Visualizing Spatial Partitions*

Penny Rheingans[1], Blazej Bulka[2], and Marie desJardins[3]

1 **University of Maryland, Baltimore County**
1000 Hilltop Circle, Baltimore MD, 21250, USA
`rheingan@cs.umbc.edu`
2 **Clark & Parsia, LLC**
926 N St NW Rear, Studio #1, Washington DC, 2001, USA
`blazej@clarkparsia.com`
3 **University of Maryland, Baltimore County**
1000 Hilltop Circle, Baltimore MD, 21250, USA
`mariedj@cs.umbc.edu`

Abstract

We describe an application of geospatial visualization and AI search techniques to the problem of school redistricting, in which students are assigned to home schools within a county or school district. This is a multicriteria optimization problem in which competing objectives must be considered, such as school capacity, busing costs, and socioeconomic distribution. Additionally, school assignments need to be made for three different levels (elementary, middle, and high school) in a way which allows children to move from one school to the next with a cohort of sufficient size. Because of the complexity of the decision-making problem, tools are needed to help end users generate, evaluate, and compare alternative school assignment plans. A key goal of our research is to aid users in finding multiple qualitatively different redistricting plans that represent different tradeoffs in the decision space. We present visualization techniques which can be used to visualize the quality of spatial partititioning plans, compare the alternatives presented by different plans, and understand the interrelationships of plans at different educational levels. We demonstrate these techniques on partitions created through both manual construction and intelligent search processes for the population data of the school district of Howard County, Maryland.

1998 ACM Subject Classification I.3.8 Applications of Computer Graphics, I.2.8 Problem Solving, Control Methods, and Search

Keywords and phrases geospatial visualization, heuristic search, color mapping, multivariate visualization

Digital Object Identifier 10.4230/DFU.Vol2.SciViz.2011.311

1 Motivation and Overview

This research focuses on developing decision support tools for the problem of school redistricting. In this domain, the goal is to assign the students from each geographic region (neighborhood or *planning polygon*) in a school district to a home school at each level (elementary, middle, and high school). We are working with the Howard County (Maryland) school system to develop tools that will aid in generating, evaluating, and comparing alternative school assignment plans. Related applications include emergency response planning, urban planning and zoning, robot exploration planning, and political redistricting.

---

* This work was supported by the US National Science Foundation ((grant #IIS-0414976).

Scientific Visualization: Interactions, Features, Metaphors. *Dagstuhl Follow-Ups, Vol. 2.*
Editor: Hans Hagen; pp. 311–321
DAGSTUHL FOLLOW-UPS Dagstuhl Publishing
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Germany

The school assignment plan should ideally satisfy a number of different goals, such as meeting school capacities, balancing socioeconomic and test score distributions at the schools, minimizing busing costs, allowing students in the "walk area" of a school to attend that home school, and keeping students together in peer groups as they move from school level to level. Since these objectives are often at odds with each other, finding the best plan is a complex multicriteria optimization problem. It is also often desirable to create several alternative plans for consideration; these plans should be *qualitatively different*—that is, they should represent different tradeoffs among the evaluation criteria. Finally, because of the complexity of the problem, it is difficult for users to fully understand these tradeoffs. Therefore, developing effective visualizations is an important challenge.

2 Redistricting Process

The Howard County Public School System (HCPSS) serves a rapidly growing county in suburban Maryland. The pace of development and population growth has necessitated the opening of 25 new schools in the last 16 years, turning the adjustment of school attendance areas into an almost annual event. Under the current process, candidate plans and feasibility studies are generated manually[1] by school system staff. These plans are evaluated and refined by a committee of citizens, then presented at regional meetings for public comment. A small set of candidate plans is forwarded to the Superintendent, who presents two or three recommended alternatives to the Board of Education. The Board has final decision-making authority, and will typically select one of the recommended plans, sometimes making minor modifications in response to concerns raised by parent groups or staff. Note that this process is specific to Howard County; other school districts may have different processes and models.

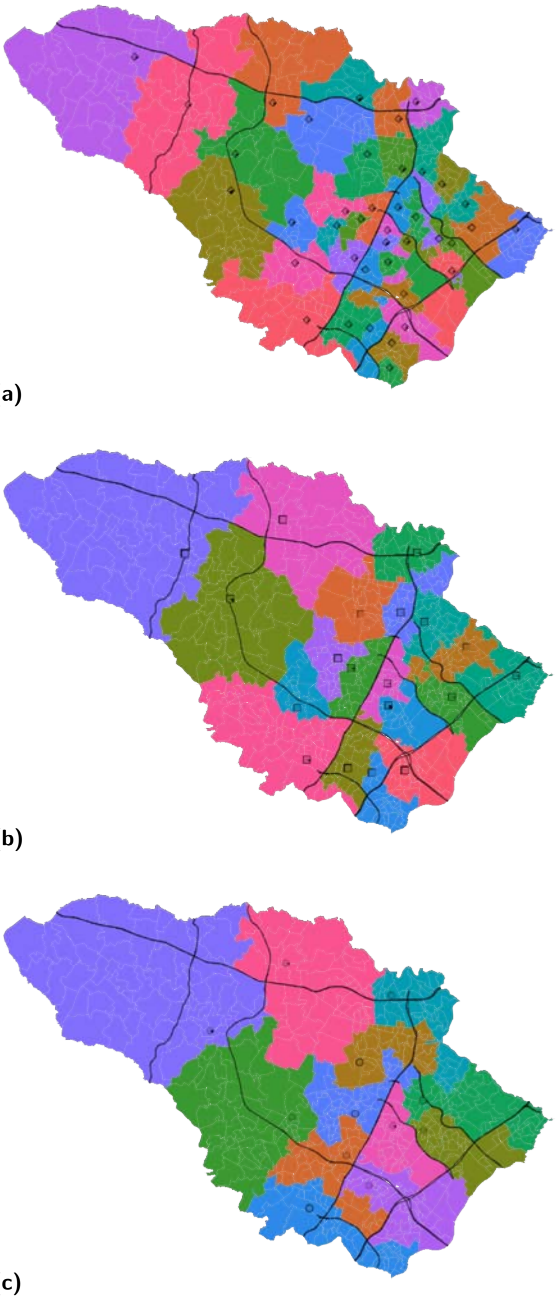
Candidate plans are evaluated according to eleven measured criteria: (1) the educational benefits for students, (2) the frequency with which students are redistricted, (3) the number and distance of students bused, (4) the total busing cost, (5) the demographic makeup and academic performance of schools, (6) the number of students redistricted, (7) the maintenance of feeder patterns (i.e., the flow of students from elementary to middle to high school), (8) changes in school capacity, (9) the impact on specialized programs, (10) the functional and operational capacity of school infrastructure, and (11) building utilization. Some of these criteria can be clearly quantified (e.g., building utilization and busing costs), while others are harder to quantify (e.g., educational benefits and impact on specialized programs).

In practice, the process is primarily driven by building utilization, but serious consideration is given to feeder patterns, the number of students redistricted, demographic makeup, busing costs, and the frequency with which students are redistricted. Ideally, building utilization should be between 90% and 110% of program capacity and should stay in that range as projected population and capacity changes occur. Desired feeder patterns ensure that there is a critical mass of students who move together from one school level (elementary, middle, and high school) to the next. For instance, the students from a particular middle school should constitute at least 15% of the population of any high school that they feed into. Consideration of the demographic makeup of schools helps to ensure that economically and academically disadvantaged children are not unnecessarily segregated into a few schools.

Figure 1 shows the partitioning of planning regions into school attendance areas at the elementary, middle, and high school levels. A glyph shows the location of each school, and

---

[1] Map-based tools are used to show the proposed school districts, and a set of spreadsheets is used to generate evaluation data. No other decision support tools are used in the current process.


(a)

(b)

(c)

Figure 1 2008-2009 plans for **(a)** elementary, **(b)** middle, and **(c)** high school assignment plans. Each planning polygon is colored according to the school to which it is assigned. Heavy black lines show major roads to give geographic context.

each planning polygon is colored according to the school attended. For instance, in the high school (bottom) image, all students in the northwest region of the county are assigned to Glenelg High School (violet region in northwest) in the current plan.

2.1 Building plans

New plans are typically generated by school system staff in response to changes in capacity, such as the completion of a new school or addition, or a localized problem of overutilization. While one might build a plan from scratch, in practice, new plans are generally derived from base plans. An expert identifies trouble spots in the base plan and reassigns polygons to address the problem. For instance, polygons might be moved out of a school which was over capacity or into another to bolster a small feed. Almost invariably, fixing one problem creates another. The schools adjacent to an overcrowded school may not directly have additional capacity, so adjustments may cascade across multiple schools. Additionally, moving a polygon at one level frequently breaks a feed at another level.

Traditionally, plan developers work at a single level and prefer not to introduce changes at another. For instance, in response to the opening of a new high school, staff prefer to contain changes to the high school level. The effect of a change across level, for instance to feed patterns, is generally evaluated as a second step. The complex nature of the interrelationships between levels makes a clear understanding of the whole pattern both difficult and essential. A clear visual representation of the plan under consideration and its characteristics at this level is a starting point, but does not necessarily indicate which potential changes will have undesired effects across levels. A visual representation which effectively conveys the cross-level interrelationships would allow for more efficient selection of polygons to move.

2.2 Searching for plans

The search space for the redistricting problem is very large, making automated search methods particularly attractive. Automated methods allow for the consideration of a larger number of potential plans than would be practical by manual methods. We have developed methods to search for plans which demonstrate good performance on the measured criteria. The result of the search process is a set of alternative plans. Ideally, these plans are very different from one another, making different compromises between criteria.

For p polygons and s schools, there are

$$s^{(p-s)}$$

possible assignments of schools to polygons (since polygons containing a school are constrained to be assigned to that school). Requiring that school attendance areas be geographically contiguous reduces the number of possible plans, but the number of plans still grows exponentially with the number of schools and polygons. Because of this complexity, we have chosen to use heuristic local search methods, which do not guarantee optimality, but which can be used to find good solutions reasonably quickly.

Our basic approach is a two-stage process: first, we generate an initial “seed” plan using one of several methods described below; second, we use local search to “hill-climb” to a local optimum. Because of the multicriteria nature of the redistricting problem, we have designed several different variations of hill-climbing search that can be used to find qualitatively different alternative plans in the solution space. Variants include basic hill climbing, biased hill climbing with blind bias, and biased hill climbing with diversity bias. Our automatic methods were able to find higher quality plans than manual construction methods, both when optimizing for a single outcome and when balancing among them. These methods are described in more detail in desJardins et al. [4, 5].

3 Scenarios

We present a series of typical exploration and analysis scenarios relevant to school redistricting in order to demonstrate the utility of our system. Different scenarios have different goals, resulting in different visualization requirements. For each, we discuss the visualization techniques developed to address those requirements.

Some visualization requirements and design choices are constant across examples. In order to display multiple values on a single polygon or region, we have chosen to combine the alternatives in the same map. While separate maps are more standard and would be easier to produce, the arbitrary shapes and small sizes of polygons would make correspondences and differences between maps very difficult to perceive. In order to show both possibilities, we have chosen to spilt each polygon or region into rings, displaying different school assignments on different rings. Two alternatives for ring aggregation are shown in the examples below.

All visualizations used the same color assignment for a given school. These colors are selected to be roughly equiluminant, in order to avoid brightness and saturation differences between school colors. In order to show assignment, colors must be distinct from those adjacent. Although it might be nice for colors to be perceptually distinguishable from one another more globally, the number of distinct schools makes that impossible. Consequently, the assignment of colors to schools is arbitrary and two schools at the same level might have similar colors. For example, Figure 1c shows two high schools in very similar shades of blue: Reservoir in the south and Mount Hebron in the north. Intended users of the system would understand the geographic constraints, resulting in no confusion.

3.1 Understanding a plan

Helping viewers understand the nature of a plan under consideration is one core function of a visual plan representation. Situations in which facilitating plan understanding are important include the initial developement of plans, the presentation of potential plans to the public, and the consideration of plans by decision-makers. Two basic types of information should be conveyed by a visual representation of a plan. First, the display should compactly represent the geographic structure of the assignments, specifically which polygons are assigned to which schools at the level(s) under consideration. Geographic properties of interest include the contiguity and compactness of attentence areas, as well as the naturalness of boundaries between schools. Next, a good display should convey the key outcomes resulting from that plan, both quantitative information about such measured criteria as utilitzation and qualitative information about how different values of those criteria are distributed across the space.

Consider a plan which assigned each neighborhood to the closest school at each level. Figure 2 shows two representations of such a plan. On the top, the three rings in each polygon show the assignment at the three school levels; high in the outer ring, middle in the center ring, and elementary in the innermost region. On the bottom, polygons with identical assignments have been grouped into homogeneous regions and each region is ringed to show assignments. The two versions emphasize different types of information. The individual polygon rings emphasize the assignments of each polygon clearly, while the region rings emphasize the makeup of the schools. Polygons with assignments distinct from all neighbors are easier to see with the region rings as small areas with an outlier appearance. One example

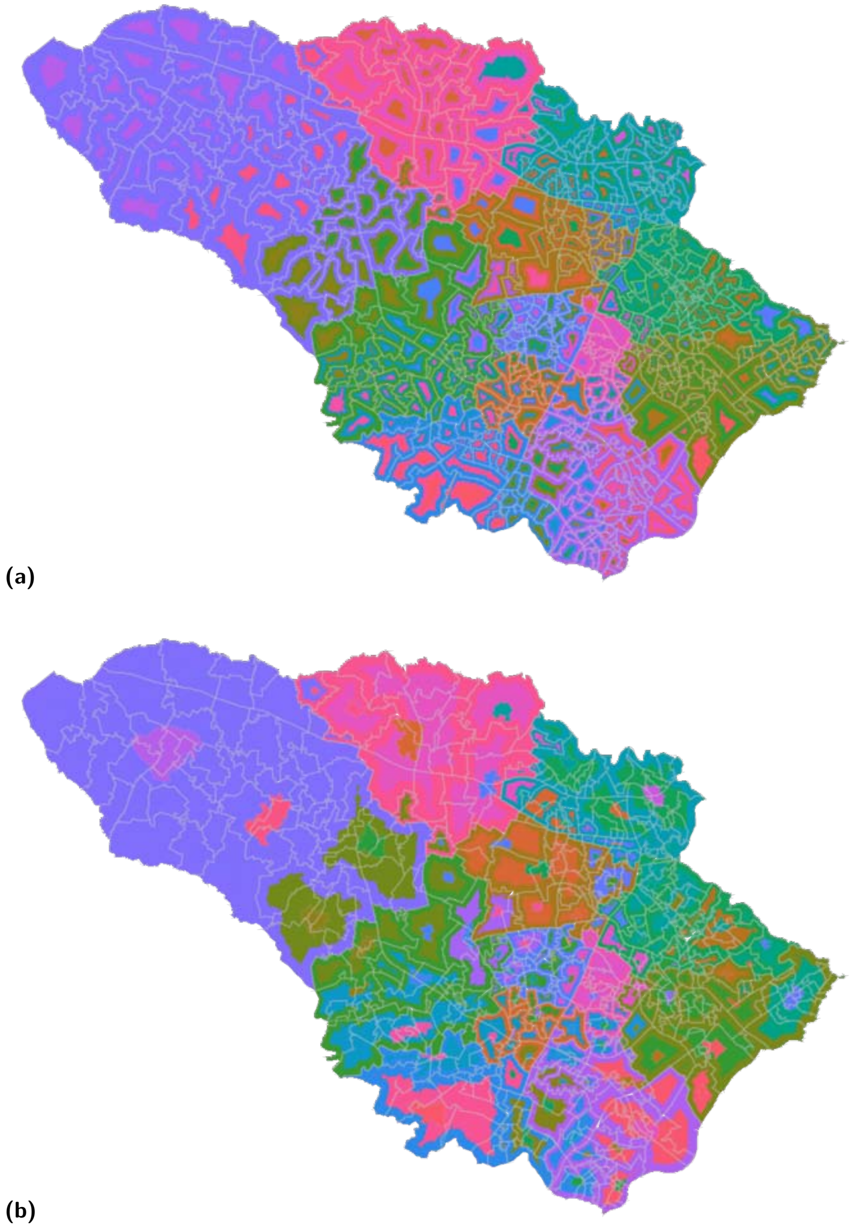
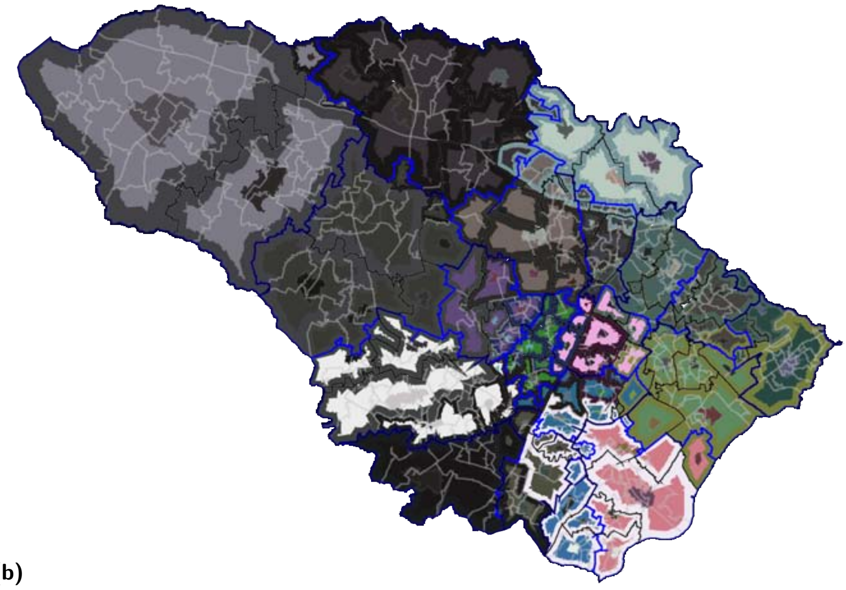


Figure 2 High school plan in which each neighborhood attends the closest school. In the top image **(a)**, each planning polygon is colored according to the schools to which it is assigned (high in the other ring, middle in center ring, and elementary in central area). In the bottom image **(b)**, regions of homogenous assignments are grouped and rings drawn in the resulting region.

can be seen in Figure 2, with a single polygon ringed in orange, green, and olive located toward the eastern side of the county. In this case, those small feeds would not be a problem, since that polygon contains a shopping center and no children (discovered by clicking on the polygon for more detailed information). There is a similar example of an isolated polygon along the northern border of the county; this one does contain children, creating a potential concern. In either view, one can see that the closest school plan makes geographic sense.

Figure 3 shows the pattern of school utilization and free and reduced meal (FARM) percentages created by this plan. Utilization is mapped to brightness (brighter colors show

(a)



(b)

Figure 3 High school plan in which each neighborhood attends the closest school. In the top image **(a)**, each planning polygon is colored according to the schools to which it is assigned (high in the other ring, middle in center ring, and elementary in central area). In each polygon, utilization rate is mapped to brightness, while FARM percentage is mapped to saturation. In the bottom image **(b)**, regions of homogenous assignments are grouped and rings drawn in the resulting region.

fuller schools, while FARM percentage is mapped to saturation (more saturated colors indicate a higher FARM percentage). Since both of these quantities are calculated at the school level, the values are constant across all of the polygons assigned to a school. For such school-level quantities, the choice between individual polygon rings and larger region rings does not make as much difference. Both representations show a great range of brightnesses, indicating that some schools would be dangerously over capacity with such a plan. In particular Hammond

High School in the southeast would be at about 190 percent of capacity, while Marriotts Ridge High School in the northwest would be at about 43 percent of capacity. These images make it clear that a plan which assigns neighborhoods to the closest schools, while intuitively appealing, is not workable given the current distribution of schools and students. FARM percentages, as well, show great disparities among schools.

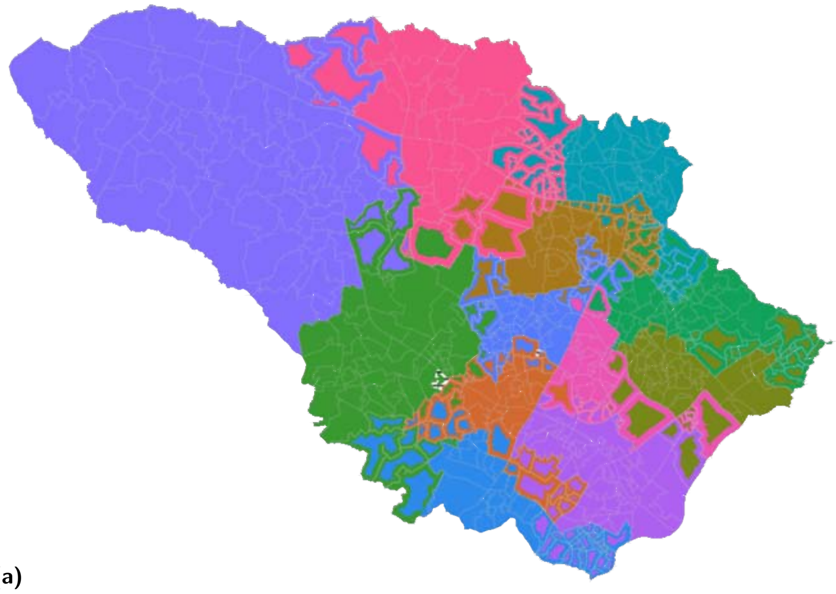
3.2 Comparing alternative plans

During the redistricting process, two or three alternative plans are prepared and presented to the public and the School Board for comment and consideration. Understanding the differences between these alternative plans, both in terms of geography and outcome, leads to better informed opinions and decisions. The goal of visual support for comparing alternative plans is to effectively convey how plans differ in their assignments, how the resulting school attendance areas correspond, and how these assignments impact the characteristics of the resulting school populations. Since the process of choosing between plans is necessarily one of balancing tradeoffs, an accurate understanding of options and costs is essential.

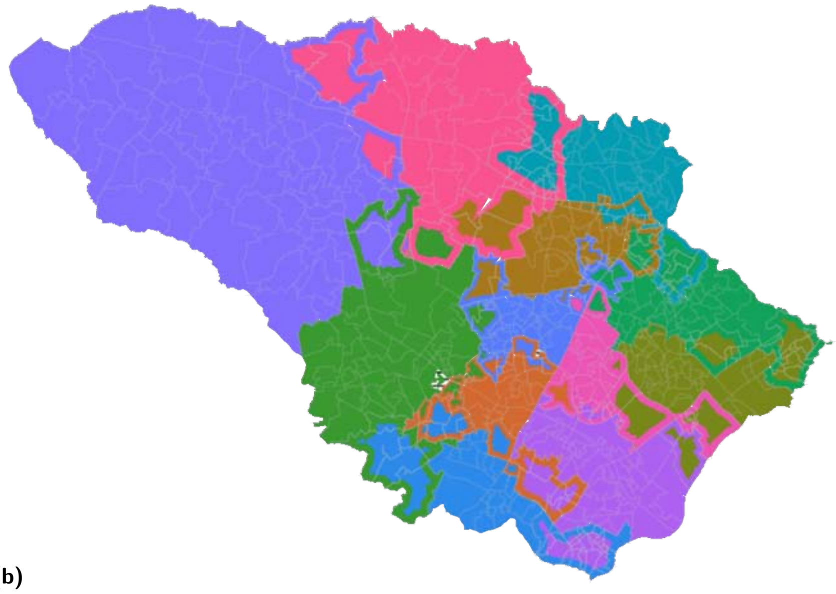
Figure 4 shows a comparison picture of two alternative plans for county high schools. The school assignments for the closest-school plan are indicated by the inner colored rings in each planning polygon. This plan provides a useful baseline, because it optimizes both walk usage and busing costs, but can be seen to be undesirable in terms of utilization and demographics. The outer ring shows the assignments for a plan which balances capacity utilization across schools. The ring effect allows the user to easily see the planning polygons where the two plans make different recommendations. Polygons with a single color are assigned to the same school in both plans.

For example, in the north center of the county, several polygons are assigned to Marriotts Ridge by the balanced-utilization plan, but to the nearby Mount Hebron, River Hill, and Centennial High Schools by the closest-school plan. The outer ring of these polygons is displayed as pink (Marriots Ridge); the inner ring (center) is either teal (Mount Hebron), green (River Hill), or brown (Centennial). This corresponds to the need to draw from a larger area in order to adequately fill Marriotts Ridge. Along the southeast border of the county, the closest-school plan assigns a number of polygons to Hammond High School (purple innner ring) that are assigned to Reservoir (blue outer ring) by the recommended plan. This difference occurs because assigning them to Hammond would cause that school to be over capacity; also, in this case, those polygons help to balance the socioeconomic distribution at Reservoir. In the bottom image of Figure 4, homogeneous regions are joined. This view emphasizes a higher-level overview of the differences.

Figure 5 shows the comparison of these plans in terms of utilization and FARM outcomes. Since the emphasis is on qualities of the resulting high schools, rather than on assignments of specific polygons, we just show the view with merged rings. From this view, one can clearly see that the outcomes at Reservoir High School (the light blue area in the south) are very different under the two plans, with the balanced-capacity plan increasing both utilization and FARM percentage. With the balanced-capacity plan (shown in the outer ring), there is moderate utilization and FARM percentage (shown by moderate brightness and saturation), while the closest-school plan has very low utilization and FARM percentage (shown by the dark grey inner ring in the area which attends the school under both plans). In the western part of the county, the plans have different outcomes in terms of utilization (utilization is generally slightly lower under the balanced-capacity plan, shown by darker rings), but similar characteristics in terms of FARM percentage (both sets of rings are very unsaturated, indicating low FARM percentage).



(a)



(b)

Figure 4 A comparison of the balanced-utilization plan to the closest-school plan. **(a)** The color of the outer and inner rings in each planning polygon indicate the school assignments for the balanced-utilization and closest-school plan, respectively. **(b)** polygons with identical assignments are grouped together into regions.

4 Related Work

The problem of school redistricting is related to that of political redistricting. Several software packages (such as Maptitude [3]) are available for building and analyzing political and school redistricting plans. These packages do not generally provide automated or interactive search methods, do not provide visual comparison techniques such as our ring comparison, and do not facilitate the discovery of qualitatively different plans. Academic research on computational

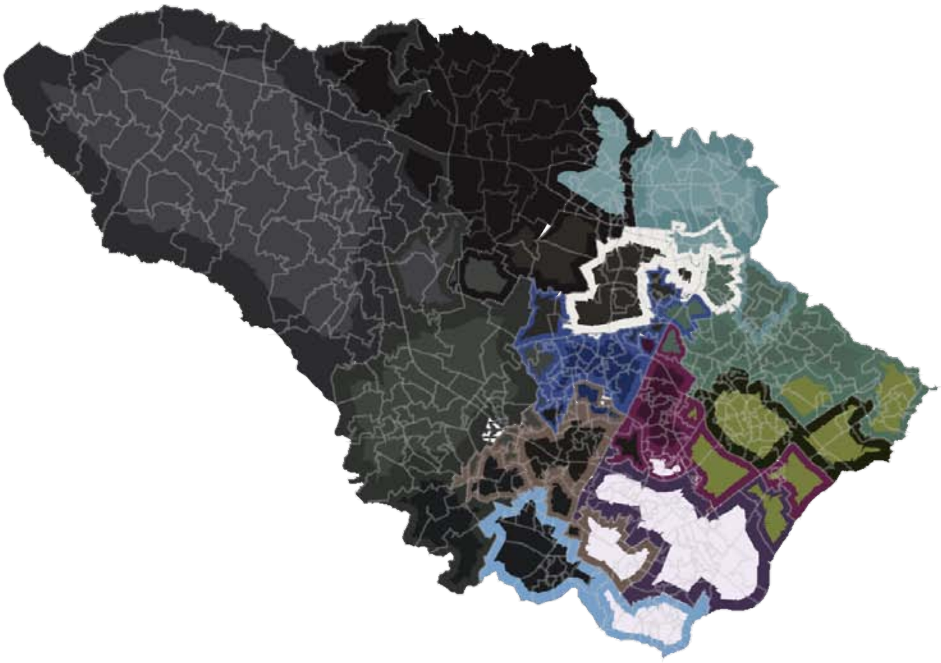


Figure 5 A comparison of the balanced-utilization plan to the closest-school plan. The color of the outer and inner rings in each group of polygons with identical assignments indicate the school assignments for the balanced-utilization and closest-school plan, respectively. Saturation shows FARM percentage, while brightness shows utilization.

approaches to political redistricting [7, 1] has concentrated on methods for constructing plans, rather than visualizing the result.

School redistricting differs from political redistricting in several important ways. First, although compactness is an important factor (both for community building and to minimize busing costs), it is not as important as in political redistricting. Second, the walk usage and feeder issues complicate the scenario for school redistricting. Third, redistricting occurs more frequently (at least in Howard County) than in most political districts, and students are greatly affected by the process. As a result, minimizing the number of students who are redistricted is also an important criterion. Finally, the nature of the decision-making process, in which alternative plans are explicitly compared and contrasted to each other, raises the desirability of generating multiple plans that represent different tradeoffs.

We draw upon previous research on how sets of discrete colors should be used effectively in data visualization. Brewer [2] characterizes different kinds of properties (e.g., numeric and a few types of categorical ones) and their visualization by mixing colors and the use of color components (saturation and brightness). Healey [6] described factors that affect the effectiveness of a particular color selection: the Euclidean distance between the colors in a perceptually uniform color space (L^*, u^*, v^*) and the geometrical positioning of color patterns. Healey also presented a method for manual selection of a fixed number of colors that can be displayed on a monitor by placing them at regular intervals in a carefully positioned circle in L^*, u^*, v^* space. Our method for color selection uses Healey's basic approach.

5 Conclusions

School redistricting is an interesting and challenging problem both computationally and from an application perspective. We have developed a prototype system that uses novel heuristic search and visualization techniques to aid an end user in generating, evaluating, and comparing alternative plans. These tools should provide end users with significant insights into the tradeoffs among alternatives.

The school redistricting problem is closely related to the resource positioning problem of deciding where to build schools, locate fire or police substations, or position emergency response equipment. Our optimization framework, search methods, and visualization tools can all potentially be applied to these other application domains.

Acknowledgements

This work was supported by the National Science Foundation (grant #IIS-0414976). We would like to thank David Drown, Joel Gallihue, and Larry Weaver of the Howard County Public School System's Superintendent's Office for their support, data, and valuable feedback. Thanks also to Poonam Shanbhag, who wrote the original visualization software, and to Andrew Hunt and Priyang Rathod, who also worked on the implementation.

References

1 Micah Altman. *Districting principles and democratic representation.* PhD thesis, California Institute of Technology, 1998.

2 Cynthia A. Brewer. Color use guidelines for data representation. In *Proceedings of the Section on Statistical Graphics*, pages 55–60. American Statistical Association, 1999.

3 Caliper Corporation. Maptitude for redistricting, 2006. `http://www.caliper.com/mtredist.htm`.

4 Marie desJardins, Blazej Bulka, Ryan Carr, Andrew Hunt, Priyang Rathod, and Penny Rheingans. Heuristic search and information visualization methods for school redistricting. In *Proceedings of the Eighteenth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-06)*, pages 1774–1781. AAAI Press, 2006.

5 Marie desJardins, Blazej Bulka, Ryan Carr, Eric Jordan, Priyang Rathod, and Penny Rheingans. Heuristic search and information visualization methods for school redistricting. *AI Magazine*, 28(3):59–72, 2008.

6 Christopher G. Healey. Choosing effective colours for data visualization. In *Proceedings of IEEE Visualization '96*, pages 263–270, 1996.

7 Robert E. Heldig, Patrick K. Orr, and Robert R. Roediger. Political redistricting by computer. *Communications of the ACM*, 15(8):735–741, August 1972.