Network Analysis - Project Report

Jimi Hytönen Hanna Holtdirk Basil Mashal

23. helmikuuta 2020

1 Introduction

This project is part of network analysis course. The purpose of this project was to analyse real-world network by applying different algorithms to extract some interesting information about the network as well as get hands-on experience with network analysis. We chose to analyse California road networks because TBD - why?

TBD-something else?

In this report we will cover the technical aspects of gathering the data, tell how the work was divided, explain our network analysis and visualizations, and finally describe our conclusions.

2 Data

We used dataset from https://www.cs.utah.edu/ lifeifei/SpatialDataset.html which was collected, cleaned and formated from multiple different sources into easy-to-use format. Network's nodes were in longitude-latitude coordinate form and network's edges contained information about start node, end node and the distance between them. In addition, the site provided information about California's points of interests such as hospitals, lakes and airports.

We used pandas for data manipulation and processing as we needed to get the data into a certain form for further analysis. We used networkx for building and visualisation of the network. This was really straightforward to do as the data was in an easy to use format. In the figure below we can see the network of California's road map. TBD-image, something else?

3 Methods

For the project we divided the original question into the following subtasks:

1. What is the general structure of the road network
2. Can we find where the (big) cities are from the road network and points of interest
3. Can we learn to make predictions about the placement of the roads and places of interest

For each of these subtasks we planned what analysis was needed to answer the question. After setting these subtasks we created a timeline for the project.

For the first question we analysed...

4 Results

5 Conclusions

6 Contributors