AN EXTREME-SCALE COMPUTATIONAL APPROACH TO REDISTRICTING OPTIMIZATION

Allocation: Illinois/0.6 Mnh
PI: Shaowen Wang[1,2,3,4,5,6]
Co-PIs: Wendy K. Tam Cho[1,6,7,8], Yan Liu[1,2,6]

[1]CyberGIS Center for Advanced Digital and Spatial Studies, University of Illinois at Urbana-Champaign
[2]Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign
[3]Department of Computer Science, University of Illinois at Urbana-Champaign
[4]Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign
[5]Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign
[6]National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign
[7]Department of Political Science, University of Illinois at Urbana-Champaign
[8]Department of Statistics, University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY:

The goal of this phase is to develop a scalable political redistricting application and use the application to integrate significant computing power from Blue Waters with our computational redistricting strategies for analyzing, contextualizing, and understanding redistricting plans. Because the redistricting problem is complex and computationally challenging, existing algorithms were feasible for only very small redistricting problems. We are at the cusp of being able to draw insights from both statistical models as well as computational models for this task.

We have developed a scalable computational approach that allows us to synthesize and organize massive amounts of computation and data to evaluate redistricting schemes and tailor them to notions of "fairness" and democratic rule. Specifically, we extended a scalable parallel genetic algorithm (PGA) to a parallel evolutionary algorithm (PEA) for redistricting. In the PEA, we developed a set of spatial evolutionary algorithm (EA) operators to handle the costly spatial configuration of redistricting maps. Based on the PEA, we formulated a novel set of statistical measures for identifying partisan gerrymanders. Preliminary results of 399,562 high-quality redistricting maps were obtained on Blue Waters using up to 32,768 processor cores. These maps were used to conduct a comparative study of existing redistricting plans and draw insights that were not previously possible.

INTRODUCTION

Redistricting or drawing electoral maps amounts to arranging a finite number of indivisible geographic units into a smaller number of larger areas (i.e., districts). A computational approach is developed to harness the massive computing power provided by Blue Waters to quantitatively study redistricting, e.g., partisan gerrymandering. Partisan gerrymandering comes in two flavors. The first is partisan gerrymandering in which one party seeks to disadvantage the other party by creating districts that pack excessive numbers of minority party voters into districts. This tactic wastes some set of the minority votes while more efficiently dispersing the majority voters among districts to ensure victory. Another form of political gerrymandering is bipartisan gerrymandering where the two parties, majority and minority, join forces to create a "sweetheart deal" wherein the incumbents of both parties are protected in safe seats, thereby preserving the status quo. Both types of gerrymanders seek to minimize the role of the voters. To identify a partisan gerrymander, we must have the ability to compare large numbers of possible plans. Creating a large number of possible plans has not previously been possible, so while this idea is theoretically attractive, it remains a fundamentally new way to conceptualize gerrymanders.

METHODS AND RESULTS

In our computational model, we formulate the redistricting problem as a combinatorial optimization problem that identifies redistricting plans as it attempts to optimize objectives (e.g., competitiveness, safe districts, incumbent protection) while simultaneously satisfying legal constraints (e.g. contiguity and equipopulous districts). The redistricting problem is NP-hard. The solution space of a large redistricting problem is not characteristic of a rugged solution space. While the space landscape is hilly in the sense that it has the usual peaks and valleys, these peaks and valleys are not a rapid succession of precipices but instead a series of vast plateaus, and hence, not rugged in the traditional sense.

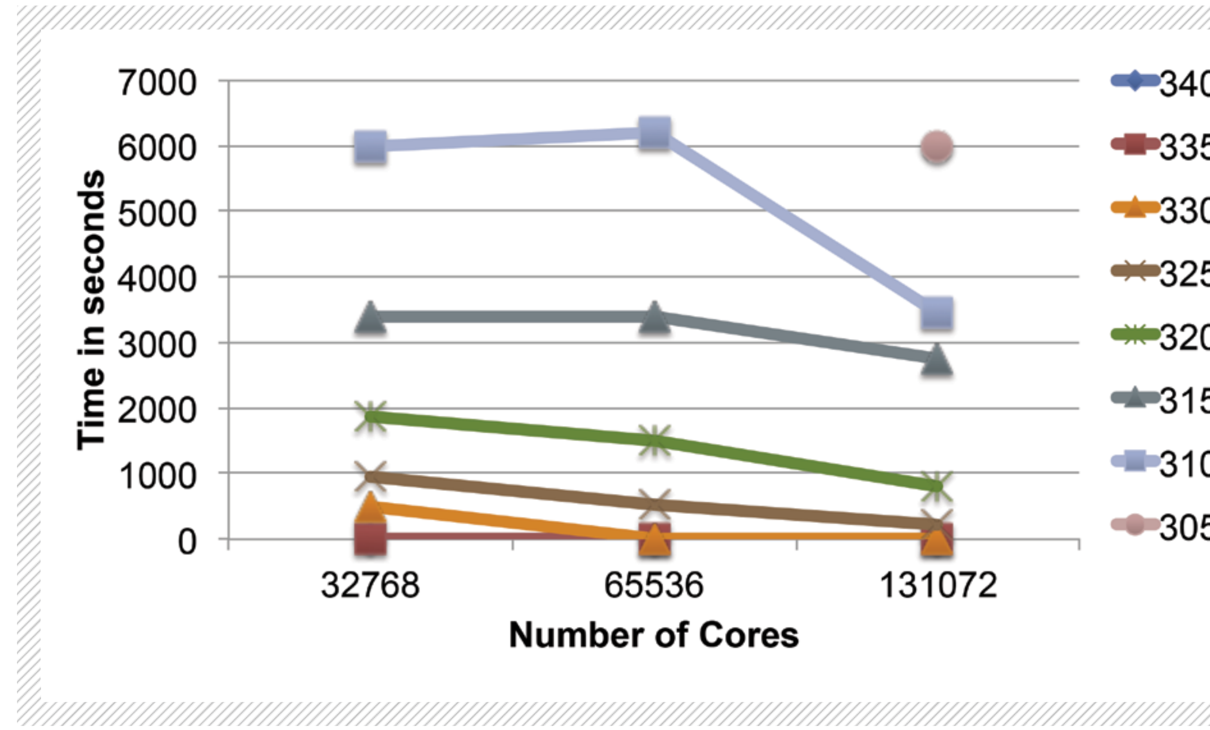

FIGURE 1: Weak scalability results. Execution time is measured at multiple solution quality thresholds. Data: North Carolina VTD.

We employ an evolutionary algorithm (EA) to efficiently find optimal or near-optimal solutions. The EA, which is a stochastic algorithm, is developed to handle spatial configuration of redistricting maps and enable the search process to avoid simple local optima. The parallelization of the EA provides a scalable computational approach for using a large number of processors to work on many plateaus simultaneously and jump from one to another through non-blocking inter-process communication for elite solution propagation and the collective but independent evolutionary searches surrounding these elite solutions.

We developed our parallel EA (PEA) by extending an existing PGA library to improve its scalability and provide explicit programming control of parallelism to overlap the computation and communication. The PEA code exhibited desirable scalability to up to 131,000 integer cores on Blue Waters with marginal communication cost (0.015% on 16,384 cores, compared to 41.38% using synchronous communication). It can conduct 22 EA iterations per second.

To demonstrate how our approach sheds light on redistricting plans, we examined North Carolina at the voting district (VTD) level using 2010 census data. On Blue Waters, 32,768 processor cores were used to identify 399,562 feasible solutions. These solutions are legally possible (contiguity requirements) and reasonable redistricting plans (with population deviation under 5%, compactness scores of 0.95 or better, and respect for political subdivisions). They form a solution pool that characterizes the underlying distribution of all solutions beyond a threshold of goodness. Based on this distribution, the 2001 redistricting plan is an outlier except on biasedness and efficiency gap, and the year 2011 plan (the current plan) is an outlier except on responsiveness.

Our computational approach has the potential to fundamentally transform our understanding of the redistricting process by providing a widely accessible objective tool that will open the redistricting process to participation by a broader and more diverse group of stakeholders. It will also provide great flexibility and enhanced capability for designing redistricting plans.

WHY BLUE WATERS

With the aid of supercomputers, no longer are we merely at the point of solely using a computer interactively to complete simple tasks, we are now contemplating the ability to employ computationally intensive models that can

FIGURE 2: An example redistricting plan found by the PEA and histograms of partisan gerrymandering measures, compared with two existing plans.

generate and evaluate alternative redistricting schemes and compare them to each other on various notions of "fairness," far exceeding human capabilities and analysis. No longer are we constrained to consider a plan in isolation. We have made considerable progress with the aid of the Blue Waters supercomputer and the superb support from the Blue Waters team, who initially created and enabled scalable computational software to support domain science research. We expect to access the next generation of Blue Waters to further develop our statistical and computational approach and achieve the capability of solving larger redistricting problems (e.g., census block-level redistricting with 0.5 million problem variables).

PUBLICATIONS

Liu, Y.Y, W.K.T Cho, S. Wang (2015). A Scalable Computational Approach to Political Redistricting Optimization. *Proceedings of the 2015 Annual Conference on Extreme Science and Engineering Discovery Environment (XSEDE'15)*, July 26-30. St Louis, Missouri.

Liu, Y. Y. and S. Wang (2015), A Scalable Parallel Genetic Algorithm for the Generalized Assignment Problem, *Parallel Computing*, 46, pp 98-119.