Expanding Transit Access Through Street Network Investments

Matthew Wigginton Bhagat-Conway
Department of City and Regional Planning
University of North Carolina at Chapel Hill
223 E Cameron Ave, CB #3140, Chapel Hill, NC 27599 USA
Email: mwbc@unc.edu
ORCiD: 0000-0002-1210-2982

Word count: 4388 + 0 tables (250 words per table) = 4388 words

Submitted 2025-08-01

ABSTRACT

Evaluating changes to transit service using accessibility metrics is increasingly common. However, these metrics are almost always used to evaluate changes to the transit service itself. Almost every transit user is also a pedestrian at some point when accessing or egressing from a stop. In this article, I apply an automated algorithm to identify locations in the pedestrian network where adding an additional link would improve transit accessibility. I then evaluate these locations in terms of their impact on accessibility to jobs via transit, combining both the pedestrian network and the transit service to evaluate the real-world impact of these changes.

Improving accessibility via transit through street network improvements is an attractive proposition, because unlike service improvements street network improvements have minimal on-going costs. Furthermore, many pedestrian network improvements (e.g. crosswalks or sidewalks) are relatively low cost compared to capital investment in transit service.

Keywords: last-mile, transit, pedestrian environment

INTRODUCTION

Accessibility metrics are increasingly used in transportation planning (*1*, *2*), and particularly in public transit planning (*3*–*9*). These metrics measure the potential access to opportunities afforded by the transportation system—in their simplest form, this might be something like "a person at this location can reach 200,000 job opportunities within 60 minutes on public transportation."

Most often, accessibility metrics are used in transit planning to evaluate the impacts of changes to transit service. However, the vast majority of transit users walk to transit (e.g., *10*), meaning the pedestrian environment around stops is also very important to understanding transit access. Modern accessibility analysis tools model the walk to transit as well as the ride time, often using OpenStreetMap data, so the pedestrian environment is not completely discounted. However, it is rare to model transit access impacts of changes to the pedestrian network.

Much post-war development in the United States is characterized by disconnected street networks, cul-de-sacs, and sparse pedestrian infrastructure. This type of street network design can significantly reduce the catchment area of bus stops. For instance, Figure 1 shows a 20-minute walk from two bus stops in Raleigh, North Carolina, USA. The dark line indicates the theoretical maximum catchment if one could walk in a straight line in any direction. The shaded area shows the actually reachable area. On the left, a stop in downtown Raleigh is shown; since downtown has a well-connected, gridded street network, the reachable area approaches the theoretical maximum. On the right, a stop in suburban north Raleigh is shown; the reachable area is significantly smaller.

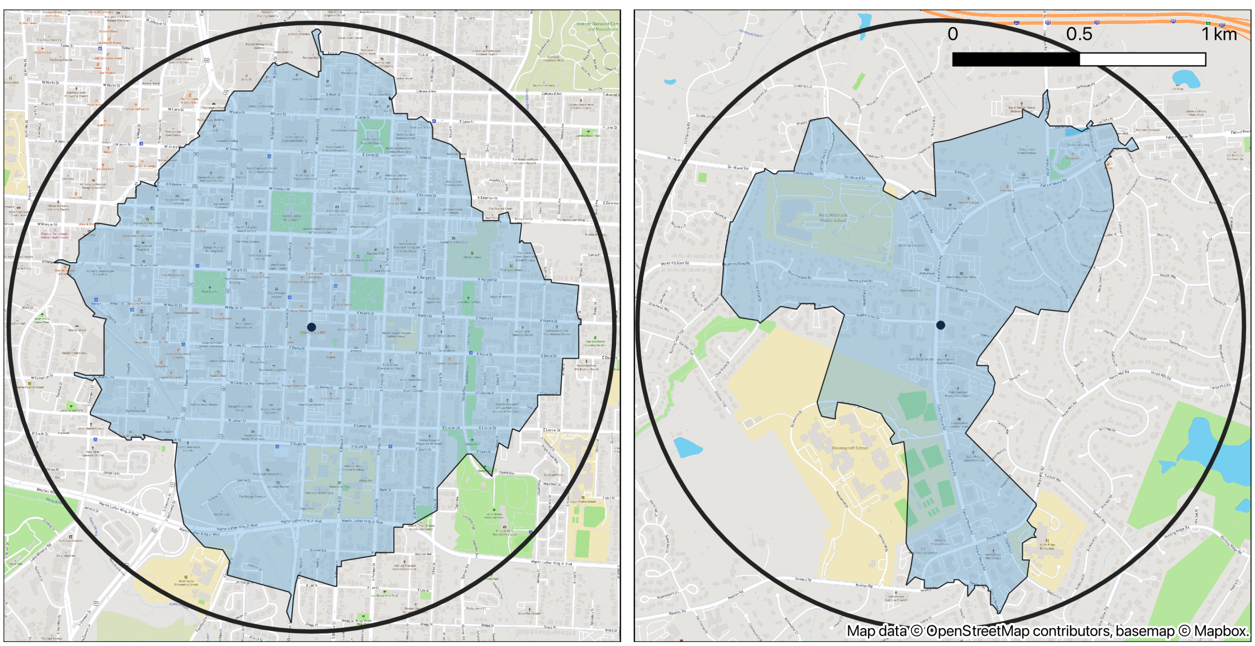

Figure 1: Catchment areas of two bus stops in Raleigh, North Carolina, USA

Bhagat-Conway, Compiano, and Ivie (*11*) introduced an algorithm to identify "low-hanging fruit" missing links in the pedestrian network. Specifically, this "Missing Links" algorithm identifies locations on the street network that are close together geographically, but far apart in network distance. It then scores those links based on their contribution to accessibility.

In this article, I apply this algorithm to identify potential links that could improve access to public transit stops, and then evaluate the most-highly-ranked links using a traditional accessibility metric generally applied to transit service changes.

LITERATURE REVIEW

A prerequisite for using public transit is that transit stops must be readily accessible from both the origin and the destination of the trip. Most transit users walk to access transit. Walking distances vary, but many transit agencies assume that riders will walk up to 400 m to access transit, possibly more for higher-quality transit (*12*). A recent review article found widely varying typical walking distances to transit for different types of service and in different geographic areas (*13*). Even within a single metropolitan area, walking distances vary based on the built-environment context as well as rider characteristics (*14*).

Considering first- and last-mile barriers in transit accessibility analyses is very important. Different methods for calculating catchment areas produce very different results, both in terms of the catchment areas identified (*15*) and in terms of overall acessibility results (*16*). However, from a ridership perspective, Guerra et al. (*17*) found that different catchment areas did not have strong effects on predictive power in a ridership model.

Lahoorpoor and Levinson (*18*) studied the impact of changes to the pedestrian environment on access to rail transit in the Sydney metropolitan area. They investigated the impacts of adding additional entrances and exits to rail stations on access to the local community, and found that catchment areas and residents/jobs near transit in some areas could be significantly increased through additional station access points.

METHODS AND DATA

I use data from the Research Triangle region of North Carolina, USA, to identify locations where additions to the pedestrian network could increase transit access. I source GTFS data from the five major transit operators in the region: GoDurham, GoRaleigh, GoTriangle, GoCary, and Chapel Hill Transit.

I source pedestrian network data from OpenStreetMap. Sidewalk representations in OpenStreetMap vary somewhat in both quality and in how they are mapped (separately from the road vs. as a tag on the road itself) (*19*). For this reason, most OpenStreetMap-based routing engines take a fairly lenient approach in terms of pedestrian path generation: they allow walking on any road where pedestrians are assumed to be legally allowed (i.e. most roads, excepting freeways/motorways and those explicitly tagged as not allowing pedestrians).

Since this work focuses on improvements to the pedestrian network, I want to use a network of links where pedestrians are likely to feel comfortable walking, rather than all places they are legally allowed to walk. To approximate this network, I filter the OpenStreetMap data to only ways that have a highway type of

- footway,
- cycleway,
- pedestrian,
- track,
- sidewalk,
- service,
- road,
- steps,
- path,

- crossing, or
- residential,

as well as those that are tagged as having sidewalks or are explicitly tagged as allowing pedestrians. This leads to a somewhat disconnected network in suburban areas. This disconnected network is a better representation of the lived experience of pedestrians in suburban areas of the Research Triangle.

I then apply the algorithm of Bhagat-Conway et al. (*11*) to this network. This algorithm identifies locations on the network that are geographically close together, but far apart or disconnected entirely in network space, and ranks them using an accessibility metric. In this work, I identify links that are no more than 100 m long, and avoid a walk of at least 500 m. As Bhagat-Conway et al. did, I remove very small "islands" in the OSM network with fewer than 10 vertices, as these are more likely to represent data errors than actual disconnected locations.

This process identifies hundreds of thousands of potential links in the region, though many provide effectively the same access. I use the deduplication method described in (*11*) to reduce this to a more manageable amount. I then score these links using an accessibility metric. Tolerance for distance to transit is usually assumed to be between 400 m and 1 km; I take the upper end of this range and use the an accessibility metric based on stops accessible within 1 km. I weight this by the number of vehicles serving the stop, as well as the number of people who the new link serves. Specifically, the overall access to transit provided by the network is

$$\sum_i \sum_s p_i w_s f(d_{is}) \tag{1}$$

where i is an index of all nodes in the graph, p_i is the population assigned to each node (based on the 2020 Census block-level population and the same assignment technique used in (*11*)), w_s is the weight assigned to stop s, and d_{is} is the network distance from node i to stop s. $f(d_{is})$ is the impedance function, which in this case is simply

$$f(d_{is}) = \begin{cases} 1 & d_{is} < 1 \text{ km} \\ 0 & \textit{otherwise} \end{cases} \tag{2}$$

The weight assigned to a stop is the number of vehicles that serve that stop on a typical weekday, so stops that are served more frequently will have a higher impact on overall accessibility. Similarly, stops that serve a larger number of people will have a greater impact on overall accessibility. I rank the identified links based on the impact each link individually has on the accessibility in Equation 1.

The accessibility algorithm requires that all origins and destinations be associated with a node, so I insert all stops in each of the GTFS into the graph. If they are near an existing node, I assign the stop to that node. If they are within 20 meters of an existing edge, but not near an existing node, I insert a node into that nearby edge. Lastly, if they are not near any edges (e.g. stops that are on streets without sidewalks), I add two nearly-coincident nodes, and a "phantom" edge between them. The algorithm only identifies missing links between edges, so this phantom edge allows the algorithm to connect these stops to the network.

The Research Triangle network is much larger than the Charlotte network used by Bhagat-Conway et al. (*11*), with 470,000 vertices and 580,000 edges (as opposed to 140,000/160,000). In initial testing, I found that identifying, deduplicating, and scoring links in this large network had unsatisfactory performance. However, since pedestrian accessibility is only affected by the network in the local area, I solve this problem by partitioning the network into 24 sub-networks in a regular grid.

To avoid edge effects, I buffer each grid cell by 1.1 km. Thus, any link within 1.1 km of a cell boundary will be more than 1.1 km crow-flies from the edge of anothe sub-network. I only retain links where at least one end is in the main part of the grid cell, rather than in the buffer, to avoid finding spurious links or scores caused by edge effects. If at least one end of the link is more than 1.1 km from the edge of the grid cell, and the links are no more than are 100 m long and based on transit stops within 1 km, there is no way a portion of the network more than 1.1 km away could affect the link or its score.

It is important to differentiate access to transit from access *via* transit. Access to transit measures the ability to get to a bus stop, while access via transit measures what destinations you can actually get to using the street network and transit system. For instance, providing access to a bus stop that is served by a bus that does not go to many destinations is much less valuable than providing access to a bus stop that connects to a large portion of the city's transit network.

Links ranked highly by the Missing Links algorithm have a strong impact on accessibility *to* transit, but accessibility *via* transit is generally more interesting. Evaluating the accessibility via transit for such a large number of links is computationally chanllenging. Instead, I take the 100 links that are ranked most highly, and I create a modified version of the input OpenStreetMap data that includes these links. I then evaluate accessibility to jobs within 60 minutes via transit, to understand how these links that increase access to transit also increase access using transit.

I use the `r5r` package (*3*) to perform accessibility via transit calculations, using the GTFS data described above as well as the modified OpenStreetMap pedestrian network data. I calculate the number of jobs accessible within 60 minutes, based on the median travel time for departures between 7:00 and 9:00 AM on a weekday. I retrieve data on block-level job locations from the 2022 US Census Bureau Longitudinal Employer-Household Dynamics.

By default `r5r` connects a stop to closest street segment, as long as there is any street segment within 1.6 km. For this project, I modified `r5r` (and the Conveyal R^5 routing engine that powers it) to instead connect stops to segments that are within 20 m, to match the logic used in finding missing links. This way, `r5r` will correctly calculate the accessibility impact of proposed links that connect a stop to a nearby street (for instance, a sidewalk connecting to a stop along an arterial that does not have pedestrian infrastructure). I also modified `r5` to retain all islands with 10 or more vertices, as was done in the link identification algorithm, rather than its default of 40 or more.

RESULTS

After deduplication, the algorithm identified 42,000 potential new links in the pedestrian network. Of these, 16,000 have an access-to-transit score of more than 0, indicating that they potentially improve access to the transit network. Identifying and deduplicating the links took 78 minutes on an M1 Macbook Pro with 16 GB RAM.

Impacts of a single link

The first step in the algorithm scores scored links based on the number of stops they provide access to, while the second step evaluates the aggregate impacts of high-scoring links on transit access. One of the links that scored highly in the first stage is shown in red in Figure 2a. This link is a crossing of Glenwood Avenue in Raleigh in front of a commercial and residential development on the northeast side of the street. Glenwood Avenue is a fairly typical high-speed suburban arterial with minimal pedestrian infrastructure. There is relatively frequent bus service along this street (15 minutes, among the most frequent service in the Triangle). For residents of the development on the northeast side of Glenwood Avenue, it is fairly easy to access the bus stop for service traveling northwest. However, there is no infrastructure to allow them to safely cross the street to travel southeast. The link identified by the algorithm is shown in red, and would provide such a crossing.

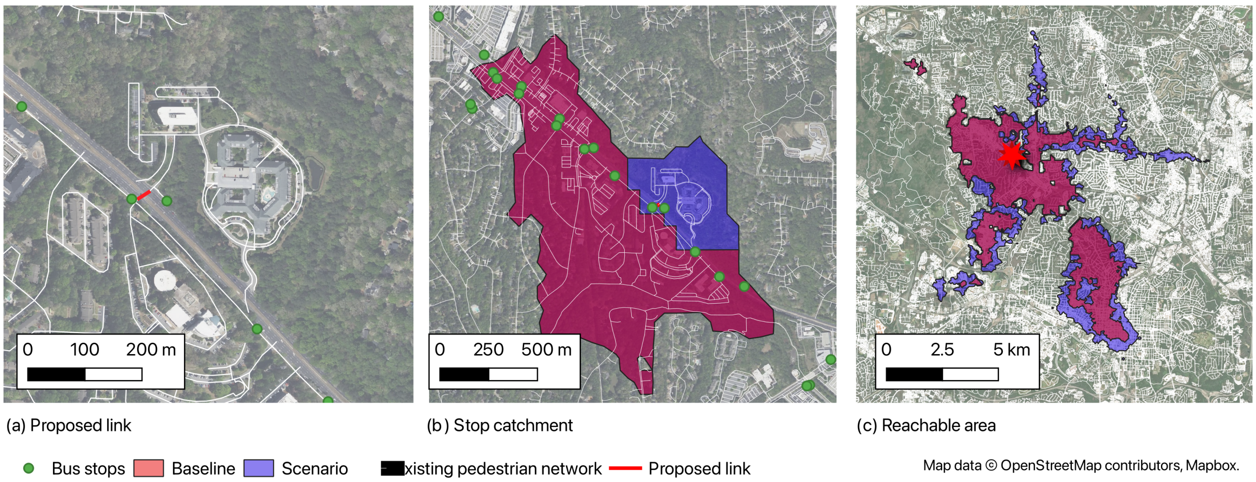

Figure 2: Link that scored highly in the first stage and enables access to a large number of additional destinations

Figure 2b shows the catchment area of the stop immediately across the street from the development, under existing conditions and with the proposed link added. Under existing conditions, the development is not in the catchment area of this stop, but adding a safe crossing would expand the catchment area to include the development.

Lastly, Figure 2c shows the reachable area via transit from the development under existing conditions and with the proposed crossing added to the network. There are no changes to the transit service in this scenario; the only change is the addition of the proposed crossing of Glenwood Avenue. The development the link connects to is highlighted with a red star. There is a significant expansion in reachable area to the north, east, and south thanks to the addition of the safe crossing. Previously it was not possible to board a southeast-bound bus from this development, as there was no way to cross the street; the reason areas to the south and east are not completely inaccessible in the baseline is because it would be possible to ride a bus the wrong way and then transfer to the opposite direction bus somewhere where it is safe to cross the street.

The link is shown in its local context in Figure 3. The location of the proposed link is highlighted in red, and the two bus stops as well as the entrance to the neighborhood are highlighted. This is clearly a challenging location to add a pedestrian crossing—Glenwood Avenue is essentially a freeway here, with a center median barrier, along with limited sight distances due to topography. There is a signalized intersection approximately 400 m north of this location, so perhaps extending sidewalk infrastructure to that location and providing a safe crossing there would be more realistic. This also demonstrates the challenges planners face in trying to reconnect and pedestrianize networks that were built for the exclusive use of private cars.

Figure 3: The location of the proposed link, looking northwest. © _jcaruso on Mapillary, CC BY-SA

Even without making any changes to the transit service, adding this one key link to the transit network has a large effect on transit accessibility for people living in this location. This clearly demonstrates that there are opportunities to improve transit access, and thus transit ridership, through physical infrastructure, and that this algorithm is a reasonable way to identify these possibilities.

Regional results

The previous section evaluated the results of adding a single link to the network. However, the Missing Links algorithm identifies thousands of potential locations to add links. Individually evaluating the contribution of each of these links to job access *via* transit is not computationally feasible. However, the Missing Links algorithm does score each of these links with their contribution to access *to* transit stops. I use this as a heuristic and select the top 100 links based on this metric. I then add the top 100 links to the OpenStreetMap network and evaluate changes in job access *via* transit.

It is computationally feasible to evaluate the effects of each of these 100 links on job accessibility via transit individually. However, for brevity, in this work I only compare existing conditions to a network where all of the top 100 links are added. I evaluate the accessibility impacts using a regional location-based access to jobs metric; for every Census block in the region, I evaluate the number of jobs accessible via transit within 60 minutes, for departures between 7 and 9 am, under existing conditions and with the links added.

These links are shown in red in Figure 4. While it appears there are less than 100 links, this is because even after deduplication links tend to cluster, something also observed by Bhagat-Conway et al. (*11*). The impact of the new links on job access via transit is shown in blue, with darker blue areas indicating a larger change in access. Increases in access range from just a few percent to significantly above 50% (for instance, in a situation where a location has no transit access at all before the link is added).

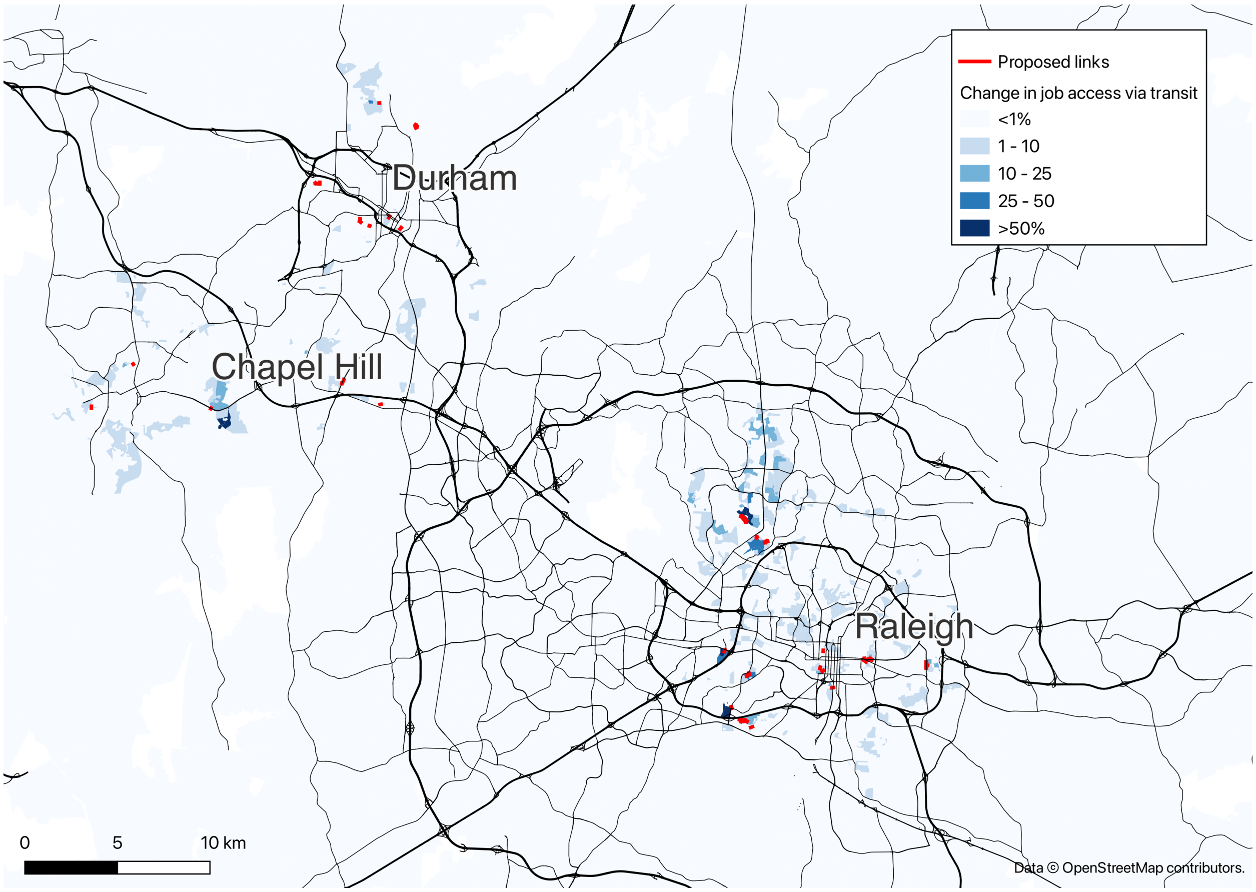

Figure 4: Change in jobs within 60 minutes via transit

As one might expect, the largest increases in accessibility are near the links. However, smaller changes are spread throughout the region. Residents who do not live near a link may still benefit from that link if it connects them to destinations, or even if it makes a transfer easier. This also demonstrates why it is not sufficient to only consider access *to* transit; closing missing links may make transit journeys easier even for people who do not live near those links. Combining the Missing Links algorithm with a jobs access *via* transit measure is critical to get the full impact of a link.

DISCUSSION

This article has demonstrated that there is potential to improve accessibility via transit through investments in the street network, and that the Missing Links algorithm provides a practical and computationally feasible way to identify candidate locations. It also demonstrates that it is important to evaluate effects on accessibility via transit, not just accessibility to transit, to understand the full impact of these investments.

Improving transit access via street network investments has several attractive fiscal properties. Unlike investing in more frequent or extensive service, street network improvements are a one-time investment. Once the network improvement is made, it will improve access indefinitely with minimal additional commitment of funds. Similarly, investing in the street network is capital rather than operating cost. Some grant funds are restricted to capital rather than operating investments. If extensive right of way acquisition or grading are not required, the cost to add connections to the network is also small compared to other transit capital investments. This is not to say that investments in service are unsound, but that it makes sense to also invest in the pedestrian network to ensure that service dollars provide the best accessibility they possibility can.

On the other hand, built environments change slowly, and once land is subdivided it can be very difficult to acquire right of way for new connections. It is much cheaper and easier to provide for these connections when street networks are first laid out. In fast-growing areas such as the Research Triangle, many new subdivision street networks are planned each year. It may be valuable to apply a tool such as this one to identify potential network changes to improve transit accessibility before plans are approved. Ideally, this exercise would be done using forecast rather than current transit service. Similarly, when planning a new bus or fixed-guideway transit route, applying this algorithm may help identify connections that could be added to the surrounding street network to maximize the impact of the new service.

FUTURE RESEARCH

While this research does evaluate job accessibility via transit, idenfication of top-scoring links is done using a much simpler access to transit stops metric. This simpler metric does not fully capture the experience of being a transit user. If those stops connect you to a bus line that doesn't run frequently or serve many destinations, they are not very useful. Similarly, gaining access to a bus stop on a line you cannot currently access is much more valuable than gaining access to a second or third bus stop on a line you can already access. A true access via transit measure would account for both of these things.

One of the key algorithmic innovations that makes Bhagat-Conway et al.'s (*11*) algorithm tractable is that it directly evaluates the difference in accessibility for each link, without needing to re-calculate accessibility estimates for the whole network. This requires that link costs not change over time–saving 5 minutes (or 500m walking) has the same effect on the trip regardless of where it was saved. This is not true for transit; a path that saved 5 minutes of walking time might save an hour overall if it let you get on an earlier bus.

An intermediate step is to measure access to routes rather than stops. Currently, the Missing Links algorithm works with undirected graphs, which is generally an appropriate way to model pedestrian networks as one-way pathways for pedestrians are rare. If the algorithm were extended to work on directed graphs, a zero-length edge could be added from every stop along a route to a node that represented the route itself, and those nodes could be considered destinations. Thus, paths that connect people to additional stops on the same routes would not score highly. This requires an algorithm implementation with a directed graph because the edges connecting to the route nodes must all go *to* the route node, rather than *from* it. Otherwise, the route nodes could be used as "shortcuts" to allow people to teleport to any location along the route from any stop (cf. *21*).

STATEMENT ON AI TOOLS

No AI/large language model tools were used in this research or the preparation of this manuscript.

REFERENCES

1. Miller, E. J. Accessibility: Measurement and Application in Transportation Planning. *Transport Reviews*, Vol. 38, No. 5, 2018, pp. 551–555. https://doi.org/10.1080/01441647.2018.1492778.
2. Committee of the Transport Access Manual. Transport Access Manual: A Guide for Measuring Connection Between People and Places. 2020.
3. Conway, M. W., A. Byrd, and M. van der Linden. Evidence-Based Transit and Land Use Sketch Planning Using Interactive Accessibility Methods on Combined Schedule and Headway-Based Networks. *Transportation Research Record*, Vol. 2653, 2017, pp. 45–53. https://doi.org/10.3141/2653-06.
4. Conway, M. W., A. Byrd, and M. van Eggermond. Accounting for Uncertainty and Variation in Accessibility Metrics for Public Transport Sketch Planning. *Journal of Transport and Land Use*, Vol. 11, No. 1, 2018, pp. 541–558. https://doi.org/10.5198/jtlu.2018.1074.
5. Conway, M. W., and A. F. Stewart. Getting Charlie Off the MTA: A Multiobjective Optimization Method to Account for Cost Constraints in Public Transit Accessibility Metrics. *International Journal of Geographical Information Science*, Vol. 33, No. 9, 2019, pp. 1759–1787. https://doi.org/10.1080/13658816.2019.1605075.
6. Santana Palacios, M. A. Unpacking Data Representation Issues in Distributional Accessibility Impact Assessments: Lessons from Bogotá's Urban Gondola. *Journal of Transport Geography*, Vol. 127, 2025, p. 104258. https://doi.org/10.1016/j.jtrangeo.2025.104258.
7. Pereira, R. H. M. Transport Legacy of Mega-Events and the Redistribution of Accessibility to Urban Destinations. *Cities*, Vol. 81, 2018, pp. 45–60. https://doi.org/10.1016/j.cities.2018.03.013.
8. Herszenhut, D., R. H. M. Pereira, L. da S. Portugal, and M. H. de S. Oliveira. The Impact of Transit Monetary Costs on Transport Inequality. *Journal of Transport Geography*, Vol. 99, 2022, p. 103309. https://doi.org/10.1016/j.jtrangeo.2022.103309.
9. Wu, H., P. Avner, G. Boisjoly, C. K. V. Braga, A. El-Geneidy, J. Huang, T. Kerzhner, B. Murphy, M. A. Niedzielski, R. H. M. Pereira, J. P. Pritchard, A. Stewart, J. Wang, and D. Levinson. Urban Access Across the Globe: An International Comparison of Different Transport Modes. *npj Urban Sustainability*, Vol. 1, No. 1, 2021, p. 16. https://doi.org/10.1038/s42949-021-00020-2.
10. Cignarella, J., H. M. Clark, R. Kraft, and D. Byrd. *An Onboard Survey of GoDurham Customers*. GoDurham, Durham, NC, 2023.
11. Bhagat-Conway, M. W., A. Compiano, and E. I. Ivie. So Close, yet so Far: A New Method for Identification of High-Impact Missing Links in Pedestrian Networks. *Computers, Environment and Urban Systems*, Vol. 120, 2025, p. 102290. https://doi.org/10.1016/j.compenvurbsys.2025.102290.
12. Walker, J. Basics: Walking Distance to Transit. http://www.humantransit.org/2011/04/basics-walking-distance-to-transit.html.
13. van Soest, D., M. R. Tight, and C. D. F. Rogers. Exploring the Distances People Walk to Access Public Transport. *Transport Reviews*, Vol. 40, No. 2, 2020, pp. 160–182. https://doi.org/10.1080/01441647.2019.1575491.

1 14. El-Geneidy, A., M. Grimsrud, R. Wasfi, P. Tétreault, and J. Surprenant-Legault. New Evidence on Walking Distances to Transit Stops: Identifying Redundancies and Gaps Using Variable Service Areas. *Transportation*, Vol. 41, No. 1, 2014, pp. 193–210. https://doi.org/10.1007/s11116-013-9508-z.

2 15. Macias, K. Alternative Methods for the Calculation of Pedestrian Catchment Areas for Public Transit. *Transportation Research Record*, Vol. 2540, No. 1, 2016, pp. 138–144. https://doi.org/10.3141/2540-15.

3 16. van Eggermond, M. A. B., and A. Erath. Pedestrian and Transit Accessibility on a Micro Level: Results and Challenges. *Journal of Transport and Land Use*, Vol. 9, No. 3, 2016, pp. 127–143.

4 17. Guerra, E., R. Cervero, and D. Tischler. Half-Mile Circle: Does It Best Represent Transit Station Catchments? *Transportation Research Record*, Vol. 2276, No. 1, 2012, pp. 101–109. https://doi.org/10.3141/2276-12.

5 18. Lahoorpoor, B., and D. M. Levinson. Catchment If You Can: The Effect of Station Entrance and Exit Locations on Accessibility. *Journal of Transport Geography*, Vol. 82, 2020, p. 102556. https://doi.org/10.1016/j.jtrangeo.2019.102556.

6 19. Omar, K. S., G. Moreira, D. Hodczak, M. Hosseini, and F. Miranda. Crowdsourcing and Sidewalk Data: A Preliminary Study on the Trustworthiness of OpenStreetMap Data in the US. http://arxiv.org/abs/2210.02350. Accessed Jan. 25, 2024.

7 20. Pereira, R. H. M., M. Saraiva, D. Herszenhut, C. K. V. Braga, and M. W. Conway. R5r: Rapid Realistic Routing on Multimodal Transport Networks with R5 in R. *Findings*, 2021, p. 10. https://doi.org/10.32866/001c.21262.

8 21. Humagain, P., and P. A. Singleton. Would You Rather Teleport or Spend Some Time Commuting? Investigating Individuals' Teleportation Preferences. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 74, 2020, pp. 458–470. https://doi.org/10.1016/j.trf.2020.09.010.