Gerrymandering and Graph Partitioning: Approaches to Population Division Models

Chau Pham
Department of Computer Science, Earlham College
801 National Road West, Richmond, IN, USA
cqpham15@earlham.edu

ABSTRACT

Gerrymandering is the practice where voting district lines in a state are drawn to benefit one party over another. To combat this practice, computer-generated district plans can be utilized. One way to generate district plans is to treat the geographical area like a graph and model its population attributes, then divide up the area according to these attributes. In this project, a graph partitioning algorithm is created to generate compact electoral districts in several states. The algorithm generates several district plans, then finds the most compact ones by evaluating several attributes measuring how spread out the population is in each district. The best district schemes are then visualized using a mapping tool. An evaluation tool calculates the same attributes for existing district plans to quantitatively compare them to the generated plans. It is found that the generated plans have more compact districts according to the comparison of these values. A graphical analysis of the visualized generated district plans compared to current district plans also shows that the algorithm generated more contiguous districts.

CCS CONCEPTS

• Information systems → Data analytics; **• Theory of computation →** *Design and analysis of algorithms*;

KEYWORDS

Gerrymandering, population modeling, graph partitioning, compactness, redistricting

1 INTRODUCTION

In the American electoral system, a voting district is a geographical delineation used in the process of electing members of legislative bodies. Existing administrative boundaries cannot be used as-is because of divergent ratios of voters to representatives in major towns and cities compared to more sparsely populated rural areas. For example, a voter living in a metropolitan area with a population of 1,000,000 people would have a vote that is 100 times weaker than a voter in a town of 100,000 people in a first-past-the-post, winner-takes-all election.[1]

Electoral district lines are meant to create proportional numbers of voters across a region. The process of creating district lines differs from state to state, but is commonly done by the state legislature. Since the state legislature is controlled by a majority party, drawing district lines is thus a partisan process. This inevitably leads to bias, both subconscious and deliberate, in the creation of voting districts.

Table 1 lists the federal and state criteria for a valid district plan.

Equal population	Legislative districts within a state must have equal population
Contiguity	All parts of the district physically adjacent to each other
Administrative boundaries	Units like counties be kept together whenever possible
Compactness	Contortion of boundaries and spread from a central core
Communities of interest	Shared social, cultural, racial, ethnic, and economic interests

Table 1: Redistricting Criteria [7]

The manipulation of voting districts to profit one political party over another is commonly known as gerrymandering. In 1812, Elbridge Gerry, then-Governor of Massachusetts, signed into law a bill that redrew the district lines of Massachusetts to benefit his party. The *Boston Gazette* coined the term *gerrymander*, referring to Governor Gerry and the shape of one of the newly created districts, which resembled a salamander [12]. The point of gerrymandering is to waste as many of the opponent's votes as possible. Wasted votes are votes which do not contribute to a candidate's win; for instance, any votes above fifty percent in a simple-majority two-candidate race. If one party's votes are concentrated in a few districts and they win those districts by a large margin with lots of wasted votes, their opponent can actually win the rest of the districts and get a majority.

A solution to gerrymandered districting is computer-generated district lines. Drawing district lines is analogous to a graph partitioning problem, where geographical areas are represented in the form of a graph G = (V,E), with V vertices and E edges. The solutions to the problem would be partition schemes which divide G into smaller sections that satisfy specific criteria. In the case of voting districts in the United States, some of these criteria are equal population, contiguity, compactness, following existing administrative boundaries, and preserving communities of interest. Solving this partitioning problem computationally can help eliminate bias from the districting process, while also reducing the human effort needed to create these complicated divisions. Furthermore, with these computer-generated districting plans, we can even train a system to spot gerrymandering in an existing plan, and thus prevent politicians from manipulating district lines to benefit their own party. The existing research on gerrymandering and population division models includes several proposals for different methods and algorithms to identify extreme districting plans, including those that indicate gerrymandering. These include a computational approach which compares one districting plan to a large set of plans in order to quantify its characteristics [7]. Many of these algorithms rely on treating the population division problem like a graph partitioning problem [2] [17].

Since the goal of dividing a geographical area into electoral districts is to group together people with common interests and stakes

[1] http://aceproject.org/main/english/es/esd01.htm

in their community, this project focuses on solving the problem of creating districts where voters would be equally spread out geographically. The graph partitioning algorithm focuses on dividing population equally amongst all the districts, by iteratively changing the groupings of population blocks and calculating how spread out the voters in each district are and how high the standard deviation of all the district populations is.

The major novel contribution of this project is the application of graph partitioning algorithms in creating electoral districts, which allows for a large number of district plans to be created and evaluated in a short amount of time. It also eliminates the potential for partisan bias in the districting process. Another contribution is the development of a visualization tool for the generated electoral districts, and the evaluation of the compactness of the generated districting plans compared to the existing districting plans for two states: Pennsylvania and Connecticut.

This paper covers related research on gerrymandering strategies, computationally generated districting plans, and graph partitioning and its application in population modeling. It describes the design and implementation of the graph partitioning algorithm and the visualization tool. Lastly, the resulting districting plans generated by the algorithm will be analyzed and evaluated in comparison to existing districting plans, both visually and quantitatively.

2 RELATED WORK

In order to create a tool which generates fair and non-gerrymandered districts, first we need to understand the common strategies and indicators of gerrymandering. By examining the different ways gerrymandered districts can be identified and how they can affect elections, we can address these issues in the districting tool built in this project. Secondly, an examination of existing tools for computationally generating districting plans is crucial in the process of developing another tool. By comparing and evaluating the different criteria and approaches that other researchers have implemented to generate districting plans, we can examine the pros and cons of the graph partitioning approach implemented in this project. Lastly, relevant work on different approaches to graph partitioning provides a mathematical examination of what graph partitioning is and how it can be applied to population division in a geographical area, which is the core of creating districting plans.

2.1 Gerrymandering Strategies

There are many different ways that a district plan can be gerrymandered. Understanding these strategies is a crucial step in finding ways to create unbiased district plans.

The perverse-effects claim is the common notion that majority-minority legislative districting helps Republican candidates. By analyzing district lines drawn by both Democrats and Republicans, and voting records, Shotts found that geographical constraints (i.e., compactness) or supermajority-minority mandates (requiring that significantly more than half of a district population be minority groups members) can waste Democratic votes [20]. This is because many minority groups typically vote Democrat, based on survey data. A model of optimal partisan gerrymandering was developed, and it is also flexible enough to incorporate factors such as a third type of voters, majority-minority federal mandates, and a new technique for analyzing geographical and informational constraints. Using this model, the effects of each of these factors on wasting Democratic votes in a district were analyzed. It was discovered that when Democrats control redistricting, geographical constraints or supermajority-minority mandates can force them to create some districts where white Democrats are grouped together with Democrat minority members, thus wasting their votes. However, this doesn't happen when Republicans control redistricting or with a bare majority-minority mandate. The perverse-effects were also found to be asymmetrical, as it only decreases the number of elected Democrats and not Republicans.

Similarly, Chen and Rodden found that geographical distribution of parties' supporters, in this case Democratic voters in urban areas, can create a lot of wasted votes in many districts while also showing a bias favoring Republicans who have more geographically scattered voters [6].

One redistricting criterion, compactness, is rarely defined clearly in legal terms, but it usually refers to how contorted a district's boundaries are, or how spread out it is from a central core. To put it simply, if a district is shaped like a regular geometric shape and its constituents live near each other, the district is likely compact. Geography compactness requirements try to combat gerrymandering, but Altman showed that it makes little difference unless the compactness is very extreme [1]. Furthermore, compactness doesn't take into account the fact that neighborhoods are still segregated and a majority party is much likely to have more geographically diffused supporters compared to minority parties; in which case compactness requirements do more harm than good.

Building on the concept of wasted votes, Stephanopoulos and McGhee proposed the concept of the efficiency gap: the ratio between a party's wasted votes and the number of total votes cast [22]. The bigger the efficiency gap, the more likely that a districting plan is gerrymandered. The researchers computed the efficiency gap in every congressional and state house plans from 1972 to 2012, and found that the typical plan was fairly balanced over this period as a whole, but in recent years the pro-Republican gaps grew larger and would continue to do so according to sensitivity testing. Lastly, this paper proposed setting a threshold for efficiency gaps in districting plans, which aimed to prevent gerrymandering districting schemes.

In a 2008 report, Wall discussed different models of voting districts that take into account opinion dynamics, community structure, and geographical distribution, all of which play major roles in whether or not a state can be easily gerrymandered [23]. By studying networks of users on the social media site Facebook, the researchers created a model of how groups of people with similar opinions and interests cluster together (i.e. neighborhoods, friends, etc.) and how these groups can be utilized in a gerrymandering districting scheme.

By studying the phenomena of gerrymandering from the perspective of a hypothetical gerrymanderer, Friedman and Holden concluded that the commonly employed strategy of throwing away unwinnable districts to waste the opponent's votes while concentrating on winnable districts by grouping as many supporters together is actually not an optimal partisan gerrymandering scheme [10]. They found that grouping Democrats and Republicans from the two extremes of the spectrum together and the moderate ones together actually allow for more opportunities of neutralizing your opponents.

On a different note, Puppe and Tasnádi studied the concept of an unbiased districting plan. It was shown via examples that given simple geographical and population constraints, an unbiased districting plan might not exist for every area. Furthermore, a proof was presented which showed that determining whether or not an unbiased districting plan exists for a given geography was an NP-complete problem [15]. If finding an unbiased districting plan is an NP-complete problem, then what would an approach for finding reasonable and valid districting plans have to compromise in so that it may still generate (albeit imperfect) partisan-bias-free solutions?

2.2 Computer-Generated Districting Plans

One novel approach to creating non-partisan, unbiased district plans is to computationally generate them. This also allows for a large number of valid district plans to be generated in a short amount of time, with quantifiable methods to evaluate them.

Chou et al. proposed an evolutionary algorithm designed and implemented to generate a large number of legally valid districting plans for the city of Philadelphia, which were then evaluated for "compactness" (defined in this paper as the largest intra-district distance in a districting plan) by both human subjects and a validated surrogate fitness (VSF) function [8]. The evolutionary algorithm generated districting plans based on a single contiguous districting plan which was then mutated by moving neighborhoods from adjacent districts around at every iteration. These plans were experimentally evaluated for fitness, which in this case was a measure of compactness. By comparing the results obtained from both the human evaluators and the VSF function, the researchers overcame a common weakness of Interactive Evolutionary Computation - the cost in time and labor, and human fatigue - by automating the subjective fitness evaluation of the districting plans. One of the main ideas of this paper is how districting plans are evaluated based on the requirements for contiguity, equal population, compactness, and preservation of political and administrative regions (neighborhoods, counties, etc.) By quantifying these criteria, the researchers successfully automated the generation of a large number of different legally valid districting plans.

Chu et al. introduced an algorithm called the Colonial Algorithm, used to draw legislative boundaries that satisfy compactness and population variances, while also retaining relatively simple and clear district lines [9]. The process is controllable and visible, and can speed up the redistricting process and prevent gerrymandering.

Using a dynamic network model that employs depth-first search and breadth-first search, Wang et al. showed that district lines that attain continuity and compactness could be drawn quickly and detect current gerrymandering in a state [24].

By conducting computer simulations of the districting process, Chen and Cottrell created non-gerrymandered districts to generate electoral results from [5]. These results were then compared with real electoral results of congressional races. It was found that while gerrymandering did have an identifiable effect in some states, the net effect was small and unlikely to be the cause of the partisan imbalance in Congress.

Liu et al. proposed a parallel implementation of an evolutionary redistricting algorithm [12]. The algorithm includes spatial evolutionary algorithm operators which take into account spatial characteristics and effectively search the solution space for configurations which are legally valid. By utilizing supercomputers, the performance of the algorithm can be improved so that a large set of viable districting plans can be generated in a short amount of time. Since generating just one of these solutions is commonly a very resource-intensive and difficult task, plagued with human bias, this large set of districting configurations can massively assist lawmakers in developing viable districting plans which are computationally proven to be free of partisan bias and gerrymandering.

Rincón-García et al. proposed and tested a multiobjective simulated annealing algorithm for districting [18]. The multiobjective approach allowed for improvement in measures of compactness and population balance, objectives which are usually ignored in a single objective approach. The simulated annealing algorithm allows for the generation of new districting plans from a seed plan, which is then iteratively improved over time with the multiobjective approach. The novel method was tested on real electoral data from Mexico, and showed better performance (higher quality districts regarding compactness and population equality) than a single objective simulated annealing algorithm. This method using a multiobjective approach can provide the missing link needed to generate districting plans which can compromise for natural constraints.

Optimization Modelling in a GIS Framework: The Problem of Political Redistricting by Macmillan et al., in the book *Spatial Analysis and GIS*, described optimization modeling and how implementing mathematical optimization models in a GIS framework could improve the redistricting process and prevent gerrymandering by individual redistricting [13].

2.3 Graph Partitioning

The major contribution of this project is the creation of a graph partition algorithm which treats a geographical area like a graph, where the graph attributes are population attributes. Many different graph partitioning schemes can then be explored to create district plans.

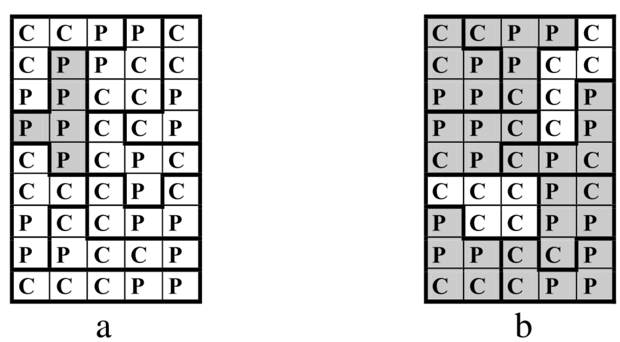

Figure 1: Example by Dixon and Plischke:
(a) Party P wins 1 seat and party C wins 8
(b) Party P wins 7 seats and party C wins 2 [2]

Apollonio et al. introduced a combinatorial model where a districting plan was represented by a partitioned rectangular grid, with red/blue nodes representing clusters of voters [2]. By analyzing models with antagonistic red and blue coloring, it was shown that after a number of graph cycles, the gap between the two colors could get extremely large (sometimes as large as a color-balanced map graph would allow). This showed that partisan bias in districting plans could create a real discrepancy in electoral results. An example of how different divisions can favor one party over another, even if the number of votes doesn't change, is shown in figure 1 [2]. The number of votes for C is 24 and for P is 21, so they almost have equal numbers of votes. But if partitioning scheme in 1.a is used, then P wins one district and C wins eight. In the scenario of 1.b, P wins seven districts and C only wins two.

Ricca et al. studied districting by modeling it with graph partition [17]. In order to obtain districts with compactness and population balance, the proposed approach utilizes weighted Voronoi regions (partitioned regions of a plane based on distance from pre-computed points). The distances were updated iteratively in order to ensure equal population as much as possible.

3 DESIGN

3.1 Framework

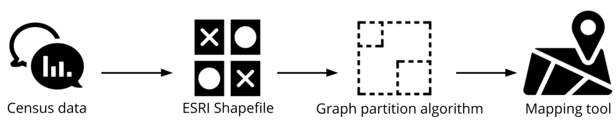

Figure 2: Design Framework for Districting Application

The software components of this project consist of five main parts. The first is a web scraping tool which retrieves the 2010 census data files from the United States Census Bureau website. The second is an ESRI Shapefile parser which reads the census data into their corresponding geospatial shapes and census attributes, adapted from work done by Brian Olson on his redistricting tool [14]. An ESRI (Environmental Systems Research Institute) Shapefile is a data format which stores geometric and attributes information of geospatial data [11]. The graph partition algorithm uses the population attributes produced by the Shapefile parser to generate several districting plans and evaluate them. The mapping tool then takes the modified Shapefiles produced by the graph partition algorithm and visualizes the districting plans. Lastly, an evaluation tool computes different measures of compactness for the existing district plans and compares them to our generated plans.

3.2 Graph Partitioning

For the graph partitioning algorithm implemented in this project, the population attributes focus on compactness. The three attributes calculated to find the most compact districting plan are the average distance between a voter and the center of their district in kilometers (km/p), the population difference between the least populous district and the most populous district (spread), and the standard deviation of district populations (std). The differences in populations of all districts measure how spread out voters are across a state, and it is an indicator of how equally distributed voters are. This is why these values also contribute to evaluating how compact districts are. These values are used by Olson to evaluate the districting plans generated by his redistricter, and are adopted to evaluate the results of this project because the graph partition algorithm implemented in this project is also designed to create compact districts, similar to the objectives of Olson's redistricter [14].

The graph partitioning algorithm creates the districting plans by treating each district as a graph region with population attributes assigned to them by the census data. These regions are represented by the tabular census blocks containing population data. In every iteration, the algorithm tries out different groupings of these blocks, and evaluates these groupings using the aforementioned values in order to arrive at the best districting plans. The calculation of the geographical center of a district is adapted from work by Olson in his redistricting tool. Essentially, the districting plan is the result of solving a graph partitioning problem, and the districts hold certain attributes which are as equal to each other as possible, while maintaining a regular shape and preserving existing administrative boundaries as much as possible. This creates a fair districting plan.

3.3 Mapping Tool

The mapping tool utilizes an ESRI Shapefile parser to overlay the district lines on top of a state map. ESRI Shapefiles can support point, line, and area features, which are parsed from the census data files [11]. ESRI Shapefiles store nontopological data, which means that each geographical feature stored are self-contained and includes data which link it to other features in the same data set. The ESRI Shapefile parser was adapted from work by Brian Olson [14]. The districting plans are displayed with different districts highlighted in different colors.

3.4 Evaluation Tool

To evaluate the generated districting plans, the aforementioned population attributes used to measure compactness in a district are analyzed and compared to existing districting plans. The tool built for this purpose parses the shapefiles of current districting plans for Pennsylvania and Connecticut (obtained from their respective state legislature websites) and calculates the aforementioned compactness values for these plans using the population data for each district available in the shapefiles. These values are then used to quantitatively compare our generated districting plans to the current districts.

4 IMPLEMENTATION

4.1 Data

The 2010 Census data is taken from the United States Census Bureau website [4]. One of the data files is the *tabular blocks* and their boundaries for each geographical areas. The *tabular block* data file has the geographic coordinates of the start and end nodes for each block. This data connects to the census *edges* and *faces* data files. The *edges* data include the geometry and attributes of each topological primitive edge. Each *edge* is delineated by a start node and an end node, and these *edges* are the boundaries that create the *tabular blocks*. Each *edge* has a unique *Line Identifier* value. Each *edge* also has identifiers for its left and right faces, which link to information

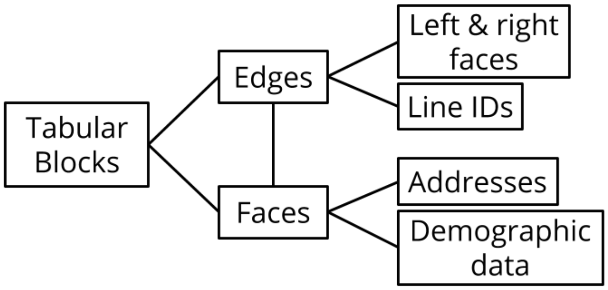

Figure 3: Census Data Structure

in the *faces* data files. Each *face* encompasses an address range, and includes population data about households in that range. Each *face* and *edge* also has an identifier for which county and state it belongs in, and the county boundaries and attributes file is useful in the map visualization. Figure 3 represents the structures and relationships between these census data files.

4.2 Evaluation

Using the evaluation tool, values measuring how spread out the population of districts in existing Congressional district plans can be calculated. In a compact district, voters should be equally spread out from the center of their district. Furthermore, they should be as close to the center as possible. A districting plan where this average distance is minimized is a compact districting plan. Equal population spread throughout different districts and a small standard deviation between the populations of all the districts also create a better districting plan. Therefore, these values will be compared between our generated plans and the existing plans.

5 RESULTS

The generated graph partitioning plans are evaluated for their compactness, measured by the three attributes of the average distance between a voter and the center of their district in kilometers (km/p), the population difference between the least populous district and the most populous district (spread), and the standard deviation of district populations (std). Since some of the worst gerrymandering schemes achieve their goal by creating unusually-shaped districts, districts with contiguous shapes are also less likely to be gerrymandered. There have been algorithms built for this evaluation process, such as the one in Cho and Liu [7]; however, in this case the evaluation for contiguous districts is done via visual inspection of the mapped district plans.

Figure 4 is a Congressional district plan for the state of Pennsylvania generated by the graph partition algorithm and visualized with the mapping tool. Pennsylvania has 18 districts.

Figure 5 is the Pennsylvania Congressional districts from 2011 to 2018, before the court-mandated redistricting of Pennsylvania [19]. Compared to our generated district plan, we can see that the three districts in the southwestern corner of the state (district 12, 14, and 18) are very irregularly shaped, especially compare to those same districts in our generated plan. In the southeastern corner of the state, the green, orange, and yellow districts (district 6, 7, and 16, respectively) are also not contiguously shaped, compared to the

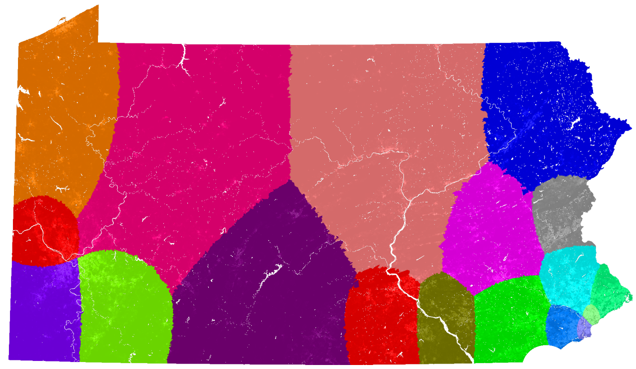

Figure 4: Pennsylvania Congressional district plan generated by the application

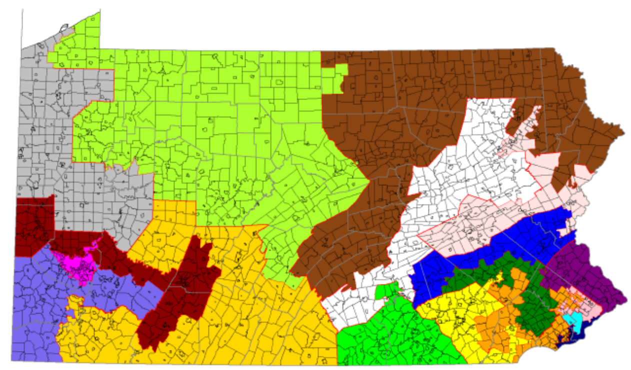

Figure 5: Congressional district map of Pennsylvania (before the 2018 court-mandated redistricting) [19]

districts in the southeastern corner of the map in figure 4. District 10 (dark brown), in the northeastern part of the state, is also much less contiguous than the same district (light pink) in our generated district plan. Generally, we can see that our map contains much more contiguous districts than the existing districts.

Table 2 contains the values measuring compactness calculated during the iterations the graph partition algorithm went through to arrive at the final district plan. The district plan visualized in figure 4 is generated on iteration number 8971, with the lowest average distance between a voter and the geographical center of their district.

Table 3 contains the values measuring compactness calculated using the evaluation tool for the Pennsylvania Congressional district plan from 2011 to 2018.[2] The shapefile that the evaluation tool uses to calculate these values is downloaded from the Pennsylvania Redistricting: The Legislative Guide to Redistricting in Pennsylvania website [16]. As we can see, the average distance between a voter and the center of their district is almost three times

[2]In February 2018, Pennsylvania went through a court-mandated redistricting after the Pennsylvania Supreme Court ruled that the Congressional Districts were an unlawful partisan gerrymander in violation of the Pennsylvania Constitution. See more: https://thehill.com/homenews/state-watch/374561-pa-supreme-court-releases-new-congressional-map

	km/p	spread	std	gen
Best km/p	32.59	6159	1647.00	8971
Best spread	32.61	4947	1609.04	6980
Best std	32.59	5173	1365.19	9855

Table 2: Population attributes for the generated Pennsylvania Congressional district plan in figure 4

km/p	spread	std
62.26	50200	12083.21

Table 3: Population attributes for the 2011-2018 Pennsylvania Congressional district plan in figure 5

larger than that of our generated district plan (62.26 and 32.59, respectively). The population difference between the most and least populous districts is almost ten times larger (50,200 and 6,159), and the standard deviation of the district populations is also almost ten times larger (12,083.21 and 1,647.00).

Figure 6 is a Congressional district plan for the state of Connecticut generated by the graph partition algorithm and visualized with the mapping tool. Connecticut has 5 districts.

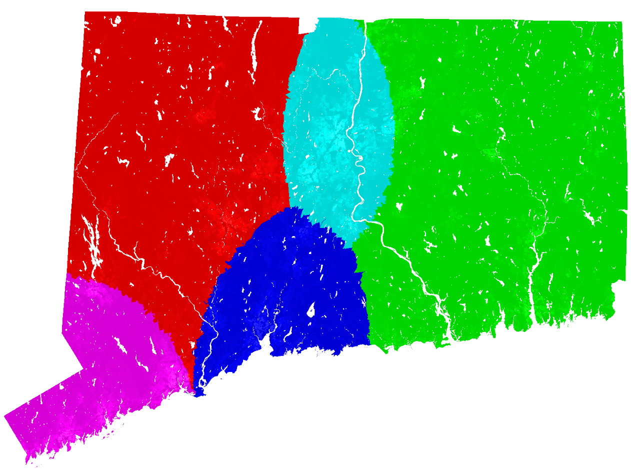

Figure 6: Connecticut Congress district plan generated by the application

Figure 7 is the current Congressional district map of Connecticut [21]. As we can see, district 1 (red) is very irregularly shaped, compared to the same district (teal) in our generated plan in figure 6. Similarly, district 3, 4, and 5 (pink, green, blue, respectively) are also not contiguous shapes, while the same districts (blue, purple, red, respectively) in figure 6 are much more regularly shaped.

Table 3 contains the values measuring compactness calculated during the iterations the graph partition algorithm went through to arrive at the final district plan. The district plan visualized in figure 6 is generated on iteration number 779, with the lowest average distance between a voter and the geographical center of their district.

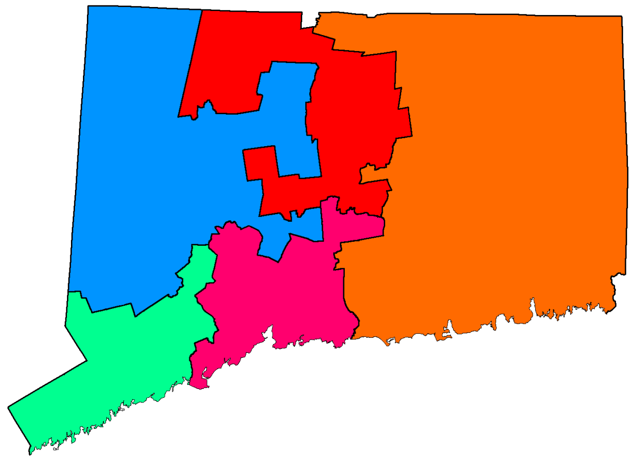

Figure 7: Current Congressional district map of Connecticut [21]

	km/p	spread	std	gen
Best km/p	22.26	5887	2320.42	779
Best spread	22.35	1265	516.78	847
Best std	22.35	1265	516.78	847

Table 4: Population attributes for the generated Connecticut Congress district plan in figure 6

km/p	spread	std
42.50	24174	9866.42

Table 5: Population attributes for the current Connecticut Congressional district plan in figure 7

Table 4 contains the values measuring compactness calculated using the evaluation tool for the current Connecticut Congressional district plan. The shapefile used by the evaluation tool is downloaded from the Connecticut General Assembly website [3]. As we can see, the average distance between a voter and the center of their district is almost double that of our generated district plan (42.50 and 22.26, respectively). The population difference between the most and least populous districts is almost four times larger (24,174 and 5,887), and the standard deviation of the district populations is also almost five times larger (9,866.42 and 2,320.42).

6 CONCLUSION

This project focuses on the implementation of a graph partition algorithm which groups different census tabular blocks together to create districts. These groupings get evaluated during every iteration, and at the end the best grouping is visualized. This visualization is graphically compared to a map of the existing district plan. An evaluation tool is created to parse the shapefiles of existing district plans for the states of Pennsylvania and Connecticut, evaluate them by calculating the compactness measures, and then compare these values to those of the generated plans.

As we can see from the comparison of our visualized district plans and the current district plans of the states of Pennsylvania and Connecticut, the generated plans for both states have more contiguous and regularly shaped districts. Using the evaluation tool to calculate the average distance between a voter and the center of their district, the average voter in our generated plans are much closer to the geographical center of their district compared to the existing district plans in both states. Similarly, the population difference between the most and least populous districts and the standard deviation of all districts' populations are both much lower for our generated district plans than those of the existing district plans. Therefore, regarding the creation of compact districts, the graph partition algorithm is better than the current plans created by state lawmakers.

7 FUTURE WORK

The evaluation tool could be modified to also visualize the current district plans given the shapefiles obtained from state legislature websites. However, not every state has these files available. Furthermore, there are many other criteria to be considered when drawing district lines, such as preserving existing administrative boundaries and communities of interest (i.e., racial or socioeconomic communities). It would also be interesting to analyze the racial and income data of each district and their political affiliation to see how the generated district plans can change the results of elections. These types of information are available in the census data and could be addressed in future work.

On another note, there has been research questioning whether an unbiased solution will always exist for a district plan [15], or if gerrymandering even significantly affects elections at all [5]. These are interesting questions to consider when studying gerrymandering.

ACKNOWLEDGMENTS

The author would like to thank Dr. Xunfei Jiang for her assistance and feedback in the preparation and execution of this project. The Department of Computer Science at Earlham College also enabled the completion of this project. Brian Olson and his work on his redistricting tool was a vital part of the implementation of this project.

REFERENCES

[1] Micah Altman. 1998. Modeling the effect of mandatory district compactness on partisan gerrymanders. *Political Geography* 17, 8 (Nov 1998), 989–1012. https://doi.org/10.1016/s0962-6298(98)00015-8

[2] N. Apollonio, R.I. Becker, I. Lari, F. Ricca, and B. Simeone. 2009. Bicolored graph partitioning, or: gerrymandering at its worst. *Discrete Applied Mathematics* 157, 17 (2009), 3601 – 3614. https://doi.org/10.1016/j.dam.2009.06.016 Sixth International Conference on Graphs and Optimization 2007.

[3] Connecticut General Assembly. 2007. Connecticut Congressional District Data. https://www.cga.ct.gov

[4] United States Census Bureau. 2010. United States 2010 Census TIGER Data. https://www2.census.gov/geo/tiger/TIGER2010/

[5] Jowei Chen and David Cottrell. 2016. Evaluating partisan gains from Congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the U.S. House. *Electoral Studies* 44 (1 12 2016), 329–340. https://doi.org/10.1016/j.electstud.2016.06.014

[6] Jowei Chen and Jonathan Rodden. 2013. Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures. *Quarterly Journal of Political Science* 8 (2013), 239–269. https://doi.org/10.1561/100.00012033

[7] Wendy Cho and Yan Liu. 2016. Toward a Talismanic Redistricting Tool: A Computational Method for Identifying Extreme Redistricting Plans. *Election Law Journal* 15, 4 (11 2016), 351–366. https://doi.org/0.1089/elj.2016.0384

[8] Christine Chou, Steven Kimbrough, John Sullivan-Fedock, C. Jason Woodard, and Frederic H. Murphy. 2012. Using Interactive Evolutionary Computation (IEC) with Validated Surrogate Fitness Functions for Redistricting. In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation (GECCO '12)*. ACM, New York, NY, USA, 1071–1078. https://doi.org/10.1145/2330163.2330312

[9] Hongjiang Chu, Yue Wu, Qiang Zhang, and Yuehua Wan. 2011. *Colonial Algorithm: A Quick, Controllable and Visible One for Gerrymandering*. Springer Berlin Heidelberg, Berlin, Heidelberg, 424–430. https://doi.org/10.1007/978-3-642-19853-3_62

[10] John N. Friedman and Richard T. Holden. 2005. *Towards a Theory of Optimal Partisan Gerrymandering*. Technical Report. Department of Economics, Harvard University, Cambridge, MA.

[11] Environmental Systems Research Institute. 1998. ESRI Shapefile Technical Description. http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf

[12] Yan Liu, Wendy Cho, and Shaowen Wang. 2016. PEAR: a massively parallel evolutionary computation approach for political redistricting optimization and analysis. *Swarm and Evolutionary Computation* 30 (2016), 78–92. https://doi.org/10.1016/j.swevo.2016.04.004

[13] William Macmillan, Todd Pierce, S. Fotheringham, and Peter Rogerson. 2013. *Optimization Modelling in a GIS Framework: The Problem of Political Redistricting*. CRC Press, 221–246. https://books.google.com/books?id=Hirdl1ZFE38C

[14] Brian Olson. 2017. Redistricter. https://bitbucket.org/bodhisnarkva/redistricter/src/default/

[15] Clemens Puppe and Attila Tasnádi. 2008. A computational approach to unbiased districting. *Mathematical and Computer Modelling* 48, 9 (2008), 1455 – 1460. https://doi.org/10.1016/j.mcm.2008.05.024 Mathematical Modeling of Voting Systems and Elections: Theory and Applications.

[16] Pennsylvania Redistricting. 2018. The Legislative Guide to Redistricting in Pennsylvania. http://www.redistricting.state.pa.us/

[17] Federica Ricca, Andrea Scozzari, and Bruno Simeone. 2008. Weighted Voronoi region algorithms for political districting. *Mathematical and Computer Modelling* 48, 9 (2008), 1468 – 1477. https://doi.org/10.1016/j.mcm.2008.05.041 Mathematical Modeling of Voting Systems and Elections: Theory and Applications.

[18] E.A. Rincón-García, M.A. Gutiérrez-Andrade, S.G. de-los Cobos-Silva, P. Lara-Velázquez, A.S. Ponsich, and R.A. Mora-Gutiérrez. 2013. A Multiobjective Algorithm for Redistricting. *Journal of Applied Research and Technology* 11, 3 (2013), 324 – 330. https://doi.org/10.1016/S1665-6423(13)71542-6

[19] JURIST Legal News Research Services. 2018. Pennsylvania 2011 District Map. https://www.jurist.org/news/wp-content/uploads/sites/4/2018/03/PA_2011_map.png

[20] Kenneth W. Shotts. 2001. The Effect of Majority-Minority Mandates on Partisan Gerrymandering. *American Journal of Political Science* 45, 1 (2001), 120–135. http://www.jstor.org/stable/2669363

[21] Ashley Smith. 2011. Connecticut District Map. https://imgur.com/ELYurYm

[22] Nicholas O. Stephanopoulos and Eric M. McGhee. 2015. Partisan Gerrymandering and the Efficiency Gap. *The University of Chicago Law Review* 82 (2015), 831–900. https://lawreview.uchicago.edu/publication/partisan-gerrymandering-and-efficiency-gap-0

[23] James F. Wall. 2008. *Considering Opinion Dynamics and Community Structure in Complex Networks: A view towards modelling elections and gerrymandering*. Technical Report. University of Oxford.

[24] De Wang, Jing Liu, and Zhi-Peng Zhou. 2009. Dynamic Modeling of Political Districting Problem. In *2009 International Conference on Computational Intelligence and Software Engineering (CiSE '09)*. IEEE. https://doi.org/10.1109/cise.2009.5365452