From Network Topology to Traffic: An investigation of the Relationship Between Network Topological Features and Traffic Data

Saman Behrouzi*[1], Panchamy Krishnakumari[2], Irene Martinez Josemaria[3], Mario Romero[4], Serge P. Hoogendoorn[5], and Hans van Lint[6]

[1]PhD student, Transport and Planning Department, Delft University of Technology, Netherlands
[2]Assistant Professor, Transport and Planning Department, Delft University of Technology, Netherlands
[3]Assistant Professor, Transport and Planning Department, Delft University of Technology, Netherlands
[4]Associate Professor, Department of Computational Science and Technology, KTH Royal Institute of Technology Stockholm, Sweden
[5]Full Professor, Transport and Planning Department, Delft University of Technology, Netherlands
[6]Full Professor, Transport and Planning Department, Delft University of Technology, Netherlands

Short summary

Traffic dynamics is a complex phenomenon. Visualizing and uncovering its hidden complexities using simple topological features with clear physical interpretations allows us to improve traffic predictions and offer valuable insights to decision-makers. This study delves into the relationship between these topological features and actual traffic data, with an emphasis on the highway network in the Netherlands. Using both traffic data and complex network analysis techniques, we explore how the betweenness centrality (BC) corresponds to patterns in traffic flow and speed. We apply Pearson correlation analysis to quantify these relationships, especially during peak traffic hours. Interestingly, while the results show that the correlation between BC and traffic flow and speed is not strong during the day, a more intricate relationship emerges during peak times. We showed that BC demonstrates a notable negative correlation with traffic speed, a finding that is statistically significant (p-value ≤ 0.05). These insights pave the way for a deeper understanding of how network topology affects traffic behavior.

Keywords: Centrality Measures, Complex Network Analysis, Traffic Congestion, Transportation Networks

1 Introduction

Traffic congestion worldwide leads to longer travel times, increased pollution, and lower living standards. Understanding its contributing factors enables transportation planners to design more efficient highway networks, improving traffic flow. Network analysis and visualization techniques have become effective tools for examining the dynamics and structure of urban transport and road networks. The aim is to comprehend their spatial structure and dynamics better, with the ultimate goal of enhancing transportation services' efficiency and fairness, as well as improving the performance of road traffic. Researchers have explored the applications of complex networks analysis in understanding mobility networks through the lens of road traffic and public transport networks.

For instance, some researchers put their focus on investigating road networks. In a study, Xu et al. developed an evaluation mechanism for Sydney's urban road network, using a binary directed weighted network to assess node degrees, traffic flow, and network efficiency Xu et al. (2022). Saberi et al. proposed a contagion model to understand traffic congestion patterns in urban areas, similar to infectious disease spread models Saberi et al. (2020). Sheikh and Regan introduced a method combining time series analysis and independent component analysis for traffic incident detection, proving effective in both simulated and real-world scenarios Sheikh & Regan (2022). Curado et al. analyzed private vehicle mobility in Rome and London, using centrality measures in a multidimensional network to identify key urban zones for mobility and tourism Curado et al. (2021).

Also, in public transport area of studies, Luo et al. found network properties to be good estimators of passenger flow in public transport systems, offering a simpler alternative to traditional models Luo et al. (2020). Wei et al. studied the impact of administrative boundaries on bus network integration in Nanjing, highlighting the need for strategic planning for spatial fairness and efficiency in public transport Wei et al. (2021). Dai et al. used complex network theory to explore the relationship between bus line structures and passenger flows in Beijing, finding similar spatial patterns and peak congestion times Dai et al. (2022). Meng et al. examined the complexity of Shenzhen's metro networks during rush hours, suggesting targeted passenger flow control for enhanced safety and resilience Meng et al. (2023).

The motivation behind this study lies in the lack of comprehensive research on motorway networks and the actionable insights it can offer to transportation planners, policymakers, and engineers responsible for optimizing traffic flow and improving the performance of the highway network. By establishing correlations between traffic data and topological features, we can gain a deeper understanding of how network properties influence traffic patterns to identify bottlenecks, optimize road capacity, and enhance network resilience. Furthermore, such insights can assist in the design and development of future transportation networks, taking into account the relationship between network topology and traffic dynamics. Moreover, there is still a need for comprehensive studies focusing specifically on motorway traffic data, providing an opportunity to uncover valuable insights and contribute to the existing body of knowledge in transportation network analysis.

In this study, we focus on the Netherlands highway network which is renowned for its advanced transportation system and has witnessed tremendous growth in vehicular traffic over the years. With increasing population, urbanization, and economic activities, it becomes imperative to comprehensively analyze the factors impacting traffic conditions and explore the role of network topology in shaping traffic dynamics. The topological characteristics of the Netherlands highway network and actual traffic statistics are compared to see how they relate to one another. For each node (intersection) in the network, we specifically look at how the Betweenness centrality (BC) of that node relates to traffic flow and speed. In order to find patterns and correlations in the traffic data, we employ complex network analysis with a preprocessed dataset from the highway network of the Netherlands.

The paper's structure includes: Section 2 describing our evaluation methods, Section 3 presenting our findings and analysis of BC, traffic flow, and speed, and Section 4 concluding with future research recommendations.

2 Methodology

Data preparation

For this study, we employed traffic data sourced from the Nationaal Dataportaal Wegverkeer (NDW) in the Netherlands, which utilizes loop detectors scattered throughout the Dutch highway network. Our study focused specifically on one week's worth of data, gathered between 6-12 June 2022 and for each day we have the data from 5:00 to 23:00.

The geographic locations of each intersection on the roads, obtained from NDW open data, were used to construct the highway network. In this network, junctions are represented as nodes, while the roads connecting them are referred to as links. To ensure better topological feature extraction, first, we extracted the largest connected component of the network. This process involved removing 750 isolated nodes from the network, and keeping the giant component of the network with 6775 nodes. Then, to focus on significant intersections, we further coarsened the network by neglecting nodes with a degree of 2 and kept the nodes with degree 3 or higher, which resulted in a final network with 1235 nodes and 1684 edges. These nodes represent big intersections, main crossroads, and facilitate the calculation of more meaningful topological features. Figure 1 shows the network before and after coarsening.

The loop detector data from NDW measures both speed and flow. In the Netherlands, there are approximately 10,000 loop detectors placed in the highways which collect data every 1 minute and are placed approximately 500 meters apart. In our study, we use a well-known estimation technique Adaptive Smoothing Method to obtain aggregated data with a uniform temporal and spatial resolutions of 5 minutes and 500 meters, respectively. The aggregation is done based on traffic flow theory and the detailed filtering method can be found in Schreiter et al. (2010).

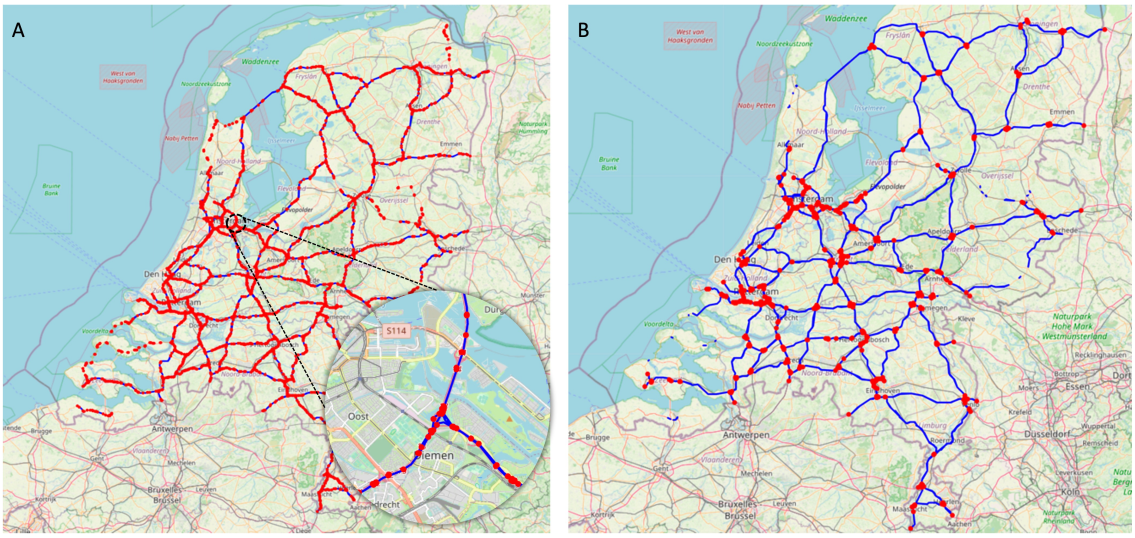

Figure 1: The highway network of the Netherlands. The blue lines represent the parts of the network with loop detectors to gather the traffic data and the red dots represent all intersections. Figure A shows nodes of the network before coarsening and Figure B represents them after coarsening.

Calculating network topological features and correlation analysis

We examined four different centrality measures for this study: degree centrality, eigenvector centrality, closeness centrality and betweenness centrality. Among all of them, the BC showed more discernible correlation patterns with the traffic data. Thus, we only present BC to keep the results concise in this paper. BC captures the importance and influence of nodes within the network. This feature quantifies the amount of control a node exerts over the interactions of other nodes in the network. Nodes with high BC can often influence the flow of traffic in the entire network, making it a global measure of centrality Brandes (2001). Nodes with high BC act as key traffic channels. Delays around nodes with higher BC can disrupt the entire network, making them priorities for traffic control. The following equation shows how the BC is being calculated:

$$c_B(V) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest paths, and $\sigma(s,t|v)$ is the number of those paths passing through some node v other than s,t. If $s = t, \sigma(s,t) = 1$, and if $v \in s,t, \sigma(s,t \mid v) = 0$.

In our study, we employed Pearson correlation analysis to quantify the degree of association between traffic data and various network topological features Meghanathan (2015). Specifically, we focused on traffic flow during peak hours, which were identified as 06:00 to 10:00 and 15:30 to 19:00. These peak hours were chosen as they represent the busiest periods of the day when traffic congestion is at its highest and understanding the dynamics is most crucial for traffic management.
To assess the strength and direction of the linear correlation between traffic flow and BC, we calculated the Pearson correlation coefficient. We are interested to see if there is a negative or positive correlation. Also, we computed the p-value to determine the statistical significance of the observed relationships. In our study, we considered a p-value of less than 0.05 to be statistically significant, which is a commonly accepted threshold in the statistical analysis.
To calculate the Pearson correlation, first, we computed the average inbound and outbound flow for each node (intersection) within the distance of two kilometers for each 5 minutes for each day. Then, for each time step, we calculated the Pearson correlation between the BC of all nodes in the network and their average inbound flow and outbound flow. We also did the same process for the inbound speed and outbound speed.

Figure 2: The selected subnetwork for correlation analysis highlighted with the shaded polygon. The examined intersections for the correlation analysis have been highlighted with red dots.

3 Results and discussion

Our analysis reveals some interesting findings about the correlation between BC and the mean traffic flow and speed throughout an entire week including work days and weekends. We performed the correlation analysis on the whole highway network of the Netherlands, and separately did the calculation for work days and weekends. We then performed the same correlation analysis for a region in South Holland province, which has the highest population among other provinces in the Netherlands. It is important to note that for this sub-network we used the same BC scores that have been calculated based on the whole network to see how the identified nodes with higher importance contribute to traffic congestion in their local area. Figure 2 shows the selected sub-network for the correlation analysis.

Figure 3 illustrates the correlation between inbound and outbound traffic speed and the BC for the whole highway network of the Netherlands and the selected sub-network in South Holland province. Figure 4 shows the same correlations for the traffic flow data. By looking at the highway network as a whole, we did not observe a significant correlation between the BC and the traffic speed and flow (the left two columns in the mentioned figures). This raised the assumption that we might have nodes with high BC all over the highway network, and most of them may not experience heavy congestion. To confirm this assumption, we picked a region around two big cities in South Holland and calculated the correlations again. As shown in the right two columns of Figure 3 and Figure 4, we observe significantly stronger correlations between traffic speed, flow, and the BC in the selected sub-network. Notably, within this sub-network, the correlation of BC with traffic speed is more pronounced than with traffic flow, particularly during peak hours. However, the p-value indicates that the observed fluctuations in our figures could be random, this could be attributed to the analysis of the aggregated data for multiple days in each case. For this, we calculated the average of correlations and p-values per time step across all days.
Therefore, we picked the selected region to look deeper into the correlation patterns during the week. Figure 5 shows the correlation between the inbound and outbound traffic speed and the BC for the selected sub-network and their p-value. As depicted in the figure, we observed a strong negative correlation in peak hours during workdays, except on Monday. However, this strong correlation was missing during the morning peak hours on Wednesday and Friday. To assess the statistical significance of these correlations, we employed p-values. Our null hypothesis stated that any perceived correlation between traffic speed or flow and BC is random. Figure 5 shows that the p-value reject this hypothesis and confirms that the correlations during peak hours are significant.

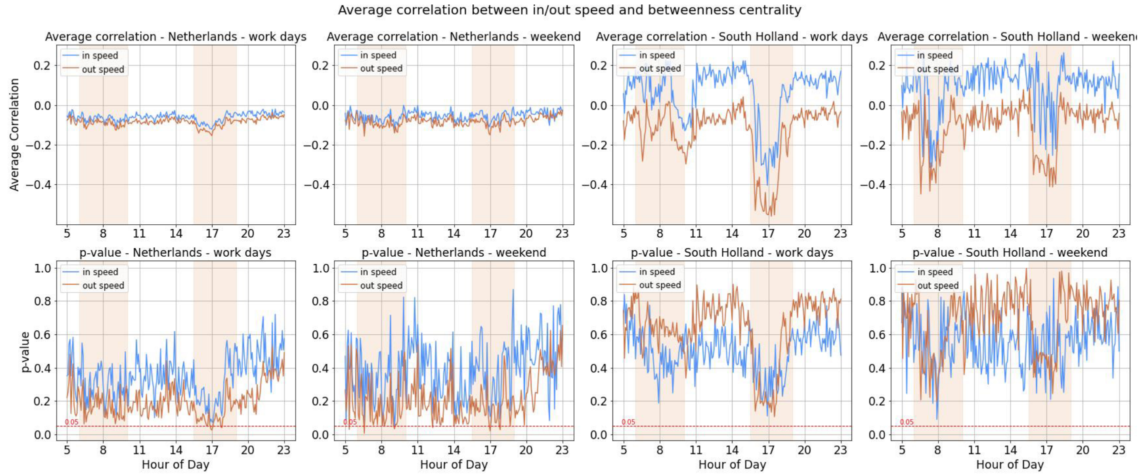

Figure 3: Average correlation between inbound and outbound traffic speed and BC for the whole highway network of the Netherlands and a selected region in the South Holland Province. The shadowed time frames represent the peak periods.

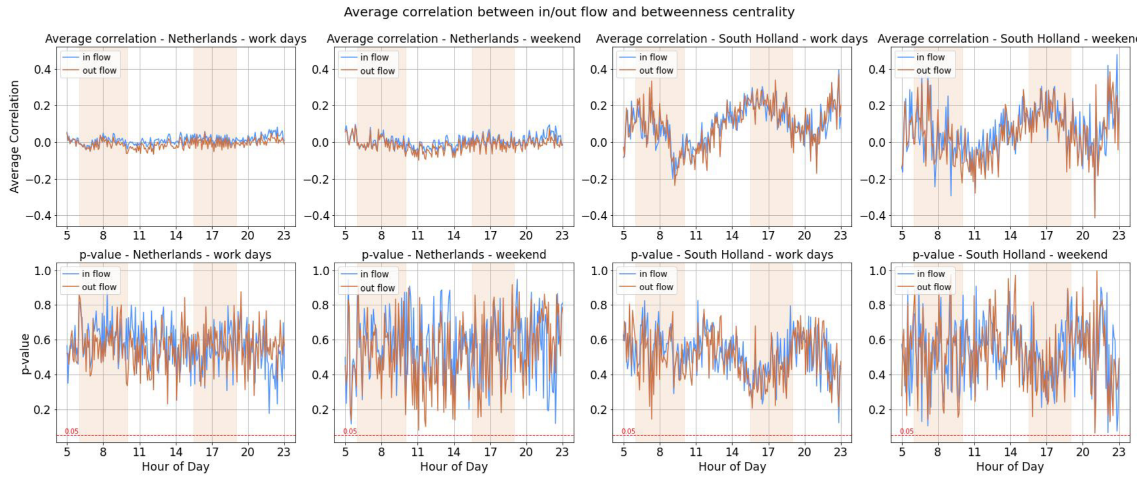

Figure 4: Average correlation between inbound and outbound traffic flow and BC for the whole highway network of the Netherlands and a selected region in the South Holland Province

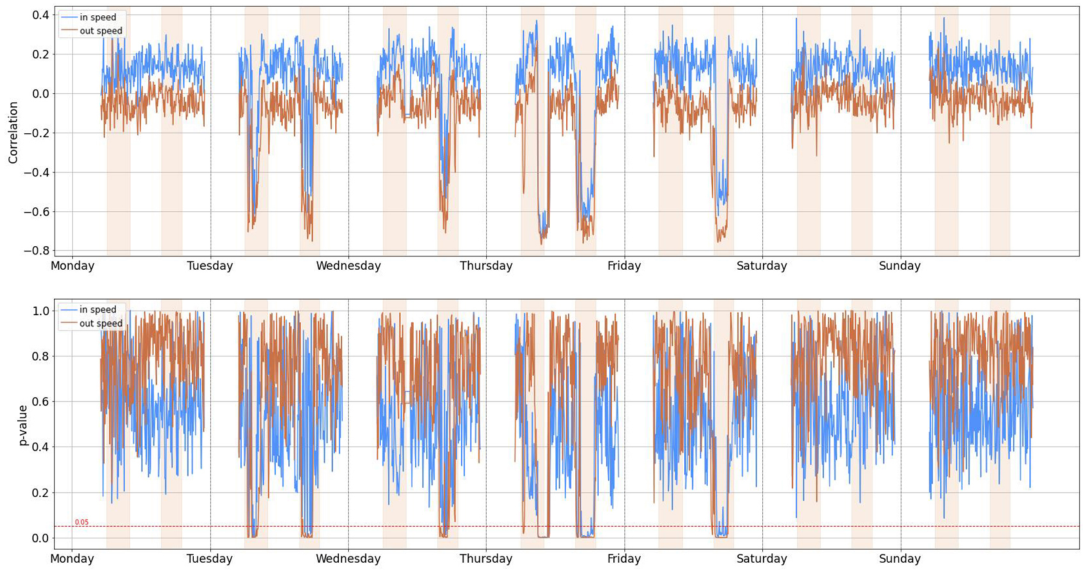

Figure 5: The correlation between the inbound and outbound speed and the BC for the selected sub-network. The data has been shown for each 5 min for June 6th and June 12th, 2022.

The presented correlation analysis shows the influence of topological features, in particular BC, on traffic flow and speed and provides valuable insights on the highway network structure, which can be crucial for traffic planners and policy managers to factor into their future analysis. When considering aggregated data across days, the correlations were not significant which emphasizes the complex nature of traffic flow dynamics, which are likely influenced by other factors beyond network topology, e.g., demand patterns.

4 Conclusions

This study is a part of a bigger work on traffic data analysis in which we delved into the intricate relationship between the topological characteristics of the Netherlands' highway network (e.g. betweenness centrality) and actual traffic conditions. Employing complex network analysis techniques, the study assessed how BC correlates with traffic speed and flow. The results indicate a generally weak linear correlation between BC and traffic parameters for the whole highway network of the Netherlands. However, when we zoomed in and looked at the South Holland province we could easily see the strong negative correlation between the BC and traffic speed during peak hours. These findings reveal that despite the complex interplay of factors influencing traffic dynamics, the network topology alone may be a decisive predictor. The insights gained provide valuable context for transportation planners and policymakers, emphasizing the need for multi-faceted strategies in traffic management and network design.
Further research could explore how other variables, such as road capacity, interact with network topology and how these variables can be exploited to make faster and more accurate traffic prediction models. Also, traffic engineers can study how targeted traffic management strategies like variable speed limits in areas identified as critical nodes (high BC) interact with the network topology to affect traffic flow.

References

Brandes, U. (2001). A faster algorithm for betweenness centrality*. *The Journal of Mathematical Sociology*, *25*(2), 163-177. doi: 10.1080/0022250X.2001.9990249

Curado, M., Tortosa, L., & Vicent, J. F. (2021, 10). Identifying mobility patterns by means of centrality algorithms in multiplex networks. *Applied Mathematics and Computation*, *406*, 126269. doi: 10.1016/J.AMC.2021.126269

Dai, T., Ding, T., Liu, Q., & Liu, B. (2022, 11). Node centrality comparison between bus line and passenger flow networks in beijing. *Sustainability 2022, Vol. 14, Page 15454*, *14*, 15454. doi: 10.3390/SU142215454

Luo, D., Cats, O., & van Lint, H. (2020, 12). Can passenger flow distribution be estimated solely based on network properties in public transport systems? *Transportation*, *47*, 2757-2776. doi: 10.1007/S11116-019-09990-W/FIGURES/8

Meghanathan, N. (2015). Correlation coefficient analysis of centrality metrics for complex network graphs. In R. Silhavy, R. Senkerik, Z. K. Oplatkova, Z. Prokopova, & P. Silhavy (Eds.), *Intelligent systems in cybernetics and automation theory* (pp. 11–20). Cham: Springer International Publishing.

Meng, Y., Zhao, X., Liu, J., Qi, Q., & Zhou, W. (2023, 1). Data-driven complexity analysis of weighted shenzhen metro network based on urban massive mobility in the rush hours. *Physica A: Statistical Mechanics and its Applications*, *610*, 128403. doi: 10.1016/J.PHYSA.2022.128403

Saberi, M., Hamedmoghadam, H., Ashfaq, M., Hosseini, S. A., Gu, Z., Shafiei, S., ... González, M. C. (2020, 4). A simple contagion process describes spreading of traffic jams in urban networks. *Nature Communications 2020 11:1*, *11*, 1-9. doi: 10.1038/s41467-020-15353-2

Schreiter, T., Lint, H. V., Treiber, M., & Hoogendoorn, S. (2010). Two fast implementations of the adaptive smoothing method used in highway traffic state estimation. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 1202-1208. doi: 10.1109/ITSC.2010.5625139

Sheikh, M. S., & Regan, A. (2022, 1). A complex network analysis approach for estimation and detection of traffic incidents based on independent component analysis. *Physica A: Statistical Mechanics and its Applications*, *586*, 126504. doi: 10.1016/J.PHYSA.2021.126504

Wei, S., Zheng, W., & Wang, L. (2021, 4). Understanding the configuration of bus networks in urban china from the perspective of network types and administrative division effect. *Transport Policy*, *104*, 1-17. doi: 10.1016/J.TRANPOL.2021.02.002

Xu, D., Wang, Y., Peng, P., Lin, L., & Liu, Y. (2022). The evaluation of the urban road network based on the complex network. *IEEE Intelligent Transportation Systems Magazine*, *14*(3), 200-211. doi: 10.1109/MITS.2021.3049351