

---

**ISyE 6740 - Spring 2021**  
**Team 350 Project Proposal**

---

**Team Member Names:** William Pickard

**Project Title:** Graph-Based Optimization  
Model for Broadband Infrastructure Expansion Planning

**Group Number:** 350

# 1 Problem Statement

In the modern digital era, reliable high-speed internet is a necessity. Research finds that increased access and usage of high-speed internet in rural areas leads to higher population and job growth and higher rates of business formation. Additionally, broadband access improves health outcomes by enabling telehealth for remote citizens<sup>1</sup>.

While 4G LTE coverage nationwide has been a great success since its rollout in 2010, it lacks the sufficient bandwidth to serve households and businesses' increasing demands.

Recognizing the substantial return on investment, state and federal programs have allocated substantial funding to improving network connectivity, especially in underserved areas. The FCC's Broadband Funding Map tracks these billions in investments. However, programs like Broadband Equity, Access, and Deployment (BEAD) and the older Rural Digital Opportunity Fund (RDoF) are far slower in expansion than their privatized counterparts.

These private-government partnerships and private expansions are contingent on clear profit incentives. Fiber expansion, in particular requires immense capital, planning, and operational costs.

This project seeks to incorporate a data-driven prioritization framework for aligning public connectivity needs with private ROI interests. By creating county-level "Opportunity Scores" and graph theory, we will identify compelling expansion opportunities along configurable planning horizons.

## 2 Data Source

The foundational data source for this experimentation will be the Ookla Global Fixed and Mobile Network Performance map. Ookla is the world leader in network intelligence, and is used by Telecom providers and regulatory bodies for network testing. The dataset aggregates results from millions of global speed measurements per month. This data provides detailed network metrics such as upload and download speeds for fixed and mobile connectivity for geographic areas roughly 610x610 meters.

To incorporate the profit incentive, we will incorporate county-level median income and population data from the U.S. Census Bureau's American Community Survey (ACS).

To add information about the current broadband infrastructure, we will use data from the FCC's Broadband Data Collection dataset, which details coverage levels for different network technologies.

To enable county-based aggregations, we will use the US Census Bureau's TIGER/Line shapefiles to form geographic data into the county baseline.

## 3 Methodology

### 3.1 Data Engineering

The first task is to combine the datasets. The Ookla data is structured into tiles sized by Web Mercator Projection. We will need to average the metrics from these tiles and project them into county boundaries using the TIGER/Line shapefiles.

From there, the unique FIPs codes can be used to match ACS income and population data and FCC broadband data to the Ookla county aggregation.

### 3.2 Data Analysis

With the combined dataset, we will perform exploratory data analysis. We will start with a correlation matrix. We hypothesize that there will be a strong correlation between high-income and high-population counties with existing broadband coverage. This would align with the known profit incentive.

We will also explore the importance of existing fiber infrastructure. The fiber backhaul is the backbone of the modern high-speed network infrastructure. While wireless backhaul technology exists, it is limited in capacity and range, and is primarily used to moderately expand network reach from an existing fiber node. We hypothesize that counties with higher fiber penetration will show persistently higher average network speeds.

We will also visualize the dataset's key variables with choropleth maps to garner general understandings of the distributions.

### 3.3 Feature Engineering

We want to establish a metric to score counties for broadband expansion potential. This "Opportunity Score" will be a weighted combination of current speeds, current coverage, population, and average income. Counties high on the index will ideally have larger, wealthier populations with low coverage and below average speeds.

The Opportunity Scores will be mapped to an undirected graph. This will model the US, with each county as a node and edges representing adjacent counties.

### 3.4 Multi-hop Planning Algorithm

We will use the Bellman-Ford algorithm to traverse the graph and exhaustively calculate the highest opportunity plans for network expansion. The Bellman-Ford algorithm is a shortest-path finding algorithm similar to Dijkstra's Algorithm, but crucially able to handle negative numbers. A shortest-path algorithm normally attempts to minimize total cost. In order to adapt this for the goal of maximizing the Opportunity Score of a multi-hop path, we simply need to make all the edge weights negative.

The action in which Bellman-Ford accounts for negative weights also requires that the algorithm is more exhaustive in its search than a greedy algorithm like Dijkstra's, which serves our purpose of collating multiple viable paths across different numbers of hops.

Since both fixed and mobile networks are fundamentally dependant on fiber backhauls, expansion must begin with existing infrastructure. As such, counties with a threshold level of fiber adoption will serve as starting points for the algorithm. For each additional hop in the desired plan, the algorithm will calculate the opportunity added by an adjacent, less-covered county.

This exhaustive search will produce all possible plans of  $N$  hops with corresponding opportunity scores, which would serve as strategic options for network planners.

## 4 Evaluation and Final Results

The primary evaluation method will be historical comparison. We will generate our optimal paths using data from prior years and compare the output plans with the actual county expansion seen in ensuing years.

We will evaluate the stability of the model by adjusting the Opportunity Score component weights. A robust model will not drastically change with minor weight adjustments.

### 4.1 Opportunities for Enhancement

Aggregating to the county level is a limitation. Certain adjacent counties may present a challenging expansion factors, such as incompatible zoning. A more viable dataset would have more granular data to more precisely assess the viability of expansion.

Network expansion planning is not a static, point-in-time consideration. Expanding the network takes time and economic considerations will change during the process. A more advanced version would incorporate time-series forecasting to predict economic growth and population changes, as well as data from the FCC Broadband Funding Map. By aligning this forecasting with expected build times, the model could provide more nuanced strategic recommendations.

## References

- [1] Sophia Campbell, Jimena Ruiz Castro, and David Wessel. *The benefits and costs of broadband expansion*. 2021. URL: <https://www.brookings.edu/articles/the-benefits-and-costs-of-broadband-expansion/>.