Identifying and Preserving Communities-of-Interest in PA Congressional Maps via Graph Theory and Markov Chain Algorithms

Mondal, Antara (School: Upper St. Clair High School)

In the US, electoral redistricting frequently divides Communities of Interest (COIs) in favor of partisan-biased seat outcomes. This research addresses COI definitions by establishing a data-driven framework using demographic and socioeconomic variables from U.S. Census data to generate map ensembles with and without COI-weighting. These indicators include income, education, ethnicity, and geography. Communities were modeled as statistically coherent groupings via hierarchical cluster analysis, generating a corresponding map of COIs. By representing precinct data as weighted adjacency graphs, with edge weights encoding demographic similarity, two Markov Chain Monte Carlo (MCMC) ensembles were generated: a control ensemble constrained by compactness, population and contiguity only, and an experimental ensemble incorporating a novel constraint to minimise community splits. Partisan symmetry, compactness, and community fragmentation metrics established baseline distributions against which enacted 2020 congressional maps were evaluated. The resulting COI-aware algorithm preserved over twice as many COIs compared to unweighted maps. Disparities between seat outcomes and COI splits calculated for the enacted 2020 PA congressional maps in comparison to the two algorithms showed voter dilution. However, results also showed a decrease in competitiveness and compactness for the COI-aware algorithm, highlighting a tradeoff. Partisan shifts in the COI-aware maps were greater than unweighted maps, and the COI-weighted algorithm exhibited higher computational complexity and longer run times. By using this framework, this project provides a non-partisan method of map generation and a quantitative tool for evaluating vote dilution in future redistricting processes.

Awards Won: