REGIONAL TRANSIT SCREENING PLATFORM

dvrpc

SURFACE TRANSIT RELIABILITY

Where should efforts to improve transit reliability be focused? A variety of factors, such as infrastructure, congestion, and enforcement, impact surface transit reliability. Find out where reliability issues are likely to impact the most passengers.

Goal

The goal of the surface transit reliability analysis was to identify corridors where surface transit service is particularly slow or delayed as places where road or transit improvements could increase reliability.

Data Sources and Preparation

The surface transit reliability analysis relied on data from 2017 INRIX XD, SEPTA, NJ Transit, PATCO, and TIM 2.3. Significant time and effort were spent on creating input datasets for the reliability analysis, most of which are useful on their own.

Input: Travel Time Index

TTI is the ratio of peak hour travel time to free flow travel time. The TTI is used in this analysis accounts for all vehicle types, not just transit vehicles. High TTI indicates congestion which negatively impacts surface transit reliability.

2017 INRIX XD data was used for the Travel Time Index (TTI) on available road segments. A model estimate of TTI was calculated using TIM 2.3 for road segments used by surface transit where INRIX data was not available.

Input: On Time Performance (OTP)

Transit agencies provided line level OTP percentages. OTP represents the percent of time the bus route is considered on time. A trip is considered on time when it arrives between 0 and 6 minutes after the scheduled time. SEPTA OTP was compiled from the 2017 Route Statistics report. NJ Transit OTP was gathered from time schedule adherence reports from trips made between September 2017 and January 2018.

Input: Average Scheduled Speed

TIM 2.3 was used to calculate the average speed of transit on road segments based on the distance between stops and the scheduled time for each stop point. If multiple bus routes use a road segment, a weighted average was calculated to ensure that the buses using the segment most throughout the day were considered more heavily. Average scheduled speed was used in the overlay analysis at the road segment level.

Input: Ridership

Transit agencies provided ridership data for each transit route. SEPTA also provided stop level ridership data. SEPTA's GTFS was used to convert stop level data to segment level passenger loads for incorporation into the analysis and for visualization. This conversion relied on the following assumptions:

- GTFS includes all patterns of surface transit trips. The longest pattern was used to visualize load data.
- Stop level load data was provided by direction (northbound, southbound, eastbound, westbound). GTFS trip directions are identified using a binary field. The most common combinations of text and binary directions were used to join ridership and GTFS data by direction.

The process for preparing all datasets for overlay analysis included cleaning the data, evaluating the distribution of its values, and creating a continuous range of scores on the scale of 0 to 100, where 100 indicates areas of highest priority. The scores are derived from each value's cumulative probability under the normal curve.

The shapefiles used to create link- and line-based scores presented spatial problems both within and between files. Within files, several links or lines can run on top of one another. Consider Market Street, which has trolleys, multiple buses, regional rail, and the subway line sharing approximately the same latitudes and longitudes, but at different elevations. This is part of the logic behind removing lines and links with exclusive rights-of-way, as including lines such as regional rail may artificially inflate the performance and ridership of lines that share space with traffic on the surface. Between files, the data came from different sources. This means that the exact same line appeared in slightly different places in every shapefile.

Rasterizing the data alleviated some of these concerns. Within datasets, it ensured that reliability metrics for links or lines that run on top of another were not accidentally eliminated—for example, by only considering the topmost layer. Instead, the key values for these overlapping links were averaged or added where appropriate (passenger loads). Between datasets, it acted as a spatial buffer, ensuring that reliability metrics pertaining to the exact same line stack on top of one another, even though each metric came from a different source and might have slightly different latitudes and longitudes. Each dataset was rasterized at a 50m cell resolution.

Methodology

TTI, OTP, and scheduled speed were used to answer questions about the road segment conditions.

1. Where is there congestion? High TTI indicates congestion which negatively impacts surface transit reliability.
2. Where is transit slow or unreliable? Low OTP and low average scheduled speed are suggestive of reliability issues.

The rasters representing the three datasets were overlaid to calculate the overall measure of reliability. The variables were weighted as shown in Table 1, where traffic data accounted for half the score and transit data accounted for the other half. In the resulting raster, cells over road segments with the highest scores indicate those with the greatest potential for reliability issues.

Table 1

Variable	Weight
TTI	50
OTP	25
Average Scheduled Speed	25

To prioritize the highest scoring road segments, the results were then overlaid with transit ridership which highlighted segments that impact high ridership surface transit service. Line level ridership was used for NJ Transit routes, while passenger load estimates were used for SEPTA routes. Rasterizing load estimates sometimes resulted in counting the same loads several times. To avoid creating false outliers, the 95th percentile of estimated loads was computed, and cells with loads exceeding this value were replaced with the 95th percentile value.

For webmap visualization purposes, the reliability raster and the ridership weighted reliability raster results were reclassified using a 20-class Jenks (natural breaks) classification scheme and converted to polyline using a series of buffers and joins.

Results

The reliability score layer shows the results of combining TTI, OTP, and scheduled speed. A high reliability score is indicative of segments that may benefit from targeted improvements to improve transit operations.

The results were then weighted by ridership (Ridership Weighted Reliability Score) to highlight segments that impact high ridership surface transit service and allow for further prioritization of improvements.