An Algorithmic and Computational Approach to Optimizing Gerrymandering

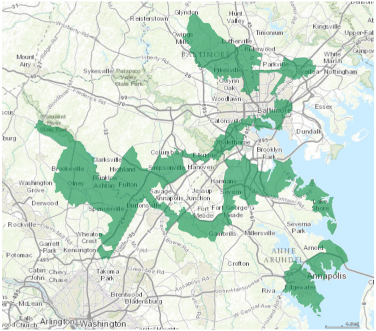
Kyle Gatesman
Mentor: James Unwin, University of Illinois

May 20, 2017

Introduction

What and Why:
Voting Districts in Democracy

- Determine elected representatives
- Equal population
- Determined periodically by
  - State legislature
  - Independent commission
- Analogous to electoral college

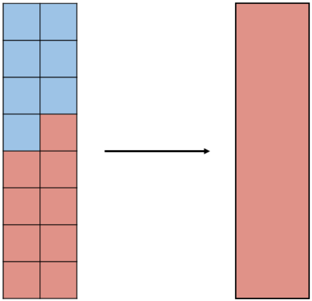


Maryland's 3rd District
(Source: Wikipedia)

Importance

Misrepresentation: districting can lead to unbalanced districts, making certain voters ineffective.

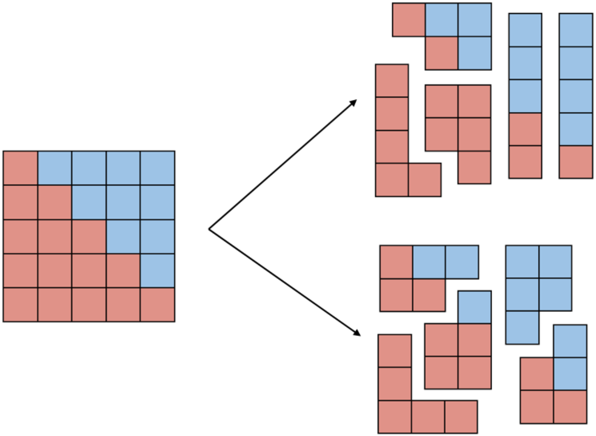
Gerrymandering exploits misrepresentation for political gain.



"We have to end the practice of drawing our congressional districts so that politicians can pick their voters, and not the other way around."

– Barack Obama

Impacts of Gerrymandering



Friedman-Holden Approach

We considered a specific approach to optimal gerrymandering:

The Friedman-Holden districting approach is based on

- Extremity
- Continuous distribution
- Shock factor model
- Population continuity
- Geography not accounted for

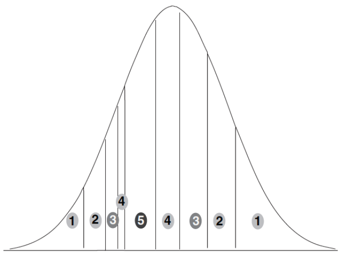


FIGURE 6. AN EXAMPLE OF THE OPTIMAL STRATEGY

Friedman, Holden, Am Econ Rev. (2008), 98:1, 113-144

We first aim to study the geographic distribution of districts that arise.

Implementation: Voter Distribution

To study the districts arising from Friedman-Holden approach we take

- Population on a lattice
- Lattice units will be associated to either Proponent or Opponent.
- Distribution from probabilistic walker method
- Randomized, but mimics a 2D Gaussian distribution

We take an 11×11 lattice and all voters "walk" from the center point. This gives an overall normal distribution, with small fluctuations:

1	4	7	4	14	15	13	8	6	3	2
4	11	15	21	26	30	31	19	12	8	3
4	13	22	38	41	44	33	42	26	6	5
7	22	28	61	70	87	60	63	42	23	7
7	17	42	63	126	108	138	68	52	40	21
7	27	57	87	99	149	95	90	40	38	13
8	16	34	76	124	110	105	69	57	27	11
12	26	37	69	79	86	75	61	34	19	6
6	10	16	38	49	47	44	34	25	11	8
4	10	12	31	34	36	19	28	19	11	3
2	2	8	10	10	11	10	14	6	2	4

Implementation: Partisanship

We assign some population unit P to be a source of partisan bias E_P.

Contribution to extremity at point Q, a distance $d(P,Q)$ away, is

$$\Delta E(Q) = E_P / d(P,Q).$$

The voter extremity at point Q is a sum over all sources

$$E_{\text{net}}(Q) = \sum_{P \in S} \frac{E_p}{\max[1, d(P,Q)]}$$

This draws on an analogy to electrostatic potential.

Take an idealized model with symmetric and proximal source points:

-0.028	-0.03	-0.03	-0.025	-0.014	0.0	0.014	0.025	0.03	0.03	0.028
-0.038	-0.044	-0.047	-0.043	-0.026	0.0	0.026	0.043	0.047	0.044	0.038
-0.051	-0.064	-0.077	-0.081	-0.056	0.0	0.056	0.081	0.077	0.064	0.051
-0.065	-0.092	-0.13	-0.17	-0.146	0.0	0.146	0.17	0.13	0.092	0.065
-0.078	-0.12	-0.205	-0.391	-0.553	0.0	0.553	0.391	0.205	0.12	0.078
-0.083	-0.133	-0.25	-0.667	-0.5	0.0	0.5	0.667	0.25	0.133	0.083
-0.078	-0.12	-0.205	-0.391	-0.553	0.0	0.553	0.391	0.205	0.12	0.078
-0.065	-0.092	-0.13	-0.17	-0.146	0.0	0.146	0.17	0.13	0.092	0.065
-0.051	-0.064	-0.077	-0.081	-0.056	0.0	0.056	0.081	0.077	0.064	0.051
-0.038	-0.044	-0.047	-0.043	-0.026	0.0	0.026	0.043	0.047	0.044	0.038
-0.028	-0.03	-0.03	-0.025	-0.014	0.0	0.014	0.025	0.03	0.03	0.028

Friedman-Holden Approach to Districting

Combining the population and parity distributions, one obtains the aggregate vote distribution, example (with net-vote +2.39):

−0.03	−0.12	−0.21	−0.1	−0.19	0.0	0.18	0.2	0.18	0.09	0.06
−0.15	−0.48	−0.71	−0.90	−0.68	0.0	0.81	0.82	0.56	0.35	0.11
−0.20	−0.83	−1.69	−3.08	−2.29	0.0	1.85	3.4	2.00	0.38	0.26
−0.46	−2.02	−3.64	−10.37	−10.22	0.0	8.76	10.71	5.46	2.12	0.46
−0.55	−2.04	−8.61	−24.63	−69.68	0.0	76.31	26.59	10.66	4.8	1.64
−0.58	−3.59	−14.25	−58.03	−49.5	0.0	47.5	60.03	10.0	5.05	1.08
−0.62	−1.92	−6.97	−29.72	−68.57	0.0	58.07	26.98	11.69	3.24	0.86
−0.78	−2.39	−4.81	−11.73	−11.53	0.0	10.95	10.37	4.42	1.75	0.39
−0.31	−0.64	−1.23	−3.08	−2.74	0.0	2.464	2.75	1.925	0.70	0.41
−0.15	−0.44	−0.56	−1.33	−0.88	0.0	0.494	1.2	0.893	0.48	0.11
−0.06	−0.06	−0.24	−0.25	−0.14	0.0	0.14	0.35	0.18	0.06	0.11

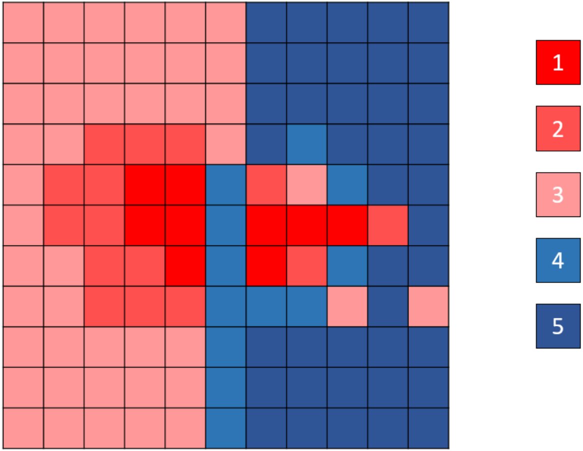
We wish to district the above such that the proponent party wins.

Our algorithmic implementation of Friedman-Holden involves:

- A preset population benchmark per district
- Chunking of territorial units
- Fine-tuning
- Recursion on subsequent, moderate districts

Program Results: Districting

Visualization of the five districts determined by our algorithm



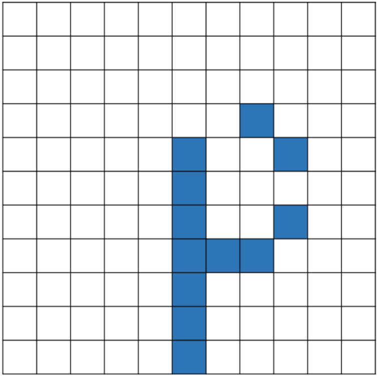
Limitations

As anticipated, application of the unmodified Friedman-Holden approach does not satisfy geographic restrictions:

- Continuity (legally required)
- Compactness
- Convexity

As example, see District 4 (right).

We are refining our algorithm to better adapt to a realistic setting.



We hypothesize that the efficacy of gerrymandering becomes more limited when more constraints need to be satisfied.

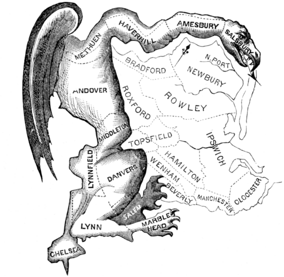
Summary and Future Directions

We have studied the Friedman and Holden approach to gerrymandering.

Our lattice study shows Friedman-Holden leads to non-continuous districts.

We are working to construct an algorithm incorporating restrictions on districts from:

- Continuity
- Compactness
- Convexity



We aim to show these constraints make gerrymandering more difficult.

Studying gerrymandering methods will aid in detecting and inhibiting it.

Acknowledgments

I am grateful to the MIT PRIMES-USA faculty for the opportunity to present at this meeting, my mentor, Prof. James Unwin, for facilitating my research, and the head mentor, Dr. Tanya Khovanova, for providing feedback.

Thank you!