Sam Burden, Aaron Dilley, and Lukas
University of Washington
2007 Outstanding Winner

Applying Voronoi Diagrams to the Redistricting Problem

May 10, 2007

Abstract

Gerrymandering is an issue plaguing legislative redistricting resulting from inadequate regulation. Here, we present a novel approach to the redistricting problem, an approach that uses a state's population distribution to draw the legislative boundaries. Our method utilizes *Voronoi* and population-weighted *Voronoiesque* diagrams, and was chosen for the simplicity of the generated partition: Voronoi regions are contiguous, compact, and simple to generate. We found regions drawn with Voronoiesque diagrams attained small population variance and relative geometric simplicity. As a concrete example, we applied our methods to partition New York state. Since New York must be divided into 29 legislative districts, each receives roughly 3.44 % of the population. Our Voronoiesque diagram method generated 29 regions with an average population of $(3.34 \pm 0.74)\%$. We discuss several refinements that could be made to the methods presented which may result in smaller population variation between regions while maintaining the simplicity of the regions and objectivity of the method. Finally, we provide a short statement that could be issued to the voters of New York state to explain our method and justify its fairness to them.

Contents

1 Introduction **4**

1.1 Current Models . . . 4

1.2 Developing Our Approach . . . 5

2 Notation and Definitions **5**

3 Theoretical Evaluation of our Model **6**

4 Method Description **7**

4.1 Voronoi Diagrams . . . 7

4.1.1 Useful Features of Voronoi Diagrams . . . 8

4.2 Voronoiesque Diagrams . . . 8

4.3 Determining Generator Points Using Population Density Distributions . . . 9

4.4 Procedure for Creating Regions using Voronoi and Voronoiesque Diagrams . 10

5 Redistricting in New York State **10**

5.1 Population Density Map . . . 11

5.2 Limitations of the Image-Based Density Map . . . 11

5.3 Selecting Generator Points . . . 11

5.4 Applying Voronoi Diagrams to NY . . . 14

5.5 Applying Voronoiesque Diagrams to NY . . . 14

5.6 Precisely Defining Boundary Lines . . . 17

6 Analysis **17**

6.1 New York State Results . . . 17

6.2 General Results . . . 18

7 Improving the Method **18**

7.1 Boundary Refinement . . . 18

7.2 Geographic Obstacles . . . 19

8 Bulletin to the Voters of the State of New York **19**

9 Conclusion **20**

List of Figures

1 Illustration of Voronoi diagram generated with Euclidean metric. Note the compactness and simplicity of the regions. . . . 7

2 Illustration of the process of ‘growing’ a Voronoiesque diagram with respect to a population density. Only three three generator points are used. Figures from left to right iterate with time. . . . 9

3 Illustration of creating divisions by first subdividing the map. Left: Population density distribution of hypothetical map with five desired districts. Middle: A subdivision of the map into two regions generated from two unshown generator points. Right: Final division of each subregion from the middle figure into desired final divisions. 10

4 New York State population density map. Data obtained from a 792-by-660 pixel raster image; color and height indicate the relative population density at each point. 12

5 Depiction of the implamentation of Voronoi diagrams with the Manhattan metric in the three step process of: assigning degeneracies to generator points, using the degenerate points to generate regions using the Voronoi diagram method, and creating subregions of the regions generated by degenerate points. Only the last two steps are depicted. The process for Voronoiesque diagrams is the same. (Dots in each region represent generator point locations.) 13

6 Voronoi diagrams generated with three distance metrics before subdivision of densely populated regions. (Dots in each region represent generator point locations.) 15

7 Districts created by the Voronoiesque diagram for New York state. Average population per region = $(3.34 \pm 0.74)\%$. (Dots in each region represent generator point locations.) 16

8 Illustration of Voronoi diagram generation which takes geographic obstacles into account. 19

1 Introduction

Defining Congressional districts has long been a source of controversy in the United States. Since the district-drawers are chosen by those currently in power, the boundaries are often created to influence future elections by grouping an unfavorable minority demographic with a favorable majority; this process is called *Gerrymandering.* It is common for districts to take on bizarre shapes, spanning slim sections of multiple cities and criss-crossing the countryside in a haphazard fashion. The only lawful restrictions on legislative boundaries stipulate that districts must contain equal populations, but the makeup of the districts is left entirely to the district-drawers.

In the United Kingdom and Canada, the districts are more compact and intuitive. Their success in mitigating Gerrymandering is attributed to having turned over the task of boundary-drawing to nonpartisan advisory panels. However, these independent commissions can take 2-3 years to finalize a new division plan, calling their effectiveness into question. It seems clear that the U.S. should establish similar unbiased commissions, yet make some effort to increase the efficiency of these groups. Accordingly, our goal is to develop a small toolbox that aids in the redistricting process. Specifically, we will create a model that draws legislative boundaries using simple geometric constructions.

1.1 Current Models

The majority of methods for creating districts fall into two categories: ones that depend on a current division arrangement (most commonly counties) and ones that do not depend on current divisions. Most fall into the former category. By using current divisions, the problem is reduced to grouping these divisions in a desirable way using a multitude of mathematical procedures. Mehrotra et.al. uses graph partitioning theory to cluster counties to total population variation of around 2% of the average district size [8]. Hess and Weaver use an iterative process to define population centroids, use integer programming to group counties into equally populated districts, and then reiterate the process until the centroids reach a limit [5]. Garfinkel and Nemhauser use iterative matrix operations to search for district combinations that are contiguous and compact [3]. Kaiser begins with the current districts and systematically swaps populations with adjacent districts [4]. All of these methods use counties as their divisions since they partition the state into a relatively small number of sections. This is necessary because most of the mathematical tools they use become slow and imprecise with many divisions. (This is the same as saying they become unusable in the limit when the state is divided into more continuous sections.) Thus using small divisions, like zip codes which on average are 5 times smaller than a county in New York, becomes impractical.

The other category of methods is less common. Out of all our researched papers and documentation, there were only two methods that did not depend on current state divisions. Forrest's method continually divides a state into halves while maintaining population equality until the required number of districts is satisfied [4, 5]. Hale, Ransom and Ramsey create pie-shaped wedges about population centers. This creates homogeneous districts which all contain portions of a large city, suburbs, and less populated areas [4]. These approaches are noted for being the least biased since their only consideration is population equality and do not use preexisting divisions. Also, they are straightforward to apply. However, they do not consider any other possibly important considerations for districts, such as: geographic freaures of the state or how well they encompass cities.

1.2 Developing Our Approach

Since our goal is to create new methods that add to the diversity of models available to a committee, we should focus on creating district boundaries independently of current divisions. Not only has this approach not been explored to its fullest, but it is not obvious why counties are a good beginning point for a model: Counties are created in the same arbitrary way as districts, so they might also contain biases, since counties are typically not much smaller than districts. Many of the division dependent models end up relaxing their boundaries from county lines in order to maintain equal populations, which makes the initial assumption of using county divisions useless, and also allows for gerrymandering if this relaxation method is not well regulated.

Treating the state as continuous (i.e. without preexisting divisions) does not lead to any specific type of approach. It gives us a lot of freedom, but at the same time we can impose more conditions. If the Forrest and Hale et.al. methods are any indication, we should focus on keeping cities within districts and introduce geographical considerations. (Note that these conditions do not have to be considered if we were to treat the problem discretely because current divisions, like counties, are probably dependent on prominent geographical features.)

Goal: Create a method for redistricting a state by treating the state continuously. We require the final districts to contain equal populations and be contiguous. Additionally, the districts should be as simple as possible (see §2 for a definition of simple) and optimally take into account important geographical features of the state.

2 Notation and Definitions

- **contiguous:** A set R is contiguous if it is pathwise-connected.
- **compactness:** We would like the definition of compactness to be intuitive. One way to look at compactness is the ratio of the area of a bounded region to the square of its perimeter. In other words

  $$C_R = \frac{A_R}{p_R^2} = \frac{1}{4\pi}\mathcal{Q}$$

  where C_R is the compactness of region R, A_R is the area, p_R is the perimeter and $\mathcal{Q}$ is the isoperimetric quotient. We do not explicitely use this equation, but we do keep this idea in mind when we evaluate our model.
- **simple:** Simple regions are compact and convex. Note that this describes a relative quality, so we can compare regions by their simplicity.

- **Voronoi diagrams:** a partition of the plane with respect to n nodes in the plane such that points in the plane are in the same region of a node if they are closer to that node than to any other point (for a detailed description, see §4.1)
- **generator point:** a node of a Voronoi diagram
- **degeneracy:** the number of districts represented by one generator point
- **Voronoiesque diagram:** a variation of the Voronoi diagram based on equal masses of the regions (see §4.2)
- **population center:** a region of high population density

3 Theoretical Evaluation of our Model

How we analyze our model's results is a tricky affair since there is disagreement in the redistricting literature on key issues. **Population equality** is the most well defined. By law, the populations within districts have to be the same to within a few percent of the average population per district. No specific percentage is given, but be assumed to be around 5%.

Creating a successful redistricting model also requires **contiguity**. In accordance with state law, districts need to be path-wise connected so that one part of a district cannot be on one side of the state and the other part on the other end of the state. This requirement is meant to maintain locality and community within districts. It does not, however, restrict islands districts from including islands if the island's population is below the required population level.

Finally, there is a desire for the districts to be, in one word, **simple**. There is little to no agreement on this characteristic, and the most common terminology for this is *compact*. Taylor defines simple as a measure of divergence from compactness due to indentation of the boundary and gives an equation relating the non-reflexive and reflexive interior angles of a region's boundary [9]. Young provides seven more measures of compactness. The *Roeck* test is a ratio of the area of the largest inscribable circle in a region to the area of that region. The *Schwartzberg* test takes ratio between the adjusted perimeter of a region to a the perimeter of a circle whose area is the same as the area of the region. The *moment of inertia* test measures relative compactness by comparing "moments of inertia" of different district arrangements. The *Boyce-Clark* test compares the difference between points on a district's boundary and the center of mass of that district, where zero deviation of these differences is most desirable. The *perimeter test* compares different district arrangements buy computing the total perimeter of each. Finally, there is the *visual* test. This test decides simplicity based on intuition [11].

Young notes that "a measure [of compactness] only indicates when a plan is more compact than another"[11]. Thus, *not only is there no consensus on how to analyze redistricting proposals, it is also difficult to compare them.*

Finally, we remark that the above list only constrains the shape of generated districts. We have not mentioned of any other potentially relevant feature. For instance, it does not consider how well populations are distributed or how well the new district boundaries conform with other boundaries, like counties or zip codes. Even with this short list, it is

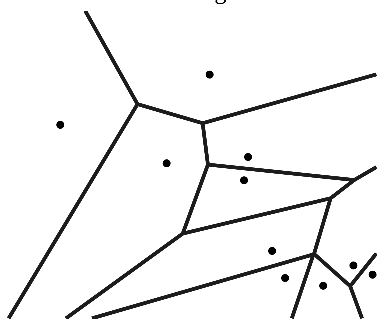


Figure 1: Illustration of Voronoi diagram generated with Euclidean metric. Note the compactness and simplicity of the regions.

clear that we are not in a position to define a rigorous definition of simplicity. What we can do, however, is identify features of our proposed districts which are simple and which are not. This is in line with our goal defined in sec. 1.2, since this list can be provided to a districting commission who decide how relevant those simple features are. **We do not explicitly define *simple*, we loosely evaluate simplicity based on overall contiguity, compactness, convexity, and intuitiveness of the model's districts.**

4 Method Description

Our approach depends heavily on using Voronoi diagrams. We begin with a definition, its features, and motivate its application to redistricting.

4.1 Voronoi Diagrams

A Voronoi diagram is a set of polygons, called Voronoi polygons, formed with respect to n generator points contained in the plane. Each generator p_i is contained within a Voronoi polygon $V(p_i)$ with the following property:

$$V(p_i) = \{q | d(p_i, q) \leq d(p_j, q),\ i \neq j\} \text{ where } d(x, y) \text{ is the distance from point } x \text{ to } y$$

That is, the set of all such q is the set of points closer to p_i than to any other p_j. Then the diagram is given by (see fig 1)

$$\mathbf{V} = \{V(p_1), \ldots, V(p_n)\}$$

Note that there is no assumption on the metric we use. Out of the many possible choices, we use the three most common:

- Euclidean Metric: $d(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$

- Manhattan Metric: $d(p, q) = |x_p - x_q| + |y_p - y_q|$
- Uniform Metric: $d(p, q) = \max\{|x_p - x_q|, |y_p - y_q|\}$

4.1.1 Useful Features of Voronoi Diagrams

Here is a summary of relevant properties:

- The Voronoi diagram for a given set of generator points is unique and produces polygons, which are path connected.
- The nearest generator point to p_i determines an edge of $V(p_i)$
- The polygonal lines of a Voronoi polygon do not intersect the generator points.
- When working in the Euclidean metric, all regions are convex.

These properties are important for our model. The first property tells us that regardless of how we choose our generator points we generate unique regions. This is good when considering the politics of Gerrymandering. The second property implies that each region is defined in terms of the surrounding generator points while in turn, each region is relatively compact. **These features of Voronoi diagrams effectively satisfy two out of the three criteria for partitioning a region: contiguity and simplicity**.

4.2 Voronoiesque Diagrams

The second method we use to create regions is a modification of the intuitive construction of Voronoi diagrams. The method does not fall under the definition of Voronoi diagrams, but since it is similar to Voronoi diagrams, we call them Voronoiesque diagrams. One way to visualize the construction of Voronoi diagrams is to imagine shapes (determined by the metric) that grow radially outward at a constant rate from each generator point. In the Euclidean metric these shapes are circles. In the Manhattan metric they are diamonds. In the Uniform metric, they are squares. The interior of these shapes form the regions of the diagram. As the regions intersect, they form the boundary lines for the regions. With this picture in mind, we define Voronoiesque diagrams to be the boundaries defined by the intersections of these growing shapes. The fundamental difference between Voronoi and Voronoiesque diagrams is that Voronoiesque diagrams grow the shapes radially outward at a constant rate like Voronoi diagrams. Their radial growth is defined with respect to some real function on a subset of $\mathbb{R}^2$ (representing the space on which the diagram is being generated). See fig.2

More rigorously, we define a Voronoi diagram to be the intersections of the $\mathcal{V}_i^{(t)}$'s, where $\mathcal{V}_i^{(t)}$ is the Voronoiesque region, or just 'region', generated by the generator point p_i at iteration t. With every iterations,

$$\mathcal{V}_i^{(t)} \subset \mathcal{V}_i^{(t+1)}$$

and

$$\int_{\mathcal{V}_i} f(x,y)dA = \int_{\mathcal{V}_j} f(x,y)dA$$

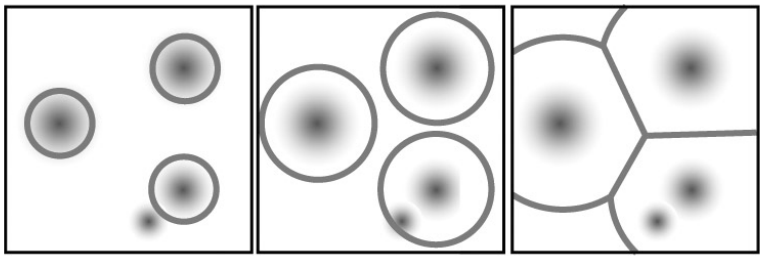


Figure 2: Illustration of the process of 'growing' a Voronoiesque diagram with respect to a population density. Only three three generator points are used. Figures from left to right iterate with time.

for all $\mathcal{V}_i$, $\mathcal{V}_j$ representing different regions. The manner in which the $\mathcal{V}_i^{(t)}$'s are grown radially from one iteration to the next is determined by the metric used.

What's useful about Voronoiesque diagrams is that their growth can be controlled by requiring that the area under the function f for each region is the same at every iteration. In our model, we take f to be the population distribution of the state. Thus the above equation is a statement of population equality. Also, when f is constant, the regions grow at a constant rate, so the resulting diagram is Voronoi.

The final consideration for using Voronoiesque diagrams is determining the location for generator points.

4.3 Determining Generator Points Using Population Density Distributions

For now, we have defined how to generate regions given a set of generator points. Here we consider how to define the generator points in order to create Voronoi and Vornoiesque diagrams. In the case of Voronoi diagrams, this is our only degree of freedom since generator points generate unique Voronoi regions. We found no well defined algorithm to do this, but instead came up with a procedure that functions decently.

Our first approach is to place generator points at the m largest set of peaks that are well distributed throughout the state, (where m is the required number of districts in that state). By choosing generator points in this way, we keep larger cities within the boundaries we will generate with Voronoi or Vornoiesque diagrams and we make sure the generator points are well dispersed throughout the state. One problem that arises is when cities are so large that in order for districts to hold the same amount of people, a city must be divided into districts. A perfect example is New York City, which contains enough people to hold 13 districts. Taking large cities into account takes extra consideration.

Our second approach is to choose the largest peaks in the population distribution and assign each peak with a weight. The weight for each generator point is the number of districts the population surrounding that peak needs to be divided into. We call this weight the degeneracy of the generator point. We begin assigning generator points to the highest populated cities with their corresponding degeneracies until the sum of all the generator points and their respective degeneracies is equal to m. In other words, until:

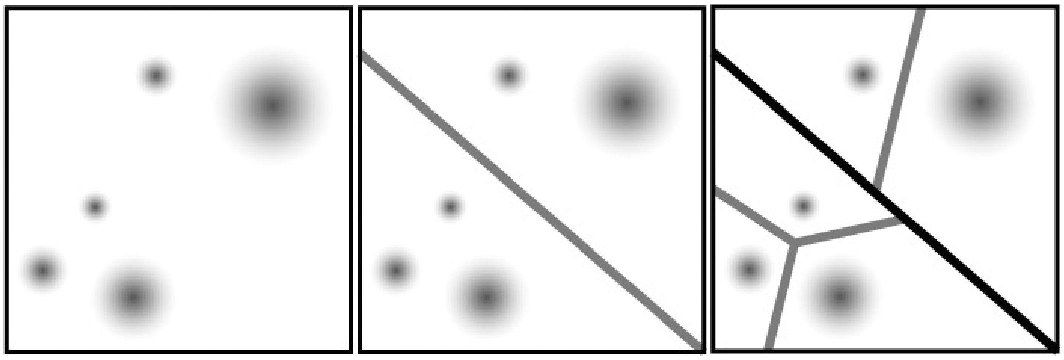


Figure 3: Illustration of creating divisions by first subdividing the map. Left: Population density distribution of hypothetical map with five desired districts. Middle: A subdivision of the map into two regions generated from two unshown generator points. Right: Final division of each subregion from the middle figure into desired final divisions.

$$\sum_{\text{all generator pts.}} \text{degeneracy of generator pt.} = m$$

As we will see when we apply our model to New York, this method works well. It should be noted, though, that this is not the only way to define the location of generator points, but it is a very good start.

4.4 Procedure for Creating Regions using Voronoi and Voronoiesque Diagrams

Once we have our generator points, we can generate our diagrams with two more steps: first generate the diagram using the given generator points. Within each generated region, called a subdivision, with some degeneracy r, create r new generator points within that subdivision by finding the r largest population density peaks and create another diagram. See fig.3

5 Redistricting in New York State

At this point, we have described a general procedure for generating political districts with Voronoi diagrams which seems effective. We now turn our attention to testing our models on New York.

- **has regions with large population density**,
- **has regions with constrained geography**,
- and **must be divided into many (29) regions**.

We begin by explaining our method for choosing generator points at population centers, since these points will uniquely determine a Voronoi diagram for the state. Then we describe several methods for generating Voronoi and Voronoiesque diagrams from these points and present the corresponding results. Finally we discuss how to use these diagrams to create actual political districts for New York state.

5.1 Population Density Map

To apply our Voronoi diagram methods to New York, we first obtain an approximate population density map of the state. The U.S. Census Bureau maintains a database [2] which contains census tract-level population statistics; when combined with boundary data [1] for each tract, it's possible to generate a density map with a resolution no coarser than 8,000 people per region [7]. Unfortunately, our limited experience with the Census Bureau's data format prevented us from accessing this data directly, and we contented ourselves with a 792-by-660 pixel approximation to the population density map [6].

We loaded this raster image into MATLAB and generated a surface plot where height represented population density at each point. To remove artifacts introduced by using a coarse lattice representation for finely-distributed data, we applied a 6-pixel moving average filter to the density map. The resulting population density is shown in fig. 4.

5.2 Limitations of the Image-Based Density Map

The population density image we used yielded a density value for every third of a square mile from the following set (measured in people per square mile):

$$\{0,\ 10,\ 25,\ 50,\ 100,\ 250,\ 500,\ 1000,\ 2500,\ 5000\}\,.$$

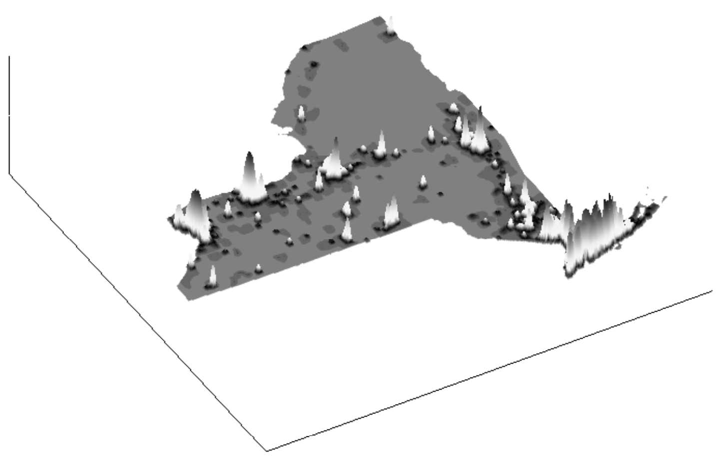
This provides a decent approximation for regions with a density smaller than $5,000\ \mathit{people/sq.mi.}$ However, since New York City's average population density is $26,403\ \mathit{people/sq.mi.}$ [10], the approximation will break down at large population centers.

5.3 Selecting Generator Points

Our criteria for redistricting the state stipulates that the regions we generate must contain equal populations. New York state must be divided into 29 congressional districts to support its share of representatives, so each region must contain $\approx 3.45\%$ of the state's population. Since a state's population is concentrated primarily in a small number of cities, we use local maxima of the population density map as candidates for generator points.

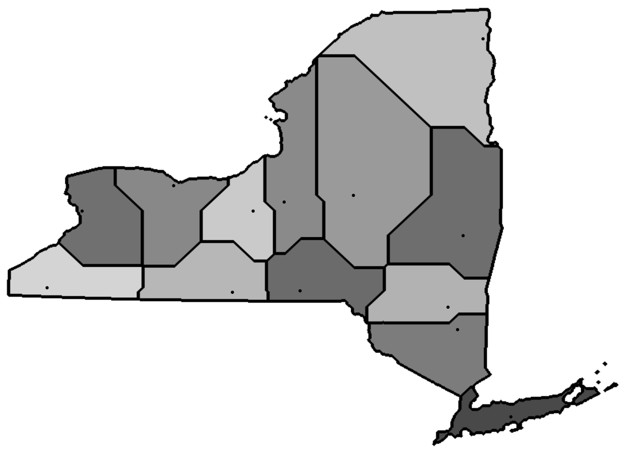
If we were to simply choose the highest 29 peaks from the population density map as our set of generator points, the resulting set would be contained entirely in the largest population centers and would not be well distributed evenly over state. For the largest population centers, we assign a single generator point with a degeneracy as described in 4.3. After all the generator points have been assigned, we generate a Voronoi diagram for the state. Then, we return to the regions with degenerate generator points and repeat the process of finding generator points for that region and generate a Voronoi diagram from them. See fig. 3 for an illustration of the decomposition before and after subdivision.

(a) Top View: White areas represent high population density over New York

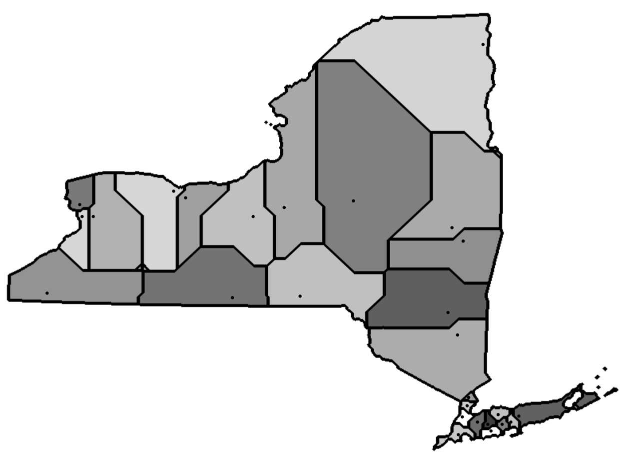


(b) Angled View: Clearer view of population distibution over New York

Figure 4: New York State population density map. Data obtained from a 792-by-660 pixel raster image; color and height indicate the relative population density at each point.



(a) Regions created using the Manhattan metric before subdivisions are implemented.



(b) Regions created using the Manhattan metric after subdivisions are implemented. Subdivisions are created in New York City, Buffalo, Rochester, and Albany

Figure 5: Depiction of the implamentation of Voronoi diagrams with the Manhattan metric in the three step process of: assigning degeneracies to generator points, using the degenerate points to generate regions using the Voronoi diagram method, and creating subregions of the regions generated by degenerate points. Only the last two steps are depicted. The process for Voronoiesque diagrams is the same. (Dots in each region represent generator point locations.)

Based on our density data for New York state, we subdivide the region around New York City into 12 subregions, Buffalo into 3 subregions, and Rochester and Albany into 2 subregions. Note that this roughly corresponds to the current allocation, where New York City receives 14 districts, Buffalo gets 3, and Rochester and Albany both get roughly 2. Here, New York City's population is underestimated since the average density there far exceeds our data's density range. With a more detailed data set, our method would have called for the correct number of subdivisions.

5.4 Applying Voronoi Diagrams to NY

The simplest method we consider for generating congressional districts is to simply generate the discrete Voronoi diagram from a set of generator points. We achieve this by iteratively 'growing' regions outward with the function f constant. That way the regions grow at a constant rate, and hence the resulting diagram is voronoi. A region's growth is limited at each step by its radius in a certain metric; we considered the Euclidean, Manhattan, and uniform metrics. Once the initial diagram has been created, a new set of generator points for dense regions are chosen and those regions are subdivided using the same method. Unrefined decompositions can be seen in fig. 6.

Each metric produces a relatively simple decomposition of the state, though the Manhattan metric has simpler boundaries and yields a slightly smaller population variance between regions.

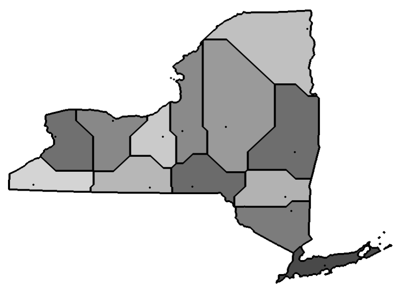
5.5 Applying Voronoiesque Diagrams to NY

Though our simple Voronoi diagrams produced simple regions with a population mean near the desired value, the population variance between regions is enormous. In this sense, the simple Voronoi decomposition doesn't meet one of the main parts of our redistricting goal. However, the Voronoi regions are so simple that we prefer to augment this method with population weights rather than abandon it entirely.

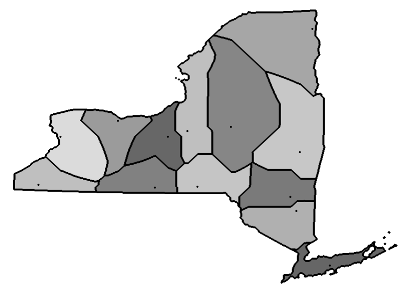
Fig. 7 shows the result of this decomposition, along with exploded views of the two regions which were subdivided more than twice in the refinement stage of the diagram generation. The population contained in each region is summarized in table 1.

Region #	Population %	Region #	Population %	Region #	Population %
1	3.03	2	3.02	3	6.15
4	3.51	5	3.43	6	3.45
7	3.43	8	3.22	9	3.50
10	3.19	11	2.78	12	3.33
13	3.21	14	3.00	15	3.43
16	3.38	17	4.43	18	4.76
19	3.24	20	3.12	21	2.97
22	3.17	23	3.17	24	3.21
25	2.44	26	2.12	27	3.66
28	3.79	29	3.71		

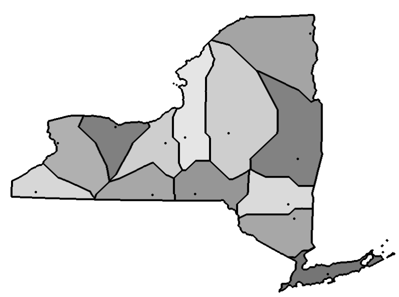
Table 1: Population Fraction in each Legislative District



(a) Regions created using the Manhattan metric before subdivisions. Average Population $= (3.5 \pm 2.2)\%$.

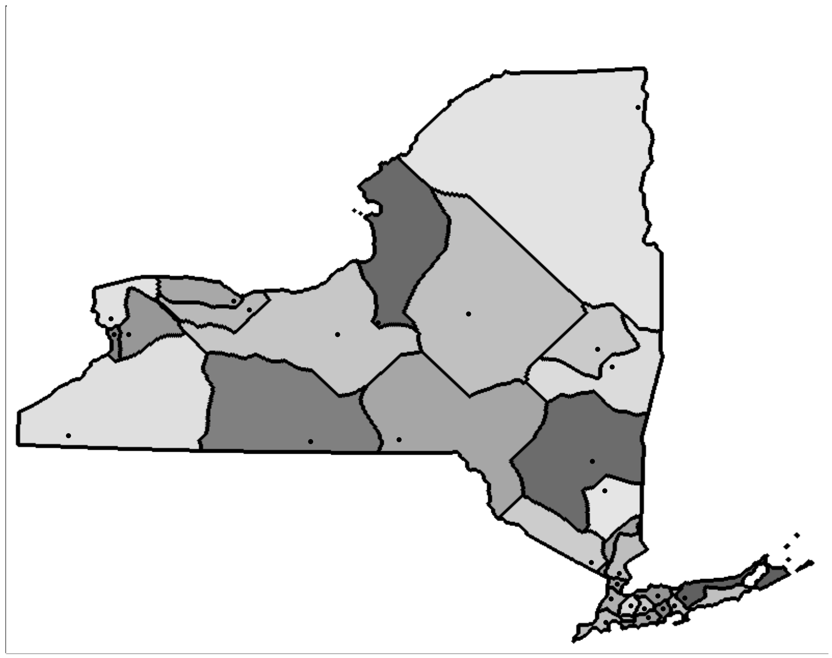


(b) Regions created using the Euclidean metric before subdivisions. Average Population $= (3.7 \pm 2.6)\%$.

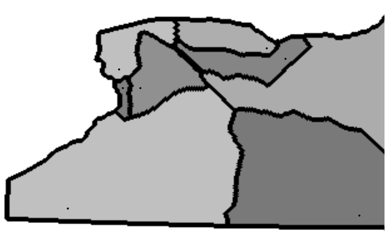


(c) Regions created using the Unifrom metric before subdivisions. Average Population $= (3.7 \pm 2.6)\%$.

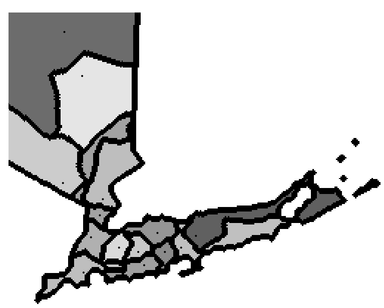
Figure 6: Voronoi diagrams generated with three distance metrics before subdivision of densely populated regions. (Dots in each region represent generator point locations.)



(a) Overall New York Voronoiesque regions



(b) Exploded view of regions around Buffalo.



(c) Exploded view of regions around Long Island.

Figure 7: Districts created by the Voronoiesque diagram for New York state. Average population per region = $(3.34 \pm 0.74)\%$. (Dots in each region represent generator point locations.)

5.6 Precisely Defining Boundary Lines

It is not satisfactory to say the regions created by our models should define the final boundary locations. In the least, boundaries should be tweaked so that they don't accidentaly divide houses into two districts. However, given the scale at which the Voronoi and Voronoiesque diagrams were drawn, it seems reasonable to assume that their boundaries could be modified to trace existing boundaries—like county lines, ZIP codes, or city streets—without changing their general shape or average population appreciably. As an example, the average area of a ZIP code in New York state is ≈ 10 *sq.mi.* and roughly 200 city blocks per square mile in Manhattan, while the minimum size of one of our Voronoi regions is 73 *sq.mi.* and the average size is $\approx 2,000$ *sq.mi.*. Therefore it seems reasonable that we could approximate the boundaries of our Voronoi and/or Voronoiesque diagrams by preexisting boundaries.

6 Analysis

6.1 New York State Results

We turn now to a discussion of how well our results from the previous section meet our original specification for redistricting. In terms of simplicity of generated districts, our Voronoi diagram method is a clear winner, particularly when applied with the Manhattan metric: the generated regions are contiguous and compact while their boundaries, being unions of line segments, are about the simplest that could be expected. However, this method falls short in achieving equal population distribution among the regions, since the variance in the average population per region is on the order of the average population itself.

As may be expected in any sort of high-dimensional optimization problem, there is an essential tradeoff in this problem between the simplicity of the legislative districts and their respective populations. Accordingly, when we modify the Voronoi diagram method to generate population-weighted Voronoiesque regions, we cut the population variance by a factor of four—from $\pm 2.8\%$ to $\pm 0.7\%$—while suffering a small loss in the simplicity of the resulting regions. In particular, regions in the Voronoiesque diagrams appear to be less compact and their boundaries are more complicated than their Voronoi diagram counterparts, though contiguity is still maintained.

Finally, we noted in the previous section that any actual implementation of a diagram generated from either of our methods would have to make small, localized modifications to ensure the district boundaries make sense from a practical perspective. Though this would appear to open the door for the same sort of politically-biased district manipulations our methods were aiming to avoid in the first place, we think the size of the necessary deviations (on the order of miles) is small enough when compared to the size of a Voronoi or Voronoiesque region (on the order of tens or hundreds of miles) to make the net effect of these variations insignificant. Therefore, though we have provided only a first-order approximation to the congressional districts, we have left little room for Gerrymandering to occur.

6.2 General Results

We already know how well our results worked for New York. How effective is our method in general? We examine the results for an arbitrary state including worst case scenarios for each criteria.

Population Equality

The largest problem with this requirement occurs when we try to make regions too simple. Typically, our Voronoi method has the most room for error here. If a state has a series of high population density peaks with a relatively uniform decrease in population density extending away from each peak, then the regions will differ quite a bit. This is because in this situation, ratios of populations are then roughly equal to the ratios of areas between regions. However, our final method focuses primarily on population so equality is much easier to regulate here.

Contiguity

Contiguity problems arise often if the state itself has little compactness, like Florida, or if the state has some sort of sound like Washington. The first two methods focus more on population density without really acknowledging the boundaries of the state itself. So it's possible for one region to be separated by some geographic obstruction like a body of water or a mountain range. Again the final method fixes this by growing in increments, this allows for state boundaries to be defined. Then regions wouldn't grow over but around specified obstacles.

Compactness

Unfortunately, the final method doesn't do everything, it is the least likely candidate for generating compact regions. The first two are most successful in this area. The first method creates all convex regions. Though the second can't guarantee convexity, its form is similar in shape and size to the first. Furthermore, one nice property of the generated regions from the first method is that there is a way to make slight adjustments to the boundaries while still maintaining convexity (see § 7.1) This is good for taking population shifts across districts into account between redistricting periods.

7 Improving the Method

Now that the problem areas have been defined, we offer some ways to reduce the effect of these problems.

7.1 Boundary Refinement

Consider the Voronoi diagram method. We know this approach is good at generating polygonal districts but not as successful at maintaining population equality. One such method that helps is vertex repositioning. Notice that adjacent districts generated by this method all share a vertex common to at least three boundaries. From this vertex extends

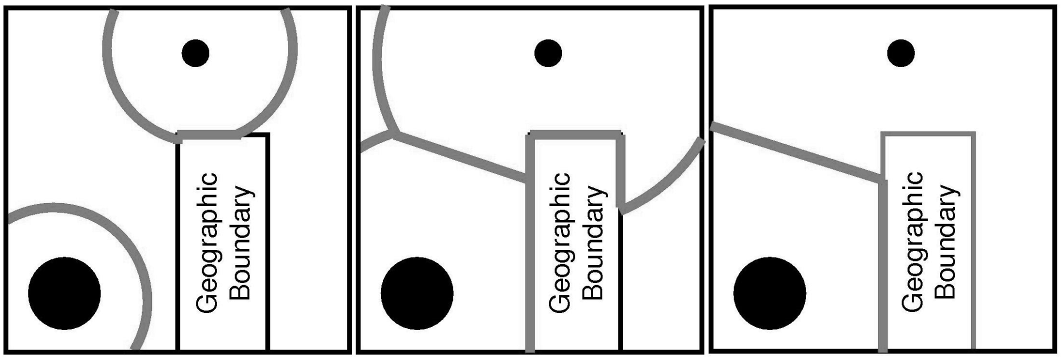


Figure 8: Illustration of Voronoi diagram generation which takes geographic obstacles into account.

a finite number of line segments that partially define the boundaries of these adjacent regions. Connecting the endpoints of these segments yields a polygon. Now we are free to move the common vertex anywhere in the interior of this polygon while still maintaining convexity. With this we can redraw boundaries between regions that are significantly different in population size and in doing so help equalize each of the regions.
There are also ways to adjust population inequality in the Voronoiesque method. Looking at the region with the lowest population, systematically increase the area of the low-population regions while decreasing the area of the neighboring high-population regions.

7.2 Geographic Obstacles

Our method doesn't implement geographic areas such as rivers, mountains, canyons, and other prominent features. The Voronoiesque method, however, has the potential to impliment these features. The same algorith that detects intersections between vornoiesque regions can detect a defined geographic boundary and stop growing in that direction. An illustration of this idea is shown in fig. 8.These geographic obstacles would be chosen by the redistricting committee.

8 Bulletin to the Voters of the State of New York

READ ON FOR IMPORTANT INFORMATION REGARDING YOUR REPRESENTATIVE GOVERNMENT

Authorities within your state's government recently realized that during reapportionment—the process by which your state's number of congressional representatives changes—the incumbent political leaders tend to *Gerrymander* the boundaries of congressional districts, redrawing them to influence future elections in their favor. As this can undermine equal representation for all citizens, the State of New York commissioned an interdisciplinary team of mathematicians and engineers to create an objective procedure for redistricting that can be applied in the future to prevent partisan influence over congressional district boundaries.

The team came to the conclusion that to be fair to all, congressional districts should:

- **be connected**,
- **contain equal populations**,
- **be as compact as possible**, and
- **not unnecessarily subdivide large cities**.

Accordingly, they created a simple method for generating districts that meet these criteria.

The method is based on a geometrical structure known as a *Voronoi diagram*, which describes a partition of your state into compact, connected regions generated from a set of initial points; see figure 1 for an example. Since the regions are supposed to envelop equal populations, the initial points were chosen at major population centers (like New York City, Buffalo, Rochester, and Albany, among others). The regions are then 'grown' out from these population centers as in figure 2 until the entire state is covered.

To ensure the districts end up with roughly equal populations, a regions' growth is limited by the population contained within it. This results in a final diagram which has connected, compact regions with small population variation. In other words, diagrams generated with this method fulfill the guidelines for creating fair legislative districts.

The new district diagram is illustrated in figure 7. This diagram is composed of 29 distinct congressional districts, each of which contains close to 3.4% of the total state population. But more important than the precise population contained in each region is the fact that the districts were generated objectively by a computerized method, so *partisan politics play no role in the result.* This ensures that the next time boundary lines are drawn, they will provide an impartial partition of our state's population, with no room for Gerrymandering.

9 Conclusion

There are many methods in existence for drawing district boundaries. Most of these models are successful in what is sets out to do. However, many of them depend on current state divisions as a starting points for creating districts. Our model differs in that we only require the use of a state's population distribution and as an option can incorporate county, property, and geographic considerations.

Our Voronoiesque model satisfies our proposed goal. We supply a model for a redistricting committee to generate district boundaries that are simple, contiguous, and produce districts with equal populations. In particular, we found that Voronoiesque diagrams redistrict New York very well. What's particularly attractive about all the methods is that generating the districts is intuitive and accessible to the general public and also the computer generation process takes less than 10 seconds to complete.

References

[1] U.S. Census Bureau. 2005 first edition tiger/line data, Feb. 2007.

[2] U.S. Census Bureau. Summary 1 file: Census 2000, Feb. 2007.

[3] Garfinkel, R. S. and Nemhauser, G. L. Optimal political districting by implicit enumeration techniques. *Management Science*, 16(8):B495–B508, apr 1970.

[4] Howard D. Hamilton. *Reapportioning Legislatures*. Charles E. Merril Books, Inc., 1966.

[5] Hess, S. W., Weaver, J. B., Siegfeldt, H. J., Whelan, J. N., and Zitlau, P. A. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998–1006, nov 1965.

[6] Jim Irwin. Wikipedia.org: New york population map, Feb. 2006.

[7] Nielsen Media. Glossary of media terms, Feb. 2007.

[8] Mehrotra, Anuj, Johnson, Ellis L., and Nemhauser, George L. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, aug 1998.

[9] Peter J. Taylor. A new shape measure for evaluating electoral district patterns. *The American Political Science Review*, 67(3):947–950, sep 1973.

[10] Wikipedia.org. New york city, Feb. 2007.

[11] H. P. Young. Measuring the compactness of legislative districts. *Legislative Studies Quarterly*, 13(1):105–115, feb 1988.