target victories, and r be the maximum ratio cap. RatioCap(G) denotes the ratio of the size of the largest district to and smallest district in G.

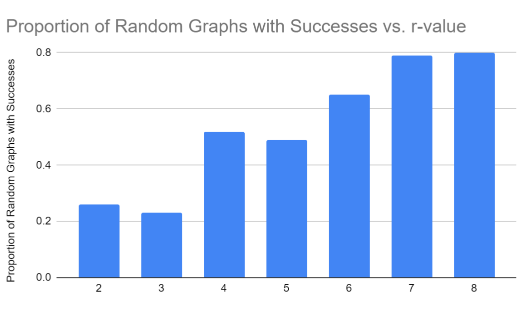

Figure 3: Ratio Cap vs. Success of Algorithm on Random Graphs

Algorithm 1: Greedy Algorithm for *PROB-GERRY*

Initialize a new voter graph G' with no edges
while *Number of districts is more than desired amount* **do**
  $e = \text{argmax}_{e \in E(G) \setminus E(G')}$ P(candidate wins $\frac{l}{k}$ proportion of current districts) given that RatioCap($G' \cup \{e\}$) $> r$
  if $e = None$ **then**
    $e = \text{argmin}_{e \in E(G) \setminus E(G')}$ RatioCap($G' \cup \{e\}$)
  $E(G') = E(G') \cup \{e\}$

Testing and Results

We tested our greedy algorithm by generating graphs of voters. Each voter was assigned a position in $[0, 1]^2$ from a normal distribution, and two candidates were created at (0.75, 0.75) and (0.25, 0.25). We generated graphs using the procedure described in (Cohen-Zemach, Lewenberg, and Rosenschein 2018). We then specified our problem to draw 5 districts and see if there is a way that candidate 1 can win 3+ districts with 70% probability and lose 3+ districts with 25% probability. We used the Plackett-Luce model to assign voters a probability to each linear order of candidates using the distance from each candidate to a voter. We analyzed the success rate of finding an assignment on 1000 random graphs of 100 voters while changing the ratio cap (figure 3). We conclude that the ratio cap strongly determined the chance of success and that improvements in how to handle this cap is important in developing a more applicable algorithm. We also analyzed running time based on number of voters and candidates to confirm that our algorithm scaled in polynomial and exponential time respectively. These results indicated that even with ideal conditions, large numbers of voters or even just a few candidates causes even this greedy algorithm to take large amounts of time.

Acknowledgements

This work began as a project in a class taught by Lirong Xia. We thank all anonymous reviewers, all students attending the class, Inwon Kang, Farhad Mohsin, and Lirong Xia for helpful suggestions and feedback.

References

Cohen-Zemach, A.; Lewenberg, Y.; and Rosenschein, J. S. 2018. Gerrymandering over graphs. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, 274–282.

Hazon, N.; Aumann, Y.; Kraus, S.; and Wooldridge, M. 2012. On the evaluation of election outcomes under uncertainty. *Artificial Intelligence*, 189: 1–18.

Joint Select Committee on Congressional Redistricting. 2016. *2016 Contingent Congressional Plan*. North Carolina General Assembly.

Wines, M. 2019. State Court Bars Using North Carolina House Map in 2020 Elections. *The New York Times*.