ENABLING REDISTRICTING REFORM: A COMPUTATIONAL STUDY OF ZONING OPTIMIZATION

Allocation: Illinois/400 Knh
PI: Wendy Cho[1]
Co-PI: Yan Liu[1]
Collaborator: Bruce Cain[2]

[1]University of Illinois at Urbana-Champaign
[2]Stanford University

EXECUTIVE SUMMARY

We have developed a scalable computational tool for redistricting that synthesizes and organizes massive amounts of computation and data to evaluate voter redistricting schemes. The tool allows one to create high-quality maps and tailor them to notions of "fairness" and democratic rule. It can also be used as an evaluation tool by courts, advocates, and the public to ensure nondiscriminatory representation. Specifically, we developed a scalable, parallel, evolutionary algorithm for redistricting that includes a set of spatial evolutionary algorithm operators to handle the costly spatial configuration of redistricting maps. These maps provide the basis for additional statistical analysis.

RESEARCH CHALLENGE

In the United States, political redistricting occurs every 10 years following the decennial census. It is intended to provide fair representation for all communities and interest groups. This process is hampered when those drawing the maps unfairly advantage a particular partisan or racial group over others. Despite broad disdain for the practice of gerrymandering, the Supreme Court has found it difficult to identify a workable standard by which gerrymandering might be regulated. Moreover, in more than two-thirds of the states, the majority and self-interested political party in the lower state house has the responsibility of devising the electoral map, creating a process that is inherently skewed toward biased outcomes.

In the context of litigation, there are insufficient tools to synthesize redistricting data to provide a sufficient basis from which to analyze and decide the legal issues. The problem is, in part, because the requisite computation is massive. Without the tools to quantify the effect of electoral maps, the court is left without the ability to issue legal and consistent judgments. As a result, despite the five decades since the Supreme Court declared gerrymandering to be capable of being decided by legal principles, the court has yet to identify manageable standards under which one could measure and identify a partisan gerrymander. The failure of the legal system in this political realm has significant ramifications for our democratic system of governance.

At the initial map drawing stage, computation has improved significantly in the last two decades, but this advancement has served only to further isolate and insulate the process since only politicians and those in their employ have had access and the knowledge to exploit these tools. In essence, technological advances have only served to increase the manipulability of the redistricting process by those with nefarious intent. Our work intends to reverse this trend.

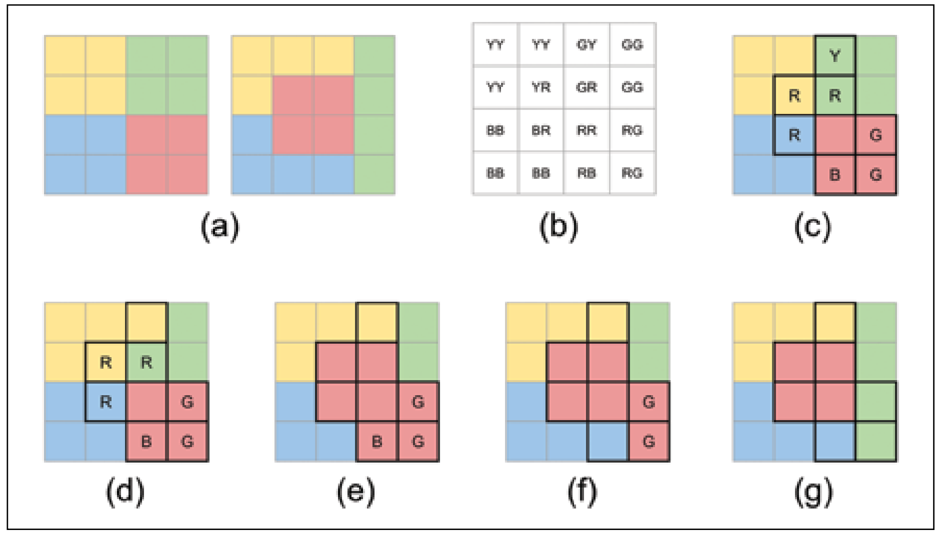

Figure 1: By applying the path relinking heuristics on spatial configurations, the spatial crossover operation is able to generate multiple new solutions that propagate alleles from.

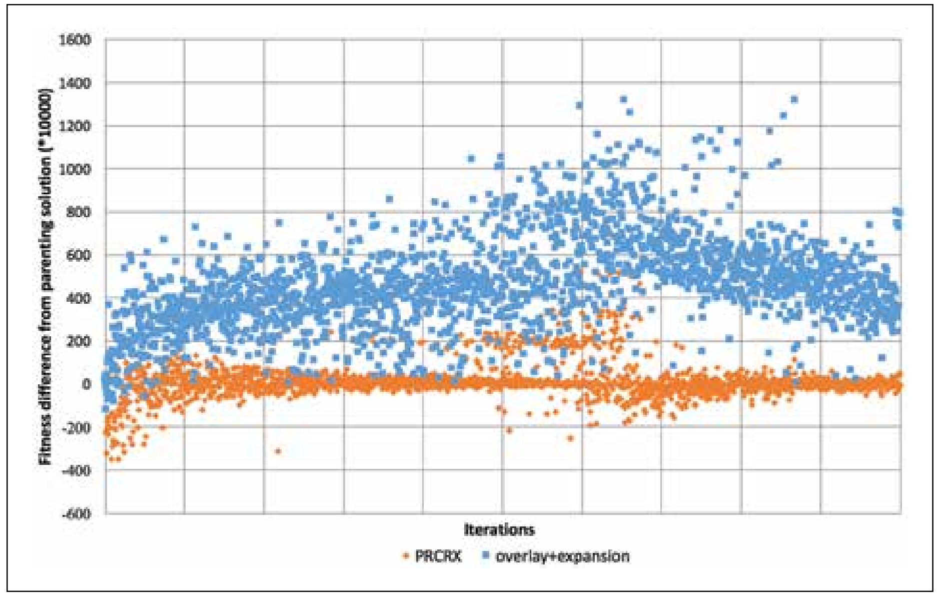

Figure 2: Performance comparison: Spatial Path Relinking crossover operator and the basic overlay + expansion crossover. Both take the same input. Fitness difference between the input and output solution is plotted. Negative difference value means fitness improvement.

Using the Supreme Court's articulated legal reasoning and mandates, we have developed a computational redistricting tool that formulates redistricting as a combinatorial optimization problem, with objectives and constraints defined to meet legal requirements. Drawing electoral maps amount to arranging a finite number of indivisible geographic units into a smaller number of districts where every unit must belong to exactly one district and no districts are empty. The redistricting problem is similar to the set-partitioning problem that is known to be *NP*-hard and, thus, the time required to solve the problem increases very quickly as the size of the problem grows. Our scalable evolutionary computational approach utilizes massively parallel high-performance computing for redistricting optimization and analysis at fine levels of granularity.

We intend for our tool to be used both at the initial map-drawing stage as well as at the litigation stage. At the map-drawing stage, the tool is far more advanced than the extant tools that politicians employ. It also incorporates the idiosyncratic legal criteria required of electoral maps so that those without this specialized knowledge are still able to utilize it. At the litigation stage, it enables the statistical analyses that the Supreme Court has found illusive. Both potential uses are intended to improve outcomes from the redistricting process.

METHODS & CODES

Our algorithm, Parallel Evolutionary Algorithm for Redistricting (PEAR), is implemented in ANSI C. It can be compiled on Linux and OS X as a standard *makefile* project. PEAR uses MPI nonblocking functions for asynchronous migration for load balancing and efficiency. It uses the C SPRNG 2.0 library to provide a unique random number sequence for each MPI process, which is necessary for running a large number of evolutionary algorithm (EA) iterations.

We devised the evolutionary algorithm operators specifically for the redistricting application. The effectiveness and efficiency are enhanced by the explicitly spatially aware mutation and crossover operators. These operators are much more efficient in navigating spatially constrained decision space where the solution space exhibits extremely high ruggedness that classic EA operators could not explore comprehensively. The crossover operator employs a path relinking heuristic on adjacency planar graphs and generates new solutions by randomized graph and geometry calculation of suitable neighborhoods. Parallelization of the algorithm further harnesses massive parallel computing power via the coupling of EA search processes and a highly scalable message-passing model that maximizes the overlapping of computing and communication at runtime.

RESULTS & IMPACT

Our approach is designed to identify redistricting maps that satisfy a set of user-defined criteria with a particular focus on addressing fine levels of spatial granularity. We leveraged and enhanced a scalable Parallel Genetic Algorithm library to develop PEAR for the computationally intensive redistricting problem. PEAR provides a powerful and computationally scalable redistricting tool that has never before existed by incorporating a set of spatial configuration operators and spatial EA operators to handle spatial characteristics and the associated computational challenges, and by harnessing massive computing power. It incorporates a novel spatial crossover operator that incorporates the optimization literature on path relinking methods for spatial constraints.

Our work has been referenced in 11 amicus briefs and discussed in oral arguments before the Supreme Court in the landmark

Continued on page 267

TN

POLICY RESPONSES TO CLIMATE CHANGE

Allocation: GLCPC/250 Knh
PI: Yongyang Cai[1]
Co-PIs: Kenneth Judd[2], William Brock[3], Thomas Hertel[4]
Collaborators: Simon Scheidegger[5], T.J. Canann[6], Carlos Rangel[7]

[1]The Ohio State University
[2]Hoover Institution
[3]University of Wisconsin
[4]Purdue University
[5]University of Zurich
[6]University of Minnesota
[7]Penn State University

EXECUTIVE SUMMARY

Cai, Brock, Xepapadeas, and Judd [1] have built a model of Dynamic Integration of Regional Economy and Spatial Climate under Uncertainty (DIRESCU), incorporating a number of important climate science elements that are missing in most integrated assessment models. These include spatial heat and moisture transport from low latitudes to high latitudes, sea level rise, permafrost thaw, and tipping points. Using this more realistic model of the world economy and climate, we study policy responses to climate change under cooperation and various degrees of competition among regions. We find that other assessment models that are missing elements of climate science lead to significant bias in important policy variables such as the social cost of carbon and adaptation.

RESEARCH CHALLENGE

Leading integrated assessment models assume that climate damages are related to the mean surface temperature of the planet. But climate science shows that when the climate cools or warms, high-latitude regions tend to exaggerate the changes seen at lower latitudes due to spatial heat and moisture transport. This effect is called polar amplification (PA). Thus, the surface temperature anomaly is differentiated across spatial zones of the globe. The low- (high-) latitude regions would be hotter (colder) if poleward heat transport were absent; hence, damages in the low-latitude regions would be higher since they are already under heat stress and transporting some of that heat poleward helps relieve this heat stress.

PA will accelerate the loss of Arctic sea ice, a potential meltdown of the Greenland and West Antarctica ice sheets, which could cause serious global sea level rise. Moreover, PA will lead to faster thawing of the permafrost, which is expected to bring about widespread changes in ecosystems and damage to infrastructure, along with the release of greenhouse gases that exist in permafrost carbon stocks. Furthermore, PA will also affect the likelihood of tipping points, such as the "nearest" three potential tipping points located in the high latitudes of the Northern Hemisphere (Arctic summer sea ice loss, Greenland ice sheet melt, and boreal forest loss).

METHODS & CODES

We developed the DIRESCU model to include spatial heat and moisture transport from low latitudes to high latitudes, sea level rise, permafrost thaw, and tipping points. To model spatial heat and moisture transport, we disaggregate the globe into two regions: Region 1 is the region north of latitude 30°N to 90°N (called the North), while Region 2 is the region from latitude 90°S (the South Pole) to 30°N (called the Tropic-South). The disaggregation also makes clear their significant economic difference, since most countries in the Tropic-South are poor and most countries in the North are rich.

To address the tipping points and solve the dynamic stochastic programming problem, we adapt the computational method in DSICE [2], developed by Cai and Judd in the past four years using GLCPC allocations on Blue Waters. The computational method is parallel backward value function iteration using the master-worker structure—the master assigns N tasks for workers to solve in parallel and then gathers the results of these tasks from workers. Our code shows high parallel efficiency, with an almost linear speed-up from 30 nodes to 5,000 nodes.

RESULTS & IMPACT

In 2017, Cai, Judd, and Steinbuks published a paper in *Quantitative Economics* [3] that develops a nonlinear certainty equivalent approximation (NLCEQ) method to solve efficiently and in parallel huge-dimensional dynamic stochastic problems without exogenous trends, by using Blue Waters resources. We also extended NLCEQ to solve high-dimensional dynamic stochastic problems with exogenous trends and applied it to analyze the effect of climate and technological uncertainty in crop yields on the optimal path of global land use in [4].

Also in 2017, Yeltekin, Cai, and Judd published a paper in *Operations Research* [5] that developed a parallel algorithm that can solve supergames with states, which models strategic interactions among multiple players, by using Blue Waters resources.

In 2018, Cai, Brock, Xepapadeas, and Judd released a National Bureau of Economic Research working paper [1] that is under review for publication in a prestigious economic journal. The paper builds the DIRESCU model, studies optimal climate policies under cooperation and various degrees of competition among regions, and finds that excluding some of the elements of climate science leads to significant bias in important policy variables such as the social cost of carbon and adaptation.

WHY BLUE WATERS

Our parallel computational package requires low-latency communications because the algorithm uses the master-worker structure and needs frequent communications between the master and workers. Our problems are large. For example, the DIRESCU model has 10 continuous state variables and one binary state variable, as well as eight continuous decision variables, and a more than 500-year horizon. It corresponds to solving a Hamilton–Jacobi–Bellman equation with 10 or 11 state variables. Using our efficient parallel algorithm, we solved it with one specification case in 3.4 wall-clock hours with 102 computer nodes on Blue Waters. Moreover, we have solved the model with many specification cases for analysis. In addition, the largest problem we solved for DSICE used 3,459 computer nodes and took 7.5 wall-clock hours on Blue Waters. Blue Waters allows us to solve these large problems efficiently as has already been shown in our previous work.

PUBLICATIONS & DATA SETS

Cai, Y., K.L. Judd, and J. Steinbuks, A nonlinear certainty equivalent approximation method for stochastic dynamic problems. *Quantitative Economics*, 8:1 (2017), pp. 117–147.

Yeltekin, S., Y. Cai, and K.L. Judd, Computing equilibria of dynamic games. Operations Research, 65:2 (2017), pp. 337–356.

Continued from page 265

partisan redistricting case, *Gill v. Whitford*. Our computational approach to redistricting reform has garnered media attention from popular outlets (Vox.com, *Chicago Inno*, *Reason*, the *Washington Post*); computing outlets (Cray Inc., TOP500, *Communications of the ACM*); and outlets aimed at the science and mathematics communities (*Quanta Magazine*, Science Node, *WIRED*, *Nature*). Most of the discussion has focused on the impact of our work on gerrymandering litigation. However, our work is also applicable to earlier stages of the redistricting process, which is now evolving and becoming clearer to practitioners and politicians as the project progresses.

WHY BLUE WATERS

The PEAR library is designed for extreme-scale redistricting applications. From the beginning, it was intended to scale to all of the processor cores on Blue Waters through nonblocking MPI communication calls. The computational approach implemented in our solution requires generating a very large number of electoral maps for quantitative study of redistricting phenomena. Identifying quality electoral maps requires significant computing in the combinatorial optimization process. Generating a large number of statistically independent maps is only feasible on a supercomputer at Blue Waters' scale.

PUBLICATIONS & DATA SETS

Cho, W.K.T., and Y.Y. Liu, Sampling from Complicated and Unknown Distributions: Monte Carlo and Markov Chain Monte Carlo Methods for Redistricting. *Physica A*, (2018), DOI:10.1016/j.physa.2018.03.096.

Cain, B., et al., A Reasonable Bias Method for Redistricting: A New Tool for an Old Problem. *William & Mary Law Review*, 59: 5 (2018).

Cho, W.K.T., and Y.Y. Liu, Massively Parallel Evolutionary Computation for Empowering Electoral Reform: Quantifying Gerrymandering via Multi-objective Optimization and Statistical Analysis. *SC17: The International Conference for High Performance Computing, Networking, Storage and Analysis* (Denver, Colo., November 12–17, 2017).

Cho, W.K.T., Measuring Partisan Fairness: How well does the Efficiency Gap Guard against Sophisticated as well as Simple-Minded Modes of Partisan Discrimination? *University of Pennsylvania Law Review Online*, 166 (2017).

Liu, Y.Y., and W.K.T. Cho, A High-Performance Evolutionary Computation Framework for Scalable Spatial Optimization. *ICCS: International Conference on Computational Science: Science at the Intersection of Data, Modelling and Computation* (ICCS, Wuxi, China, June 11–13, 2018).