Pareto Frontier - What Are the Compromises in Redistricting?

Sonali Durham, Parker Rule, Si Wu, and Hamidreza Validi

August 2019

1 Introduction

At the Voting Rights Data Institute, we look at a range of criteria about districting plans - partisan metrics (efficiency gap, mean-median score), demographics (population balance, black voting age population), geographic information (Polsby-Popper score), and graph-based statistics (number of cut edges). In this project, we apply Pareto optimization to attempt to optimize two or more of these factors simultaneously. The trade-offs between two or more of these factors therefore become important.

Here's an example of Pareto optimization. For a screwdriver, we want to optimize its beauty and strength. Therefore the red points are on Pareto front, and the grey points are Pareto-dominated. There's never a reason to pick screwdriver N or K - there are screwdrivers that are both stronger and more beautiful than them.

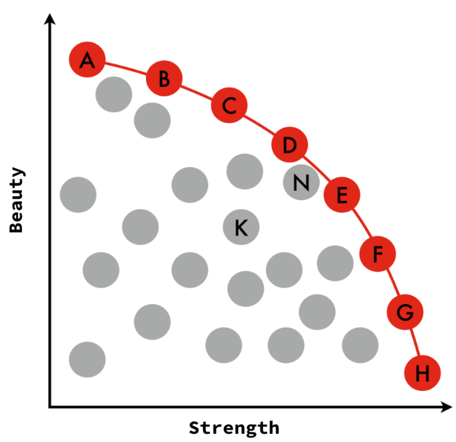

Figure 1: Example of Pareto optimization - maximizing the beauty and strength of a screwdriver

2 Methodology and Data Analysis

The analysis is done on Iowa, Virginia and Pennsylvania shapefiles. The shapefiles were obtained from the U.S. Census Bureau and processed by members of the Metric Geometry and Gerrymandering Group (MGGG). The data can be found at https://github.com/mggg-states.

2.1 Ensemble Generation

Ensembles of districting plans are generated by the GerryChain software.

2.2 Trade-offs between criteria on the state of Iowa

The figure below shows the trade-offs between districting criteria (absolute population deviation and number of cut edges), using a dataset of 5,000 county-level Iowa districting plans collected with GerryChain's random ReCom algorithm. An algorithm was written to compute the Pareto front (colored in blue in figure 2). Unlike the screwdriver example, we want to minimize the percentages of cut edges and absolute population deviation.

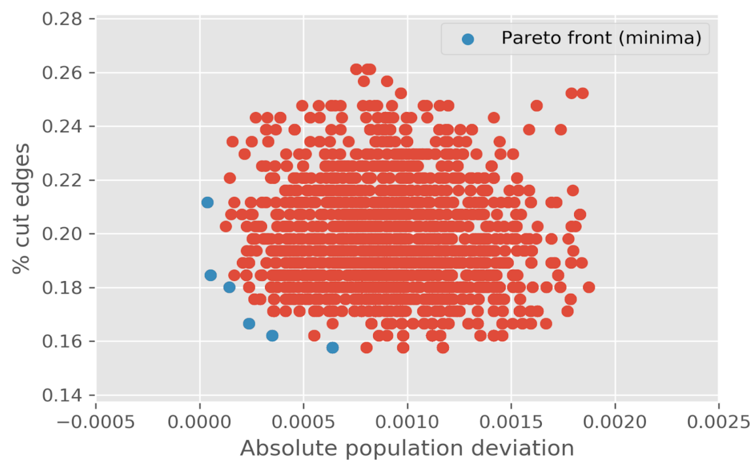

Figure 2: Percentage of cut edges versus absolute population deviation in Iowa

2.3 Trade-offs between BVAPs of Virginia

Figure 3 below shows the percentage of Black Voting Age Population (BVAP) of 100 House Districts of Virginia, using a dataset of 10,000 county-level Iowa districting plans collected using GerryChain's random ReCom algorithm. Based on the figure, District 87 and District 88 are opportunity districts, because the BVAPs of those two districts are located around the lower bound of the opportunity range (indicated by horizontal red lines). The trade-off between the BVAPs of these two districts therefore become important.

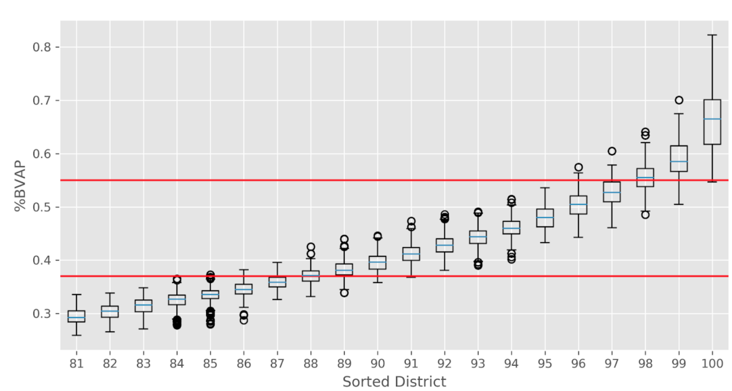

Figure 3: Virginia House Opportunity Districts by percentages of BVAPs

The figure below shows the trade-offs between BVAPs of District 87 and District 88.

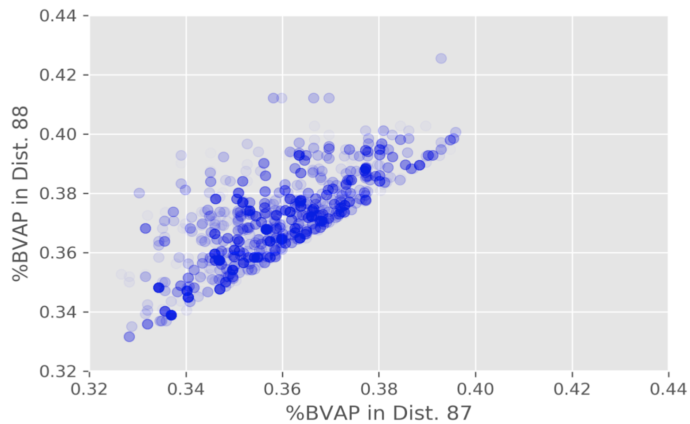

Figure 4: Trade-off between BVAPs of District 87 and District 88

2.4 Biased Recom

We want to study trade-offs that come from moving a district into or out of the BVAP opportunity range. However, we only expect a few districts to ever fall in the opportunity range, so the probability that ReCom will perturb an opportunity district in any one step is small.

This is why we came up with the biased Recom algorithm. Biased recom works in five steps: 1) choose a district-level metric (e.g. percentage of Black Population (BPOP)); 2) choose a set of target districts based on that metric (e.g. the districts ranked 86th and 87th by percentage of BPOP); 3) rank the districts of the last plan by the metric at each step of the chain; 4) choose a target district and one of its neighbors; otherwise, it chooses any pair of districts; 5) apply ReCom as usual.

Figure 5 below shows the percentage of Black Population of 18 Congressional Districts of Pennsylvania, using a dataset of 10,000 precinct-level Pennsylvania districting plans collected using GerryChain's random ReCom algorithm. Based on the figure, District 16 and 17 are opportunity districts, because the BPOPs of these districts are located around the oppotunity range (indicated by horizonal red lines).

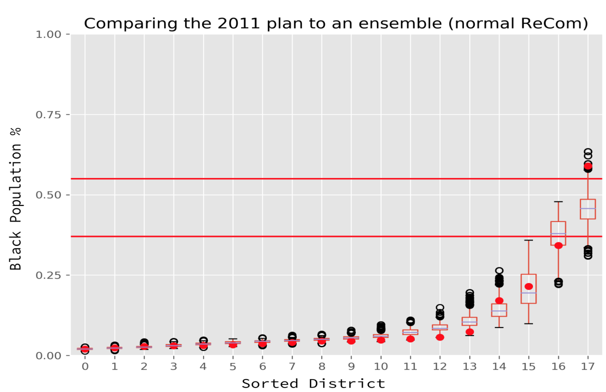

Figure 5: Pennsylvania Congressional Opportunity Districts by percentages of BPOPs

Figure 6 below shows the distribution of Black Population (BPOP) of District 16 (opportunity district for BPOP) of Pennsylvania, for 10,000 plans generated from normal and biased Recom.

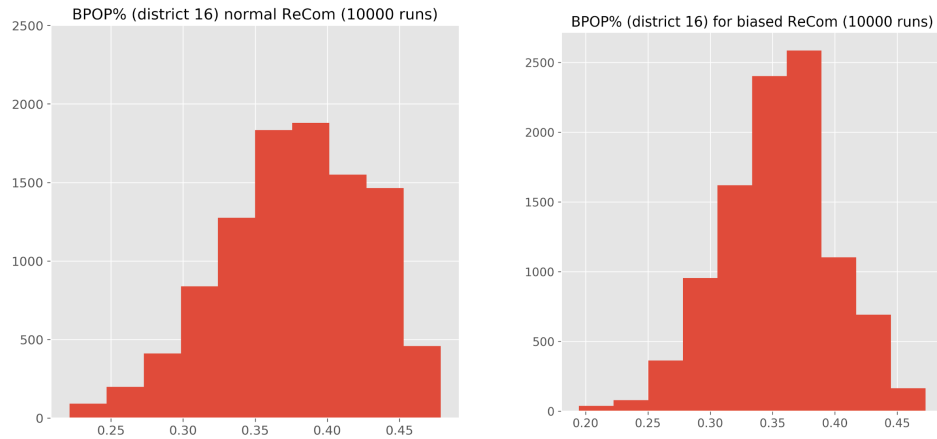

Figure 6: Comparison of normal and biased Recom

In the biased case, the distribution is closer to normal, and it also hits more extreme values. This suggests that we have sampled from more of District 16's possibility space in the same number of ReCom steps.

2.5 Challenges and Future Work

1) Dealing with state-level and district-level metrics; 2) properties of biased Recom algorithm.