ECONOMIC OPTIMIZATION OF FIBER OPTIC NETWORK DESIGN IN ANCHORAGE

By

Jasen Kintner

RECOMMENDED:

Martin Cenek, Ph.D.

Todd Petersen, Ph.D.

Seong Dae Kim, Ph.D., PMP
Chair, Advisory Committee

LuAnn Piccard, M.S., PMP
Chair, Engineering, Science, and
Project Management Department

APPROVED:

Kenrick Mock, Ph.D.
Associate Dean, College of Engineering

Date

ECONOMIC OPTIMIZATION OF FIBER OPTIC NETWORK DESIGN IN ANCHORAGE

A

PROJECT

Presented to the Faculty

of the University of Alaska Anchorage

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE,

ENGINEERING MANAGEMENT

By

Jasen Kintner, B.S.

Anchorage, Alaska

May 2016

Abstract

The wireline telecommunications industry is currently involved in an evolution. Growing bandwidth demands are putting pressure on the capabilities of outdated copper based networks. These demands are being meet by replacing these copper based networks with fiber optic networks. Unfortunately, telecommunications decision makers are tasked with figuring out how best to deploy these networks with little ability to plan, organize, lead, or control these large projects.

This project introduces a novel approach to designing fiber optic access networks. By leveraging well known clustering and routing techniques to produce sound network design, decision makers will better understand how to divide service areas, where to place fiber, and how much fiber should be placed. Combining this output with other typical measures of costs and revenue, the decision maker will also be able to focus on the business areas that will provide the best outcome when undertaking this transformational evolution of physical networks.

Table of Contents

Page

List of Figures

Page

List of Tables

Page

List of Equations

Page

1 Introduction

In recent years, the telecommunications industry has seen a shift of products and services that have traditionally been served via copper cabling to products and services that require higher throughput offered by fiber optic cabling. The demand for these products and services, often called triple play services (voice, video, and data), is what is driving the industry to change from copper based networks that transmit electromagnetic signals to fiber optic networks that transmit light. As more information and new services drive growth, this bandwidth demand is expected to increase year over year.

Further to this point of demand, copper based solutions are inferior to fiber optic solutions in technical areas such as distance, bandwidth, and number of voice channels (or a measurable equivalent). Copper is an adequate solution for voice as it makes designed use of the 3000 Hz bandwidth available on the medium. Fiber optics on the other hand, provides much greater spectrum allowing for transmission of the triple play services much further and much faster.

In order for service operators to keep up with this demand, they have to retire traditional copper plant for a more futureproof solution. While some wireless solutions offer a nominal stop-gap, the demand growth is clear and wireless will not be able to sustain the growth. Fiber optic cabling as a "last-mile" access network is the future.

Currently decision makers are faced with a complex problem. They have to decide how to build out a fiber optic network while making a strong business case to do so. New physical plant builds can be complex and time consuming efforts. Often times there is a lack of expertise in-house to help guide them.

This research proposes a solution that will empower the decision maker with a viable build model and resulting metrics to help the decision making process. It also provides a framework to manage project performance. The model begins by looking at a service area, consisting of a collection of service locations, as a graph problem. Treating each service location, or sink, as a point on a map I am able to use clustering techniques, such as k-means, to group sinks together with the intent to minimize the amount of distance to each location. Once the clusters are defined, routing algorithms such as, Dijkstra's and Nearest Neighbor Search, can be used to create pathways from each service location to a source. This clustering and routing will produce a suitable local optimum with which economic analysis techniques can be applied. From here, the model creates an output of path distance that costs can be applied to. Costs and potential revenue are calculated to gain an understanding of the clustering and routing algorithm impacts. Finally, a sensitivity analysis is performed so as to empower the decision maker with a view that allows them to focus on the more critical components of cost reduction and revenue generation.

The remainder of the paper is organized as follow. I will provide background on related works and industry specific terminology in section 2. Section 3 will describe, in detail, how the model works to produce a local optimum of routed pathways. Section 4 will provide results and analysis using Anchorage, Alaska as an area to run simulations. The results of these simulations will be fed into a cost benefit model to assist in the creation of the sensitivity analysis. Finally, section 5 will conclude the paper with a short supposition and discussion about possible future works.

2 Background

In this section, I provide an overview of telecommunications fiber optic networks by describing the different physical parts of a traditional fiber optic network. In addition to this, I will show some related work that has taken place in academia and the industry. This review of work focus on two outcomes. First, it will show that the industry is attempting to solve this problem in new and creative ways. Second, this related works analysis will show where technical improvements are still achievable.

2.1 Fiber Optic Networks

In the telecom industry, fiber optic networks are generally referred to as access networks. These access networks are made up of three physical outside plant sections. These sections are feeder cable, distribution cable, and drop cable. Figure 2.1 (INET Telecom - Networking Technology Company LTD, 2008) shows a typical fiber optic access network. The feeder cable connects the electronic source(s), found in the Central Office, with fiber distribution hubs (FDH). These FDHs act as aggregation points for the rest of the network. In passive optical networks (PON), the FDH contains passive splitters that can split a single fiber hair into 4, 8, 16, or 32 more hairs. This provides the ability to remain fiber lean on the feeder cable portion of the network and fiber rich on the distribution portion. The distribution portion of the network connects the FDH to further access terminals, such as pedestals and drop closures. These pedestals and drop closures act as connection points for the last portion of the network known as the drop cable. This drop cable is typically installed at the time of service order or delivery, if a drop cable does not already exist. The drop cable is then used to connect the customer into the network. (INET Telecom - Networking Technology Company LTD, 2008)

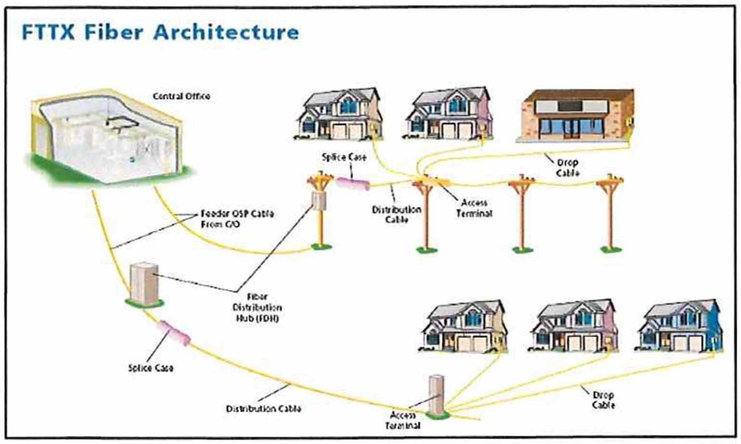

Figure 2.1 Traditional FTTx Network Layout (INET Telecom - Networking Technology Company LTD, 2008).

There are varying kinds of fiber optic deployment in service operator networks. This is commonly referred to as FTTx or "Fiber to the x" where "x" can be replaced with the destination type. Some possible types include "N" for node, "C" for curb, "P" for premise, or "H" for home. In the case of this research, FTTH or "Fiber to the Home" is the fiber architecture of focus. Business and residential locations are considered, therefore FTTP is also suitable and interchangeable.

2.2 Related Work

The study area of computational modeling of telecommunications access networks is relatively sparse at the moment. The industry has been dominated by wireless research in the past 20 years. That said, there have been some interesting concepts and ideas put forward.

One particularly interesting approach presents a novel heuristic algorithm to create a cost minimization solution to reduce total deployment costs (Li & Shen, 2009). The authors considered aspects such as maximal split ratio of fiber optics, transmission, distance, and differential distance. The problem is well defined and understood and the optimization model appears sound. They applied Weiszfeld's algorithm as a means of clustering, or sectoring, the selected service locations. While this pie shaped wedging of service locations might be useful, it ultimately will not reduce the distribution distance in the network. In some extreme cases, the solution may create unreachable areas in the network due to region and terrain constraints. Finally, this work creates straight-line connections from source to distribution hubs and from distribution hubs to sinks. This should not be used as a measure of distance as traditional wireline network routes follow indirect paths.

Another approach to this problem is presented in (Finn, 2011). This source shares a very brief overview of how Verizon is approaching the problem of converting 28 million service locations from copper based services to fiber based services. They present the idea of "modules" to analyze the network and the associated finances. The work presents a very industry focused approach. While the idea is presented, no analysis or results are shared. The paper simply suggests an approach to the problem. Finally, some of the assumptions, such as bandwidth use predictions, are underestimated when compared to common measures today.

Yet another approach uses stochastic programming to forecast demand evolution in wireless networks (Eisenblatter & Schweiger, 2012). While this paper diverges from the wireline problem presented in this research, it has some similarities in that it attempts to solve an industry problem using computational optimization. One downside to this approach is that the authors' solution suggest an optimum is not reachable given the continual change in the network. This gives the impression that forecasting cannot be performed.

All three of these works support the industry's desire to find an economic solution to deploying new physical fiber optic networks. The rest of this paper will present my approach that builds on these ideas. The goal is to provide key metrics and useful insights to the decision maker when it comes to building a new and forward-looking network.

3 Clustering, Routing, and the Model

This research is interested in programmatically clustering potential service locations with the idea of minimizing the feeder and distribution distance in the fiber optic access network. Once these clusters are formed, the density center is identified as the ideal location for the FDH. After all the FDH locations have been identified, routing algorithms will be used to generate edges or pathways to connect the FDHs to the corresponding sources, thus producing the feeder network. With the feeder network formed, the application then creates the distribution network for each FDH by finding the shortest routes from the sink to the FDH. This step results in the creation of the distribution network. This model does not include the drop portion of the network as it is typically placed at the time of first service delivery and is typically a minimal cost when compared to the large network buildout.

3.1 K-means Clustering

The first input into the model is a set of service locations identified on Earth via a latitude and longitude. The prevailing idea is to tightly group these locations in an attempt to minimize the distribution distance in the network. Research into appropriate clustering algorithms leads to the classic k-means clustering algorithm presented by J. MacQueen in 1967. K-means allows clustering a set of n observations into k clusters that are reasonably efficient (MacQueen, 1967). Additionally, k-means is easily programmed and is computationally economical (MacQueen, 1967). The clustering's objective is to minimize the average squared Euclidean distance from the cluster center and the observations. This technique provides a reasonable solution to clustering service locations.

FDH distribution cabinets come in counts sized as follows: 144, 288, 432, 576, 720, and 864. This means that if a 576 sized FDH is placed, it can serve 576 service locations at 100% fill capacity. These FDH counts and fill capacities are used to help determine the *k* count of clusters given *n* service locations. This resulting *k* is rounded up to the nearest whole number and provided as an input to the algorithm. Exact FDH size and fill capacity are discussed later during results and analysis.

With the clusters created and the density centers identified, a serving area is now generated. This is done by drawing boundaries around service locations belonging to the FDH. For this Fiber Serving Area (FSA) I create Voronoi diagrams (Voronoi, 1908) or polygons.

Both the k-means algorithm and Voronoi polygons are foundational to the model as they are used to create the cluster, identify the density center, define the service area, and pinpoint the future FDH location. The output of k-means clustering provides the input for routing along both the feeder and distribution portions of the network. The Voronoi polygons become the visual representation of the clusters.

3.2 Routing Distribution Points and Service Locations

The next step in the model is to take the results of the clustering operation and preform two sets of routing, one for each portion of the network. Two suitable routing algorithms are analyzed and used in the upcoming Results and Analysis section. The feeder portion of the network, connecting the source(s) to all FDH locations are treated as a traveling salesman problem. For this, the greedy and simple, Nearest Neighbor Search (Knuth, 1973) (NNS) is employed. For the distribution network, Dijkstra's algorithm (Dijkstra, 1959) is engaged to find the shortest path from the service location (sink) back to the FDH. Unlike Li and Shen's work where straight lines are used to create the network (Li & Shen, 2009), a vehicle road network is more desirable for routing in this project.

Routing along a street network gives two distinct advantages. First, common access or "rights of way" are found along nearly all roads. This provides path access for utilities, such as telecommunications networks. Routing in this manner, also yields a closer approximation to actual pathway footage. This approximation becomes critical when performing sensitivity analysis and making determination about how to best reduce costs or increase revenues around a fiber optic network build.

3.2.1 Nearest Neighbor Search for Feeder Networks

The nearest neighbor search algorithm, as presented by Donald Knuth (Knuth, 1973) provides a greedy and simple solution to connecting FDH locations along an well-organized route. Each FDH point is connect by the closest next point, where closest is a measure of distance. NNS is defined as: given a set S of points in a space M and a starting point, find the closest point in S (Knuth, 1973). The new point is set to the current vertex, the starting point is marked as visited, and the iteration is executed again. Once all points have been visited, the algorithm terminates. This results in a distance efficient feeder network that can be measured and used as an input for analysis and tuning.

3.2.2 Dijkstra's Algorithm for Distribution Networks

For the distribution network, a slightly different approach is used. Each service location is treated as a sink looking for the shortest path back to its FDH. For this portion of network I lean on the very well know Dijkstra's algorithm (Dijkstra, 1959). Once all locations, served by a distribution hub are connected, common routes are merged and total pathway distance is calculated. This action of merging and calculating provides two critical outputs; total pathway distance and fiber cable sizing. While there is nominal cost difference between cables with different counts fiber, from a planning perspective it is important to know how much fiber should be placed in a given pathway. Limiting the number of construction operations during network builds lowers labor costs.

3.3 Cost and Revenue Modeling

The remaining portion of the model involves costs and revenue. Cash flows will be created based on assumptions about cost variables such as, linear cost per foot for fiber builds and fixed costs on FDH placements. Revenue variables such as penetration rates, average revenue per unit (ARPU), and yearly growth projections, will be applied to the model as well. This will allow a picture to be formed about how design changes can impact decision making.

These cash flows will drive a measure of Internal Rate of Return (IRR). In the telecom industry, IRR is traditionally selected to evaluate different sets of projects against each other. For the purposes of this project, IRR is preferred as the sensitivity analysis measure. A base, or control, case is then established. Selected variables within the design model and outside the model are changed. Seeing the changes in the resulting IRR will allow the decision maker to assess the impacts of changing model variables.

The resulting measure is not meant to be an exhaustive analysis of all costs and revenues that go into a telecommunications business model. The selected variables and measure are intended to show how programmatic network design could have an effect on the direction of the decision maker.

3.3.1 Cost Minimization Model

With the focus of this project in mind, I have created the following cost minimization model that considers clustering and routing optimization attempts. The model is as follows:

Let:

- c_{fdh}: the cost factor for placing and FDH cabinet.
- x: observations $\{x_1,\ldots,x_n\}$. This is the set of service locations.
- n: *total number of service locations*
- k: seed centroids $\{k_1,\ldots,k_m\}$ for k-means clustering.
- m: total number of centroids
- c_{fiber}: the cost factor per linear foot of fiber placement.
- F: the set of FHD locations from k-means clustering. $\sum_{i=1}^{n} \min_{k} \|x_i - k\|^2$
- N: the distance resulting from Nearest Neighbor Search algorithm.
- s: the source location for electronics
- D: the distance resulting from Dijkstra's algorithm.

Objective: minimize

$$Cost = c_{fdh} \sum_{i=1}^{n} F_i + c_{fiber} \left(\sum_{i=1}^{n} N(F_i, s) + \sum_{i=1}^{n} \sum_{j=1}^{m} D(F_i, x_j) \right)$$

Equation 1 Cost Minimization Objective Function.

Constraints:

- $n, m, s \geq 1$
- $m, s \leq n$

In this objective function, costs originate from two primary sources. The first source is the count of cluster centers produced from k-means clustering. A cost factor is applied to the count of centers and the cost is produced. The second cost source is derived from the total summed distance of the feeder and distribution network. Similar to the clustering portion of the cost equation, a linear footage cost factor is applied to the summed distance. The function has couple constraints. Count of service locations (n), count of centroids (m), and source of electronics (s) must all be greater than or equal to 1. In addition, both m and s must be less than n. Violation of these constraints would result in no network design and thus no resulting costs.

4 Results and Analysis

In this section, I will evaluate the proposed model using Anchorage, Alaska as the subject location. All simulations are run with ESRI's ArcGIS Network Analyst supported by a custom built application in Python.

An initial test, with base conditions and assumptions, is performed to demonstrate the model output. After this, nine more runs are executed to test for variance of the output. Descriptive statics are then calculated to quantitatively describe the model product. From this, one of the ten runs is selected to act as the control point in the sensitivity analysis.

During the sensitivity analysis, I will adjust the base conditions so as to study the uncertainty in the result. This will help demonstrate how changes to model inputs will empower the business leader when making decisions about how to proceed with the proposed design. This sensitivity analysis will conclude this portion of the report.

4.1 Model Variables and Assumptions

As previously described, the study area of this project is Anchorage, Alaska. The area is not for the entire Municipality of Anchorage. The study areas is shown as the blue cross hatched area in Figure 4.1. It can be further described by the following bounds. The region is bounded on the north by Joint Base Elmendorf Richardson and confined on the south by Potters Marsh. The east boundary is defined by the Chugach Mountains. The west boundary is defined by the Cook Inlet.

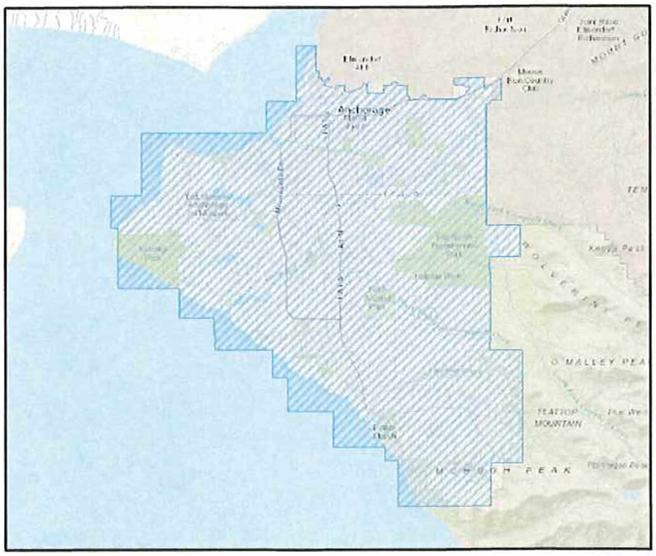

Figure 4.1 Study Area, Anchorage Alaska.

4.1.1 Revenue Assumptions

This study area consists of approximately 115,000 service locations. These service locations are further classified as business or residential. Of the 115,000 service locations, approximately 90,000 locations are residential. The remaining 25,000 are businesses. Understanding this distribution of service locations allows us to assign revenue potential values to each. The 25,000 business locations have an annual revenue value of $176M. The 90,000 residential locations are valued at $1320 per unit (ARPU) or $123M per year.

For the purposes of this project, the following revenue assumptions are made and can be seen in Table 1. The maximum annual revenue potential for the study area is $299M. This is a sum of the business and residential revenue. The max penetration rate is set at 30%. Penetration rate is defined as the percentage of service locations that will become paying customers. A 30%

penetration rate is an arbitrary number but should be taken to indicate there is a competing company in the area. Looking at penetration rate and maximum annual revenue, we can see that this gives us potential annual market revenue of $89M. In this project, I use Gompertz's curve to describe the rate of change in market penetration. For year one, I expect a 2% penetration, worth $1.8M. For year two, I would expect that penetration to grow to 6%, worth $5.4M.

Table 1 Revenue assumptions.

Revenue Potential	$299,889,688
Penetration	30%
Projected Max Revenue	$89,966,906
Year 1 Projection (2%)	$1,799,338
Year 2 Projection (6%)	$5,398,014

Looking closer at the projected revenue curve, shown in Figure 4.2 we can see that year 10 is when the revenue starts to approach its maximum. The x-axis is representative of years and the y-axis shows annual revenue. In the telecom industry, this curve is typically expected for the deployment of new products or services. Aggressive marketing campaigns can speed this up but this curve shape typically holds over the long term.

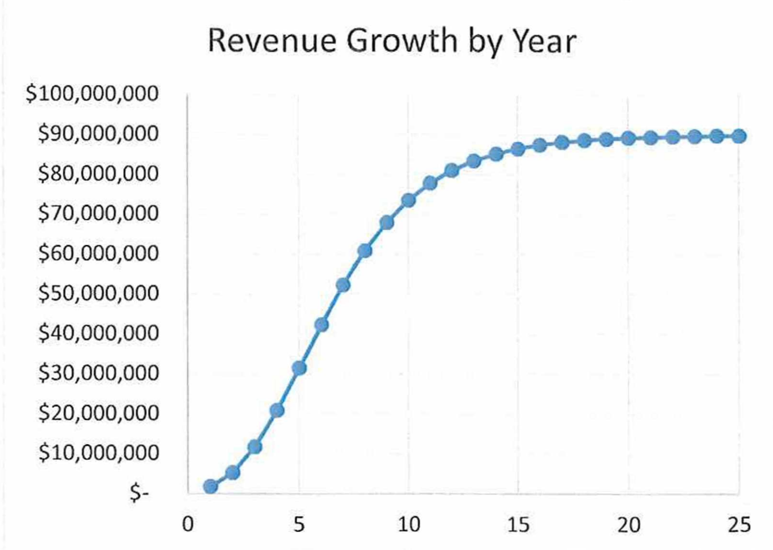

Figure 4.2 Revenue Growth Curve.

4.1.2 Cost Assumptions

For the cost portion of the project, two main areas of focus are studied. These two areas are the fixed cost of placing a Fiber Distribution Hub and the average linear cost per foot for placing fiber. Both of these costs can be subject to regional and seasonal modifiers.

FDH costs range from $8,000 per cabinet for a 288 sized cabinet to $19,500 per cabinet for 864 sized cabinets. This cost is not linear in growth from the 288 up through 864. The first three sizes are traditionally cheaper as they can be placed in rights-of-ways without incurring additional permitting costs. For sizes 720 and 864, there are additional costs associated due to extra permitting work. This change in cost is reflected in Table 2.

Table 2 FDH fixed costs.

FDH Size	Fixed Cost
288	$ 8,000
432	$ 9,500
576	$ 11,000
720	$ 18,000
864	$ 19,500

In addition to these FDH fixed costs, I consider the average cost per foot for placing fiber. There are three ways that telecommunications wireline, such as fiber, are placed. They can be placed aerially along telephone or power pole lines. They can be placed in conduit or duct systems. Finally, they can be directly buried in the ground. Each of these placement methods have different costs associated with them. For this project, I looked at an industry average cost and determined that $15 per foot is suitable. Future work on the model could include weighted routing based on specific costs for each of the three situations previously described.

4.2 Initial Simulation

For the initial simulation, I selected one source location to serve all subsequent FDH's and 115,000 service locations. An FDH fill rate of 80% is used to allow for future growth in an FSA. For example, I look to serve 460 locations from a 576 sized cabinet. Cabinet locations are placed at the density center. NNS is designated to provide the routing for the feeder portion of the network. Dijkstra's is used for routing on the distribution side of the network.

The first run produced the results found in Table 3. Looking at the results, we see that the total count of FSAs, or FDHs, is 246. The feeder distance equates to 889,643 feet. The total distance is 4,805,756 feet. This gives us a total of 3,916,113 feet for the distribution portion of the network. Figure 4.3 shows the model output for the feeder portion of the network. The purple lines represent the proposed feeder route. The red triangles show the proposed FDH locations. Figure 4.4 shows a very zoomed in portion of the a couple FSAs. Like the previous figure, the purple lines represent feeder network and the red triangles show the FDH locations. The blue dashed lines show the distribution portion of the network. The double red lines represent FSA boundaries. The output of this first run shows a robust result that will be useful for further economic and sensitivity analysis.

Table 3 Initial simulation results.

RUN	FSA Count	Feeder Distance	Total Distance
1	246	889643	4805756

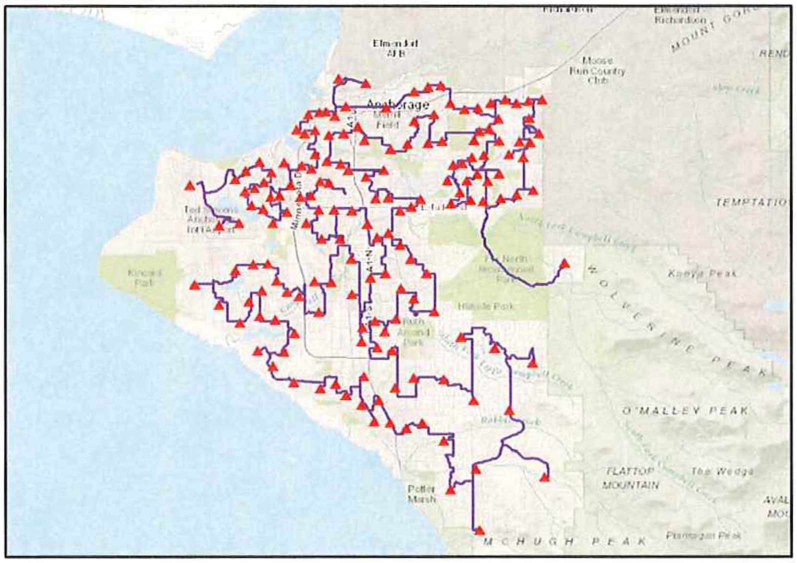

Figure 4.3 Feeder Network from Initial Run

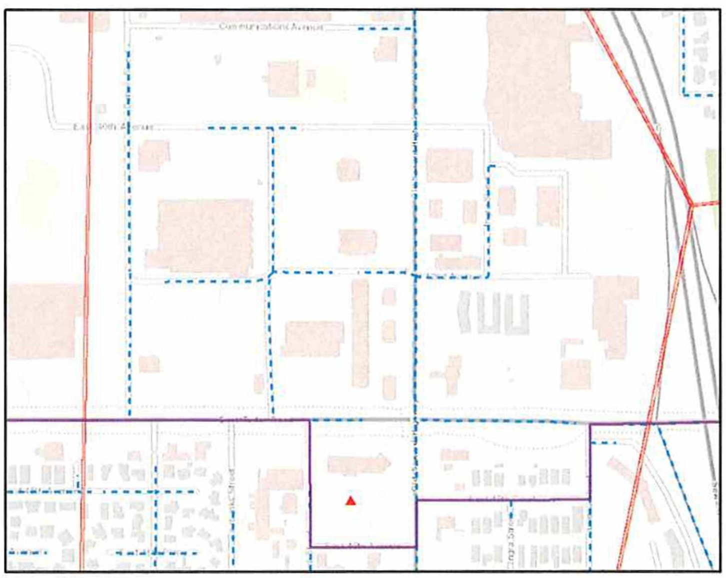

Figure 4.4 Portion of the Distribution Network from Initial Run

4.3 Control Base

The next step in the project is to create a control point around the model output from which a sensitivity analysis can be performed. For this, nine more runs were produced using the same base conditions that were described in the previous section. A 576 sized FDH, 80% fill, 115,000 service locations, and one central source of electronics. Table 4 shows the results of runs one through ten.

Table 4 Results of ten simulation runs.

RUN	FSA Count	Feeder Distance	Total Distance
1	246	889643	4805756
2	244	873269	4815824
3	245	859339	4781810
4	245	885351	4807318
5	243	893282	4814954
6	246	895951	4809482
7	248	849809	4857058
8	245	944775	4842226
9	242	955214	4822292
10	244	875672	4800710

This table provides insight into how the clustering and routing are attempting to find optimums. By examining the FSA count, we can see results ranging from 242 to 248. Understanding that k-means is not guaranteed to produce a global optimum, this is to be expected. The same holds true for the total distance output, we see a range from approximately 478,000 to approximately 485,000. From these results I performed a descriptive statistics analysis to better understand the data set. These statistics are shown in Table 5.

Table 5 Descriptive statistics on control runs.

Mean	4815743
Standard Error	6699.218437
Median	4812218
Mode	#N/A
Standard Deviation	21184.7888
Sample Variance	448795276.7
Kurtosis	0.862784309
Skewness	0.65834576
Range	75248
Minimum	4781810
Maximum	4857058
Sum	48157430
Count	10
Confidence Level(95.0%)	15154.68497

The statistics shows that the module will produce a fiber optic network design that will total 4,815,743 feet plus or minus 15,155 feet with a confidence level of 95%. This provides a solid estimate of how much footage will be involved in a fiber optic build in Anchorage. With Kurtosis and Skewness both very small and under a one, we can consider this set of data to be Gaussian in distribution. For the purposes of the upcoming sensitivity analysis, I selected RUN #2 as the control case.

4.4 Sensitivity Analysis

With the control case defined, I set out to show how different variables inside and outside the model can affect decision making with regards to a project or portfolio of project related to a fiber optic build. For the sensitivity analysis, I selected FDH size, FDH Fill (or Saturation) Rate, Fiber Build Cost per foot, Penetration, and ARPU for residential service locations. This gives a decent mix of model inputs, cost variables, and revenue variables in the analysis. Future work could include additional variables that could impact service delivery, lifecycle costs, or route selection.

The base case is defined as an FDH size of 576. The FDH Saturation is set to 80%. The cost per foot is set at $15 per foot. Penetration is at 30% and APRU for residential service locations is $1320. IRR is calculated at year eight. This is a typical telecom industry measure. The product from this project is estimated to have a useful life of up to 25 years as that is considered the useful life of fiber optic cables. Year eight represents possible changes to electronics that could affect products and services. For this base case, we end up with an IRR of 20.86%.

Each variable is given a lower and upper limit of percentage change. In order to stick to the appropriate FDH sizes, the percentage changes are distinct at 50%, 75%, 125%, and 150%. This ensures FDH sizes of 288, 432, 720 and 864 respectively. FDH Saturation is limited to ranges between 80% and 120% in 10% increments. This ensures the FDH does not saturate past 100% using 80% as the base condition. The rest of the variables, Cost per foot, Penetration, and APRU range from 50% to 150%.

Table 6 shows the IRR results of the sensitivity analysis. The lowest IRR result is found when reducing Penetration from 30% to 15% (or a 50% reduction). This takes year 8 IRR from 20.86% to 7.14%. Conversely, the biggest gain in IRR comes when the fiber build Cost per foot is reduced from $15 per foot to $7.5 per foot. This resulting IRR is 36.74%.

Table 6 Sensitivity analysis results displaying changes to IRR.

% Change	FDH Size	FDH Sat.	Cost(per ft)	Penetration	ARPU (Res)
50%	20.60%		36.74%	7.14%	16.00%
75%	20.76%		27.12%	14.84%	18.53%
80%		20.59%	25.67%	16.15%	19.01%
90%		20.73%	23.10%	18.60%	19.95%
100%	20.86%	20.86%	20.86%	20.86%	20.86%
110%		20.97%	18.88%	22.96%	21.74%
120%		21.04%	17.11%	24.93%	22.60%
125%	20.74%		16.30%	25.88%	23.02%
150%	20.69%		12.75%	30.25%	25.04%

Looking at plot of these results, found in Figure 4.5, it is easy to see how each change in variable impact IRR. Interestingly, the change in FDH Sizes or FDH Saturation Rates has the least amount of impact on overall IRR. This is an important discovery, as build decisions can often be decided by deployment size or technique. This graphs shows that the decision maker would be better served by minimizing fiber per foot costs and maximizing penetration rates.

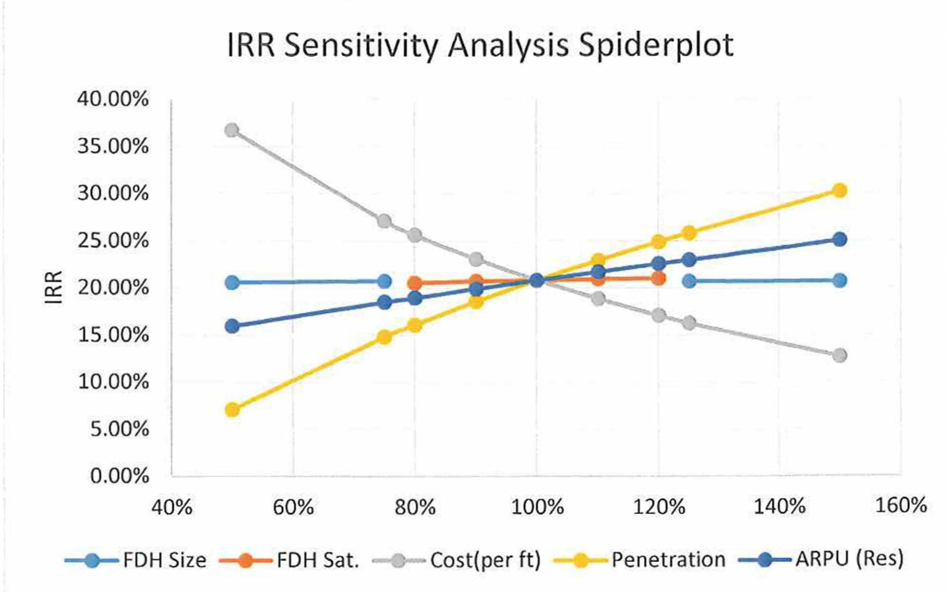

Figure 4.5 Spiderplot for Fiber Optic Network Modeling.

5 Conclusion

As the telecom industry continues to explore different solutions to deploying fiber networks, applying this programmatic approach to network design can provide a solid directional foundation to preparing a project approach. This project has shown a new approach to solving this problem.

Using clustering and routing techniques, I have shown that deployment costs can be defined and understood from a strategic perspective. The cost minimization function allows the decision maker to vary input parameters and analysis the results.

The current implementation of the model provides opportunity for applying other clustering algorithms. Using k-means, the expectation is that all service areas, or clusters, are rigidly defined as one size. Exploring other clustering algorithms such as affinity propagation or DBSCAN, the output could result in dynamic sizing of clusters. These different size clusters would mean different sized FDH's in the network. This might increase the need for inventory of various sizes of FDH cabinets, but this might also provide for more efficiency in routing, thus reducing costs in the long run.

On the topic of routing, the model can also easily support the evaluation of other routing algorithms. For example, applying Prim's algorithm to create minimum spanning tree routes, might find efficiencies not currently realized with Nearest Neighbor Search or Dijkstra's algorithms. Again these efficiencies in routing are related directly to reducing expected costs.

The ability of testing different approaches is success in itself. While it was not discussed in the paper, traditional methods of exploring alternatives would require a team of engineers and months of time. Typically, the result of this traditional design process would produce a sub-optimal solution that was driven by tribal knowledge and gut feeling. The design output resulting from this project takes a fraction of that time. For the Anchorage Alaska study area, a simulation run was reduced to two hours from start to finish. This time savings provides the decision maker with a huge advantage. They can alternate variables and observe the outcome without the worry of losing precious time waiting for a response.

Overall this model and resulting project open up new avenues for approaching fiber optic builds. Telecom service operators can apply a mathematical and scientific approach to building futureproof networks. This will help them better plan and organize the workforce to help ensure a successful portfolio of projects.

6 References

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 1*, 269-271.

Eisenblatter, A., & Schweiger, J. (2012). Multistage stochastic programming in strategic telecommunication network planning. *Computational Management Science, 9*(3), p. 303.

Finn, J. (2011). Investment optimization planning for the access network. *National Fiber Optic Engineers Conference and Optical Fiber Communication Conference and Exposition (OFC/NFOEC)*, (pp. 1-3, 6-10).

INET Telecom - Networking Technology Company LTD. (2008). FTTx Tutorial. Ha Noi: INET Telecom - Networking Technology Company LTD.

Knuth, D. (1973). *The Art of Computer Programming* (Vol. 3). Addison-Wesley Professional.

Li, J., & Shen, G. (2009). Cost Minimization Planning for Greenfield Passive Optical Networks. *IEEE/OSA Journal of Optical Communications and Networking*, 17-19.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1*, 281-297.

Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie de formes quadratiques. *Journal für die reine und angewandte Mathematik, 134*, 198-287.

Zhou, M., & Chen, Z. (2005). Network planning from telecom provider's perspective. *Network Architectures, Management, and Applications III. 6022.* SPIE Proceedings. doi:10.1117/12.639938