Supplementary Material

For manuscript titled:

A link criticality approach for pedestrian network design to promote walking

Process overview

The process of deriving the pednet score and identifying critical walkways is illustrated in Figure 1. We used publicly available geographic data to create the base pedestrian network ("pednet") comprising sidewalks and crosswalks, both existing and candidate/proposed. By comparing different pedestrian networks obtained by the different combinations of segment type and existence, we assessed the changes occurring in the accessibility of pedestrians to their destinations. The pedestrian movement was simulated using origin-destination sampling and all-or-nothing assignment. Pednet edge weights were computed based on their popularity derived from this route assignment. Pednet scores were then computed and used to identify the critical links.

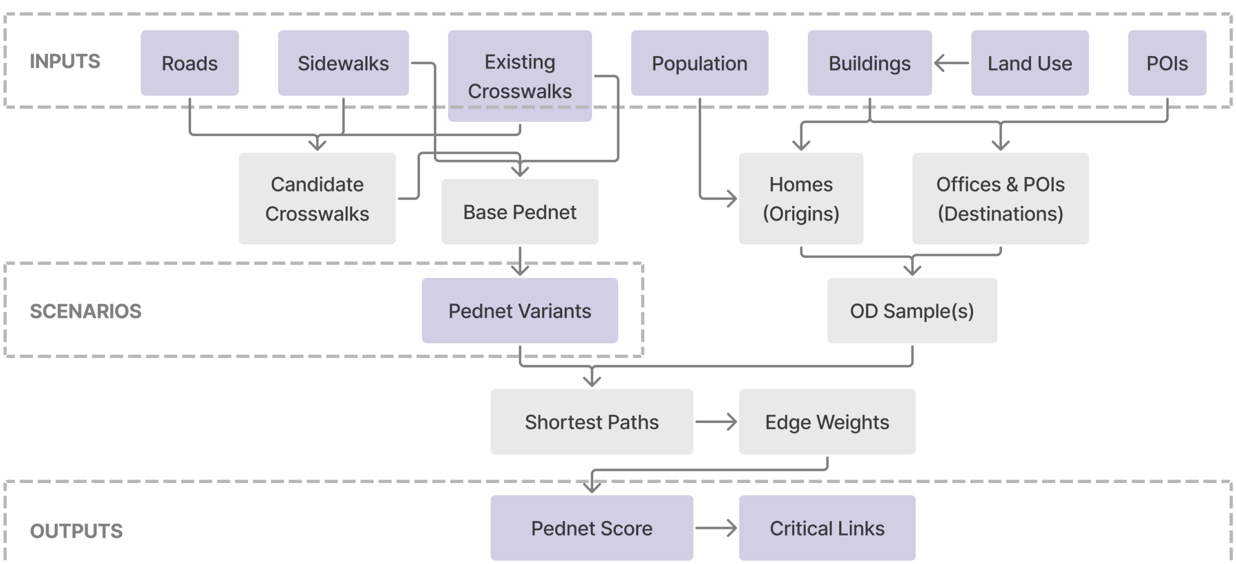


Figure 1: Overview of the pednet score computation process.

Pednet creation

Data acquisition

The geometric data layers of the cities, including city boundary, neighborhoods, road network, sidewalks, crosswalks, and land use, were obtained from the open data portals of the study cities and other sources, all of which are listed in Table 1.

Table 1: Sources of the geometry data layers for the study cities.

City [geocode]	Layer	Layer data URL
Austin [Austin, TX]	Boundary	Taken from city boundary with city label "Austin, TX" from the US TIGER shapefile of US metropolitan statistical area boundaries (https://www2.census.gov/geo/tiger/GENZ2018/shp/cb_2018_us_ua10_500k.zip)
	NPAs	https://data.austintexas.gov/dataset/boundaries-austin-neighborhood-planning-areas/nz5f-3t2e
	Roads	https://data.austintexas.gov/locations-and-maps/street-centerline/m5w3-uea6
	Sidewalks	https://data.austintexas.gov/locations-and-maps/sidewalks/pc5y-5bpw
	Crosswalks	https://data.austintexas.gov/transportation-and-mobility/transportation-markings_short_line/3p2i-pqdc
	Land use	https://data.austintexas.gov/locations-and-maps/land-use-inventory-detailed/fj9m-h5qy
Cambridge [Cambridge, MA]	Boundary	https://www.cambridgema.gov/GIS/gisdatadictionary/boundary/boundary_cityBoundary
	Roads	https://www.cambridgema.gov//GIS/gisdatadictionary/trans/trans_centerlines
	Sidewalks + crosswalks	https://www.cambridgema.gov/GIS/gisdatadictionary/trans/trans_sidewalkCenterlines

	Land use	https://www.cambridgema.gov/GIS/gisdatadictionary/cdd/cdd_landuse
Toronto [City of Toronto]	Boundary	https://open.toronto.ca/dataset/regional-municipal-boundary
	Roads	https://open.toronto.ca/dataset/toronto-centreline-tcl
	Sidewalks + crosswalks	https://open.toronto.ca/dataset/pedestrian-network
	Land use	https://borealisdata.ca/dataset.xhtml?persistentId=doi:10.5683/SP3/1VMJAG

The information of buildings was obtained from OpenStreetMap (OSM) using the Python library `OSMnx`. All geometries within the geocodes listed in Table 1 with the tag ‘`building`’ were extracted.

Similarly, the information of points of interest (POIs) was extracted from `OSMnx` by extracting all geometries with one or more of the following tags – `amenity`, `healthcare`, `leisure`, `office`, and `shop`. For the OSM POIs that were polygons instead of points, their centroids were considered.

All the GIS layers were cropped by the city boundaries given in Table 1.

Sidewalk and crosswalk labeling

The sidewalk and crosswalk information obtained from the open city data portals vary significantly in their details. Therefore, different rules were used to categorize existing and proposed walkway segments.

In Austin’s sidewalk layer, the attribute `pedestrian` was used to distinguish existing and proposed/candidate segments. The tags `EXISTING_SIDEWALK`, `DRIVEWAY`, and `SHARED_USE_PATH` were labeled ‘existing’, while `PLANNED_SIDEWALK`, `PLANNED_SHARED_STREET`, `POTENTIAL_SIDEWALK`, `ABSENT_SIDEWALK`, and `PROTECTED_STREET_PATH` were considered ‘proposed’ segments. For crosswalks, only the segments tagged `short_line = CROSSWALK` were considered. Moreover, these crosswalk segments were snapped to the nearest sidewalk segment endpoints to ensure the topological connectivity of Austin’s pednet.

For Cambridge, the pednet links tagged `TYPE = "CWALK-CL"` were considered crosswalks, while the other two types (`PFOOTPATH-CL` and `SIDEWALK`) were considered sidewalks.

In Toronto, the pednet links were already categorized with the attribute `crosswalk` (= 0 for sidewalks and 1 for crosswalks). For sidewalks, however, the existence labels were difficult to binarize. Existing segments were identified as those whose attribute `sdwlk_desc` showed the presence of a sidewalk on both sides of the road, on one side of the road in any of the four cardinal directions (with no or partial sidewalk on the other side), or having a laneway with sidewalk in any of the four directions. City walkways and recreational trails were also considered existing. Other labels, such as `no sidewalk on either side` and `not applicable` were considered proposed.

Candidate crosswalk generation

We created a set of candidate crosswalk segments at street intersections using the following procedure. First, each sidewalk segment was matched to its closest road segment. Then, a circular buffer of 20 m was drawn around each road intersection. This figure was based on the worst-case assumption of six 12-ft lane intersections with sidewalks and road clearance, and was also deemed suitable based on visual inspection. Then, all those combinations of sidewalk endpoint pairs within an intersection circle were identified which share the same road but not the same sidewalk and intersect the road centerline. Finally, the shortest candidate corresponding to each intersection leg was chosen. Existing crosswalk segments were removed from the resulting set. In this way, 35,804 candidate crosswalk segments were created in Austin, 733 in Cambridge, and 3,590 in Toronto. Though not intersection legs are covered by this technique (as illustrated in Fig. 1 of the manuscript), they are very comprehensive in increasing the network connectivity, as illustrated in the Results section of the manuscript.

Pednet score computation

The pednet edge weights were computed by the count of trip paths going through the edges. To this end, pedestrian trips were simulated based on land use and population distribution in each of the study cities. This involved the following data preparation and trip assignment procedure.

Building categorization

First, the location and attributes of buildings were prepared, primarily using the Overpass application programming interface (API) of OpenStreetMap (OSM) using the `OSMnx` Python library. Next, they were categorized as either trip generators and attractors. Since this study primarily concerns home-based walking trips, residential buildings were considered the demand generators while non-residential buildings were considered the attractors. To associate buildings with their residential status, building locations were overlaid with the city's land use layer (sources given in Table 1) and then labeled as generators or attractors. Inactive land use categories such as water, forests, and street right of way were not included in either category. The details of the labeling criteria of buildings based on land use category for each city is given in Table 2.

Table 2: Land use category labels used to categorize buildings as trip origins and destinations.

City	Land use category column	Trip generator categories	Trip attractor categories
Austin	`land_use`	Single Family; Mobile Homes; Duplex; Large-lot Single Family; Three/Fourplex; Apartment/Condo; Group Quarters; Retirement Housing; Semi-institutional Housing	Commercial; Mixed Use; Office; Manufacturing; Warehousing; Miscellaneous Industrial; Resource Extraction (Mining); Landfills; Hospital; Government Services; Educational; Meeting and Assembly; Cemetery; Cultural Services; Parks/Greenbelts; Golf Courses; Common Areas; Preserves; Railroad Facilities; Transportation Facilities; Aviation Facilities; Marinas
Cambridge	`CATEGORY`	Residential; Mixed Use Residential; Assisted Living/Boarding House; Education Residential	Commercial; Office; Transportation; Higher Education; Charitable/Religious; Office/R&D; Education; Industrial; Mixed Use Commercial; Government Operations; Utility; Health; Mixed Use Education
Toronto	`Class_name`	ApartmentNeighbourhoods; Neighbourhoods	Core Employment Areas; General Employment; Institutional; Mixed Use; Parks; Utility Corridors

POI data

Points of interest (POIs) are usually non-residential places which people travel to for different purposes such as goods stores, services, entertainment, education, etc. and thus form an important component of attractors (but not necessarily generators) of home-based trips. While OSM has a large collection of building footprints, the collection of POIs is usually small, which would limit the number of trip destinations. Therefore, we augmented the OSM POI dataset with additional POI data made available to us from SafeGraph Inc. as part of an academic data sharing contract. Duplicate POIs from SafeGraph and OSM were removed. Finally, all POIs were added to the set of candidate trip attractors.

OD weighting

Since buildings and POIs do not contribute equally to generating and attracting foot traffic, it is important to weigh them by their relative importance. To this end, each residential building was assigned an estimated population. This was based on the assumption of a uniform distribution of population in neighborhood. The population figures were obtained from the 5-year American Community Survey estimates of 2020 [1] for the cities of Austin and Cambridge at census block group (CBG) level, while they were derived from the 2021 crime rate estimates in Toronto by neighborhood [2].

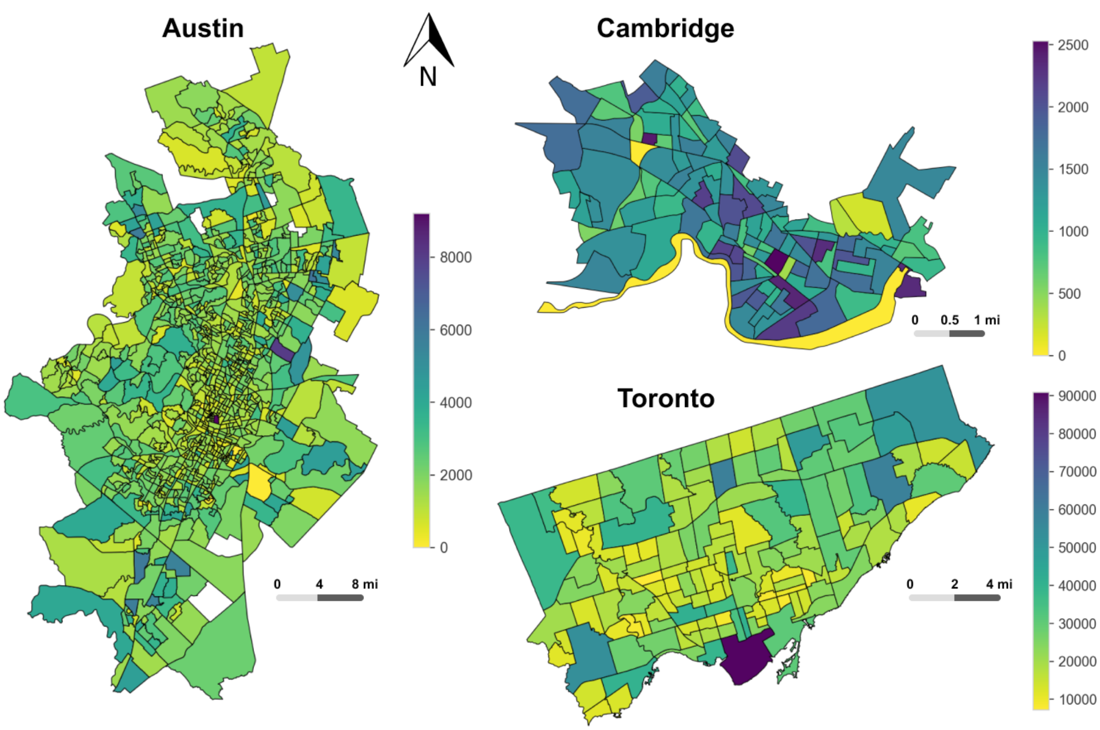


Figure 2: Population distribution by city neighborhoods (2020 estimates by census block groups for Austin and Cambridge and 2021 estimates by neighborhoods for Toronto).

Similarly, POIs were weighted by the mean foot visit count to them in September 2020 as provided by SafeGraph Inc. POIs from OSM and non-residential buildings do not have this information and were thus weighted uniformly, with the weight given by the median visit count to all SafeGraph POIs within the cities.

OD sampling

Once the candidate origin and destination points were prepared, origin-destination (OD) pairs were randomly sampled from them with repetition based on their weights. To ensure that walking trips are short, it was necessary to bound the distance between the origin and the destination. For regular walking trips, a distance of about 400 m (0.25 mi or about 5 minutes) to 1000 m (0.62 mi or 13 min) is typically considered acceptable [3], [4]. In this study, only OD pairs with a geographical (haversine) distance of less than 1000 m were allowed. With this constraint, a total of 1 million OD pairs were sampled with repetition for each city. Since

Shortest paths and edge weights

First, all origin and destination points were mapped to their closest node of a given pednet graph. Then, the shortest path between each OD pair was computed along the pednet using Dijkstra's algorithm. The total number of shortest paths incident on each edge yielded its pedestrian trip count that would then serve as the edge weight.

Pednet score distribution

Fig. 3 of the manuscript shows the spatial distribution of the edge weights and the pednet scores of the four main pednet variants of Austin. The pednet score distribution in the corresponding variants of Cambridge and Toronto are shown in Figure 3. It can be seen that there is not substantial difference in the overall distribution of the pednet scenarios where all candidate crosswalks are added to the base pednet in both Cambridge and Toronto. This is in line with the description in the manuscript as well as its figures 2 and 4.

An interesting observation in the case of Cambridge is the presence of concentric rings around two regions – the central area around Harvard University and the East Cambridge-Kendall Square area in the east, which are two high commercial and educational activity centers. North Cambridge and the region between Harvard University and Massachusetts Institute of Technology do not generate enough pedestrian demand according to the given population distribution and land use patterns.

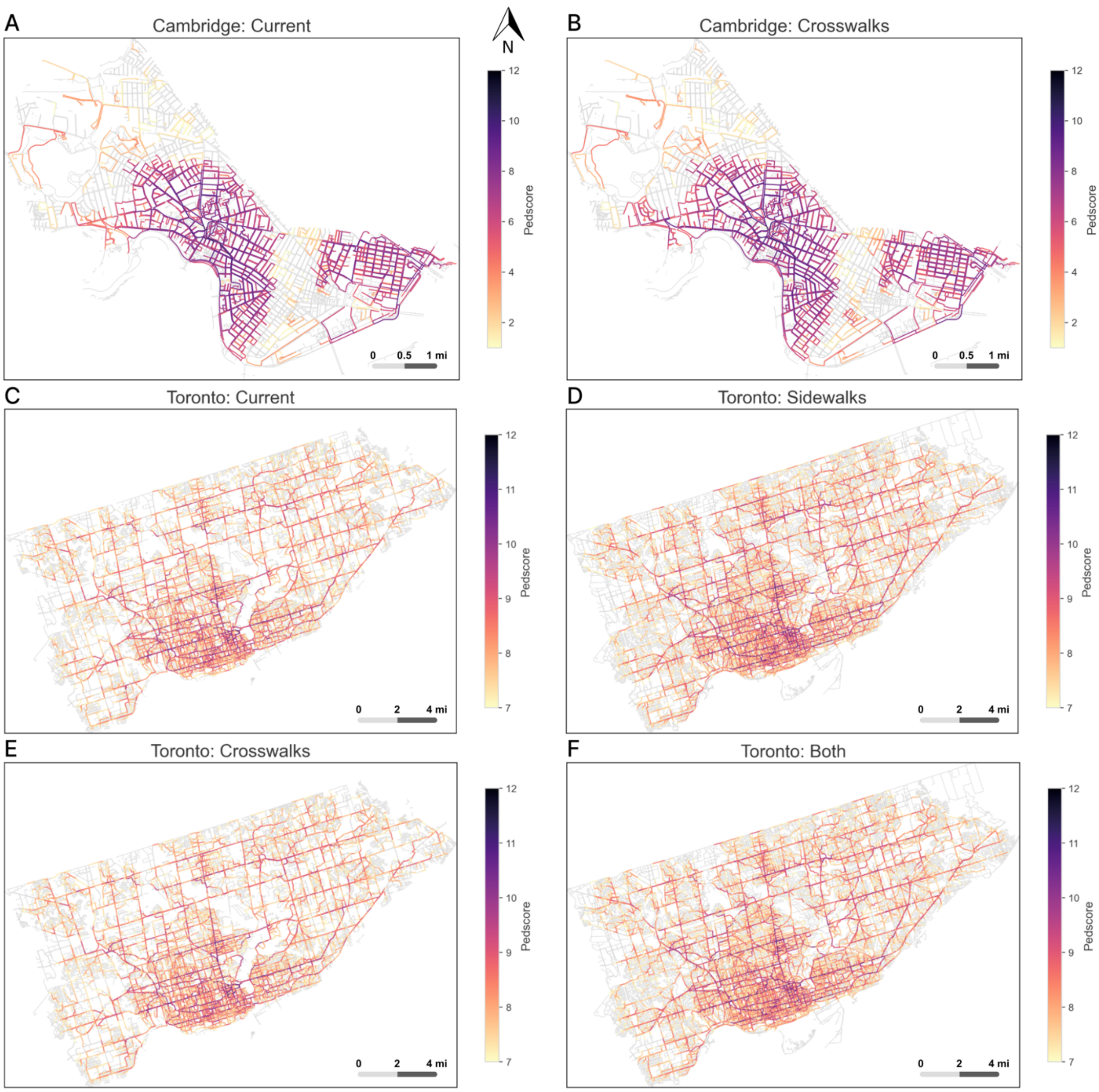
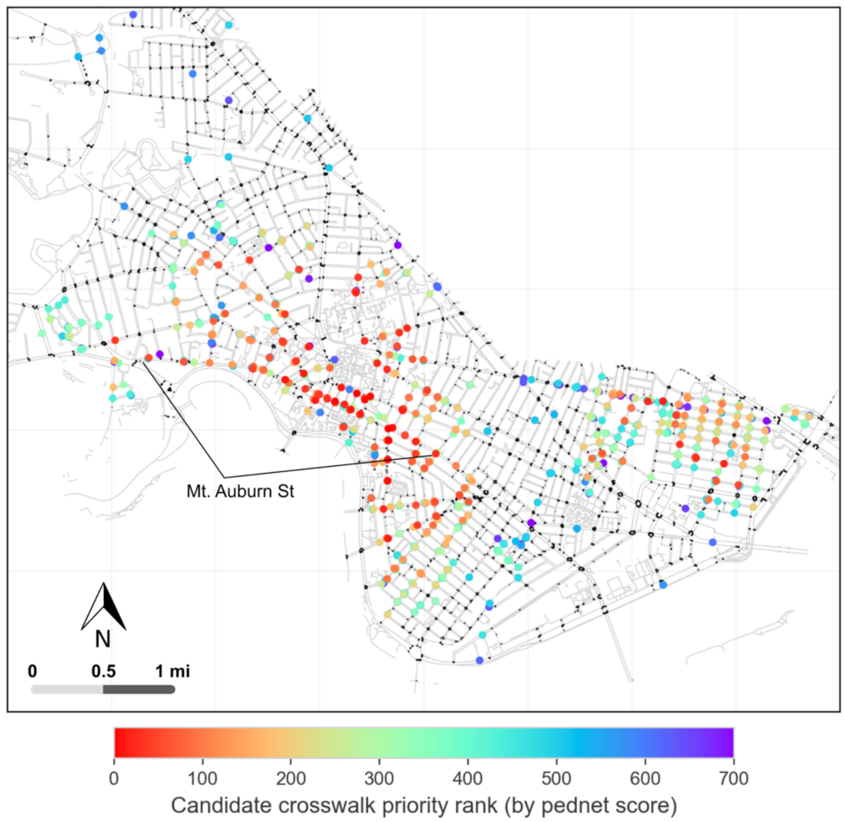


Figure 3: Pednet score distribution of the pednet variants of Cambridge (panels A and B) and Toronto (panels C through F). The pednet scores below the lower limit on the color-bars are excluded. The base pednet links are shown in light grey.

Critical links

Similar to the top 2000 most critical links of Austin shown in Fig. 5 of the manuscript, the locations of the most critical crosswalks of Cambridge (top 700) and Toronto (top 3500) are shown in Figure 4. These are ranked by their pednet scores. Similar to Austin, many candidate crosswalks near the downtown region are seen as more critical, though the downtown core itself is highly connected. In Cambridge, a continuous stretch of Mt. Auburn St passing through Harvard University shows a sequence of highly critical crosswalks whose implementation could benefit the nearby pedestrians.


Mt. Auburn St
N
0
0.5
1 mi
0
100
200
300
400
500
600
700
Candidate crosswalk priority rank (by pednet score)

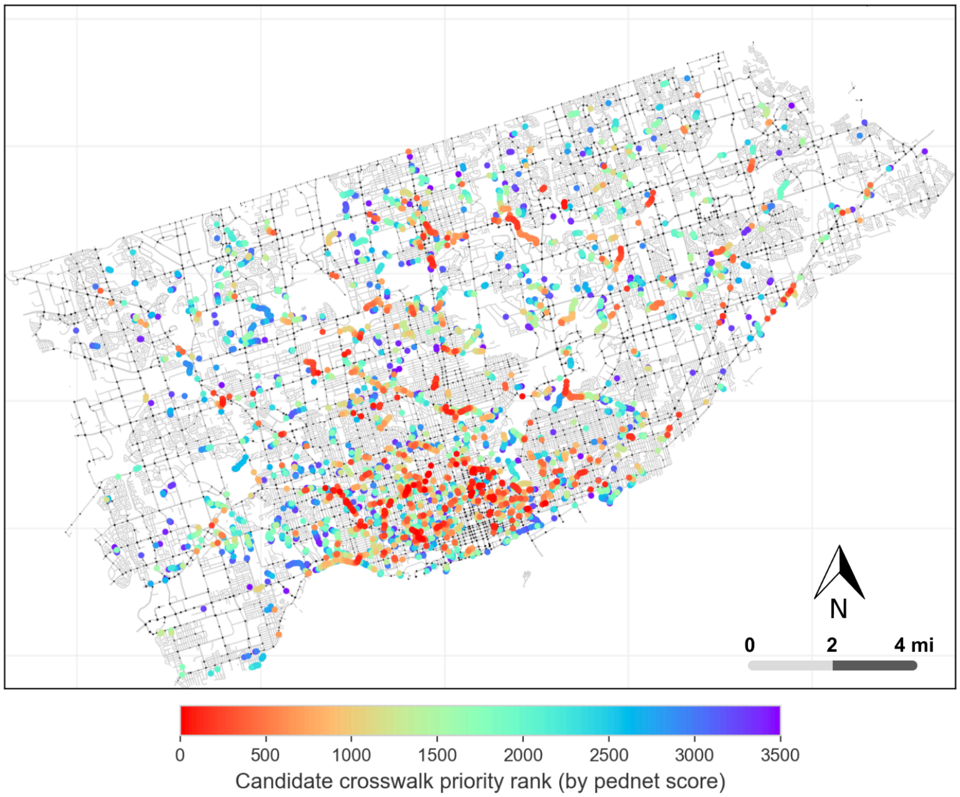


Figure 4: Criticality ranks of candidate crosswalks in Cambridge (top) and Toronto (bottom) shown as colored points. The existing sidewalks are shown as grey lines and the existing crosswalks shown as black lines.

Additional references

[1] "American Community Survey Data via API." https://www.census.gov/programs-surveys/acs/data/data-via-api.html (accessed Oct. 23, 2022).

[2] "Neighbourhood Crime Rates - City of Toronto Open Data Portal." https://open.toronto.ca/dataset/neighbourhood-crime-rates/ (accessed Oct. 23, 2022).

[3] C. M. Hoehner, L. K. Brennan Ramirez, M. B. Elliott, S. L. Handy, and R. C. Brownson, "Perceived and objective environmental measures and physical activity among urban adults," *Am J Prev Med*, vol. 28, no. 2, pp. 105–116, Feb. 2005, doi: 10.1016/J.AMEPRE.2004.10.023.

[4] A. W. Agrawal and P. Schimek, “Extent and correlates of walking in the USA,” *Transportation Research Part D: Transport and Environment*, Dec. 01, 2007.