Experimental Evaluation of Network Properties in Public Transportation Systems

Alan Joyce
Stanford University
CS 224W: Social and Informational Network Analysis

ABSTRACT

Public transportation systems across the United States serve millions of riders every day, forming crucial links between people and their daily destinations. In this work, we develop a method for modeling public transportation systems as network graphs. We use the resultant graphs to study the properties of 18 public transportation systems from cities and counties in the United States and Canada. Furthermore, we develop a simple model for passenger behavior and use it to test the relative robustness of these networks.

1. INTRODUCTION

Recently, public transportation agencies across the United States, Canada, and a handful of other countries have been making information about their systems publicly available online in the General Transit Feed Specification (GTFS) format. This allows applications such as Google Transit to be built, which use this data to help passengers plan their trips. At the same time, this data provides a wealth of raw material for an analysis of the properties and behavior of public transportation systems.

In basing this study on GTFS data, we have developed a set of routines that interpret an arbitrary GTFS feed as a transportation network graph. The routines are then able to measure static properties of the network, including degree distribution, average path length, clustering coefficient, degree-degree correlation, and betweenness. Next, the routines can run passenger simulations over the network and gather data on how riders are affected when the network is modified. These modifications include removing stops randomly, removing stops based on node degree, and simulating vehicle lateness across the network.

The advantage of this general set of routines is that it can be applied to any transportation network specified in GTFS format. This has allowed us to collect data on 18 networks (including bus, rail, and ferry lines), but also holds the promise of being extendable to other networks across the world as more transportation agencies release GTFS feeds for their systems.

2. RELATED WORK

Previous work in public transportation network research has focused heavily on analyzing the network properties of specific transportation systems and on building models that facilitate comparisons between systems in different cities.

The study most similar to ours, Sienkiewicz and Hołyst [8], examines the properties of public transportation networks in 22 Polish cities. The authors measure these networks with universal tools of complex network analysis and find that networks of varying size exhibit common features, such as degree distributions and a power-law decay of clustering coefficients for large node degrees.

In Seaton and Hackett [6], the Boston subway system is compared with one in Vienna and it is found that both networks satisfy small-world criteria. The authors are able to use bipartite graph theory to predict the value of average degree, but are unable to make reliable predictions with regard to other network properties.

Latora and Marchiori [4] examine the properties of the Boston subway system and propose a measurement of local and global network efficiency. The authors find that while the system is efficient on a global scale, the local efficiency is very low, indicating that damage to (or removal of) a single station will be highly detrimental to the connections between its adjacent stations.

Sen et al. [7] study the Indian railway network, defining nodes as corresponding to individual stations and links as pairs of stations connected by a single train. The authors find that mean distance is a suitable measure for the network's connectivity and use it to highlight the Indian railway network's small-world properties.

While several studies have analyzed the network properties of public transportation systems, we believe that no study has used precise stop time data (as is available in the GTFS format) for analysis and simulation. Furthermore, we believe that the incorporation of a passenger behavior model with simulations of network effectiveness is a novel approach in this area of inquiry.

3. NETWORK REPRESENTATIONS

The initial problem posed by transportation network analysis is one of representation. How should we interpret the real-world transportation system as a network graph in order to facilitate analysis and simulation? To meet the needs of this study, we present two representations: one which will allow us to measure general network properties (GN), and one which will allow us to simulate the effect of network modifications on passengers (G).

GN: A general representation

The purpose of GN is to allow us to measure degree distribution, clustering coefficient, degree-degree correlation, and betweenness in a network. In GN, each stop or station in the transportation network is represented as a single node. A weighted, directional edge is created between two nodes if there is a route that connects the two stops directly without any intermediary stops. This edge is weighted by the average amount of time in seconds it takes to travel between the two stops across all routes that connect them in the relevant direction.

G: A detailed representation

While GN serves our purposes well for the analysis of general network properties, it does not provide the level of granularity necessary to accurately view the network as a passenger would while trying to navigate it. In order to achieve this, we propose the specific representation G, which fully models the existence of individual trips along each route throughout a day.

In G, each node represents a single stop at a specific point in time. Each node is identified by the stop (*s*) along with a count of seconds since midnight (*t*). Each edge represents a single trip along a single route that connects stops s_0 and s, occupying the time between t_0 and t. Each edge is directional (forward in time) and is weighted by the value $t - t_0$.

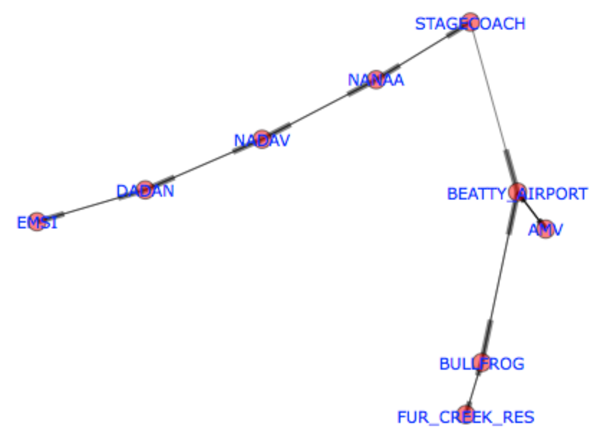


Figure 1: GN for the arbitrary Demo Transit Agency

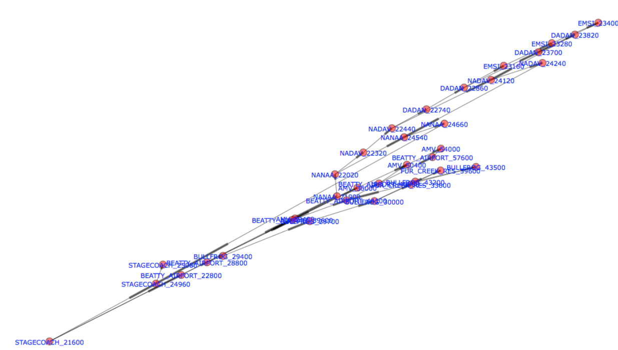


Figure 2: G for the arbitrary Demo Transit Agency

Modeling Passengers

If the purpose of G is to provide an accurate representation of the network as a passenger would see it, our next step is to develop a model for passenger behavior across G. Existing models for passenger actions can be seen in [3] and [5], but we propose a simplified model for the purposes of this study:

1. For each passenger p_i among j total passengers
   a. Choose a random source node u and a random destination node v, where $u \in$ G and $v \in$ GN.
   b. Find the set of nodes V in G such that for all $n \in$ V, $s(n) = s(v)$ and $t(n) >= t(u)$, where $s(n)$ is the stop ID and $t(n)$ is the number of seconds past midnight for node n.
   c. Calculate the weight, $w(u,n)$ of the shortest path between u and n in G for all $n \in$ V. Assign p_i[weight] = min($w(u,n)$, $n \in$ V). If there is no path between u and any $n \in$ V, assign p_i[weight] = 0.
2. Obtain a measurement for the average trip time: $w_{avg} = avg(\, p_i[\text{weight}], \text{ where } p_i[\text{weight}] \neq 0\,)$
3. Obtain a measurement for the fraction of stranded passengers: $s = count(\, p_i, \text{ where } p_i[\text{weight}] = 0\,) / j$

The above model represents each passenger as being currently located at a node in G and planning to travel to a node in GN. This means that there is a time variable associated with each passenger's source position, but not with their destination. The model proceeds by finding the shortest path in G from the source position to any node later in time with the same stop ID as the destination.

The result is that each passenger will either complete their journey in some number of seconds (their trip time), or will become stranded because it is impossible to reach their chosen destination before the end of the day. We draw from this the general network properties for average trip time, w_{avg}, and fraction of passengers stranded, s.

4. MEASUREMENTS

We define measurements for network properties, including average path length, clustering coefficient, degree-degree correlation, and betweenness.

Average path length

Choose node u with the maximum degree of all nodes in G. Define $short(v)$ to be the length of the shortest path between u and node v in G. Compute the average path length,

$$L = avg(\, short(v), \text{ for all } v \in \text{G where } v \neq u\,)$$

Clustering coefficient

For each node v in GN, compute the clustering coefficient,

$$c_v = [2\, T(v)]\, [deg(v)]^{-1}\, [deg(v) - 1]^{-1}$$

where $T(v)$ is the number of triangles through node v. We can then compute the average clustering coefficient across the network,

$$C = avg(\, c_v, \text{ for all } v \in \text{GN}\,)$$

Degree-degree correlation

We use the method specified in [8] to calculate the assortativity coefficient,

$$r = \frac{\sum_i j_i k_i - \frac{1}{M}\sum_i j_i \sum_i k_i}{\sqrt{\sum_i j_i^2 - \frac{1}{M}(\sum_i j_i)^2}\sqrt{\sum_i k_i^2 - \frac{1}{M}(\sum_i k_i)^2}}$$

where i iterates over all pairs of nodes in GN. j_i and k_i are the degrees of the nodes.

Betweenness

We use the algorithm from [1] for each node v in GN to calculate betweenness centrality,

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where V = GN, σ_{st} is the number of shortest paths from $s \in$ V to $t \in$ V, and $\sigma_{st}(v)$ is the number of shortest paths from s to t that v lies on. We can then compute the average betweenness,

$$B = avg(\, C_B(v), \text{ for all } v \in \text{GN}\,)$$

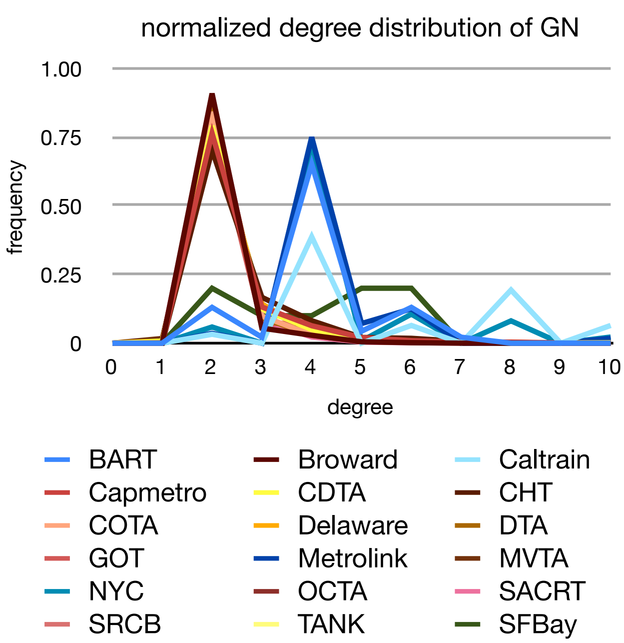
Table 1: Information and network properties for the 18 public transportation networks considered in this study. n is the number of stops (nodes in GN), n' is the number of stop times (nodes in G), L is the average path length, C is the clustering coefficient, D is the degree-degree correlation, and B is the betweenness.

network	location	type	*n*	*n'*	L	C	D	B
BART	San Francisco, CA	rail	46	17047	14301.9	0.0688	0.0897	0.1936
Broward	Broward County, FL	bus	4847	338818	10884.5	0.0064	0.2646	0.0211
Caltrain	San Francisco, CA	rail	31	2232	25036.1	0.4869	0.1022	0.1045
Capmetro	Austin, TX	bus	2870	349597	27940.0	0.0332	0.2663	0.0156
CDTA	Albany, NY	bus	2997	160203	13512.7	0.0390	0.3899	0.0214
CHT	Chapel Hill, NC	bus	623	33309	8510.7	0.0367	0.4430	0.0455
COTA	Columbus, OH	bus	4236	298093	6947.5	0.0089	0.4519	0.0137
Delaware	Wilmington, DE	bus	2816	123948	24982.0	0.0441	0.3311	0.0183
DTA	Duluth, MN	bus	1681	102668	8653.3	0.0236	0.4112	0.0530
GOT	Toronto, Ontario, Canada	bus	1526	61021	17805.1	0.0175	0.1606	0.0287
Metrolink	Los Angeles, CA	rail	56	1926	13969.8	0.1149	0.0006	0.1520
MVTA	Burnsville, MN	bus	941	24443	28434.5	0.0335	0.2826	0.0421
NYC	New York, NY	rail	492	395033	27604.8	0.0606	0.3542	0.0118
OCTA	Orange County, CA	bus	6233	399261	17019.9	0.0120	0.2152	0.0127
SACRT	Sacramento, CA	bus	3065	136496	6765.6	0.0054	0.1807	0.0232
SFBay	San Francisco, CA	ferry	10	484	15937.1	0.6603	-0.7273	0.0944
SRCB	Santa Rosa, CA	bus	459	18782	8453.0	0.0059	0.0880	0.0381
TANK	Ft. Wright, KY	bus	1517	61249	11971.7	0.0439	0.3220	0.0350

5. ANALYSIS

Data gathered in this study includes degree distributions, the network properties detailed in Table 1, and the results of three network modification simulations.

Degree distribution



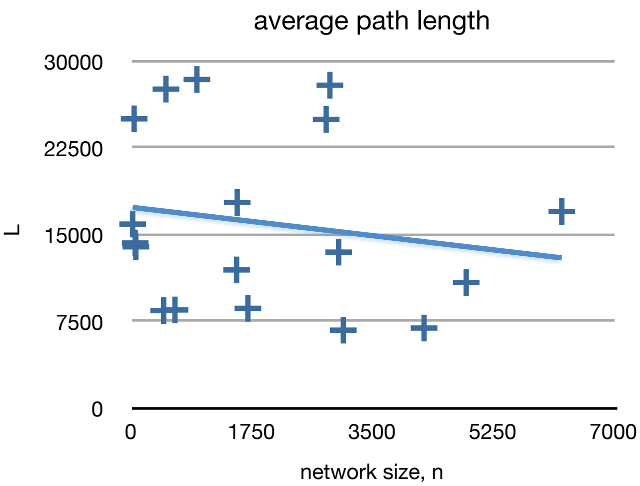
The above graph shows degree distributions across GN for each of the 18 public transportation networks. Lines are colored based on network type, with red representing bus lines, blue representing rail systems, and green representing the single ferry network.

We observe that degree distributions assume distinct and predictable forms that depend directly upon a network's type. Bus networks contain a majority of nodes with degree 2, while rail networks contain a majority of nodes with degree 4. It is difficult to determine a similar pattern for ferry networks based on the single data point we have, but it is reasonable to imagine that they would possess a characteristic degree distribution as well. These results suggest that it would be feasible to build a classifier to determine network type based solely on degree distribution.

Average path length

According to [7], mean distance or average path length should be a suitable measure of network connectivity, and thus also network effectiveness. Based on this measure alone, we would have SACRT and COTA as top performers and NYC, Capmetro, and MVTA at the bottom.

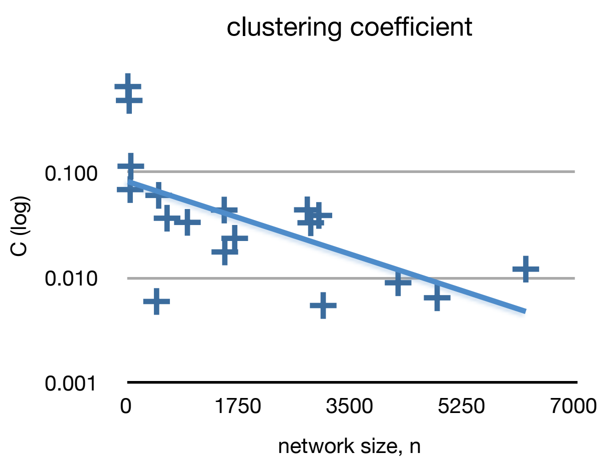
Bus networks have an average path length of 14760, while rail networks have an average of 20228. At the same time, bus networks have an average of 2601 stops, while rail networks average 156 stops.



We observe a slight downward trend in average path length with increased n, but this alone does not appear to fully represent the difference between the values for bus and rail networks.

Clustering coefficient

The clustering coefficient, C, represents the probability that two arbitrary neighbors of a node share a common link. In [8], a small decrease in C was observed with increasing network size.



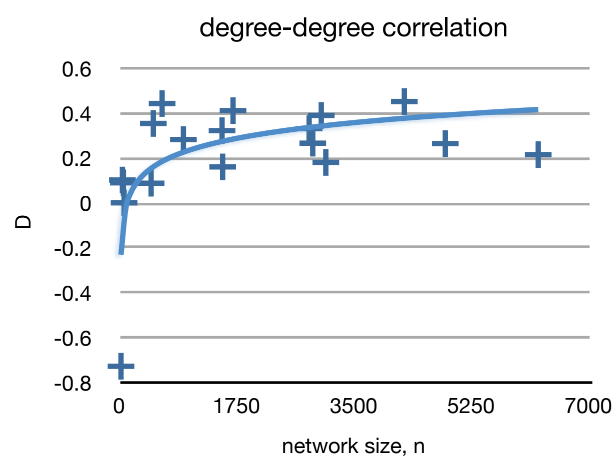
We observe a dramatic decrease in C between very small networks ($n < 100$) and slightly larger ones, as well as a small decrease in C among larger networks with $n > 100$.

Average C is 0.0239 for bus networks and 0.1828 for rail networks, representing a portion of the gap between small networks and large ones.

Degree-degree correlation

The degree-degree correlation, D, is shown to be positive in all cases except for the SFBay ferry network (see Table 1). Transportation networks have few nodes with high degree and these nodes are usually linked with each other, resulting in a positive D value. The SFBay network defies this trend because it is actually a collective representation of several individual ferry agencies in the San Francisco Bay (Blue & Gold Fleet, Harbor Bay Ferry, Baylink, and Golden Gate Ferry). This structure results in weakly linked network subsets that reduce D.

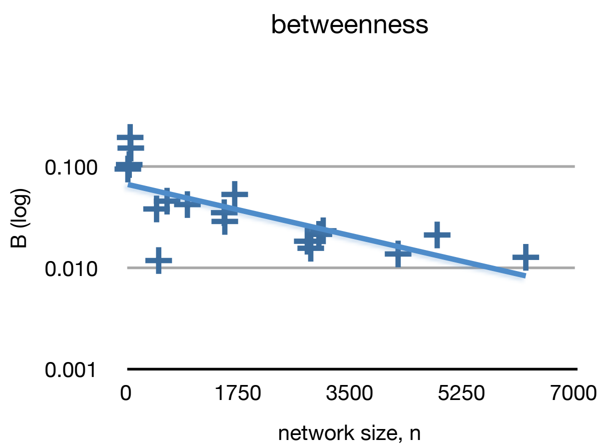
Average D is 0.2929 for bus networks and 0.1367 for rail networks, with SRCB having an unusually low D for a bus system and NYC having a high D for a rail system. Bus networks tend to have fewer high-degree nodes than rail networks, which makes it easier for these nodes to be linked and increases D.



Betweenness

Betweenness, B, allows us to measure the average importance of a node. In general, rail networks have high B (> 0.1), although NYC has the lowest B of any network studied. SFBay has a high B as well at 0.0944, significantly higher than all bus networks, which have B < 0.06. Averages are 0.0283 for bus networks and 0.1155 for rail networks.

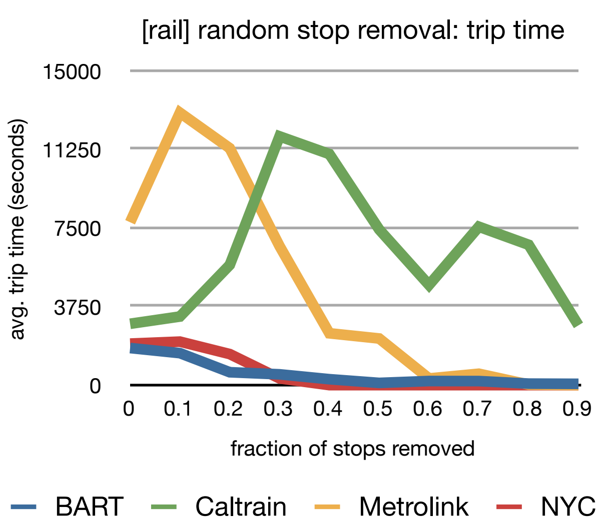
NYC continues to be an outlier for rail networks, likely due to its size and its segmented structure across a large area.

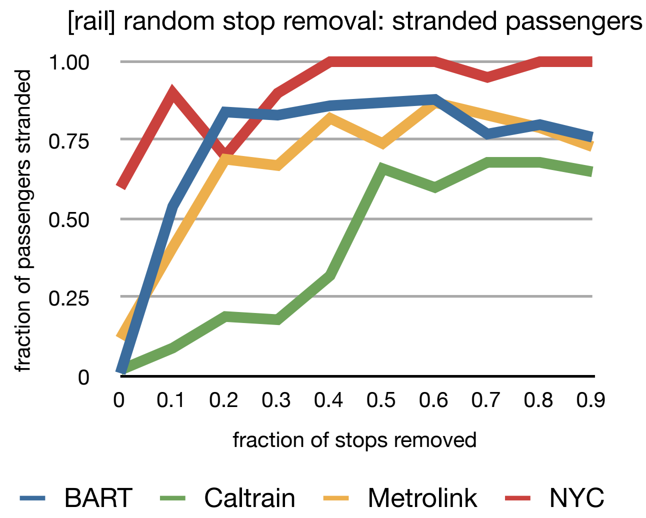


Network size plays a similar role for betweenness as it does for clustering. We observe a dramatic decrease between very small networks ($n < 100$) and other networks. Beyond this, B exhibits a small decrease as network size increases above $n = 100$.

Random stop removal

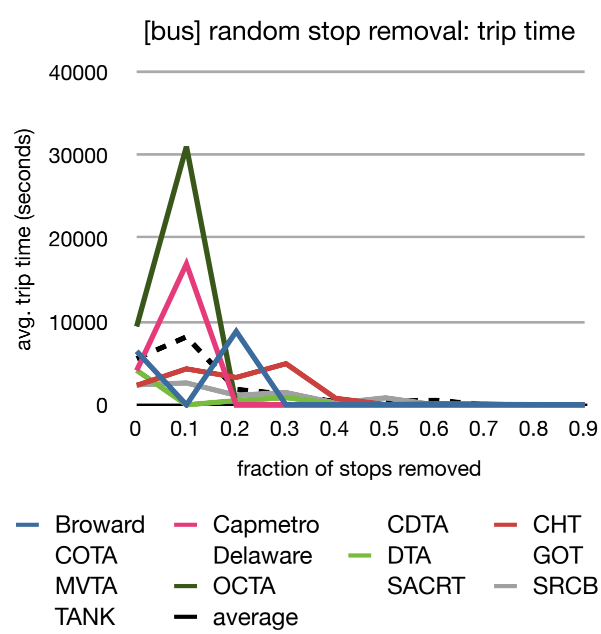
In this study, our goal is to investigate the characteristics of transportation networks both as static incarnations of complex networks and as dynamic real-world entities working to meet the needs of passengers on a daily basis. To pursue the latter area of study, we have subjected each network to three robustness tests and monitored the effect on passengers. The first of these tests is a random removal of stops.

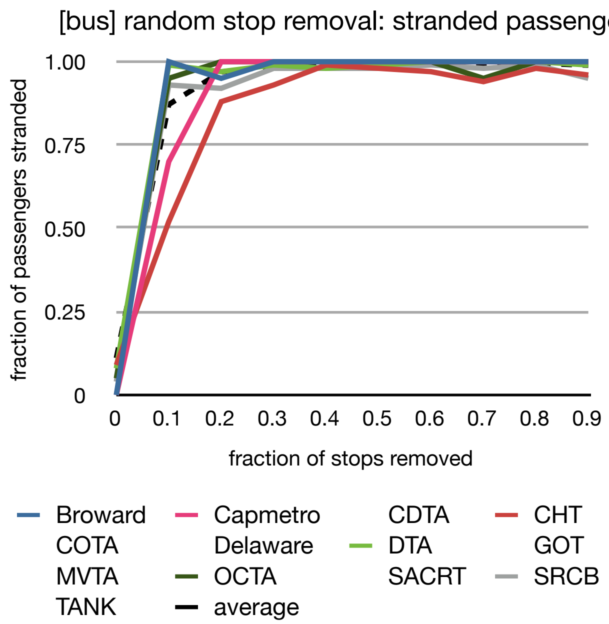




We see the effects of random stop removal on rail networks. These effects are measured both in terms of average passenger trip time and fraction of passengers stranded, in accordance with the passenger action model presented earlier.

As predicted in [4], removing stops is highly detrimental to the network. After only 20% of stops are removed, the majority of rail networks become fragmented to the point where reaching one's destination is very difficult. At the same time, we see that some networks are more resilient toward this decomposition than others.



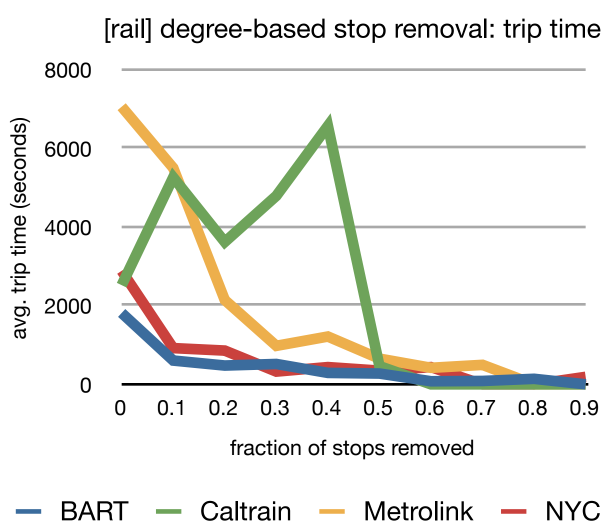


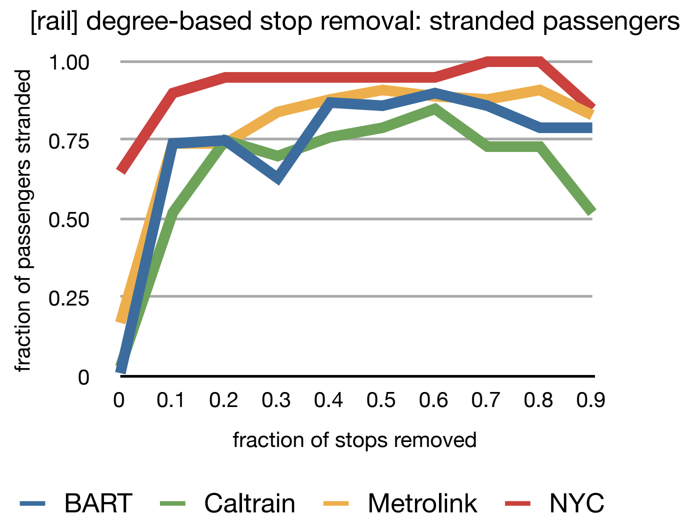
The above graphs display the average network behavior as a dotted black line and show specific data only for networks that are outliers for this test.

We observe that removing 10% of stops from almost all bus networks has the effect of crippling the network's ability to serve passengers to an extent not seen with rail networks at 10% removal. CHT is shown to be the most resilient bus network in this test.

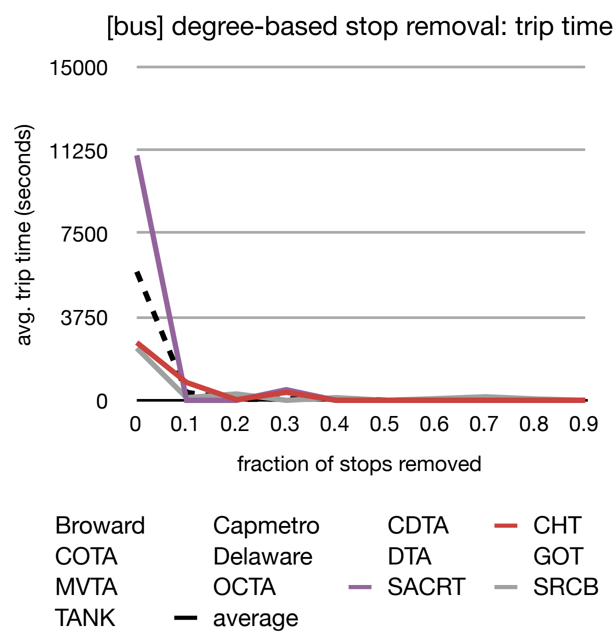
Degree-based stop removal

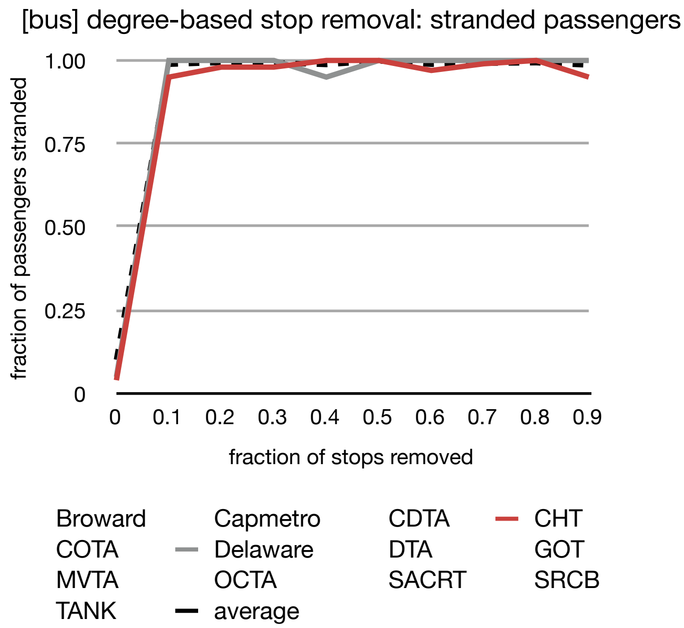
We now perform a second test on each network by removing stops in order of decreasing degree.





We observe that removing stops in order of decreasing degree does not produce a dramatically different result than removing stops randomly for most rail networks. The exception to this trend is Caltrain, which displayed a relatively high level of resilience toward random stop removal, but exhibited average resilience toward degree-based stop removal. This is likely because Caltrain is a linear system of standard and express trains. By removing the high-degree stops first, we remove the local transfer junctions and fragment the system to an extent not seen with random stop deletion.



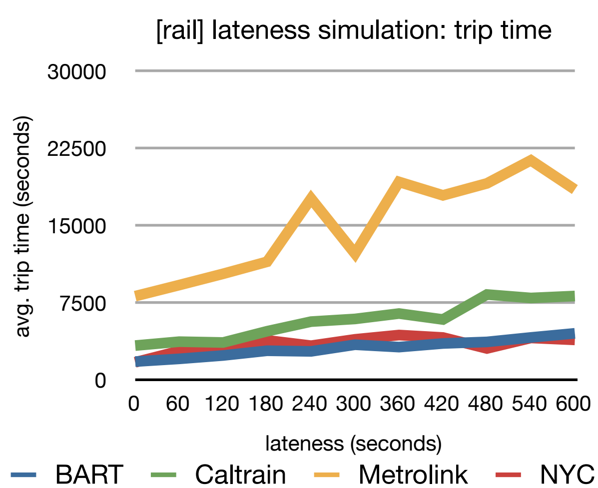


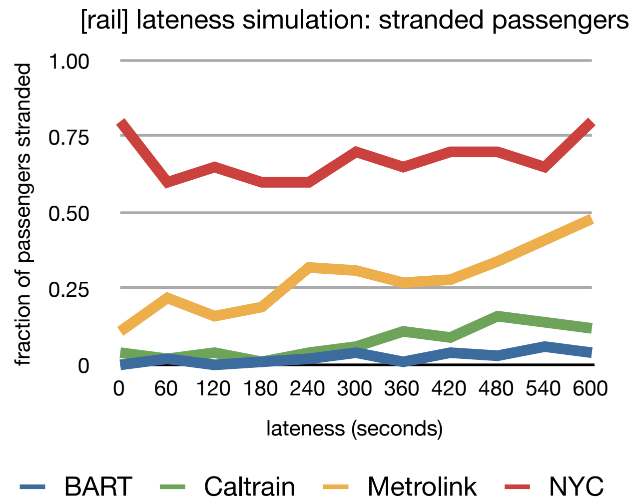
As with random stop removal, the above graphs display an average of all bus networks and any outliers.

With bus networks, degree-based stop removal is found to not be an effective measure of resilience. By the time 10% of stops are removed, every bus network tested will become fragmented and unusable by passengers. As with random stop removal, CHT is slightly more resilient than other bus networks.

Lateness simulation

The final test of network resilience performed in this study is a simulation of vehicle lateness. At each step in the simulation, 10% of vehicle trips are made 60 seconds late to their destination.





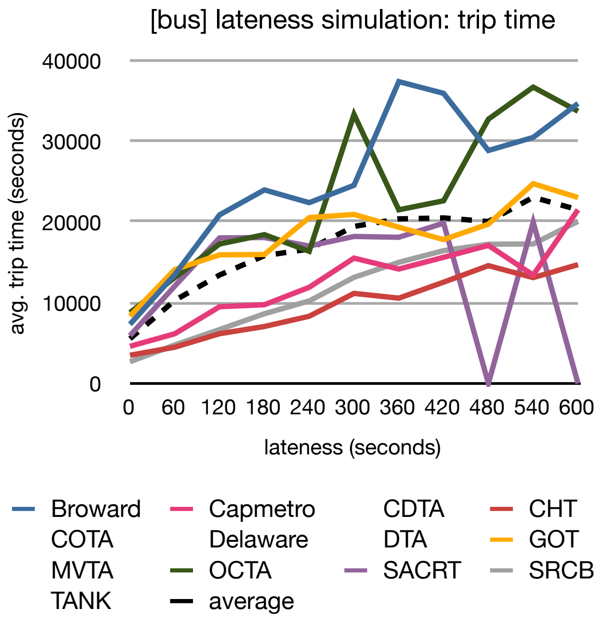
Because this test does not result in the widespread fragmentation of the network, we can observe more gradual changes in network utility. Linear regression across these results yields a trip time growth factor, k_t, and a stranded passenger growth factor, k_s (Table 2). These factors provide a standard metric for robustness, which we can use to compare networks.

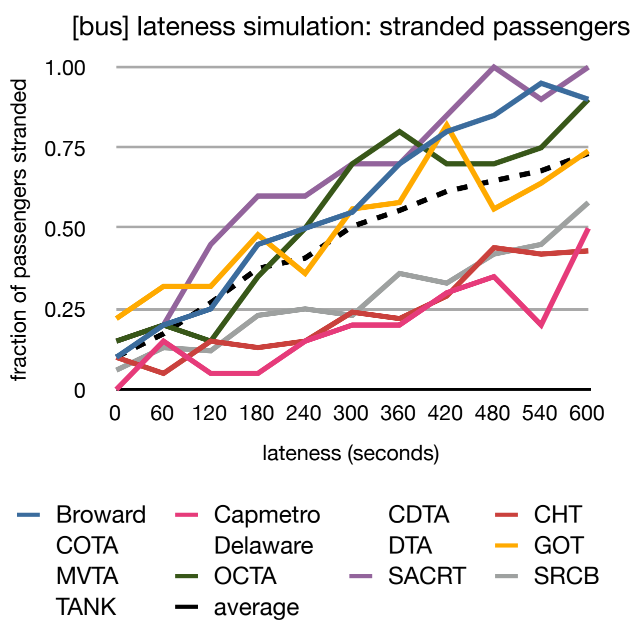
network	k_t	k_s
BART	4.21	0.76
Caltrain	8.80	2.23
Metrolink	21.37	4.97
NYC	2.78	0.91

Table 2: Trip time growth factor (k_t) and stranded passenger growth factor (k_s) for rail networks across lateness simulation.

We observe that Metrolink is significantly less robust than other rail networks in response to late trains, while BART and NYC avoid significant disruption.

Despite NYC's above-average performance in the lateness simulation, it exhibits a much higher fraction of stranded passengers than any other network, even in its unperturbed state. This is likely due to a limitations of this study where a passenger is considered stranded if they cannot reach their destination by the end of the day. The New York City Subway system is one of a few public transportation agencies in the United States to operate continuously, 24 hours per day. Because our passenger action model cuts off at midnight, it is unable to fully represent NYC.





As in previous tests, the above graphs display an average of all bus networks along with any outliers.

We expect trip time and stranded passengers to monotonically increase during the lateness simulation, but many networks defy this expectation. This could be an artifact of the random elements of the passenger action model, but it is also likely that there are cases where one or more trips running late actually represent an advantage for passengers.

Consider a stop that is visited by three buses (b_a, b_b, and b_c) at times t=0, t=2, and t=3, with a fourth bus, b_x, scheduled to visit at t=1 before continuing on to a structurally important destination. As scheduled, b_x will only be able to transport passengers from b_a to the important destination. However, if b_x runs late and visits the stop at t=4 instead, it will be able to transport passengers from b_a, b_b, and b_c to the important destination.

network	k_t	k_s
Broward	40.54	14.70
Capmetro	22.79	6.29
CDTA	20.11	8.41
CHT	19.56	6.65
COTA	27.85	10.91
Delaware	17.75	9.92
DTA	28.31	12.44
GOT	19.69	8.33
MVTA	28.51	11.53
OCTA	42.27	13.03
SACRT	-7.19	14.09
SRCB	28.65	7.71
TANK	34.02	12.70

Table 3: Trip time growth factor (k_t) and stranded passenger growth factor (k_s) for bus networks across lateness simulation.

Corresponding to results from previous tests, CHT continues to be more resilient than the majority of other bus networks. It is also notable for having the second-highest degree-degree correlation of all networks studied, with the highest being for COTA, a much larger network. Indeed, CHT is the second-smallest of the bus networks at n = 623, larger than only SRCB. CHT also has a below-average average path length, above-average clustering coefficient, and above-average betweenness for a bus network. All of this information portrays a small, strongly connected, and efficient transportation network.

Overall, we see greater k_t and k_s with bus networks than we do with rail networks. This follows the trend seen with random and degree-based stop removal, and is somewhat surprising. Intuitively, rail networks would seem to be less flexible with regard to missing stations or late trains, due to the sparsity of their stops and stop times relative to bus networks.

6. CONCLUSIONS

Public transportation systems present a wealth of opportunities for analysis and simulation. This is becoming increasingly true as more transit agencies make data about their systems available to developers in the GTFS format. The analysis presented in this work is intended to be a starting point for general analysis of arbitrary transportation networks, which we believe is possible through the extension and refinement of the tools and routines developed here.

One such refinement is necessary in relation to the stop removal simulations presented above. Because nearly all bus networks were severely fragmented by the time 10% of their stops were removed, a more delicate approach is called for. Such a refined approach would ideally allow for more effective differentiation among bus networks than was possible in this study's stop removal simulations.

Additionally, we envision multiple extensions to our passenger action model. One such extension would be to incorporate power-law human movement distributions into the destination selection process, as detailed in [2]. Another extension would use fare information available in GTFS datasets to incorporate financial cost into the shortest path calculation. Finally, it could be beneficial to extend the stranded passenger metric to account for the distance remaining to the destination, which would provide more information than the current binary stranded/not-stranded implementation.

7. REFERENCES

[1] U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematics Sociology*, 25, 2001.

[2] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439, 2006.

[3] P. Gruer, V. Hilaire, and A. Koukarn. Multi-agent approach to modeling and simulation of urban transportation systems. *IEEE International Conference*, 2001.

[4] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Physical Review Letters*, 87, 2001.

[5] D. Meignan, O. Simonin, and A. Koukam. Simulation and evaluation of urban bus networks using a multiagent approach. *Simulation Modelling Practice and Theory*, 15, 2007.

[6] K.A. Seaton and L.M. Hackett. Stations, trains and small-world networks. *Physica A*, 339, 2004.

[7] P. Sen, et al. Small-world properties of the Indian railway network. *Physical Review E*, 67, 2003.

[8] J. Sienkiewicz and J.A. Hołyst. Statistical analysis of 22 public transport networks in Poland. *Physical Review E*, 72, 2005.