Efficiently Growing Rochester Bike Infrastructure Based on Network Theory

Nate August, Margaret Brennan, Blaise Derenze, Quang Hoang
University of Rochester Goergen Institute for Data Science
May 1st 2022

***Abstract*—In this work, we present a network approach for providing recommendations for growing bicycle infrastructure from an existing network in the City of Rochester. We propose improvements on previous work that begins with a fully-connected network between common origins and destinations (points of interest). Our work allows for the consideration of existing infrastructure through prioritized routing and iterative pruning. From this process we corroborate some bike infrastructure priorities identified by subject matter experts at the City of Rochester and propose other, new priority-roadways for consideration.**

I. Introduction

The City of Rochester has seen sustained growth in downtown areas and many neighborhoods in addition to the increase in demand for better transportation [6]. In response, the City of Rochester has a vision to increase the percentage of Rochester residents that rely on biking or public transit as their primary mode of transportation to 50% by 2050. Therefore the City has a strong commitment to expand mobility as they understand the importance of a multi-modal transportation system to the quality of life and economic competitiveness of the community [6]. In order to achieve this, the city is seeking recommendations for the enhancement of the bike infrastructure based on network theory. As they are creating the future bike infrastructure plans, the City of Rochester is taking a data science perspective into consideration on where to prioritize bike infrastructure growth across the city. Using network theory to provide insight on the future growth of the bicycle infrastructure there are many points of interest from the city to research. The basis of the network is built off of points of interest that the city prioritizes. This includes locations such as parks, convenience stores, museums, and different census data that are evenly distributed across the city to ensure all residents of Rochester are able to use the infrastructure. Once points of interest are decided, past literature provided insight on how network theory can create insightful recommendations. The goal is for these recommendations to build and enhance Rochester's bike infrastructure and the Rochester community..

II. Related Work

Previous work on both Rochester and applying network theory to bicycle infrastructure were considered. The City of Rochester has commissioned several analyses of bicycle infrastructure over the past few decades as part of various planning processes. In 2018 the Bikeable City Report was created as part of the Comprehensive Access Mobility Plan (CAMP), a project focused on planning the Rochester infrastructure of 2034 [1]. In this plan, existing biking infrastructure was inventoried and evaluated according to Level of Traffic Stress (LTS) [5], accident frequency, and community input. The recommended priorities identified included replacing sharrows with dedicated bike lanes in high traffic areas, prioritizing connected corridors and emphasizing investment in low income areas. These priorities are addressed within our framework later in the report. The authors also suggest specific roadways to prioritize for new infrastructure which are used in this work as a baseline against which new network-based approaches can be compared.

The primary inspiration for our graph methodology is Growing Urban Bicycle Networks by Szell et al. [2]. In this work, the authors investigate strategies for growing synthetic bike networks from scratch. They outline a four part process for prioritizing new bike infrastructure. The process begins by identifying points of interest (POIs) that represent common origins and destinations for travel. Next, candidate connections are identified through 32 weight triangulation–a process of connecting all POIs without overlapping paths. The candidate connections are ordered by the growth strategy and finally routed through the urban street network to identify concrete routes.

Szell et al., explore three prioritization strategies: 1) the betweenness strategy prioritizes new connections based on the number of shortest paths between POIs that route through the connection; 2) the closeness strategy begins with the most central node and iteratively connects the most central adjacent nodes; 3) a random strategy is also employed to serve as a baseline. Using a variety of metrics that are discussed in greater detail in the METRICs section, the authors conclude that the different strategies optimize different metrics, but the betweenness strategy is strong for fast area coverage and moderate connectedness and directness.

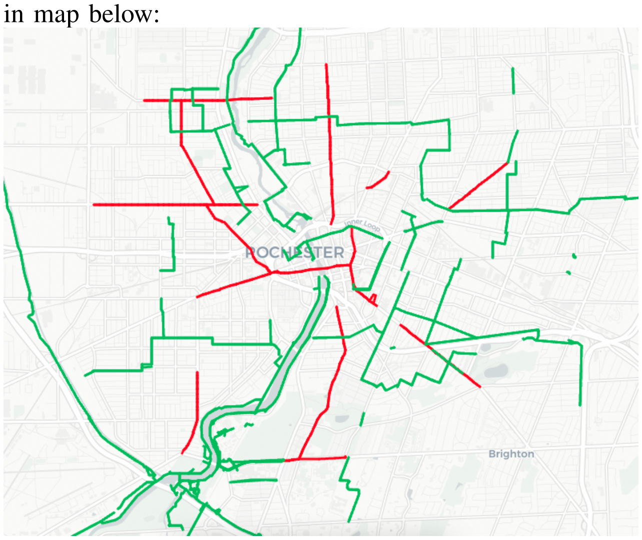
Upon applying these strategies to 62 major cities across the globe, the authors observe a tendency for network metrics to not increase monotonically and even decrease during the early stages of network growth. The worsening effect is especially pronounced for metrics that take into account well-connectedness, e.g. directness and efficiency. It was argued that this is due to a growth strategy that does not directly incentivize the development of connections between components. Once the network is sufficiently developed, the routes between disconnected subgraphs are filled and those become consolidated, at which point the metrics improve rapidly.

III. Street Network Data

Data on the Rochester urban street network was sourced through OpenStreetMap (OSM) [3] and via bicycle infrastructure maps supplied by the City of Rochester. In response to conclusions in the Bikeable Cities Report, published research [4], and input from subject matter experts at the City of Rochester, only three types of existing bike infrastructure were considered: 1) protected cycletracks, 2) bike boulevards – streets with low traffic volumes and speed limits below 25 miles per hour and 3) off-street bicycle paths and trails. Unprotected bike lanes and sharrows were excluded due to their limited effects on encouraging new bicycle ridership. In many instances, it was necessary to update OSM to match the bike infrastructure maps provided by the City of Rochester. Table below shows the query tags that were used to identify various infrastructure types.

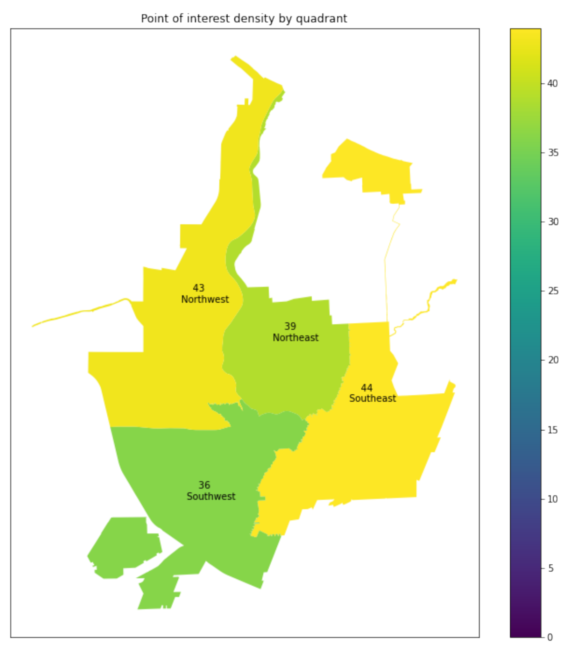
Bike Infrastructure	OSM Tags
Cycletrack	cycleway=track
Bike boulevard	highway!=cycleway, surface!=path, bicycle=designated
Off-Street Bike Path	highway=cycleway and highway=path|footway, bicycle=yes|designated, surface != ground|unpaired

New infrastructure proposals from Section 8. Priority Projects in the Bikeble Cities Report [1] were included as a baseline scenario against which new approaches could be compared. These proposals include converting sharrows to dedicated infrastructure in some very specific roadway segments and more general recommendations of creating continuous corridors along other roadways. The proposals were interpreted to produce the proposed infrastructure shown in map below:



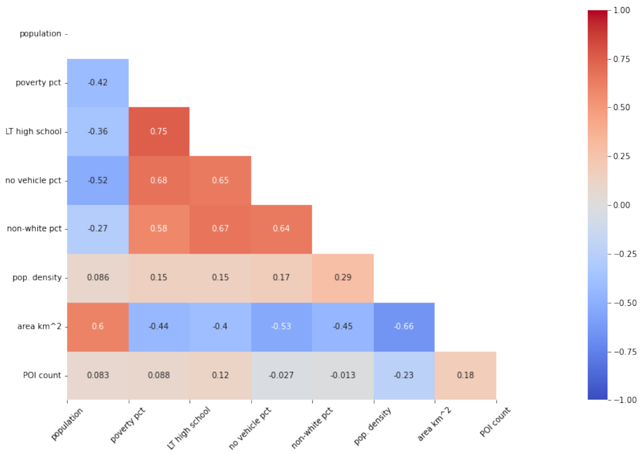
IV. Points of Interest

The points of interest are extremely important to build an optimal bike network that would be ideal for the City of Rochester. Points of interest are the basis of the network as these are the nodes that are connected in the graph, hence determining the future bike network. These points are places that the city recommends and prefers accessibility to by Rochester residents. After research and discussion with our sponsors we used the City of Rochester Commercial Corridor Business Data. This was published as a part of the Rochester 2034 Full Plan and is able to highlight the occupied businesses in the Rochester Area. This data contained roughly 1800 points which is a very large amount for our points of interest. Therefore, we reduced the points down to 86 by looking at clusters of points and intersections. I would take one point from a cluster instead of including all the businesses in one area. The places represented are parks, museums, convenience stores, schools, and other businesses around Rochester that are in populated areas. This helps the city's goal of increasing public transportation as more areas and destinations are accessible by bike path. In order to make sure there is no bias and fair representation, the points of interest were distributed across the four quadrants of Rochester. From the map and chart below, the amount of points of interest are fairly balanced across the quadrants in terms of area and population.



Quad	POI Count	Population	POI per capita	POI per km^2
Southeast	44	61858	0.000711	0.000161
Northeast	39	47043	0.000829	0.000238
Southwest	36	56396	0.000638	0.000123
Northwest	43	55050	0.000781	0.000147
Total	163	220347		

In addition, the density of points of interest against different census statistics was evaluated to see if there are any correlations between where the points of interest are placed and socioeconomic data.



The bottom row exhibits the correlation between the amount of points of interest (POI) with different census statistics. All values are ranging from (-.23, .18) indicating that there is very little to no correlation between population, poverty percentage, non white percentage, and no vehicles owned.

However, when looking at different census statistics, median income level stands out as one that should be favorited to the bike network. As bikes are a low cost of transportation, the bike network should be able to prioritize low income areas to ensure this network is accessible to those who need it. Comparing the current bike infrastructure to the low income areas there appears have no increased amount of bike lanes in low income area versus the medium to high income areas. To combat this, in our project we implemented a selection method that will remove a certain percentage of the POI's in high income areas. This helps the algorithm connect more low income areas and create more bike lanes in them.

V. Methods

In our research, we consider the existing Rochester infrastructure when generating networks as opposed to generating networks from scratch. Outright replicating the methods used by Szell et al. would present us with the following issues. 1) A comparison of a network generated using the methods in Growing Urban Bicycle Networks to the existing infrastructure without any changes would be too dependent on the selection of POIs– a greater the amount of POIs along the routes of existing infrastructure would lead to a greater similarity independent of geometry or metrics. 2) Failing to consider existing networks could lead to redundant paths, or paths whose inclusion in the network are dependent on the the inclusion of redundant paths.

A. Growing New Infrastructure from Existing

Below is the pseudocode for the algorithm used in Growing Urban Bicycle Networks using the betweenness growth strategy:

The parameters for the algorithm are the drivable roads C, an OSMnx directed graph; list of points of interest ids pois, a python List; prune factor, a float 0-1.

```
function GTFromScratch(C, pois, prune_factor)
    poi_dists← List()
    A← DiGraph()
    A.nodes ← pois
    for all i, j ∈pois do
        if i ≠ j then
            pois_dists.append((i, j, dist =C.shortest_path_dist(i, j)))
    pois_dists.orderby(dist)
    for all (i, j, d) ∈pois_dists do
        if not_intersects((i, j),A.edges) then
            A.add_edge(i, j, weight = dist)
    BW← List(e.betweennessfore ∈A)
    bw_threshold ← BW.quantile(prune_factor)
    for all e ∈A.edges do
        if e.betweenness ≤ bw_threshold then
            A.remove_edge(e)
    B← DiGraph()
    B.nodes ←C.nodes
    for all (i, j) ∈A.edges do
        B.add_edges(C.shortest_path_edges(i, j))
    return B
```

B. Betweenness Pruning

Szell et al. consider a driveable map of the city as possible areas to add bike paths. In our analysis we also include Rochester's bikeable paths and take the union of the two graphs. In order to encourage routing through existing infrastructure, we apply a discount on the weights of currently bikeable edges by a routing factor, where routing_factor = 0 corresponds with not weighting existing routes and routing_factor = 1 corresponds assigning a weight of zero to all existing routes(not practical). Because Szell et al. do the pruning on an abstract graph of POI distances, it is possible even with a routing factor that some of the existing routes are pruned. After routing, any existing infrastructure pruned is added back to the network.

In this code the parameters for the algorithm are again the drivable roads C, an OSMnx directed graph; list of points of interest ids pois, a python List; prune factor, a float 0-1. Now we've added the current bike infrastructure E, an OSMnx directed graph; route factor r, a float 0-1.

```
function GTWithExisting(C, pois, prune_factor)
    C ← C ∪ E
    for all e ∈C.edges do
        if e ∈ E then
            e.weight ← e.weight − e.weight∗ route_factor
    B ← GTFromScratch(C,pois, prune_factor)
    B ← B ∪ E
    return B
```

C. Iterative Pruning

This solution encourages new infrastructure to use existing routes, but pruning away the existing routes just to add them back after a pass rings unintuitive. In order to fully consider the existing network in our growth strategy, we may do pruning iteratively on the full (non-abstract) graph and recalculate the betweenness for all edges after each removal. By doing pruning after the routing on the full graph, we can mark the components of the existing network such that they won't be removed. By recalculating the betweenness at each step and keeping the existing infrastructure in at each step, we further consider it in our growth strategy. For both this iteration of the algorithm and the next we use the same parameters as the previous one.

```
function GTWithExistingIter(C, E, pois, prune_factor, route_factor)
    C ← C ∪ E
    for all e ∈C.edges do
        if e ∈ E then
            e.weight ← e.weight − e.weight∗ route_factor
    poi_dists← List()
    A← DiGraph()
    A.nodes ← pois
    for all i, j ∈pois do
        if i ≠ j then
            pois_dists.append((i, j, dist =C.shortest_path_dist(i, j)))
    pois_dists.orderby(dist)
    for all (i, j, d) ∈pois_dists do
        if not_intersects((i, j),A.edges) then
            A.add_edge(i, j, weight = dist)
    for all (i, j) ∈A.edges do
        B.add_edges(C.shortest_path_dist(i, j))
    B← DiGraph()
    B.nodes ←C.nodes
    B ← B ∪ E
    while len( B.edges ≥ len( C.edges ∗ prune_factor do
        BW← List(e.betweennessfore ∈B)
        e_min ← e ∈B|e = min(BW) ∩ e ∉E
        B.remove(e_min)
return B
```

D. Hybrid Pruning

A disadvantage of the above approach is that to get betweenness values in the full graph, we must consider all nodes and not just POIs. Therefore we introduce a hybrid method, that does a majority of the pruning using the betweenness strategy from the paper, and does the remaining with an iterative approach. More precisely, for prune factor p, we do betweenness growth with prune factor 2*p, add the existing infrastructure back into the graph, then do iterative betweenness pruning as if prune factor was p. Since the full greedy triangulation introduces 405 km of bike paths, we typically use a prune factor less than 0.5, meaning that the majority of the pruning occurs using the abstract graph.

```
function GTHybrid(C, E, pois, prune_factor, route_factor)
    C ← C ∪ E
    for all e ∈C.edges do
        if e ∈ E then
            e.weight ← e.weight − e.weight∗ route_factor
    poi_dists← List()
    A← DiGraph()
    A.nodes ← pois
    for all i, j ∈pois do
        if i ≠ j then
            pois_dists.append((i, j, dist =C.shortest_path_dist(i, j)))
    pois_dists.orderby(dist)
    for all (i, j, d) ∈pois_dists do
        if not_intersects((i, j),A.edges) then
            A.add_edge(i, j, weight = dist)
    BW← List(e.betweennessfore ∈A)
    bw_threshold ← BW.quantile(prune_factor*2)
    for all e ∈A.edges do
        if e.betweenness ≤ bw_threshold then
            A.remove_edge(e)
    B← DiGraph()
    B.nodes ←C.nodes
    for all (i, j) ∈A.edges do
        B.add_edges(C.shortest_path_dist(i, j))
    B ← B ∪ E
    while length( B.edges) ≥ length( C.edges) ∗ prune_factor do
        BW← List(e.betweennessfore ∈B)
        e_min ← e ∈B|e = min(BW) ∩ e ∉E
        B.remove(e_min)
return B
```

E. Software

All analysis, computation, and visualization, for this work were completed in Python 3.9 [7] and JupyterLab [11]. The library OSMNX [12] was leaned on heavily for querying and managing OpenStreetMap data. Cenpy was used for accessing and summarizing census statistics. NetworkX [13], iGraph [14], Pandas [9] and Numpy [10] libraries were used in data manipulation and analysis. Matplotlib [8] and Plotly [15] were used for visualization. All the code is available at https://github.com/BakerAugust/roc-bike-growth.

VI. Metrics

We utilized a variety of metrics to assess the quality of the grown bike networks under different settings. The metrics cover both geometric attributes and real world quantities, giving us different ways to assess the proposed network. These metrics include density, size of largest component, directness, coverage, length, cohesion, global efficiency, and local efficiency.

Resilience: The ideal network needs to be well-connected and robust, i.e. unforeseen disruptions to some path would not severely degrade the whole network. We measure this through 2 different aspects: - Density: the number of paths present out of all possible paths. Having higher density equates to having readily available alternatives to disrupted paths.

$$Density = \frac{m}{n(n-1)}$$

with m the number of edges and n the number of nodes

- Size of largest component: Since it is possible to achieve high density with many disconnected component, we use the size of largest component as a proxy to check if the network on a whole is well-connected

Directness (Szell 2022): Measures the average ratio of euclidean distance over network distance of every edge in the graph. Having a high directness would mean it is easy to traverse between different points, i.e. not many detours.

$$Directness = \frac{\sum_{i \neq j} d_E(i.j)}{\sum_{i \neq j} d_G(i.j)}$$

Coverage (km2) (Szell 2022): Measures how much of Rochester's physical area is included. A buffer of 500 m is drawn around the constructed edges and the area within is considered covered by the network.

Length (km): In a similar vein to coverage, we measure the total physical length of the network. The length is calculated assuming an unidirectional graph.

Cohesion: cohesion is calculated as coverage adjusted for the number of disconnected components. Ideally, we want the network with high coverage but low disconnected components.

$$Cohesion = \frac{Coverage}{n\_components^{\beta}}$$

with β a regularization term.

Efficiency: Another measure commonly seen for networks is efficiency, which can be calculated as global or local efficiency. - Global: measures the sum of pairwise efficiency (defined as the inverse of distance), normalized by the sum of pairwise efficiency when the graph is fully connected.

$$Efficiency_G = \frac{\sum_{i \neq j} \frac{1}{d_{ij}}}{n(n-1)}$$

with n the number of nodes and d_{ij} distance between 2 nodes.

- Local: the average of global efficiencies on the subgraphs of each node and its neighbors (the node itself is not included in this subgraph)

VII. Results

With the aim of pragmatism, we focus our analysis on settings that produce additional infrastructure that are equivalent in length to the proposed biking plan. We offer these as alternative plans of development optimized for a number of metrics. Additionally, there are more aggressive growth plans, i.e. longer paths, that can be considered when going beyond the proposal plan length limit, as well as less developed networks for comparison. The prune factor that gives similar additional length is around 0.07. Comparing different scenarios across prune factors, we see that the differences are minimal between having income adjustment on POIs and otherwise.

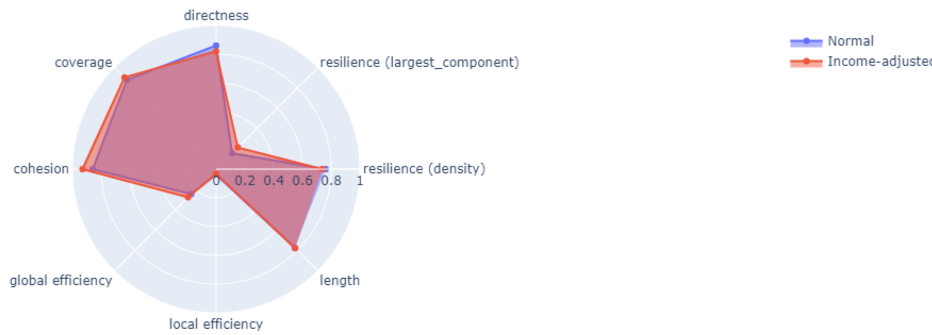
1. Not Income Adjusted Results

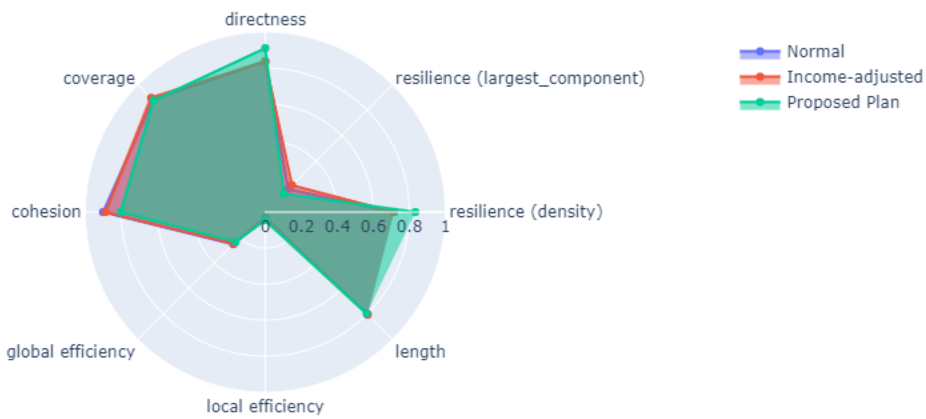
Prune Factor	Density	Component Size	Directness	Coverage	Length (km)	Cohesion	Global Efficiency	Local Efficiency
.01	.42	97	.94	79.35	126.66	0.55	0.25	.06
.02	.39	97	.91	80.25	130.59	0.55	0.25	.05
.03	.35	97	.89	82.11	135.99	0.55	0.24	.04
.04	.33	107	.87	83.89	139.99	0.62	0.26	.04
.05	.32	107	.86	84.02	142.30	0.61	0.25	.04
.06	.31	122	.84	84.23	146.17	0.63	0.25	.03
.07	.3	123	.84	84.24	147.64	0.64	0.25	.03
.08	.29	143	.82	86.03	152.18	0.69	0.26	.03
.09	.24	275	.69	88.54	157.43	0.51	0.27	.03
.10	.23	275	.70	88.97	162.58	0.57	0.26	.02
.11	.21	283	.75	89.59	166.03	0.56	0.28	.02
.12	.2	538	.57	89.81	168.14	0.59	0.35	.02
.13	.18	624	.63	90.76	177.72	0.57	0.41	.02
.14	.16	633	.62	91.01	180.71	0.4	0.37	.02
.15	.16	684	.61	91.22	183.96	0.45	0.38	.02

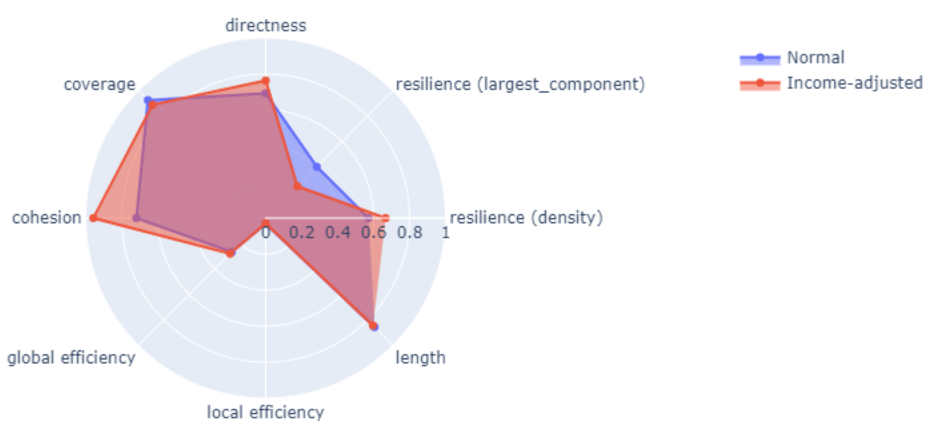
2. Income Adjusted Results

Prune Factor	Density	Component Size	Directness	Coverage	Length (km)	Cohesion	Global Efficiency	Local Efficiency
.01	.42	97	.94	79.35	126.66	0.55	0.25	.06
.02	.38	97	.9	80.47	130.95	0.56	0.24	.05
.03	.37	97	.9	81.2	133.08	0.57	0.24	.04
.04	.34	100	.87	83.29	137.22	0.59	0.24	.04
.05	.31	146	.83	86.28	143.19	0.66	0.28	.03
.06	.32	97	.87	81.65	143.5	0.62	0.24	.02
.07	.3	144	.84	85.83	148.6	0.63	0.25	.04
.08	.28	264	.72	86.9	154.56	0.71	0.31	.03
.09	.28	171	.76	85.1	155.07	0.68	0.28	.03
.10	.24	370	.57	86.38	155.2	0.42	0.28	.02
.11	.23	441	.62	88.75	166.2	0.57	0.36	.03
.12	.19	372	.59	95.03	170.9	0.31	0.27	.04
.13	.19	513	.64	86.29	169.63	0.32	0.34	.04
.14	.18	525	.67	90.05	170.13	0.27	0.27	.02
.15	.14	570	.61	95.65	183.5	0.17	0.33	.04

Comparative Radar Charts For Prune Factor of 0.5, 0.7(proposal equivalent), 0.9 respectively

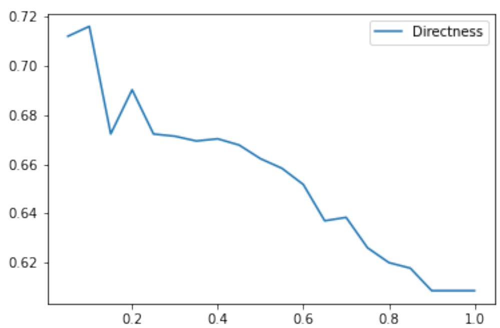


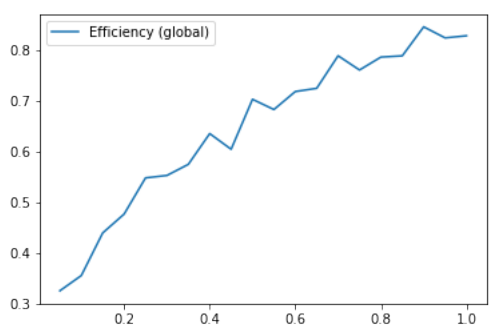




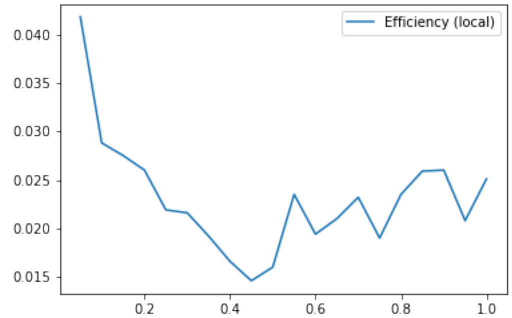
A. Monotonic Rate of Change in Metrics

We have adopted a mathematical framework to synthesize a robust and well-connected bike network for the city of Rochester based on existing infrastructure. Contrary to results obtained when growing bike networks from scratch by Szell et al., we observe a monotonic increase in network metrics in Rochester when growing the network from existing infrastructure. As shown below for a full growing process (prune factor from 0 to 1), both Directness and Global Efficiency decrease and increase, respectively, at a relatively stable rate.





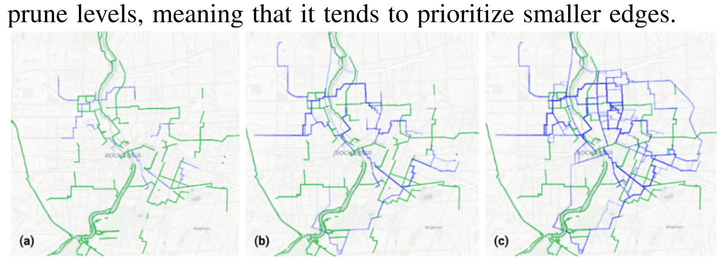
We argue that this might be the effect of growing on top of existing infrastructure, which was not taken into account in the aforementioned paper. The imperfect existing network both gives a different starting point compared to growing from scratch and has an influence in the growing process, due to that, the metrics behave in a much more predictable manner. - Small world phenomenon It is notable from the radar graphs that local efficiency tends to be low. This is consistent across different scenarios (growth strategies, POIs setting, prune and route factors).



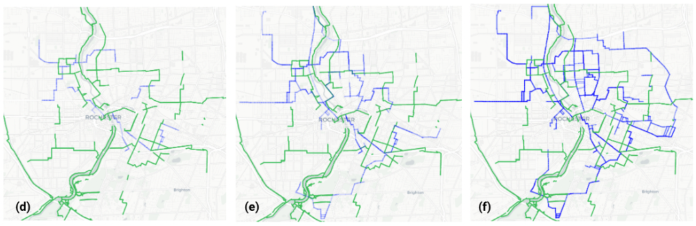
However, this is desirable given a high enough global efficiency. Having high efficiency on a global scale signals that the bicycle network is well-connected, however, with respect to economic and other real world constraints, it is not desirable to have redundant connections for the sake of connectedness. In this regard, local efficiency is a proxy for measuring cost efficiency and a low local efficiency with high global efficiency indicates a small-world network, where most points are not directly connected, but can be traversed through a neighboring point.

B. Pruning Strategies

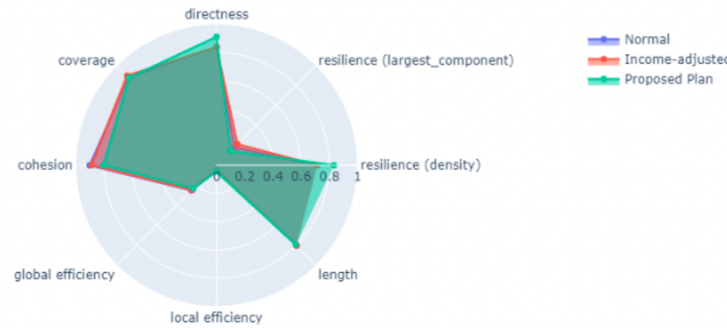
At 98km, we observe some overfitting to the POIs using the betweenness pruning strategy compared to the hybrid pruning strategy at equal lengths. At 60km, we observe this to a lesser extent. And at 20km, the networks are nearly identical. The similarity of the networks at lower millage added demonstrates the efficacy of using a routing factor of 0.25. We retained edges that would have higher betweenness given the existing infrastructure. In addition to the effects of the routing factor, more similar graphs are somewhat expected as 1) at a lower level a majority of the pruning is done on the abstract graph and 2) with a smaller graph, we are less likely to observe overfitting to a single POI, even using the 'betweenness' strategy. It is notable that the hybrid pruning strategy generally generates smaller networks for equal prune levels, meaning that it tends to prioritize smaller edges.



Results from the infrastructure growth algorithm algorithm using the hybrid pruning strategy, evenly-distributed POIs and route factor of 0.25 under pruning levels increasing from 0.05 (a) to 0.15 (b) and 0.35 (c). New infrastructure is shown in blue with existing infrastructure in green. Total of added infrastructure is 20 km (a), 60 km (b), and 98 km (c).



Results from the infrastructure growth algorithm algorithm using the betweenness pruning strategy, evenly-distributed POIs and route factor of 0.25 under pruning levels increasing from 0.05 (d) to 0.1 (e) and 0.19 (f). New infrastructure is shown in blue with existing infrastructure in green. Total added infrastructure is 20 km (d), 60 km (e), and 98 km (f).

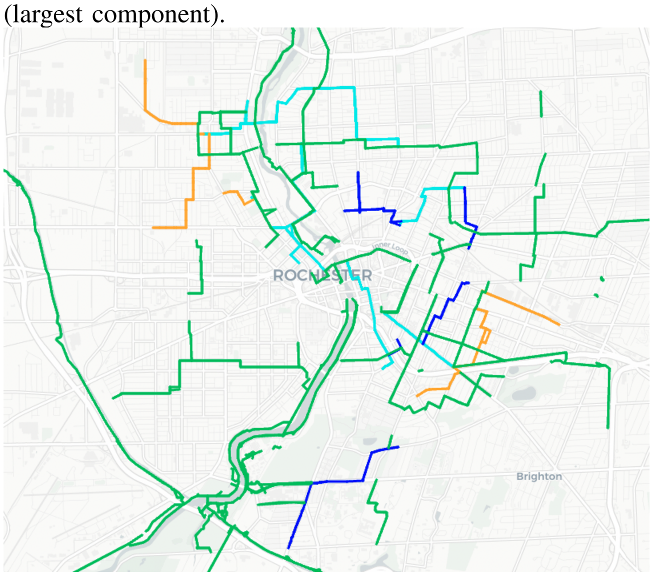


Scaled metrics comparing even POIs (blue) and income-adjusted POIs (red). Income-adjusted POIs have higher cohesion, slightly higher directness and resilience (density), but lower resilience (largest component)

Prune Factor	Density	Component Size	Directness	Coverage	Length (km)	Cohesion	Global Efficiency	Local Efficiency
Even POIs	0.3	123	0.84	84.24	147.64	**0.64**	0.25	0.03
Income-adjusted POIs	0.3	**144**	0.84	**85.83**	**148.06**	0.63	0.25	0.04

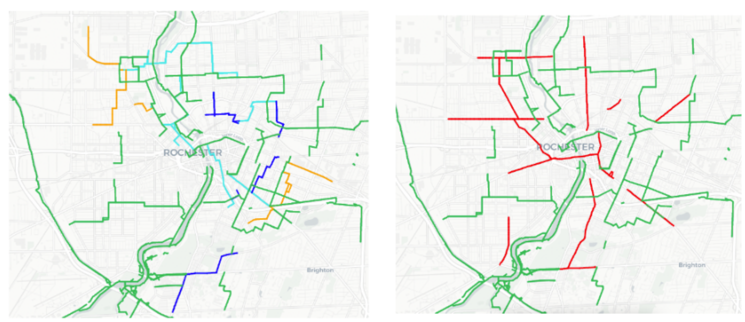
C. Income Adjusted POIs

To evaluate the impact of the income-adjusted POIs, 25 kilometers of new infrastructure were proposed under two scenarios: 1) using evenly-distributed POIs and 2) using income-adjusted POIs. The results show considerable overlap in the infrastructure produced. The two scenarios share 15.5 km of their total 25 km. The even-income scenario produces additional infrastructure in the northwest and southeastern areas while the income-adjusted scenario adds more infrastructure just northeast of the city center and south of the city. Looking at the metrics, we can see the income-adjusted POIs have higher cohesion, slightly higher directness and resilience (density), but lower resilience (largest component).



Comparing 25 km of new infrastructure under evenly-distributed POIs (orange) and income-adjusted POIs (blue). Infrastructure suggested in both POI scenarios (light blue) and existing infrastructure (green) also shown.

D. Comparing with Proposed Infrastructure



Left: 25 km of new infrastructure under evenly-distributed POIs (orange) and income-adjusted POIs (blue). Infrastructure suggested in both POI scenarios (light blue) and existing infrastructure (green) also shown.
Right: Priority infrastructure from Section 8. Priority Projects in the Bikeable Cities Report (red) and existing infrastructure (green).

Comparing the proposed infrastructure from the Bikeable City Report ("proposed") with even ("even-POI") and income-adjusted POIs ("adj-POI") 25-km output, there are areas of agreement and conflict. All scenarios prioritize Monroe Avenue, northern Joseph Avenue and a north-south corridor running through downtown. Adj-POI finds further agreement with proposed on the expansions to Elmwood Avenue infrastructure and central segments of Joseph Avenue. Even-POI shows a bit more activity around and north of Lyell Avenue, but does not suggest very much infrastructure on Dewey or Lyell Avenues directly, possibly because it is routing through existing infrastructure on N Plymouth and Fulton Avenues. Both POI scenarios suggest improved connections to existing infrastructure along the Genesee river northwest of the inner loop. Looking at suggested infrastructure in the 60-km even-POI scenario (Figure b), most of the proposed infrastructure is covered. Notably the proposed additions to Webster Avenue, W. Main Street and South Avenue are not found, even in 60-km even POI. The lack of emphasis on this area could be due to surrounding existing infrastructure that limits the number of shortest paths traveling along these routes.

VIII. Recommendations for Rochester

A. Incremental Expansion of the Network

The observed missing substantial threshold noted by Szell et al. (2022) means that additional development on top of the existing infrastructure would show monotonic changes across metrics. However, to take advantage of the small world phenomenon, it is also advisable to bring the bike network to a highly developed state to increase overall cost efficiency.

B. Priority Roadways

Roadways that are found both in the proposed infrastructure and the results of this work should be considered high priority as their importance is supported by both subject matter expertise and network theory. These roadways include:

- Monroe Avenue between Culver Road and Howell Street
- Elmwood Avenue between Mount Hope Avenue and S. Goodman Street
- Driving Park Avenue between La Grange Avenue and Saint Paul Street
- Joseph Avenue between Cumberland Avenue and Norton Street
- Some connection between North Street and Central Park, either Davis and Scio Streets or Portland Avenue

Additional priorities that have been identified in this work focus on creating continuous corridors connecting Rochester's rich existing bike infrastructure:

- State Street between Andrews Street and Smith Street
- Smith Street between Lake Avenue and Saint Paul Street
- South Clinton Avenue between Gregory Street and East Broad Street
- Stone Street between East Broad Street and East Main Street
- Saint Paul Street between East Main Street and Andrews Street

IX. Additional Work

The approach and recommendations presented in this work could be improved in several ways. A major challenge of this approach is the sensitivity of results to the selection of points of interest. This could be studied further to understand which roadways are prioritized across many different POI specifications. Furthermore, investigation of how to better bias network expansion towards areas with lower income or lesser access to vehicles when fewer POIs are present in those areas could be helpful for advancing the cities goals around equitable development. It may also be beneficial to consider other data sources in this approach. Data on accident frequency and road width is available and has important implications for the feasibility of infrastructure development, but were not included in this work.

The "iter"/"hybrid" approaches don't consider POI locations which might shift them towards part of the graph with more nodes, although the hybrid method didn't appear to significantly impact the generation. In our approach to the "hybrid" method, we simply did abstract pruning for 2x prune level, then iterative pruning. There potential in more creatively mixing pruning abstractly with pruning on the full street network than simply doing them in series. These iterative approaches could also be further optimized in their recalculations of betweenness for a faster runtime.

X. References

1 https://www.cityofrochester.gov/article.aspx?id=21474837742
2 https://arxiv-export3.library.cornell.edu/pdf/2107.02185
3 https://wiki.openstreetmap.org/wiki/ResearcherInformation
4 Dill, J and McNeil, N. Four Types of Cyclists? Examining a Typology to Better Understand Bicycling Behavior and Potential. Transportation Research Record: J. of Transportation Research Board, 2387, 1, pp 129-138. 2013. https://doi.org/10.3141/2387-15
5 Mekuria, M. C., Furth, P.G. and Nixon, H. Low-Stress Bicycling and Network Connectivity. Mineta Transportation Institute Publications. 2012.
6 https://www.cityofrochester.gov/camp/
7 G. Van Rossum, F.L. Drake, "Python 3 Reference Manual" Scotts Valley, CA: CreateSpace, 2009.
8 J.D. Hunter "Matplotlib: A 2D graphics environment" Computing in science amp; engineering, 9(3), pp.90–95, 2007
9 "Data structures for statistical computing in python", McKinney, Proceedings of the 9th Python in Science Conference, Volume 445, 2010.
10 C. R. Harris, K. J. Millman, S. J. van der Walt, "Array programming with NumPy" Nature vol. 585, pp. 357–362
11 T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, "Jupyter Notebooks – a publishing format for reproducible computational workflows", Ebook: Positioning and Power in Academic Publishing: Players, Agents and Agendas pp. 87-90, 2016
12 G. Boeing, "OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks." Computers, Environment and Urban Systems. 65, 126-139, 2017
13 A. A. Hagberg, D. A. Schult and P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX", in Proceedings of the 7th Python in Science Conference (SciPy2008), Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008
14 G. Csardi, T. Nepusz "The igraph software package for complex network research." InterJournal, Complex Systems, 1695. https://igraph.org. 2006
15 Plotly Technologies Inc. Title: Collaborative data science Publisher: Plotly Technologies Inc. Place of publication: Montréal, QC Date of publication: 2015 URL: https://plot.ly