Districting problems

1 Introduction

Districting is the practice of splitting a (geographical) area into multiple smaller regions. This may be necessary for a variety of reasons. Consider for example a company with multiple offices. To clearly divide responsibilities, they might assign a geographical region to each office. Each customer within those regions must then be served by the respective offices. In this case, the size of the district (or at least the number of customers within it) should be in proportion to the office size, and locations are preferably in the district of a nearby office. Another particularly visible topic in districting is related to American elections. The American congress, and also many state level legislatures, are based on districts. States are divided into districts, and the voters of each of these elect one representative which will take a seat in congress/state legislature.

Unsurprisingly, the drawing of political districts has at times been controversial. Since the political preferences of areas are more or less known, this could be taken into account to draw maps that are favourable for certain political parties. As a result, there is considerable interest in ways to draw maps fairly. In particular, we will be interested in drawing districts that are in some sense compact. The idea being that maps where districts are relatively small better represent actual communities, and not specifically chosen groups of people to maximize some objective.

2 Problem Description

You are given a graph $G(V, E)$ and a required number of districts K. In the graph, each vertex $v \in V$ represents a small area of land. The edges show which of these land areas border one another, i.e., if $(v, v') \in E$, then these two vertices are neighbours. For each of the vertices, there is a population p_v, the number of people living in that area. Finally, there is a number $d_{v,v'}$ for all $v, v' \in V$, this represents the distance between two areas.

The goal is to divide the vertices into districts A_i, such that $\bigcup_{i=1}^{K} A_i = V$. Districts must be connected, i.e., if two nodes are in the same district, there has to exist a path from one to the other that only uses nodes also in the district. The division into districts should also take the following two things into account.

1. Populations of the different districts should be relatively equal.
2. Distances between areas in the same district should be small.

You are free to make your own choices on how to weigh these different goals against one another.

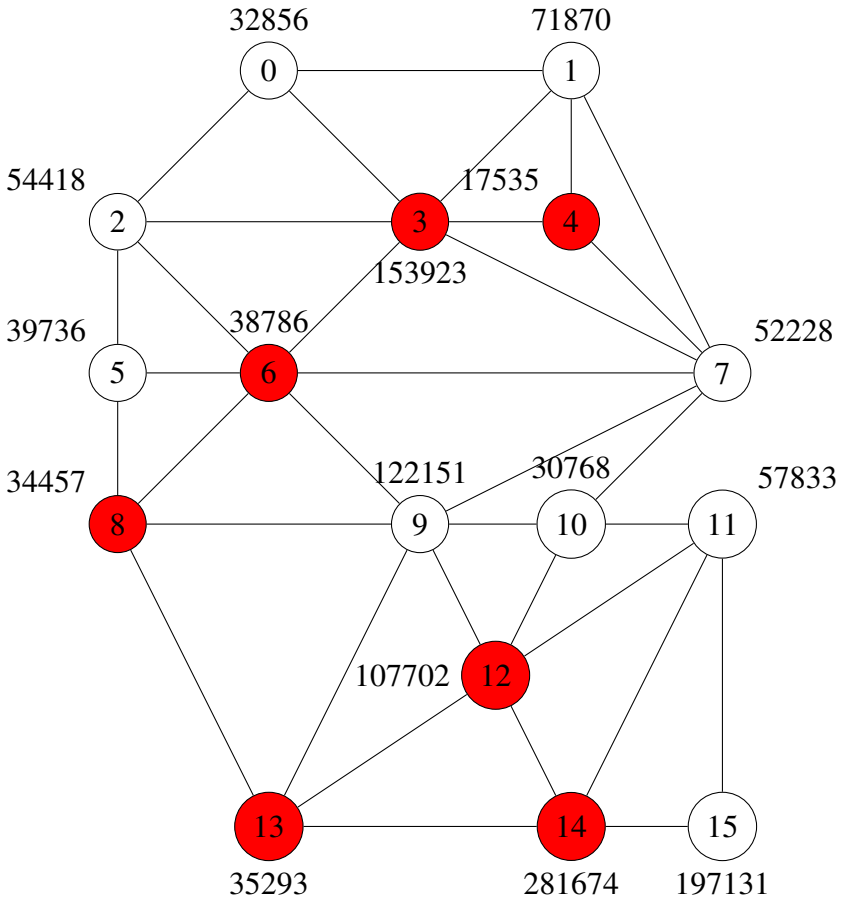

Figure 1: A graph split into two districts of relatively equal population. Numbers next to each vertex show their population. Note that for any combination of red or white vertices, there exists a path only passing vertices of that same color.

3 Data Files

Data files can be found at https://lykhovyd.com/files/public/districting/2020/. Here you will find data for every American state, at 2 levels of detail. There is county level data, with each state having as little as 5, up to 250 counties. There is also more granular census tract data, with on average 1400 tracts per state. The number of required districts per state does not seem to be given here, see Canvas for a file containing this information.

Due to the number of instances, you will likely not be interested in studying all instances you can obtain from this database. To make sure you look at a variety of instances, you should report results on census tract data for at least.

- Rhode Island (RI)
- Arkansas (AK)
- Washington (WA)

Enjoy the puzzles!