Path score robustness in metro system public transportation networks

Sarah Kosic[a], **Jason Lim**[a, 1], **Arnav Saxena**[a], **and Grace Chow**[a]

[a]Department of Mathematics, University of California, Los Angeles, CA 90025

This manuscript was compiled on December 9, 2022

Many people depend on public transportation networks to reliably complete daily trips. Several facets of public transportation can be studied using network science, including robustness. Robustness metrics on a transportation network typically aim to measure the availability of alternative paths in the event of node or edge failure. In this paper, using 14 global subway networks, we apply the idea of transport-based core-periphery structures and path score as introduced by Lee, Cucuringu, and Porter as a measure of robustness. (1) We find that the core-periphery structure of transportation networks is crucial to robustness from a path score perspective, with networks with larger cores and smaller peripheries being more robust, regardless of network size. We then compare this path score as a robustness metric with the robustness measure previously proposed by Derrible and Kennedy, which is based on the amount of cycles in a network. (2) We find that path score as a robustness measure, due to its structural nature, may provide a more complete picture of robustness as it applies to metro networks. As a result, transportation networks should prioritize developing a large, well-connected central area before expanding outward in order to maximize their resistance to failure disruptions.

public transportation | robustness | core-periphery structure | subway networks

Public transportation systems make life in dense urban areas possible by shuttling millions between destinations reliably and safely. During the three-month period of 2022 from July to September, public transportation riders took 1,627,982 distinct trips in the U.S. (3) Despite a drastic fall-off of ridership during the COVID-19 pandemic, more riders are returning to public transportation. Furthermore, public transportation in cities is more efficient than automobiles and is increasingly viewed as an important step towards carbon neutrality.

Public transportation systems are a common area of study in network science, as they naturally lend themselves to representation in graph format. Commuter rail systems (or subways, or metros) tend to have a well-defined structure that is smaller, and thus more manageable for analysis, than other public transportation networks such as buses, ferries, or ride-sharing. For these reasons, our analysis will focus on analyzing metro systems.

In existing network science studies of public transportation networks, some common attributes of metro networks have been detected. Namely, most metro networks are "scale-free" because the degree distribution of their stations follows a Power Law distribution. (2) Furthermore, metro networks exhibit high clustering (and increasing clustering with size) and small average geodesic length, meaning that they are small-world networks as defined in Watts and Strogatz. (4) For public transportation networks, geodesic length is substituted for longest trip length. (2)

Specifically, the concept of public transportation network robustness is a highly relevant measure for transportation experts to consider. Robustness can be thought of as the resilience of a public transit system to a failure of a line or station, which in a network representation is the removal of an edge or node. This is a very practical problem facing many municipalities globally, as it is important that public transit systems continue to serve users in the event of a line breakdown or station closure. Previous studies of transportation network robustness involve measuring the availability of alternative backup paths that can be used in the event of a failure. In this spirit, Derrible and Kennedy determine robustness by measuring the amount of cycles in a network, scaled to the size of the network. (2) The intuition behind this measure is that the presence of a cycle ensures the existence of multiple paths to travel between the same two nodes.

In this paper, our objective is to explore "path score" as a measure of transport-based core-periphery structure, as introduced by Lee, Cucuringu, and Porter. (1) The traditional hub and spoke structure (2) of metro networks is similar to a core-periphery structure, making path score a useful measure to investigate in this context. The path score for a node i measures the amount of backup paths between all pairs of adjacent nodes that contain i. By measuring the availability of backup paths, we intend to use this measure to evaluate the robustness of 14 metro systems around the globe. We can then compare our robustness analysis with that of Derrible

Significance Statement

This paper expands discussion of public transportation network robustness: specifically, the ability of a subway network to continue to function in the case of station or line failures. We observe that many subway networks have a "hub and spoke" structure with an interconnected core and tree-like spokes. Thus, we apply the concept of core-periphery structure to measuring robustness by applying path score, a novel betweenness measure proposed by Lee et al. (1) and compare this with Derrible and Kennedy's cycle-based robustness metric. (2) We assess this measure empirically, comparing the results of both metrics on 14 metro networks in major cities around the world. Our results contribute to existing public transportation network literature by providing an application of path score to metro networks and comparing robustness measures for this application.

Author contributions: A.S. helped implement path score code and helped create visualizations. G.C. formatted metro network data and implemented path score code. J.L. helped create visualizations, edited the paper draft and synthesized the final publication. S.K. helped create visualizations and wrote the paper draft.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: jmlim@ucla.edu

and Kennedy, evaluating the consistency of our findings from various lenses.

This analysis will contribute to transportation robustness literature in multiple ways. First, Lee, Cucuringu, and Porter previously applied their "path score" metric to road networks, which have different qualities than metro networks. For example, metro networks are much sparser than most road networks, and often have a hub and spoke structure rather than a grid-like structure, which could give interesting results. Additionally, our paper will provide a more complete picture of robustness on a transportation network. While Derrible and Kennedy presented a global robustness metric, our analysis looks at individual nodes in addition to the network as a whole, allowing us to provide validation and new insights on alternative paths as measures of transportation network robustness.

Other robustness measures of public transportation networks in the networks analysis literature are based on the concept of "attacks," that is, how a network's properties change when nodes/edges are removed either randomly or systematically. For example, Berche et al. applies this methodology to city transit networks and finds that the most "effective" node attacks targeted the highest degree nodes or nodes with the highest betweenness centrality measures. (5) In this context, an "effective" attack is one that quickly negates the effectiveness of the network at performing its intended purpose, which the authors represented as the size of the giant connected component (GCC) in the network, S. The size of the GCC does not translate smoothly into our study of robustness as a path-related property. Without cycles or backup paths, removing an edge creates two components, decreasing the size of the GCC but not necessarily the overall number of paths in the network. However, the notion that high degree or high betweenness nodes are most effective to remove gives these nodes a sense of "importance" to the robustness of the network.

Comparing Robustness Measures

We will use the following mathematical representation of a public transportation network: $G(V, E)$, where V is the set of metro stations represented as nodes and E is the set of train lines connecting stations, represented as edges. We assume G is undirected (common in similar studies) and connected. These assumptions are based on the real-world observations that 1) most trains between stations run both ways, and 2) a station is only considered part of the public transportation network if it is connected to other stations.

Derrible and Kennedy r^T. Derrible and Kennedy (2) propose a measure of network robustness specifically tailored to the characteristics of a public transportation network. For the purpose of their calculations, Derrible and Kennedy divide the nodes into two sets, terminus stations with one connection (V_m for monotonic) and those with more than one connection (V_d for diatonic). Thus, $V = \{V_m, V_d\}$. They divide the edges similarly: $E = \{M, D\}$. Their approach to measuring robustness is based on the cyclomatic number μ defined by Berge et al. (6), which calculates the number of cycles in a graph by subtracting the number of edges in the transportation network as a tree, $|V_d| - 1$, from the total number of edges $|D|$. The authors don't consider terminus nodes and edges because these components can't be involved in cycles. Thus,

$$\mu = |D| - |V_d| + 1. \quad [1]$$

The final robustness measure subtracts multi-edges (not present in our data) and considers the higher possibility of node/edge failure with a larger network.

$$r^T = \frac{\mu - |D^m|}{|V|} \quad [2]$$

where D^m is the set of multi-edges. Thus, the Derrible and Kennedy robustness metric r^T gives a count of the number of cycles relative to the size of the network. Although larger networks tend to have more cycles, this metric seeks to measure robustness based on the structure of the public transportation system rather than solely its size. (2) Based on the author's analysis of 33 metro systems around the world, the robustness of smaller systems is determined mostly by μ, so introducing cycles is critical to robustness for small train networks. However, for larger systems, both terms play an equal role in determining r^T. (2)

Lee et. al Path Score. On a more local scale, one measure of robustness is "path score" (PS), which is a measure of centrality for each node. (1)

Given our edgelist

$$\mathbb{E} = \{(j, k) | \text{an edge exists between nodes } j \text{ and } k\},$$

and $\{p_{jk}\}$ as the set of shortest paths from node j to k when the edge (j, k) is removed, we can define path score as follows:

$$\mathrm{PS}(i) = \frac{1}{|\mathbb{E}|} \sum_{(j,k)\in\mathbb{E}} \sum_{\{p_{jk}\}} \sigma_{ijk}[\mathbb{E} \setminus (j, k)]. \quad [3]$$

We have that $\sigma_{ijk}[\mathbb{E} \setminus (j, k)] = \frac{1}{|\{p_{jk}\}|}$ if $i \in \{p_{jk}\}$, and $\sigma_{ijk}[\mathbb{E} \setminus (j, k)] = 0$, otherwise. Therefore, for node i, the path score is based on the number of "optimal backup paths" that i is present on. In this sense, path score removes edges in order to determine a node's importance, an approach shared by many of the other "attack"-metrics. (5, 7)

The "optimal backup path" was defined as the shortest path based on network distance, or the fewest number of edges in the path. In the context of public transportation, the path score for a station would be the number of backup paths that the station was a part of in the full network, and the optimal backup path would be the route with the fewest number of stops. In a more thorough study that perhaps uses a weighted network, this optimal backup path could use the "cost" of the path, which could consider the distance or time of the route, but for simplicity, fewest number of stops works well to explain the networks.

The path score measure was derived from the structural model of a core-periphery structure. This model divides a network into a core and a periphery, in which nodes in the core are densely connected to one another and also peripheral nodes, but nodes in the periphery are only sparsely connected to one another. Path score and the idea of backup paths allow us to investigate these core-peripheral edges and connections further.

Many nodes in the network will have a path score of zero, as the network is primarily comprised of line graphs. Several peripheral nodes will not be involved in any backup paths, or oftentimes backup paths do not exist between two nodes.

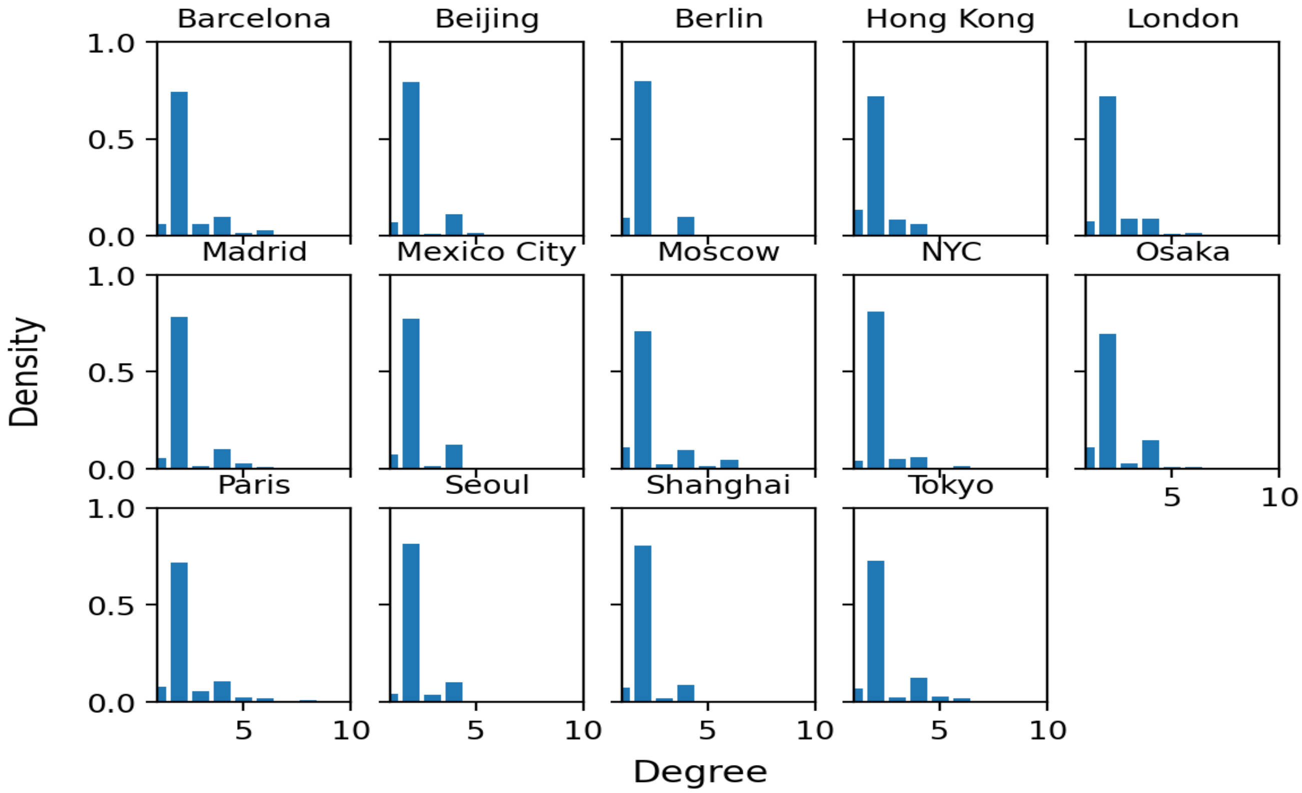

Fig. 1. Degree distributions for each metro system, displaying the proportion of nodes with each degree. Most stations connect to 2 other stations.

However, looking at the proportion of nonzero path scored nodes is telling about the robustness of the network, as the proportion of nodes with a nonzero path score indicates the quality of a transportation network's structure and how well it can operate with station outages. Additionally, looking at the nodes with the highest path scores indicates "hubs" within the network, as these stations are often at major intersections and connect several of the line graphs within the network. Visualizing the distribution of path scores on a network can be useful to identify robust and non-robust areas.

One challenge in using path score is that it is a local measure, so using it to characterize a full network is not entirely natural. One measure that is useful is the mean path score, which is the average of each node's individual path score across the full network. The mean path score then becomes a global metric and can be used across networks to compare robustness or to compare with other global metrics, such as Derrible and Kennedy's r^T, in order to gain a wider view of robustness. However, the mean path score may be biased by a skewed distribution of path scores and outliers. The theoretical rigor of taking an average of path scores has not been studied and would need further examination before concrete conclusions can be made based on our results.

Another drawback of path score is that it fails to consider the individual importance of a node or edge, and instead considers its relation to paths to and from other nodes. Other studies of network robustness consider the concept of "bridge connections," which are nodes or edges that are the sole connection between two communities. (8) Thus, a node is important for robustness for connecting these communities. However, path score would miss this importance, since it only considers "optimal" backup paths that would only include the bridge connection for nodes in the two connected communities. Thus, intra-community paths are not counted, underrepresenting node importance. Thus, a better robustness measure would be able to consider both "types" of importance.

Results

Degree Distribution. The degree distributions of the networks in Figure 1 are telling of each network's structure. The degrees seem to be more heavily distributed on even-numbered degrees, with the mode being 2 for all distributions. The network structure for the subways is largely composed of line graphs connected at "hubs." Since all nodes on a line graph besides the two ends of the line have a degree of 2, this explains the high frequency of the degree 2. It also explains why there is a higher frequency of even-numbered degrees, since unless a node is the end point on a line graph, or the terminus on a transportation route, then it will connect a multiple of 2 nodes per route. Any "hubs" can be defined as nodes that connect more than one transportation route, or serve as an intersection of multiple line graphs. These nodes will have a degree higher than 2. This directly relates to the ease of transport within a city: in cities with more evenly distributed degree distributions, it will be easier to change transport lines and reach more nodes. A city with a degree distribution with almost all nodes at a degree of 2, such as Berlin, may be difficult to get around or change direction.

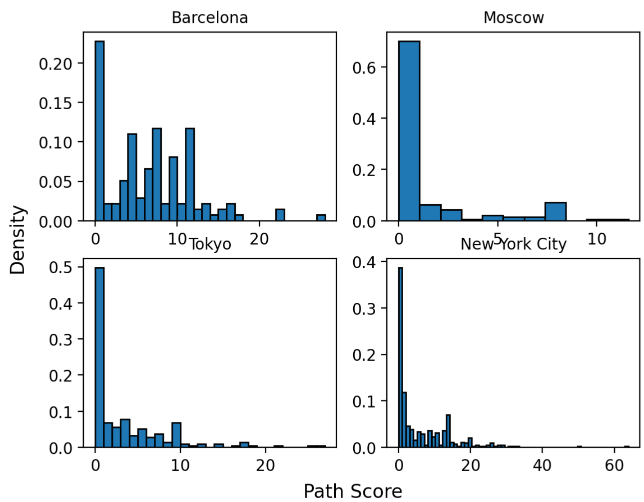

Fig. 2. Path score distributions for 4 noteworthy cities. Each graph displays the proportion of nodes with a given path score.

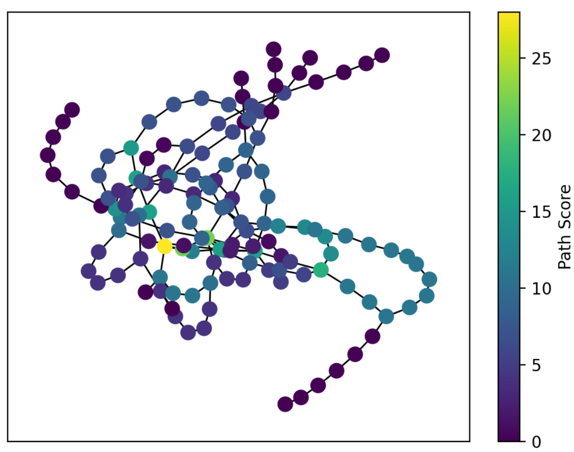

Fig. 3. Spring representation of Barcelona metro system. Nodes are colored according to path score value. Barcelona has a large, interconnected core, leading to a high mean path score.

Path Score Distribution. For all 14 metro systems studied, we find that the largest proportion of nodes have a path score of 0, and this proportion ranges from 20.6%, as in Barcelona, to 70.9%, as in Moscow (see Figure 2). We can compare the structure of the transportation networks in Barcelona and Moscow in conjunction with their path score with the heatmap representations in Figure 3 and Figure 4. We can see that Moscow has many lines with 0 path score extending out from a dense non-zero path score center. These outer lines have no access to transfer stations. This results in these peripheral nodes having a path score of 0, indicating that they are not present on any backup paths. Failures along these lines will leave riders without any alternative paths to their destination. Conversely, Barcelona has a much larger, sparser center and fewer, shorter peripheral lines than Moscow. Within the non-zero path score center, there will be alternative paths available in the event of a failure, allowing riders to continue to their destination. Therefore, based on path score distribution and network structure, we can conclude that Barcelona appears to

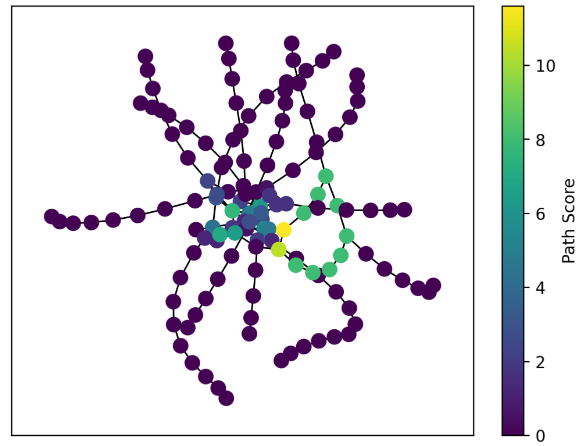

Fig. 4. Spring representation of Moscow metro system. Nodes are colored according to path score value. Moscow has a small core and large periphery, leading to a low mean path score.

be more robust than Moscow. More broadly, transportation networks with a smaller proportion of nodes with a path score of 0, or a smaller periphery size, will likely be more resilient to failure.

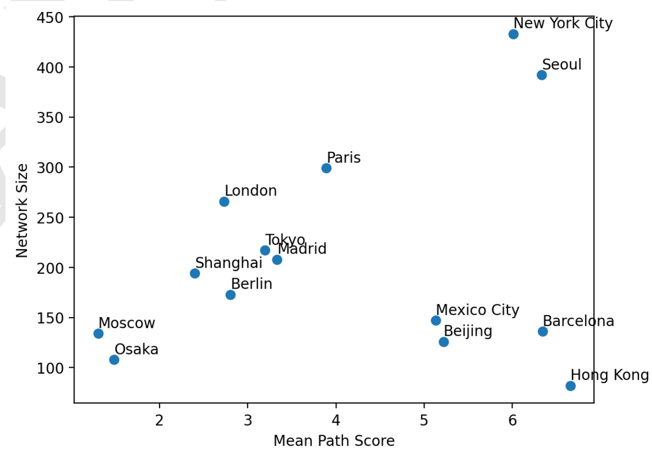

Fig. 5. Comparing mean path score to the number of nodes in a network, there is a general positive correlation with interesting exceptions.

Network Size vs. Path Score. To further investigate path score as a means of measuring transportation network robustness, we explored the relationship between network size and the mean path score over the network. In this context, network size is the total number of stations in a metro network. From Figure 5, there appears to be a positive correlation between mean path score and network size, especially for mean path scores less than 4. This is an interesting, but not necessarily surprising result, as larger networks have more stations and lines, and thus, more backup paths available.

However, there is a cluster of 4 networks, Mexico City, Beijing, Barcelona, and Hong Kong, which have a high mean path score, but a relatively small network size. Looking at these networks' structures in comparison with the similarly sized Moscow and Osaka, which have low mean path scores, we can see that Mexico City, Beijing, Barcelona, and Hong Kong have larger, well-connected central cores with higher path scores, allowing for higher robustness despite their relatively smaller size. A visualization of the Hong Kong system is provided in Figure 6.

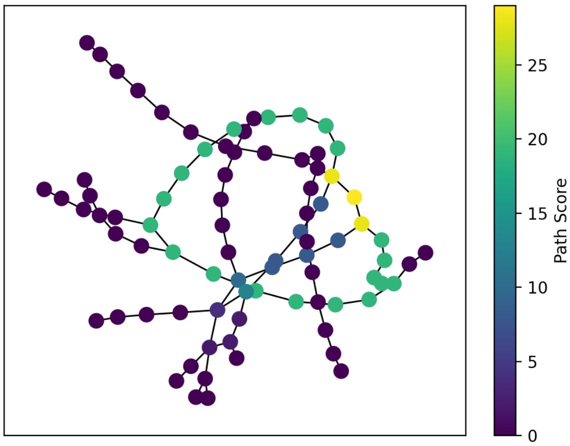

Fig. 6. Hong Kong's metro system has a large core and relatively short peripheral lines. A circle line contains many nodes with high path score. Hong Kong has a high mean path score, especially for its size, but low r^T.

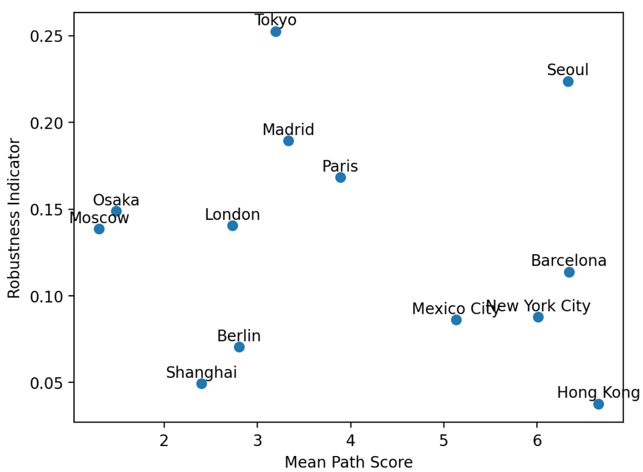

Fig. 7. Comparing mean path score to Derrible and Kennedy's metric r^T, there is little correlation. Hong Kong (Fig. 6) is noteworthy for it's high path score but low r^T.

r^T vs. Path Score. We compared our robustness measure of path score with the Derrible and Kennedy metric (2), in order to see if our results were consistent. Similar to our analysis of mean path score and network size, below a mean path score of 4, there is a positive correlation between Derrible and Kennedy's robustness metric and path score, providing evidence for consistency between these two measures. However, above a mean path score of 5, our results seem to diverge from Derrible and Kennedy's, with networks receiving a large mean path score receiving low robustness scores from Derrible and Kennedy. For example, Hong Kong (see Figure 6) has one of the highest mean path scores despite also having the lowest robustness measure value. The reason can be seen in the overall topology of Hong Kong, which contains one large, primary loop which branches out to several transfers. This decreases the robustness metric, which is based on the volume of cycles, while maintaining high path scores for the large proportion of stations in the primary loop. The exception to this is Seoul, which received both a high robustness score from Derrible and Kennedy and a high mean path score.

Fig. 8. Comparing mean path score to fraction of nodes with $k > 2$ (transfer stations) reveals less correlation than the same comparison with r^T.

Transfer Stations vs. Path Score. Furthering this comparison, Derrible and Kennedy found a strong positive correlation between fraction of transfer stations and robustness. (2)

We defined transfer stations as nodes with a degree greater than 2. This was motivated by the line structure of a public transportation network, leading terminal nodes to have degree 1, and interior nodes on the line having degree 2. Any node with a degree greater than 2 will thus connect multiple lines, indicating a transfer station. We then calculated the fraction of nodes of degree greater than 2 in order to determine the fraction of stations in the network that are transfer stations.

As seen in Figure 8, Derrible and Kennedy's positive correlation is not consistent with our results using mean path score. Revisiting Moscow and Barcelona, both cities have similar fractions of transfer stations, but Barcelona has a significantly higher mean path score. However, due to their different structures, this suggests effects on path score and robustness may also be dependent on where the transfers are located, rather than solely the fraction. Barcelona has a much larger core area in which these transfers occur (see Figure 3), allowing backup paths to exist over a larger network area. Conversely, most of Moscow's transfers occur in a condensed central area (see Figure 4). These structural differences do not impact the total number of cycles in a graph, making this an additional insight gleaned from path score that could explain the discrepancies with Derrible and Kennedy's robustness metric.

Conclusion

In conclusion, the core-periphery structure of public transportation networks is important to consider when investigating network robustness. By using the measure of path score for a node, we studied the structure of 14 public transportation networks. We determined that having a large proportion of

City	$\langle$PS$\rangle$	r^T	Prop. Nodes PS = 0	Prop. Transfer
Barcelona	6.34	0.11	0.21	0.20
Beijing	5.22	Missing	0.41	0.13
Berlin	2.80	0.07	0.61	0.11
Hong Kong	6.66	0.04	0.57	0.15
London	2.73	0.14	0.46	0.21
Madrid	3.33	0.19	0.38	0.16
Mexico City	5.13	0.09	0.39	0.15
Moscow	1.30	0.14	0.71	0.18
NYC	6.01	0.09	0.35	0.15
Osaka	1.48	0.15	0.43	0.19
Paris	3.89	0.17	0.38	0.21
Seoul	6.33	0.22	0.38	0.15
Shanghai	2.40	0.05	0.61	0.12
Tokyo	3.20	0.25	0.43	0.20

Table 1. A summary of results for each city

nodes with a path score of 0 indicates a large periphery without alternative paths between nodes, leading to a lower mean path score, indicating lower network robustness. Additionally, we found that larger networks tend to be more robust due to the larger availability of backup paths, but a robust network need not be large. We found that smaller networks with larger, although sparser core structures still had high mean path scores, indicating their structural characteristics impact their robustness quality. Path score seems to better represent these structural differences than Derrible and Kennedy's cycle-based robustness score, providing a more complete picture of robustness when looked at together.

Therefore, it seems that transportation networks should prioritize developments to expand the well-connected center of the network, such as creating more transfer stations and overlapping lines that can provide more alternative paths. Expansions of this nature, instead of extending developments exclusively outward, will better protect the network against disruptions.

This paper has many limitations, which can be addressed in future work. We studied a small sample of only 14 public transportation networks, so performing a similar analysis on a larger collection of metro networks could further validate or dispute our findings. Within our existing data, we could compare the metro networks over time to see what additions/modifications to systems improved path score. We acknowledge that current robustness work considers a highly simplified version of a subway network. Future work could tie in ridership data, population density, and failure rate data to create a more accurate representation of subway networks. Furthermore, the mathematical properties of existing robustness metrics have not been thoroughly studied. Understanding such properties for mean path score would give us a better view of how this metric assesses robustness. Additionally, this paper only examines one measure of robustness, path score, and compares it with one other proposed metric. In order to develop a more complete analysis of robustness, future work should examine more network metrics as measures of robustness, providing more insight into different structural characteristics that are important to a transit network's resistance to failure.

Materials and Methods

We calculated path score values in MATLAB. We created our visualizations using `matplotlib` in Python.

The code we used for this paper can be found in the public GitHub repository: https://github.com/jlimcode/pathscore-robustness.

Data. The data used in this paper is sourced from Barthelemy et al. (9) The original authors of the data constructed the topologies of the subway networks from 2009 and 2010 network maps, the most current available at the time of the article. In order to preserve the accuracy of the network data and use the most reliable source possible, this paper studies subway information from the most recent years available. This differs by city, and a known limitation for the application of our results is that our subway network maps are at least ten years old.

ACKNOWLEDGMENTS. We would like to thank Marc Barthelemy for providing us with the data used in his 2012 paper "A long-time limit for world subway networks." We would also like to thank Mason A. Porter for his guidance in our research process and for providing us with MATLAB code to calculate path score.

1. SH Lee, M Cucuringu, MA Porter, Density-based and transport-based core-periphery structures in networks. *Phys. Rev. E* **89** (2014).
2. S Derrible, C Kennedy, The complexity and robustness of metro networks. *Phys. A: Stat. Mech. its Appl.* **389**, 3678–3691 (2010).
3. Public transportation ridership report (2022).
4. DJ Watts, SH Strogatz, Collective dynamics of 'small-world'networks. *nature* **393**, 440–442 (1998).
5. B Berche, CV Ferber, T Holovatch, Y Holovatch, TRANSPORTATION NETWORK STABILITY: A CASE STUDY OF CITY TRANSIT. *Adv. Complex Syst.* **15**, 1250063 (2012).
6. C Berge, *Theory of graphs.* (Methuen young books, London, England), (1962).
7. S Wandelt, X Shi, X Sun, Estimation and improvement of transportation network robustness by exploiting communities. *Reliab. Eng. & Syst. Saf.* **206**, 107307 (2021).
8. A Masoumzadeh, T Pechlivanoglou, Network-based analysis of public transportation systems in north american cities (2020).
9. C Roth, SM Kang, M Batty, M Barthelemy, A long-time limit for world subway networks. *J. The Royal Soc. Interface* **9**, 2540–2550 (2012).