

Vulnerability of Urban Road Systems

A Network Analysis of the Baltimore Bridge Collapse

Gavin Engelstad

Spring 2024

Abstract

This paper used network techniques, especially centrality analysis, to examine the implications of the Francis Scott Key Bridge collapse. We analyze these effects on the road network for the entire Baltimore Metro Area, using intersections as nodes and road segments as edges. We found the global network effects of the collapse were limited, but that there significant local effects on intersections near the bridge on the ability to traverse the network and on routes extending from the bridge on traffic. We also employ a utilization-weighted analysis to find similar but even more localized results.

1 Introduction

On March 26, 2024, the container ship Dali struck one of the concrete supports holding up the Francis Scott Key Bridge and caused it to collapse. The incident caused six deaths¹ and caused trade through the port of Baltimore, one of the largest ports in the country, to grind to a halt costing an estimated \$15 million per day (**Hassan24AlJazeera24**).

Pre-collapse, the Francis Scott Key Bridge was an essential part of the Baltimore road system. It was the final link completed in Interstate 695, or the Baltimore Beltway, a major highway that wraps around the City of Baltimore. Spanning the Port of Baltimore, the bridge was 1.6 miles (2.6

¹. Only four of the six bodies have been found. Still, the remaining two were on the bridge during the collapse and are presumed dead (**Hassan24**).

km) across and the third-longest continuous truss bridge in the world (**Parry24**). Every day, approximately 30 thousand vehicles crossed the bridge, making it an important piece of infrastructure for thousands of commuters every day (**MDOT24**). Post collapse, the Maryland Department of Transportation issued a traffic warning and has observed a 7-11% increase in traffic along alternate major highway routes across and around the port (**MDOT24Domen24**).

This paper will examine the effects of the bridge collapse on the Baltimore Metro Area (BMA) road system using tools from complex network analysis. The rest of the paper will proceed as follows: Section ?? will explore existing literature on the use of complex networks to analyze road network structures and robustness. Section ?? will outline the methods we'll use to analyze the effects of the bridge collapse on the road network. Section ?? will present the data sources that we'll use to create and examine the network. Section ?? will give the results to our analysis. Finally, section ?? will conclude.

2 Literature Review

Road systems have a long history of being represented as a network. The most popular method of this, the primal approach, uses intersections as nodes and roads as links in the network (**Porta06Ding19**). This method is seen as more comprehensive, realistic, and feasible than the alternative dual approach to modeling the network (**Porta06Porta06b**). This approach is implemented into many software packages and databases (**EsriBoeing17**) and has the advantage that it allows for analysis that considers distance, versus solely looking at the topological properties of the network (**Ding19Erath09Wang11Barthelemy09**).

Using this network, centrality measures can be very predictive of the real world context around roads and intersections. Typically, multiple centrality assessment (MCA) methods that analyze a number of centrality measures are used to give a comprehensive view of the network (**Porta06Porta07**). Closeness centrality, when done globally, is highly correlated with land-use, but fails to be meaningful looking at only a subset of a larger network (**Wang11Porta06Barthelemy09**). Betweenness centrality is very effective at identifying important parts of longer routes to traverse large urban networks (**Porta06**). Straightness centrality, a modification of closeness centrality

where the distance along the network is divided the Euclidean distance between nodes, identifies both areas with high land-use and important routes while being robust on network subsets (**Porta06Wang11**). Information centrality, which measures the drop in global efficiency when a piece of the network is removed, is a comprehensive measure of node importance and identifies nodes that have high centrality from all other measures **Porta06**. Eigenvector centrality is able to identify high-traffic areas and road connectivity (**Jayaweera17Ando21**).

These tools are useful to analyze network robustness. On its own, eigenvector centrality can identify nodes in a network that are vulnerable to the removal of other nodes (**Ando20**). Analysis of how stable centrality measures are when nodes are added or removed, a technique called centrality interference or centrality robustness, can act as a measure of road network robustness (**Scardoni13**).

Other tools to analyze road network robustness look at the topological properties of the network after an attack, defined as the removal of edges or nodes in the network (**Holme02**). The clustering coefficient, serves as a measure of reliability of a network, and how much it changes after an attack can show how much the attack affects network connectivity (**Xeumei10**). The change in the average shortest path also serves to demonstrate how robust people's ability to travel along the network is to an attack (**Xeumei10Kaub24**).

This paper will implement these methods on the Baltimore road network, utilizing MCA combined with centrality interference methods and analyzing how the average shortest path changes before versus after the bridge was destroyed. Most studies that implement these methods do so at a theoretical level and attempt to identify the edges and nodes where the network is most vulnerable to an attack (**Xeumei10Ando21Masuccia16Julliard15Holme02**). The handful of studies that implement these methods that are inspired by real events tend to analyze the effects of disasters that have affects that span the entire network (**Kaub24Sakakibara04**). Therefore, this paper's application of these methods to the real-world removal of a single, important edge of a city's road network is, to my knowledge, new.

3 Model

3.1 Network Structure

The first way we'll analyze the robustness of the network is through comparing the structure of the network before versus after the destruction of the bridge.

At a global scale, that means we'll be comparing the average path between nodes. To calculate this, we'll use

$$\bar{d} = \frac{1}{n(n-1)} \sum_{i \in N} \sum_{j \in N, j \neq i} d_{i,j}$$

where n is the number of nodes in the network, N is the set of all nodes in the network, and $d_{i,j}$ is the distance between nodes i and j traveling along the network. This will tell us how far an individual has to travel to get from one randomly chosen intersection to another. Any change in this that we observe would suggest the bridge collapse has made it harder to travel across the network, though, since the event only affected a small part of the road network, we don't expect any observed change to be particularly significant.

Because of this likely smaller global effect, we also want to study the local effects of the bridge collapse within the network. Therefore, we'll also look at node level centrality measures and node-node pair level travel lengths.

For centrality measures, we'll look at the eigenvector centrality (C_i^E), betweenness centrality (C_i^B), closeness centrality (C_i^C), and straightness centrality (C_i^S). The first three of these are standard in complex network analysis and tell us whether a node is connected to important nodes (C_i^E), between other nodes (C_i^B), and close to other nodes (C_i^C). For C_i^E , we'll follow **Ando20**, where roads were weighted based on capacity, and weight edges based on the type of road they represent. For C_i^B and C_i^C , we'll weight edges by their lengths.

Straightness centrality is a measure specific to the study of spacial networks, networks where the nodes and edges have a physical location, that tells us how directly a node can get to other

nodes along the network. Introduced in **Porta06**, it's calculated using

$$C_i^S = \frac{1}{n-1} \sum_{j \in N, j \neq i} \frac{d_{i,j}^{\text{Eucl}}}{d_{i,j}}$$

where $d_{i,j}^{\text{Eucl}}$ is the Euclidean distance between the points. Intuitively, a straightness centrality of 1 implies a node has a straight connection to all other nodes along the network, and anything lower tells us how inefficient traveling along the network is relative to just taking the straightest route.

To use these centrality measures to analyze robustness, we'll analyze how much they change when the Key Bridge is removed from the network. This gives us a method of analyzing how individual nodes on the network were affected by the bridge collapse and where the most affected nodes are located.

To get an even more specific look at the implications of the collapse, we'll also look at how travel lengths between individual nodes are affected. To do this we'll compare $d_{i,j}$ for all $i, j \in N$ to see which routes were the most affected and what areas of the city these routes go through. This will help give us an idea of which trips the collapse will impact, and, therefore, the types of people that will be the most affected by the collapse.

3.2 Network Utilization

Analysis of the overall structure of the Baltimore road system will give us a general idea of how the network, nodes, and paths were affected by the bridge collapse, but it doesn't take into account how people actually use the road network. People aren't evenly spread throughout a city and equally likely start and end their journeys at all nodes.

To fix this, we'll use commute data to weight trips based on how often the routes are actually being traveled. Using this new measure, we'll look at the average shortest path length and equivalences to betweenness, closeness, and straightness centrality. Because this new model changes the paths taken instead of the structure of the network, there's no useful analogue to eigenvector centrality.

Our utilization-sensitive shortest path length will be calculated as

$$\bar{d} = \frac{1}{\sum_{i \in N} \sum_{j \in N, j \neq i} n_{i,j}} \sum_{i \in N} \sum_{j \in N, j \neq i} n_{i,j} d_{i,j}$$

where $n_{i,j}$ is the number of people traveling from node i to j .

Use-weighted betweenness centrality will be calculated as

$$C_i^B = \frac{1}{\sum_{j \in N, j \neq i} \sum_{k \in N, k \neq i, j} n_{j,k}} \sum_{j \in N, j \neq i} \sum_{k \in N, k \neq i, j} n_{j,k} \frac{p_{j,k}(i)}{p_{j,k}}$$

where $p_{j,k}$ is the number of shortest paths between j and k and $p_{j,k}(i)$ us the number of shortest paths between j and k that cross i . Therefore, we're measuring the number of traveled paths that go through i , not just possible paths like with normal betweenness centrality.

Use-weighted closeness centrality will be

$$C_i^C = \frac{\sum_{j \in N, j \neq i} n_{i,j}}{\sum_{j \in N, j \neq i} n_{i,j} d_{i,j}}.$$

The interpretation of this is a measure of how far individuals starting at i need to travel.

Finally, use-weighted straightness centrality will be

$$C_i^S = \frac{1}{\sum_{j \in N, j \neq i} n_{i,j}} \sum_{j \in N, j \neq i} n_{i,j} \frac{d_{i,j}^{\text{Eucl}}}{d_{i,j}}.$$

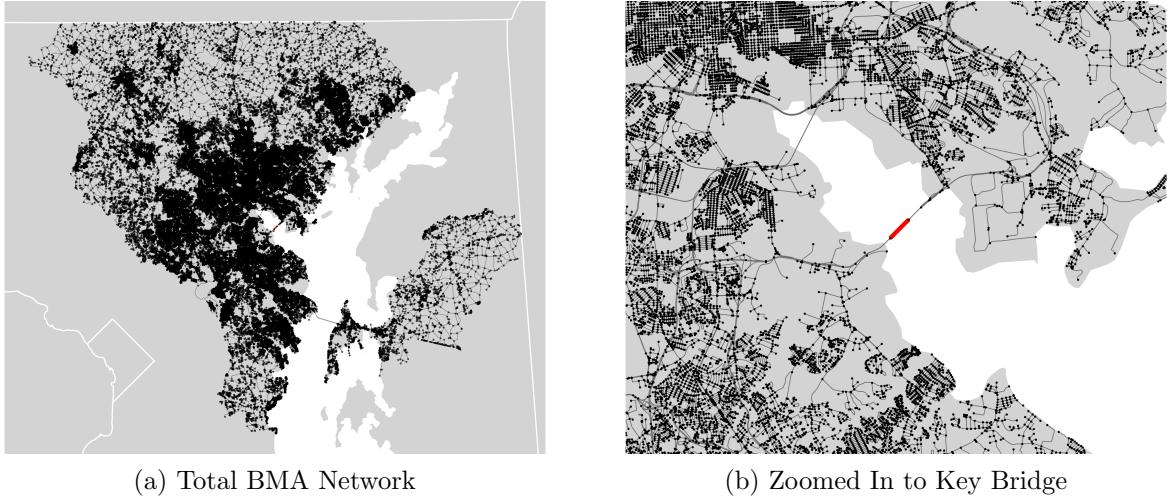
This can be interpreted as how direct the paths of individuals starting at i are.

Together, this will give us tools to analyze the ways the collapse affected how people use the road system in the Baltimore Metro Area, not just the structure of the network.

4 Data

For our analysis, we'll use a road system from Open Street Maps provided by the Python API from **Boeing17**. This API allows us to get a directed network of roads and intersections for almost anywhere in the world, along with basic statistics about each section of road, like the exact

Figure 4.1: Baltimore Metro Area Road Network



Note: Key Bridge highlighted in red.

coordinates, length, road type, and number of lanes.

In our study, we'll focus on the 7 counties that make up the Baltimore metro area: Anne Arundel, Baltimore City, Baltimore County, Carroll, Harford, Howard, and Queen Anne's. To make our analysis feasible and meaningful, especially the eigenvector centrality and average shortest path, we exclude any intersections that aren't strongly connected to the largest connected component. This results in us excluding 272 intersections in the network, mostly near the county boundaries.

The resulting network has 91,300 intersections (nodes) and 218,842 road segments (edges), 4 of which are on the Key Bridge and were affected by the collapse. A visualization of this network is shown in Figure ???. We can see the road network varies substantially in density, and is much denser near the city and to the South nearer to Washington DC. The zoomed in view of the Key Bridge shows the bridge plays an important role in the network. When the bridge is removed, paths that normally go across it will need to be diverted moderately far West towards downtown. Therefore, we can expect a significant number of shortest paths to be rerouted and affected.

For data on network utilization, we'll use census tract level journey-to-work flows from the US CENSUS (**Bolton20**). This data gives bilateral flow estimates for the number of daily commuters from each census tract to every other census tract from 2012 to 2016. This isn't the time period

we're interested in, but is the most recent data available. Due to the bounds of the road network we're working with, we only include commuters who both live and work within the BMA. This excludes 11.1% of commuters who work in the BMA and 12.2% of commuters who live in the BMA.

The dataset includes 678 tracts within the BMA. 672 of these are home to at least one commuter and all 678 are the workplace of at least one commuter. Together, there are 41,272 residence-workplace pairs which 1.02 million commuters travel between. The maps in Figure ?? shows the commuter volume into and from census tracts in the BMA. We can see that commuter residences are somewhat evenly distributed throughout the area, but workplaces are much more central. When we calculate paths commuters are taking along the network, this will mean that the paths will point to a much more central location, which may limit how many go across the relatively off to the side Key Bridge. The density plot in Figure ?? shows that commuter flow has a significant right-skew which even appears on a log-scale plot. This means that commuter flows are, typically, fairly low, but there are tract pairs with much higher flow.

We are limited by the granularity of this data. Census tracts are very large compared to the intersection/road segment scale of the network, which means we can't perfectly estimate the paths commuters are taking.² In our analysis, we'll estimate paths with commuters starting and ending at the intersection closest to the center-of-mass of the census tracts they're traveling between and assuming all commuters travel along the road network. This should give us a reasonably accurate estimate for how commuters are traveling, although will have some errors, especially closer to the center of mass for the higher-volume census tracts.

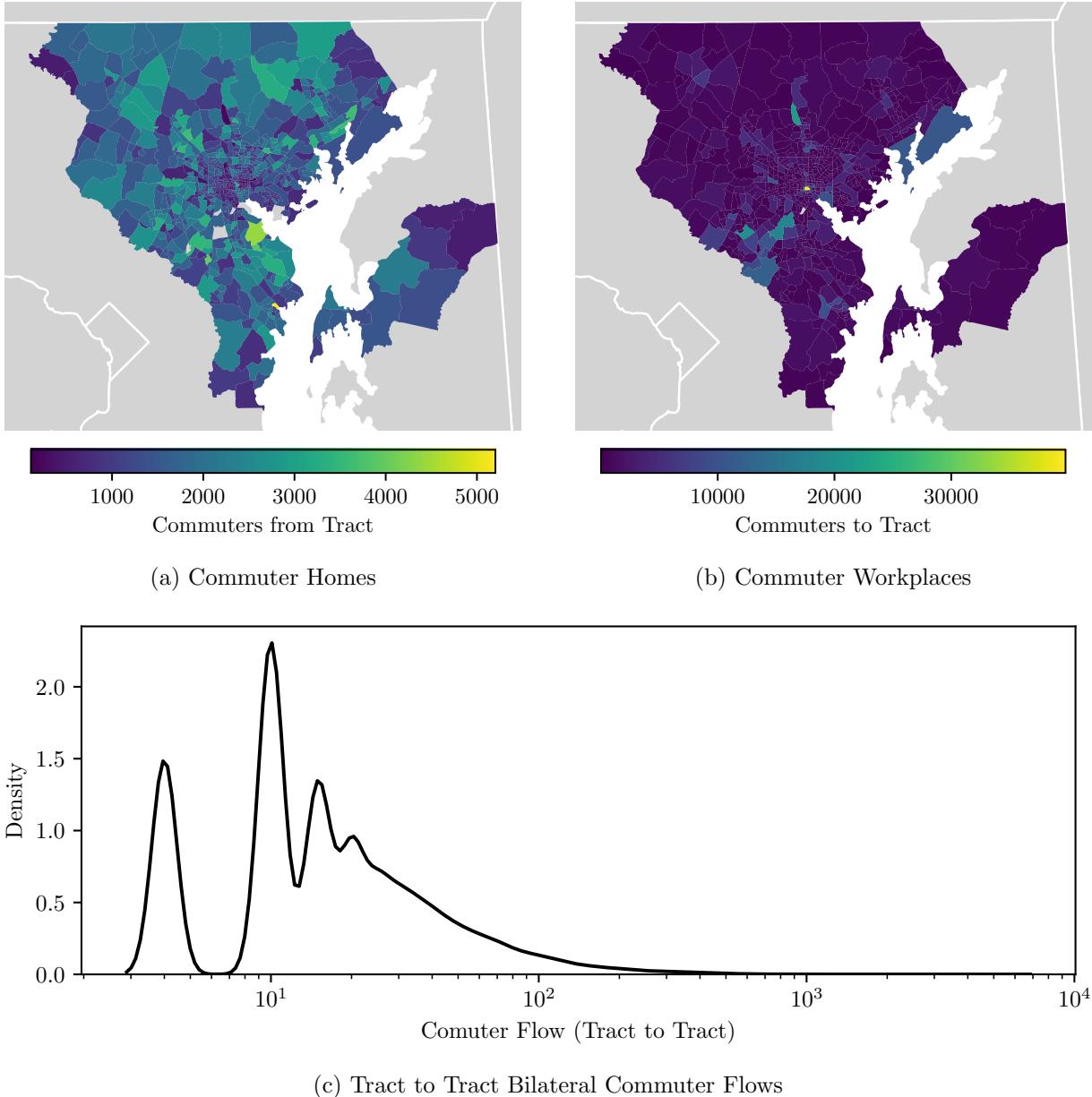
5 Results

5.1 Network Structure

The average shortest path length along the network is 26.3 miles (42.3 km) with the bridge and 26.4 miles (42.5 km) without the bridge. That means the destruction of a single road increased the

2. But definitely is a positive from a privacy standpoint.

Figure 4.2: Baltimore Metro Area Census Tracts by Commuter Volume and Flow

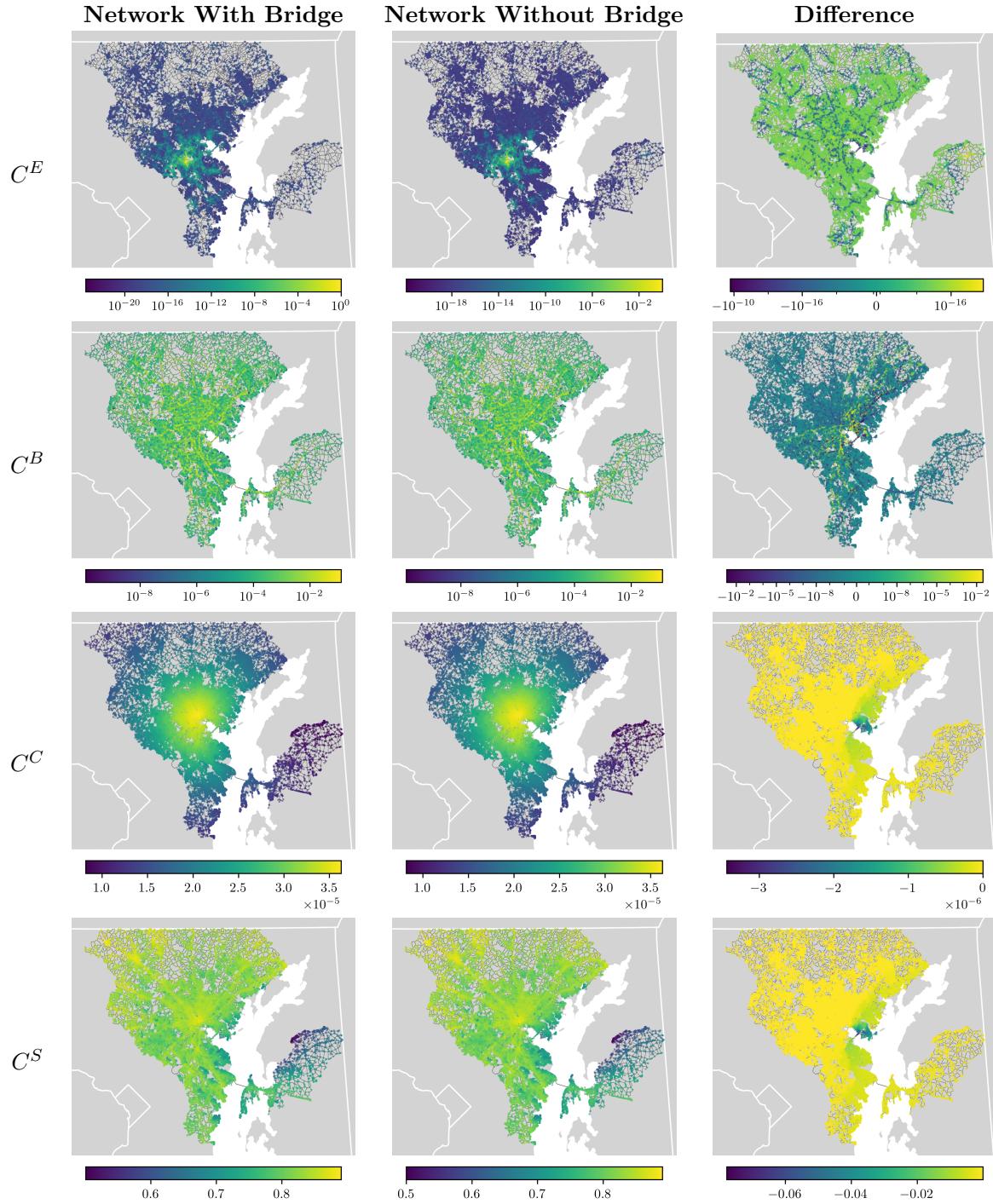


average shortest path length by 0.09 miles (0.14 km). This difference of only 0.003% is relatively small, and suggests the network-wide implications of the event were small.

At a node level, Figure ?? shows the calculated centralities for intersections in the BMA with the bridge, without the bridge, and the difference between the two.

Both with and without the bridge, the eigenvector centralities are the highest just south of

Figure 5.1: Node centrality before and after the bridge collapse.



Difference calculated a $C_i^{\text{Without Bridge}} - C_i^{\text{With Bridge}}$. Log-scale used for Eigenvector and Betweenness centrality and a symmetric log-scale used for their difference. This meant any nodes with 0 weren't plotted, which was between 9,000 and 35,000 nodes depending on the measure.

downtown Baltimore. This area has a handful of larger neighborhoods and workplaces (check out the maps on Figure ??) and BWI airport, consistent with the idea that eigenvector centrality is correlated with areas of high use from **Jayaweera17**.

Without the bridge, the centrality effects are spread throughout the network. The nodes near the bridge are the most affected, but throughout the network a connected web of nodes also decreases in centrality, stretching out in all directions. Aside from these, nodes generally face a slight increase in centrality, suggesting the importance of these affected nodes was spread throughout the network, though there are a handful of nodes in Queen Anne's county to the east and downtown that face a disproportionate increase in centrality.

The betweenness centrality is highest for nodes at the center of the network and at the ends of bridges, especially the one connecting Queen Anne's county to the East with the rest of the BMA. From these, there are clear paths of high centrality that extend through the rest of the city. Following **Porta06**, these are likely important paths that are frequently used to traverse the network.

Before it collapsed, the nodes along the Key Bridge were all very important, and paths extended from it along the shore. After the collapse, these nodes became much less important, shown in the path of purple nodes on the difference graph that crosses the bridge and stretches up and down the coast. Instead, paths of increased centrality extend from the Fort McHenry Tunnel (I-95) and the Harbor Tunnel (I-895), which crosses the port slightly farther inland than the key bridge, and downtown, which sits at the end of the port.

The closeness centrality for both graphs is centered around the downtown area, likely because it sits at the center of the city. From here, it gradually decreases, and is lowest in corners of the city, especially in Queen Anne's county to the East. This highlights the problems with using closeness centrality on a subset of a network, since the inclusion of other counties surrounding the BMA would likely change where this center of mass is. Still, the change in the closeness centrality will still be a meaningful indicator of the network implications of the bridge collapse, since nodes are all located in the same place relative to the boundary of our network before and after the event meaning the bias from this boundary effect should be equivalent on both networks.

The effects of the bridge collapse are almost entirely concentrated near the end of the bridge toward the bay. This makes sense, since, like we saw with the minimal change in average shortest path, places aren't that much farther in general, which would suggest closeness centrality should change by a minimal amount for most nodes. The nodes that were affected, however, now have to take a route around the port to get to nodes on the other side, when before they would have just taken the bridge. The nodes to the north were especially affected, likely because their routes south across the port now need to start going north to get around it.

The straightness centrality is fairly uniform across both graphs, except in areas where water makes it hard to get to a large portion of the network. This is especially apparent for Queen Anne's county, where the lowest centrality nodes are. Within the rest of the network, there are paths with slightly higher straightness extending out from the center of the network, suggesting these might be either important routes for traversing the network or important areas within the network (**Porta06Wang11**).

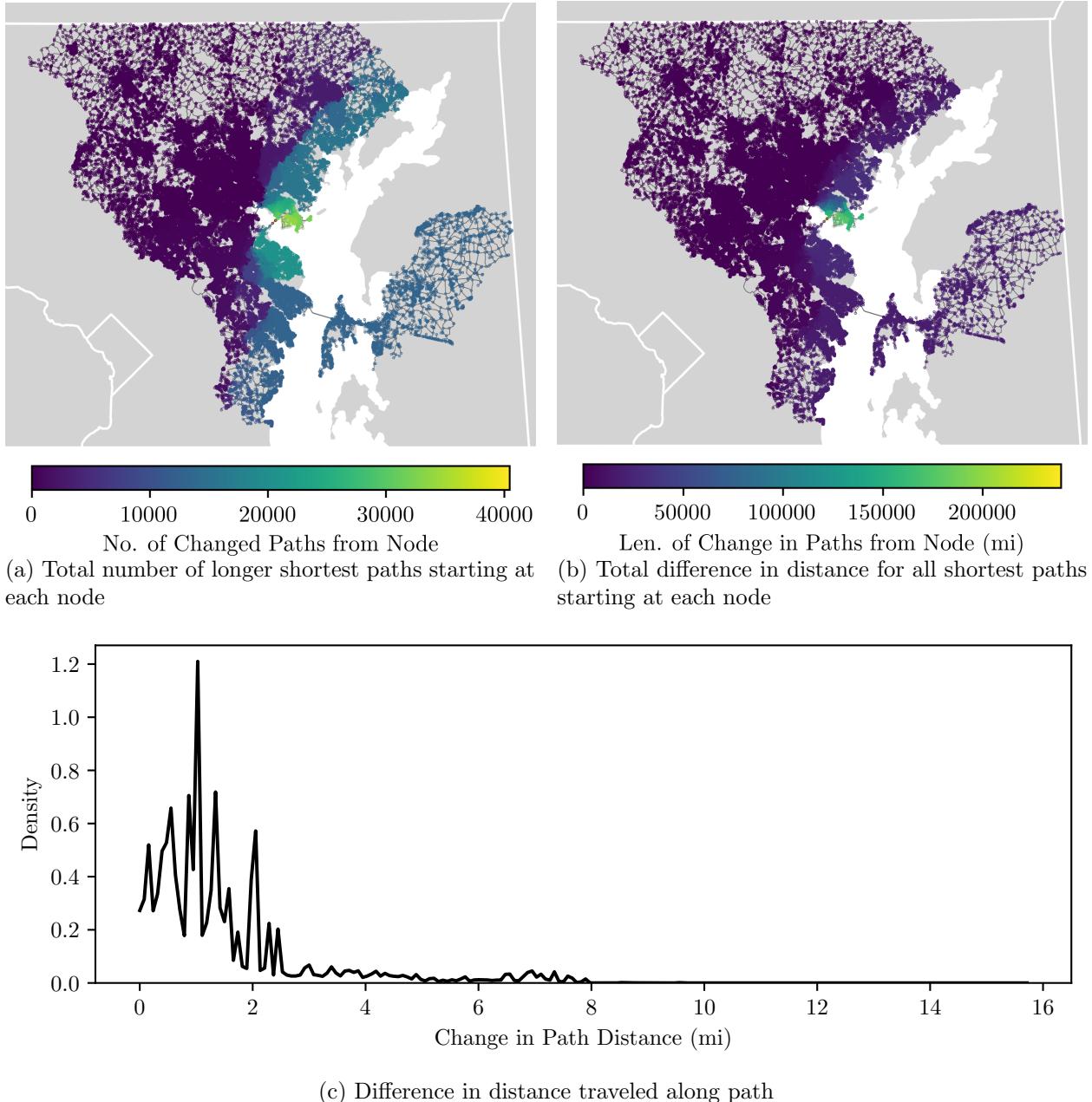
Like closeness centrality, the effects of the event on straightness centrality are almost entirely felt by nodes near the event, and especially to the north. The rationale behind this effect is the same: paths that previously would have taken the bridge now have to take a much less direct route around the port, making their path less straight.

Importantly, the bridge collapse didn't benefit the closeness of straightness for any intersections in the network. This makes sense, since the shortest path can't ever be made shorter by removing an edge from the network, but does contrast the other two centralities where there are winning and losing nodes.

Altogether, our MCA analysis suggests that the effects to the bridge collapse on the ability to travel from a node across the network are concentrated closer to the event, but global effects are seen in the specific paths taken that affect traffic through would-be-unaffected intersections.

To analyze the affected paths, we'll focus on the 475 million paths that were impacted, roughly 0.57% of the possible paths in the total network. This is done both for tractability reasons, there are more than 80 billion node-to-node paths in the network as a whole which would pose significant problems for computability, and analysis reasons, since the effect of these 99.43% of paths that

Figure 5.2: Changes in shortest paths



are unaffected is seen in the minuscule change in the average shortest path length. These shortest paths are shown in Figure ??.

The maps show that nodes all along the coast had significant numbers of paths that would have gone over the bridge and were affected, but these only significantly affect travel distances to other

Table 5.1: Summary statistics for the difference in affected paths (miles).

Count	Mean	St. Dev.	Min	25%	Median	75%	Max
4.750×10^8	1.525	1.514	1.864×10^{-6}	0.559	1.049	2.018	15.716

nodes closer to the bridge and the port, since by cutting in at a slight angle someone traveling from farther north or south can get around the port without using the bridge while adding little distance to their route, while someone traveling from closer has to cut straight west first.

The distribution of path distances has a strong right skew and a lot of noise, this suggests that most affected paths weren't lengthened significantly, but a handful were. This suggests the impacts of the bridge collapse are going to be localized to people traveling a handful of routes, and not necessarily felt by everyone in the BMA area.

Summary statistics for this distribution are shown in Table ???. The mean and standard deviation are similar, evidencing the wide spread of the distribution. The mean is larger than the median, which, again, suggests there is a right skew.

5.2 Network Utilization

Based on our model, the average commute to work is 11.43 mi (18.39 km) before the bridge collapsed and 11.44 mi (18.40 km) after. This is less than half of the average shortest path along the network, and suggests that people are choosing to live and work closer together than would be expected if both choices were independent, meaning that our structural analysis might overweight long trips. That means the bridge collapse caused the average commute distance to increase by 0.010 miles or 0.016 km, more than we predicted without the weighting by use, but still a fairly insignificant amount.

For our centrality measures, two of them, closeness and straightness, require flow from the node, meaning we can only find them at a tract level, while the other one, betweenness, is still found at a node-level. These statistics are shown in Figure ??.

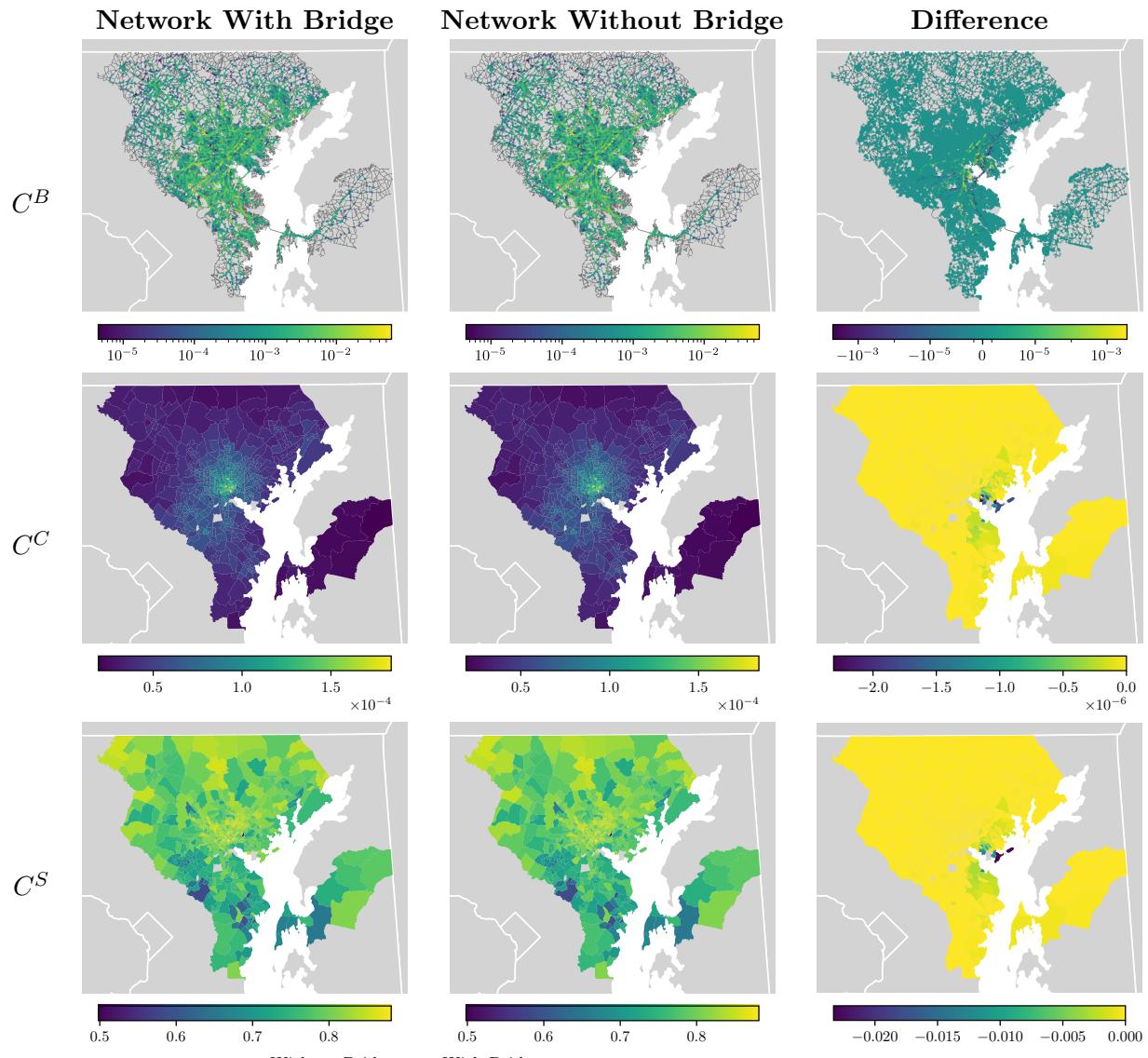
As expected by excluding most nodes as possible origin points, there are quite a few more nodes with 0 betweennesses, making the network more sparse. Still, we see similar patterns as the

standard betweenness centrality since high-betweenness paths extend from high-betweenness nodes in the center and near bridges, likely representing important routes to travel along.

The effects of the bridge collapse are much more pronounced on the nodes they affect, but affect less nodes. In Figure ??, there are faint routes of different centrality that span the whole network, but here only specific routes, namely along the coast are affected.

Closeness centrality is, again, concentrated around the center, though it tapers off much more

Figure 5.3: Node centrality before and after the bridge collapse.



Difference calculated as $C_i^{\text{Without Bridge}} - C_i^{\text{With Bridge}}$. Log-scale used for Betweenness centrality and a symmetric log-scale used for its difference. This meant approximately 57,000 nodes with 0 centrality weren't plotted.

quickly. This could be caused by people who live downtown valuing living and working closer together more than those that live in other parts of the city or the suburbs, causing them to be closer to where they work.

The effects of the bridge collapse are even more localized when accounting for use than without. In Figure ??, the effects of the collapse can be seen in a wide range of nodes near the bridge, but when accounting for use quite a few census tracts in area with affected nodes are unaffected. The couple tracts that area affected, however, see an effect that's similar in magnitude to before. This suggests only significant amount of people from a couple, nearby tracts cross the bridge in their commute.

The use-weighted straightness centrality exhibits far more variability than before. In Figure ??, there is a little variability between nodes within geographic areas, but the variability between geographic areas is far more significant. Here, much of the variability between geographic areas goes away, possibly because people are working near where they live, but within nearby tracts in geographic areas varies significantly more.

Like with closeness, the effects of the collapse are very localized to the area just north of the port near the bridge relative to the network-wide estimations in Figure ???. Again, this suggests that only a handful of tracts are home to significant numbers of commuters that used the bridge.

Altogether, our use based analysis suggests that the effects of the collapse are still significant, but very localized to commuters that live in certain areas and that the traffic going through intersections is only affected in a handful of routes, not in a web that spans the network like we found without using use-weighted statistics.

6 Conclusion

Through our network analysis of the Key Bridge collapse, we found significant local impacts of the Bridge Collapse. On average, paths across the network only increase slightly in length. However, in local areas, especially near the bridge, we see significant impacts. This is evident in the change in closeness and straightness centrality, especially when weighted by use where the impacts become even more localized than would be suggested in an analysis of the network as a whole. The

betweenness and eigenvector centralities suggest that the implications of the blockage will propagate and affect traffic and importance for intersections in a web throughout the network, especially along the coast. The origins of affected path lengths again suggest a very localized effect, and the right-skew of the distribution suggests, again, that only specific paths are the most impacted.

The results of our analysis, especially in regard to betweenness centrality, are very consistent with behavior patterns we're observing after the collapse. The Fort McHenry Tunnel and Harbor Tunnel, which we predicted would increase in betweenness due to the event, has experienced a significant increase in traffic (**Domen24**). This consistently with the real world suggests the use of this type of network analysis to examine the road network is valuable and paints an accurate picture of the implications of the event.

6.1 Limitations

Our analysis has a handful of limitations. Mostly because of the small size of the shock to the network, we don't examine the the change in the size of the largest connected component, the clustering coefficients, or modularity classes and community divisions, factors which are often used in the literature to analyze road network shocks (**Xeumei10**). These simply weren't affected by the removal of a single bridge, so analysis of them that's more in line with other network attack literature would be uninformative.

The network itself also has an arbitrary boundary. We chose the network to include all counties in the BMA area, but the BMA road network isn't completely disconnected from that of the rest of Maryland and United States. In that way, our analysis might not be able to analyze how far the collapse effects propagate, intersections with affected betweenness centrality extended all the way to the north and south of the network, and ignores the significant number of paths that travel between the BMA area and the surrounding counties.

6.2 Future Work

Future work could work on incorporating other use-related data sources, for example average annual daily traffic, with our network. This would help get a better view of how people's use of the network

was affected by the collapse, but isn't feasible until the end of 2024 when these numbers will be released for the time of the collapse.

Future work could also look more at what types of nodes were impacted. Knowing whether these were highways, arterial roads, or side streets could be valuable and help inform how we interpret the effects of the shock.

Our utilization analysis also only looked at paths from work to home, since in this paper we only analyzed the trips people are taking to get to work.