Nate Bottman, Wes Essig, and Sam Whittle
University of Washington
2007 Outstanding Winner

Why Weight? A Cluster-Theoretic Approach to Political Districting

February 17, 2007

Abstract

Political districting has been one of the most contentious issues within American politics over the last two centuries. Since the landmark case of *Baker v. Carr*, in which the United States Supreme Court ruled that the constitutionality of a state's legislated districting is within the jurisdiction of a federal court, many within academia have attempted to produce a rigorous system for determining a set of districts for a given state. In this paper, we attempt to improve upon these past efforts. We propose both a modified form of classical **K**-means clustering and an interesting algorithm called the shortest-splitline algorithm to accomplish impartial redistricting. As an example, we apply our methods to redistricting the state of New York, and, as further examples, to Texas and Colorado. Both methods use only population density data and state boundaries as inputs and run in a feasible amount of time. Our criteria for successful redistricting include contiguity, compactness, and sufficiently uniform population. The **K**-means method produces districts similar to convex polygons and the splitline method guarantees that the resulting districts have piecewise linear boundaries. The **K**-means method has the advantage of allowing seeding of the district centers. The centers of the generated districts then roughly correlate to the existing districts, by proper seeding, but the resulting boundaries are vastly simpler.

Contents

1 Introduction

In the words of Ted Harrington, chair of political science at the University of North Carolina,

> There is no issue that is more sensitive to politicians of all colors and ideological persuasions than redistricting. It will determine who wins and loses for eight years [You88].

The writers of the constitution created the House of Representatives with the intention that it would be the branch of government most responsive to the people. The reality is just the opposite. Though representatives are elected every 2 years, instead of every 4 or 6 years, almost 400 of the 435 seats of the House are not contested as a result of the extraordinary power of gerrymandering. With the immensely detailed amount of data and unlimited computing power available to politicians today, gerrymandering has been elevated to an art. With only the requirements that districts be connected and all have equal population, it is possible to pinpoint candidates and place them in a different district than their neighbors [Too03].

Though undemocratic, gerrymandering is nearly always legal (see, for instance, [Bac86]) and has been used to obtain striking results. In 2002 only 4 incumbent representatives lost their bid for reelection — the lowest total ever [Too03]. We will argue that it is certainly true that any attempt to fairly restructure legislative districts needs to ignore the human factors that overwhelmingly determine the current redistricting process. Defining some measure of compactness is essential to ensure fair districts. Both methods we describe produce districts that at first glance are clearly simpler than the existing ones.

Restructuring the districts with no regard to the current layout would be more difficult to implement. We will use the centers of the existing districts as seeds for our clustering algorithm. Thus, the new districts have some correlation to the existing districts, but their boundaries will be determined in a fair manner. The core of many districts will be roughly the same, while the boundaries will be dramatically simpler. This will effectively counteract the effects of gerrymandering, without being overly difficult to put into use immediately.

1.1 Plan of Attack

Our goal is to develop an algorithmic process for dividing an arbitrary region into k legislative districts, which satisfy some heuristic definition of *fairness*. In order to do so, we must do the following:

- **Define terms.** Crucial to creating a model is defining the somewhat ambiguous terms *fairness* and *simpleness*.
- **Define metrics for comparing algorithms.**

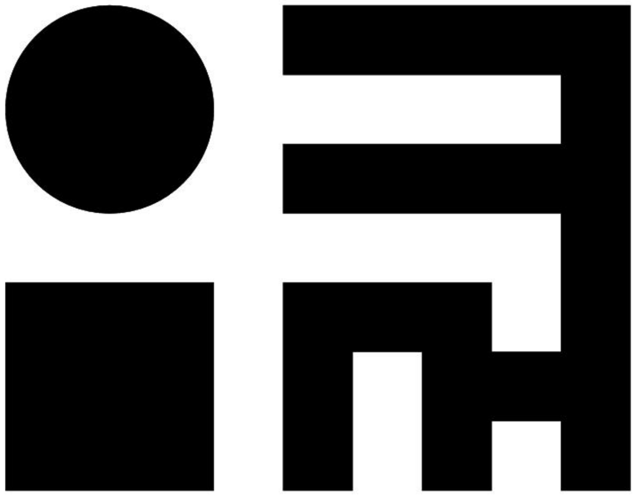

Figure 1: The compactness quotients of the circle, square, and gerrymander are 1, $\pi/4 \approx 0.79$, and $23\pi/576 \approx 0.13$, respectively.

1.2 Defining Simpleness

We say that district A is more simple than district B if *district A is contiguous*, and *district A is more compact than district B*.

- **Contiguity.** We say that a district is contiguous if it is arcwise-connected; that is, if one can travel from any point a to any other point b in district A while remaining entirely within district A. If A contains regions separated by bodies of water, A is contiguous if all regions are connected by water and each region is arcwise-connected.
- **Compactness.**

  Intuitively, we say that a district is compact if it does not meander excessively. This is a hard concept to formalize; many authors give only a hasty definition of compactness, and some have even argued that compactness is ambiguous to the point of being irrelevant in a serious treatment of districting. Nonetheless, we will now attempt to develop a suitable definition of compactness.

1.3 Towards a Suitable Conception of Compactness

In [You88], Young gives compelling reasons for abandoning all of the definitions of compactness mentioned in the second appendix.. Interestingly enough, Young does not consider the following adjusted version of the Schwartzberg Test, which is alluded to in [GN70]:

Definition 1. *We say that district A is more compact than district B if*

$$\frac{4\pi\,\mathrm{Area}_A}{(\mathrm{Perimeter}_A)^2} > \frac{4\pi\,\mathrm{Area}_B}{(\mathrm{Perimeter}_B)^2}.$$

Call the quantity $4\pi\,\mathrm{Area}\,/\,\mathrm{Perimeter}^2$ *the* compactness quotient.

For a circle of radius r, this ratio is equal to

$$4\pi\cdot\frac{\pi r^2}{(2\pi r)^2} = 1.$$

It is well-known that the shape with the largest ratio of area to squared perimeter is the circle (see, for instance, [Fol02]). Because of this, the quantity

$$4\pi\cdot\frac{\mathrm{Area}}{\mathrm{Perimeter}^2}$$

is restricted to the interval $[0,1]$.

As seen in figure 1, a compactness quotient of 0.13 is visually quite bad. Using the fact given in [Bou88] that the area of a non-self-intersecting closed N-gon (with the k-th vertex taken in counterclockwise order equal to (x_k, y_k)) is equal to

$$\frac{1}{2}\sum_{i=1}^{N-1}\left(x_i y_{i+1} - x_{i+1} y_i\right),$$

we have calculated the compactness quotients of several actual districts by approximating their boundaries by piecewise linear segments. The results illustrate the inappropriate nature of the districts currently in place. Two of New York's more sprawling districts, the 8th and 28th, produced compactness quotients of 0.097 and 0.101, respectively — even worse then the gerrymander shown in figure 1! The two most compact districts in New York, the 26th and 21st, had compactness quotients of 0.406 and 0.498, respectively. We decided that the mean for any state should be at least .6. With this condition the average district in every state would be better than the best districts currently in New York. Furthermore we insist that .25 should be more than 2 standard deviations from the mean. It is not possible to require that all districts be greater than .25 as several districts will inevitably end up having most of their border coincide with the border of the state.

1.4 Defining Fairness

Almost all unfairness occurs when political and social measures factor into redistricting decisions. Practices such as concentrating supporting voters in a single district, diluting opposing voters over several districts, placing two incumbents in the same district and forcing them to run against each other, and isolating minorities have been seen many times before (see [Too03] and [Hay96]), and are all the result of districing being controlled by those who attempt to skew voting patterns. In general, one can summarize past districting patterns in the following way:

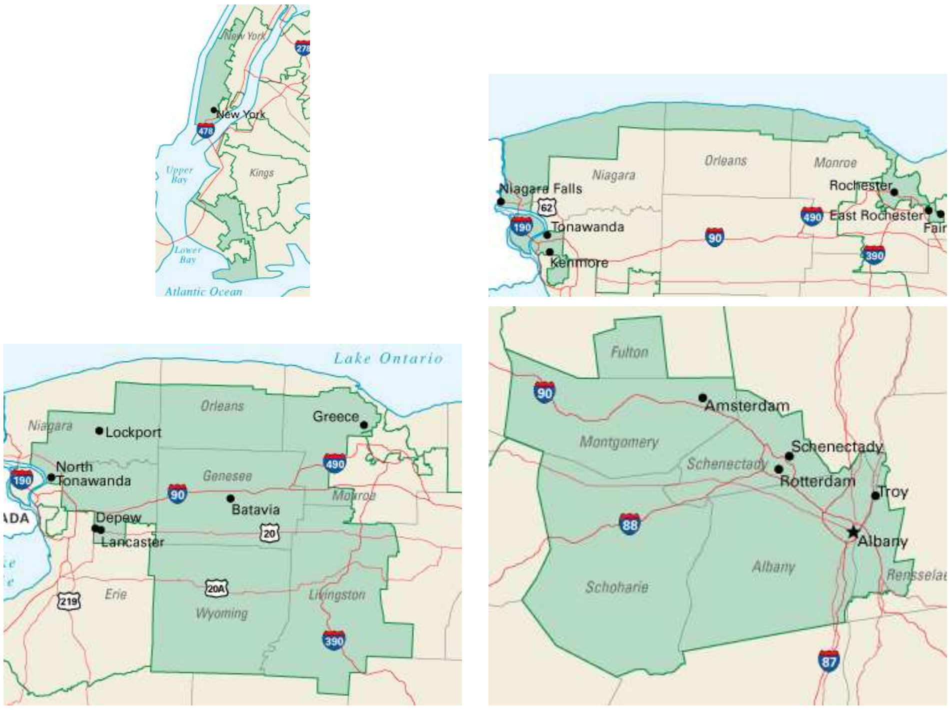

Figure 2: Current districts 8, 28, 26, and 21, from left to right and top to bottom, with compactness quotients of 0.097, 0.101, 0.406, and 0.498, respectively.

- **Unfair districting stems from either human biases or poorly designed algorithms.**

Our computer simulations do not use any of this extraneous data. The only data that we have used is population density and the boundary of the state. Therefore, the determination of districts is completely unbiased. While it may hold a district may be unfair on a local scale, in that it divides up a community with a common interest — for instance, a community of apple-growers may be split between two districts — on the national scale, such imbalances will even out. Because of this, there will be no pathological examples of disproportionate representation.

2 Applying the Theory of Data Clustering

The theory of data clustering is the theory of classifying n *observations* (or objects) into m groups — for instance, placing two carts' full of groceries into into paper sacks. There are two main benefits of applying a cluster-theoretic algorithm to a given data set:

- **Data clustering often reveals an internal structure that may not have been initially apparent.**
- **It is often much easier to work with a small number of clusters than with a large number of raw data.**

The philosophy of data clustering is that we should be able to divide our data into a (not necessarily fixed) number of clusters, and that the elements of a given clusters should be somehow *similar*. In general, data clustering is applied to problems that deal with a large number of variables. For instance, when data clustering is used to create an animal taxonomy, there are a myriad of variables — mode of reproduction, mode of transportation, presence and type of spine, ideal diet, preferred habitat, and so forth [And73]! Because of this, it is usually very difficult to determine the "proper" way to cluster data [AC84].

In the case of attempting to draw up simple and fair congressional districts, we can apply data clustering in the following way:

- **Split the state into small, discrete units.** Our units correspond to geographic locations of census population measurements [fIESIN].
- **Determine some partition of these units, such that the subsets of this partition can be viewed as clusters.** Note that the only variables present are the location and population of each unit.

After defining a method for ordering the preference of cluster partitions, we may suppose we are done with the problem: all that is left is to look at all possible cluster partitions and choose the best one! However, this turns out to be not feasible. In [AS68], Abramowitz and Stegun give a proof of the fact that the number of ways of sorting n observations into m groups is a Stirling number of the second kind:

$$S_m^{(n)} = \frac{1}{m!} \sum_{k=0}^{m} (-1)^{m-k} \begin{pmatrix} m \\ k \end{pmatrix} k^n.$$

For instance, there are more than 10^{15} ways to sort 25 objects into 5 groups. It is clear that we need some sort of algorithmic process in order to determine an appropriate partition of clusters.

3 The K-means Algorithm

3.1 Standard Algorithm

The K-means algorithm is an iterative method for data clustering. Let $D = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{R}^n$ be the data to be clustered, and let $S = \{\mathbf{s}_j\}_{j=1}^{K}$ be a set of *seeds*. Suppose we desire D to be partitioned into K clusters; let the i-th cluster be denoted by C_i. Associate to the i-th cluster a *geographical center*, denoted by $\mathbf{c}_i$. Given an distance function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the K-means algorithm proceeds as follows.

- **Initialization:** for all C_i, let $\mathbf{c}_i = \mathbf{s}_i$.
- **Iteration:**
  - **Assign points to clusters:** For all $\mathbf{x} \in D$, associate $\mathbf{x}$ to the center $\mathbf{c}_i$ such that $f(\mathbf{x}, \mathbf{c}_i)$ is minimized.
  - **Update cluster centers:** Redefine
$$\mathbf{c}_i = \frac{\sum_{\mathbf{x} \in C_i} f(x, c_i)}{\sum_{\mathbf{x} \in C_i}}.$$
- **Repetition:** If updating cluster centers changes at least one cluster center, repeat the iteration step. Otherwise, stop.

3.2 Weighted Algorithm

To generate districts of appropriate population, we have added a weighting system to the standard algorithm. Let each cluster correspond to a legislative district. Let $D = \{\mathbf{x}_j\}_{j=1}^{N} \subset \mathbb{R}^2$ be the set of census coordinates. Thus, $\mathbf{x} \in D$ corresponds to the position of a population measurement. Define a population function $p : D \to \mathbb{R}$ such that p_i is the population at the coordinates specified by $\mathbf{x}_i$. A cluster C_j is defined by its points $\mathbf{x} \subset \mathbb{R}^2$, its center $\mathbf{x}_j \in \mathbb{R}^2$, and some *weight* α_j. Define f to be the Euclidean distance function in $\mathbb{R}^2$. Our weighted K-means algorithm proceeds as follows:

- **Initialization:** Using the *standard* **K**-means algorithm, assign points to clusters and centers to appropriate positions.
- **Iteration:**
  - **Assign points to clusters:**
    For all $\mathbf{x} \in D$, associate $\mathbf{x}$ to the center $\mathbf{c}_i$ such that $\alpha_i f(\mathbf{x}, c_i)$ is minimized.
  - **Update cluster centers:**
    Redefine
$$\mathbf{c}_i = \frac{\sum_{\mathbf{x} \in C_i} p_i f(x, c_i)}{\sum_{\mathbf{x} \in C_i}} p_i.$$
  - **Update cluster weights:**
    Redefine
$$\alpha_i = g\left(\sum_{\mathbf{x} \in C_i} p_i\right),$$
    where g is defined below.

- **Repetition:** If the properties of the clusters are within our tolerance levels we stop. Otherwise, repeat the iteration step.

By adjusting the weights, we are able to control the growth or decay of the clusters. If the weight of a cluster increases, data points are more likely to be grouped in other clusters. Similarly, decreasing the weight helps to increase the population of a cluster. Thus the weight function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is crucial in the performance of the algorithm. We define:

$$g(p, w) = w\sqrt{\frac{i}{i_0}} + w \cdot \frac{p}{p_0} \cdot \sqrt{1 - \frac{i}{i_0}},$$

where i is the current iteration, i_0 is the maximum number of iterations, and p_0 is the desired population for each cluster. Towards the beginning of the algorithm, i/i_0 is low causing the term $w * p/p_0 * \sqrt{1 - i/i_0}$ to dominate the weight function. As the i increases, the weight fluctuates less because $w * sqrt i/i_0$ begins to dominate $w * p/p_0 * \sqrt{1 - i/i_0}$. This enables the weights to change rapidly at the beginning of the iterative process causing the clusters to vary greatly between iterations. However, by the end of the algorithm, the weights do not change as readily, allowing stabilization over a optimal clustering. This is somewhat similar to the process of simulated annealing where initial negative actions allow the algorithm to escape local optimums and the probability a negative action is taken decreases over time.

4 Splitline Algorithm

Recently, a very elegant algorithm for districting has been proposed by applied mathematician Warren B. Smith [Smi].

4.1 Method

The idea behind the splitline algorithm is quite simple:

- Start with the number of districts for the state. Divide that number in two as evenly as possible, using integers (for instance, $18 = 9 + 9$ and $35 = 17 + 18$).
- Find the *shortest* line that divides the state into two parts such the ratio of their populations is the same as the ratio determined in the previous step.
- Repeat this process recursively on the subdivided parts until the number of parts is the same as the number of districts. At every step, the division is just a line, and so the resulting districts have piecewise linear boundaries. Using the *shortest* line ensures that the districts will have a good compactness quotient.

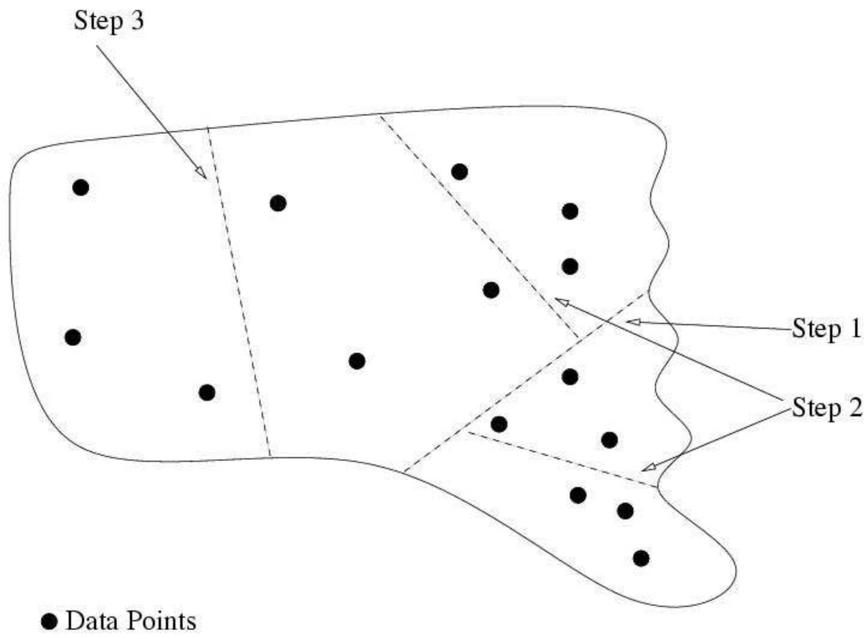

Figure 3: An illustration of the splitline method.

4.2 Demonstration

Figure 3 is a demonstration of the splitline algorithm creating 5 districts from a simple data set of 15 points. With 15 points and 5 districts there need to be 3 points in each district. 3:2 is the most balanced integer ratio that 5 can be divided into. At step 1 the algorithm divides the state into two regions with 9 and 6 people respectively, the correct ratios for 3 and 2 districts.

At step 2, it acts recursively on the 2 subdivisions. Thus the region that had 6 people is divided into regions that have 3 people each, with no more subdivision needed. The other region is divided into regions with 6 and 3 people, the appropriate numbers for 2 and 1 districts respectively.

At the third and final step, the last region is split in two and the process is complete. By using the shortest line at each step, none of the shapes end up with an unsatisfactory compactness quotient.

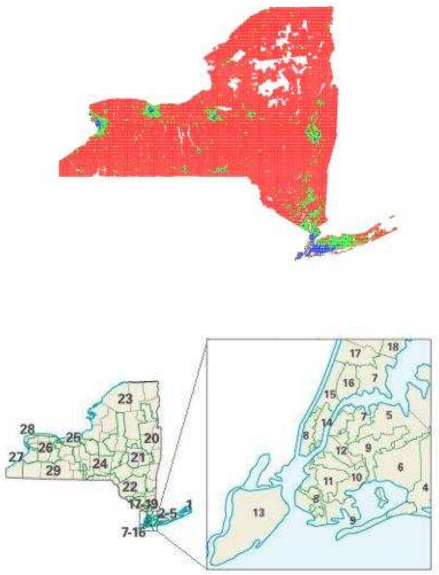

Figure 4: Plots of the population and current districting of New York state, from top to bottom.

5 An Application: Considering the Congressional Districting of New York State

5.1 K-means Algorithm

The results given by the **K**-means algorithm are generally quite good. Traditionally, when applying cluster-theoretic algorithms, it is common practice to split off any regions with particularly high population density, and to apply the algorithm to those regions separately (see, for instance, [GN70]). This was not needed for the **K**-means algorithm: even though the maximum population density of New York City is roughly $2,000$ times the mean population density of the state of New York, the **K**-means algorithm produced results within our tolerance levels.

To confirm that the weighted **K**-means algorithm we developed was an effective aid for determining districts, we used it to also redistrict Texas. Texas is a good choice because it is large and contains a variety of population densities. The **K**-means algorithm worked overall well with only a few districts being outside our desired tolerance. Because the data set was much larger, the computing time for each iteration of the **K**-means was greater. With increased run time of the algorithm, perhaps even better results could be achieved.

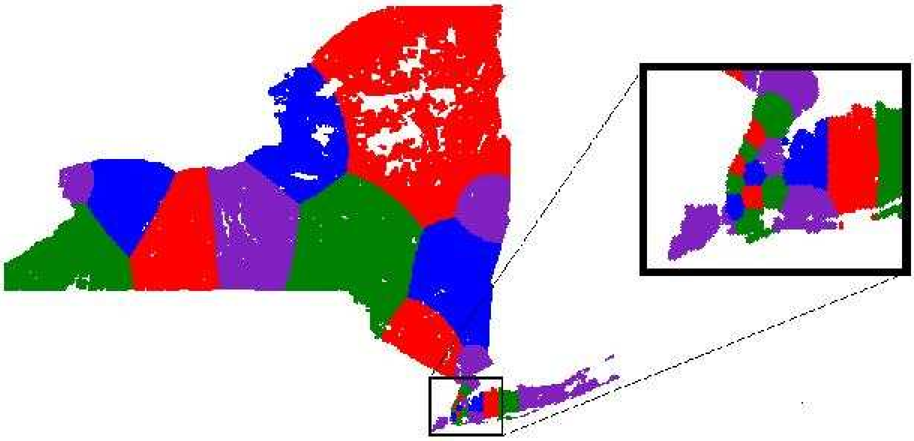

Figure 5: A proposed redistricting of New York, using the **K**-means algorithm.

5.2 Splitline Algorithm

To obtain results within our desired tolerance it was necessary to calculate the districts of New York City separately from the remainder of the state. One limitation of the current splitline algorithm is that it does not guarantee the contiguity of districts, see Figure 6. However,it will produce contiguous (and, furthermore, convex) districts for a convex state.

6 Conclusions

We conclude that both **K**-means and the splitline algorithm are viable methods for fair and simple redistricting. **K**-means produced much better results in our application to New York: the greatest value of

$$\max_{\text{alldistricts}} (1 - (\text{clusterpopulation}) \,/\, (\text{targetpopulation}))$$

, when **K**-means is applied to the whole of New York state, is no more than 2.5%. As an interesting note, while the unweighted **K**-means method clusters data into regions with piecewise linear boundaries, inclusion of the weight function effectively rounds the boundaries of the produced districts. These rounder districts have superb compactness coefficients. **K**-means also has a visually appealing output and meets all other criteria.

The splitline algorithm results are not quite as satisfactory; however, we believe that is a result of our implementation and not the algorithm. Even our flawed version of splitline produced districts simpler than the current districts in New York. Our implementation could achieve districts with either even population or high compactness coefficients, but not both simultaneously. It was also difficult to enforce the contiguity requirement in regions posessing a highly

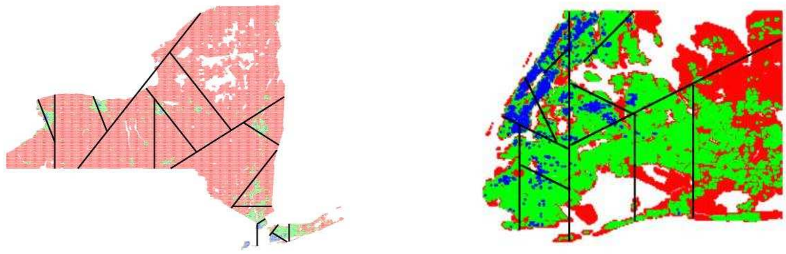

Figure 6: A proposed redistricting of New York state, using the splitline algorithm and calculating the districts within New York City separately from the remaining districts

irregular border. We note that when the splitline algorithm is applied to states with convex boundaries — say,

6.1 Towards a Suitable Definition of Compactness

As discussed in [You88], the following definitions of compactness are often used or cited in the literature.

- **The Visual Test.** A district is more compact if it appears to be more compact.
- **The Roeck Test.** Find the smallest circle containing the district and take the ratio of the district's area to that of the circle. This ratio is always between 0 and 1; the closer it is to 1, the more compact is the district.
- **The Schwartzberg Test.** Construct the adjusted perimeter of the district by connecting by straight lines those points on the district boundary where three or more constituent units (*i.e.*, census tracts) from any district meet. Divide the length of the adjusted perimeter by the perimeter of a circle with area equal to that of the district.
- **Length-width Test.** Find a rectangle enclosing the district and touching it on all four sides, such that the ratio of length to width is a maximum. The closer the ratio is to 1, the more compact is the district.
- **Taylor's Test.** Construct the adjusted perimeter of the district by connecting by straight lines those points on the district boundary where three or more constituent units (*i.e.*, census tracts) from any district meet. At each such point the angle formed is "reflexive" if it bends away from the

district and “non-reflexive” otherwise. Subtract the number of reflexive from the number of non-reflexive angles and divide by the total number of angles. The resulting number is always between 0 and 1; the closer to 1, the more compact the district.

- **The Moment of Inertia Test.** Locate the geographical center c_i of each census tract i in the district. Select an arbitrary point x and calculate the square of the distance from x to c_i, multiplied by the population of tract i. The sum of these numbers is the district's moment of inertia about he point x. That point which gives the minimum moment of inertia is the center of gravity of the district. The smaller the moment of inertia about the center of gravity the more compact is the district.

- **The Boyce-Clark Test.** Determine the center of gravity of the district and measure the distance from the center to the outside edges of the district along equally-spaced radial lines. Compare the percentage by which each radial distance differs from the average radial distance, and find the average of the percentage deviations over all radials. The closer the result is to 0, the more compact is the district.

- **The Perimeter Test.** Find the sum of the perimeters of all the districts. The shorter the total perimeter, the more compact is the districting plan.

— there are no discontiguities; furthermore, every district is then convex. In the case of simple states, the splitline algorithm works well — perhaps even better than the **K**-means algorithm. Its intuitive simplicity is also likely to make shortest splitline more appealing to the public.

Both **K**-means and splitline are deterministic: that is, when each algorithm is applied to a fixed problem, and all parameters are constant, the final result is unique. Some authors have expressed the opinion that any good districting algorithm is deterministic [Hay96]. There is one human element involved in the **K**-means algorithm: the choice of seeds is made, in some sense, subjectively, by the person implementing the algorithm. This factor could be completely eliminated by randomly picking the seeds, but this is not the most desirable solution. Random seeds can produce solutions far from the global optimum of the optimization function, and require many more iterations to get an answer within a given tolerance level. The natural choice is to use the approximate centers of existing districts as seeds. At first, this may seem contrary to our goal of reversing the effects of gerrymandering. A closer analysis of gerrymandering shows this is not true. Gerrymandering relies on intricately carving districts based on data that is invisible to our algorithm — say, distribution of people of various ethnicities, or distribution of people of varying income level or political affiliation. Correlation of district centers also has the benfit of easing the transition phase.

As of now, the **K**-means algorithm clearly performs better on more complex data sets. The splitline algorithm works well on simple data sets, but struggles with more complex sets — for instance, the state of New York. We maintain that the splitline algorithm should not be abandoned, but our final recommendation is that **The K-means algorithm quickly and deterministicly produces systems of districting that satisfy all requirements on simplicity and fairness, and applying this algorithm would produce a drastic improvement over the current district plans of any state in the United States.**

A Various Definitions of Compactness

The following definitions of compactness are said in [You88] to be representative of those definitions favored in past and present scholarship.

As discussed in [You88], the following definitions of compactness are often used or cited in the literature.

- **The Visual Test.** A district is more compact if it appears to be more compact.
- **The Roeck Test.** Find the smallest circle containing the district and take the ratio of the district's area to that of the circle. This ratio is always between 0 and 1; the closer it is to 1, the more compact is the district.
- **The Schwartzberg Test.** Construct the adjusted perimeter of the district by connecting by straight lines those points on the district boundary where three or more constituent units (*i.e.*, census tracts) from any district meet. Divide the length of the adjusted perimeter by the perimeter of a circle with area equal to that of the district.
- **Length-width Test.** Find a rectangle enclosing the district and touching it on all four sides, such that the ratio of length to width is a maximum. The closer the ratio is to 1, the more compact is the district.
- **Taylor's Test.** Construct the adjusted perimeter of the district by connecting by straight lines those points on the district boundary where three or more constituent units (*i.e.*, census tracts) from any district meet. At each such point the angle formed is "reflexive" if it bends away from the district and "non-reflexive" otherwise. Subtract the number of reflexive from the number of non-reflexive angles and divide by the total number of angles. The resulting number is always between 0 and 1; the closer to 1, the more compact the district.
- **The Moment of Inertia Test.** Locate the geographical center c_i of each census tract i in the district. Select an arbitrary point x and calculate the square of the distance from x to c_i, multiplied by the population of tract i. The sum of these numbers is the district's moment of inertia about he point x. That point which gives the minimum moment of inertia is the center of gravity of the district. The smaller the moment of inertia about the center of gravity the more compact is the district.

- **The Boyce-Clark Test.** Determine the center of gravity of the district and measure the distance from the center to the outside edges of the district along equally-spaced radial lines. Compare the percentage by which each radial distance differs from the average radial distance, and find the average of the percentage deviations over all radials. The closer the result is to 0, the more compact is the district.
- **The Perimeter Test.** Find the sum of the perimeters of all the districts. The shorter the total perimeter, the more compact is the districting plan.

B Numerical Results

In this section we give a more detailed overview of our numerical results in districting New York state, along with the results of applying the **K**-means algorithm to Texas and Colorado. The purpose in applying our algorithm to these states is to give some semblance of proof that the **K**-means algorithm will apply to essentially every state in the United States.

B.1 Compactness quotient results

Note that in the long table of compactness quotients, the column on the left is used only to enumerate these quotients. Quotients in the same row should not be compared in a pairwise fashion.

	N	mean	std	range
K-means	29	.658	.177	.775
Splitline	29	.480	.167	.695

District No.	**K**-means	Splitline
1	0.603467	0.20358
2	0.912486	0.609967
3	0.71772	0.650431
4	0.767607	0.520665
5	0.816948	0.436889
6	0.699194	0.447266
7	0.720281	0.517044
8	0.741351	0.586529
9	0.880247	0.371716
10	0.744034	0.096317
11	0.724836	0.44827
12	0.651049	0.528087
13	0.428957	0.359168
14	0.811315	0.35745
15	0.737009	0.709886

District No.	**K**-means	Splitline
16	0.646914	0.592954
17	0.672443	0.552759
18	0.672738	0.755555
19	0.695763	0.511479
20	0.807295	0.605663
21	0.695065	0.641265
22	0.819779	0.424846
23	0.39237	0.611651
24	0.705848	0.610452
25	0.441429	0.434587
26	0.452181	0.636516
27	0.682806	0.251353
28	0.306814	0.27772
29	0.13779	0.156226

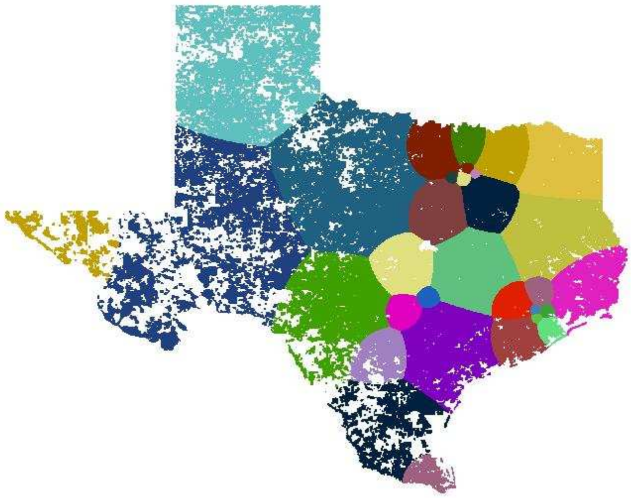

Figure 7: A proposed redistricting of Texas, using the **K**-means algorithm.

B.2 Population distribution error

The population distribution error is measured as the sum of the squares of

$$1 - \text{(current cluster population)} / \text{(target cluster population)}$$

divided by the number of clusters.

	K-means	Splitline
New York (29 districts)	1.799e-004	8.5032e-005
Colorado (7 districts)	5.6784e-006	1.7361e-004
Texas (32 districts)	4.4160e-004	-

Figure 8: Two proposed redistrictings of Colorado, using the **K**-means algorithm and the splitline algorithm, from left to right.

References

[AC84] A. A. Afifi and Virginia Clark. *Computer-aided Multivariate Analysis.* Lifetime Learning Publications, Belmont, CA, 1984.

[And73] Michael R. Anderberg. *Cluster Analysis for Applications.* Academic Press, New York, NY, 1973.

[AS68] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* U.S. Gov't Printing Office, Washington, D.C., 1968.

[Bac86] Charles H. Backstrom. The supreme court prohibits gerrymandering: A gain or a loss for the states? *The State of American Federalism*, 17(3), 1986.

[Bou88] Paul Bourke. Calculating the area and centroid of a polygon. 1988. Accessed on February 11, 2007.

[fIESIN] Center for International Earth Science Information Network. Socioeconomic data and applications center. Located at http://sedac.ciesin.columbia.edu/. Accessed on February 8, 2007.

[Fol02] Gerald B. Folland. *Advanced Calculus.* Prentice Hall, Upper Saddle River, NJ, 2002.

[GN70] R. S. Garfinkel and G. L. Nemhauser. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8), 1970.

[Hay96] Brian Hayes. Machine politics. *American Scientist*, 84(6), 1996.

[Smi] Warren B. Smith. Examples of our unbiased district-drawing algorithm in action. Located at http://rangevoting.org/RangeVoting.html. Accessed on February 12, 2007.

[Too03] Jeffrey Toobin. The great election garb. *The New Yorker*, 2003.

[You88] H. P. Young. Measuring the compactness of legislative districts. *Legislative Studies Quarterly*, 13(1), 1988.