When Topologists Are Politicians...

Nikifor C. Bliznashki
Aaron Pollack
Russell Posner
Duke University
Durham, NC

Advisor: David Kraines

Summary

According to Supreme Court Justice William Brennan, former Supreme Court Justice Sandra Day O'Connor once noted that any politician who did not do everything to secure power for the party "ought to be impeached" [Toobin 2003]. Though Congress argues that wild election districts such as "a pair of earmuffs" are inherently fair and reasonable, they are so counterintuitive that such claims can be exceedingly difficult to believe.

Defining the big picture of what is "fair" can be left to philosophers—or computers. Using a novel method, we divide states into districts of equal population, with each district as compact and elementary as possible, where compactness is defined as the moment of inertia of the district with respect to the population density. By not examining any other demographic data in the grouping, we avoid many of the biases that people may impose. We obtain districts considerably more compact than current congressional districts for Ohio and New York.

Since it is constitutionally sound to group people into congressional districts by "shared interests," we extend the problem by allowing other demographic data to be considered in the formation of such districts. to form revised districts that seek to preserve uniformity of these qualities. We identify how suitable these solutions are and determine their advantages and disadvantages.

Finally, we propose alternative districting techniques that take into account county boundaries and natural boundaries.

The UMAP Journal 28 (3) (2007) 249–259.

Introduction

When topologists are politicians ... the dimension of New York's 13th congressional district might not even be an integer. Districting is usually handled by a political and partisan body, such as the state legislature or the governor. Unsurprisingly, partisans district to maximize the number of representatives of their party will in Congress. This process is called *gerrymandering*, after Elbridge Gerry, governor of Massachusetts in 1812, who famously approved a congressional district that resembled a salamander.

Gerry's salamander of 1812 could not even hold a candle to such well-known districts of today such as Louisiana's "the 'Z' with drips," or Pennsylvania's "supine seahorse" and "upside-down Chinese dragon" districts. Such awkward and complex districts can lose sight of the primary goal of the House of Representatives as outlined in the U.S. Constitution: to provide regional representation to the people. We seek a "fair" and "simple" districting that maximizes accessibility of all people to regional representation, while providing a partitioning insensitive to partisan motives.

We define an objective function F whose minimum gives what we define to be the best districting. We apply this method to New York and to Ohio.

Definitions

- **Block:** A unit of area that corresponds to a fixed number of people. Since population densities vary, block sizes vary. A block is marked in the plane by a pair of (x, y) coordinates.

- **District:** A collection of a fixed number of blocks (thus having a constant population).

- **Capitol:** The average of the coordinates of each block in the district, an approximation of the center of its population.

- **Fairness:** In *Shaw v. Reno* [1993], the U.S. Supreme Court mentioned that acceptable ways of districting a state include "compactness, contiguity, [and] respect for political subdivisions or communities defined by actual shared interests." By compactness, the justices were alluding to a vague notion that congressional districts should be more like squares or circles than "spitting amoebas" (read: Maryland's Third District).

- **Compactness:** Suppose that a district D contains n blocks, $z_1, \ldots, z_n$, with capitol c. Compactness C is the variance of the spatial distribution of the population:

$$C = \sum_{i=1}^{n} \|z_i - c\|^2.$$

When C is small, we conclude that the district's constituents live within a relatively small area.

- **Shared Interests** (S)**:** We assign a vector to each block, giving one component to each interest, and minimize the sum of the variances of the components over the district. That is, given vectors v_i associated with blocks z_i and mean interest vector $\vec{\mu} = \frac{1}{n}\sum_{i=1}^{n} \vec{v}_i$, the shared interest is

$$S = \sum_{i=1}^{n} \|\vec{v}_i - \vec{\mu}\|^2.$$

A Note on Shared Interest

Though race, gender, age, and religion are important issues for many, legal ambiguities exist in the use of such measures. Because the benefit is yet unproven of either grouping together or dispersing such groups among congressional districts, we do not use these data. Though measuring political affiliation is an entirely legal and often implemented districting tool, we remain nonpartisan to avoid inadvertently favoring one party over another.

Specific Formulation of the Problem

We measure fairness of a partition of a state into congressional districts. Let $P = \{D_1, \ldots, D_k\}$ be a partition of blocks into congressional districts, such that each district is contiguous and has the same number of blocks. We define

$$f(D_i) = w_1 C_i + w_2 S_i,$$

where w_1 and w_2 are positive weights that are the same for all districts. We define globally

$$F(P) = \sum_{i}^{k} f(D_j).$$

We seek a partition that minimizes F.

Assumptions

- We have accurate data about a state's population, geographical layout, and other relevant factors.
- The population represented by one block is small enough to ensure that districts have negligibly different numbers of people.
- The initial assignment of blocks to districts is random enough to assure that the districting to which our algorithm converges is near the global minimum.

Background and Goals

Weaver and Hess [1963] set the standard for computerized nonpartisan districting. Using integer programming methods, a set of capitols (called "LDs" in their paper) were matched with blocks ("EDs") so as to minimize moment of inertia, i.e., our C. Repeatedly, the LDs were relocated to the appropriate centers of mass and then the EDs were redistributed to each LD until the moment of inertia hit a local minimum. By repeating many times with a large number of initial conditions, they hoped to approximate the global minimum and thus derive the most compact districting. Though precise, integer programming algorithms on large sets of data are extremely time-consuming. Though Weaver and Hess's methods found applicability at the county and small state level at best, their landmark work paved the way for the development of a variety of approaches.

We create a model that expands on theirs with the following goals:

- Find the ideal partition P^*, i.e., the one that minimizes F globally.
- The method to find P^* should be versatile—able to find the ideal partition for a wide variety of shared interest functions S.
- The method should be scalable—able to handle large quantities of data quickly.

Friendly Trader Method

Our method starts with an initial arrangement of blocks into districts and moves blocks between districts to decrease F. Let the blocks be arranged into n districts. By our method, district D attempts to trade blocks to reduce its $f(D)$. However, our districts are "friendly traders"—they conduct only trades that make the districts as a whole better off (i.e., reduces F). Our districts are so friendly, in fact, that they will execute trades that raise $f(D)$ so long as F decreases. Since the composition of our districts changes after each trade, capitols must be recalculated at each step. The problem of finding a minimum then reduces to finding and executing all trades that reduce F until no more exist.

How Are Blocks Determined?

We obtained demographic data are obtained from the 2000 U.S. Census [U.S. Census Bureau n.d.]. New York State is partitioned into roughly 5,000 tracts, each with a specific population and coordinates in latitude and longitude. For each minor civil division, we assign one block per 250 people, rounding population to the nearest 250. We spread these blocks evenly within their minor civil division. Thus, each block has the same population, and population density corresponds to block density.

How Are Districts Initialized?

We try to devise an initial partition with low F. First, we arbitrarily choose 29 blocks to be district capitols. Each capitol's position is assumed to be the center of population and its interests the mean interests of the district. One by one, each capitol picks districts that "fit well" with the capitol's location and interests. The process is like a professional sports draft, where teams take turns picking players who suit each team well. After all the blocks are assigned to capitols, trading of blocks begins.

How Do We Maximize Compactness?

Let D_1 and D_2 be two districts and b_k a block in D_1. By moving b_k from D_1 to D_2, we form districts D_1' and D_2'. Define

$$\Delta F(D_1, D_2, b_k) = f(D_1') + f(D_2') - f(D_1) - f(D_2).$$

To determine which trades to make, we first find a block $b*_{12}$ in D_1 such that $\Delta F(D_1, D_2, b*_{12}) \leq \Delta F(D_1, D_2, b_k)$ for all blocks b_k in D_1. Let us call $b*_{12}$ the *best block* from D_1 to D_2.

Now we can define a fully connected directed graph G, with the vertices of G the districts D_j and the edge $v_i \to v_j$ having length $\Delta F(D_i, D_j, b*_{ij})$, where $b*_{ij}$ is the best block from D_i to D_j, where the length is negative if the trade decreases F. To reduce F by trades, we search for cycles of negative length in G. (The length of a cycle is the sum of the lengths of the edges composing the cycle, counting multiplicity if an edge appears more than once.) A cycle with negative length corresponds to a group of trades that reduce F. Hence, finding good trades of blocks between districts reduces to finding cycles of negative length in the digraph.

That problem in turn reduces to the simpler problem of finding a shortest path between any two vertices—that is, a path of minimum length in which no edge is used more than once. The Bellman-Ford-Moore algorithm modifies a standard shortest-path algorithm to find any negative cycle [Cherkassy and Goldberg 1999]. (This algorithm finds only the *first* negative cycle encountered and thus gives no choice of cycle.) Once a trade is found, it is completed and capitols are recalculated. We reject any trade that, after recalculation of capitols, actually increases F. Since we are making only trades that strictly reduce F, when F can no longer decrease through trading, we have achieved a local minimum.

Results

We implemented the algorithm in a computer program and simulated the process for New York and for Ohio. No matter the starting configuration for the state, we always ended up with the same shape for the districts, making us believe that these data sets are big and diverse enough to converge always to a unique global minimum.

Figure 1 shows our apportionment of New York, calculated to make the regions most compact ($w_2 = 0$). We then redid New York using compactness as a guide but with the objective function F weighted toward preserving population density. The results are shown in **Figure 2**.

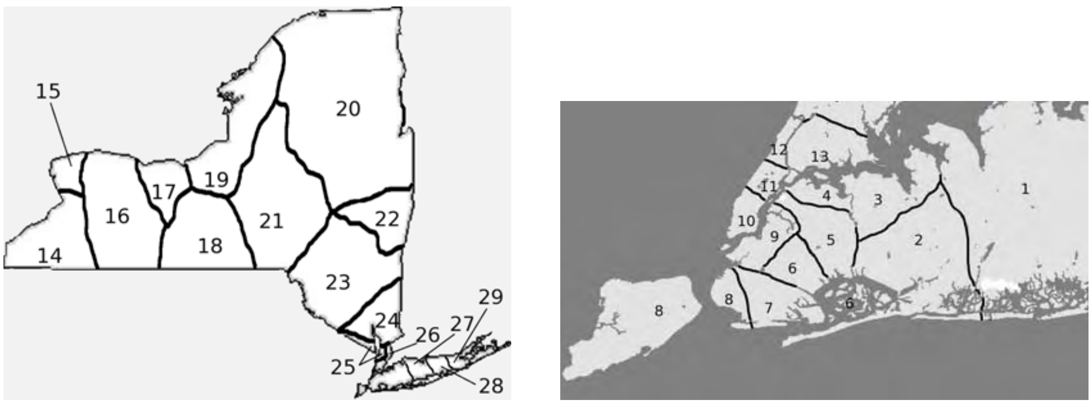

Figure 1. The most compact apportionment of New York, with close-up of New York City on right.

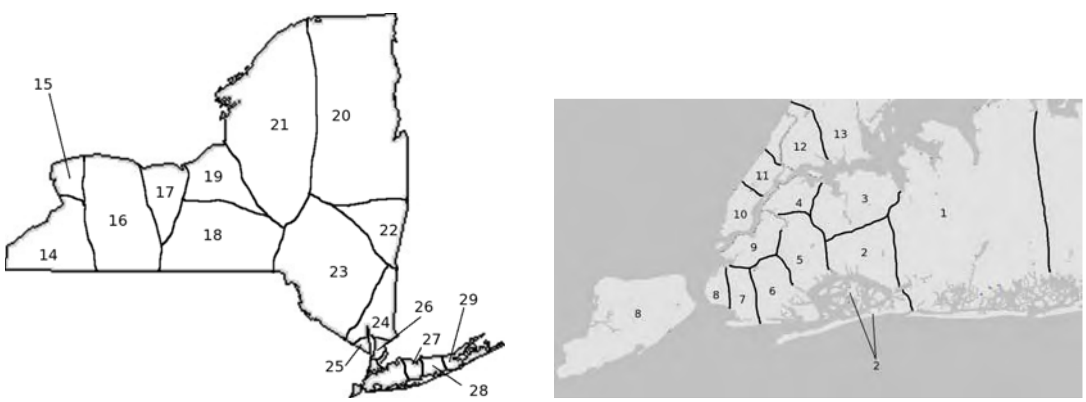

Figure 2. Apportionment of New York based both on compactness and population density.

We also tested our algorithm on Ohio to check that our method is applicable in other circumstances. In **Figure 3**, Ohio is partitioned using only compactness as a guide; **Figure 4** use the same function F as **Figure 2**.

In Ohio, congressional districts are designed with preserving county lines in mind. In **Figure 6**, we attempt not to split counties between congressional districts. We thus add a term in $f(x)$ to take into account county separation. This idea can easily be extended to natural boundaries such as rivers and highways.

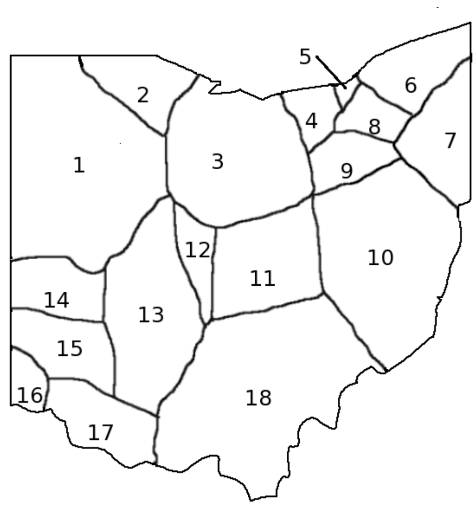

Figure 3. Districting of Ohio, based solely on compactness.

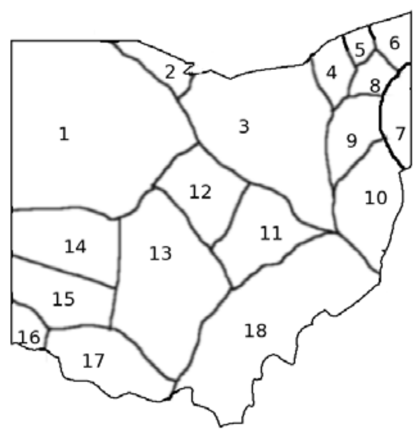

Figure 4. Districting of Ohio, based on both compactness and population density.

Figure 6. Districting of Ohio, based solely on preserving county boundaries.

Analysis of Results

Our results are primarily visual and not numerical. One weakness of our method is the lack of quantifiable data for comparison. Since F itself can be varied, normalizing it for use from one application to another is far from trivial. In addition, the remarkable reproducibility of our results given a wide variety of initial conditions almost entirely eliminates the need for separate numerical results.

In **Figures 2** and **4**, after the sorting of blocks, we connected a line surrounding all boundary blocks in a district and softened the line to make it smooth. Compared to current congressional districts, those produced by our algorithm are a vast improvement in simplicity, corresponding to a reduction in C by a factor of 7 for Ohio and by a factor of about 22 for New York. Since our model evaluates districting through distance, it promotes star-shaped districts (in which the capitol is connected to every point in the district by a straight line) rather than fully concave districts, improving accessibility and reducing complexity of the districts. Our model handles the problems of districting New York and Ohio wonderfully, achieving very simple, reasonable results.

However, interest-weighted models are where our model really shines (**Figures 3** and **5**). Since we chose not to gauge common interest by controversial factors such as race or age, we chose the tamest quantity possible: population density at each block. We felt that this quantity would be useful to group districts by, since urban issues tend to differ from rural issues, and thus both city slickers and farmers alike could obtain representation for their grievances.

In some ways, our model already favors uniform population density across districts. Since blocks in urban areas are more densely packed, they naturally migrate to the same district. Our adjustment then merely increases the population density component, producing a noticeable change for Ohio, with tightened districts around the major cities of Columbus in the center, Cincinnati in the southwest, and Cleveland in the northeast, as well as a tightening of the districts around the densely populated Bronx and Queens in New York City. These solutions improve compactness and uniformity of population density compared to current congressional districting (quite unsurprisingly); they also demonstrate the existence of a range of reasonable solutions that satisfy the goals of compactness, simplicity, and fairness.

Since many states have independent county governments (including New York and Ohio), we ran an alternative solution set for Ohio in which we tried to preserve county lines. Our mean interest vector $\vec{\mu}$ assigned a direction for each county tabulating the number of blocks in each. We then weighted our solution with the goal of preserving county lines. **Figure 6** shows the great success of this solution. In most districts, boundaries coincide almost perfectly with county lines. The advantage of the county-based districting solution is clear: Since citizens pay taxes to and receive services from county governments, allowing counties to have exclusive congressional representation allows for easier handling of issues on the local level. Surprisingly, such lines can be taken into consideration with little consequence on compactness or simplicity.

Why Our Model is Fair

The question of fairness is much more difficult. To give an example of this challenge, we consider the Fourteenth Amendment to the U.S. Constitution, which states that all races are strictly equal under the law. However, the Voting Rights Act of 1965 states that the government will assist in facilitating the voting of minority areas. Thus, even the government itself has trouble deciding whether "fair" involves helping the often-disadvantaged to realize their rights or involves giving every person exactly the same treatment. We argue that our model is fair because it remains passive and uninvolved. It only takes a set of directives (i.e., the function F) and produces a solution that divides the region into relatively uniform districts with respect to F. If nonpartisan goals are incorporated, such as uniform population density or compactness, a nonpartisan solution arises.

However, any component of data can be (and likely has been) misused. For example, African Americans tend to be affiliated with the Democratic Party, while those in rural areas tend to be Republican. Thus, race and population density can be delicately used to achieve partisan aims. Our model is fair because it assigns no judgment to any of these considerations. Rather, it is designed to improve the citizen's accessibility to attentive and diligent representation in order to maximize every individual's rights and powers.

Strengths and Weaknesses of Model

Our model achieves all of the initial goals. It is fast, and could handle large quantities of data, but also has flexibility. Though we did not test all possibilities, we showed that our model optimizes state districts for any of a number of variables. If we had input income, poverty, crime, or education data into our interest function, we could have produced high-quality results with virtually no added difficulty. As well, our method is robust. Moreover, we can divide areas into fairly simple, contiguous, and uniform regions. Our model also consistently leads to useful minima.

The primary weakness of our model is the absence of good nicknames for our districts—somehow districts such as "egg" and "sort of diamond-shaped thing" don't raise any excitement.

Though we achieved solid equilibria, our model in no way guarantees that it will ever find a global solution. To see this, consider a rectangle with different sides, and assume that we have 10 points at each of its vertices. Moreover, let the districts initially be the long sides. It is easy to check that no trade will occur, and thus this configuration is a local, but not a global, minimum.

The other primary weakness of our model is our lack of metrics for comparison. Though compactness and shared interest levels are appropriate measures for comparison of two models within a state, we lack invariant metrics for assessing the quality of one districting versus another.

Food for Thought

Given the proper data, our model can do much more than merely political districting. At heart, it simply attempts to group regions into smaller parts, unified by whatever characteristics desired. For example, if a governing body wanted to determine where to build police stations or hospitals, it could input weighted crime, health, poverty, and/or age statistics into the model. The model could then partition the area into small regions united not only by spatial relations but also by needs and desires. Thus, our model could help politicians and authorities most effectively deploy public resources and services. Ironically, our nonpartisan partitioning method could be a politician's best friend! Of course, by inputting political affiliation data, politicians could identify partisan strongholds in order to plan campaigns.

Conclusion

We set forth an algorithm to determine congressional districts, given data on location, population, and any other factors desired. The algorithm is intended to be fair, or nonpartisan, in stark contrast to the political process of gerrymandering. Characteristics that we consider to be fundamental in the division of a state into congressional districts include contiguousness, compactness, and shared interests or concerns among a district's citizens.

We assume that a state can be divided into blocks of small constant population and interpret the problem of congressional district apportionment as the distribution of these blocks to the districts so that each district contains a fixed number of blocks (and therefore all districts have the same population). Furthermore, we define an objective function F that measures the quality of a distribution. Finding good partitions is equivalent to finding distributions with low values of F. Our goal, therefore, was to find a partition that minimizes F. This is a useful formulation of the problem because, if all agree to use this method beforehand, the existence of a global minimum of F (our problem is finite) guarantees that if this minimum is found, there can (should) be no partisan squabbling as to the legitimacy of the solution obtained.

Admittedly, our algorithm is only guaranteed to find local minima of F. However, simulations with random initial starting values seem to converge to the same final apportionment, suggesting that the local minima that our algorithm finds are close to the global minima. Additionally, while we have implemented certain particulars to quantify shared interests of citizens in a district, our procedure for determining congressional districts is flexible; with a simple change in the particulars, it can partition blocks into districts under other criteria.

References

Cherkassy, B.V., and A.V. Goldberg. 1999. Negative-cycle detection algorithms. *Mathematical Programming* 85: 277–311.

Toobin, Jeffrey. 2003. The great election grab. *New Yorker* 79 (38) (8 December 2003): 63–80.

Shaw v. Reno, 509 U.S. 630 (1993).

U.S. Census Bureau. n.d. FactFinder. `http://factfinder.census.gov/home/saff/main.html?_lang=en` .

Weaver, James B., and Sidney W. Hess. 1963. A procedure for nonpartisan districting: Development of computer techniques. *Yale Law Journal* 73: 288–308.