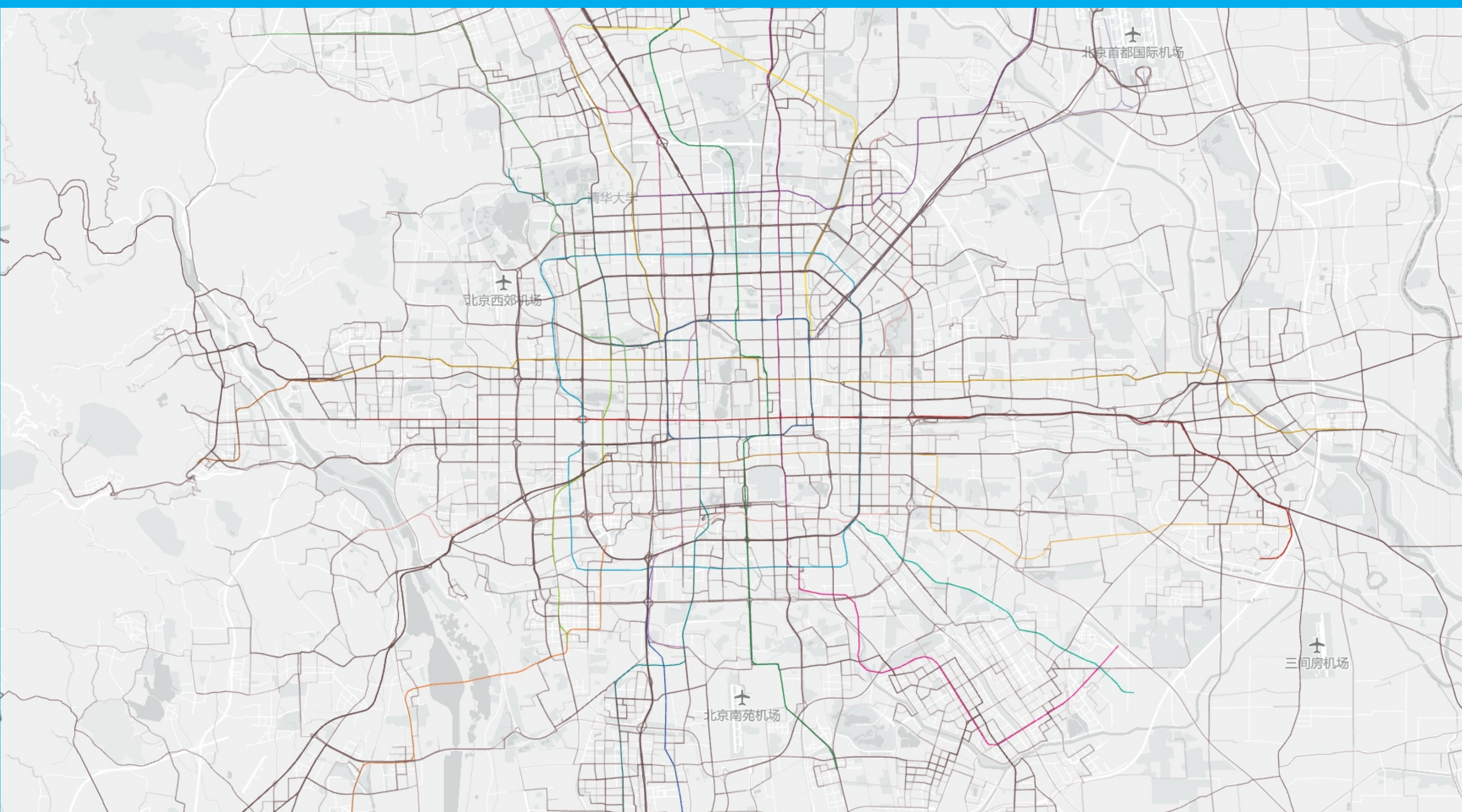
Quantification and Comparison of Hierarchy in Unimodal Public Transport Networks

Ketong Huang

Department of Transport & Planning
Faculty of Civil Engineering & Geoscience
Delft University of Technology



TU Delft

Quantification and Comparison of Hierarchy in Unimodal Public Transport Networks

by

Ketong Huang

to obtain the degree of Master of Science
at the Delft University of Technology,

Student number:	5266688	
Date:	May 8, 2023	
Thesis committee:	Prof.dr. O. Cats,	TU Delft, chair
	Ir. Z. Wang,	TU Delft, daily supervisor
	Dr. R. Massobrio,	TU Delft, secondary supervisor
	Dr. A. Bombelli,	TU Delft, external supervisor

TU Delft Delft University of Technology

Preface

The whole process of the graduation thesis is like a ten-month voyage, fulfilled with adventures and harvest. I will cherish this experience and regard it as a treasure of my life. Time rewinds to December 2019, when I was applying for the master's course at TU Delft. I still remember what I wrote in the motivation letter. I said that I was impressed by the graduation thesis reports published on the information pages of the Traffic & Planning track. The thesis report's novel ideas, rigorous argumentation and beautiful graphics drove my desire to study at TU Delft. I also wanted to produce such excellent work. I was not a straight-A student in my bachelor, and my grades barely passed the minimum requirements for application, so I didn't expect the offer from TU Delft. I don't know how this motivation letter affected my application results. Still, from the very first day that I received the offer, I knew that I got the opportunity to realise my goal of finishing a good thesis work. After these ten months of hard work, I believe I have made it.

Throughout the journey, my supervisors guided me to the finish line. First, I want to thank my committee chair, Prof.dr. O. (Oded) Cats. Thank you for your wisdom. You always provide constructive opinions and insightful feedback, guiding the direction of the development of the entire topic, like a captain steering the course. Secondly, I'd like to thank my daily supervisor, Ir. Z. (Ziyulong) Wang. Thank you for your guidance and collaboration, from the inspiration of the project to its completion. In the topic, you are a dedicated guide who always puts in your best effort, patiently answers my questions and discusses all the details of the study with me. In life, we are good friends who care about each other and share our joys and sorrows, encouraging me and inviting me to LIT Coffee during my moments of depression. Thirdly, my thanks go out to my secondary supervisor, Dr. R. (Renzo) Massobrio. You have provided a lot of help in data processing and case studies. In reviewing my reports, you are always careful to give comments. When I am stuck with problems, you will appear with your Mate tea to support me. To my external supervisor from Aerospace Engineering (AE) faculty, Dr. A. (Alessandro) Bombelli. You always provide me with suggestions on the methods of network topology. You are always patient when I visit your office at the AE faculty to seek advice. It is a pleasure to communicate with you. My sincerest gratitude to my supervisors for all the help you have given me!

Finally, I would like to thank everyone who has helped me during this voyage. Special thanks to my parents for your unwavering support. No matter the distance, your love is always warm. I love you. Big thanks to all my friends. We are neighbours, soccer teammates, marathon-running friends, etc. The time spent with you all is precious. Thanks for accompanying me along the way: Li Xu, Zelin Xu, Yiting Li, Jiaxuan Zhang, Heqi Wang, Zhexin Lyu, Tianyi Deng, Xin Tan, Qingxin Liu and all the DCF teammates. Thanks, everybody!

Ketong Huang
Delft, April 2023

Executive Summary

Hierarchy is a network property that identifies the organisation and importance of network elements. Public transport systems consist of stops organised by connections, which can be regarded as networks. In a Public Transport Network (PTN) with a high hierarchy, the number of elements gradually decreases as their importance increases, where the majority of elements have low importance and a few high-important elements. The hierarchical organisations in PTNs contribute to the network performance by efficiently allocating resources based on the elements' importance. In the field of transport, the hierarchy has been studied with different methods and data sources. However, the literature related to quantification and comparison methodologies of PTN hierarchy and its mode-wise and continent-wise effects is scarce, particularly for high-capacity PTNs that serve the majority of public transport demands. A unified PTN hierarchy definition and the quantification methodology which is necessary to enable PTN comparisons in terms of the network organisations, and reflect the relative PTN performance.

Background and Research Objectives

The PTN is the backbone of urban public transport. The performance of PTN greatly influences the performance of urban transport networks. The high performance of a PTN is achieved when the public transport demands are assigned with a capable organisation of elements in the PTN. The high-performance organisation of elements can be reflected by the PTN hierarchy. Hierarchy is a network property indicating the organisation of network elements, where the number of elements gradually descends as their importance increases. The high-importance elements are few, and the elements with low importance are the majority. By quantifying the PTN hierarchy, the PTN performance can be assessed from the perspective of network element organisations. Based on the unimodal PTN hierarchy comparison between different modes and continents located, the mode-wise and continent-wise effects on PTN hierarchy can be revealed.

In the previous work, there are some main PTN hierarchy identification methods. With the PTN element attribute-based method, the levels of network elements' importance are coarse and limited, making the PTN hierarchy hard to quantify and compare. The passenger flow data-oriented method is more accurate than the PTN element attribute-based method but only studies the observed empirical patterns. Compared to these two methods, the network topology-based method is closer to the essence of the PTN hierarchy for quantification and comparison. The PTN elements are classified into importance levels based on element topological attributes. Notwithstanding, there are still some research gaps that need to be filled. First, although the previous research proved that the hierarchy exists as a property of a PTN, a unified network topology-based PTN hierarchy definition is missing, which is the base of the hierarchy comparison. Second, there is a shortage of identification of the PTN hierarchy's topological characteristics from multiple scales. With the identified multi-scale characteristics, the PTN hierarchy can be more comprehensively assessed and compared. Third, lack of PTN hierarchy quantification and comparison methodologies that endeavours to utilize network topology. Fourth, research about the effects of PTN attributes on the hierarchy is lacking, such as the modes and continents. To fill the research gap, the formulated research questions are presented below, beginning with the main question:

How to quantify and compare unimodal Public Transport Networks hierarchy?

The main research question is divided into four sub-questions in this study. First, referring to the hierarchy-related literature in different fields, a unified definition of the PTN hierarchy is proposed, which focuses on the PTN elements organisations. Second, by exploring the existing literature, the multi-scale topological characteristics of the PTN hierarchy are identified, both in network and element scales. The third step develops the network topology-based PTN hierarchy quantification and comparison methodology. Based on the topological characteristics, the hierarchy of high-capacity unimodal PTNs worldwide is quantified and compared along the dimensions of the corresponding topological indicators. Fourth, according to the hierarchy of these high-capacity unimodal PTNs, the mode-wise and continent-wise comparisons are conducted, and the effects of modes and continents on the PTN hierarchy are investigated.

Research Methodology

In general, the methodology of this study has four steps. The workflow of the methodology is shown in Figure 1. First, clearly define the PTN hierarchy, then identify and select its topological characteristics based on the literature review. The second step selects quantified indicators for the topological characteristics and develops the quantification methodology for the PTN hierarchy. Third, the methodology is applied to a GTFS data-based case study. A database consisting of 63 high-capacity unimodal PTNs worldwide is investigated, obtaining the six-dimension PTN hierarchy. The final step carries out PTN hierarchy comparisons with visualised radar charts, analysing the PTN hierarchy in each dimension and the mode-wise and continent-wise effects.

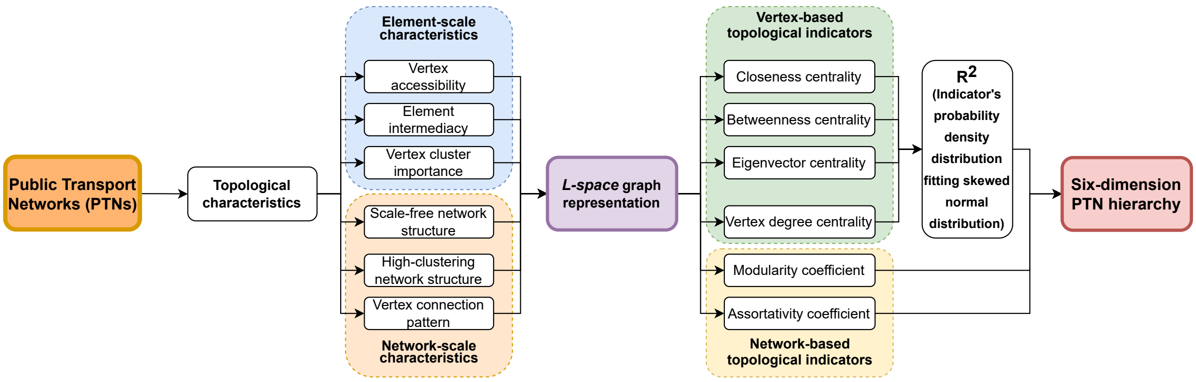


Figure 1: Workflow of the research methodology

First, originating from the PTN hierarchy defined as a network property, six topological characteristics of the PTN hierarchy on element and network scales are identified and selected by summarising the literature reviewed. In the element scale, vertex accessibility, element intermediacy and vertex cluster importance are selected to represent the PTN elements' importance. In a high-hierarchy PTN, the elements follow hierarchical organisations by their importance. In the network scale, scale-free network structures, high-clustering network structures and vertex connection patterns by vertex degrees are selected to reflect the high-hierarchy networks' structures.

Second, six vertex-based and network-based topological indicators are selected to quantify the six topological characteristics. The selected vertex-based indicators include vertex degree, closeness centrality, betweenness centrality and eigenvector centrality. The selected network-based indicators are the optimal modularity coefficient and the assortativity coefficient. Among the vertex-based topological indicators, the vertex degree is used for quantifying a network-scale characteristic, the scale-free network structures. In vertex-based quantification, each

PTN is first assumed to have the probability density distribution of elements' importance following a skewed normal distribution. The goodness of fit (R square) of vertex-based indicator probability density distribution fitting skewed normal distribution represents the quantified PTN hierarchy. The network-based coefficients can directly quantify the PTN hierarchy.

Next, apply the PTN hierarchy quantification methodology to a GTFS data-based case study. The case study database consists of 63 high-capacity unimodal PTNs worldwide. The GTFS data are collected from online mobility data platforms and public service providers. After data processing, the GTFS data are filtered and converted to L-space-represented PTN graphs for topological quantification. With the hierarchy quantified, each PTN has a six-dimension hierarchy based on indicator values.

Finally, based on the 63 high-capacity unimodal PTNs' hierarchy, the variation of the hierarchy in each dimension is evaluated. Supported by the six-dimension radar charts, the normalised PTN hierarchy comparison is visualised and intuitive, reflecting the relative network performance. Plus, the mode-wise and continent-wise effects are comprehensively discussed with influencing factors' typical values based on the median hierarchy of PTNs with the same modes or located in the same continent. Three high-capacity public transport modes are included in the mode-wise comparison, the metro, tram and BRT. The continent-wise comparison considers the six inhabited continents, Africa, Asia, Europe, Oceania, North America and South America.

Results

First of all, the worldwide high-capacity unimodal PTNs' hierarchy varies in the six dimensions. Figure 2 presents the normalised PTN hierarchy in six dimensions. In vertex degree and network assortativity dimension, the PTNs in the database show a tendency to have a relatively low hierarchy. It appears that the scale-free network structures and vertex connection patterns by degrees are not significant amid the investigated PTNs. By contrast, the R square value distributions in the closeness centrality and betweenness centrality dimensions indicate tendencies for PTN in the database to have a relatively high hierarchy. The organisations of PTN elements by the vertex accessibility and element intermediacy tend to be hierarchical. The distributions of PTN hierarchy in the eigenvector centrality and network modularity dimensions both present a slightly left-skewed distribution, indicating a similarity in the numbers of PTNs exhibiting relatively low and high hierarchy. The hierarchical organisation of vertex clusters by their importance and the high-clustering structures are relatively moderate for PTNs in the database.

Second, in the PTN hierarchy comparison with radar charts, the implications of hierarchy in six dimensions are different and exhibit priority distinction. Figure 3 offers an example of a normalised PTN hierarchy comparison with a radar chart. According to the PTN hierarchy in the six dimensions, evaluations in the closeness centrality and betweenness centrality dimensions hold a high priority across the six dimensions. The hierarchy in these two dimensions evaluates the organisation of PTN elements from the perspectives of stop accessing and traffic flow intermediacy, which could serve as important references for passengers and public service providers during PTN operations. The hierarchy in the eigenvector centrality and network modularity dimensions assesses the mono-centric and multi-centric structures, respectively. The hierarchy in the vertex degree and network assortativity dimensions is less apparent than in the former four dimensions, because of the insignificant scale-free structures and vertex connection patterns by degrees in PTNs of the database.

Next, the mode-wise and continent-wise comparison of the median high-capacity unimodal PTN hierarchy in the database revealed the effects of public transport modes and located continents. Figure 4(a) presents the radar chart of mode-wise median PTN hierarchy comparison.

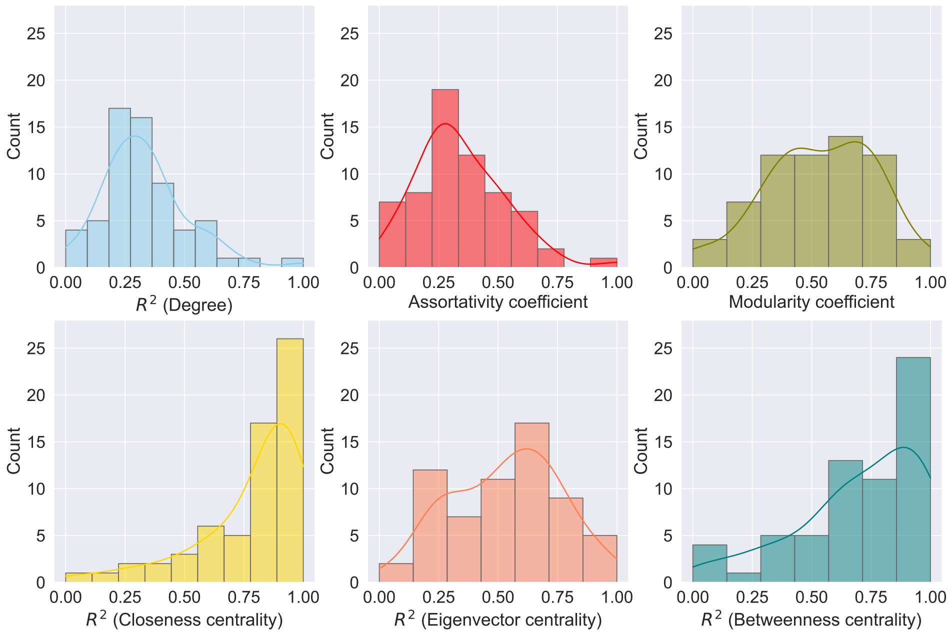


Figure 2: Histograms of normalised topological indicators in the six dimensions

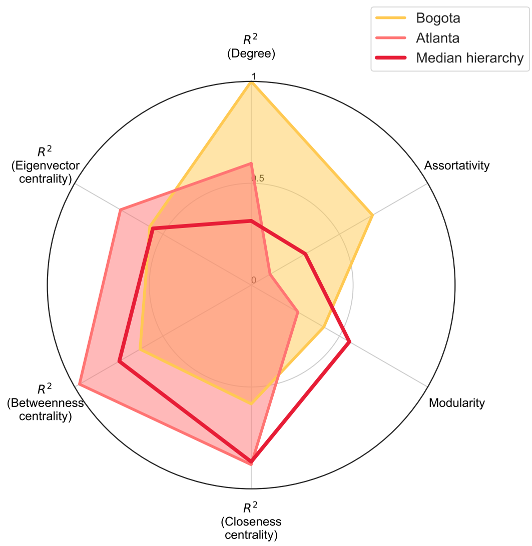
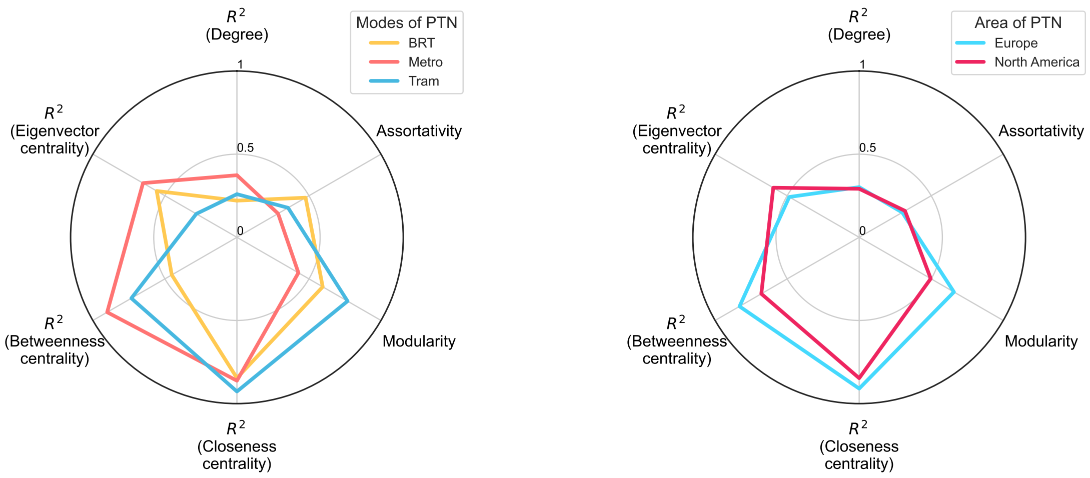


Figure 3: PTN hierarchy comparison example with six-dimension radar chart

The results show that the order of modes having PTN hierarchy from high to low is metro, tram, and BRT. None of the modes has the highest PTN hierarchy in all dimensions. The PTN hierarchy is mainly influenced by stop spacing and line spacing, the stop infrastructure, the operating speed and the stopping patterns. These influencing factors affect the number of vertices in the PTNs, the number of connections between vertices and the weights of edges. Affecting the topology of PTNs by the number of vertices and edges, the modes bring effects to the PTN hierarchy. The European PTNs tend to have a higher hierarchy than North American PTNs. Neither of the two continents has a higher PTN hierarchy in all dimensions. The continent-wise effects mainly result from the different modal compositions, especially the proportions of tram and BRT networks in the two continents.



(a) Mode-wise radar chart (b) Continent-wise radar chart

Figure 4: Mode-wise and continent-wise median PTN hierarchy comparison with six-dimension radar charts

Conclusions and Recommendations

This thesis presented an assessment of the hierarchy in high-capacity unimodal PTNs and how modes of transport and geographical location influence this hierarchy. It fills the research gap where hitherto there is a rare unified topological-based definition of PTN hierarchy and a lack of quantification and comparison methodology. With regard to the organisation of network elements, the PTN hierarchy is defined as a network property. A six-dimension topology-based quantification methodology is developed on both network and element scales. A database that relies on GTFS data and contains the network topology information of 63 high-capacity unimodal PTNs worldwide is used as a case study. The PTN hierarchy is interpreted and analysed along the dimensions of topological indicators. Furthermore, the effects of modes and continents are discussed by comparing the PTN hierarchy.

This research develops a network topology-based methodology for PTN hierarchy comparison and quantification. First, a unified definition of the PTN hierarchy is established. In this definition, the PTN hierarchy is determined by the organisation of the elements in PTNs. Then, it has been found that the PTN hierarchy is embodied in multiple aspects and different scales based on the literature review. Six topological characteristics of the PTN hierarchy are identified and selected in element and network scales. To reflect the organisation of elements based on their multi-aspect importance, the R square values between element-scale indicators' probability density distributions and skewed normal distributions are regarded as the PTN hierarchy. Based on the normalised PTN hierarchy in the six dimensions, the PTN hierarchy is visually compared through radar charts. To apply the methodology to a case study for high-capacity unimodal PTNs worldwide, a database that depends on GTFS data and contains L-space topology information is constructed with a data pipeline. Relying on this data pipeline, 63 high-capacity unimodal PTNs worldwide are processed and included in the database. With the quantified hierarchy of the high-capacity unimodal PTNs in the database, the analysis and discussion of the PTN hierarchy are conducted. It has been found that the PTN hierarchy in dimensions exhibits a priority distinction. For example, PTN hierarchy in the closeness centrality and betweenness centrality holds priority and could serve as important references for passengers and service providers during operation. In addition, metro networks present a tendency to have a high hierarchy than the tram and BRT networks. The mode-wise effects result from factors influencing PTN topologies, such as stop spacing and line spacing. For different modal compositions, the European PTNs have a higher hierarchy compared to North American PTNs.

For future research, an integrated and quantified multi-dimension PTN hierarchy metric is recommended. Based on the found PTN hierarchy priority of dimensions, future work can work on quantifying the weights of dimensions and integrating the six-dimension PTN hierarchy with one single quantified metric. Based on the single quantified PTN hierarchy metric, the assessment of the PTN hierarchy can be direct, simplified and accurate. Besides, the PTN hierarchy can be further studied with the relationship between the vulnerability of PTNs. Both PTN hierarchy and PTN vulnerability focus on the element organisations in PTNs, especially the elements with high importance. In past research, some topological indicators in this study are usually associated with PTN vulnerability analysis, such as closeness centrality and betweenness centrality. Benefiting from it, the planning PTNs can be evaluated and optimised by balancing the PTN hierarchy and vulnerability. Moreover, if data source and volume allow, the relation between the mode-wise and continent-wise effects on PTN hierarchy and their influencing factors can be quantified and analysed.

Contents

1 Introduction

The public transport network (PTN) is the backbone of urban public transport. The PTN accounts for the general and high-capacity public transport services, greatly influencing the urban transport network. Particularly, high-capacity PTNs undertake the majority of people's mobility with the public transport modes having high-capacity and exclusive right of way, such as metro, tram and bus rapid transit (BRT) (Raicu et al., 2009). Therefore, the performance of the high-capacity PTN largely affects the performance of a PTN, such as the total travel times, travel cost and level of service. The high performance of a high-capacity PTN is achieved when the public transport demands are assigned with a capable organisation of elements in the PTN.

To investigate the element organisations of high-capacity PTNs, the hierarchy can be an evaluation indicator. Hierarchy is a network property indicating the organisation of network elements, where the number of elements gradually descends as their importance increases. The high-importance elements are few, and the elements with low importance are the majority. Hierarchy is studied across multiple disciplines, such as neuroscience (Zhao et al., 2015), social networks (Rowe et al., 2007), and in particular PTNs (Buijtenweg et al., 2021; Pumain, 2006; Tsiotas and Polyzos, 2015; Van Nes, 2002a; Wang et al., 2020). Hierarchy in PTNs indicates the organisation of elements in the PTN, such as stops, route sections, or stop clusters. Therefore, by understanding the PTN hierarchy, the PTN elements with high importance can be identified. Thus, the public transport service can be designed to cope with the demand efficiently, improving the performance of PTN. Various modes in a high-capacity PTN lead to different effects on the hierarchy due to their unit capacity, stop spacing, and operating speed (Aston et al., 2021; Pomykala, 2018). Quantifying and comparing the hierarchy of different high-capacity unimodal PTNs can help to understand the effects of public transport modes on the hierarchy, which is valuable for optimising the existing PTN and evaluating future PTN planning. Previous works show three main methods for PTN hierarchy identification: the PTN element attribute-based method, the passenger flow data-oriented method, and the network topology-based method.

In the PTN element attribute-based method, the PTN element importance levels are based on one or several PTN attributes, for example, the modes of PTN or the location of network elements, such as urban, suburban and rural. Van Nes (2002b) discussed an optimisation method for PTN, classifying the network into urban and interurban networks, referring to the locations of travel origins and destinations. Similarly, Gao et al. (2012) classified the PTN links into three types: mass route, feeder route and local route, based on the distance between stations and the transfer routes. Yap et al. (2018) classified PTN into three levels, where each level utilises one or several transport mode networks. For instance, the train network represents an inter-region level, the metro and tram networks represent an agglomeration level, and the bus network stands for an urban level. With the PTN element attribute-based method, the levels of network elements' importance are coarse and limited, making the PTN hierarchy hard to quantify and compare.

The second main method to identify hierarchies in PTNs is the passenger flow data-oriented method. Passenger flow data is recorded by the Automatic Fare Collection (AFC) system when passengers travel in the PTN, indicating passenger flow volume on PTN elements, such as stops or route sections. The volume of passenger flow is used for classifying the PTN element importance levels. The higher the passenger flow undertaken, the higher the importance levels of the elements. Bassolas et al. (2019) studied the urban mobility hierarchy and proposed a flow-based hierarchy metric, reflecting the magnitude of trips between trip flow-based travel hot spots and the hot spot hierarchy. Yap et al. (2019) used smart card data to identify and cluster the significant transfer hubs with passenger flow in the PTN. The PTN hierarchy identified by this method is more accurate than the PTN element attribute-based method, focusing on the attributes of PTN elements, such as transfer volumes of stops. However, the passenger flow data-oriented method only studies the observed empirical patterns, like the distribution of passenger flow, which is a consequence of the passenger demand and the network. Therefore, the essential reasons for the PTN hierarchy behind the passenger flow are not explored. In addition, the passenger flow data of PTNs are large-volume and not easily accessible, making it difficult to conduct a comparison between a large number of PTNs.

The network topology-based method is the third main method of PTN hierarchy identification. The network topology theory describes the organisation of network elements. Using topological network indicators, the PTN elements are classified into importance levels based on network attributes, for example, the stops' number of connections, or the minimum number of links between two stops. Yerra and Levinson (2005) found that the emergence of transport network hierarchy is a self-organised process, and the properties of vertices and edges reflect the hierarchy. Investigating the relation between passenger flow data with the intrinsic attributes of networks, Luo et al. (2020) proved that the topological properties of PTN have a linear correlation with the passenger flow distribution, which implies the possibility of correlating topological attributes with the PTN hierarchy. Some research fuses network topological analysis with the passenger flow data for PTN hierarchy identification. For example, Buijtenweg et al. (2021) proposed an integrated hierarchy metric to quantify the importance of stations in the PTN by three topological vertex attributes and passenger flow data. Wang et al. (2020) used passenger transfer flow as the weights of PTN edges. Fusing with network topological coefficient modularity, a four-level time-dependent PTN hierarchy is identified. Compared to the other two methods, the network topology-based method also focuses on the attributes of PTN elements, which can identify and quantify the PTN hierarchy from a more fundamental perspective.

Reviewing the previous research on PTN hierarchy identification methods, the network topology-based method is the closest to the essence of PTN hierarchy for quantification and comparison. However, there are still some research gaps that need to be filled. First, although the previous research proved that the PTN hierarchy exists as a property of a PTN, the representation and interpretation of the PTN hierarchy are different, and a unified and clearly defined PTN hierarchy is missing. Second, there is a shortage of identification of the PTN hierarchy's topological characteristics from multiple scales. With the identified multi-scale characteristics, the PTN hierarchy can be more comprehensively assessed and compared. Third, lack of methodology of PTN hierarchy quantification and comparison methodologies that endeavours to utilize network topology. Fourth, research about the effects of other attributes of PTN on PTN hierarchy is lacking, such as the continents of PTNs.

In response to the above research gaps, this research proposes a network topology-based method for quantifying and comparing the high-capacity unimodal PTN hierarchy. First, a clear definition of the PTN hierarchy is given, focusing on the PTN elements and reflecting the organisation of these PTN elements. Second, a network topology-based interpretation of the PTN hierarchy is offered, in which the multi-scale topological characteristics of the PTN hierarchy are identified, both in network and element scales. Third, the network topology-based PTN hierarchy quantification and comparison methodology is developed. Based on the topological characteristics of the PTN hierarchy, the hierarchy of high-capacity unimodal PTNs worldwide is quantified and compared along the dimensions of the corresponding topological indicators. Fourth, according to the hierarchy of these high-capacity unimodal PTNs, the mode-wise and continent-wise comparisons are conducted, and the effects of modes and continents on the PTN hierarchy are investigated.

To fill the research gap, the main research question is stated as follows:

How to quantify and compare unimodal Public Transport Networks hierarchy?

The main research question is divided into four sub-questions:

1. What are the definition and the topological characteristics of the PTN hierarchy?
2. What indicators can be used to quantify the network topology characteristics of PTN hierarchy?
3. Based on the selected indicators, how do the high-capacity unimodal PTNs worldwide quantify and compare in terms of hierarchy?
4. Based on high-capacity unimodal PTNs' hierarchy, what are the mode-wise and continent-wise effects on PTN hierarchy?

The rest of the thesis is structured as follows: Chapter 2 reviews the related works and discusses the definition and topological characteristics of the PTN hierarchy. In Chapter 3, based on the topological characteristics of the PTN hierarchy, the PTN hierarchy quantification and comparison methodologies are developed. Chapter 4 applies the network topology-based PTN hierarchy quantification methodology to 63 high-capacity unimodal PTNs worldwide. Based on the results, the PTN hierarchy along each dimension of topological indicators and the mode-wise and continent-wise effects on PTN hierarchy are discussed. Finally, Chapter 5 provides the conclusions and recommendations for future research.

2

Literature Review

This chapter clarifies the research scope and the two main topics of the literature review in Section 2.1. In Section 2.2, the PTN hierarchy definition-related research is reviewed, and two main definitions of the hierarchy are discussed. The related literature about the PTN hierarchy's topological characteristics is reviewed from the network and element scales in Section 2.3, which summarizes the corresponding topological indicators for analysis.

2.1. Scoping

There are two topics of the literature review from top to bottom: the definition of PTN hierarchy and the topological characteristics of PTN hierarchy. The following literature review is organised following these two topics:

- **Hierarchy definition-related research**: The topic focuses on different definitions of "hierarchy" in the previous research. Since the definition of hierarchy is not unified, hierarchy-related research from different fields is reviewed to fill this research gap. This review builds a better understanding of the network hierarchy for further analysis.
- **Topological characteristic-related research**: The topological characteristics of the PTN hierarchy are embodied on multiple scales. For example, the network scale or element scale, such as vertices, edges or vertex clusters. Topological indicators can quantify each topological characteristic. Different topological indicators are reviewed and discussed based on the PTN hierarchy's topological characteristics. The review helps to select appropriate topological indicators for building the PTN hierarchy quantification methodology.

2.2. Hierarchy definition-related research

Multiple definitions for hierarchy can be found in different research fields including, such as sociology (Garandeau et al., 2014; Pattiselanno et al., 2015; Pumain, 2006; Rowe et al., 2007), neuroscience (Meunier et al., 2009; Zhao et al., 2015), aviation transport (Fernandes et al., 2019). Based on the definitions of hierarchy, past works have two main streams: "hierarchy is an order of items", and "hierarchy is a property of a network". The related work of each stream is reviewed below.

2.2.1. Hierarchy is an order of items

The first definition of hierarchy is an order of a series of same-type items based on one or several quantifiable indicators. Thus, the hierarchy indicates the series' high or low order of items. For example, in social network research, Garandeau et al. (2014) discussed the social hierarchy in groups of students by surveying the indices of popularity and social impact of students in small groups. Rowe et al. (2007) identified the social hierarchy based on the number of exchanged emails between persons. In city hierarchy-related research, Godfrey and Zhou (1999) investigated the global city hierarchy, based on the number of headquarters and first-level subsidiaries of the world's top 500 companies in cities. Fang et al. (2017) investigated the hierarchy of cities based on the population for city urbanisation evaluation. Research with this hierarchy definition is also common in transport, especially aviation-related hierarchy research. For instance, Grubesic et al. (2008) identified the worldwide airport hierarchy based on the passenger flow data, such as the airport flight and passenger throughput, and the airports' maximum exchange volumes. Fernandes et al. (2019) ranked the Brazilian airports annually based on an original indicator "networkability". This indicator depends on airports' connectivity and mutual passenger flow. By comparing the change in the airport hierarchy over ten years, the development of the regional economy is discussed. Chen et al. (2022b) identified the airport hierarchy based on passenger throughput and the accessibility of airports to high-speed railway facilities. Research with this hierarchy definition on other transport facilities is also common. Tsigdinos et al. (2022) studied the hierarchy of urban roads based on traffic flow and the road infrastructure, for example, the flow of the city's central areas and the coverage of cycling or bus facilities. Acton et al. (2022) identified the hierarchy of BRT networks based on the total line length and the independence of the BRT networks' right-of-way. In a word, in the research defining hierarchy as an order, a common point is shared: the hierarchy identification is based on the ordering by one or several quantifiable indicators.

2.2.2. Hierarchy is a property of a network

The second definition considers hierarchy as a property of a network. A hierarchical network has an organisation of network elements, in which the minority of elements have high importance, while the majority of elements have low importance. Compared to the first definition, the scale of hierarchy changes from element to network. When hierarchy is a network property, the hierarchy is not only the order of elements based on their importance, but also a more macroscopic network-scale organisation shown by these elements. When hierarchy is an order of items, it cannot reflect the network-scale organisation of elements. Yerra and Levinson (2005) discussed the self-organised emergence of hierarchy in the transport network by assigning unbalanced traffic demands to a theoretical grid network. The differences in road use create the hierarchy of the transport network, performing an organisation of edges in highly differentiated levels. Lai and McDysan (2002) claimed that the hierarchy is the abstraction of the network's topology mechanism. The network hierarchy reflects a universal law followed by vertex connections in the network. Mones et al. (2012) discussed the commonalities of various complex networks and offered a complex network hierarchy quantification methodology. The quantification of hierarchy is based on an original network-scale indicator, the global reaching centrality (GRC), which denotes the heterogeneity of the reaching centrality of vertices. The higher GRC of a network, the higher heterogeneity between vertices' reaching centrality, i.e., a higher network hierarchy. Loginova et al. (2022) studied the national economy network hierarchy. In the economy networks, cities act as nodes, and the connections between multination firms' headquarters and branch offices are edges. Regarding the nations as subplots in the network, the k-core method is used to evaluate the national city importance heterogeneity. Furthermore, the Gini coefficient of k-core values is used to reflect the national economy network hierarchy.

The network hierarchy can be affected by multiple aspects, for example, the traffic flows and the network topology characteristics. Bassolas et al. (2019) associated the urban mobility network hierarchy with the traffic flow between city hotspots. They used the trip flow between city hotspots and assigned the city hotspots to five levels, in which a high network hierarchy is derived from more connections within similar-level hotpots but fewer connections between different levels.

In PTN hierarchy research, network hierarchy is also regarded as a network property and is evaluated by topological network characteristics. Von Ferber et al. (2009) developed the essential methodology for evaluating the PTN hierarchy. Four PTN representation spaces were developed to interpret the PTN characteristics better. The four spaces are L-space for infrastructure, P-space for service, C-space for transfer availability, and B-space for public transport line coverage. With the four-space PTN representation, Wang et al. (2020) fused the PTN hierarchy with passenger transfer flow data and identified the time dependence of the PTN hierarchy by optimising the network modularity in C-space. The PTN achieves optimal modularity when it reaches the highest inter-group passenger flow heterogeneous on edges, indicating the PTN hierarchy. Buijtenweg et al. (2021) quantified the city's PTN hierarchy in L-space and proposed a hierarchy degree metric, regarding the stops as vertices and the connection between stops as edges. The hierarchy degree in this research is affected by three vertex-based topological attributes: topological influence, transfer redundancy, and transfer potential. Incorporating with passenger flow data, each attribute is quantified based on a vertex-based topological indicator: eigenvector centrality, clustering coefficient and vertex degree. Eventually, the Gini coefficient of vertex "hierarchy degrees" is regarded as the quantified PTN hierarchy.

In a word, the network hierarchy as a network property is consistent with the research. When regarding the hierarchy as a network property, it reflects not only a numerical relationship between the importance of network elements but also the organisation of network elements based on the heterogeneity of their importance.

2.3. Topological characteristic-related research

The network hierarchy can be reflected in the network characteristics from multiple aspects, especially by using the topological characteristics. The topological characteristics of network hierarchy describe the organisation of network elements (Groth and Skandier, 2005). The organisation of network elements is evaluated from the network and element scales, reflecting corresponding network hierarchy topological characteristics.

2.3.1. Element-scale topological characteristics

The element-scale characteristics are the topological characteristics embodied in network elements, which reflect and quantify the importance of network elements. The network hierarchy is high when the elements in a network follow an organisation, where the number of elements gradually descends with their importance ascends, and the majority of elements have low importance, while a few elements have high importance.

The importance of network elements can be embodied in different aspects. Four main element-scale topological characteristics have been studied in past research, which are the accessibility of vertices (Háznagy et al., 2015; Huang et al., 2014; Luo et al., 2019), the intermediacy of elements (Derrible, 2012; Kanrak and Nguyen, 2022; Zhang et al., 2013), the importance of sub-networks or vertex clusters (Háznagy et al., 2015; Hong et al., 2019; Pattiselanno et al., 2015) and the connectivity of vertices (Hong et al., 2019; Shanmukhappa et al., 2018). Among the four element-scale topological characteristics, the connectivity of vertices is particular. In the previous works, two main topological indicators are used for quantifying the connectivity of vertices, the hub and authority centrality (Kleinberg, 1999) and the vertex degree (Newman, 2010). The hub and authority centrality originates from a web page ranking algorithm, identifying the hub and authority vertices based on large numbers of in or out connections. Shanmukhappa et al. (2018) applied the hub and authority centrality to the L-space bus networks in Hong Kong, London and Bengaluru, identifying the hub and authority vertices in the network. Because most of the connections between the bus stops are bi-directional, it was found that the identified hub and authority vertices have a large overlap. Although the hub and authority centrality is a quantified indicator, it qualitatively identifies the hub or authority vertices, which is not appropriate to reflect the heterogeneity of vertices' importance for PTN hierarchy assessment. As for vertex degree, which is a basic indicator for quantifying the connectivity of single vertices. Applying the vertex degrees in the network scale, the indicator can further quantify the extent of scale-free network structures. Thus, this study categorises the vertex degree as a network-scale characteristic. Other than vertex connectivity, the accessibility, intermediacy and importance of sub-networks or vertex clusters comply with the definition of the PTN hierarchy and are discussed in detail in the following paragraphs.

Vertex accessibility
The accessibility of vertices denotes the difficulty of reaching a vertex from other elements in the network. In PTNs, the difficulties of reaching a vertex are shown by travel impedance, such as travel time, distance, and ticket fare. The vertices with the lower impedance to reach indicate higher importance. There are three main topological indicators used in the previous work quantifying vertex accessibility, the average travel impedance (Luo et al., 2019), eccentricity centrality (Hage and Harary, 1995), and closeness centrality (Bavelas, 1950). Luo et al. (2019) defined a travel time-based travel impedance metric, evaluating the vertex accessibility in tram networks worldwide by average travel impedance between vertices. The research verified that the travel-time information from other data has benefits for vertex accessibility evaluation, such as the AFC data. However, the average travel impedance focus on the impedance between vertex pairs in P-space networks, which does not comply with L-space methods in this research. Háznagy et al. (2015) utilised the distribution of the vertices' eccentricity centrality in the comparative analysis of PTNs in five Hungarian cities. The networks were represented in L-space and weighted by the vehicle capacity of edges in the morning peak. It has been found that a large city area decreases vertex accessibility in PTNs. The eccentricity centrality of vertices can partly denote the vertex accessibility by calculating the greatest impedance to other vertices, but it is not comprehensive to reflect the general accessibility of vertices by considering the impedance to all other vertices in the network. Luo et al. (2020) investigated the relationship between the quantified element properties by centrality indicators in L-space and the passenger flow distribution in P-space. The closeness centrality is used as the indicator for quantifying vertex accessibility. The case study on the Hague and Amsterdam tram networks shows that the PTN topological properties can be used to approximate the global passenger flow distribution on the PTNs. Hong et al. (2019) compared L-space elements' topological characteristics between the unimodal PTNs, such as bus and metro networks, with the integrated PTN in Soul, South Korea. The closeness centrality of vertices is utilised as the indicator for vertex accessibility. They found that the vertex accessibility is increased in the integrated PTN compared to the unimodal PTNs. In this research, the closeness centrality is appropriate to quantify the vertex accessibility in L-space by comprehensively considering the impedance between all vertex pairs in the networks.

Network element intermediacy
The intermediacy of network elements describes the importance of elements in forming connections between other elements and further denotes the flow passing through network elements (Zhang et al., 2013). The intermediacy can be embodied on vertices or edges. In past research, the widely used topological indicator for quantifying elements' intermediacy is

betweenness centrality (Freeman, 1977). Derrible (2012) applied the betweenness centrality as the indicator for vertex intermediacy in 28 L-space metro networks worldwide. It was found that the vertex intermediacy becomes more evenly distributed with the increasing sizes of metro networks. Zhang et al. (2013) studied the L-space topological characteristics of urban rail networks worldwide and used betweenness centrality as the intermediacy indicator. They found the average betweenness centrality of vertices and edges linear increases with the increase of vertices in urban rail networks. Besides, when the vertex with the highest intermediacy is attacked, the network failure is the most severe. Shanmukhappa et al. (2018) used the L-space betweenness centrality of vertices to identify the "supernodes" in the bus networks of Hong Kong, London and Bengaluru. They found that removing the vertices with high intermediacy would significantly increase the average path length and change the routing behaviour of both passengers and buses. Thus, the betweenness centrality can well present the intermediacy of network elements, fitting well with the definition of PTN hierarchy in this research.

Sub-network or vertex cluster importance

Sub-networks and vertex clusters are part of the networks, consisting of several vertices and connecting edges. The number of sub-networks and vertex clusters with different importance can also be quantified and reflect the PTN hierarchy. In the previous works, the main topological indicators for quantifying the sub-networks or vertex clusters are clique (Bron and Kerbosch, 1973), Pagerank centrality (Page et al., 1999), and eigenvector centrality (Bonacich, 1987). The maximum clique are frequently used in social network assessment. Rowe et al. (2007) formed an undirected social network in a company by email communications, where the weights of edges are the number of emails and the average response time between two people. The cliques in the network are identified and assigned scores based on the sizes of cliques. With the clique scores, the social hierarchy is detected. However, the clique is not appropriate for this research. First, the clique is applicable for undirected networks, such as social networks, rather than the directed PTNs. Second, although cliques can be applied to identify the important sub-networks in the network, the quantification of importance is limited, which is usually reflected by the sizes of cliques but cannot embody the difference of sub-network importance in detail. Pagerank centrality and eigenvector centrality both define that a vertex's importance depends on the vertices it connects with. When the Pagerank centrality or eigenvector centrality value of a vertex is high, then the importance of the cluster consisting of the vertices and all the connecting vertices is high. Háznagy et al. (2015) applied Pagerank centrality to identify the important vertices in L-space PTNs of Hungarian cities and found the distributions of the Pagerank centrality in the five PTNs are similar. Soh et al. (2010) applied daily passenger flow-weighted eigenvector centrality for quantifying the importance of the stations in the L-space rail transit network and bus network in Singapore. A dynamic analysis of the two networks' topological characteristics is conducted by calculating the eigenvector centrality values for stations with weekday and weekend passenger flow data. The case study results showed that the rail transit network's topological characteristics are less affected by the time-dependent passenger flow, while the bus network was deeply affected. Compared to the eigenvector centrality, the Pagerank centrality only considers the in-degree of vertices. In this research, the PTNs focus on the connections between infrastructures, so the difference between in and out vertices connections is unimportant. Thus, the eigenvector centrality is more applicable for quantifying the importance of vertex clusters.

2.3.2. Network-scale topological characteristics

The network-scale topological characteristics are reflected in the overall network structures. Some researchers investigated the association between network hierarchy and network structures. Ravasz and Barabási (2003) found that scale-free networks are strongly associated with high network hierarchy and high-clustering network structures. The researchers calibrated a “hierarchical network” model with high-hierarchy and high-clustering networks to prove the association. The model results indicate whether or not the input network is high-hierarchy and high-clustering. When applying the proven scale-free networks to this model, such as the actor co-starring network, the World Wide Web network, and the language network, the model results show these networks are high-hierarchy and high-clustering. The results proved that the scale-free network is associated with a high-clustering network structure and a high network hierarchy. Besides, Czégel and Palla (2015) discussed some specific network structures and the network hierarchy with the random walking method. This research discusses three specific network structures: the star, tree and chain-shape network structures. These network structures in specific shapes result from specific connection patterns of vertices. It was found that in directed networks, the tree network structure has a higher network hierarchy than the chain or star-shaped network structure, for vertices in tree network structures connect to vertices with similar importance. When the random walker density in the random walking method reaches convergence, the tree-shaped network structure shows the highest heterogeneity of elements’ importance and indicates a higher network hierarchy.

Scale-free network structure

The scale-free network structure is strongly associated with a high network hierarchy. In the scale-free network structure, a few vertices take up the majority of connections. The vertices with many connections are the minority, while the majority has a few connections. Solé and Valverde (2004) associated scale-free networks with high heterogeneity of network elements. The research summarised that complex networks have three topological characteristics: network heterogeneity, modularity and randomness. Visualised by a 3-dimensional complex network space, scale-free networks always associate with high network heterogeneity of network elements, consistent with high-hierarchy networks’ characteristics. To identify and analyse scale-free networks, vertex degree is the most widely used topological indicator. Tsiotas and Polyzos (2015) and Von Ferber et al. (2009) claimed that the distribution of vertex degrees reflects the network structures. Solé and Valverde (2004) found that the vertex degrees always exhibit a long-tail distribution. Moreover, Barabási and Bonabeau (2003); Newman (2005) found that the degree distribution in scale-free networks is usually associated with the power-law distribution. However, according to Broido and Clauset (2019), scale-free networks are rare in reality, and the degree distribution rarely follows the power-law distribution in real-life networks. Clauset et al. (2008) found that the real-life scale-free networks usually follow the skewed normal distribution. In particular, being PTNs a real-life example, the scale-free network structures can be identified and analysed by vertex degrees’ skewed normal distribution. The better goodness of fit between the vertex degree distribution and the skewed norm distribution can illustrate the higher network hierarchy.

High-clustering network structure

The high-clustering networks have a high potential to be deconstructed and rebuilt into vertex clusters with a higher density of connections within clusters than between clusters. There are two main topological indicators for quantifying the extent of high-clustering structures, the global clustering coefficient (Luce and Perry, 1949) and the modularity coefficient (Newman, 2006). The global clustering coefficient is a topological indicator that measures the degree of vertices in a network that tends to form triplet clusters. For example, Dimitrov and Ceder (2016) examined the structure and topological properties of the L-space bus route network in Auckland, New Zealand. The case study results showed that the bus route network is high-clustering but not “small-world”, which is a type of network having most nodes not directly connected, but vertices can have indirect connections with other vertices through a few vertices. Since the empirical value for the average path length is longer than the theoretical length in a random graph, which cannot meet the requirement of a small-world network. Although the global clustering coefficient can quantify the degree of a network being high-clustering, it only focuses on the triplets or full-connected cliques. The connections not in vertex triplets or cliques are not included. The optimal modularity coefficient also quantifies the degree of high-clustering structures. Some work has been done on applying the optimal modularity coefficient in network hierarchy identification. Guimerà and Amaral (2005a) proposed an optimal modularity-based hierarchy identification methodology. In high-clustering networks, when the optimal modularity coefficient is reached, the roles of vertices are identified based on the vertex's connections within the cluster and between clusters. Then they applied the methodology and identified the network hierarchy in the high-clustering organism's metabolic networks (Guimerà and Amaral, 2005b) and worldwide aviation networks (Guimerà et al., 2005). The optimal modularity coefficient does not confine to the full-connected clusters. The vertices with localised denser connections can be identified as clusters, which contributes to the high-clustering structure. Thus, the optimal modularity coefficient is more appropriate for this research.

Vertex connection pattern

Vertex connection following a unified pattern contributes to the high network hierarchy. For example, the vertices in a network are prone to connect the vertices having a similar number of connections. If the vertex connections follow some patterns, the network structure shows a specific shape. Jiang et al. (2017) studies the two types of high-hierarchy network structures following the vertex connection based on vertex degree. The two network structures are named core/periphery network and star-like network, as shown in Figure 2.1. In a network following vertices connect to vertices with similar degrees, namely high-degree vertices prone to connect high-degree vertices, and low-degree vertices prone to connect low-degree vertices pattern, then the network is concentrated and shows a core/periphery structure. In a core/periphery network, the vertices in the core area have higher importance than the vertices in the periphery areas, which can be regarded as a variant of the tree-shape networks studied by (Czégel and Palla, 2015). By contrast, in a network following low-degree vertices prone to connect high-degree vertices patterns, the vertices with similar degrees are dispersed. The network is decentralised and shows a star-like structure.

(a) Core/periphery network (b) Star-like network

Figure 2.1: Network structures with vertex connection patterns (Jiang et al., 2017)

Both the centralised and decentralised network structures show heterogeneity of vertex importance. In core/periphery networks, the vertices in the core areas have higher importance than the periphery areas. The importance of vertices is gradually decreased from the core to the periphery, which complies with the definition of hierarchical networks. In star-like networks, the vertices with more local connections show higher importance than the low-degree vertices connected to them. The two types of networks are the two extremes of networks having vertex connection patterns based on vertex degrees. To evaluate the extent of a network having a vertex connection pattern, the assortativity coefficient is widely chosen. Assortativity is a concept derived from the definition of the assortative network, where well-connected vertices are prone to connecting other well-connected vertices (Newman, 2002). Kanrak and Nguyen (2022) applied the assortativity coefficient to the Asian-Australasian cruise shipping network. Regarding ports as the vertices, the port connection patterns were evaluated. The case study results showed that ports with a similar number of connections tend to connect to each other. Chopra et al. (2016) assessed the resilience of the London metro network by the network topology with passenger flow data. The results of the case study showed there is no significant connection pattern where high-degree vertices connect high-degree vertices, and more connections are low-degree vertices connect low-degree vertices. Plus, the results also show the London metro network does not show a small-world structure, and the conclusion is the London metro network is vulnerable. With the assortativity coefficient, the extent of a network following a connection pattern can be quantified, so that this network-scale topological characteristic of different PTNs can be compared.

2.3.3. Topological indicator summary

Based on the topological characteristics of PTN hierarchy from network and element scales, the related research is reviewed. Considering the applicability to the PTN hierarchy, not all topological indicators are included in the content. Therefore, Table 2.1 offers an overview of all these topological indicators referenced in the research. In the column of other topological indicators, there are element-scale or network-scale topological indicators which do not comply with PTNs or embody the network organisation. For vertex-based indicators, they are less suitable for reflecting the elements' importance in PTN networks, for example, the information centrality. For network-based indicators, they focus more on a specific aspect of networks, rather than reflecting overall element organisations, such as the network diameter and average shortest path length. With the review of the indicators in the previous research, the topological indicators selected in this study can reflect the importance of PTN elements with element-based indicators and interpreter the overall structures of high-hierarchy networks with network-based indicators.

Table 2.1: Summary of studies on network hierarchy using topological indicators

Literature	Network objects	PTN representation	Directed/ Undirected	Weighted/ Unweighted	Closeness centrality	Betweenness centrality	Eigenvector centrality	Node degree	Assortativity coefficient	Modularity coefficient	Others topological indicators
Cats (2017)	Urban rail	-	Directed	Unweighted	√	√		√			Total network length, network diameter, network connectivity, network meshedness, network directness
Gattuso and Miriello (2005)	Metro	L-space	Directed	Unweighted				√			Total network length, network diameter, network connectivity, number of loops
Luo et al. (2019)	Tram	L-space, P-space	Directed	Travel times	√						Number of nodes and links
Cats and Birch (2021)	Multi-modal PTN	L-space	Directed	Travel times		√		√			Total network length, network connectivity, network directness, average link length
Silva et al. (2022)	Air	-	Directed	Number of flights	√			√	√		Clustering coefficient, network density, authorities and hubs
Huang et al. (2014)	Social	-	Undirected	Unweighted	√	√	√	√			Clustering coefficient, information centrality, vertex efficiency
Tanglay et al. (2022)	Neurosurgery	-	Undirected	Weighted			√	√			Pagerank centrality
Von Ferber et al. (2009)	PTN	L-space, P-space, C-space	Undirected	Unweighted	√	√		√			Shortest path length
Tsiotas and Polyzos (2015)	Air	-	Directed	Number of flights	√	√		√		√	Mobility centrality, clustering coefficient
Chen et al. (2022b)	Air	-	Undirected	Total carriers	√	√		√			
Guan et al. (2020)	International trade	-	Directed	Unweighted	√						
Solé and Valverde (2004)	-	-	Undirected	Unweighted				√	√		
Ravasz and Barabási (2003)	-	-	Undirected	Unweighted				√		√	
Czégel and Palla (2015)	-	-	Directed	Weighted				√			Pagerank centrality
Clauset et al. (2008)	-	-	Undirected	Unweighted				√	√		
Kanrak and Nguyen (2022)	Cruise shipping	-	Directed	Unweighted	√	√		√			Network density, average path length, clustering coefficient
Guimerà et al. (2005)	Air	-	Directed	Number of flights		√		√		√	Average shortest path, clustering coefficient
Guimerà and Amaral (2005b)	Metabolic	-	Undirected	Unweighted				√		√	
Bienenstock and Bonacich (2021)	-	-	Directed	Unweighted			√	√	√		
Meunier et al. (2009)	Brain function	-	Undirected	Unweighted				√		√	
Pattiselanno et al. (2015)	Social	-	Undirected	Unweighted							Clique
Loginova et al. (2022)	City	-	Undirected	Number of firms				√			K-core
Mones et al. (2012)	-	-	Directed	Unweighted				√			Local reaching centrality
Badhrudeen et al. (2022)	Road	-	Undirected	Unweighted				√			
Dimitrov and Ceder (2016)	PTN	L-space	Directed	Unweighted				√			Average shortest path, clustering coefficient
Hong et al. (2019)	PTN	L-space	Directed	Inversed distance	√	√	√	√	√	√	Network diameter, average path length clustering coefficient, eccentricity centrality
Soh et al. (2010)	Rail, bus	L-space	Bi-directed	Daily travel			√	√	√		Clustering coefficient
Háznagy et al. (2015)	PTN	L-space	Directed	Peak-hour capacity	√	√		√			Pagerank centrality, average path length
Cats et al. (2020)	PTN	L-space	Undirected	Distance				√			Ringness
Fernandes et al. (2019)	Air	-	Directed	Number of flights				√			
Luo et al. (2020)	PTN	L-space, P-space	Directed	Passenger flow, travel time	√	√		√			
Rowe et al. (2007)	Social	-	Undirected	Number of mails		√		√			Clique, clustering coefficient
Buijtenweg et al. (2021)	PTN	L-space, P-space	Directed	Passenger flow			√	√			Clustering coefficient
Wang et al. (2020)	PTN	L-space, C-space	Directed	Passenger flow							
Shanmukhappa et al. (2018)	Bus	L-space	Directed	Number of lines		√	√	√			Clustering coefficient, average shortest path length, hub and authority centrality
Zhang et al. (2013)	Urban rail	L-space	Undirected	Unweighted		√		√			Clustering coefficient, average shortest path length, topological efficiency
Derrible (2012)	Metro	L-space	Undirected	Unweighted		√					
Chopra et al. (2016)	Metro	L-space	Directed	Passenger flow				√	√	√	Clustering coefficient, average shortest path length, global efficiency
Robson et al. (2021)	Infrastructure	-	Undirected	Unweighted		√		√	√		Cycle basis
This study	PTN	L-space	Directed	Average travel times of trips Inverse of average travel times of trips	√	√	√	√	√	√	

3
Methodology

In Section 3.1, the definition of the PTN hierarchy and the PTN hierarchy's topological characteristics are clarified before proceeding to the quantification. Section 3.2 introduces the graph representation of the topological PTN hierarchy quantification. Section 3.3 and Section 3.4 present the topological quantification methods of PTN hierarchy's topological characteristics, including vertex-based indicators and network-based coefficients.

3.1. Problem description

Before starting with the PTN hierarchy quantification, the definition of the PTN hierarchy and the topological characteristics of the PTN hierarchy are clarified and determined.

3.1.1. PTN hierarchy

Firstly, the PTN hierarchy is a network property of PTNs, which defines an organisation of PTN elements, such as vertices, edges and vertex clusters. The organisation has the following features: PTN elements have high heterogeneity of importance, and the number of PTN elements gradually descends with the importance ascending. The PTN elements with high importance are the minority, and the majority has low importance. The importance of PTN elements is multi-aspect, which can be reflected by the PTN hierarchy's topological characteristics. Element organisations having these features are defined as hierarchical organisations.

3.1.2. Topological characteristics of PTN hierarchy

PTN hierarchy has topological characteristics from element and network scales. Six topological characteristics of the PTN hierarchy are chosen in this study and summarised as follows:

1. Network element-scale characteristics
    - PTN elements follow the hierarchical organisation by vertices' accessibility
    - PTN elements follow the hierarchical organisation by elements' intermediacy
    - PTN elements follow the hierarchical organisation by vertex clusters' importance
2. Network-scale characteristics
    - The PTN structure is scale-free
    - The PTN structure has a high-clustering level
    - Vertex connections follow a pattern based on vertex degrees

Each topological characteristic can interpret the PTN hierarchy from one aspect, forming a dimension of the PTN hierarchy. The higher extent a PTN meets one of the characteristics, the higher hierarchy the PTN has in the dimension of this indicator. The following PTN hierarchy quantification methodology is based on the six topological characteristics. Topological indicators quantify the extent to which a PTN has these topological characteristics.

3.2. PTN representation

Based on the PTN representation forms developed by Von Ferber et al. (2009), PTNs in topological hierarchy quantification are presented in L-space. L-space is the representation form of infrastructure in the PTN. A stop represents a vertex v_i. The set of all stops v in a PTN is represented as V. If at least one public transport line directly connects two stops i and j, the two stops are connected with an edge $e(i,j)$. Every two connected stops can have a maximum of an edge. The self-loop edges having the endpoints as the same vertex is not permitted. The set of all the edges in a PTN is represented as E.

The edges in PTNs are directed as it is how in reality operates. The edge $e(i,j) \neq e(i,j)$. The directed edges can represent the directions of connections between stops. In metro and tram networks, most connections between stops are in two opposite directions, represented by the bi-directed edges in the network, for which both $e(i,j)$ and $e(j,i)$ exist. PTNs like BRT networks sometimes have stops that have connections with other stops only in one direction, presented by single-directed edges. The $e(i,j)$ is feasible for the single-directed edges, but $e(j,i)$ is not.

The edges in PTNs are also weighted. The weight of the edge $e(i,j)$ is represented as $W(i,j)$. In the topological characteristics of the PTN hierarchy, the impedance of connections and the strength of connections are represented by two types of edge weights. The impedance of connections is weighted by the average travel time of the trips on the edges. The average travel time of edge $e(i,j)$ is mathematically represented as $t(i,j)$. By contrast, the strength of connections is represented as the inverse of the average travel time of the trips on the edges. The connection strength of edge $e(i,j)$ is mathematically represented as $s(i,j)$.

In a word, the PTNs are weighted and directed with the L-space PTN representation. The PTNs are represented as $G = (V, E, W)$. The PTN G consists of stops as vertices V and connections as edges E. Two types of edge weights of an edge $W(i,j)$, the connection impedance weight $t(i,j)$ and the connection strength $s(i,j)$.

3.3. Vertex-based indicators

The vertex-based indicator dimensions use the goodness of fit to quantify the PTN hierarchy topological characteristics, where the goodness of fit is between the vertex-based topological indicators' probability density distribution and the skewed normal distribution. The following sub-sections detail the goodness of fit and the four vertex-based indicator dimensions.

3.3.1. Goodness of fit

The goodness of fit is used to represent the quantified PTN hierarchy in the dimensions of vertex-based indicators. When the PTN elements follow the hierarchical organisation by their importance, the distribution of the PTN elements' importance follows a skewed normal distribution (Clauset et al., 2008). The goodness of fit measures the extent of PTN elements following the skewed normal distribution by their importance. The importance of PTN elements is multi-aspect, and each aspect is reflected by one vertex-based topological indicator.

The coefficient of determination, also known as R square (R^2), is selected as the measurement of the goodness of fit for two reasons. First, the R square can deal with PTNs having different sizes of elements, ensuring the hierarchy comparison between the PTNs. Second, the difference between R square values of PTNs is meaningful, it denotes the ratio that the fitted skewed normal distribution explains the probability density distribution of the vertex-based indicators. The difference in R square values can quantify and reflect the difference between the PTN hierarchy.

The R square denotes the proportion of data explained by the regression distribution (Rao, 1973). The equations of the coefficient of determination are shown as Equation 3.1.

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{3.1}$$

$$SS_{\text{res}} = \sum_i (y_i - f_i)^2 \tag{3.2}$$

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2 \tag{3.3}$$

A data set has n values denoted as $y_1, ..., y_n$, collectively known as y_i. The regression values of the data are denoted as $f_1, ..., f_n$, collectively known as f_i. The average value of the data set is $\bar{y}$. The sum of squares of residuals (SS_{res}) indicates the differences between the data values y_i and the regression values f_i. The total sum of squares (SS_{tot}) indicates the difference between the data values and the average data value $\bar{y}$. The R square equals the difference between 1 and the fraction of SS_{res} and SS_{tot}. The value range of R square is from negative infinity to 1. If the regression values all equal the data value, the R square value is 1, indicating the best goodness of fit. When the R square equals 0, the goodness of fit is as good as the average value of the data set. The R square can also be negative when the regression values have a worse fit than the average data value. The more the R square value between a data set and the regression distribution, the better goodness of fit they have.

The probability histograms reflect the probability density distributions of vertex-based topological indicators. In the histogram, the x-axis indicates the bins of topological indicators values, while the y-axis indicates the probability density of vertices allocated to the bins. The number of bins in the histogram is determined with Equation 3.4 with Scott's normal reference rule (Scott, 1979), which considers the variance of data and the data size. In Equation 3.4, h indicates the bin width, $\hat{\sigma}$ is the standard deviation of the data, and n is the size of the data. In the R square value calculations and PDF representations, the midpoints of bins in histograms are applied as the representative values of probability density distributions.

$$h = \frac{3.49\hat{\sigma}}{\sqrt[3]{n}} \tag{3.4}$$

The logistic regression for the skewed normal distribution uses the coordinates composed of the median values of bins as the x values and the height of the bins as the y values. The probability density function (PDF) $\phi(x)$ of the skewed normal distribution is shown as Equation 3.5.

$$\phi(x) = \frac{2}{\omega\sqrt{2\pi}} e^{-\frac{(x-\xi)^2}{2\omega^2}} \int_{-\infty}^{\alpha(\frac{x-\xi}{\omega})} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \tag{3.5}$$

In the equation, x is the independent variable, ranging from negative infinity to positive infinity. Real numbers ξ, ω and α denote the distribution's location, scale and shape parameters.

In this study, the comparison between the PTN hierarchy in each dimension is relative. Considering the difference in topological characteristics of the PTN hierarchy, the distribution ranges of R square values in topological dimensions vary, and R square values in different indicator dimensions are not directly comparable. The comparison of all PTN hierarchy indicators is introduced in the following sections. In this study, R square is the method quantifying the PTN hierarchy, rather than the parametric test of the skewed normal distribution, so no non-parametric test is required. For PTNs investigated in this study, the probability density distribution of vertex-based indicators is pre-assumed following the skewed normal distribution. R square values denote the extent that the PTN element organisation follows the hierarchical organisation, namely the PTN hierarchy. The higher R square value of a PTN indicates a higher hierarchy the PTN has compared to other PTNs in the dimension of a vertex-based indicator.

3.3.2. Vertex degree

The vertex degree dimension is the PTN hierarchy dimension that uses vertex degree as the topological indicator. The R square value of the vertex degree PDF fitting the skewed normal distribution indicates the degree of PTN's scale-free structure and further denotes the PTN hierarchy. Besides, the dimension is the vertex-based indicator dimension that reflects the network-scale PTN hierarchy topological characteristic.

The vertex degree, also known as degree centrality, is the number of edges a vertex connects to (Newman, 2010). The degree is one of the most basic indicators in network topology, describing the vertex's connectivity. In directed networks, the degree includes in-degree and out-degree, and the vertex degree equals the sum of the in-degree and out-degree. To exclude the differences in the number of vertices between PTNs, the vertex degree values are normalised by dividing by the maximum possible degree of the network. For a network, $G = (V, E)$, the network's adjacency matrix is A. A_{ij} is one of the elements in the adjacency matrix, showing whether a connection exists from vertex i to vertex j. When it exists, A_{ij} equals 1. Otherwise, the value is 0. Thus, the degree of the vertex k_i is equal to:

$$k_i = \sum_j A_{ij} + \sum_j A_{ji}, \quad \forall i, j \in V \tag{3.6}$$

3.3.3. Closeness centrality

The closeness centrality dimension is the PTN hierarchy dimension using the closeness centrality of vertices as the topological indicator. The R square of the closeness centrality PDF of vertices fitting the skewed normal distribution denotes the degree of vertices following the hierarchical organisation by their vertex accessibility and further denotes the PTN hierarchy.

In a PTN consisting of N vertices, the closeness centrality C_i of vertex i is defined as the harmonic mean value of the weighted shortest path lengths d_{ij} from vertex i to every other vertex j, shown as Equation 3.7. Apart from the endpoints of the shortest path from i to j, there are P vertices located in the middle of the shortest paths are $k_1, \dots, k_P$. The lengths of the shortest paths d_{ij} are defined as the sum of the weights of edges located on the paths, namely the average travel time of trips on the edges, shown in Equation 3.8. To exclude the differences in the number of vertices between PTNs, the closeness centrality of vertices is normalised by the sum of all the shortest path lengths in the PTN.

$$C_i = \frac{1}{N-1} \sum_{j(\neq i)} \frac{1}{d_{ij}}, \quad \forall i, j \in V \tag{3.7}$$

$$d_{ij} = s(i, k_1) + \sum_{p=1}^{P-1} s(k_p, k_{p+1}) + s(k_P, j) \tag{3.8}$$

A better understanding of the closeness centrality distribution patterns in specific network configurations helps analyse the organisation of vertices by their accessibility. As an example, the four network configurations are shown in Figure 3.1. The network configurations are undirected and unweighted. In the examples, vertices with more connections have higher closeness centrality than their connected vertices with fewer connections.

Linear network segment is defined as a network configuration where each vertex connects one after the other sequentially from one end to another. Figure 3.2 offers two examples of linear network segments from vertex a to b. The linear network segment consists of two endpoints a and b, which are the vertices connecting one vertex or more than two vertices. The endpoints that connect only one vertex are called the *dead ends*, for example, the vertex b in Figure 3.2(b). The other vertices located on the linear segments are denoted as k_1 to k_p. There is a pattern of the closeness centrality value changes on linear network segments. For the vertices on the same linear network segment, the closer the vertices are to the endpoint with more connections, the higher their closeness centrality.

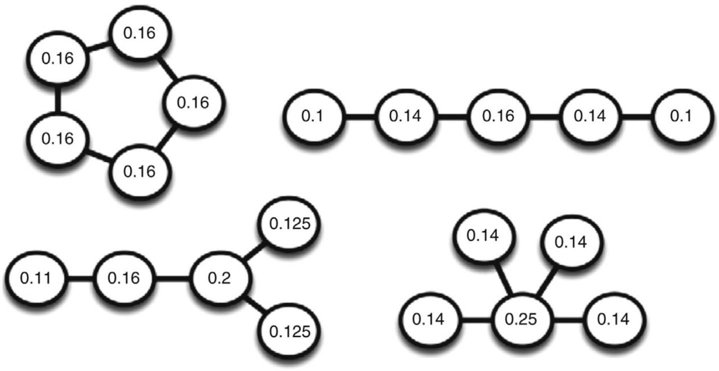
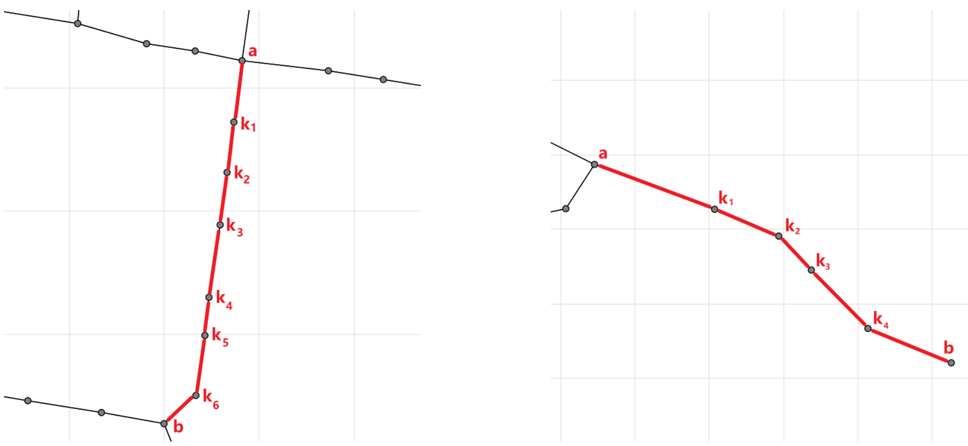


Figure 3.1: Closeness centrality values in network configurations (undirected & unweighted) (Charles and Rony, 2016)



(a) (b)

Figure 3.2: Linear network segments from endpoints a to b via vertices $k_{1,\ldots,p}$

3.3.4. Betweenness centrality

The betweenness centrality dimension is the PTN hierarchy dimension that uses the vertex's betweenness centrality as the topological indicator. The betweenness centrality denotes the intermediacy of network elements. The R square between the betweenness centrality PDF fitting the skewed normal distribution indicates the degree of the vertices following the hierarchical organisation by their intermediacy and further denotes the PTN hierarchy.

The betweenness centrality B_i of a vertex measures the times a vertex lies on the shortest paths between other vertex pairs (Newman, 2010). In the identification of the shortest paths, the average travel times of trips on the edges are considered as the weights of edges. The equation of betweenness centrality is shown as Equation 3.9.

$$B_i = \sum_{st} n_{st}^i, \quad \forall i, s, t \in V \tag{3.9}$$

In the equation, i, s and t are three vertices in the network. n_{st}^i denotes whether having one shortest path between vertices s and t that vertex i lies on. Once one shortest path between vertices crossing vertex i, the n_{st}^i equals 1; otherwise, it equals 0. To exclude the vertex amount differences of PTNs, the betweenness centrality values are standardised by the number of possible shortest paths not ending at vertex i in the network, which equals $(n-1)(n-2)$, where n is the number of vertices in the PTN.

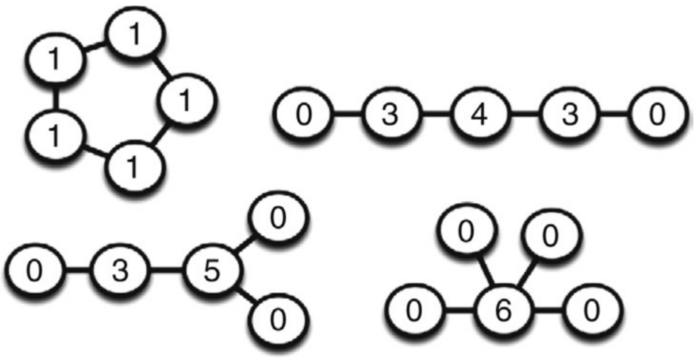


Figure 3.3: Betweenness centrality values in network configurations (undirected & unweighted) (Charles and Rony, 2016)

The betweenness centrality of vertices shows patterns in some network configurations. Take examples of simplified unweighted and undirected network configurations in Figure 3.3. The examples show that the betweenness centrality of vertices with more connections is higher than the connected vertices with fewer connections. The betweenness centrality of linear network segments' endpoints with one connected vertex is 0, lying on no shortest path. Vertices on the same linear network segment have higher betweenness centrality when they are to endpoints with more connections. Another pattern is that the betweenness centrality of vertices with the same degree and on the same linear network segment is similar. Except for the shortest paths that end at the vertices on the linear network segment, the times that shortest paths pass through these vertices are the same. Because of this pattern of betweenness centrality values, the vertex-based betweenness centrality can reflect the intermediacy of not only the vertices but also the linear network segments.

3.3.5. Eigenvector centrality

The eigenvector centrality dimension is the PTN hierarchy dimension that uses the eigenvector centrality of vertices as the topological indicator. The eigenvector centrality values denote the importance of the vertex clusters consisting of the vertices and the vertices they connect to. The R square between the vertex's eigenvector centrality PDF fitting the skewed normal distribution denotes the degree of the vertex clusters following the hierarchical organisation by their importance and further denotes the PTN hierarchy.

The eigenvector centrality E_i of vertex i is proportional to the sum of the eigenvector centrality of vertices connecting vertex i (Newman, 2010). The calculation of eigenvector centrality considers the strengths of the connections between vertices as the weights of edges. The strength of connections is represented by the inverses of the average travel times of the trips on the edges. The following Equation 3.10 shows the equation of eigenvector centrality.

$$E_i = \kappa_1^{-1} \sum_j A_{ij} E_j, \quad \forall i, j \in V \tag{3.10}$$

In the equation, i and j are two vertices in the network. E_i denotes the eigenvector centrality of vertex i. κ_1 is a constant, denoting the network's largest eigenvector value of the adjacency matrix A. For directed networks, the calculation of eigenvector centrality only considers the in-edges of vertices to avoid duplication. The results of eigenvector centrality are inherently normalised. The determination of eigenvector centrality to all vertices in the network is iterative. Each vertex is first evenly assigned the same centrality. Then the total eigenvector centrality is reassigned based on the adjacency matrix A_{ij}, until reaching the convergence.

In practical implication, a PTN with a high hierarchy in the dimension of eigenvector centrality indicates a small number of concentrated high-degree stops. Because the importance of vertex clusters is deeply influenced by the importance of the vertex connecting to the centre vertex, a high hierarchical PTN has the important vertex clusters concentrated in a small area and shows a mono-centric structure.

3.4. Network-based indicators

The network-based indicator dimensions use the network-based coefficient as the topological indicator. The values of the network-based coefficients denote the interpretation of the network-scale PTN hierarchy topological characteristics. The details of the network-based coefficient dimensions are introduced in the following sub-sections.

3.4.1. Network modularity

The network modularity dimension is the PTN hierarchy dimension that uses the optimal value of the network-based modularity coefficient as the topological indicator. The value of the optimal modularity coefficient denotes the degree of a PTN's high-clustering structure and further denotes the PTN hierarchy.

The modularity coefficient Q describes a network's potential to be separated into clusters with high similarity within cluster (Newman, 2006). The modularity coefficient is the fraction of the expected connections between vertices that fall within the same clusters minus the fraction of the connections in the situation where the connections are randomly distributed. The equation of the modularity coefficient is shown as Equation 3.11.

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{k_i k_j}{2m}) \delta(c_i, c_j) \tag{3.11}$$

k_i and k_j are the vertex degree of vertex i and j. m is the total number of edges in the network. Thus, $2m$ is the number of endpoints of all the edges. $\frac{k_i k_j}{2m}$ is the possible number of edges between vertex i and j. c_i is the cluster of vertex i. $\delta(c_i, c_j)$ is the Kronecker delta of the clusters of vertex i and j. When c_i and c_j are the same clusters, $\delta(c_i, c_j)$ equals 1, otherwise equals 0. $\frac{1}{2m}$ is used for calculating the fraction of the connections.

The value range of the modularity coefficient is $[-\frac{1}{2}, 1]$ (Brandes et al., 2007). If the modularity coefficient is closer to 1, the vertices have more connections in the same cluster, showing a high-clustering network structure. A modularity coefficient close to 0 denotes the vertices are connected randomly. A negative modularity coefficient means the vertices connect more with vertices in different clusters.

The modularity coefficients are influenced mainly by the compositions of vertex clusters. A heuristic method called the Leiden algorithm is used to find the optimal vertex cluster composition (Traag et al., 2019). Developed from the Louvain algorithm (Blondel et al., 2008), the Leiden algorithm is more suitable for directed networks. The algorithm starts by assigning each vertex in the network to a cluster. Then the vertex clusters are aggregated or reassigned to new clusters until they converge to the maximum modularity coefficient.

In practical implication, a PTN having a high hierarchy in the dimension of network modularity indicates a large number and widely distributed high-degree stops. The widely-distributed high-degree stops have more connections with adjacent stops locally, which increases the value of the optimal network modularity coefficient, and the PTN tend to present a multi-centric structure.

3.4.2. Network assortativity

The network assortativity dimension is the PTN hierarchy dimension that uses the assortativity coefficient as the topological indicator. The assortativity coefficient of a PTN denotes the degree of the vertex connections following a unified pattern and further denotes the PTN hierarchy.

Assortativity coefficient r describes the connection pattern of vertices in a network based on the vertex degree (Newman, 2002). The vertex i and j are two vertices in the network. The covariance value of all connections on the edges between two vertices i and j is $cov(k_i, k_j)$. The calculation equation is shown as Equation 3.12. A high covariance value means the similarity of the connected vertex pairs' vertex degrees is high. The assortativity coefficient r is the normalised covariance of vertices degree by the maximum value of the covariance. The maximum covariance happens where the degree values of connecting vertices are the same (Newman, 2010). The normalisation equation of the assortativity coefficient is shown as Equation 3.13.

$$cov(k_i, k_j) = \frac{1}{2m}\sum_{ij}(A_{ij} - \frac{k_i k_j}{2m})k_i k_j, \quad \forall i, j \in V \tag{3.12}$$

$$r = \frac{\sum_{ij}(A_{ij} - k_i k_j/2m)k_i k_j}{\sum_{ij}(k_i \delta_{ij} - k_i k_j/2m)k_i k_j}, \quad \forall i, j \in V \tag{3.13}$$

In Equation 3.12 and 3.13. δ_{ij} is the Kronecker delta indicating whether vertex i and j have the same vertex degrees. When the degrees of two connected vertices are the same, δ_{ij} equals 1, otherwise equals 0.

The value range of the assortativity coefficient is from -1 to 1. The assortativity coefficient is positive when the vertex connections in a network follow a pattern where high-degree vertices connect to high-degree vertices and low-degree vertices connect to low-degree vertices. In contrast, the assortativity coefficient is negative when the vertex connections follow a pattern where the high-degree vertices tend to connect with low-degree vertices. The assortativity coefficient of a PTN is close to either +1 or -1 have a vertex connection pattern in terms of degrees. However, considering the reality of the PTN operation, the assortativity coefficient close to -1 is not feasible. In a PTN where the vertex connections follow a pattern where low-degree vertices tend to connect with high-degree vertices, the PTN has a star-like network structure, as shown in Figure 2.1(b), which is contrary to the reality. For this reason, the connection pattern in this study only considers the pattern where high-degree vertices connect high-degree vertices, and low-degree vertices connect low-degree vertices.

According to the equations of the two network-based indicators, both indicators are derived from the covariance of elements and indicate the similarity or homophily between elements. However, the two indicators assess network elements' similarity between clusters and vertices respectively. The modularity coefficients assess the similarity from the perspective of the clusters, evaluating whether connections exist within the same cluster or between clusters. In contrast, the assortativity coefficients assess similarity from the perspective of vertices, whether connections exist between vertices having similar degrees.

3.5. Six-dimension PTN hierarchy

Hierarchy is a multi-dimension property of PTNs. Apart from the single-dimension comparison of each PTN hierarchy indicator, the multi-dimension comparisons of the PTN hierarchy can lead to a comprehensive assessment. The six PTN hierarchy indicators have different distribution ranges or units. The indicators are scaled to the same range to ensure the comparison is meaningful. The min-max normalisation is used as the scaling method for this study. For a database consisting of multiple PTNs, the original PTN hierarchy in each dimension is represented as H. The normalised PTN hierarchy is H', calculated with Equation 3.14. $max(H)$ and $min(H)$ represent the maximum and minimum values of the PTN hierarchy in each dimension respectively. After the normalisation, the PTN hierarchy in the six dimensions is scaled to $[0, 1]$.

$$H' = \frac{H - \min(H)}{\max(H) - \min(H)} \tag{3.14}$$

The radar chart is developed to offer a comprehensive and intuitive presentation of the six-dimension PTN hierarchy. In the radar chart of a PTN, the six-dimension PTN hierarchy values are presented on a chart with multiple axes originating from the same centre of the radar chart. Figure 3.4 shows an example of the radar chart.

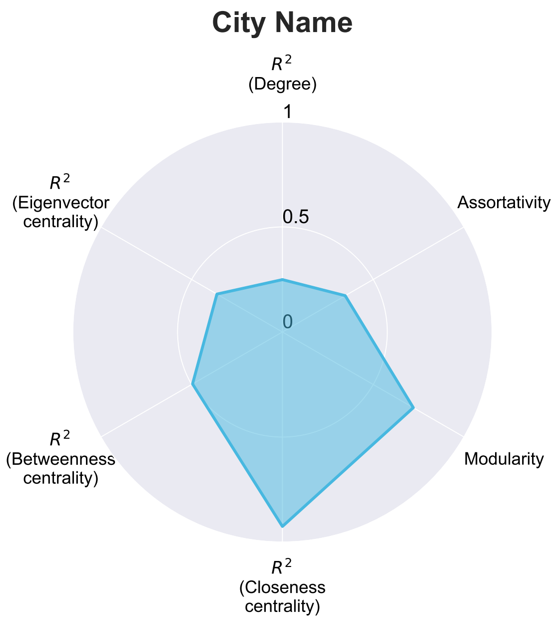


Figure 3.4: An example of a six-dimension radar chart

In the radar chart, each axis represents one PTN hierarchy dimension. For clear presentation, the names of the topological indicators of the PTN hierarchy dimension are simplified as *R^2(Degree), Modularity, Assortativity, R^2(Closeness centrality), R^2(Betweenness centrality)* and *R^2(Eigenvector centrality)*. On the top of the radar chart is the city name of the PTN, and the colour of the enclosed area indicates the mode. The ranges of the axes are all set from 0 to 1. The values on the axes decrease from the edge to the centre of the radar chart. The size of the area enclosed by the connections of value points denotes the overall six-dimension hierarchy of a PTN. The larger size of the enclosed area, the higher the six-dimension hierarchy of the PTN.

4
Case Study

This chapter introduces a case study implementing the methodology, based on GTFS data of high-capacity unimodal PTNs. Section 4.1 provides an overview of the data preparation process for the case study. Section 4.2 presents the quantification and comparison of 63 high-capacity unimodal PTNs worldwide in terms of their hierarchy. Additionally, Section 4.3 and Section 4.4 discuss the mode-wise and continent-wise effects based on the hierarchy of the PTNs.

4.1. Data preparation

The data preparation for the case study includes three processes, data selection, data collection and data processing, in order to obtain suitable high-capacity unimodal PTN data for hierarchy quantification and comparison.

4.1.1. Data selection

The data for this case study are selected from high-capacity unimodal PTNs worldwide. The data should meet the following requirements. First, the mode of the PTNs is one of the three high-capacity public transport modes: metro, tram or BRT. These three modes are commonly used high-capacity public transport modes around the world. Second, the selected unimodal PTNs need to have limited impacts from other modes, being the main unimodal PTN in the whole city. Therefore, the selected unimodal PTNs are mostly the unimodal PTNs with the highest annual ridership.

Following the requirements, 63 high-capacity unimodal PTNs from 62 cities worldwide are selected. Most cities have one unimodal PTN selected, except Berlin. Because of the historical reasons in the post-war period, Berlin's metro and tram network are independently developed, and their mutual impacts are limited. Thus, the metro and tram networks of Berlin are both selected. The PTNs are distributed in six continents: Africa, Asia, Europe, Oceania, North America and South America. The distribution map of the 63 high-capacity unimodal PTNs is shown in Figure 4.1. Each dot represents a PTN, where the colours of the dots indicate its mode. Figure 4.2 is a nested pie chart indicating both modes and continents of the selected high-capacity unimodal PTNs.

4.1.2. Data collection

The data used in this case study is General Transit Feed Specification (GTFS) data, which follows the standard format of transit data published by Google in 2006 (Google, 2022a). The transit data are generated during the public transport operation and collected by the service

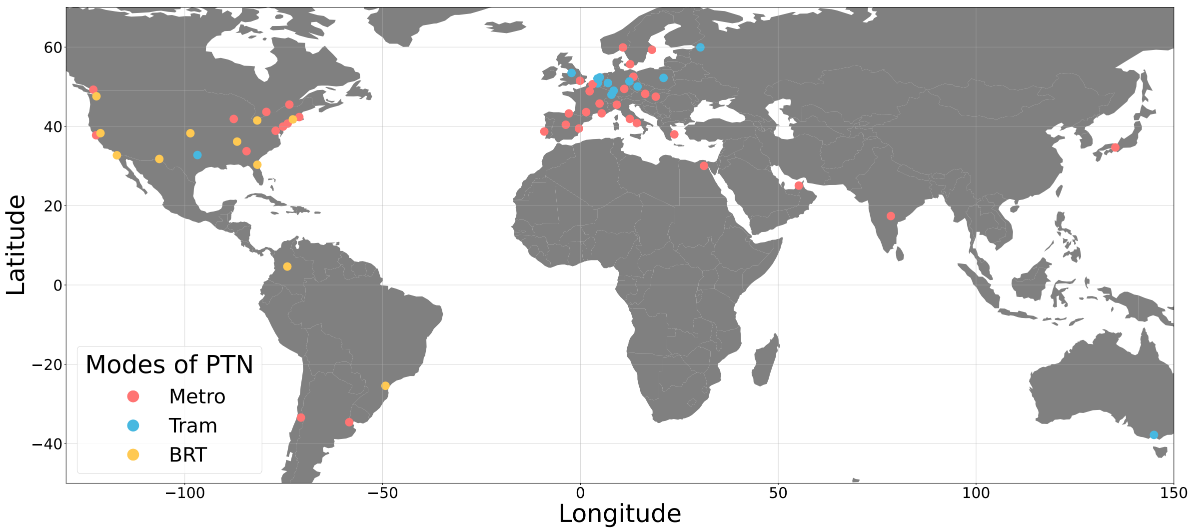


Figure 4.1: Distribution of the selected high-capacity unimodal PTNs worldwide

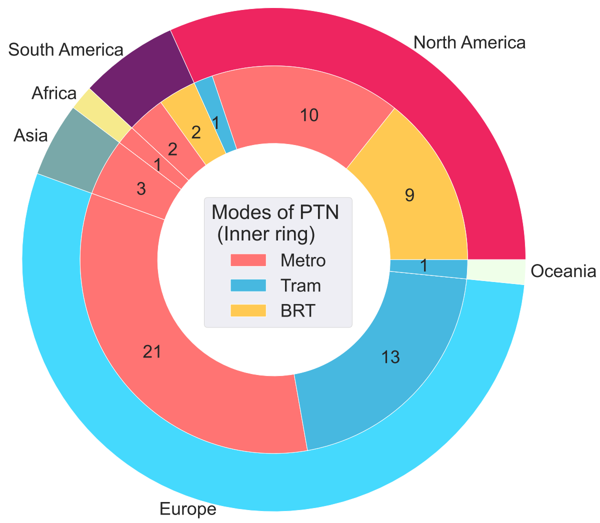


Figure 4.2: Nested pie chart of modes and continents of the selected high-capacity unimodal PTNs

provider. The data includes transit information, such as the stop geographic coordinates, the operating frequency, and the arriving and departing times of stops. The raw GTFS data are compressed as zip files and stored as text or comma-separated values (CSV) files. The source of the GTFS data are mostly from the public GTFS data collection platform, TransitFeeds (OpenMobilityData, 2022), and the rest are requested from the authorities or the public transport service providers.

4.1.3. Data processing

The data processing has four steps, processing the raw GTFS data to analysable PTNs: database generation, data filtering, draft network generating and network revision. The data preparation workflow of the high-capacity unimodal PTNs is shown in Figure 4.3.

Through the database generation step, the raw GTFS data are converted into the SQLite database (Hipp, 2020). The GTFS data analysing library GTFSPY offers effective tools to associate different transit data together based on the route IDs and trip IDs (Kujala et al., 2018).

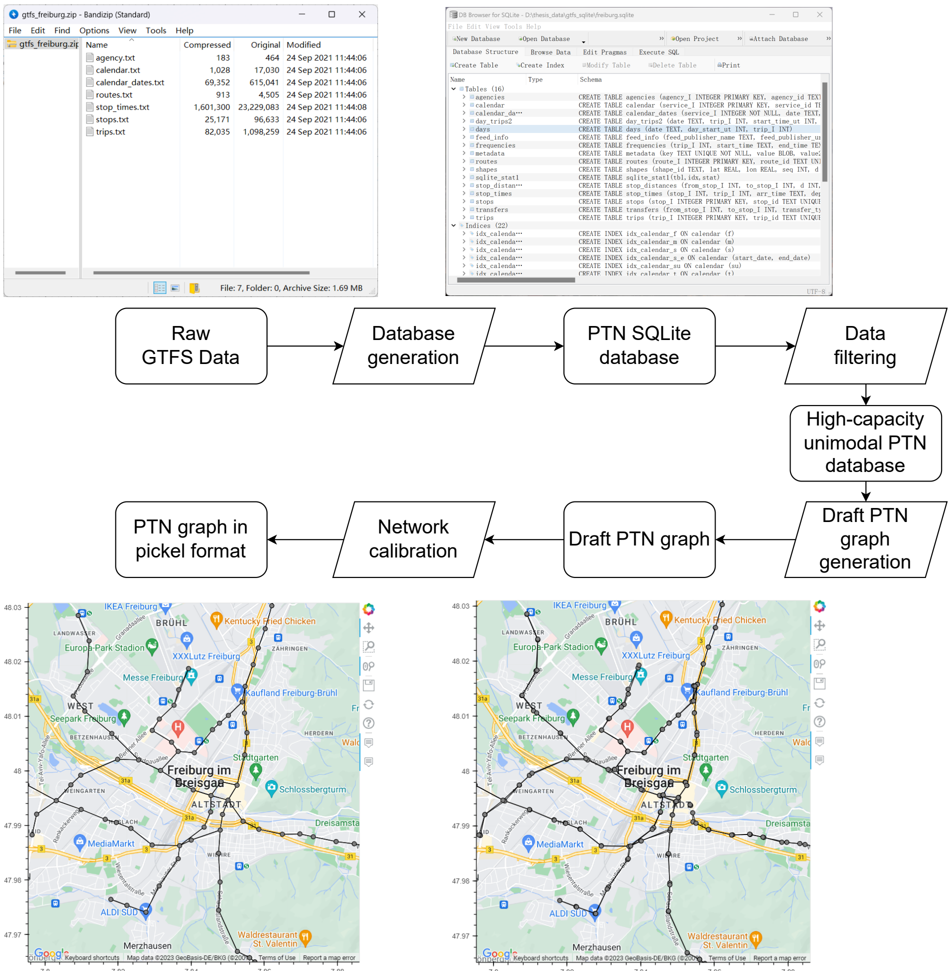


Figure 4.3: Flowchart of the GTFS data processing

The next step is filtering the data of high-capacity unimodal PTN from the database. The raw GTFS data usually includes the transit information of all the modes in the city over a period of time. To select data of the main high-capacity unimodal PTN from the whole database, the filtering is based on the mode ID of the high-capacity modes, such as metro and tram. However, BRT doesn't have the exclusive mode ID for filtering, which needs additional filtering after filtering the data of the bus. The additional filtering is based on the route ID or the operator ID of BRT. After the mode filtering and getting the data of high-capacity unimodal PTN, the date filtering is needed to reduce the volume of data and the calculation time of the following steps. The date filtering selects the representative date, which is closest to the data release date and contains over 90% of the maximum single-day trips. The operating hours of the representative day are from 5 AM till 12 AM of the next day, namely 19 hours, excluding the night public transport services.

The following step uses the filtered data and the GTFSPY library to generate graph representations of PTNs. The generated PTN graphs include topological and operating information, such as stops' coordinates, the trip arrival and departure times of stops, and route directions. Supported by the python library Bokeh (Bokeh Development Team, 2018) and Google maps API (Google, 2022b), the graphs can be visualised and interactive with maps. The stops in graphs are located based on their coordinates as vertices. The stops are connected by directed edges when trips happen between two stops. The directed edges are labelled with two types of weight information (see Section 3.2), one of which can be selected in the calculation of each topological indicator.

Although the draft networks have the essential vertices and edges, there are still flaws in the networks that need to be calibrated. For example, the unrecognised duplicated stops or transfer stops. These flaws are calibrated in two rounds. The first round automatically merges the identified duplicated stops by setting thresholds of distances and the similarity of stop IDs. In the second round, the thresholds of distances and ID similarity are reduced, and the merge recommender algorithm suggests the possible remaining duplicated stops. With manual checking, the confirmed duplicated stops are merged, and the denied duplicated stops are maintained. Figure 4.4 offers an example of the merge recommender during the second round calibration of the Karlsruhe tram network. Apart from the two rounds of calibration, the manual merge of stops is also possible.

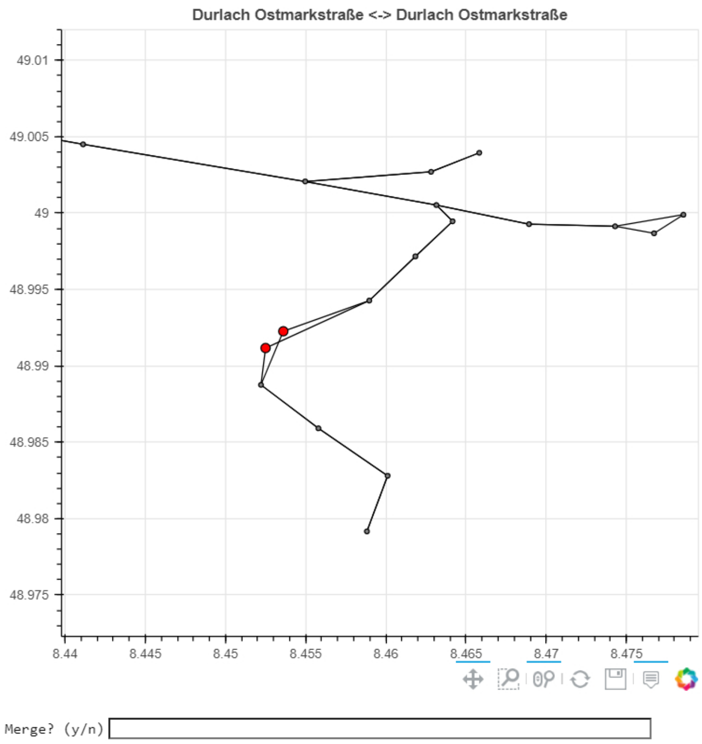


Figure 4.4: An example of a prompt from the merge recommender algorithm

Another type of flaw is the missing information on the edges, which usually happens in BRT networks. In some BRT networks, the stopping pattern is not strictly following the operating map, and the BRT can skip some stops with no demand. The stop-skipping results in missing information on the edges between two stops. For this kind of situation, the skipped stops are merged with the closest upstream stops to them. The coordinate of the merged vertex is the upstream vertex's coordinate.

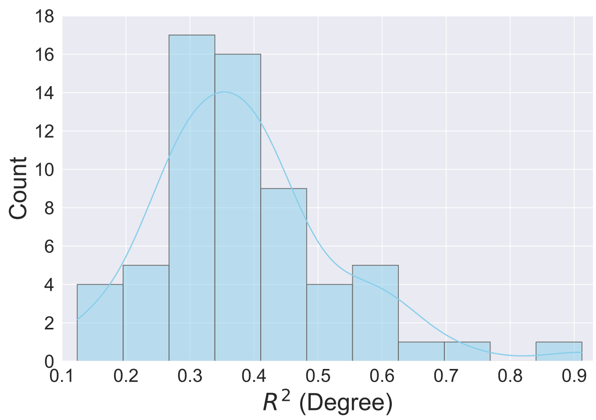
Through the four steps of data processing, the high-capacity unimodal PTNs are generated and saved as Python graph pickle format files. Networks saved in this format can be topologically analysed by Python packages iGraph (Csardi and Nepusz, 2006).

4.2. PTN hierarchy in dimensions of topological indicators

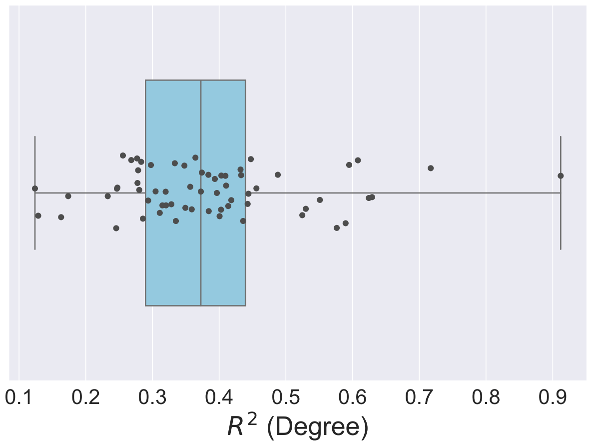
In the section, the 63 high-capacity unimodal PTNs worldwide are compared and quantified in terms of hierarchy with the selected topological indicators. The similarities and differences between the PTN hierarchy in the six dimensions of topological indicators are discussed.

4.2.1. Vertex degree dimension

Vertex degree denotes the number of connections of a vertex. The R square values in the vertex degree dimension quantify the degrees of the PTNs' scale-free network structure and denote the PTN hierarchy. The histogram and scattered box plot presented in Figure 4.5 shows the distribution of the R square values of the high-capacity PTNs in the database.



(a) Histogram

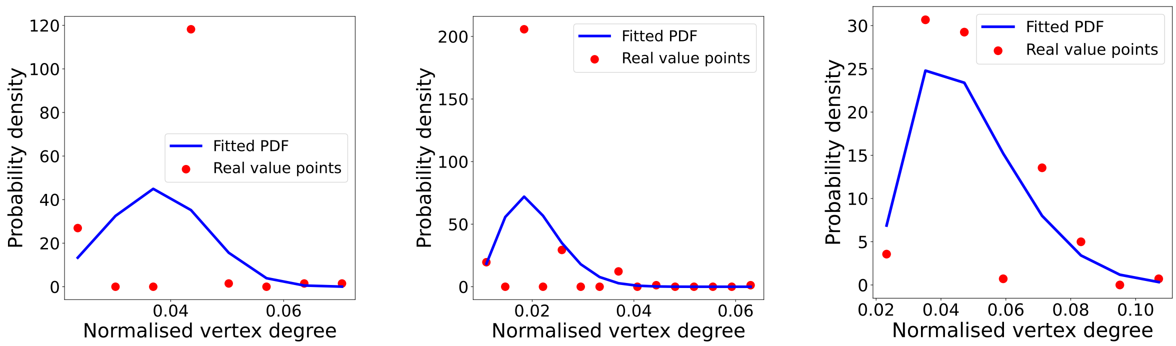


(b) Scattered box plot

Figure 4.5: Distribution of R square values in the vertex degree dimension

According to the plots and statistics, the range of the R square in the vertex degree dimension is from 0.124 to 0.912. The majority of the R square values are located in the range from 0.25 to 0.45. There are 66.7% of PTNs in the database distributed over a quarter of the value range. The mean value and median values are 0.386 and 0.373. The R square values of PTNs show a right-skewed distribution, and more PTNs tend to have a relatively low R square value. The R square value distribution of the 64 PTNs shows that scale-free structures are not common for PTNs. More PTNs in the database tend to have insignificant scale-free structures.

The reasons that result in the right-skewed distributed R square values are the majority of low-degree vertices and the lack of high-degree vertices in most PTNs. These reasons make the heterogeneity of low-degree vertices low, and the numbers of vertices in the histogram do not gradually decrease when the degree increases. As examples, PDFs with the fitted skewed normal distribution curves of three PTNs are shown in Figure 4.6. They are the BRT network in Kansas (US), the tram network in Den Haag (Netherlands) and the metro network in Santiago (Chile). The three PTNs are representative of the PTNs with low, medium and high R square values in the case study database, whose R square values are 0.129, 0.364 and 0.717, respectively.



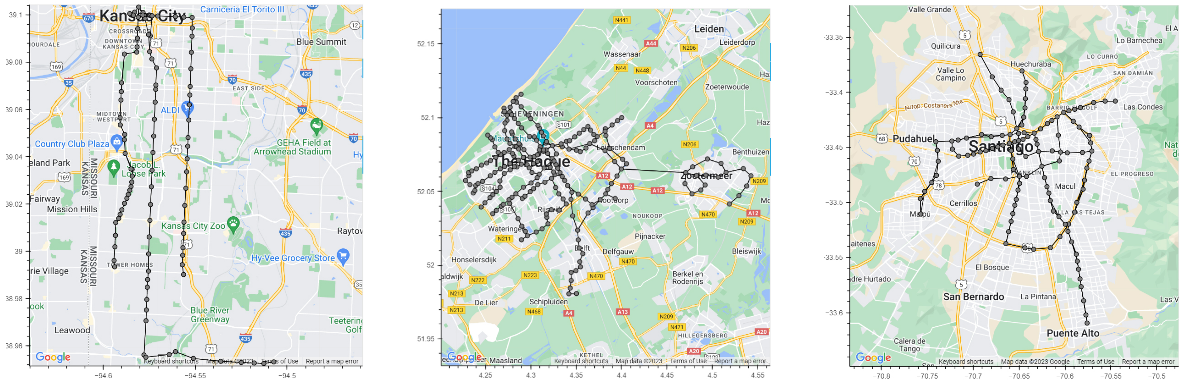
(a) BRT, Kansas (US) (b) Tram, Den Haag (Netherlands) (c) Metro, Santiago (Chile)

Figure 4.6: Vertex degree PDFs with fitted skewed normal distribution curves

In the three plots shown in Figure 4.6, the probability densities of the second lowest vertex degrees are the largest. The differences between the three PDFs are the largest probability density and the highest degree. In the plot of the Kansas BRT network, the vertices with low and medium degrees are the vast majority. In contrast, the vertices with high degrees are few, and thus the curve's goodness of fit is poor. In the plot of the Den Haag tram network, although the majority of the vertices have low degrees, there are a few vertices having medium and high degrees, the decrease of the probability density is smoother than in the plot of the Kansas BRT, and the curve has better goodness of fit. As for the Santiago metro network, the proportion of the low-degree vertices is lower than in the other two PTNs. Besides, there are more vertices with medium and high degrees, and the probability densities of vertex degrees are gradually decreased. The curve's goodness of fit is high in the Santiago metro network.

Through the analysis of the maps of PTNs, the majority of low-degree vertices and the lack of high-degree vertices are related to the large proportion of linear network segments and the lack of the high-degree transfer stop in the PTNs. The maps of the three PTNs samples are provided in Figure 4.7 for illustration.

In maps, the proportions of vertices on the linear network segments are different. The vertices on the linear network segments have two adjacent connected vertices, embodied as the vertices in the second lowest histogram bins, whose degrees are higher than the dead ends of networks. In the maps of the Kansas BRT network and the Den Haag tram network, most of the stops are located on the linear network segments connecting the city centre and the suburban areas. As for the map of the Santiago metro network, the stops located on linear network segments are fewer. The stops in the Santiago metro network are not only sequentially connected but can also connect with non-adjacent stops, reducing the number and length of linear network segments. Another difference is the number of stops having high vertex degrees. In the Kansas BRT network, only a small number of stops have high degrees in the central areas. The Den Haag tram network has more high-degree vertices, and a few transfer stops for multiple lines have higher degrees. In the Santiago metro network, the connections between non-adjacent vertices increase the connections of stops, having more high-degree stops.



(a) BRT, Kansas (US) (b) Tram, Den Haag (Netherlands) (c) Metro, Santiago (Chile)

Figure 4.7: Maps of example PTNs for comparison in the vertex degree dimension

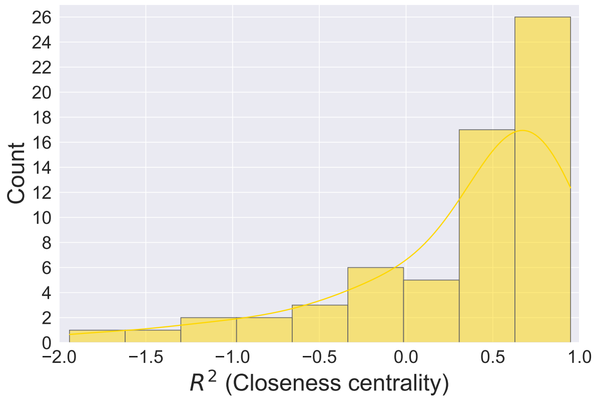
4.2.2. Closeness centrality dimension

The closeness centrality of vertices denotes the vertices' accessibility. In the closeness centrality dimension, the R square values quantify the degree of vertices following the hierarchical organisation by accessibility and indicate the PTN hierarchy. For the investigation of the R square value distribution, the histogram for positive R square values, and the scattered box plot are presented as Figure 4.8.

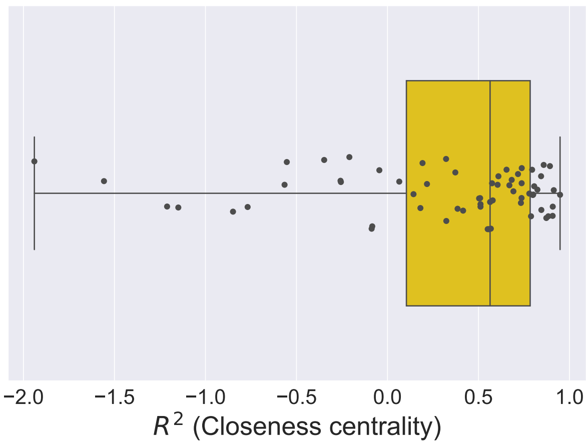
According to the plots and the statistics, the R square values in the closeness centrality dimension have a wide distribution range, from -1.940 to 0.948. The median and average R square values for the database are 0.564 and 0.321, and the distribution shows a significant left-skewed distribution, with more PTNs prone to have a high hierarchy. Among the PTNs in the case study database, 15 PTNs (23.81%) have negative R square values and low hierarchy, taking up nearly one-quarter of PTNs in the database. Over half of the studied PTNs have R square values spread over the range from 0.5 to 0.95. The numbers of PTNs increase roughly as the R square values of histogram bins increase. According to the R square value distribution, more PTNs tend to have a hierarchical organisation of stops in terms of their accessibility.

The reasons that result in the difference in R square values of PTNs are the number of vertices with different closeness centrality and the heterogeneity of vertices' closeness centrality. In a PTN having similar numbers of vertices with different closeness centrality, the number of bins in the PDF of closeness centrality is small and the shape of the PDF is flat. The flat PDF cannot meet the gradually decreased number of vertices when the closeness centrality increases, thus the skewed normal distribution curve's goodness of fit is not good and the R square value is low. As examples, in Figure 4.9, three closeness centrality PDFs with the fitted skewed normal distribution curves of PTNs are shown: the Philadelphia metro network in the US, the Nuremberg metro network in Germany and the Rotterdam tram network. The three PTNs represent the low, medium and high values of R square values in the database, whose R square values are -1.940, 0.564 and 0.948, respectively.

In the three PDFs of vertices' closeness centrality, the goodness of fit of the fitted skewed normal distribution curves increases from the Philadelphia metro network to the Rotterdam tram network. The PDF of the Philadelphia metro network contains four bins of closeness centrality values. The probability densities of the bins with high and low closeness centrality values are similar and do not gradually decrease. The fitted skewed normal distribution curve cannot fit the flat PDF, having worse goodness of fit than the horizontal average value

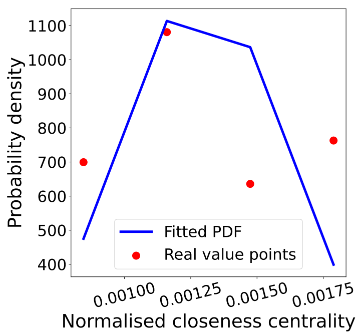


(a) Histogram

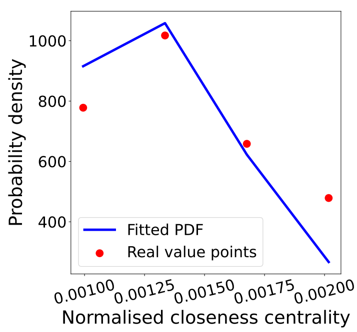


(b) Scattered box plot

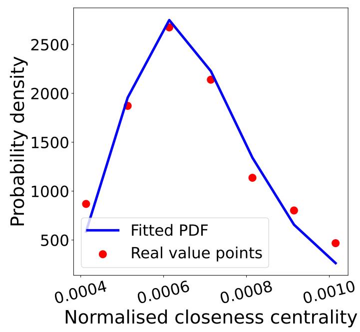
Figure 4.8: Distribution of R square values in the closeness centrality dimension



(a) Metro, Philadelphia (US)



(b) Metro, Nuremberg (Germany)



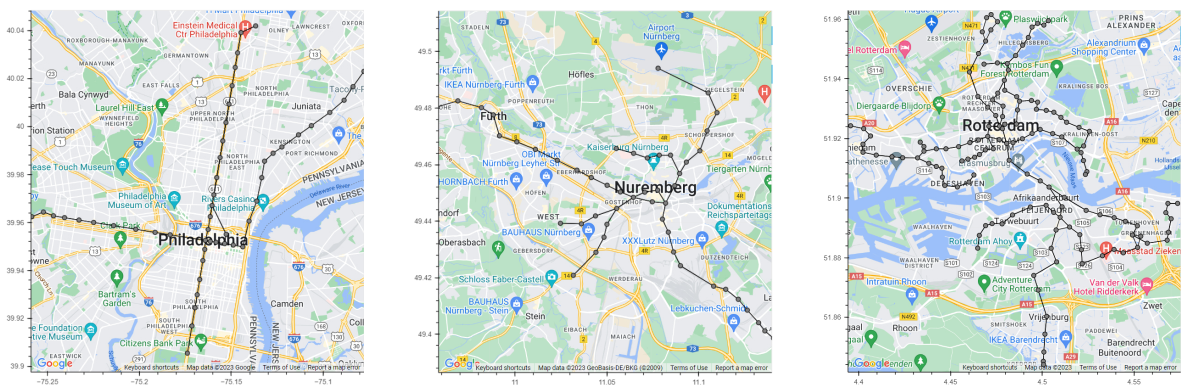
(c) Tram, Rotterdam (Netherlands)

Figure 4.9: Closeness centrality PDFs with fitted skewed normal distribution curves

line of the probability densities. Thus, the R square value of the Philadelphia metro network is negative, showing no PTN hierarchy. In the PDF of the Nuremberg metro network, the heterogeneity of the probability densities of closeness centrality values is not high enough.

In addition, in the PDF of the Rotterdam tram network, there are seven closeness centrality values, and the differences between the probability densities of closeness centrality values are apparent. More vertices in the PTN have low and medium closeness centrality, and the number of high closeness centrality vertices gradually decreases with the indicator increase. Thus, the goodness of fit of the skewed normal distribution curve is high, and the Rotterdam tram network has a high R square value and a high PTN hierarchy.

By analysing the PTN element organisations in the maps of PTNs, the heterogeneity of vertices' closeness centrality is related to the number of high-degree transfer stops and the proportion of vertices on the linear network segments starting from the dead ends of the network. In Figure 4.10, the maps of the three sample PTNs are provided for illustration.



(a) Metro, Philadelphia (US) (b) Metro, Nuremberg (Germany) (c) Tram, Rotterdam (Netherlands)

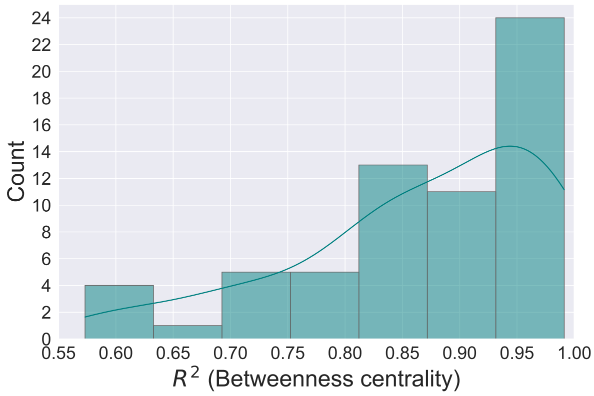
Figure 4.10: Maps of example PTNs for comparison in the closeness centrality dimension

The vertices on the linear network segments starting from the dead ends have the lowest accessibility in the network. The vertices on the same linear network segment have similar closeness centrality. The closeness centrality noticeably changes at the high-degree ends of the linear network segments (see Section 3.3.3). Therefore, more high-degree transfer stops and a modest proportion of vertices on the dead-end started linear network segments result in high heterogeneity of vertices' closeness centrality. In the metro network of Philadelphia, most of the vertices are located on the dead-end-started linear network segments. The metro lines only cross three times in the central areas. Besides, the Nuremberg metro network has a lower proportion of vertices on the dead-end started linear network segments than Philadelphia's. The metro lines have more crossings in the central areas. As for the tram network in Rotterdam, the vertices on the dead-end-started linear network segments are lower than the other two PTNs, and the tram lines have many crossings in the network. The R square values of the three PTNs compliance the analysis of the PTN element organisations.

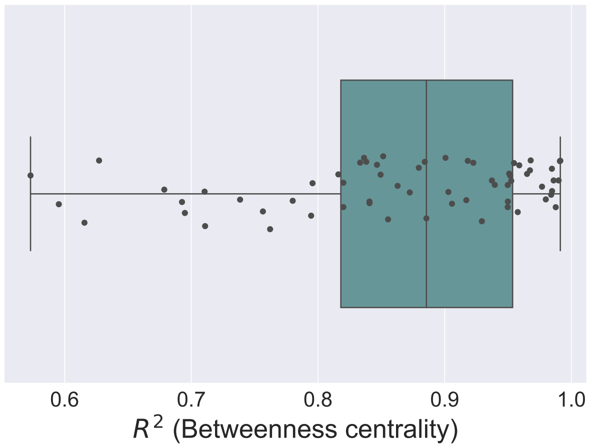
4.2.3. Betweenness centrality dimension

The betweenness centrality of a vertex denotes its intermediacy and can reflect the intermediacy of the linear network segment it locates. In the betweenness centrality dimension, the R square values indicate the degree of the hierarchical organisation of vertices or linear network segments by their intermediacy. In Figure 4.11, the histogram and box plot of the R square values of the high-capacity unimodal PTNs in the case study database is presented for distribution analysis.

According to the plots and statistics, the range of the R square values is from 0.57 to 0.99. The median and average R square values of PTNs are 0.886 and 0.866. The R square values



(a) Histogram



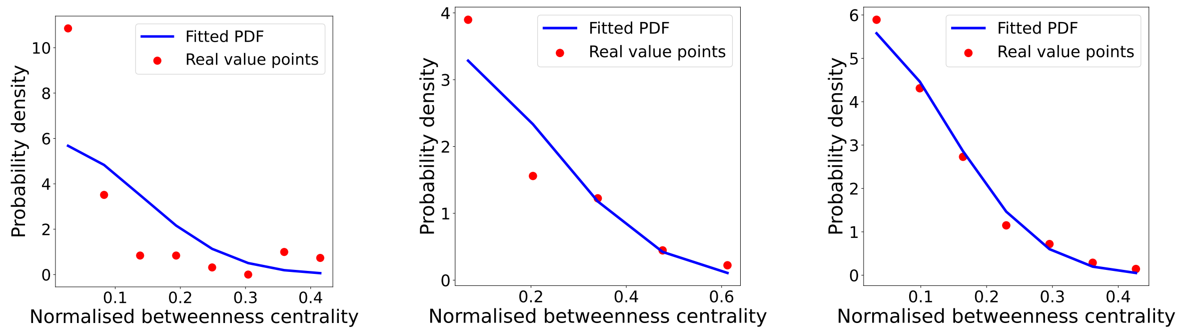
(b) Scattered box plot

Figure 4.11: Distribution of R square values in the betweenness centrality dimension

show a left-skewed distribution, indicating the majority of PTNs are prone to have a high R square value in the betweenness centrality dimension. The density of scatters in the box plot witnesses an increasing trend with the R square values increase. There are 48 PTNs having R square values over 0.8, taking up 76.2%, over three-quarters of the PTNs in the database. In a word, more assessed PTNs tend to have a relatively high hierarchy in the betweenness centrality dimension, indicating the network elements follow a hierarchical organisation in terms of traffic load intermediacy in these PTNs.

The reason resulting in the R square value differences is the probability gaps between betweenness centrality values. For instance, the three sample PTNs have relatively low, medium and high R squares in the betweenness centrality dimension: the Berlin tram network (0.573), the Dallas tram network (0.884) and the Milan metro network (0.991). The PDFs with the fitted skewed normal distribution curves in Figure 4.12 can show the reason that makes their R square values different.

In the three PDFs of different R square values, the shapes of the PDFs are similar. The probability densities of the leftmost betweenness centrality are the highest, and the probability densities decrease with the betweenness centrality increases. The differences between the PDFs are the probability density gaps between the adjacent points in each PDF. In the PDF of

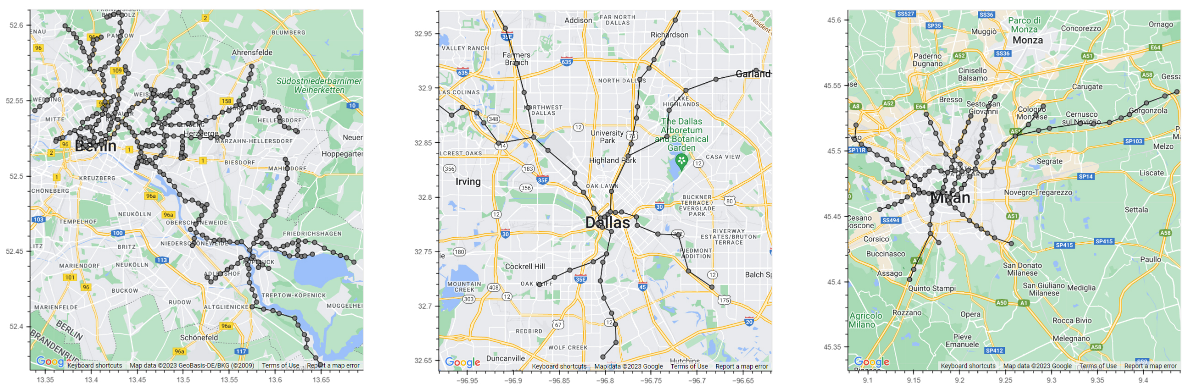


(a) Tram, Berlin (Germany) (b) Tram, Dallas (US) (c) Metro, Milan(Italy)

Figure 4.12: Betweenness centrality PDFs with fitted skewed normal distribution curves

the Berlin tram network, the probability density gaps are not smoothly changed. The excessive leftmost probability density result in a big gap to points on the right. Besides, the probability density of the second rightmost betweenness centrality has a slight increase, breaking the trend of decreasing with the betweenness centrality increase. Thus, the curve's goodness of fit in the Berlin tram network is relatively low. The Dallas tram network has a better goodness of fit for the skewed normal distribution curve than Berlin. However, the probability density of the second leftmost point is not large enough, having a large gap to the leftmost point, and a small gap to the point on the right. The curve's goodness of fit is high but lower than the Milan metro network. In the PDF of the Milan metro network, the probability densities are gradually decreased, the probability density gaps between points are smoothly decreased, and the goodness of fit is high.

In the maps of PTNs, the reasons that trigger the probability density gaps between adjacent points in PDFs are mainly embodied in the large numbers of dead ends or the long linear segments in the network. The maps of the three sample PTNs are presented in Figure 4.13 as examples.



(a) Tram, Berlin (Germany) (b) Tram, Dallas (US) (c) Metro, Milan(Italy)

Figure 4.13: Maps of example PTNs for comparison in the betweenness centrality dimension

In Section 3.3.4, the betweenness centrality change patterns on linear network segments are discussed. The dead ends of a PTN are not part of any shortest path between other pairs of vertices in the network, and their betweenness centrality is 0, having the lowest betweenness centrality in the PTN. So if a PTN has a large number of dead ends in the network, the number of vertices with low betweenness centrality will be large and result in the high leftmost bins in the PDF. Besides, the betweenness centrality values of vertices on the same linear segment are similar and slightly increase when closer to the endpoints with more connections. The long linear network segments make the probability density of betweenness centrality similar, and cannot smoothly decrease in the PDF. In the map of the Berlin tram network, there are a large number of dead ends, so the number of vertices with low betweenness centrality is large, resulting in the leftmost bins in the PDF being high. In addition, the number of vertices on the linear network segments connecting the central areas with the southeast Köpenick area is large, resulting in small gaps between the probability density of the medium and high betweenness centrality and relatively medium hierarchy. By contrast, the maps of the Dallas tram network and the Milan metro network do not have many dead-ends and long linear network segments, and their PTN hierarchy is relatively high.

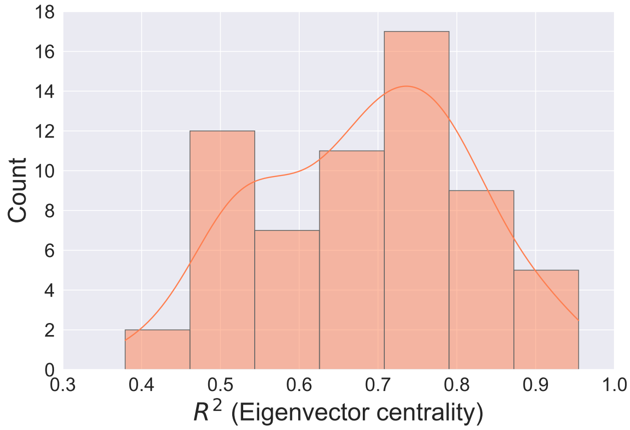
4.2.4. Eigenvector centrality dimension

The eigenvector centrality denotes the importance of a vertex by the importance of its connecting vertices and the importance of the vertex cluster consisting of the vertex and the connecting vertices. The R square values in the eigenvector centrality dimension quantify the degree of hierarchical organisation of vertex clusters in a PTN by their importance. To investigate the distribution of the R square values of the PTNs in the case study database, the histogram and box plots are presented as Figure 4.14.

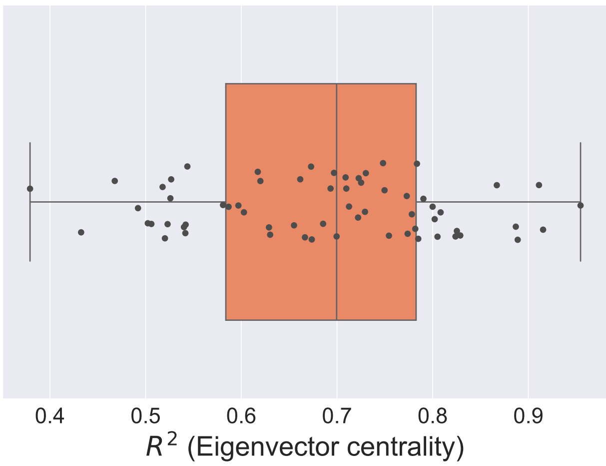
The range of R square values is from 0.379 to 0.955. The median and average values are 0.700 and 0.685. The R square values show a slightly left-skewed distribution. The majority of the PTNs have R square values from 0.5 to 0.8, concentrated in the middle right of the value range. The distribution of R square values shows around half of PTNs in the database tend to show a hierarchical organisation of vertex clusters based on their importance and show mono-centric network structures.

The reasons that result in the differences in R square values in the eigenvector centrality dimension are the number of bins in the histograms and the proportion of vertices having low eigenvector centrality. As examples for interpretation, the PDFs with the fitted skewed normal distribution curves of three PTNs with low, medium and high R square values of the database are presented in Figure 4.15. The three PTNs are the tram network of Melbourne, Australia (0.379), the BRT network of Seattle, the USA (0.700) and the metro network of Lyon (0.955).

The method used to determine the bin width for the PDF is Scott's rule (see Section 3.3.1), which is influenced by the standard deviation and the size of the data. If the number of vertices in the PTN is large and the eigenvector centrality values of vertices have a small standard deviation, the bin width of the PDF would be small, and the number of bins in the PDF would be large. Therefore, the number of bins in the PDF denotes the standard deviation of data, and the heterogeneity of the eigenvector centrality. If the number of bins in the PDF is very large, the heterogeneity of eigenvector centrality in the PTN is low, resulting in a low R square value. In the three PDFs, the number of bins for the eigenvector centrality values is different, in which the number of bins in the PDF of the Melbourne tram network is significantly higher than the other two PTNs. In the PDF of the Melbourne tram network, the number of bins is large, and the probability density of most points is low, but the point with the lowest eigenvector centrality value has a high probability density. In the Melbourne tram network, the number of high-degree vertices is large and their distribution is wide, resulting in the eigenvector centrality of vertices being various, but no vertex shows significant importance. Thus, for the vertex clusters consisting of a vertex and all its connecting vertices, most of their importance is low and lacks heterogeneity. In contrast, in the PTNs with a few and concentrated high-degree

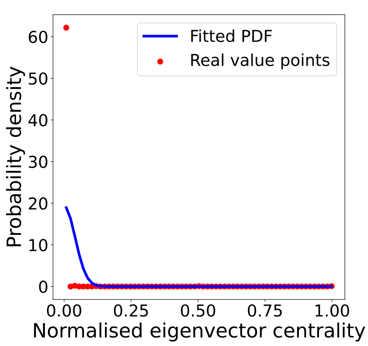


(a) Histogram

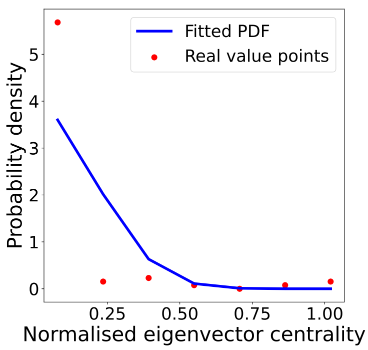


(b) Scattered box plot

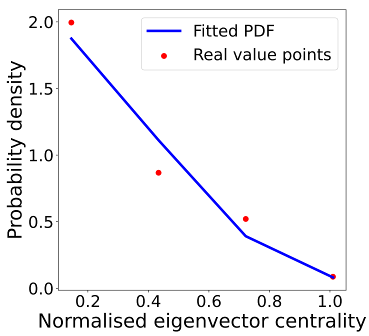
Figure 4.14: Distribution of R square values in the eigenvector centrality dimension



(a) Tram, Melbourne (Australia)



(b) BRT, Seattle (US)

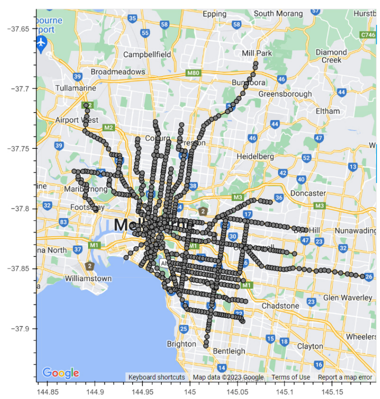


(c) Metro, Lyon (France)

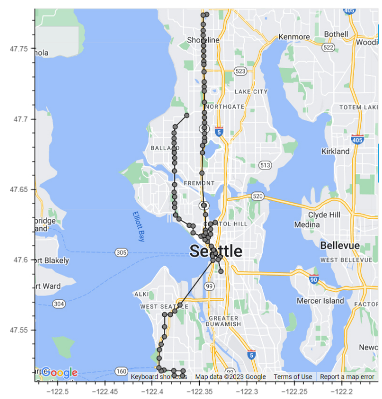
Figure 4.15: Eigenvector centrality PDFs with fitted skewed normal distribution curves

vertices, the number of bins for eigenvector centrality values is small. The high-degree vertices and their connecting vertices have high importance. For example, in the Lyon metro network, the high-degree vertices are concentrated and form a ring. The vertices on the ring can be adjacent to multiple high-degree vertices, having significant high eigenvector centrality values. Another difference between the PDFs is the number of vertices with low eigenvector centrality. In the PDFs of the Melbourne tram network and Seattle BRT network, the proportions of the vertices in the leftmost point of PDFs are high, and the vertices with medium and high eigenvector centrality are few. The excessive number of vertices with low eigenvector centrality makes the probability density of the leftmost bin high and results in low goodness of fit of the skewed normal distribution curves. Compared to the other two PTNs, the PDF of the Lyon metro network have a small number of bins, a modest number of vertices with low eigenvector centrality, and vertices with medium and high eigenvector centrality. The heterogeneity of vertices' eigenvector centrality is high, the curve's high goodness of fit and thus the high R square value.

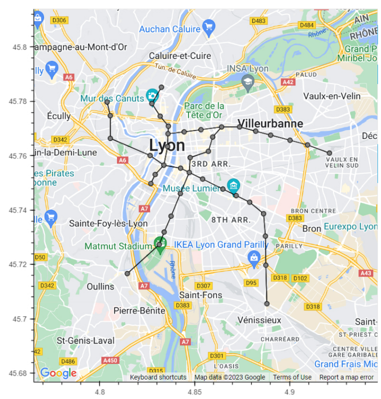
Reflected on the maps of PTNs, the high R square values in the eigenvector centrality dimension are related to the vertex organisation having a small number of concentrated high-degree vertices. The maps and zoomed-in maps of the three sample PTNs are provided in Figure 4.16 for illustration.



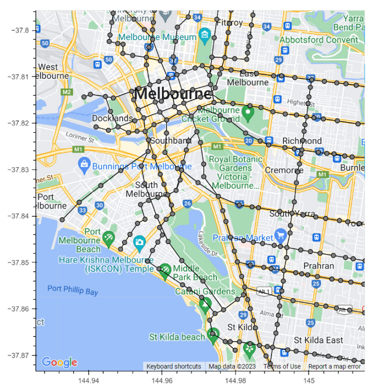
(a) Tram, Melbourne (Australia)



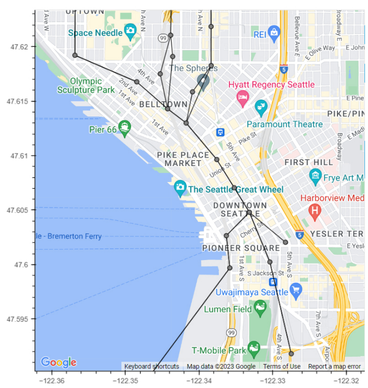
(b) BRT, Seattle (US)



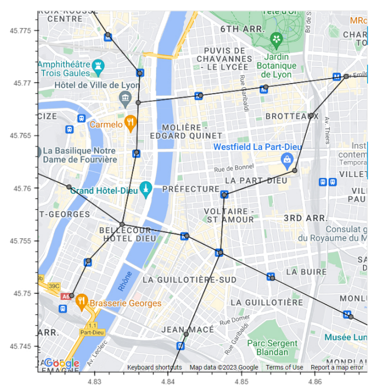
(c) Metro, Lyon (France)



(d) Tram, Melbourne (Zoomed in)



(e) BRT, Seattle (Zoomed in)



(f) Metro, Lyon (Zoomed in)

Figure 4.16: Maps of example PTNs for comparison in the eigenvector centrality dimension

In the maps of the three PTNs, the numbers of the high-degree transfer stop decrease, and the distributions are more concentrated with the R square values increase. In the map of the Melbourne tram network, the tram lines have multiple crossings in central areas and suburban areas. The high-degree transfer stops are distributed over a wide area of the whole network. In the BRT network in Seattle, the BRT lines extend in three directions, and the crossings of lines form several high-degree transfer stops in the city's central area. As for the Lyon metro network, the high-degree transfer stops form a ring and connect each other by a small number of vertices, which have a concentrated distribution in the city's central area, showing a mono-centric network structure. The stops connecting with the high-degree stops have high eigenvector centrality. Thus, a small number and concentrated high-degree vertices are assigned with higher eigenvector centrality values for their higher importance than other vertices. By contrast, the large number and widely distributed high-degree vertices make the vertex clusters centred with these high-degree vertices have similar importance, and the heterogeneity is reduced. The importance of the vertex clusters centred on these important vertices is significantly higher than other vertex clusters, forming the heterogeneity in the vertex clusters' importance and further resulting in high R square values.

4.2.5. Network modularity dimension

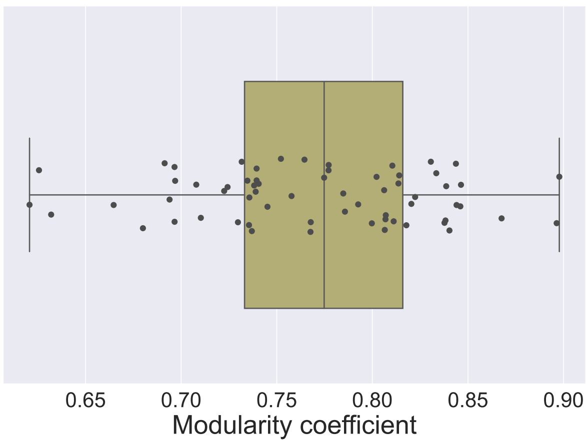
The network optimal modularity coefficient quantifies the degree of a PTN having a high-clustering structure, where the network has a higher potential to be divided into vertex clusters with high intra-similarity and low inter-similarity. The high optimal modularity coefficient of a PTN indicates the high degree of the PTN's structure being high-clustering. Figure 4.17 presents the histogram and the box plot of the PTNs' optimal modularity coefficients.

In the histogram of the optimal modularity coefficients, the distribution range is from 0.621 to 0.898. The median and average optimal modularity coefficients are 0.775 and 0.770, indicating a slightly left-skewed normal distribution. Nearly half of the PTNs in the database tend to have a relatively high hierarchy. In the box plot, most PTNs have optimal modularity values between 0.70 to 0.85, taking up around 79.4% of the PTNs in the case study. The distribution of optimal modularity coefficients shows that over half of the PTNs in the database show a relatively high hierarchy, indicating a high-clustering network structure and presenting multi-centric network structures.

According to the definition of the modularity coefficient, the optimal modularity coefficient is reached when the connections within vertex clusters are more than the connections with vertices outside the clusters. In other words, these clusters are formed around the vertices with the local most connections, i.e. the high-degree vertices. The number of high-degree vertices in a PTN affects the potential of forming the high intra-similarity and inter-difference clusters and influences its optimal modularity coefficient. As examples, the vertex degree histogram and the visualised vertex clusters with optimal modularity coefficients of three PTNs are respectively presented in Figure 4.18 and Figure 4.19. The three sample PTNs are the metro network of Marseilles, the metro network of Vienna and the tram network in Melbourne, Australia, having optimal modularity coefficients equal to 0.621 (low), 0.777 (medium) and 0.898 (high).

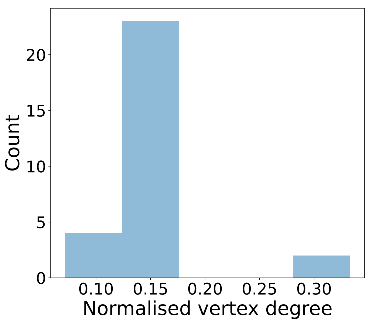
In the vertex degree histograms of the three PTNs, the Marseilles metro network has the minimum high-degree vertices among the three PTNs, with only two high-degree vertices. The number of high-degree vertices in the Vienna metro network is more than in Marseilles, having ten high-degree vertices. As for the Melbourne tram network, there are seventy-six high-degree vertices in the network. Since the optimal modularity coefficient is related to the local high-degree vertices, the more and wider distributed high-degree vertices result in a higher potential of forming clusters with high intra-similarity and inter-difference. The visualised vertex clusters with optimal modularity coefficients can also reflect this relation between the numbers

(a) Histogram

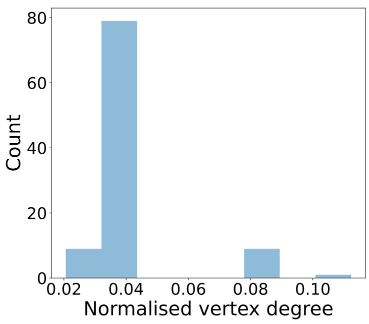


(b) Scattered box plot

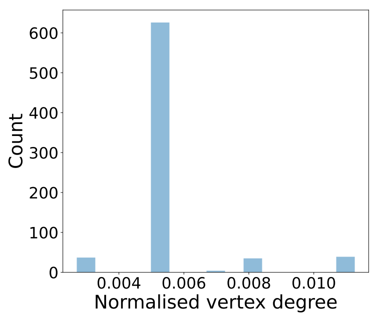
Figure 4.17: Distribution of optimal modularity coefficients in the network modularity dimension



(a) Metro, Marseilles (France)



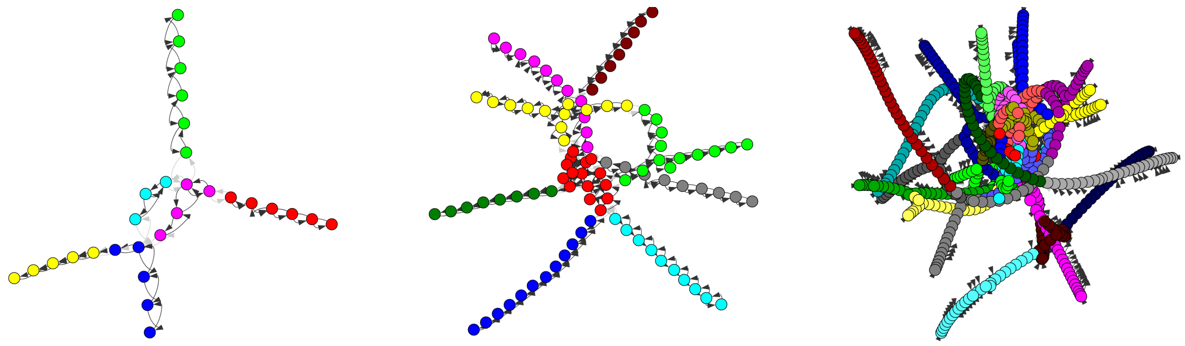
(b) Metro, Vienna (Austria)



(c) Tram, Melbourne (Australia)

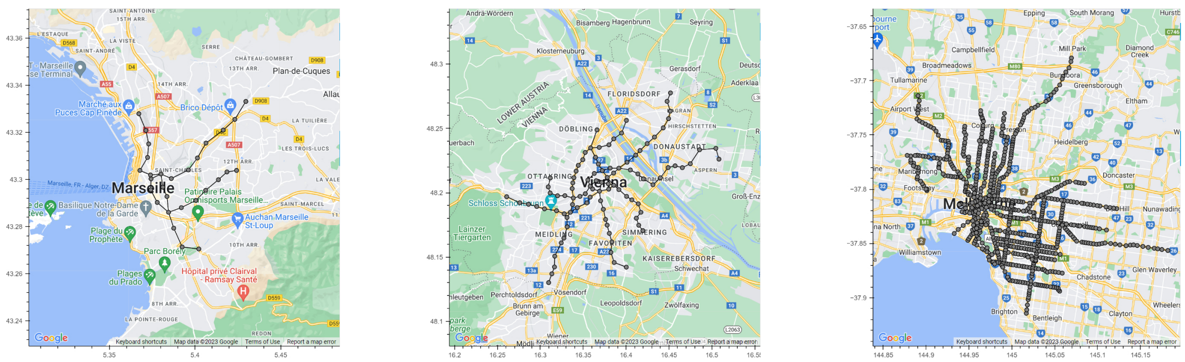
Figure 4.18: Degree histograms of the example PTNs for the comparison of optimal modularity coefficients

of high-degree vertices with optimal modularity coefficients. In Figure 4.19, the vertices with the same colours are in the same clusters and the arrows indicate the connections between vertices. It can be seen that the vertices around high-degree vertices are grouped in the same clusters. The more high-degree vertices in the PTNs, the number of vertex clusters formed are higher.



(a) Metro, Marseilles (France) (b) Metro, Vienna (Austria) (c) Tram, Melbourne (Australia)

Figure 4.19: Visualised clusters of the example PTNs with optimal modularity coefficients



(a) Metro, Marseilles (France) (b) Metro, Vienna (Austria) (c) Tram, Melbourne (Australia)

(d) Metro, Marseilles (Zoomed-in) (e) Metro, Vienna (Zoomed-in) (f) Tram, Melbourne (Zoomed-in)

Figure 4.20: Maps of example PTNs for comparison in the network modularity dimension

The differences in optimal modularity coefficients can be reflected in the numbers of transfer stops and their distribution ranges in maps. In a PTN with a large number of wide-distributed high-degree vertices, there is no clear centre of the network, and the PTN shows a multi-centric structure. The maps of the three sample PTNs are presented in Figure 4.20 for illustration. In the map and zoomed-in map of the Marseilles metro network, the two high-degree vertices and the vertices in between formed a clear centre of the network, and the metro lines are extended from the centre. In the maps of the Vienna metro network, the high-vertices stops are distributed in a wider range of the network. The metro lines have more crossings in the centre areas of the city, but the stops of the network are not concentrated in a small area and the centre of the network is unclear. As for the Melbourne tram network, the number of high-degree stops is large. The high-degree stops are distributed over the network, not only in the downtown area of Melbourne but also on the north and east sides of the city. There is no clear centre of the network, and the structure of the tram network is multi-centric.

4.2.6. Network assortativity dimension

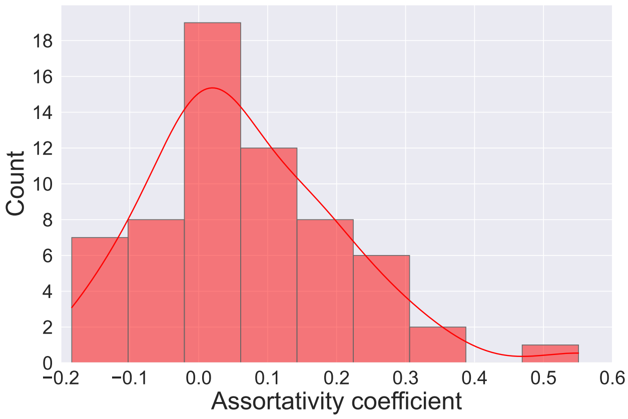
The assortativity coefficient quantifies the extent of the vertex connections in a PTN following a unified pattern, with which high-degree vertices connect high-degree vertices, and low-degree vertices connect low-degree vertices. The closer an assortativity coefficient to 1, the higher extent of vertex connections in the PTN following the pattern, indicating higher PTN hierarchy in the dimension. Figure 4.21 presents the histogram and the box plot of assortativity coefficients of the high-capacity unimodal PTNs in the case study database.

The values of the assortativity coefficients range from -0.184 to 0.551. One-third of the PTNs in the database have negative assortativity coefficients. The median and average assortativity coefficients are 0.041 and 0.066, showing a right-skewed distribution, with more PTNs having relatively low assortativity coefficients. According to the box plot, most of the PTNs have coefficients from -0.1 to 0.2, taking up 74.6% of PTNs in the database. According to the distribution of the assortativity coefficients, more PTNs in the database tend to have a low hierarchy in the network assortativity, presenting insignificant vertex connection patterns by vertex degrees and core/periphery network shapes.

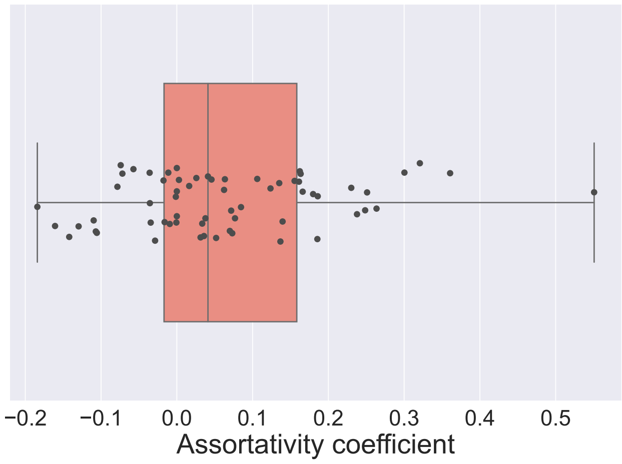
From the perspective of the organisation of PTN elements, the reasons resulting in the majority of PTNs tending to have a low hierarchy in the dimension are the lack of both high-degree and low-degree stops and the infeasible continuous connections between high-degree stops. It has also been found that public transport operating with different stop spacing on the same infrastructure can increase the assortativity coefficient. Three PTNs are presented as examples for interpreting the findings: the metro network of Vancouver (Canada), the tram network of Amsterdam (Netherlands) and the BRT network of Cleveland (US). The three PTNs have assortativity coefficients equal to -0.184, 0.070 and 0.551, respectively. The histograms of vertex degree are presented in Figure 4.22, and their maps and zoomed-in maps of the three PTNs are presented in Figure 4.23.

According to the vertex degree histograms of the three PTNs, either the high-degree or low-degree vertices take up a small proportion of the vertices in the network. As discussed in Section 4.2.1, the variety of the vertex degree in high-capacity unimodal PTNs is limited, and the high-degree stops with multiple connections are hard to achieve. By contrast, the majority of vertices are on the linear network segments with two adjacent connections, their degree neither low nor high. The connections between the vertices on the linear network segments have less contribution to the assortativity coefficient. For example, in the Vancouver metro network, only five high-degree transfer stops in the network.

Another reason for the low assortativity coefficient is the connection between low and low-degree vertices and high and high-degree vertices are infeasible. The low-degree vertices in PTNs are the dead ends of the network, with only one connected vertex. The connection between low-degree vertices will increase their degree. As for the connections between high-degree vertices, the connections between two transfer stops are unusual, and cannot form a

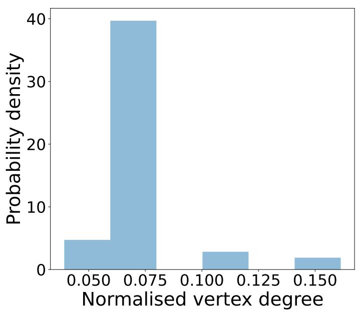


(a) Histogram

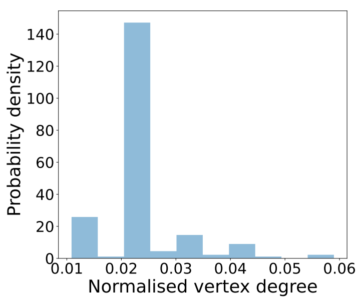


(b) Scattered box plot

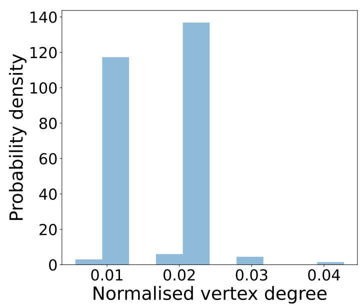
Figure 4.21: Distribution of assortativity coefficients in the network assortativity dimension



(a) Metro, Vancouver (Canada)



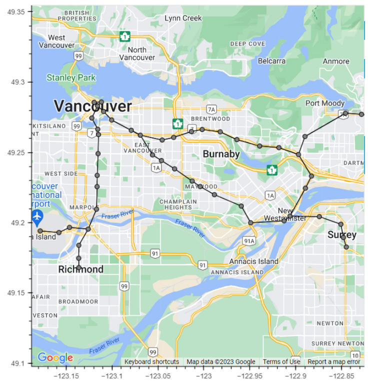
(b) Tram, Amsterdam (Netherlands)



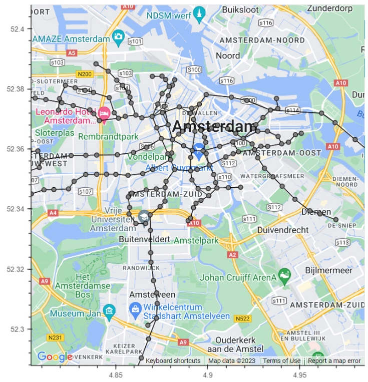
(c) BRT, Cleveland (US)

Figure 4.22: Degree histograms of the example PTNs for the comparison of assortativity coefficients

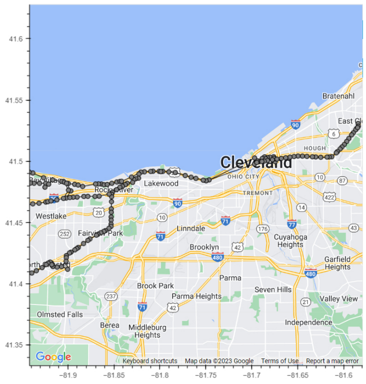
unified pattern, especially for the metro stops. Transfer stops of multiple lines need more land use for the infrastructures than normal stops or transfer stops for two lines (Chen et al., 2022a).



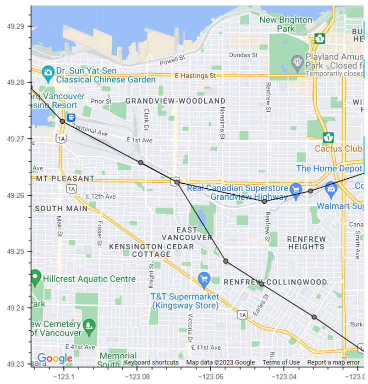
(a) Metro, Vancouver (Canada)



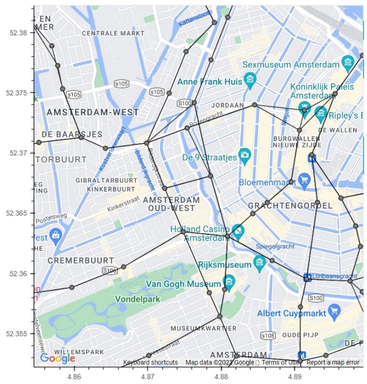
(b) Tram, Amsterdam (Netherlands)



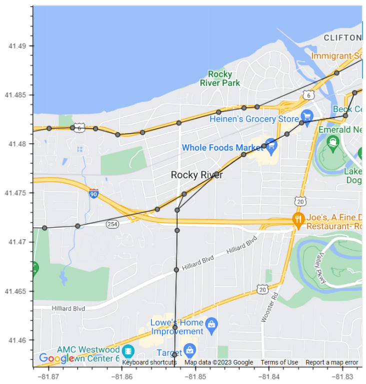
(c) BRT, Cleveland (US)



(d) Metro, Vancouver (Zoomed-in)



(e) Tram, Amsterdam (Zoomed-in)



(f) BRT, Cleveland (Zoomed-in)

Figure 4.23: Maps of example PTNs for comparison in the network assortativity dimension

The land use limits many direct connections between multiple-line transfer stops and cannot form a connection pattern. In the map of the Amsterdam tram network, the tram lines cross in the city's central areas. Some high-degree transfer stops are connected, which increases the assortativity coefficient value but cannot show a unified connection pattern.

Other than the reasons reducing the assortativity coefficient, it has been found that when public transport operates with different stop spacing on the same infrastructure, the assortativity coefficient increases. For example, in the Cleveland BRT network, many vertices are located on the linear network segments connecting the central areas with suburbs. However, the number of high-degree vertices is not small. In Figure 4.22(c), two bins of vertex degree have a high probability density, which is because of the flexible stop spacing during the operation. In the operation of some BRT networks, the stopping strategy is flexible. The drivers don't need to follow the map and stop at every stop strictly. Flexible changing the stop spacing is possible due to the passenger demand for getting on and off. This stopping strategy results in non-adjacent stops being connected, and increase their vertex degrees. Plus the original connections following the operating map, the connections between high-degree vertices increase. For instance, in Figure 4.23(f), the three stops around Rocky River have higher vertex degrees and connect each other. The organisation of vertices like this is common in the Cleveland BRT network, which makes the high-degree vertices have more connections with high-degree vertices and increases the assortativity coefficient. For the same reason, the lines with different stop spacing operating on the same infrastructure, like the Santiago metro network in Figure 4.7(c), also have positive effects on the assortativity coefficient value.

4.2.7. Discussion

The above six dimensions evaluate the organisation of the PTN elements from different perspectives to quantify and compare the hierarchy. The discussion about the similarities and differences between the normalised indicators of the six dimensions helps to have an effective and comprehensive utilisation of the methodology.

Figure 4.24 offers an overview comparison of the distribution of the normalised six indicators. To ensure the comparison are meaningful, PTN hierarchy indicators in all six dimensions are scaled to the range from 0 to 1, whose method is introduced in Section 3.5. The ranges of x-axes are set from 0 to 1, and the y-axes ranges are set from 0 to 28.

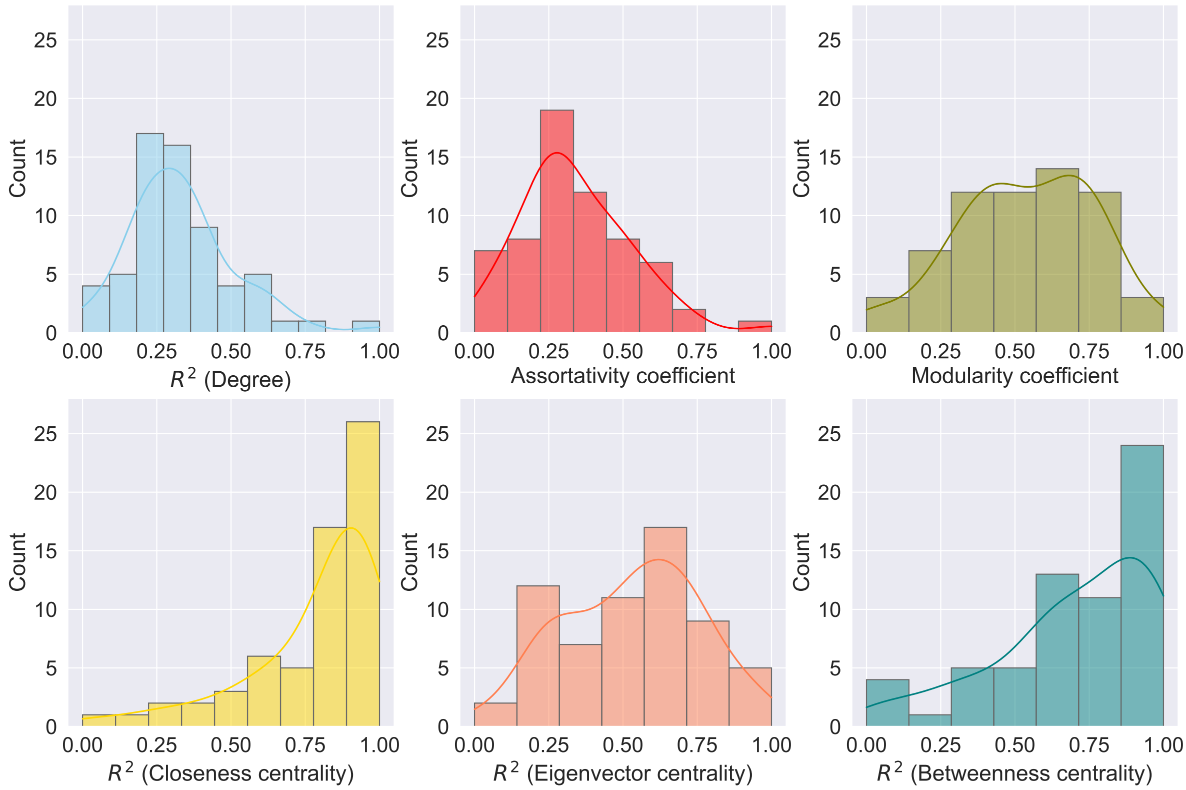


Figure 4.24: Histograms of normalised topological indicators in the six dimensions

In the plots, the R square values in the vertex degree dimension and the assortativity coefficients both show right-skewed distributions around 0. More PTNs in the database have a low hierarchy. As for the closeness centrality and betweenness centrality dimensions, the R square values are left-skewed distributed, and more PTNs in the database have a high hierarchy. In the eigenvector centrality and the network modularity dimensions, the indicator values both show approximate normal distributions. The similarities and differences between the three pairs of PTN hierarchy dimensions are discussed as follows.

Vertex degree & network assortativity dimensions The PTN hierarchy in the vertex degree and network assortativity dimensions are both prone to be relatively low. The reasons that result in the relatively low hierarchy are similar, which are the excessive low-degree vertices and rare high-degree vertices in the PTNs. These reasons result in the heterogeneity of the vertices' importance being low, and no significant vertex connection pattern in PTNs. In hierarchy-related research in other fields, high-hierarchy networks are usually associated with scale-free network structures. By contrast, the results of the case study show that scale-free network structures are not common in PTNs.

However, the generally low PTN hierarchy in the two dimensions does not mean the topological indicators are not capable of the PTN hierarchy comparison and quantification. The topological indicator values in the two dimensions are distinguishable for PTN hierarchy comparison.

Closeness centrality & betweenness centrality dimensions The majority of the assessed PTNs are prone to have a high hierarchy in both the closeness centrality and betweenness centrality dimensions. The closeness centrality and betweenness centrality are both based on the shortest paths between vertex pairs. The PTN hierarchy in the closeness centrality dimension reflects the network element organisation from the perspective of accessing stops. For example, if a PTN has a high hierarchy in the closeness centrality dimensions, there are stops with different levels of accessibility distributed in the network, so the destinations can be reached with less detouring by passing through the stops with different closeness centrality. The PTN hierarchy in the betweenness centrality dimension reflects the network element organisation from the perspective of the intermediacy and loads on infrastructures. For instance, in a PTN with a high hierarchy in the betweenness centrality dimension, the infrastructures with high intermediacy take more load. The PTNs in the case study tend to have a high hierarchy in the two dimensions for two reasons, the heterogeneity of the elements' importance and the gradually changing probability density of different importance in PDFs, which ensure the high goodness of fit for the fitted skewed-normal distribution curves.

Eigenvector centrality & network modularity dimensions Around half of the PTNs in the database have a relatively high hierarchy in both eigenvector centrality and network modularity dimensions. The distributions of the PTN hierarchy in the two dimensions are both concentrated in the middle of the histograms, showing approximate normal distributions. The PTN hierarchy in both two dimensions is related to the number of high-degree vertices and their concentration, but the relations are opposite. The PTN hierarchy in the eigenvector centrality dimension is high when the PTN has a small number and concentrated high-degree vertices. By contrast, the high PTN hierarchy in the network modularity dimension is related to the large number and widely distributed high-degree vertices. Moreover, the PTN hierarchy in the two dimensions is found to have a negative linear relation. Figure 4.25 presents the scatter plot for the PTN hierarchy in the two dimensions with the regression line. The R square value of fitting is 0.832, indicating the negative linear relation is credible.

The PTN hierarchy in the eigenvector centrality dimension is more suitable for evaluating the mono-centric PTNs, while the network modularity dimension reflects the multi-centric PTNs' hierarchy better. The two dimensions complement each other in the application objects.

Evaluation and comparison With the radar chart representation, the PTN hierarchy can be comprehensively evaluated and compared from six perspectives (see Section 3.5). Taking the comparison between the Bogota BRT network and the Atlanta metro network as an example, Figure 4.26 presents the six-dimension radar chart of the two PTNs and the referenced median PTN hierarchy of the case study database. The methods of using the radar chart for PTN hierarchy evaluation and comparison are introduced as follows.

In Figure 4.26, the enclosed yellow and pink areas denote the six-dimension hierarchy of the Bogota BRT network and the Atlanta metro network, respectively. Since the PTN hierarchy

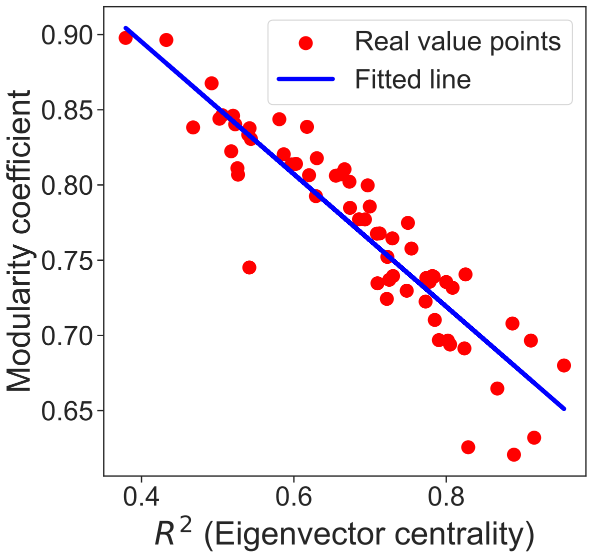


Figure 4.25: Scatter plot of eigenvector centrality R square and modularity coefficient with the fitted line

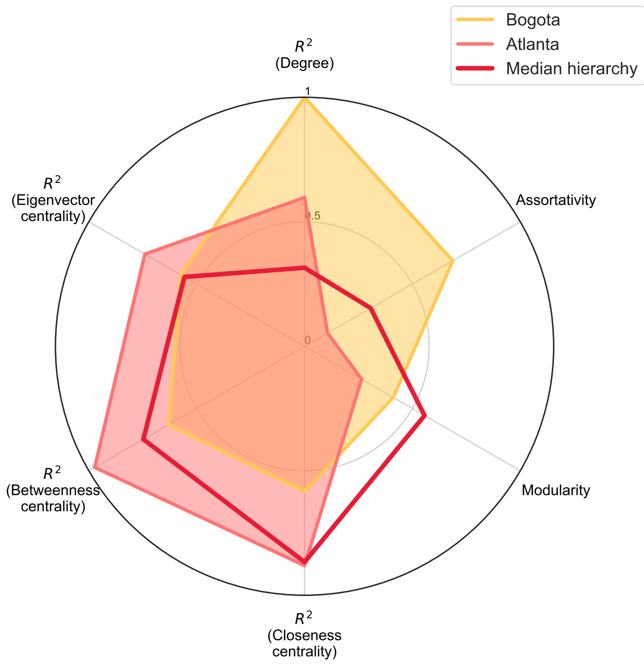


Figure 4.26: PTN hierarchy comparison of the Bogota BRT network and the Atlanta metro network with six-dimension radar chart

in each dimension is not evenly distributed, the median values represent the reference values of the six-dimension PTN hierarchy, shown in the radar chart as the enclosed red lines.

According to the previous analysis of the PTN hierarchy in each dimension, it has been found that differing from high-hierarchy networks in other fields, the scale-free network structure and the vertex connection pattern are insignificantly embodied in PTNs. Thus, in the assessment of the six-dimension radar chart, the R square values and the assortativity coefficients are not used as the primary indicators. By contrast, the hierarchy in closeness centrality and betweenness centrality dimensions has better effects. The PTN hierarchy is sufficiently differentiated in both dimensions, and most PTNs in the database tend to have a relatively high hierarchy. The hierarchy in the two dimensions assesses the PTN element organisation from the stop accessibility and the intermediacy of traffic load, which are important for passengers and service providers in the PTN operation. As for the PTN hierarchy in eigenvector centrality and network modularity dimensions, the distributions both show a nearly normal distribution, indicating the mono-centric or multi-centric network structure of the PTN.

In the radar chart of the Bogota BRT network, the hierarchy in the closeness centrality and betweenness centrality are both lower than the Atlanta metro network and the median PTN hierarchy. The lower hierarchy in the two dimensions indicates that the Bogota BRT network performs worse in the stop accessing and the traffic load intermediacy, and may affect the network performance in the PTN operation, especially in stop accessibility for passengers. In the eigenvector centrality and network modularity dimensions, both the two PTNs have a relatively high hierarchy in the eigenvector centrality dimension and a relatively low hierarchy in the network modularity dimension, presenting mono-centric network structures and having important stop clusters concentrated in an area. Compared to the Atlanta metro network, the Bogota BRT network has a less obvious mono-centric structure. As for the vertex degree and network assortativity dimensions, the Bogota BRT network has a higher hierarchy in both dimensions compared to the Atlanta metro network, showing a more significant scale-free network structure and a stop connection pattern that high-degree stops connect with high-degree stops.

Radar chart shape patterns By investigating the radar charts of the 63 PTNs in the database (Figure 4.27), some PTNs show radar chart shape patterns, which indicates the similarity between PTNs. The following are two identified patterns for discussion.

Figure 4.28 shows a radar chart shape pattern in four PTNs, including metro networks of Copenhagen, Kobe, Marseilles and Naples. The shapes of the four radar charts are mainly distributed on the left sides. In the four radar charts, the PTN hierarchy is relatively high in the closeness centrality, betweenness centrality and eigenvector centrality dimensions and moderate in the vertex degree dimension. In the dimensions of network assortativity and network modularity, the PTN hierarchy is relatively low, especially in the network modularity dimension. This pattern of radar shapes indicates the four PTNs show hierarchical element organisation in terms of stop accessibility and traffic flow intermediacy, which tend to have good network performance during operation. The four PTNs also show relatively significant mono-centric structures, having a few concentrated important vertex clusters in the central area. The scale-free structures are modestly embodied in the four PTNs, while the vertex connection patterns are relatively insignificant.

The maps of the four PTNs in Figure 4.29 confirm the assessment of the radar charts. The ring structures in the networks enable hierarchical organisations of elements in terms of stop accessibility and traffic flow intermediacy. The four PTNs show mono-centric structures. The PTNs show modest scale-free network structures. Few direct connections between high-degree stops, showing an insignificant vertex connection pattern.

Figure 4.30 shows three PTN hierarchy radar charts having another shape pattern, including the Amsterdam tram network, the Hartford BRT network and the London metro network. In the radar charts of the three PTNs, the PTN hierarchy in the closeness centrality and network modularity dimensions is relatively high and moderate in the betweenness centrality dimension. In the vertex degree, network assortativity and eigenvector centrality dimensions, the PTN hierarchy is relatively low. This radar chart shape pattern indicates that the three PTNs have high hierarchical element organisations in terms of stop accessibility, and passengers can access stops with less detouring. As far as traffic flow intermediacy, the three PTNs' element organisation is less hierarchical, which may be due to excessive low-intermediacy ele-

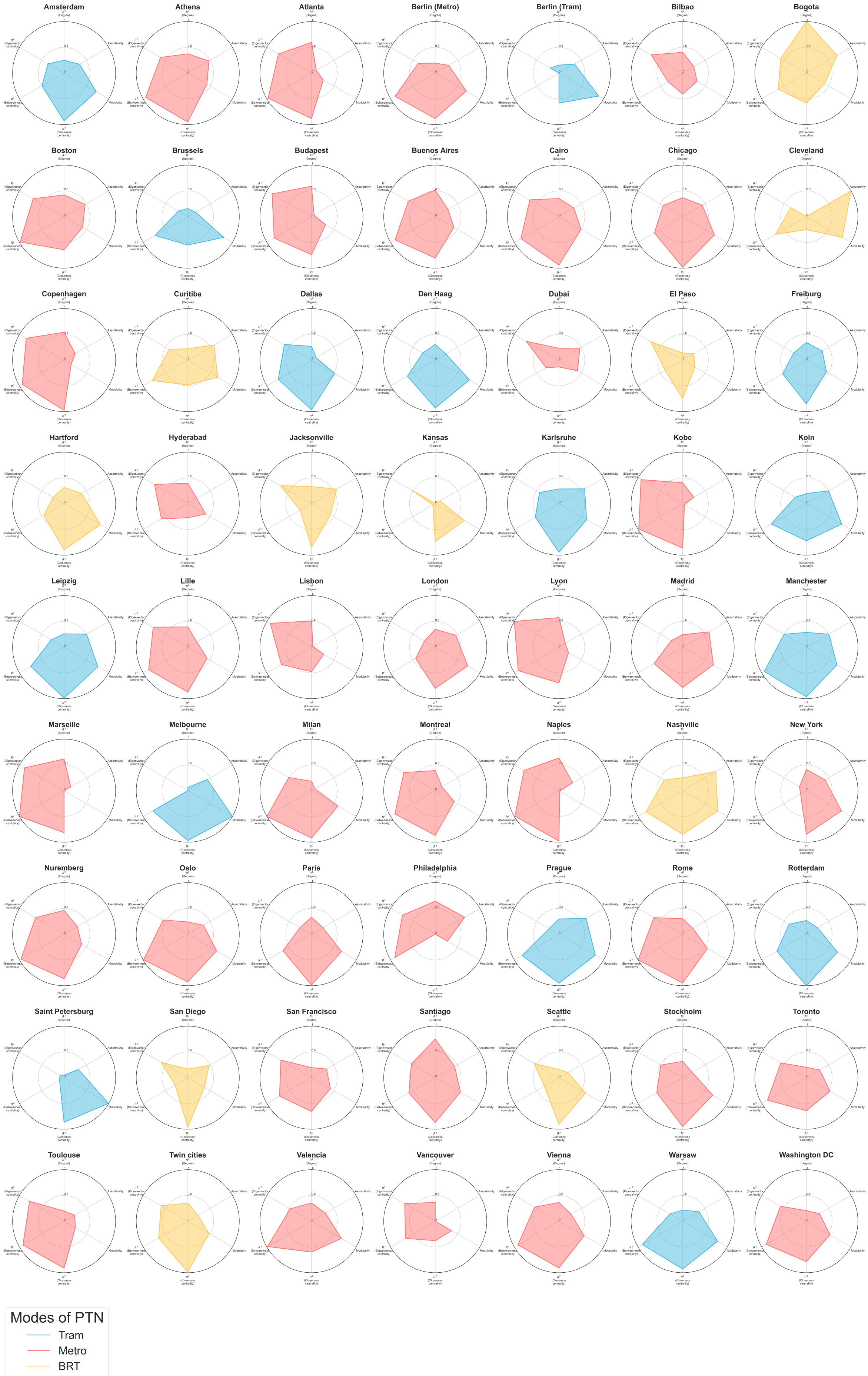


Figure 4.27: Radar charts of all selected high-capacity unimodal PTNs

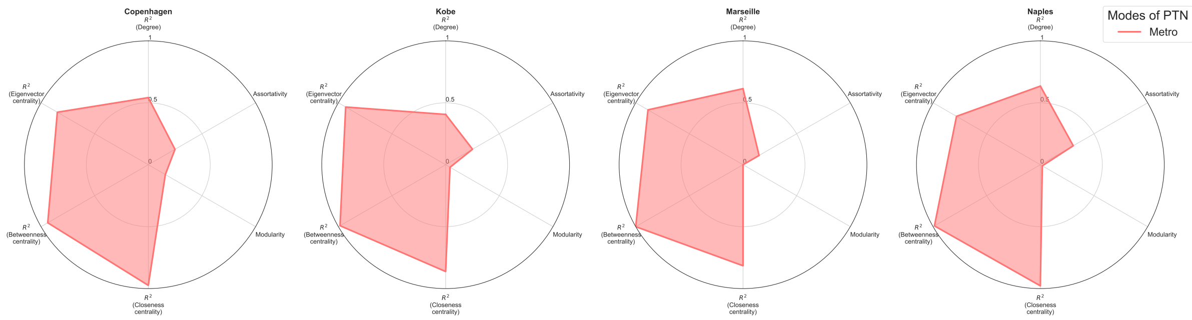
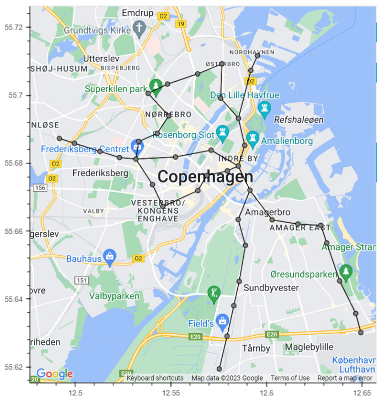
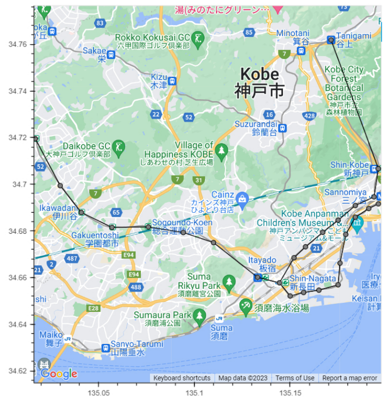


Figure 4.28: Radar charts of PTNs with radar chart shape pattern 1

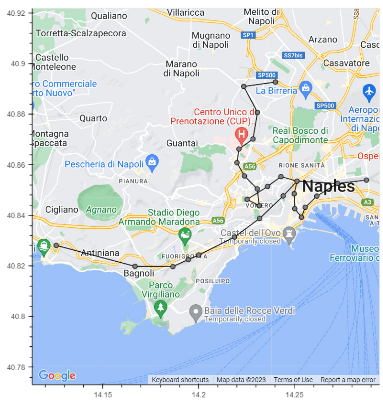


(a) Copenhagen (Metro)



(b) Kobe (Metro)

(c) Marseilles (Metro)



(d) Naples (Metro)

Figure 4.29: Maps of the PTNs with radar chart shape pattern 1

ments. The three PTNs have relatively significant multi-centric structures, and the high-degree vertices are widely distributed. The scale-free structures and vertex connection patterns by vertex degrees are also relatively less obvious.

The maps of the three PTNs in Figure 4.31 confirm the assessment of the radar charts. The transfer stops in the three PTNs are widely distributed, resulting in hierarchical organisations of stops by their accessibility. The PTNs are high-clustering, and show multi-centric network structures. Because of the relatively large number of dead ends, the low-intermediacy elements in the three PTNs are excessive, and the hierarchical structures in terms of traffic flow intermediacy are less significant. Due to the long linear network segments, the three PTNs are

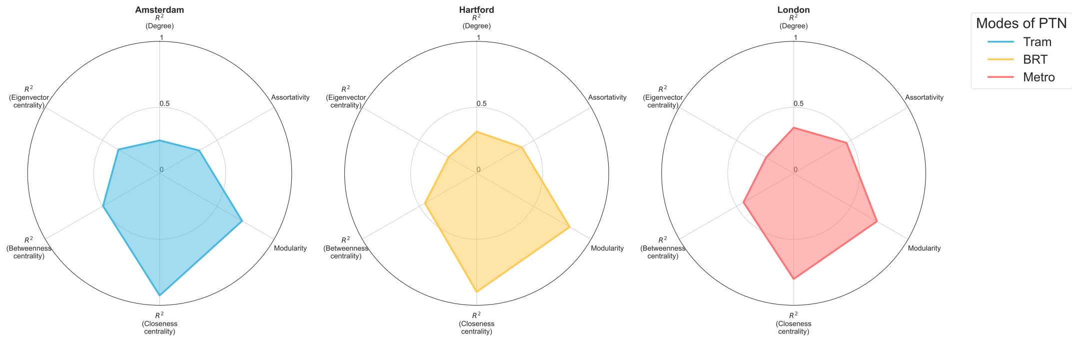
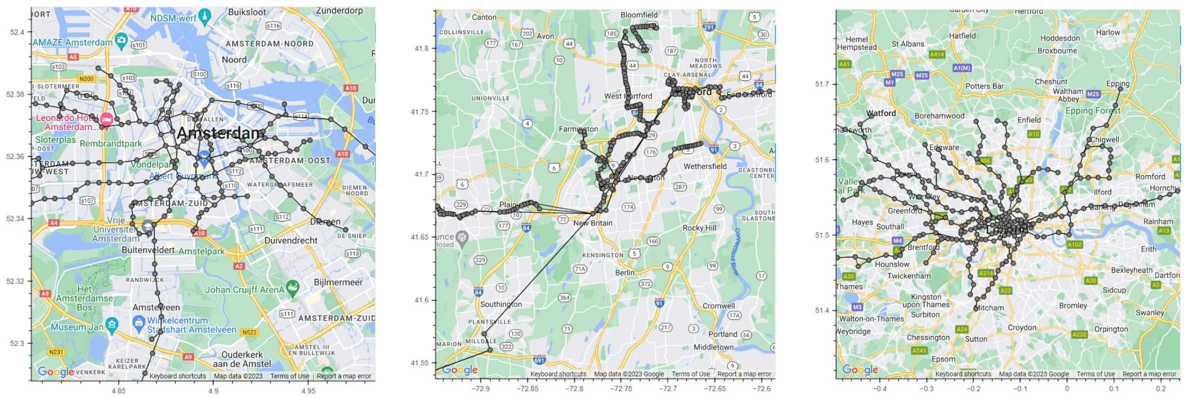


Figure 4.30: Radar charts of PTNs with radar chart shape pattern 2

less scale-free. The small number of direct connections between high-degree vertices results in insignificant vertex connection patterns in the PTNs.



(a) Amsterdam (Tram) (b) Hartford (BRT) (c) London (Metro)

Figure 4.31: Maps of the PTNs with radar chart shape pattern 2

4.3. Mode-wise comparison

In this study, three public transport modes are involved: the metro, tram and BRT. The pie chart in Figure 4.32 shows the proportion of each mode. Among the 63 high-capacity unimodal PTNs in the database, the metro networks take up over 50%, with 37 metro networks. The number of tram networks is slightly larger than the BRT networks, with 15 tram networks and 11 BRT networks.

4.3.1. Mode-wise effects

In Figure 4.33(a), the six-dimension PTN hierarchy radar chart for the three high-capacity modes is presented. Since the distributions of the PTN hierarchy indicators are uneven, the median hierarchy in each dimension of each mode is selected for representation. The PTN hierarchy in the six dimensions is scaled to the range from 0 to 1 in the radar chart. The scaled median values of the PTN hierarchy in each dimension are listed in Table 4.33(b) as reference.

In the six-dimension radar chart, no mode has the highest hierarchy in all dimensions, and the differences between modes are insignificant. Supported by Table 4.33(b), the order of modes having a high PTN hierarchy in the overall six dimensions is metro, tram and BRT. The metro networks have the highest median hierarchy in the dimensions of vertex degree,

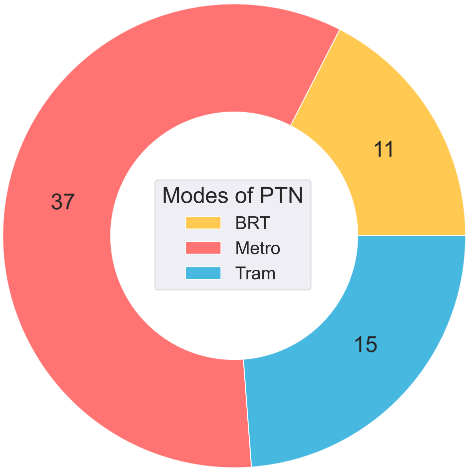
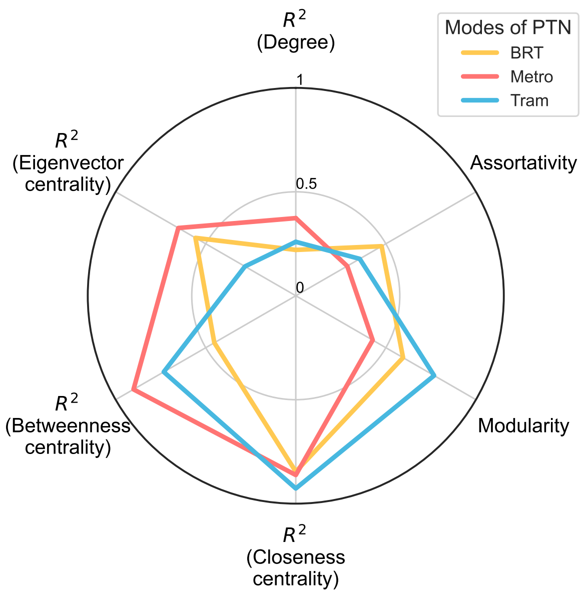


Figure 4.32: Pie chart of modes of the selected PTNs



(a) Radar chart

	R^2 (Degree)	Assortativity coefficient	Modularity coefficient	R^2 (Closeness centrality)	R^2 (Betweenness centrality)	R^2 (Eigenvector centrality)
BRT	0.221	0.477	0.595	0.846	0.452	0.557
Metro	0.373	0.286	0.427	0.862	0.901	0.652
Tram	0.259	0.355	0.768	0.927	0.733	0.282

(b) Scaled median hierarchy

Figure 4.33: Mode-wise comparison of PTN hierarchy indicators

betweenness centrality and eigenvector centrality. The tram networks have the median highest PTN hierarchy in the closeness centrality and network modularity dimensions. In the network assortativity dimension, the BRT networks have a higher median hierarchy than the other two modes.

4.3.2. Influencing factors

The three high-capacity modes have differences in multiple characteristics, such as the stop and line spacing, stop facilities, operating speeds and stopping patterns. These factors have effects on PTNs of modes and influence their hierarchy.

The first influencing factor is the stop spacing and line spacing of modes. According to previous studies, the average stop spacing of metro networks is larger than the other two modes. The average stop spacing of BRT networks usually ranges from 800 to 1600 metres (Walker, 2012). For tram networks, the typical average stop spacing ranges from 400 to 1,500 meters. And the typical average stop spacing of metro networks is from 600 to 2,500 metres (Kolks et al., 2003). The metro and BRT networks have a higher stop spacing and line spacing than the tram networks (Walker, 2012).

Another influencing factor is the operating speed of modes. The typical operating speeds of metro networks are 50 to 80 kilometres per hour (Feng et al., 2011). For tram networks, the typical operating speed range from 20 to 25 kilometres per hour (Spårvagnsstäderna, 2023). The typical operating speed of BRT networks is similar to tram networks, ranging from 23.8 to 24.8 kilometres per hour (Jain et al., 2022). The metro networks usually have higher operating speeds than tram networks and BRT networks.

The third influencing factor is the facilities of the stops. Metro stops generally have larger land use than the tram and BRT networks. The larger land use of metro stops enables metro transfer stops can have more lines cross at the same stop (Walker, 2012).

The fourth influencing factor is called the flexible stopping pattern, which is observed in the connected non-adjacent stops in the PTNs of the case study database. The reasons that result in flexible stopping patterns are stop-skipping and the lines with different stop spacing operating on the same infrastructures (Liu et al., 2013). The flexible stopping patterns are more likely to occur in the BRT and metro networks.

4.3.3. Vertex degree dimension

In the scattered box plot in Figure 4.34, the metro networks generally have a higher hierarchy than tram networks and BRT networks. The distribution of tram networks' PTN hierarchy is concentrated from 0.2 to 0.4. As for BRT networks, the variance between networks is high. Although one BRT network has the highest vertex degree R square in the database, most of the BRT networks have relatively low R square values.

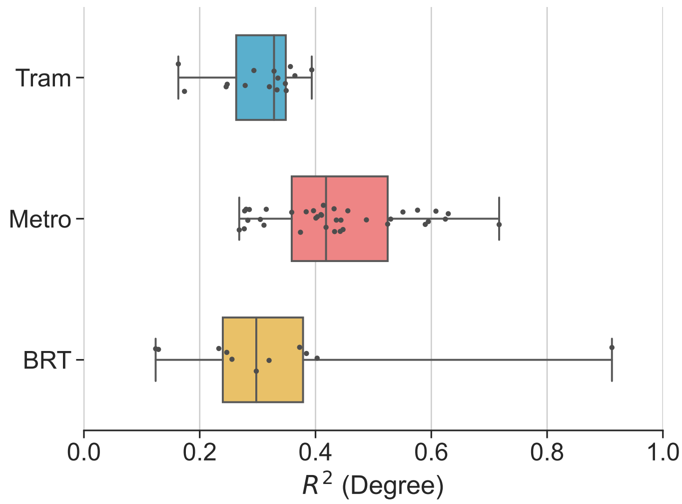


Figure 4.34: Mode-wise PTN hierarchy box plot in the vertex degree hierarchy dimension

The low PTN hierarchy in the vertex degree dimension is related to the excessive low-degree vertices and lack of high-degree vertices. Most of the low-degree vertices are the vertices located on linear network segments, so PTNs with fewer vertices on linear network segments have positive effects on hierarchy. On linear network segments with the same length, the low-degree vertices in metro networks are fewer than in tram and BRT networks for the larger stop spacing. Besides, the high-degree vertices are usually the transfer stops for multiple lines. The larger land use of metro stops usually enables metro transfer stops to have more lines cross at the same stop, resulting in more high-degree vertices. In the tram and BRT networks, the numbers of lines crossing at the same stop are limited.

4.3.4. Closeness centrality dimension

Figure 4.35 presents the scattered box plots of the mode-wise R square values in the closeness centrality dimension. The order of modes by the median R square values in the closeness centrality dimension is tram, metro and BRT. The highest hierarchy of all modes is similar. The distribution range of the metro networks is the largest among the three modes, nearly two times larger than tram networks'.

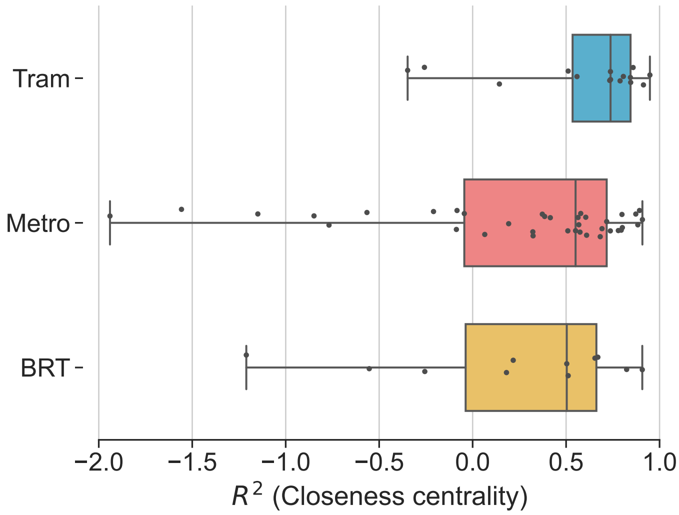


Figure 4.35: Mode-wise PTN hierarchy box plot in the closeness centrality dimension

In some PTNs, the high stop and line spacing makes some metro and BRT networks less likely to have connections between lines. In these PTNs, the metro and BRT lines only have connections in the central areas. These connection vertices have high closeness centrality, but the number of vertices with medium closeness centrality is few. Therefore, the number of vertices does not gradually descend with the closeness centrality increase, and the R square values of these PTNs are low or negative. By contrast, tram networks have small line spacing, and are more likely to form connections between lines. So there are fewer tram networks with relatively low R square values.

4.3.5. Betweenness centrality dimension

According to Figure 4.36, the order of modes having high median PTN hierarchy is metro, tram and BRT. Most of the metro networks are concentrated in the range from 0.9 to 1.0, higher than the other two modes. Compared to BRT networks, more tram networks have a higher hierarchy.

The stop and line spacing have effects on the PTN hierarchy in the betweenness centrality dimension. The betweenness centrality indicates the intermediacy of the elements in

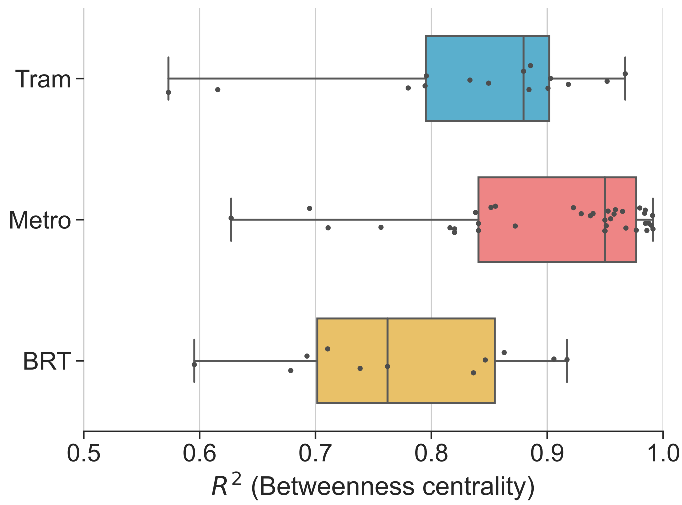


Figure 4.36: Mode-wise PTN hierarchy box plot in the betweenness centrality dimension

the PTNs. For the metro networks, the stop and line spacing is larger than the other two modes, making the structures of the metro networks simple. The intermediacy of the network elements is high in the central areas and gradually decreases in the suburbs. For the tram networks, the stop and line spacing is smaller than the metro networks, and the density of lines is higher, making the importance of network elements less concentrated and more elements with low intermediacy. Differing from rail-bound networks, such as metro and tram networks, BRT networks usually have fewer requirements on the infrastructure. The BRT lines are relatively more independent and have fewer transfer stops to other lines, resulting in excessive low-intermediacy elements and relatively low hierarchy in the betweenness centrality dimension.

4.3.6. Eigenvector centrality dimension

Based on the median R square values, The order of the modes having a high PTN hierarchy is metro, BRT and tram. In Figure 4.37, metro networks tend to have a relatively high hierarchy, and over a quarter of metro networks have a higher hierarchy than all tram and BRT networks. The distribution ranges of tram and BRT networks are similar and smaller than metro networks. Tram networks tend to have a relatively lower hierarchy than the other two modes.

The reasons for the PTN hierarchy differences between modes in the eigenvector centrality dimension are related to stop spacing and line spacing. In metro networks, the large stop and line spacing enables a small number of high-degree vertices in the central areas. These vertices are concentrated and form a centre. So the metro networks are more likely to be mono-centric and have a relatively high PTN hierarchy. For tram networks, the small line and stop spacing result in the high-degree vertices being widely distributed. The tram networks usually do not perform mono-centric structures and tend to have a lower hierarchy. In BRT networks, the distribution of high-degree vertices shows no pattern to be concentrated or widely distributed, so the PTN hierarchy is relatively moderate. Besides, the operating speed has effects on the PTN hierarchy. Since the weights of the edges are the inverse of the average travel time of the trips on the edges, the higher operating speed can reduce the average travel times on the edges and increase the importance of the vertex clusters. The operating speed of metro networks is the highest among the three modes, having positive effects on the PTN hierarchy in the eigenvector centrality dimension.

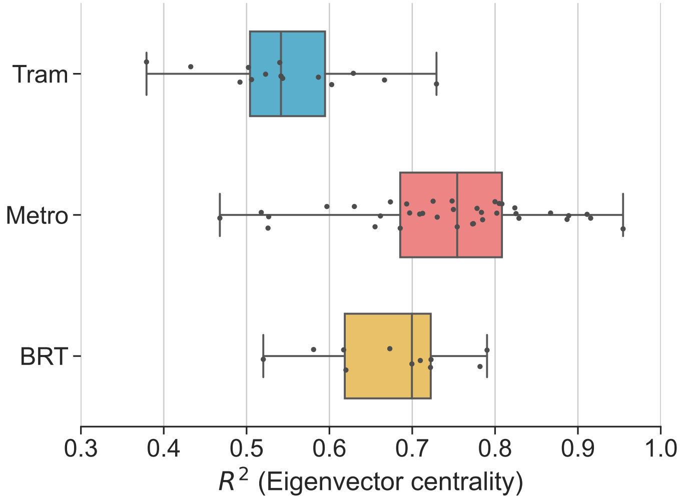


Figure 4.37: Mode-wise PTN hierarchy box plot in the eigenvector centrality dimension

4.3.7. Network modularity dimension

The order of modes having a high PTN hierarchy is tram, BRT and metro in the network modularity dimension. Figure 4.38 shows that tram networks tend to have a higher PTN hierarchy than PTNs with the other two modes. The distribution ranges of the three modes are similar, and the range of metro networks is slightly wider than the other two. BRT networks tend to have a moderate PTN hierarchy between tram and metro networks. Around a quarter of metro networks in the database have a lower hierarchy than all tram and BRT networks.

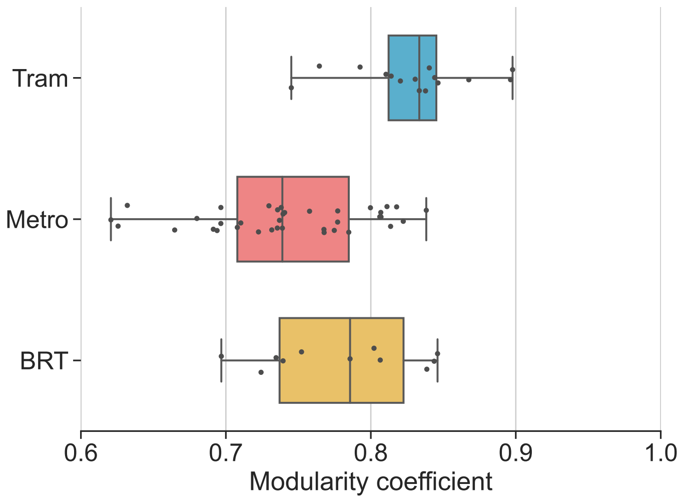


Figure 4.38: Mode-wise PTN hierarchy box plot in the network modularity dimension

Opposite to the eigenvector centrality dimension, the small line and stop spacing increase the PTN hierarchy and are more likely to form multi-centric structures. The optimal modularity dimensions are more effective in evaluating multi-centric networks. Because of the small line and stop spacing, the tram networks have large numbers of and widely distributed high-degree vertices, showing high-clustering and multi-centric structures, so the tram networks have a high PTN hierarchy. By contrast, not all metro networks are multi-centric, and some of them have a small number of high-degree vertices concentrated. Additionally, the flexible stopping patterns in some BRT networks result in a wider distribution of high-degree vertices and have positive effects on the PTN hierarchy in the network modularity dimension.

4.3.8. Network assortativity dimension

The order of modes having a high PTN hierarchy is BRT, tram and metro in the network assortativity dimension. In the scattered box plot as Figure 4.39, the assortativity coefficient values of BRT networks have the widest distribution range, from negative values to over 0.5. Compared to tram networks, more metro networks have a lower hierarchy.

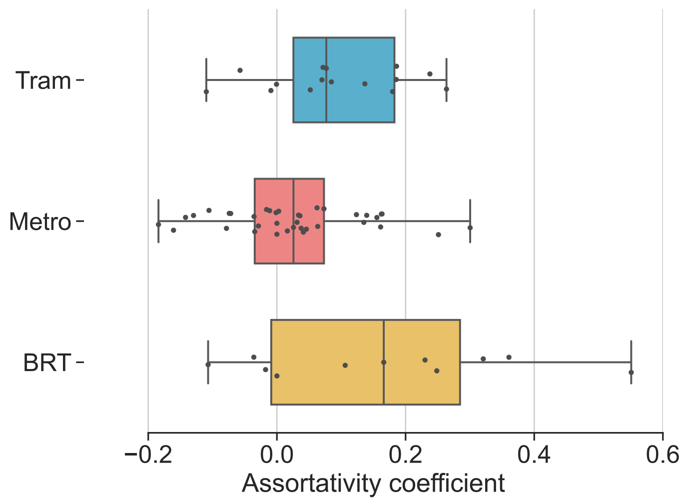


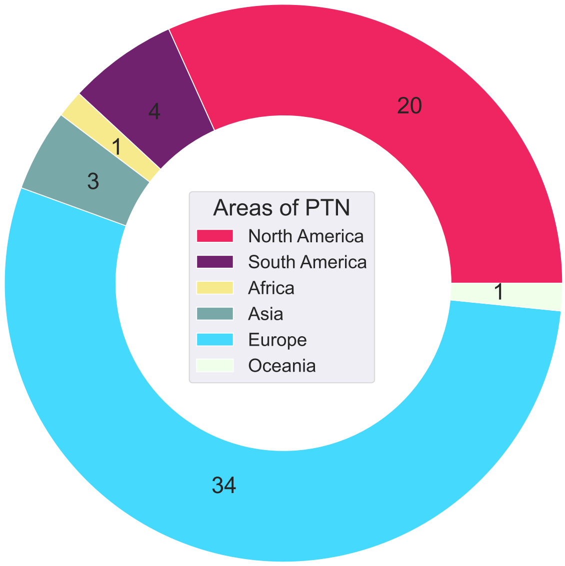
Figure 4.39: Mode-wise PTN hierarchy box plot in the network assortativity dimension

In PTNs, the high-degree transfer stops usually have a larger coverage of areas, so there is no need to have multiple high-degree transfer stops directly connected. The larger coverage of areas and land use for metro transfer stops reduce the possibility for metro networks to have high-degree vertices connected with high-degree vertices patterns. In a PTN with a high assortativity coefficient, the high-degree vertices are directly connected and form a connection pattern. Moreover, the stopping patterns during operations have effects on the assortativity coefficient. The flexible stopping pattern can increase the number of connected high-degree transfer stops and increase the assortativity coefficients. These stopping patterns are more likely to occur in the BRT networks, thus the distribution ranges of assortativity coefficients of metro and BRT networks are wider than tram networks.

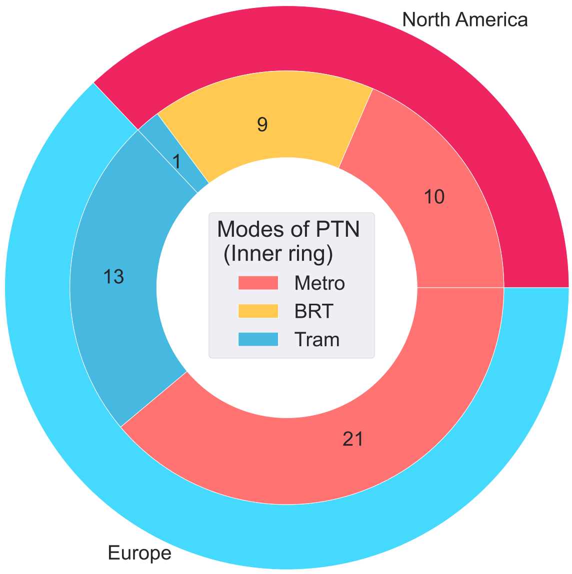
4.4. Continent-wise comparison

In the database of this case study, the high-capacity unimodal PTNs are distributed in six continents: Africa, Asia, Europe, Oceania, North America, and South America. Figure 4.40(a) shows the pie chart of the continents of high-capacity unimodal PTNs.

In the pie chart, PTNs from two continents occupy the majority of the proportion: North America and Europe. They have 20 and 35 high-capacity unimodal PTNs each. By contrast, the numbers of PTNs in other continents are too small to represent the continents and have credible conclusions about the continent-wise effects. Thus, the continent-wise comparison only considers the high-capacity unimodal PTN located in North America and Europe. The nested pie chart in Figure 4.40(b) presents the mode compositions of the PTNs in North America and Europe, where the outer ring represents the proportion of continents, and the inner ring represents the mode components of corresponding continents. PTNs in Europe consist of 13 tram networks and 21 metro networks. Besides, 10 metro networks, 9 BRT networks and a tram network consisting of the North American PTNs.



(a) Pie chart of continents of the selected PTNs



(b) Nested pie chart of modes of selected PTNs in Europe and North America

Figure 4.40: Pie charts of continents and modes of the selected PTNs

4.4.1. Continient-wise effects

Figure 4.41(a) and Table 4.41(b) present the radar chart and the table of the median continent-wise six-dimension PTN hierarchy. In the radar chart, the PTN hierarchy in the two continents is similar in vertex degree and network assortativity dimensions. In the network modularity, closeness centrality and betweenness centrality dimensions, the European PTNs have a higher hierarchy than North American PTNs. North American PTNs have a higher hierarchy in the eigenvector centrality dimension. In general, the European PTNs have a higher hierarchy and a larger enclosed area by the indicator points than North American PTNs.

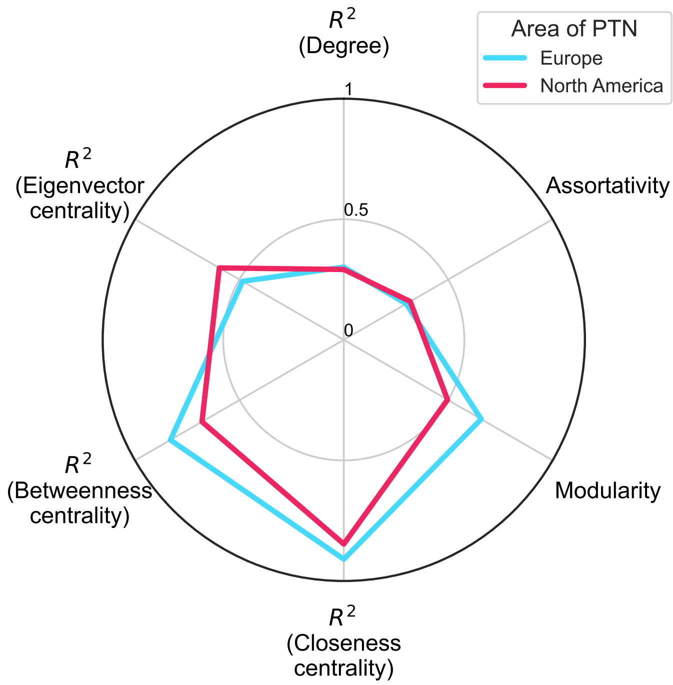
4.4.2. Box plots by dimensions

Box plots of continent-wise PTN hierarchy by the six dimensions are presented. Associated with historical reasons, urban density, and public transport-related policies, a discussion about the reasons that result in the continent-wise difference is proposed in Appendix B.

In Figure 4.42, the PTN hierarchy of the two continents has a strong similarity. The differences between the median values and the distribution range are small. Most PTNs in both continents have a low hierarchy from 0.2 to 0.4.

In Figure 4.43, the median hierarchy of the European PTNs is higher than the North American PTNs. The majority of PTNs in the two continents have a medium and above hierarchy. In addition, North American PTNs' distribution range is larger than European PTNs and has a higher proportion of negative values. The larger number of tram networks increases the general PTN hierarchy in Europe, and the European networks tend to have a higher hierarchy than North Americans.

In Figure 4.44, the hierarchy distribution ranges of the two continents are similar, both from around 0.6 to 1.0. Most PTNs in these two continents show a high hierarchy in the betweenness centrality dimension. Besides, most European PTNs have a hierarchy in the range higher than 0.8, while the North American PTNs show less concentration. It has been found previously that metro networks usually have a higher hierarchy in the betweenness



(a) Radar chart

	R^2 (Degree)	Assortativity coefficient	Modularity coefficient	R^2 (Closeness centrality)	R^2 (Betweenness centrality)	R^2 (Eigenvector centrality)
Europe	0.302	0.299	0.658	0.910	0.831	0.485
North America	0.292	0.318	0.497	0.847	0.679	0.596

(b) Scaled median hierarchy

Figure 4.41: Continent-wise comparison of PTN hierarchy indicators

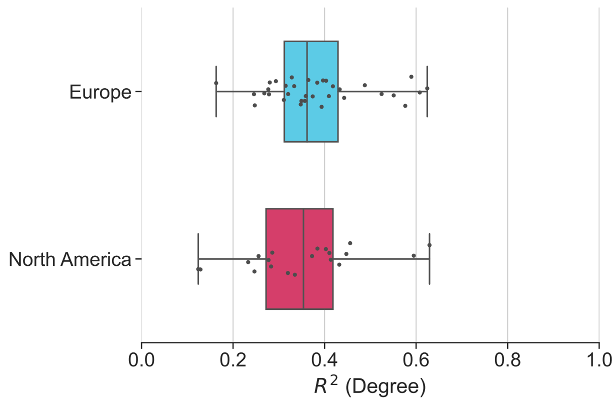


Figure 4.42: Continent-wise PTN hierarchy box plot in the vertex degree dimension

centrality dimension. Hence, the larger number of metro networks in Europe has positive effects on the PTN hierarchy, resulting in European PTNs tending to have a higher hierarchy than North Americans.

In Figure 4.45, the majority of PTNs in the two continents have a medium and above hierarchy. Besides, the distribution range of European PTNs' hierarchy is wider than the North American PTNs. Most of the North American PTNs have a hierarchy in the range from 0.6 to 0.8, while the European PTNs' hierarchy is evenly distributed from around 0.43 to 0.95.

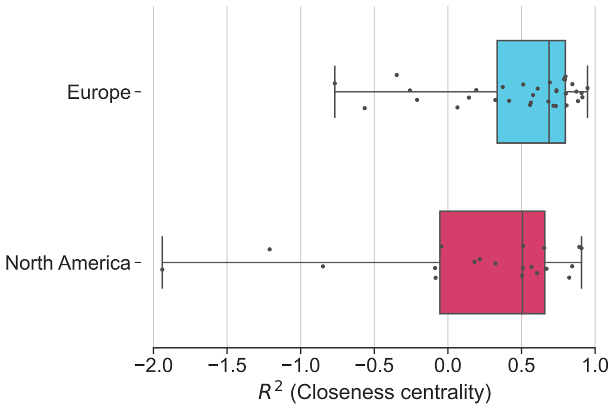


Figure 4.43: Continent-wise PTN hierarchy box plot in the closeness centrality dimension

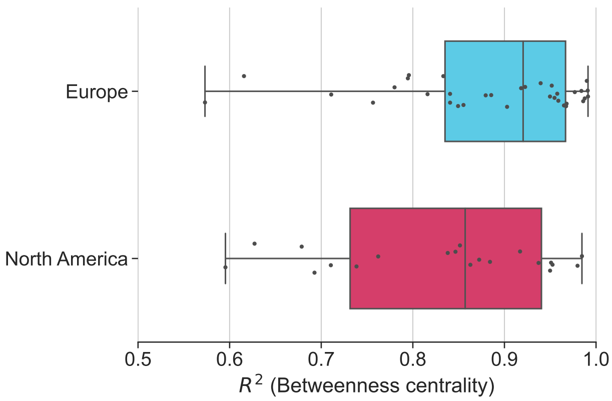


Figure 4.44: Continent-wise PTN hierarchy box plot in the betweenness centrality dimension

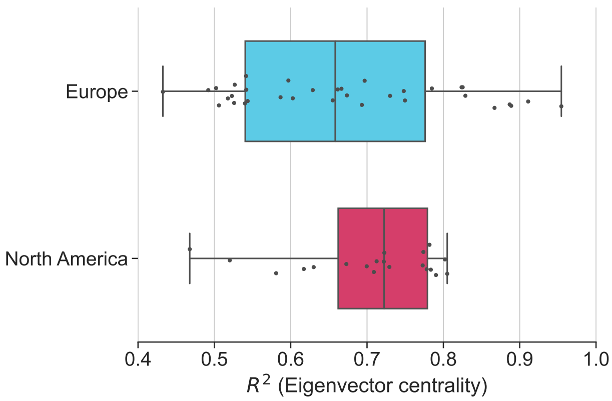


Figure 4.45: Continent-wise PTN hierarchy box plot in the eigenvector centrality dimension

Previous findings show that metro networks tend to have a relatively high hierarchy, and tram networks generally have a relatively low hierarchy in the eigenvector centrality. The larger proportions of metro and tram PTNs result in a wider range of PTN hierarchy in Europe than in North America.

In Figure 4.46, most of the PTNs in the two continents have a high hierarchy. The distribution range of North American PTNs' hierarchy (around 0.70 to 0.85) is smaller than the European PTNs (around 0.62 to 0.90). Similar to the eigenvector centrality dimensions, the higher proportions of metro and tram networks in Europe enlarge its hierarchy distribution

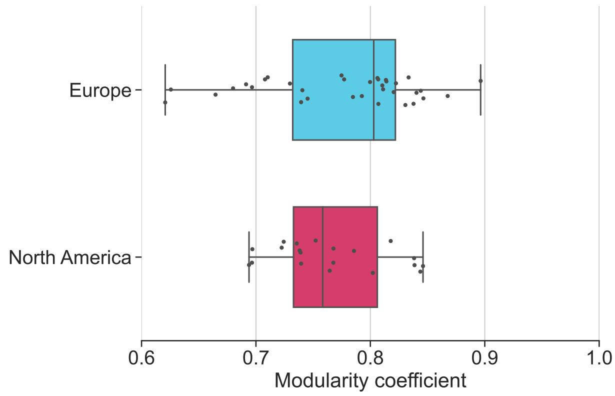


Figure 4.46: Continent-wise PTN hierarchy box plot in the network modularity dimension

range.

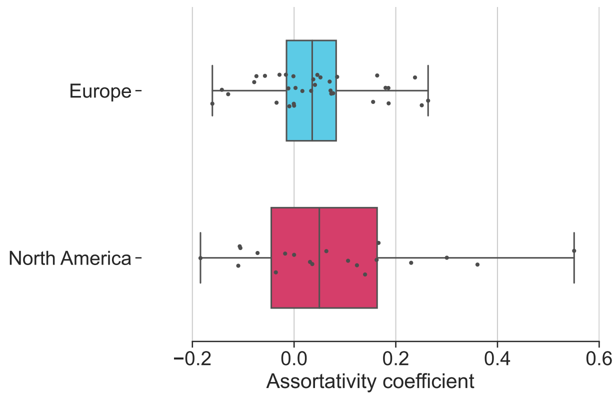


Figure 4.47: Continent-wise PTN hierarchy box plot in the network assortativity dimension

In Figure 4.47, the PTN hierarchy in the two continents is generally low, both concentrated around 0. The lowest values of the two boxes are similar, around -0.2, but several North American PTNs have a hierarchy higher than 0.3. The distribution range of the European PTNs is smaller than the North American PTNs. Due to the more commonly flexible stopping patterns, the BRT networks usually have a higher hierarchy in the network assortativity dimension. Since no European BRT network is included in this study, the distribution range of European PTNs is smaller than North Americans.

4.4.3. Discussion

In the continent-wise comparison, the modal composition of the two continents influences the PTN hierarchy. Apart from metro networks, the number of European tram networks is larger than in North America, and only North American BRT networks are included in the database. For example, European PTNs in the study have a generally high hierarchy in the closeness centrality dimension and a generally low hierarchy in the assortativity dimension. The results of the continent-wise comparison confirm the previous assessment of the mode-wise effects on PTN hierarchy.

5
Conclusion

This chapter presents the conclusions and recommendations based on the research conducted. In Section 5.1, the key findings are summarized along with the answers to the research questions. Section 5.2 and Section 5.3 outline the contributions and practical applications of the research. Finally, in Section 5.4 and Section 5.5, the limitations and future work recommendations are presented.

5.1. Key findings

This thesis presented an assessment of the hierarchy in high-capacity unimodal PTNs and how modes of transport and geographical location might influence this hierarchy. With regard to the organisation of network elements, the PTN hierarchy is defined as a network property. A six-dimension topology-based quantification methodology is developed on both network and element scales. A database that relies on GTFS data and contains the network topology information of 63 high-capacity unimodal PTNs worldwide is used as a case study. The PTN hierarchy is interpreted and analysed along the dimensions of topological indicators. Furthermore, the effects of modes and continents are discussed by comparing the PTN hierarchy.

Through the research presented in this thesis, the proposed research questions are answered. The following contents provide a summary of the answers to the sub-research questions.

1. *What are the definition and the topological characteristics of the PTN hierarchy?*

   The PTN hierarchy is defined as a property of a PTN, indicating the organisation of the elements in the PTN, where the number of elements gradually descends when the importance increases, with very few elements of high importance and the majority of elements in the network with low importance.

   By reviewing and synthesising relevant literature, this research identified six topological characteristics of the PTN hierarchy with two scales. There are three element-scale characteristics, including accessibility, intermediacy, and clusters' importance, and three network-scale characteristics, including scale-free network structure, high-clustering network structure, and vertex connection pattern.

2. *What indicators can be used to quantify the network topology characteristics of PTN hierarchy?*

   The PTN hierarchy quantification method uses six topological indicators in two types, vertex-based and network-based.

The vertex-based indicators include the vertex degree, closeness centrality, betweenness centrality, and eigenvector centrality, which are used to quantify the scale-free network structure, the vertex's accessibility, the elements' intermediacy, and the vertex cluster's importance, respectively. The goodness of fit indicator, R square value, between the indicators' PDF and the regressed skewed normal curves quantify the PTN hierarchy.

Apart from scale-free network structures using the vertex-based degree for hierarchy quantification, other network-scale PTN hierarchy characteristics are quantified with network-based indicators. The network-based indicators include the modularity coefficient and assortativity coefficient, which represent the PTN hierarchy by quantifying the high-clustering network structure and vertex connection pattern of PTNs.

Moreover, based on the normalised hierarchy in dimensions, six-dimension radar charts comprehensively show the hierarchy of each PTN and enable visualised comparison.

3. *Based on the selected indicators, how do the high-capacity unimodal PTNs worldwide quantify and compare in terms of hierarchy?*

   According to the values of the selected indicators, the 63 high-capacity unimodal worldwide PTNs' hierarchy are evaluated in the six dimensions.

   (a) Vertex degree dimension: The $R^2(Degree)$ value distribution indicates a tendency of PTNs to have a relatively low hierarchy. It appears that for most PTNs in the study, the scale-free network structure is not significant. The excessive low-degree vertices and the lack of high-degree vertices reduce the hierarchy.

   (b) Closeness centrality dimension: The $R^2(Closeness)$ value distribution indicates a tendency towards a relatively high hierarchy of PTNs in the database, denoting hierarchical organisations of stops in terms of their accessibility. The low heterogeneity of vertex numbers with different closeness centrality can reduce the PTN hierarchy.

   (c) Betweenness centrality dimension: The $R^2(Betweenness)$ value distribution indicates a tendency towards a relatively high hierarchy of PTNs in the database, denoting hierarchical organisations of stops or links in terms of their intermediacy of traffic load. The lack of elements with medium intermediacy can reduce the goodness of fit.

   (d) Eigenvector centrality: The $R^2(Eigenvector)$ values present a slightly left-skewed distribution, indicating a similarity in the numbers of PTNs exhibiting relatively low and high hierarchy. High PTN hierarchy is characterised by a hierarchical organisation of vertex clusters based on their importance and the presence of mono-centric structures.

   (e) Network modularity dimension: The optimal modularity coefficients present a slightly left-skewed distribution, indicating a similarity in the numbers of PTNs exhibiting relatively low and high hierarchy. Wide-distributed high-degree vertices and the presence of multi-centric structures characterise high PTN hierarchy.

   (f) Network assortativity dimension: The assortativity coefficient distribution indicates a tendency towards a relatively low hierarchy, indicating the vertex connection patterns with high-degree vertices connect to high-degree vertices are not significant in most PTNs of the database. The small number of high-degree vertices and limited direct connections between them reduce the coefficient values.

According to the PTN hierarchy in the six dimensions, evaluations in the closeness centrality and betweenness centrality dimensions hold a high priority across the six dimensions. The hierarchy in these two dimensions evaluates the organisation of PTN elements from the perspectives of stop accessing and traffic flow intermediacy, which could serve as important references for passengers and public service providers during PTN operations. The hierarchy in the eigenvector centrality and network modularity dimensions assesses the mono-centric and multi-centric structures, respectively. The hierarchy in the vertex degree and network assortativity dimensions is less apparent than in the former four dimensions, because of the insignificant scale-free structures and vertex connection patterns by degrees in PTNs of the database.

4. *Based on high-capacity unimodal PTNs' hierarchy, what are the mode-wise and continent-wise effects on PTN hierarchy?*

   The mode-wise effect analysis considers three high-capacity public transport modes, the metro, tram and BRT. In continent-wise comparison, due to the limited data sizes of PTNs in some continents, such as Asia and South America, only PTNs in Europe and North America are included. The mode-wise and continent-wise effects are based on the median PTN hierarchy in each of the six dimensions of modes and continents.

   The order of modes having PTN hierarchy from high to low is metro, tram, and BRT. None of the modes has the highest PTN hierarchy in all dimensions. Metro has the highest PTN hierarchy in the dimensions of vertex degree, betweenness centrality and eigenvector centrality dimensions. Tram has the highest PTN hierarchy in the dimensions of network modularity and closeness centrality. BRT has the highest PTN hierarchy in the dimension of network assortativity. The PTN hierarchy is mainly influenced by stop spacing and line spacing (vertex degree, closeness centrality, betweenness centrality, eigenvector centrality and the network modularity dimensions), the stop infrastructure (vertex degree, betweenness centrality and network assortativity dimensions), the operating speed (eigenvector centrality dimension) and stopping pattern (network modularity dimension and network assortativity dimension). These influencing factors affect the number of vertices in the PTNs, the number of connections between vertices and the weights of edges, which affect the topology of PTNs in different modes, and cause the mode-wise effects on the PTN hierarchy.

   The European PTNs tend to have a higher hierarchy than North American PTNs. Neither of the two continents has a higher PTN hierarchy in all dimensions. In the vertex degree and network assortativity dimensions, the PTN hierarchy of the two continents is almost the same. Europe has a higher PTN hierarchy in the dimensions of network modularity, closeness centrality and betweenness centrality. North America has a higher PTN hierarchy in the eigenvector centrality dimension. The continent-wise effects mainly result from the different modal compositions, especially the proportions of tram and BRT networks in the two continents.

By answering the sub-research questions above, the main research question, "**How to quantify and compare unimodal PTN hierarchy?**", is answered.

5.2. Contributions

This research contributes to the field of PTN hierarchy comparison and quantification in methodology, GTFS data processing pipeline and results analysis.

This research develops a methodology for PTN hierarchy comparison and quantification. First, a unified definition of the PTN hierarchy is established. In this definition, the PTN hierarchy is determined by the essential organisation of the elements in PTNs. Then, based on the literature review, six topological characteristics of the PTN hierarchy are identified in element and network scales. Each characteristic is quantified by either vertex-based or network-based topological indicators, in order to quantify the PTN hierarchy from six dimensions. For vertex-based indicators, the R square values between element-scale indicators' PDF and skewed normal distributions are regarded as the PTN hierarchy. Besides, the network-based coefficients are direct as the quantified PTN hierarchy. Based on the normalised PTN hierarchy in the six dimensions, the PTN hierarchy is visually compared through radar charts. Since the methodology is solely based on network topology theories, it also has the potential to be applied to multi-modal PTNs.

To apply the methodology to a case study for high-capacity unimodal PTNs worldwide, a database that depends on GTFS data and contains L-space topology information is constructed with a data pipeline. First, the GTFS data are collected from online open resources. Next, the high-capacity unimodal PTN data are filtered from the whole PTN. After representing and calibrating filtered data in L-space, a high-capacity unimodal PTN is converted to a graph file, and visualised with its map. Relying on this GTFS data processing pipeline, 63 high-capacity unimodal PTNs worldwide are processed and included in the database.

With the quantified hierarchy of the high-capacity unimodal PTNs in the database, the analysis and discussion of the PTN hierarchy are conducted. It has been found that PTN hierarchy in the closeness centrality and betweenness centrality holds priority across the six dimensions, which could serve as important references for passengers and service providers during operation. The PTN hierarchy in eigenvector centrality and network modularity dimensions evaluates the mono-centric or multi-centric structures. The scale-free structures and vertex connection patterns by degree implicated by PTN hierarchy in vertex degree and network assortativity dimensions are generally not significant for PTNs in the database, and relatively less apparent. In addition, metro networks present a tendency to have a high hierarchy than the tram and BRT networks. The mode-wise effects result from factors influencing PTN topologies, such as stop spacing and line spacing. For different modal compositions, the European PTNs have a higher hierarchy compared to North American PTNs.

5.3. Practical applications

The contributions of the research can practically support public transport authorities. Supported by the six-dimension radar charts, the PTN hierarchy can be compared to PTNs with similar sizes or populations. In the PTN planning stages, the related authorities can set several PTNs with good network performance and mature operations as comparators. Applying the quantification methodology to the planned PTN and compared PTNs, the normalised six-dimension PTN hierarchy can be visually compared in the same radar chart. The PTN hierarchy in the closeness centrality and betweenness centrality dimensions needs more attention, which can reflect the PTN performance in terms of stop accessing and traffic flow intermediacy, and are important for passengers and public transport service providers. When the hierarchy in the two dimensions is significantly lower than the compared PTNs, it indicates that the topology of the planned PTN needs improvement. The hierarchy in the eigenvector centrality and network modularity dimensions indicates the mono-centric or multi-centric network structure of the planned PTN, guiding the authorities to plan the related facilities with different strategies, such as parking lots. The planned PTN also need to achieve the average hierarchy of the compared PTNs in the vertex degree and network assortativity dimensions. Based on the contributions of this study, the planning can be optimised before committing to construction. The PTN hierarchy methodology has the following advantages in practical use. Initially, the data requirement of the methodology is not extensive, and only GTFS data are needed, which are generated during the daily operation and easily accessible. Next, the assessment of the PTN hierarchy is multi-dimension, including network-scale characteristics and element-scale characteristics from different aspects. Moreover, the PTN hierarchy is presented as standardised quantified values, making the results clear and intuitive.

5.4. Limitations

While some insights are gained from this research, there are several limitations to the research scope.

The first limitation is the coverage of the database used in the case study. Due to a lack of publicly-available data or the low quality of data, some iconic PTNs are not included in this case study, especially some PTNs in Asia and America, such as the metro networks of Beijing, Tokyo and Shanghai, and the BRT networks of Jakarta and Mexico City. This limitation affects the coverage of the continent-wise comparison, and only European and North American PTNs are included, limiting the application scope of conclusions about the continent-wise effects.

Second, the calibration of PTN representation brings limitations. When representing the GTFS data of high-capacity unimodal PTNs in L-space, the calibration of vertices and edges in the PTNs with the operating maps is semi-automatic and needs manual confirmation, which is time-consuming and can easily result in human errors. The errors of vertices and edges in the PTNs can lead to different PTN hierarchy results. For example, if wrongly merged two stops in the calibration of PTN graphs, the topological indicator values would change, and change the quantified PTN hierarchy in the six dimensions.

The third limitation is the analysis of mode-wise effects on PTN hierarchy is based on the typical values of the influencing factors due to the data limitation. For instance, the stop spacing and operating speed of modes, rather than the actual values of each PTN in the database. Although the PTNs with the same mode have similar characteristics, they can also be adjusted to suit the needs or features of cities. By applying the typical influencing factor values, the special cases of PTNs are excluded, and the conclusion applicability of the mode-wise effects can be reduced in some PTNs.

5.5. Recommendations for future research

Based on the limitations, recommendations are given for future research.

First, to improve the current methodology, future work can focus on the progress of converting GTFS data to PTNs with L-space representation. For example, the manual calibration of the draft PTN graphs. The calibration with the PTN operating maps can be optimised with automatic methods. The machine learning algorithms can be used for studying the operating maps of PTNs and automatically calibrating the visualised PTNs. The optimised data calibration can avoid human errors, and improve the accuracy of results and manual processing time.

Second, the six-dimension PTN hierarchy can be attempted to integrate into one single metric. In the case study, it has been found that the PTN hierarchy in the six dimensions has different levels of priority. For example, PTN hierarchy in the closeness centrality and betweenness centrality dimensions hold a higher priority than the other four dimensions. Future work can work on quantifying the weights of dimensions and integrating the six-dimension PTN hierarchy with one single quantified metric. Based on the single quantified PTN hierarchy metric, the assessment of the PTN hierarchy can be direct, simplified and accurate.

Third, the PTN hierarchy can be further studied with the relationship between the vulnerability of PTNs. In past research, the topological indicators are usually associated with PTN vulnerability analysis, such as closeness centrality and betweenness centrality (Berche et al., 2009; Cats et al., 2017; Chopra et al., 2016). For example, a PTN can be robust to random failures, but vulnerable to disruptions on some critical stops (Chopra et al., 2016). Both PTN hierarchy and PTN vulnerability focus on the element organisations in PTNs, especially the elements with high importance. Thus, the relationship between the PTN's hierarchy and vulnerability can be an interesting topic to explore in the future.

Fourth, the influencing factors for evaluating the PTN hierarchy mode-wise effects can use the actual data of each PTN. By applying the actual data of influencing factors, the relation between the mode-wise effects and influencing factors can be quantified analysis.

Fifth, the continent-wise effects on PTN hierarchy can be further discussed. To begin with, the PTNs in different continents can be unimodally compared to exclude the influence of modes. Next, the analyse of continent-wise effects can be evaluated with more related quantified factors, such as population density and GDP. These factors offer new perspectives on continent-wise effects and help to build deep understanding.

Bibliography

Acton, B., Le, H.T.K., Miller, H.J., 2022. Impacts of bus rapid transit (BRT) on residential property values: A comparative analysis of 11 US BRT systems. Journal of Transport Geography 100, 103324.

Aston, L., Currie, G., Kamruzzaman, M., Delbosc, A., Brands, T., Van Oort, N., Teller, D., 2021. Multi-city exploration of built environment and transit mode use: Comparison of Melbourne, Amsterdam and Boston. Journal of transport geography 95, 103136.

Badhrudeen, M., Derrible, S., Verma, T., Kermanshah, A., Furno, A., 2022. A geometric classification of world urban road networks. Urban Science 6(1), 11.

Barabási, A.L., Bonabeau, E., 2003. Scale-free networks. Scientific American 288(5), 60–69.

Bassolas, A., Barbosa-Filho, H., Dickinson, B., Dotiwalla, X., Eastham, P., Gallotti, R., Ghoshal, G., Gipson, B., Hazarie, S.A., Kautz, H., et al., 2019. Hierarchical organization of urban mobility and its connection with city livability. Nature communications 10(1), 1–10.

Bavelas, A., 1950. Communication patterns in task-oriented groups. The journal of the acoustical society of America 22(6), 725–730.

Berche, B., Von Ferber, C., Holovatch, T., Holovatch, Y., 2009. Resilience of public transport networks against attacks. The European Physical Journal B 71, 125–137.

Bienenstock, E.J., Bonacich, P., 2021. Eigenvector centralization as a measure of structural bias in information aggregation. The Journal of Mathematical Sociology , 1–19.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment 2008(10), P10008.

Bokeh Development Team, 2018. Bokeh: Python library for interactive visualization. https://bokeh.pydata.org/en/latest/. Last visited 2023-03-25.

Bonacich, P., 1987. Power and centrality: A family of measures. American journal of sociology 92, 1170–1182.

Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., Wagner, D., 2007. On modularity clustering. IEEE transactions on knowledge and data engineering 20(2), 172–188.

Broido, A.D., Clauset, A., 2019. Scale-free networks are rare. Nature communications 10(1), 1–10.

Bron, C., Kerbosch, J., 1973. Algorithm 457: finding all cliques of an undirected graph. Communications of the ACM 16(9), 575–577.

Buehler, R., Pucher, J., 2012. Demand for Public Transport in Germany and the USA: An Analysis of Rider Characteristics. Transport Reviews 32(5), 541–567.

Buijtenweg, A., Verma, T., Cats, O., Donners, B., Wang, H., 2021. Quantifying the hierarchy of public transport networks, in: 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE. pp. 1–6.

Cats, O., 2017. Topological evolution of a metropolitan rail transport network: The case of Stockholm. Journal of Transport Geography 62, 172–183.

Cats, O., Birch, N., 2021. Multi-modal network evolution in polycentric regions. Journal of Transport Geography 96, 103159.

Cats, O., Koppenol, G.J., Warnier, M., 2017. Robustness assessment of link capacity reduction for complex networks: Application for public transport systems. Reliability Engineering & System Safety 167, 544–553.

Cats, O., Vermeulen, A., Warnier, M., Van Lint, H., 2020. Modelling growth principles of metropolitan public transport networks. Journal of Transport Geography 82, 102567.

Charles, P., Rony, G., 2016. Chapter 7 - graph creation and analysis for linking actors: Application to social data, in: Automating Open Source Intelligence. Syngress, pp. 103–129.

Chen, E., Stathopoulos, A., Nie, Y.M., 2022a. Transfer station choice in a multimodal transit system: An empirical study. Transportation Research Part A: Policy and Practice 165, 337–355.

Chen, Y., Fuellhart, K., Zhang, S., Witlox, F., 2022b. Airport classification in Chinese multi-airport regions: An interaction network perspective between aviation and high-speed rail. European Journal of Transport and Infrastructure Research 22(2), 1–21.

Chopra, S.S., Dillon, T., Bilec, M.M., Khanna, V., 2016. A network-based framework for assessing infrastructure resilience: a case study of the London metro system. Journal of The Royal Society Interface 13(118), 20160113.

Clauset, A., Moore, C., Newman, M.E.J., 2008. Hierarchical structure and the prediction of missing links in networks. Nature 453(7191), 98–101.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal Complex Systems, 1695. https://igraph.org. Last visited 2023-03-25.

Czégel, D., Palla, G., 2015. Random walk hierarchy measure: What is more hierarchical, a chain, a tree or a star? Scientific reports 5(1), 1–14.

Derrible, S., 2012. Network centrality of metro systems. PloS one 7(7), e40575.

Dimitrov, S.D., Ceder, A.A., 2016. A method of examining the structure and topological properties of public-transport networks. Physica A: Statistical Mechanics and its Applications 451, 373–387.

Directorate-General for Mobility and Transport, 2022. EU invests €5.4 billion in sustainable, safe, and efficient transport infrastructure. https://transport.ec.europa.eu/news/eu-invests-eu54-billion-sustainable-safe-and-efficient-transport-infrastructure-2022-06-29_en. Last visited 2023-01-22.

Fang, C., Pang, B., Liu, H., 2017. Global city size hierarchy: Spatial patterns, regional features, and implications for China. Habitat International 66, 149–162.

Feng, X., Mao, B., Feng, X., Feng, J., 2011. Study on the maximum operation speeds of metro trains for energy saving as well as transport efficiency improvement. Energy 36(11), 6577–6582.

Fernandes, V.A., Pacheco, R.R., Fernandes, E., Da Silva, W.R., 2019. Regional change in the hierarchy of Brazilian airports 2007-2016. Journal of Transport Geography 79, 102467.

Freeman, L.C., 1977. A set of measures of centrality based on betweenness. Sociometry , 35–41.

Gao, J., Zhao, P., Zhuge, C., Zhang, H., 2012. Research on public transit network hierarchy based on residential transit trip distance. Discrete Dynamics in Nature and Society 2012.

Garandeau, C.F., Lee, I.A., Salmivalli, C., 2014. Inequality matters: Classroom status hierarchy and adolescents' bullying. Journal of youth and adolescence 43(7), 1123–1133.

Gattuso, D., Miriello, E., 2005. Compared analysis of metro networks supported by graph theory. Networks and Spatial Economics 5(4), 395–414.

Godfrey, B.J., Zhou, Y., 1999. Ranking world cities: multinational corporations and the global urban hierarchy. Urban Geography 20(3), 268–281.

Google, I., 2022a. General Transit Feed Specification Reference. https://developers.google.com/transit/gtfs/reference. Last visited 2022-12-28.

Google, I., 2022b. Google Maps. https://www.google.com/maps. Last visited 2023-03-25.

Groth, D., Skandier, T., 2005. Network+ study guide. SYBEX Inc.

Grubesic, T.H., Matisziw, T.C., Zook, M.A., 2008. Global airline networks and nodal regions. GeoJournal 71(1), 53–66.

Guan, J., Li, Y., Xing, L., Li, Y., Liang, G., 2020. Closeness centrality for similarity-weight network and its application to measuring industrial sectors' position on the global value chain. Physica A: Statistical Mechanics and its Applications 541, 123337.

Guimerà, R., Amaral, L.A.N., 2005a. Cartography of complex networks: modules and universal roles. Journal of Statistical Mechanics: Theory and Experiment 2005(02), P02001.

Guimerà, R., Amaral, L.A.N., 2005b. Functional cartography of complex metabolic networks. nature 433(7028), 895–900.

Guimerà, R., Mossa, S., Turtschi, A., Amaral, L.N., 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. Proceedings of the National Academy of Sciences 102(22), 7794–7799.

Hage, P., Harary, F., 1995. Eccentricity and centrality in networks. Social networks 17(1), 57–63.

Háznagy, A., Fi, I., London, A., Németh, T., 2015. Complex network analysis of public transportation networks: A comprehensive study, in: 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE. pp. 371–378.

Hipp, R.D., 2020. SQLite. https://www.sqlite.org/index.html. Last visited 2023-03-25.

Hong, J., Tamakloe, R., Lee, S., Park, D., 2019. Exploring the topological characteristics of complex public transportation networks: Focus on Variations in both single and integrated systems in the Seoul metropolitan area. Sustainability 11(19), 5404.

Huang, S., Lv, T., Zhang, X., Yang, Y., Zheng, W., Wen, C., 2014. Identifying node role in social network based on multiple indicators. PloS one 9(8), e103733.

Jain, G.V., Jain, S., Parida, M., 2022. Evaluation of travel speed of conventional buses and bus rapid transit service in Ahmedabad city, India using geo-informatics. Journal of Public Transportation 24, 100034.

Jiang, J., Wen, S., Yu, S., Xiang, Y., Zhou, W., Hassan, H., 2017. The structure of communities in scale-free networks. Concurrency and Computation: Practice and Experience 29(14), e4040.

Kanrak, M., Nguyen, H.O., 2022. An analysis of connectivity, assortativity and cluster structure of the Asian-Australasian cruise shipping network. Maritime Transport Research 3, 100048.

Kleinberg, J.M., 1999. Hubs, authorities, and communities. ACM computing surveys (CSUR) 31(4es), 5–es.

Kolks, W., Fiedler, J., et al., 2003. Verkehrswesen in der kommunalen Praxis, Band I, Planung - Bau - Betrieb.

Kołoś, A., Taczanowski, J., 2016. The feasibility of introducing light rail systems in medium-sized towns in Central Europe. Journal of Transport Geography 54, 400–413.

Kujala, R., Weckström, C., Mladenović, M.N., Saramäki, J., 2018. Travel times and transfers in public transport: Comprehensive accessibility analysis based on pareto-optimal journeys. Computers, Environment and Urban Systems 67, 41–54.

Lai, W., McDysan, D., 2002. Network hierarchy and multilayer survivability. RFC 3386. RFC Editor. https://www.rfc-editor.org/rfc/rfc3386.html. Last visited 2023-03-25.

Liu, Z., Yan, Y., Qu, X., Zhang, Y., 2013. Bus stop-skipping scheme with random travel time. Transportation Research Part C: Emerging Technologies 35, 46–56.

Loginova, J., Sigler, T., Searle, G., O'Connor, K., 2022. The distribution of national urban hierarchies of connectivity within global city networks. Global Networks 22(2), 274–291.

Luce, R., Perry, A.D., 1949. A method of matrix analysis of group structure.

Luo, D., Cats, O., Van Lint, H., 2020. Can passenger flow distribution be estimated solely based on network properties in public transport systems? Transportation 47(6), 2757–2776.

Luo, D., Cats, O., Van Lint, H., Currie, G., 2019. Integrating network science and public transport accessibility analysis for comparative assessment. Journal of Transport Geography 80, 102505.

Meunier, D., Lambiotte, R., Fornito, A., Ersche, K.D., Bullmore, E.T., 2009. Hierarchical modularity in human brain functional networks. Frontiers in neuroinformatics 3, 37.

Mones, E., Vicsek, L., Vicsek, T., 2012. Hierarchy measure for complex networks. PloS one 7(3), e33799.

Newman, M.E.J., 2002. Assortative mixing in networks. Physical review letters 89(20), 208701.

Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary physics 46(5), 323–351.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the national academy of sciences 103, 8577–8582.

Newman, M.E.J., 2010. Networks: an introduction. Oxford: Oxford University Press.

OpenMobilityData, 2022. OpenMobilityData - public transit feeds from around the world. https://transitfeeds.com/. Last visited 2022-12-28.

Page, L., Brin, S., Motwani, R., Winograd, T., 1999. The PageRank citation ranking: Bringing order to the web. Technical Report. Stanford InfoLab.

Pattiselanno, K., Dijkstra, J.K., Steglich, C., Vollebergh, W., Veenstra, R., 2015. Structure matters: The role of clique hierarchy in the relationship between adolescent social status and aggression and prosociality. Journal of youth and adolescence 44(12), 2257–2274.

Pomykala, A., 2018. Effectiveness of urban transport modes, in: MATEC web of conferences, EDP Sciences. p. 03003.

Pumain, D., 2006. Hierarchy in Natural and Social Sciences. volume 3. Dordrecht: Springer.

Raicu, Ş., Dragu, V., Popa, M., Burciu, Ş., 2009. About the high capacity public transport networks territory functions. Urban Transport XV-Urban Transport and the Environment , 41–51.

Rao, C.R., 1973. Linear statistical inference and its applications. volume 2. Wiley New York.

Ravasz, E., Barabási, A.L., 2003. Hierarchical organization in complex networks. Physical Review E 67(2), 026112.

Robson, C., Barr, S., Ford, A., James, P., 2021. The structure and behaviour of hierarchical infrastructure networks. Applied Network Science 6(1), 1–25.

Rowe, R., Creamer, G., Hershkop, S., Stolfo, S.J., 2007. Automated social hierarchy detection through email network analysis, in: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, Association for Computing Machinery. p. 109–117.

Scott, D.W., 1979. On optimal and data-based histograms. Biometrika 66(3), 605–610.

Shanmukhappa, T., Ho, I.W.H., Tse, C.K., 2018. Spatial analysis of bus transport networks using network theory. Physica A: Statistical Mechanics and its Applications 502, 295–314.

Sidloski, M., Diab, E., 2020. Understanding the effectiveness of Bus rapid transit systems in small and medium-sized cities in North America. Transportation Research Record 2674(10), 831–845.

Silva, T.C., Dias, F.A.M., dos Reis, V.E., Tabak, B.M., 2022. The role of network topology in competition and ticket pricing in air transportation: Evidence from Brazil. Physica A: Statistical Mechanics and its Applications , 127602.

Soh, H., Lim, S., Zhang, T., Fu, X., Lee, G.K.K., Hung, T.G.G., Di, P., Prakasam, S., Wong, L., 2010. Weighted complex network analysis of travel routes on the Singapore public transportation system. Physica A: Statistical Mechanics and its Applications 389(24), 5852–5863.

Solé, R.V., Valverde, S., 2004. Information theory of complex networks: on evolution and architectural constraints, in: Complex networks. Springer, pp. 189–207.

Spårvagnsstäderna, 2023. Trams are efficient. https://www.sparvagnsstaderna.se/en/tramways/trams-are-efficient. Last visited 2023-01-22.

Tanglay, O., Young, I.M., Dadario, N.B., Taylor, H.M., Nicholas, P.J., Doyen, S., Sughrue, M.E., 2022. Eigenvector pagerank difference as a measure to reveal topological characteristics of the brain connectome for neurosurgery. Journal of Neuro-Oncology 157(1), 49–61.

Traag, V.A., Waltman, L., Van Eck, N.J., 2019. From Louvain to Leiden: guaranteeing well-connected communities. Scientific reports 9(1), 1–12.

Tsigdinos, S., Paraskevopoulos, Y., Kourmpa, E., 2022. Exploratory evaluation of road network hierarchy in small-sized cities: Evidence from 20 Greek cities. Transportation Research Procedia 60, 480–487.

Tsiotas, D., Polyzos, S., 2015. Decomposing multilayer transportation networks using complex network analysis: a case study for the Greek aviation network. Journal of Complex Networks 3(4), 642–670.

Van Nes, R., 2002a. Design of multimodal transport networks: A hierarchical approach. Ph.D. thesis. Technische Universiteit Delft.

Van Nes, R., 2002b. Multilevel network optimization for public transport networks. Transportation research record 1799(1), 50–57.

Von Ferber, C., Holovatch, T., Holovatch, Y., Palchykov, V., 2009. Public transport networks: empirical analysis and modeling. The European Physical Journal B 68(2), 261–275.

Vuchic, V.R., 2007. Urban transit systems and technology. John Wiley & Sons.

Walker, J., 2012. Human Transit: How clearer thinking about public transit can enrich our communities and our lives. Island Press.

Wang, Z., Luo, D., Cats, O., Verma, T., 2020. Unraveling the hierarchy of public transport networks, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), IEEE. pp. 1–6.

Yago, G., et al., 1984. The decline of transit: urban transportation in German and US cities, 1900-1970. Cambridge University Press.

Yap, M., Luo, D., Cats, O., Van Oort, N., Hoogendoorn, S., 2019. Where shall we sync? Clustering passenger flows to identify urban public transport hubs and their key synchronization priorities. Transportation Research Part C: Emerging Technologies 98, 433–448.

Yap, M.D., Van Oort, N., Van Nes, R., Van Arem, B., 2018. Identification and quantification of link vulnerability in multi-level public transport networks: a passenger perspective. Transportation 45(4), 1161–1180.

Yerra, B.M., Levinson, D.M., 2005. The emergence of hierarchy in transportation networks. The Annals of Regional Science 39(3), 541–553.

Zhang, J., Zhao, M., Liu, H., Xu, X., 2013. Networked characteristics of the urban rail transit networks. Physica A: Statistical Mechanics and its Applications 392(6), 1538–1546.

Zhao, L., Guan, M., Zhu, X., Karama, S., Khundrakpam, B., Wang, M., Dong, M., Qin, W., Tian, J., Evans, A.C., Shi, D., 2015. Aberrant topological patterns of structural cortical networks in psychogenic erectile dysfunction. Frontiers in Human Neuroscience 9.

A

Pairplot of the PTN hierarchy indicators

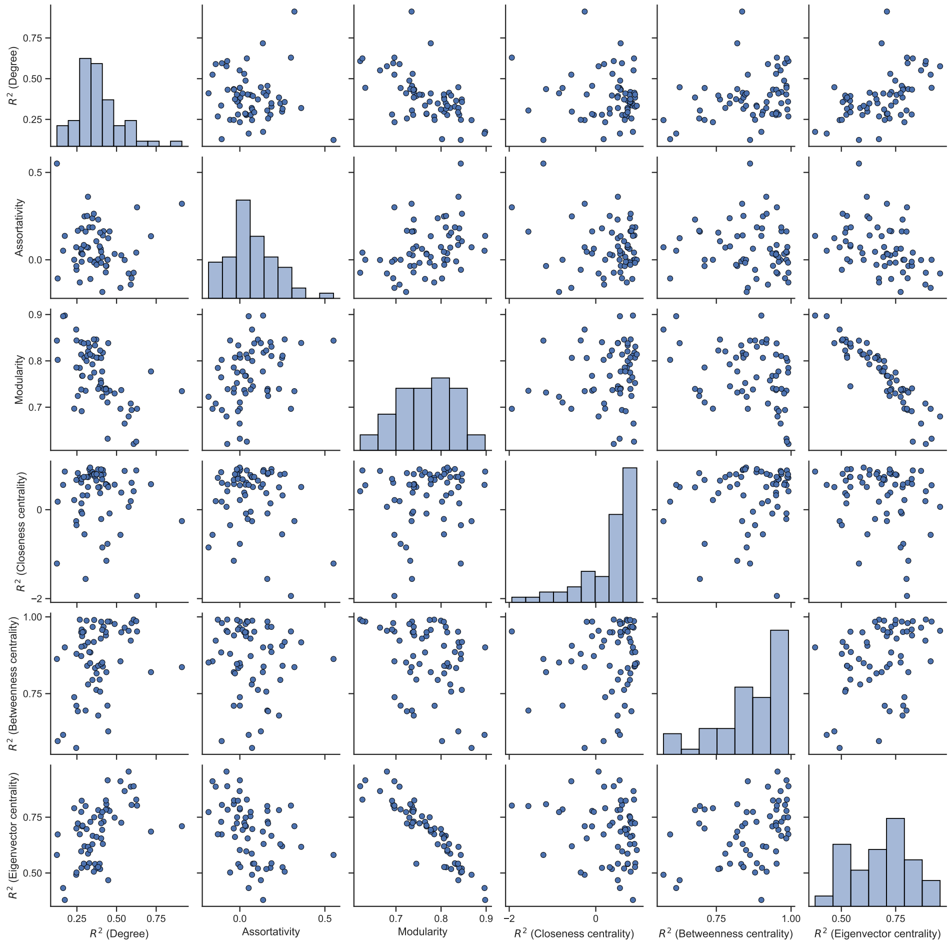


Figure A.1: Pairplot of the six topological indicators of PTN hierarchy

B

Continent-wise effects discussion

In addition to addressing the main research questions, an analysis of the factors relating to the continent-wise effects on high-capacity unimodal PTN hierarchy is of significant interest. Therefore, this section provides a detailed discussion of this question. Three factors related to the continent-wise effects on the PTN hierarchy are proposed: historical reasons, urban density, and public transport-related policies.

The development of the automotive industry after World War II had a profound impact on the development of PTN. Back in time to the post-World War II period, the economic development of Europe and North America are different. Since no war happened in North America, North American industry recovered and developed earlier than European's, especially the modern automotive industry (Buehler and Pucher, 2012). Plus the automotive industry-friendly policies of the time, car ownership witnessed a rapid rising in North America. By contrast, the North American financial allocation for public transport was undermined and resulted in the abandonment of the dilapidated public transport infrastructures (Yago et al., 1984). These historical reasons make North American transport car-oriented, and the modes of PTNs are mainly low-cost BRT, rather than metro and tram.

Urban density is another critical factor in determining the mode of high-capacity PTN. Because of the massive passenger demand in metropolises, both Europe and North American metropolises intend to use the metro as the mode of high-capacity PTN. The differences in mode choices for high-capacity PTNs are mainly in small and medium-sized cities. Due to the high urban density in Europe, the central areas of median and small-sized European cities usually have limited space and concentrated traffic demands. Tram is the appropriate mode for these cities for its higher unit capacity and the smaller line spacing requiredKołoś and Taczanowski (2016). As for the North American median and small-sized cities, BRT is the suitable mode for high-capacity PTNs. For the lower urban density with the lower demand for public transport, the revenue cannot meet the high cost of rail-bound PTN, such as tram (Kołoś and Taczanowski, 2016; Vuchic, 2007). BRT is the mode with the highest cost-effective for medium and small-sized North American cities (Sidloski and Diab, 2020).

Public transport-related policies are another essential factor in the differences between European and North American public transport. Different from the North American policies tilt towards car travel, the European policies are more public transport-friendly. In Europe, the fuel and car purchase taxes, and the restrictions on car parking and car use are more strict than in North America Buehler and Pucher (2012). The policies discourage travel by car. On the other hand, the EU invests a lot every year to develop sustainable transport infrastructures (Directorate-General for Mobility and Transport, 2022). The electrically powered trams meet the environment-friendly policies and the high public transport demand of Europeans (Kołoś and Taczanowski, 2016).

In summary, the differences in public transport between Europe and North America have contributed to the different preferences for public transport modes, with rail-bound PTNs more prevalent in Europe, while BRT is more common in North America, and further influence the continent-wise effects on PTN hierarchy.