A High Performing and Scalable Model for   
Computing and Visualizing Urban Transit Accessibility

**Yuxuan Cui**  
Computer Science and Statistics  
Faculty of Science  
University of British Columbia

[cuiyuuxuan@gmail.com](mailto:cuiyuuxuan@gmail.com)

**Luka Vukovic**  
Computer Science and Statistics  
Faculty of Science  
University of British Columbia

[luka.vuko@outlook.com](mailto:luka.vuko@outlook.com)

**Rain Shen**  
Computer Science and Statistics  
Faculty of Science  
University of British Columbia

[rain.ya1213@gmail.com](mailto:rain.ya1213@gmail.com)

**Paul**  
Computer Science and Statistics  
Faculty of Science  
University of British Columbia

[graham.kerford@gmail.com](mailto:graham.kerford@gmail.com)

June 25th, 2021

**3 METHODOLOGY**

In this section, we detail how the research problem was approached computationally. We explain what data a scalable model would require, and how transit accessibility scores are computed and visualized from the data in an efficient manner.

**3.1 Data**

The data.

**3.2 Computing Travel Time Matrices**

To evaluate transit accessibility across an urban landscape, many-to-many point travel times need to be computed into a travel time matrix. This requires a street network (Open Street Maps), a transit network (GTFS), origin-destination coordinates, and a routing engine.

Popular open-source routing engines include Open Trip Planner 1 (OTP1), Open Trip Planner 2 (OTP2), Conveyals R5 (Rapid Realistic Routing on Real-world and Reimagined networks), and GraphHopper. OTP1 and OTP2 are generally focused on passenger facing journey planning. OTP1 has analysis functionality but performs slow relative to other engines, using a generalized cost A\* algorithm. OTP2 is better optimized with a Multi-criteria range-RAPTOR algorithm but does not support one-to-many point routing analysis. Most importantly however, OTP optimizes routing on generalized cost instead of on minimizing travel time. For example, OTP may opt for a single long bus ride over one with a few transfers that yields a shorter travel time. In reality, transit users typically aim to optimize time.

R5 supports and is optimized for time-window trip planning which better reflects how people use the transportation system. R5, being implemented in Java, was also intended for analysis applications being magnitudes faster and less memory intensive than engines that are similar to OTP1, particularly for one-to-many point routing and travel time matrix generation. (OTP website) As such, travel times were computed using the open source r5r library, an R implementation of Conveyal’s R5 routing engine, with multimodal transit networks built from the GTFS and OpenStreetMap data detailed in 3.1. (Pereira, 2021) The final travel time matrix was an aggregation of 36 trips, departing every hour from 7:00am to 7:00pm with a 30-minute departure window on a weekday, a Saturday, and a Sunday (12 x 3). This allowed us to average travel times across changing bus schedules throughout the week.

**3.3 Measuring Transit Accessibility**

Transit accessibility to cultural amenities was measured in three fashions. The first was via isochrones, which indicates the shortest time from a given point to the nearest amenity. This was the most interpretable of all three measures.

The second measure involved scoring each city block using the mean transit time to the nearest amenity and the standard deviation of that mean time. The score was computed by taking the inverse of mean transit time plus two standard deviations of that time. This is essentially the worst-case scenario transit time scaled from 0 to 1, where the shortest transit times yield scores closer to 1, while the longest transit times yield scores closer to zero.

Since scores were not very interpretable, particularly because the scores were positively skewed, the percentile of the scores were computed as a third measure. This uniformly distributes the scores allowing for more interpretability than raw scores alone.

***This is more discussion in nature. For methods simply state what was done and nothing else.***

- there are multiple ways to measure accessibility:

- 0-1 scoring (less interpretable than quantiles)

- quantiles (less interpretable than isochrones)

- isochrones (most interpretable)

1) only raw scoring requires some basic mathematical intuition. Since we want a final accessibility score to be between 0 and 1, with most accessible areas having scores close to 1, we can take the inverse of the specific travel time from the travel time matrix.

- Smallest travel times will have the greatest values. Normalization was then done to fit all scores to the 0-1 range.

- since we also want to account for the average uncertainty in transit time (standard deviation of the average trip time), we chose to add 2 standard deviations to the average trip. This effectively generates a best trip worst case scenario similar to real world situations where individuals may opt for the best possible trip, but often the actual trip length will be greater than what is expected.

- (insert equations and notation for this calculation)

2) quantiles prove advantageous if looking to uniformly spread accessibility scores. It’s more interpretable in a sense that we can directly see the accessibility percentile of a particular region, however many areas with similar scores will become spread across larger percentile ranges. (show example in discussion / map of raw scores vs map of percentiles)

3) isochrones involved simply grouping the nearest destination travel times for each blocks into a time group ranging from 15 minutes and under, to 90 minutes and over.

* Nearest 1, 2, 3, accessible destinations were considered in the scoring. All accessible destinations were also visualized but generally yields the best average accessibility across a city however it also provides less information about transit network isolated communities.

**References**

Pereira, Rafael H. M., Marcus Saraiva, Daniel Herszenhut, Carlos Kaue Vieira Braga, and Matthew Wigginton Conway. 2021. “r5r: Rapid Realistic Routing on Multimodal Transport Networks with R5 in R.” Findings, March. https://doi.org/10.32866/001c.21262.