Results in Engineering 27 (2025) 106126

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

Assessing Public Transit Network Efficiency and Accessibility in Johor Bahru and Penang, Malaysia: A Data-Driven Approach

Mahbub Hassan [a], Hridoy Deb Mahin [b], Ferdoushi Ahmed [c], Md. Maruf Hassan [d], Atikur Rahaman [e], Masuk Abdullah [f,g,*]

[a] Faculty of Civil Engineering & Technology, Universiti Malaysia Perlis (UniMAP), Arau, Perlis, 02600, Malaysia
[b] Sylhet Engineering College (SEC), School of Applied Sciences & Technology, Shahjalal University of Science and Technology (SUST), Sylhet 3100, Bangladesh
[c] Faculty of Economics, Prince of Songkla University (PSU), Hat Yai, Songkhla, 90110, Thailand
[d] Department of Computer Science & Engineering, Southeast University, Dhaka, 1208, Bangladesh
[e] School of Economics and Management, Jiujiang University, 551 Qianjin Donglu, Jiujiang, Jiangxi 332005, PR China
[f] Department of Vehicles Engineering, Faculty of Engineering, University of Debrecen, Ótemető Street 2-4, 4028, Debrecen, Hungary
[g] Faculty of Data Science and Information Technology, INTI International University, Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

ARTICLE INFO

Keywords:
GTFS
Public Transit Networks
Network Connectivity
Pedestrian Accessibility
Urban Mobility

ABSTRACT

This study conducts a comprehensive analysis of public bus transit networks in Johor Bahru and Penang, Malaysia, using General Transit Feed Specification (GTFS) data. It focuses on evaluating network efficiency, spatial equity, and accessibility within the myBAS Johor and Rapid Bus Penang systems. As the first academic application of Malaysia's newly released GTFS Static datasets, this research develops an integrated methodological workflow that combines GTFS with OpenStreetMap points of interest, global walking friction surfaces, and population density rasters. The analysis employs graph-theoretic centrality metrics, kernel density estimation, the Jaccard Similarity Index, and raster-based cost surface modeling to assess transit performance across structural, spatial, and temporal dimensions. Graph-based network analysis identify key urban hubs—such as Larkin Sentral in Johor Bahru and Terminal Komtar in Penang—as critical nodes for multimodal connectivity and efficient passenger transfers. Notable spatial inequities remain, with suburban and peripheral areas facing limited transit coverage, long walking distances, and poor access to key services like healthcare, education, and commerce. Service frequency drops significantly during off-peak hours, restricting mobility for transit-dependent users. Additionally, route overlap in dense urban corridors contrasts with inadequate service in outer areas, revealing operational inefficiencies. To address these challenges, the study recommends expanding stop coverage in underserved regions, optimizing route configurations in high-demand corridors, and piloting demand-responsive transit services to enhance adaptability and inclusivity. Beyond its contributions to Malaysian transport planning, the research offers a replicable, open-data-driven approach that can inform transit modernization efforts in cities undergoing digital transitions in transit network management and accessibility planning.

1. Introduction

Public transportation plays a vital role in promoting sustainable urban development by addressing challenges related to traffic congestion, environmental degradation, and socio-economic inequality [1–4]. In many developed countries, more than half of urban commuters rely on public transit, whereas in Asia, dependency is even higher due to rapid urbanization and high population density. Urbanization across Asia is occurring at an unprecedented pace, with cities absorbing millions of new residents annually a trend that poses significant implications for infrastructure, governance, and sustainability [5,6].While public transportation, walking, and cycling account for a significant portion of urban trips, the increasing rate of motor vehicle ownership has intensified congestion and pollution, contributing to substantial economic losses and environmental concerns [7,8]. Addressing these issues requires improved transit network performance, efficiency, and accessibility to promote public transportation as a viable alternative to private vehicle use.

* Corresponding author. Masuk Abdullah
E-mail address: masuk@eng.unideb.hu (M. Abdullah).

https://doi.org/10.1016/j.rineng.2025.106126
Received 8 February 2025; Received in revised form 22 June 2025; Accepted 3 July 2025
Available online 4 July 2025

Malaysia's urban mobility system has evolved considerably in recent decades, incorporating MRT, LRT, BRT Sunway, Rapid KL, myBAS, Rapid Penang, and Rapid Kuantan as part of a broader national strategy for integrated public transport [9,10]. In 2024, these services recorded over 14.5 million passenger trips, with myBAS Johor Bahru alone serving 6.4 million commuters [11]. Despite this demand, transit systems in Malaysia continue to exhibit critical performance issues. Spatial disparities in coverage, irregular service frequencies, limited first- and last-mile infrastructure, and gaps in connectivity between peripheral and urban core areas persist as barriers to equitable access. While policy initiatives from Ministry of Transport (MOT) and Agensi Pengangkutan Awam Darat (APAD) have focused on modernizing infrastructure through electrification, digital fare integration, and corridor upgrades, these interventions have often lacked empirical grounding in geospatial or operational data [12].

A significant institutional development occurred in November 2023, when Malaysia released its first General Transit Feed Specification (GTFS) Static datasets for public use [11]. These datasets provide machine-readable, globally standardized representations of transit services detailing route alignments, stop locations, scheduling patterns, and service calendars [13]. This release marks a pivotal transition point for Malaysian transit planning. Prior to GTFS availability, most evaluations were based on proprietary maps, consultant reports, or survey-based user perception studies, which lacked transparency, spatial specificity, and reproducibility. With the availability of GTFS data, Malaysia now has the capacity to move toward a data-integrated planning paradigm that aligns with international best practices in transport science.

The present study conducts a comprehensive, dual-city assessment of Malaysia's urban bus systems using GTFS Static data. The analysis focuses on two key networks myBAS Johor and Rapid Bus Penang which differ markedly in spatial configuration, demographic structure, and modal integration. Johor Bahru presents a decentralized and automobile-oriented pattern, while Penang features a denser, more compact network centered around key transit corridors. This contrast facilitates comparative analysis and enhances the generalizability of findings across heterogeneous urban contexts.

Methodologically, this research advances a multi-layered analytical approach by integrating GTFS with OpenStreetMap points of interest, population density rasters, and global walking friction surfaces. Centrality metrics such as degree, betweenness, and closeness are employed to identify network bottlenecks and structural vulnerabilities [14]. Spatial clustering is analyzed using kernel density estimation [15], while route overlap and system redundancy are quantified through the Jaccard Similarity Index [16]. Furthermore, pedestrian accessibility is modeled using accumulated cost surfaces [17] to estimate walking time from key urban amenities—including schools, hospitals, residential areas and commercial centers to the nearest transit stop under realistic conditions. These techniques, while individually well established, are here applied in concert to the Malaysian context for the first time using standardized, reproducible methods.

The contribution of this research lies not only in its empirical findings, but also in its demonstration of how open-format data and open-source tools can be adapted for transit evaluation in emerging economies. By operationalizing Malaysia's newly available GTFS datasets, the study introduces an integrated methodological workflow that is reproducible, scalable, and capable of guiding service redesign, investment prioritization, and accessibility enhancement. It also addresses a notable gap in Malaysian transit research, which has thus far lacked the integration of computational spatial analysis into policy-relevant transit evaluation. Moreover, the study contributes to broader national goals of achieving inclusive and resilient transport systems, as articulated under Sustainable Development Goal 11.2, by proposing actionable strategies to enhance mobility for all segments of society.

The structure of the paper is as follows. Section 2 reviews the conceptual foundations and global applications of GTFS and contextualizes the Malaysian research landscape. Section 3 details the data sources, preprocessing workflows, and methodological tools used for network and accessibility analysis. Section 4 presents empirical findings, while Section 5 discusses their planning implications, limitations, and alignment with equity and sustainability goals. Additionally, outlines directions for future research. Section 6 concludes with a synthesis of the study's contributions and limitation. A visual overview of the paper's structure is presented in Figure 1.

2. Literature Review

2.1. Background of GTFS

GTFS was originally developed in the mid-2000s by Google, in collaboration with Portland's TriMet transit agency, to standardize the way public transportation data is shared and utilized across digital platforms. Since its inception, GTFS has evolved into a globally adopted framework for structuring transit schedules, stop information, and route geometries in a machine-readable, comma-separated value (CSV) format [18]. Its primary purpose is to enable interoperability between transit agencies, software developers, researchers, and mobility platforms, thereby streamlining both real-time navigation and long-term service planning across multimodal transport systems [19,20].

Over the past two decades, GTFS has become integral to modern public transit ecosystems. By standardizing the structure of transit datasets, it facilitates seamless integration across a range of applications, including journey planners, accessibility models, routing engines, and intelligent transportation systems (ITS) [21–23].Within the GTFS ecosystem, transit agencies are responsible for producing and maintaining the datasets, software developers utilize these feeds to build user-facing applications such as Google Maps or Moovit, and researchers analyze these structured datasets to evaluate transit performance, identify service gaps, and inform evidence-based policy development [21]. As a result, GTFS has significantly improved the transparency, accessibility, and efficiency of public transit systems worldwide.

The specification is currently distributed in two major variants: GTFS Static and GTFS Real-Time (GTFS-RT). GTFS Static provides detailed data on scheduled transit services, including routes, stop locations, and planned arrival/departure times. This static dataset forms the analytical foundation for most transit accessibility modeling, network optimization, and equity assessment studies [24]. In contrast, GTFS-RT offers live data streams, enabling real-time monitoring of service conditions such as delays, vehicle positions, and service disruptions. Together, these formats form a comprehensive ecosystem for both predictive planning and real-time operations management [25].

GTFS-RT enhances the responsiveness and reliability of transit systems by delivering three critical data feeds: trip updates, which inform passengers of deviations from the schedule such as delays and cancellations; vehicle positions, which provide real-time geolocation of transit vehicles; and service alerts, which notify users of operational disruptions due to road closures, adverse weather, or system outages [25].The integration of GTFS Static and GTFS-RT supports a seamless passenger experience by combining anticipated service plans with real-time updates. This capability is central to the success of intelligent routing applications and mobile trip planners, which rely on GTFS-RT to dynamically recalibrate itineraries and improve commuter satisfaction [26].

The core structure of GTFS datasets is modular and consists of multiple interlinked text files, each defining a specific aspect of the transit system. These files vary in necessity based on their role in service modeling and compliance requirements. Table 1 presents a summary of the primary files that constitute a GTFS dataset, organized by requirement level and functional description.

This standardized structure enables GTFS to serve as a foundational data layer in transit analytics. It supports cross-platform integration and allows researchers and agencies to conduct complex analyses on service frequency, spatial coverage, network redundancy, and pedestrian

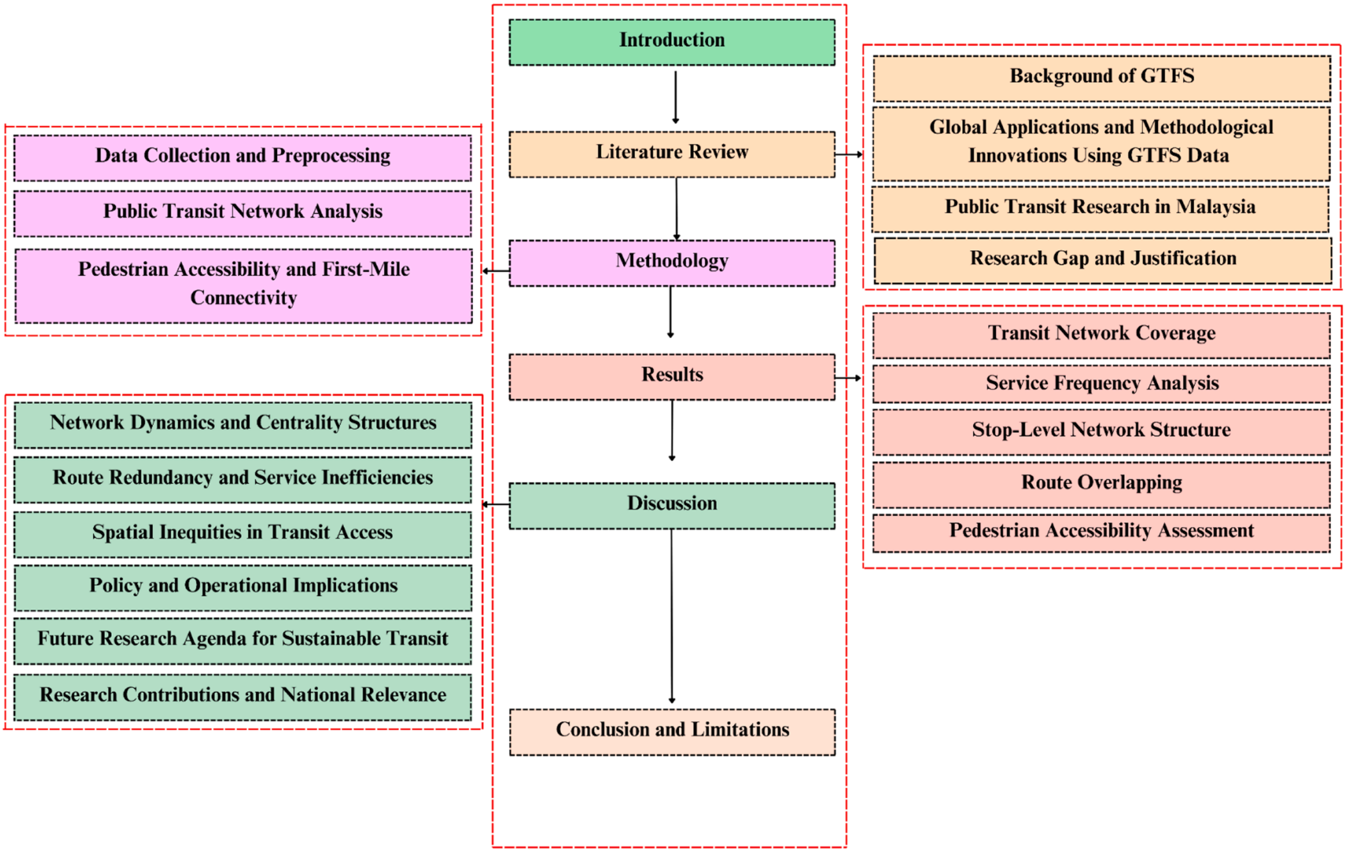


Figure 1. Structural Overview of the Paper.

Table 1
Summary of GTFS Dataset Files Used in the Study.

File Name	Requirement Level	Description
agency.txt	Required	Contains metadata on the transit agency publishing the dataset.
stops.txt	Conditionally Required	Defines geolocated points where passengers board or alight from vehicles.
routes.txt	Required	Lists transit routes, grouping trips that appear as a single public service.
trips.txt	Required	Links routes to scheduled trips, identifying stop sequences and vehicle blocks.
stop_times.txt	Required	Specifies scheduled arrival and departure times at each stop for each trip.
calendar.txt	Conditionally Required	Defines the days of service using date ranges and weekly patterns.
shapes.txt	Optional	Provides detailed spatial geometries for route paths between stops.

accessibility with a high degree of replicability.

To ensure the reliability, usability, and longevity of GTFS datasets, best practices in dataset publication must be rigorously followed. Transit agencies are advised to maintain consistent file formatting using Unicode Transformation Format-8 bits (UTF-8) encoding and to preserve stable unique identifiers across dataset versions to maintain referential integrity [27]. Public hosting of GTFS feeds on accessible servers is also strongly recommended to support transparency and data integration in external applications. For real-time applications, GTFS-RT feeds should accurately reflect the current operational status of transit services, including delays, vehicle locations, and service interruptions. These best practices collectively ensure that GTFS datasets remain robust, interoperable, and conducive to both operational use and scholarly inquiry [28–30]. To further clarify the operational and methodological value of GTFS, Table 2 presents a structured comparison between GTFS-based datasets and the non-standardized data sources commonly used in earlier public transport research. This comparison highlights the technical limitations of non-GTFS data formats and emphasizes the rationale for adopting GTFS in this study.

GTFS-based datasets offer significant advantages in terms of data structure, analytical flexibility, and real-world applicability [31]. In contrast, non-GTFS sources despite their historical role in transit planning lack the standardization and spatial resolution required for advanced modeling [32,33]. This affirms the appropriateness of employing GTFS as the foundational dataset for the study's analytical framework, representing a methodological advancement over previous transit research in Malaysia, which has predominantly relied on unstructured, non-standardized, or manually collected data sources.

In summary, GTFS extends beyond its function as a data format to serve as a globally recognized standard for transport data interoperability. Its bifurcated structure comprising both static scheduling information and real-time operational feeds renders it particularly well-suited for comprehensive, multi-dimensional transit system analysis. As urban mobility systems increasingly adopt integrated and data-driven planning paradigms, GTFS offers the foundational infrastructure necessary for designing, evaluating, and iteratively refining modern, intelligent, and equitable public transportation networks.

2.2. *Global Applications and Methodological Innovations Using GTFS Data*

Over the past decade, GTFS has emerged as a foundational data standard in global transit research, evolving from a static trip-planning tool into a comprehensive platform for analyzing public transport systems. Its standardized format encompassing spatial, temporal, and structural attributes of transit systems has enabled researchers and

Table 2
Comparison of GTFS-Based vs. Non-GTFS Transit Data for Public Transport Evaluation.

Dimension	GTFS-Based Data	Non-GTFS Data
Structure	Structured, machine-readable (CSV text files conforming to GTFS schema)	Unstructured or semi-structured (e.g., PDF timetables, printed route maps, manual logs)
Standardization	Globally standardized format developed and maintained by Google/OpenTransit	Agency-specific or project-specific formats, lacking global interoperability
Spatial Integration	Includes coordinates for stops and route geometries; compatible with GIS	Often lacks precise geospatial referencing; requires manual geocoding
Temporal Resolution	Encodes scheduled departure/arrival times, frequencies, and service calendars	Typically includes basic departure times without consistent time schemas or recurrence patterns
Multimodal Modeling	Supports multi-modal systems	Typically segmented by mode and not easily integrated
Real-Time Capability	Supports GTFS-RT: trip updates, vehicle positions, service alerts	Does not support real-time integration or passenger information systems
Reproducibility	Enables replicable analysis across time and geography	Difficult to reproduce due to inconsistency, formatting errors, and data inaccessibility
Usability in Tools	Compatible with software platforms like OpenTripPlanner, Conveyal, QGIS, R, Python	Requires significant preprocessing; limited support in analytical toolkits
Public Accessibility	Usually hosted openly online (e.g., transit agency websites, GTFS Data Exchange)	Often restricted or unpublished; if available, shared only as PDFs or spreadsheets
Suitability for Research	High: Enables accessibility modeling, network diagnostics, equity analysis	Low to moderate: Limits analytical scope and inhibits data-driven planning

practitioners to conduct high-resolution assessments of service availability, travel times, network structure, and real-time system dynamics.

A series of landmark studies have significantly extended the analytical frontiers of GTFS. Wong [34] was among the first to demonstrate GTFS's utility in system-wide analysis across 50 North American agencies. By automating stop- and route-level performance evaluations, Wong's work laid the foundation for reproducible, scalable transit diagnostics. Building on this, Carleton et al. [35] address a significant gap in public transit data infrastructure by introducing GTFS-Ride, a standardized data schema for capturing fixed-route ridership. While GTFS has become a global standard for transit schedule and geographic information, it lacks a formal mechanism for incorporating detailed ridership data. GTFS-Ride fills this void by structuring ridership records in a way that directly references the state of the transit system at the time data were collected, thereby enhancing the longitudinal validity of service evaluations. The study details the schema's five constituent files and documents pilot implementations in three Oregon transit agencies, illustrating its applicability across systems of varying complexity. Importantly, GTFS-Ride is designed for both small and large agencies, thereby promoting data standardization and interoperability across heterogeneous institutional contexts. This contribution is particularly relevant for researchers and practitioners interested in comparative transit analytics, demand modeling, and performance evaluation . Similarly, Eros et al. [36] demonstrated GTFS's adaptability to fragmented and informal transport environments in the Global South, successfully generating a standardized GTFS feed for Mexico City's semi-structured transit network. Their schema modifications enabled integration with open-source tools, providing access to analytics in contexts previously excluded due to data scarcity.

GTFS-RT data has been increasingly used for urban traffic monitoring. Jiang et al. [37] leveraged GTFS RT feeds to estimate segment-level traffic speeds across an urban network, proposing novel algorithms to extract clean movement patterns from noisy bus operations by filtering out boarding/alighting events. Their validation, using Bluetooth detectors, revealed strong correlation between GTFS-derived speeds and actual traffic flow, highlighting the viability of transit vehicles as real-time traffic probes. Elliott and Lumley [29] proposed a real-time road state estimation framework by integrating GTFS-defined transit paths with vehicle GPS data using a particle filter. Their model infers dynamic traffic conditions across road segments and provides probabilistic travel time estimates, advancing GTFS usage into the domain of real-time Bayesian state estimation for transit operations. Recent work by Santhosh Kumar et al. [38] utilized GTFS-RT data to evaluate bus travel time variability in Delhi, India, revealing weekday and hourly reliability patterns. Their findings underline GTFS's utility in diagnosing temporal inefficiencies and optimizing scheduling in high-density urban contexts.

In parallel, Bannur et al. [39] recast urban congestion prediction as a machine learning classification task using GTFS-derived features. Their comparative evaluation revealed that traditional decision tree models outperformed more complex deep learning architectures, such as RNNs, in both interpretability and predictive performance. This indicates GTFS's viability in supporting non-intrusive, intelligent congestion diagnostics. GTFS is increasingly employed in operational and predictive frameworks. For example, Zhou et al. [40] addressed the implications of latency in GTFS-RT feeds by integrating machine learning models to accurately predict current bus locations and dwell times, demonstrating improved alignment with transit signal priority systems. Such approaches enhance real-time responsiveness using GTFS metadata as a structural baseline. Further advancing temporal diagnostics, Sim and Cho [41] integrated GTFS with real-time APTS data to evaluate dynamic accessibility under both routine and adverse traffic conditions in Busan, South Korea. Their introduction of Standardized Integral Accessibility (SIA) and Rate of Integral Accessibility Change (RIA) enabled quantification of spatiotemporal variability due to both recurring congestion and unexpected weather disruptions. This study highlights the expanding utility of GTFS in resilience analytics and climate-adaptive transit planning.

An expanding body of research utilizes GTFS for bus and rail delay prediction through spatiotemporal modeling and machine learning. Gaudette et al. [42] illustrated the integration of GTFS schedule data with tap-in-only smart card data to build high-resolution microsimulation models using TRANSIMS. By inferring boarding and alighting locations algorithmically, they extended the analytical scope of GTFS toward operational modeling, offering a robust approach for simulating transit demand and evaluating network interventions with fine temporal and spatial granularity. Vijaya et al. [43] benchmarked five supervised models including SVM, Gradient Boosting, and Neural Networks for delay prediction using both GTFS Static and Real-Time feeds. The study identified SVM as the most accurate method, highlighting GTFS's potential for driving reliable and interpretable performance forecasting in public transit systems. Moreover, Chondrodima et al. [44] benchmarked multiple machine learning models for GTFS-based bus arrival prediction and found that neural network approaches consistently outperformed traditional algorithms across multiple urban datasets. Their work established a foundational PT-ATP benchmarking framework for GTFS users, advancing the methodological robustness of AI-driven transit analytics. Aemmer et al. [45] explored the generalization of bus travel time prediction models across global transit networks by leveraging GTFS-Static, GTFS-RT, and OpenStreetMap data. Their findings emphasize the potential of deep learning to transfer knowledge between cities with minimal fine-tuning, thus promoting scalable and transferable GTFS-based analytics for urban mobility systems. Wu et al. [18] developed a GTFS processing and fusion framework that supports multivariate multistep delay prediction using Long Short-Term Memory (LSTM) networks. It Applied to the Sydney train network, their model successfully anticipated minor anomalies in schedule adherence, highlighting the utility of deep learning in GTFS-based predictive analytics. Moreover, the study addressed a key barrier to broader GTFS adoption by proposing a standardized, researcher-friendly data transformation pipeline. Meanwhile, Lopes et al. [46] advanced spatiotemporal prediction by integrating Graph Neural Networks with LSTM to forecast bus arrival times using both GTFS Static and Realtime data. Their model outperformed traditional baselines by structurally encoding network topology and leveraging dynamic updates, illustrating the emergent power of deep learning architectures in transit informatics. Notably, Iovino et al. [47] advanced the frontier of GTFS data availability by applying recurrent neural networks to reconstruct GTFS-compliant transit graphs from raw GPS traces in the city of L'Aquila, Italy. Their approach offers a scalable solution for regions lacking formal GTFS feeds, facilitating broader data-driven transit analysis. To improve the replicability of accessibility metrics under uncertain conditions, Javanmard et al. [48] introduced a standardized framework leveraging GTFS-RT data to produce dispersion- and median-corrected accessibility measures. Their findings indicate that relying solely on scheduled data significantly overestimates accessibility and masks equity outcomes in healthcare access evaluations. Li and Fan [49] developed a Transit Gap Index (TGI) that fuses GTFS-derived service attributes demographic data to map and quantify spatial inequities in Charlotte's public transit system. Their GIS-based analysis demonstrates GTFS's utility in not only capturing supply-side patterns but also guiding policy interventions for equitable service distribution, thus aligning GTFS with principles of social justice and inclusive urban planning.

In a recent study, Wu et al. [50] introduced a Primary Delay Prediction System (PDPS) that integrates GTFS data, Critical Point Search (CPS), and a ConvLSTM Encoder-Decoder model for predicting primary delays in urban railway networks. Their work demonstrates the potential of data fusion and spatiotemporal deep learning in building scalable, generalizable frameworks for predictive rail operations. Complementing GTFS-based analyses, Ceccato et al. [51] developed a dual-indicator framework that compares traveler-perceived accessibility with scheduled transit service in Turin, Italy. Their findings underline the divergence between 'active' accessibility based on user-reported travel times and 'passive' accessibility based on GTFS or operator schedules, highlighting the value of integrating subjective mobility experiences in urban accessibility diagnostics. This reinforces the importance of contextualizing GTFS-derived measures within the broader experiential landscape of urban transport systems. In the area of route planning, Vágner [52] proposed a route planning framework based on Neo4j, a graph database optimized for path-finding and topological analytics. By modeling GTFS data as a navigable graph structure, the study illustrated how native algorithms such as Dijkstra and A* can be applied to compute efficient transit itineraries, thereby expanding GTFS application into graph-native database environments.

GTFS has also been instrumental in detecting system anomalies and reliability deviations. Wessel and Widener [53] pioneered the use of GTFS and GTFS-RT data to diagnose schedule padding and stochastic delay at the stop-segment level, revealing spatial mismatches between planned and experienced travel times. Their method provides transit agencies with actionable insights into where operational reforms could reduce inefficiencies without compromising reliability.

Boudabous et al. [54] conducted a detailed deviation analysis in Montreal by comparing GTFS planned and real-time feeds to classify delays as systematic or stochastic. Their study highlights the role of machine learning in predicting stochastic disruptions and highlights the operational value of GTFS data in quantifying schedule reliability from both spatial and temporal perspectives. Aemmer et al. [25] introduced a network-wide diagnostic framework to classify transit delays from GTFS-RT feeds as either systematic or stochastic, enabling proactive and reactive planning interventions and infrastructure optimization. Visual analytics tools such as *PubtraVis* [19] utilizing GTFS Static data to assess transit system dynamics in Calgary across multiple dimensions including mobility, headway, speed, density, and flow. The tool enabled both spatial and temporal visualization, offering insights into station-level performance, route similarity via cosine metrics, and service clustering using algorithms. These functionalities illustrate the potential of GTFS data to go beyond static mapping and support dynamic transit diagnostics, planning scenario evaluation, and agency-level decision support. When coupled with spatiotemporal clustering and comparative analytics, such GTFS-enabled systems significantly extend the analytical value of transit data infrastructures. Chondrodima et al. [31] developed a neural prediction framework combining a CR-GTFS preprocessing pipeline with PSO-optimized RBF neural networks, significantly improving public transport ETA predictions over existing methods.

Recent methodological innovations have focused on data transformation and user interaction. Vuurstaek et al. [55] proposed an automated mapping algorithm that aligns GTFS stops with the OSM network, enabling network-consistent simulation and accurate estimation of transit-based travel distances for all trips in a feed. Godfrid et al. [56] demonstrated the integration of GTFS-Static and GTFS-RT data within MobilityDB, a moving object database built on PostgreSQL/PostGIS, enabling scalable spatiotemporal queries and visualizations. Their case study of Buenos Aires represents the first documented application of moving object databases to GTFS, unlocking new capabilities for trajectory analysis and real-time network diagnostics.

Large language models and natural language interfaces constitute a transformative advancement in enhancing system accessibility and user interaction. Devunuri and Lehe [57] introduce *TransitGPT*, a generative AI-based framework that significantly advances human-computer interaction with GTFS data by enabling natural language querying. Recognizing the technical barriers faced by non-expert users in accessing and manipulating transit datasets. Additionally, Devunuri et al. [58] further evaluated the semantic comprehension and query accuracy of GPT-based LLMs against GTFS feeds, revealing high accuracy for simple queries and notable limitations under logical complexity. While such AI-driven approaches are still nascent, they represent a promising frontier for user-friendly, automated transit analytics built on top of the GTFS standard.

GTFS data now supports environmental diagnostics and resilience analytics. Vieira et al. [59] introduced *gtfs2emis*, a bottom-up emissions modeling tool that combines GTFS feeds with fleet-level data to estimate pollutants and energy use across spatial and temporal scales. It Demonstrated on São Paulo's public transit system, the tool enables scalable, policy-relevant climate diagnostics, illustrating GTFS's growing utility in sustainable transport assessment. Nishino et al. [60] extended the application of GTFS-RT data to the disaster resilience domain by developing a GIS-based model that estimates the spatial coverage of audible disaster warnings disseminated via public transit vehicles. Their method demonstrates the potential of GTFS as a proxy for mobile infrastructure in emergency response, offering a novel intersection of transit analytics and hydrometeorological hazard management. Hoover et al. [61] utilized GTFS-Ride and Automated Passenger Counting (APC) data to identify transfer opportunities beyond formal hubs. Their novel transfer metric combined schedule adherence, spatial coverage gain and observed rider behavior to uncover latent transfer patterns within a mid-sized Oregon transit system. This approach enables smaller agencies to conduct advanced transfer analysis using open standards, reinforcing the adaptability of GTFS-derived metrics in both structural and user-oriented planning contexts.

While GTFS has enabled an expansive array of applications globally, a foundational and enduring area of use remains the evaluation of network efficiency and accessibility using GTFS Static data the core focus of this study.

GTFS facilitates disaggregated assessments of transit coverage by integrating spatial (stop locations), temporal (headways), and structural (routes) data [49,62,63]. These elements support accessibility modeling by quantifying how easily users can access key destinations such as employment, healthcare, and education [64,65].Travel time to the nearest stop, computed through GIS-based routing engines and enhanced with friction surfaces, serves as a proxy for pedestrian connectivity. Isochrone-based surfaces visualize reachability within given time thresholds, aiding in the identification of underserved zones.

Recent work has introduced spatiotemporal accessibility models, such as travel time cubes, which account for fluctuations in service frequency throughout the day, offering more realistic assessments [48, 49]. These tools have become essential in scenario testing under urban expansion and climate stress. GTFS has also enabled the application of complex network theory in transit research. Node-link representations of transit systems allow for the computation of graph-theoretic indicators such as degree, betweenness, and closeness centrality, which reveal the structural properties and functional vulnerabilities of transit networks [23,66].When combined with accessibility metrics, these tools help diagnose equity deficits and operational bottlenecks.

Furthermore, a growing body of work focuses on transport equity, particularly how public transit systems serve low-income, elderly, or mobility-impaired populations [21,23,64,67]. Spatial analyses such as Kernel Density Estimation (KDE) are frequently applied to visualize transit supply disparities, guide service redistribution, and identify transit deserts or overlaps [38]. Software ecosystems such as OpenTripPlanner, Conveyal, and Python/R-based analytical libraries have significantly expanded access to GTFS-based transit analytics, enabling broader participation in data-driven transport planning.These platforms support multimodal trip planning, dynamic equity audits, and robust policy simulations [68,69].

In the Malaysian context, the public release of GTFS Static feeds by operators such as myBAS Johor and Rapid Penang marks a critical shift toward empirical diagnostics. GTFS now enables Malaysian agencies to move beyond perception-based evaluation and implement spatially rich, data-driven methodologies for service planning. As Malaysia confronts urban sprawl, traffic congestion, and modal fragmentation, GTFS-based assessments are particularly suited to guiding equitable investments and adaptive service designs [70]. These applications directly align with Sustainable Development Goal 11.2, which calls for universal access to safe, affordable, and accessible transport for all. To provide a synthesized overview of the wide-ranging and evolving landscape of GTFS applications in contemporary global transit research, Table 3 provides a synthesized matrix of recent peer-reviewed studies, offering a structured inventory of methodological frameworks, data types utilized, key contributions, and noted limitations. The studies span a diverse range of research domains, including traffic monitoring, delay prediction, machine learning benchmarking, accessibility assessment, spatiotemporal analytics, environmental impact evaluation, and real-time anomaly detection. This compilation illustrates the methodological breadth and evolving scope of GTFS-based transit research, underscoring its significance as a foundational tool in contemporary transportation analytics. Collectively, these studies exemplify the global scope and methodological diversity of GTFS-based research.

In sum, while global research has demonstrated GTFS's power in traffic prediction, emissions modeling, real-time diagnostics, and AI-based automation, the specification's foundational value remains in its ability to critically assess network structure and accessibility. This study builds upon that tradition by leveraging GTFS Static data to examine service coverage, spatial equity, and structural connectivity—offering grounded, scalable insights that complement and contribute to the global GTFS research landscape.

2.3. Public Transit Research in Malaysia

The academic discourse on public transit in Malaysia remains largely descriptive, fragmented, and underpinned by perception- or policy-based methodologies, offering limited analytical leverage for diagnosing and optimizing transit systems at the network scale. A comprehensive review of the literature reveals that although several studies have investigated commuter behavior, service satisfaction, infrastructure integration, and institutional planning, none have leveraged computational, spatial, or reproducible diagnostics grounded in standardized datasets such as GTFS. This absence is not trivial—it reflects a deeper gap in the application of advanced spatial science, network theory, and transport informatics within the Malaysian urban mobility context.

The dominant paradigm in Malaysian transit research is commuter perception analysis. For instance, Bachok et al. [71] used survey-based tools to examine user satisfaction with KTM Komuter services, focusing on qualitative dimensions such as ticketing experience, punctuality, and seating comfort. Similarly, Tomas and Rahman [72] assessed user perception in Batu Pahat through structured surveys and field observations, highlighting issues such as affordability and schedule reliability. While valuable for gauging public sentiment, these studies provide no empirical assessment of transit system structure, route efficiency, spatial equity, or inter-modal connectivity. Critically, none utilize standardized transit datasets or spatial analytics, precluding comparative replication or integration with planning models.

Several scholars have turned their attention to institutional and strategic planning issues, focusing on governance, integration, and infrastructure development. Yusoff et al. [73] examined Malaysia's transport policy framework in the Klang Valley, identifying institutional fragmentation and financing inefficiencies as barriers to sustainability. Masirin et al. [9] critiqued the management of rail systems and their operational disjunctions, comparing Malaysia with advanced transit systems in Japan and the UK. While these studies contribute to policy understanding, they lack analytical rigor. They are not supported by spatial performance metrics, transit service models, or empirical accessibility data.

Other studies have focused on infrastructure design, multimodal integration, and transit-oriented development (TOD). Ramlan et al. [74] evaluated land use composition within 800 meters of MRT Putrajaya stations, benchmarking against Singapore's TOD model, while Bachok et al. [75] proposed sustainability indicators for Malaysian transport systems based on expert consensus.

A notable exception is the study by Ding et al. [76], which applies complex network theory to evaluate the robustness of Kuala Lumpur's urban rail network. Using metrics such as degree, betweenness, and clustering coefficient, the study simulates network growth and cascading failures. While conceptually robust, the approach remains narrowly focused on rail-based systems and relies on manually curated datasets in the absence of GTFS or standardized inputs, thereby constraining both its reproducibility and broader applicability across diverse transit contexts. Moreover, it neglects key performance dimensions, including spatial accessibility, network redundancy, service frequency, and integration with pedestrian infrastructure.

Behavioral and design-oriented contributions further illustrate the thematic breadth but analytical limitations of the field. These studies provide important user-centered perspectives on the transit experience but remain disconnected from network-level diagnostics or equity-based planning [77]. Likewise, qualitative works such as those by Borhan et al. [78], Ustadi and Shopi [33] identify service barriers and infrastructure weaknesses but do not model these spatially or quantitatively.

Overall, this review shows a significant methodological gap: the absence of standardized, scalable, and spatially integrated approach for evaluating Malaysia's urban bus transit systems. Specifically, no existing study GTFS-based or otherwise—has undertaken a multi-scalar analysis that simultaneously incorporates service frequency, spatial accessibility, pedestrian connectivity, and network redundancy. This substantive shortfall in both methodological approach and analytical scope directly informs the rationale for the present study.

2.4. Research Gap and Justification

The existing body of Malaysian transit research, as demonstrated in sub-section 2.3, fails to engage with the methodological paradigm shifts

Table 3
Overview of Global GTFS Applications in Transit Research.

Study	Focus Area	Data Type	Methods/Models	Key Contribution & Limitation
[29]	Estimating real-time road traffic conditions from bus location data	GTFS Static, Real-time GPS (Auckland)	Particle filter, Real-time traffic state estimation, R + C++ implementation	Developed a generalizable framework using particle filters to model road state and inform transit travel time prediction. Limitation: focused on initial traffic state modeling; arrival time prediction reserved for future work.
[37]	Estimating space mean speeds and enabling citywide traffic monitoring using GTFS Realtime	GTFS RT, Bluetooth validation data	Segment-trip extraction, Segment speed estimation, Buffer zone correction	Demonstrated GTFS Realtime's validity for urban traffic speed estimation. Provided a privacy-preserving, scalable, and low-cost alternative to traditional ITS data. Limitation: validation limited to Gainesville, FL context.
[25]	Classification of systematic and stochastic transit delays using GTFS-RT	GTFS-RT (King County Metro, Seattle)	Delay classification framework, Segment-level aggregation, Interactive visualization	Proposed a generalizable framework to classify and map delay types from GTFS-RT data. Enabled proactive vs. reactive intervention planning. Limitation: relies on GTFS-RT availability and coverage fidelity.
[40]	Real-time transit prediction and GTFS-RT integration	GTFS-RT, historical GPS, GTFS schedules	Machine Learning (RNN, LSTM), Spatiotemporal Outlier Analysis	Improved prediction of bus dwell times and locations; supports real-time transit signal priority. Limitation: operational orientation; does not address structural/ systemic evaluation.
[19]	Interactive transit system visualization using GTFS Static	GTFS Static	Visualization platform (PubtraVis), Cosine similarity, X-means clustering, Temporal-Spatial analytics	Developed a dynamic platform for visualizing transit system metrics (mobility, speed, headway, etc.). Enabled agency-level decision support. Limitation: emphasizes dynamic diagnostics; our study provides structural analysis of network efficiency.
[51]	Urban accessibility analysis based on perceived and actual travel experiences	Travel diary survey (2016–2017, Turin), Transport network data	Travel time and distance-based accessibility indices, Active vs. Passive accessibility comparison	Assessed intra-city accessibility using user-perceived and scheduled service data. Highlighted mode quality and travel perception effects. Limitation: context-specific to Turin; based on historical survey data rather than real-time feeds.
[61]	Identification of transfer opportunities in mid-sized transit networks	GTFS-Ride, GTFS Schedule, Automated Passenger Counter (APC) data, Rider survey	Statistical transfer metric, Probability models (transfer time/distance), Coverage gain measure	Developed a transfer metric using GTFS-Ride to uncover latent transfer locations. Aided network planning for small-to-mid-size agencies. Limitation: dependent on availability of GTFS-Ride and APC integration.
[58]	Evaluating LLMs' ability to interpret and retrieve transit data from GTFS	GTFS (CTA subset), GPT-3.5 Turbo, GPT-4	Benchmarking (GTFS Semantics & Retrieval), Prompt Engineering, Zero-shot, One-shot, Program Synthesis	Introduced a novel evaluation of LLMs for transit data interpretation using GTFS. Demonstrated potential of LLMs in GTFS query automation. Limitation: accuracy drops in complex queries and ambiguous multiple-choice conditions.
[31]	Accurate ETA prediction for public transit using GTFS	GTFS (real-world transit data)	CR-GTFS pre-processing pipeline, RBF Neural Networks, PSO-NSFM optimization	Introduced a robust ETA prediction framework combining PSO-enhanced RBF networks with optimized GTFS data pre-processing. Outperformed state-of-the-art methods. Limitation: may require significant computational resources for training.
[18]	GTFS data processing and train delay prediction	GTFS Static and Realtime, Train schedule data (Sydney)	GTFS data acquisition and fusion framework, Multivariate Multistep LSTM	Proposed a data fusion framework to convert raw GTFS data into analysis-ready format; applied LSTM for delay prediction in Sydney transit network. Limitation: Framework complexity may limit adoption without technical expertise.
[41]	Temporal variability of public transit accessibility under real traffic conditions	GTFS, APTS (Busan, South Korea), Weather data	Door-to-door travel time estimation, SIA, RIA	Developed new metrics (SIA and RIA) to quantify accessibility fluctuations due to recurring and non-recurring congestion (e.g., heavy rain). Limitation: local case study; may require extensive sensor integration for broader applicability.
[44]	Arrival time prediction using GTFS and machine learning models	GTFS Static and Realtime (real-world datasets)	Comparative analysis of ML models for PT-ATP; Neural Networks	Proposed a framework for PT arrival time prediction using GTFS; found neural networks to be most accurate among tested ML models. Limitation: general performance comparisons; lacks detail on operational constraints or integration challenges.
[54]	Analysis and classification of bus arrival deviations in Montreal	GTFS Realtime and Scheduled data (Montreal)	Deviation classification (systematic vs. stochastic), ML models for stochastic delay prediction	Identified and categorized delay types using GTFS data; highlighted impact on commuter experience and system reliability. Limitation: preliminary study; deeper predictive modeling and broader validation are future work.
[46]	Bus arrival time prediction integrating static and dynamic data	GTFS Static, GTFS-RT, vehicle position updates	GNN for spatial features, LSTM for temporal modeling, GNN+LSTM fusion	Developed a hybrid GNN+LSTM model that outperforms standalone LSTM and static baselines in bus ETA prediction. Limitation: practical implementation complexity; challenges in aligning GTFS and GTFS-RT identifiers noted.
[59]	Estimation of emissions from public transport using GTFS	GTFS Static, Vehicle fleet characteristics	gtfs2emis R package, Emission factor models (EU, US, Brazil)	Developed an open-source tool (gtfs2emis) for estimating multi-pollutant emissions using GTFS data and fleet profiles. Limitation: accuracy depends on availability and quality of fleet data inputs.
[39]	Traffic congestion prediction using GTFS data in urban contexts	GTFS Static, derived traffic attributes	Decision Trees, RNNs, model comparison (incl. ARIMA, Hybrid DNNs), classification framing	Reframed traffic congestion prediction as a classification problem using GTFS data; Decision Tree model achieved 81% accuracy. Limitation: lacks real-time validation and focuses primarily on model benchmarking.

(*continued on next page*)

Table 3 (*continued*)

Study	Focus Area	Data Type	Methods/Models	Key Contribution & Limitation
[45]	Cross-network generalization of bus travel time prediction models	GTFS Static, GTFS-RT, GPS, OpenStreetMap (Seattle, Trondheim, 33 cities)	Deep learning models, Heuristic benchmarks, Feature ablation, Fine-tuning strategies	Developed generalizable models for bus travel time prediction using standardized GTFS and OSM data. Demonstrated high intra-network accuracy and improved inter-network transferability with fine-tuning. Limitation: Requires GTFS-RT and geospatial data integration.
[56]	Spatial-temporal analysis of GTFS data using a moving object database	GTFS Static and GTFS Realtime (Buenos Aires)	MobilityDB (PostgreSQL/PostGIS extension), Spatio-temporal functions, Visual analytics	First application of MobilityDB for analyzing GTFS data; enabled complex spatio-temporal queries and visualizations. Limitation: complex setup and need for ad hoc data loading procedures.
[60]	Estimating dynamic spatial coverage of mobile disaster sirens using transit data	GTFS Realtime, Hazard maps, GIS layers	GIS-based spatial analysis, Bus location tracking, Population exposure modeling	Developed a general-purpose model to estimate real-time coverage of audible disaster warnings from mobile sources (e.g., buses). Demonstrated high applicability using Brisbane flood scenario. Limitation: assumes accuracy and availability of real-time GTFS data during emergencies.
[52]	Application of graph databases for GTFS-based route planning	GTFS Static (routes, trips, stops)	Graph database modeling, Neo4j algorithms, NoSQL structure design	Demonstrated the feasibility of storing and querying GTFS data using Neo4j for efficient route planning. Limitation: focuses on static GTFS data; lacks consideration of real-time dynamics or multimodal integration.
[43]	Comparative machine learning approach to delay prediction using GTFS	GTFS Static and Realtime	Gradient Boosting, Random Forest, SVM, Neural Networks, kNN	Evaluated five ML models for transit delay prediction; SVM achieved best performance (lowest MAE and RMSE). Emphasized balance between accuracy and interpretability. Limitation: model performance varies with data complexity and generalizability.
[38]	Analysis of travel time variability (TTV) and reliability in urban bus systems	GTFS-RT (Delhi, 1-month dataset)	Temporal statistical analysis by day of week and time of day	Quantified variability in travel times using GTFS-RT; showed that TTV varies by weekday and time-of-day. Useful for reliability improvements and planning. Limitation: focused on short-term, city-specific dataset.
[55]	Automated integration of GTFS stop locations with OpenStreetMap network	GTFS Static, OpenStreetMap (OSM)	Bus stop mapping algorithm, Distance minimization heuristic, Perturbation analysis	Developed an automated algorithm to map GTFS bus stops onto OSM road networks. Validated with speed consistency and perturbation robustness. Limitation: Sensitive to inaccuracies in both GTFS and OSM datasets.
[53]	Detection and spatial analysis of schedule padding and stochastic delay	GTFS Scheduled, GTFS Realtime (Toronto)	Stop-to-stop segment analysis, Spatiotemporal comparison of schedule vs. actual	Identified spatial and temporal distribution of schedule padding and delays in a major transit system. Aimed to inform targeted interventions. Limitation: exploratory case study with focus on a single agency; generalizability untested.
[42]	Development of detailed bus network microsimulation using GTFS and smart card data	GTFS Static, Tap-in-only Smart Card Data (no alighting info)	TRANSIMS microsimulation, Boarding/alighting inference algorithms	Demonstrated feasibility of integrating GTFS with smart card data for high-resolution bus network simulation. Enabled precise transit planning. Limitation: tap-in-only data required inference methods, introducing potential uncertainty.
[49]	Assessment of transit equity and accessibility using GTFS and GIS	GTFS Static, Demographic data (Charlotte, NC)	TGI, GIS-based spatial analysis, Sensitivity analysis	Developed a GIS-based TGI incorporating spatial, temporal, and demographic factors to evaluate equity in transit service. Identified suburban transit gaps. Limitation: Single-city case study; TGI depends on data granularity and local context.
[58]	Evaluating LLMs' ability to interpret and retrieve transit data from GTFS	GTFS (CTA subset), GPT-3.5 Turbo, GPT-4	Benchmarking (GTFS Semantics & Retrieval), Prompt Engineering, Zero-shot, One-shot, Program Synthesis	Introduced a novel evaluation of LLMs for transit data interpretation using GTFS. Demonstrated potential of LLMs in GTFS query automation. Limitation: accuracy drops in complex queries and ambiguous multiple-choice conditions.
[36]	Adoption and adaptation of GTFS in developing cities with semi-structured transit	GTFS Static (Mexico City)	Case study analysis, Open data ecosystem evaluation, Format adaptation	Demonstrated feasibility of rapid GTFS deployment in complex megacity transit systems. Identified limitations of standard GTFS in accommodating informal transit modes common in the Global South. Highlighted need for adaptive data models.
[35]	Development of a standardized format for fixed-route ridership data	GTFS-Ride (extension of GTFS), Pilot agency data (Oregon)	Schema development, File structure specification, Pilot implementation analysis	Introduced GTFS-Ride as a scalable, standardized format for ridership data linked to GTFS static feeds. Supported usability across agencies of varying size. Limitation: early-stage adoption; reliant on widespread institutional uptake.
[34]	Demonstrating GTFS utility for operational and comparative transit analysis	GTFS Static (50 North American agencies)	Scripting, Database querying, Stop/route/system-level analytics	Showcased GTFS as a scalable and efficient source for system-level transit analyses. Developed open-source tools to aid in stop-, route-, and agency-level evaluation. Limitation: Focus on static GTFS; real-time and ridership extensions not addressed.
[62]	National-level assessment of GTFS data coverage and transit availability	GTFS Static, Census data (Canada), 213 Census Subdivisions	Systematic data collection, Spatial density benchmarking, Province-wise comparison	Developed a comprehensive GTFS inventory for Canadian cities >20K population. Identified regional disparities (e.g., underrepresentation in Québec). Limitation: availability bias; some regions lack official GTFS feeds despite having services.

(*continued on next page*)

Table 3 (*continued*)

Study	Focus Area	Data Type	Methods/Models	Key Contribution & Limitation
[64]	Integration of accessibility into federally required equity assessments for transit	GTFS Static, US Census LEHD data (USA)	Route-level accessibility metrics, Equity analysis framework	Proposed a framework combining GTFS and census data to improve FTA-mandated transit equity assessments. Demonstrated that multiple metrics are needed to avoid oversimplification. Limitation: U.S.-specific context may limit direct transferability elsewhere.
[69]	Spatiotemporal trajectory analysis of public transit using GTFS	GTFS Static	{gtfs2gps} R package, Parallel computing, Time geography analysis, Visualization	Introduced an open-source R tool to convert GTFS into GPS-like trajectories for high-resolution space–time analysis. Enabled analysis of transit access inequalities. Limitation: based solely on static GTFS; does not integrate real-time dynamics.
[67]	Evaluation of GTFS-based accessibility accuracy versus actual transit performance	GTFS Static, Automatic Vehicle Location (AVL) data (4 NA cities)	Comparative analysis of scheduled vs. actual travel times, Spatial-temporal accessibility metrics	Found that GTFS-based schedule data often overestimates accessibility, particularly in low-access zones, by 5–15%. Highlighted spatial patterns of inaccuracy. Limitation: focuses on North American systems; calls for enhanced methodologies using real-time data.
[21]	Multimodal and policy applications	GTFS Static	Review of best practices and tools	Summarized GTFS use cases in trip planning, mobile apps, IVR; promoted open-data adoption.
[23]	Graph-based analysis and comparison of transit networks using GTFS	GTFS Static	Time-expanded graph modeling, Graph database, Shortest path algorithms	Introduced a graph-oriented tool to characterize and compare transit networks using GTFS. Applied shortest-path indicators via a graph database. Limitation: current implementation requires optimization for large and complex networks.
[48]	Generating reproducible transit accessibility metrics under travel time uncertainty	GTFS Scheduled, GTFS Realtime (Columbus, Ohio)	Realtime P50 and P85 GTFS corrections, Median- and Dispersion-Corrected Accessibility Metrics	Developed a standardized framework for uncertainty-aware accessibility analysis. Found scheduled GTFS overestimates healthcare accessibility by ∼11%. Limitation: case study limited to one city; method dependent on availability of GTFS-RT feeds.
[57]	Natural language interaction with GTFS data via generative AI	GTFS Static (user-provided), LLMs (GPT-4o, Claude 3.5 Sonnet)	Prompt-based code generation, Natural language interface, Interactive data querying tool	Introduced TransitGPT, a framework enabling users to query GTFS data via natural language. Eliminates need for GTFS or coding expertise. Limitation: relies on prompt quality; no direct LLM access to live GTFS feeds; scalability in production remains to be tested.

that now define contemporary public transport evaluation globally namely, the transition from static, descriptive models to data-integrated, spatially grounded, and computationally reproducible diagnostics. This gap is not merely contextual (i.e., the absence of GTFS usage in Malaysia) but fundamentally methodological, and it significantly undermines the capacity of current research to inform efficient, equitable, and sustainable transit planning.

The use of GTFS data in our study is not presented as a novelty in itself. Rather, it is employed as the enabling data infrastructure for a reproducible, multi-dimensional evaluation approach one that allows the structured analysis of spatial equity, pedestrian access, service efficiency, and connectivity. Although transit evaluations can be conducted using manual or proprietary datasets, such approaches suffer from three critical limitations: (1) lack of public accessibility and transparency, (2) incompatibility with automated spatial analysis tools, and (3) poor generalizability across urban contexts. GTFS, by contrast, supports temporal modeling, route optimization, and multi-city comparative analysis, making it a foundational tool for open, interoperable transit science.

Beyond GTFS, our study incorporates walking friction surfaces, OSM, POIs, and population rasters to analysis a comprehensive transit accessibility. The inclusion of these layers enables us to measure not just route availability but access to opportunity including schools, hospitals, commercial zones, and transit terminals under realistic pedestrian travel conditions. Methodologically, our study applies graph-theoretic centrality metrics (degree, closeness, and betweenness) to identify structurally critical transit nodes and corridors. It also employs KDE to assess spatial service intensity and distribution uniformity and utilizes the Jaccard Similarity Index to quantify route redundancy and evaluate network efficiency. Raster-based pedestrian accessibility modeling, using global travel friction surfaces, is incorporated to evaluate first- and last-mile connectivity under realistic walking conditions. These tools are not simply imported from other contexts they are tailored and validated for Malaysia's emerging data landscape, characterized by the recent availability of GTFS feeds and the absence of real-time multimodal APIs. By demonstrating that robust transit evaluations can be conducted even in data-constrained environments, this study offers a globally transferable workflow for cities in the Global South with limited data infrastructure. This approach enables these contexts to modernize their mobility systems without dependence on costly proprietary platforms, thereby promoting scalable, open-source, and evidence-based planning practices.

Furthermore, our dual-city analysis of Penang and Johor Bahru offers a unique opportunity to contrast different urban typologies a spatially constrained island metropolis versus a sprawling mainland regional capital. This comparative approach enhances the generalizability of the findings and provides planning authorities with a transferable approach for applying analytical techniques across heterogeneous urban environments. In conclusion, this study addresses a clearly defined and empirically grounded research gap by introducing a technically robust, integrated data-driven, scalable and reproducible methodology for transit network analysis.

3. Methodology

3.1. Data Collection and Preprocessing

This study relies on a multi-source geospatial dataset, anchored by GTFS Static feeds, to evaluate transit network performance and accessibility in Johor Bahru and Penang. Supplementary data layers, including POIs, global walking friction surfaces, and Malaysia's population density raster, are integrated to support spatial analysis of service equity, walkability, and stop accessibility. Figure 2 illustrates the overall research workflow, encompassing the problem statement, study area selection, data acquisition and preprocessing, as well as the analytical components involving transit network evaluation and pedestrian accessibility assessment for the myBAS Johor and Rapid Bus Penang.

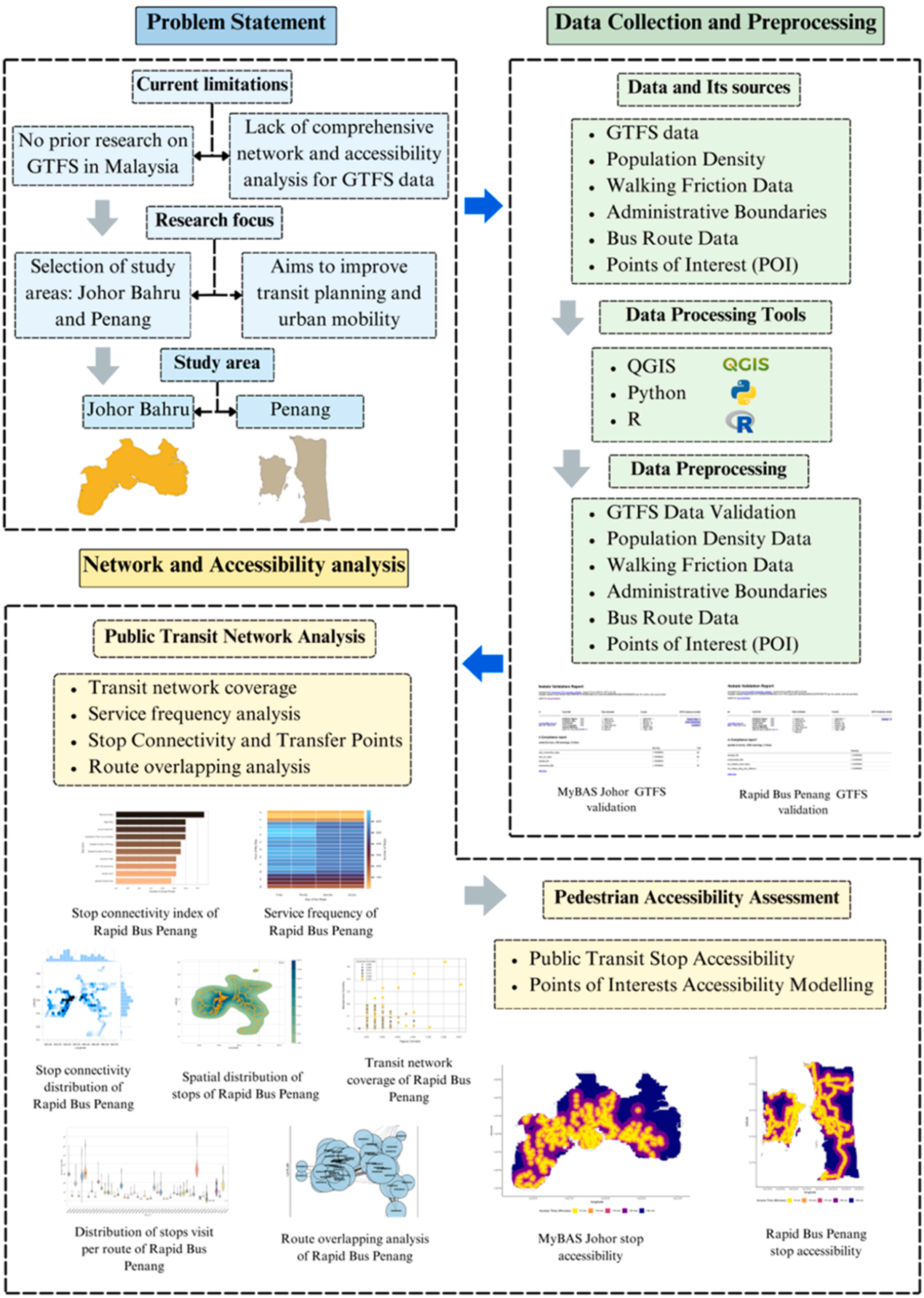


Figure 2. Workflow of the research methodology.

3.1.1. GTFS Data Acquisition and Structure

GTFS datasets for myBAS Johor and Rapid Bus Penang were obtained from Malaysia's open government portal *www.data.gov.my* on August 10, 2024.These feeds contain structured transit data including routes, trips, scheduled arrival and departure times, stop locations, and service calendars. Table 4 summarizes the key characteristics of each dataset.

The files processed included *stops.txt, routes.txt, trips.txt, stop_times.txt,* and *calendar.txt*, which collectively define the spatial and temporal structure of each city's transit system. These datasets were parsed and validated using Python-based GTFS libraries, ensuring compliance with formatting standards and internal consistency across linked files.

3.1.2. Supplementary Spatial Datasets

To complement the GTFS analysis, several spatial datasets were integrated. Administrative boundary shapefiles were acquired from the Global Administrative Areas (GADM) database to delineate the study areas. POIs used for the accessibility analysis were extracted from OSM using structured key-value filters corresponding to key functional domains [79]. Educational facilities were identified with *amenity=school*, healthcare services with *amenity=hospital*, and commercial destinations with *shop=mall* and *amenity=marketplace*. Additionally, residential areas were included using *building=apartments*, representing high-density housing clusters relevant for evaluating population-level accessibility. To ensure semantic accuracy, residential POIs were systematically reviewed to identify and correct potential misclassifications, particularly where tagging inconsistencies may have conflated residential and commercial land uses. This approach ensures land use precision while preserving functional relevance in mixed-use zones.

Two supplementary raster datasets were incorporated to support the assessment of pedestrian accessibility. The first is the global walking friction surface, which provides estimated walking travel time based on terrain and land cover [80,81]. The second is Malaysia's gridded population density raster, sourced from WorldPop [82], which offers high-resolution demographic data to contextualize accessibility in relation to population distribution.

3.1.3. Data Cleaning and validation

All datasets underwent preprocessing to ensure spatial alignment, completeness, and analytical compatibility. GTFS validation, conducted using the Canonical GTFS Schedule Validator, included checks for referential integrity (e.g., alignment between trips.txt and stop_times.txt), detection of malformed timestamps, and identification of missing or redundant identifiers. Additionally, Python scripting was used to harmonize time formats and correct logical inconsistencies [83].

The raster datasets were reprojected to the EPSG:4326 coordinate reference system (WGS 84) to ensure geospatial consistency and compatibility with vector datasets. Subsequently, they were spatially clipped to the administrative boundaries of Johor Bahru and Penang, effectively excluding extraneous regions and enhancing the analytical precision within the defined urban extents.

For POIs, attribute tables were reviewed to remove entries with missing name or type fields. Only locations with verifiable categorical definitions were retained. This step was critical to ensure that proximity-based accessibility modeling was not biased by misclassified or incomplete records. Through this integrated data preparation pipeline, the study ensures a high level of geospatial fidelity, data integrity, and analytical readiness for evaluating network structure, service frequency, and pedestrian accessibility. GTFS validation summaries for myBAS Johor and Rapid Bus Penang are presented in Figure 3(a) and 3(b), respectively.

Table 4
Summary of GTFS Datasets for Johor Bahru and Penang.

Metric	myBAS Johor	Rapid Bus Penang
Number of Stops	924	1969
Number of Routes	18	49
Number of Trips	1047	7596
Average Stops per Trip	47.82	50.74

3.2. Public Transit Network Analysis

To evaluate the performance, spatial structure, and user accessibility of public transit in Johor Bahru and Penang, we develop and implement an Integrated Transit Network and Accessibility Analysis, which forms the central methodological backbone of this study. This unified approach is designed to simultaneously assess network robustness, service coverage, route redundancy, and pedestrian-level accessibility. By combining graph-theoretic modeling, spatiotemporal service estimation, and geospatial cost-surface analysis, this approach enables a multi-dimensional and data-driven assessment of transit efficiency and equity.

The analysis is organized into two interdependent components. First, the Public Transit Network Analysis module evaluates structural connectivity, transfer potential, and service intensity by applying GTFS-derived graph models and route overlap diagnostics. Second, the Pedestrian Accessibility and First-Mile Connectivity module models walking impedance, spatial equity, and POI-to-stop access using raster-based travel cost modeling and proximity estimation. This integrated structure supports a reproducible, scalable, and policy-relevant analytical methodology particularly valuable in urban environments characterized by limited or fragmented planning data. The following subsections present the specific analytical techniques employed within each component.

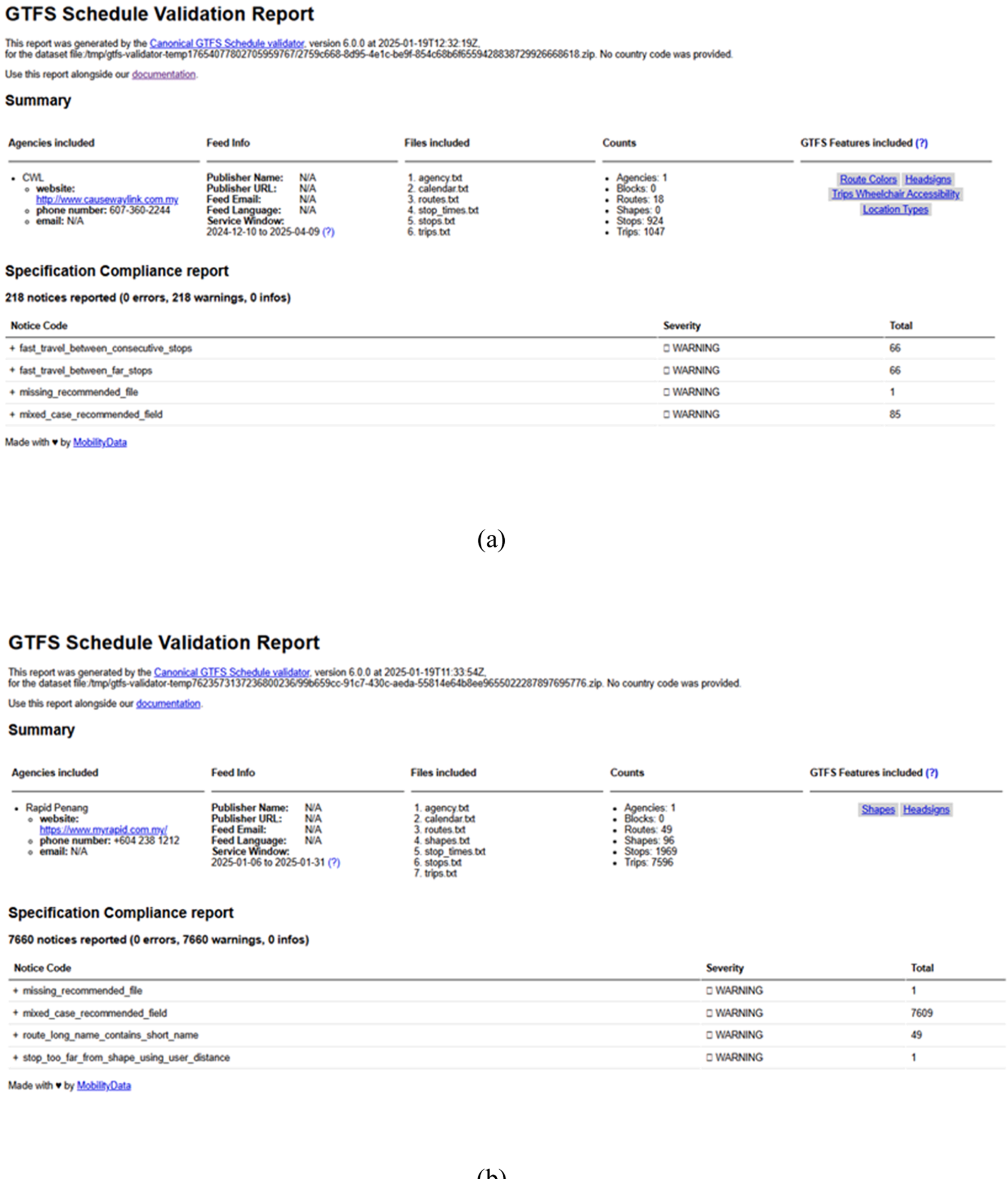
The selection of metrics in this study is guided by their theoretical foundations in transport network science and their empirical applicability to evaluating urban transit systems. Each metric serves a distinct analytical purpose within the integrated framework, contributing to a comprehensive understanding of transit performance from structural, temporal, and spatial perspectives.

Graph-theoretic centrality measures specifically degree, betweenness, and closeness centrality were employed to evaluate the topological configuration and functional significance of individual transit stops. Degree centrality quantifies the number of direct connections at a node, thereby identifying stops that function as local hubs with high transfer potential. Betweenness centrality assesses the extent to which a stop serves as an intermediary along the shortest paths between other nodes, highlighting critical junctions and potential bottlenecks in the network. Closeness centrality captures the average topological distance from a given stop to all others in the network, offering a measure of network-wide accessibility and overall efficiency [66]. These centrality indices have been widely validated in the literature as effective tools for diagnosing connectivity and resilience in public transport networks [84].

Temporal dynamics were modeled using a service frequency function that accounts for variations in operational intensity across different periods of the day. This approach allows the study to distinguish between pre-peak, peak, and post-peak service conditions, thereby capturing temporal inequities and identifying scheduling inefficiencies [85]. The piecewise functional form of the model, parameterized using headway data extracted from GTFS feeds, enables time-sensitive analysis of service patterns and supports empirical planning interventions aligned with temporal travel demand patterns.

To assess the spatial distribution of transit infrastructure, KDE was applied. KDE provides a continuous surface representation of stop density, facilitating the identification of service-rich areas, transit deserts, and redundant spatial clustering [86]. Unlike discrete point-based methods, KDE offers a flexible, non-parametric approach to visualizing spatial intensity, making it particularly well-suited to heterogeneous urban environments where service provision is unevenly distributed. The choice of a bivariate Gaussian kernel and the use of Silverman's rule for bandwidth selection ensure statistical validity and consistency with spatial analysis standards in transport research [87–89].

Route redundancy and structural overlap were evaluated using the



GTFS Schedule Validation Report

This report was generated by the Canonical GTFS Schedule validator, version 6.0.0 at 2025-01-19T12:32:19Z,
for the dataset file:/tmp/gtfs-validator-temp17654077802705959767/2759c668-8d95-4e1c-be9f-854c68b6f6559428838729926668618.zip. No country code was provided.

Use this report alongside our documentation.

Summary

Agencies included	Feed Info	Files included	Counts	GTFS Features included (?)
• CWL ∘ **website:** http://www.causewaylink.com.my ∘ **phone number:** 607-360-2244 ∘ **email:** N/A	**Publisher Name:** N/A **Publisher URL:** N/A **Feed Email:** N/A **Feed Language:** N/A **Service Window:** 2024-12-10 to 2025-04-09 (?)	1. agency.txt 2. calendar.txt 3. routes.txt 4. stop_times.txt 5. stops.txt 6. trips.txt	• Agencies: 1 • Blocks: 0 • Routes: 18 • Shapes: 0 • Stops: 924 • Trips: 1047	Route Colors Headsigns Trips Wheelchair Accessibility Location Types

Specification Compliance report

218 notices reported (0 errors, 218 warnings, 0 infos)

Notice Code	Severity	Total
+ fast_travel_between_consecutive_stops	WARNING	66
+ fast_travel_between_far_stops	WARNING	66
+ missing_recommended_file	WARNING	1
+ mixed_case_recommended_field	WARNING	85

Made with ♥ by MobilityData

(a)

GTFS Schedule Validation Report

This report was generated by the Canonical GTFS Schedule validator, version 6.0.0 at 2025-01-19T11:33:54Z,
for the dataset file:/tmp/gtfs-validator-temp7623573137236800236/99b659cc-91c7-430c-aeda-55814e64b8ee9655022287897695776.zip. No country code was provided.

Use this report alongside our documentation.

Summary

Agencies included	Feed Info	Files included	Counts	GTFS Features included (?)
• Rapid Penang ∘ **website:** https://www.myrapid.com.my/ ∘ **phone number:** +604 238 1212 ∘ **email:** N/A	**Publisher Name:** N/A **Publisher URL:** N/A **Feed Email:** N/A **Feed Language:** N/A **Service Window:** 2025-01-06 to 2025-01-31 (?)	1. agency.txt 2. calendar.txt 3. routes.txt 4. shapes.txt 5. stop_times.txt 6. stops.txt 7. trips.txt	• Agencies: 1 • Blocks: 0 • Routes: 49 • Shapes: 96 • Stops: 1969 • Trips: 7596	Shapes Headsigns

Specification Compliance report

7660 notices reported (0 errors, 7660 warnings, 0 infos)

Notice Code	Severity	Total
+ missing_recommended_file	WARNING	1
+ mixed_case_recommended_field	WARNING	7609
+ route_long_name_contains_short_name	WARNING	49
+ stop_too_far_from_shape_using_user_distance	WARNING	1

Made with ♥ by MobilityData

(b)

Figure 3. GTFS validation summaries for the two transit networks (a) myBAS Johor (b) Rapid Bus Penang.

Jaccard Similarity Index (JSI), which quantifies the degree of stop overlap between pairs of transit routes. The JSI offers a simple yet powerful measure for detecting parallel service corridors, identifying inefficiencies in network design, and informing route optimization strategies [16]. Collectively, the selected metrics offer a methodologically coherent and technically robust foundation for evaluating structural integrity, temporal adequacy, spatial equity, and operational redundancy in urban transit networks. Their integration within this study's workflow ensures a multidimensional analysis that is both theoretically grounded and practically actionable, particularly in urban environments where planning data may be fragmented or constrained.

3.2.1. Transit Network Coverage

- **Degree Centrality:** Degree centrality is a fundamental graph-theoretic measure used to assess the local connectivity of a node within a network [90]. In the context of public transportation, each node represents a transit stop, and edges denote direct route connections between stops. Degree centrality quantifies the number of immediate, direct connections a stop has with other stops in the network, thus serving as a proxy for its transfer potential and accessibility within the system. This metric is particularly valuable for identifying key local hubs nodes that facilitate high-frequency transfers and enhance the operational flexibility of the network. A higher degree centrality implies that a transit stop is better integrated with adjacent nodes, allowing passengers more options for route switching and potentially reducing total journey time. Degree centrality for a node v is defined in Equation (1) as:

$$C_D\ (v) = \frac{\deg(v)}{n-1} \tag{1}$$

where, $deg(v)$ represents the number of direct connections a node v has, and n denotes the total number of nodes in the network.

The rationale for employing degree centrality in this study stems from its simplicity, interpretability, and effectiveness in capturing the local structural role of transit stops. It allows planners and analysts to pinpoint highly connected nodes that may serve as pivotal elements in multimodal integration, service frequency planning, and infrastructure investment decisions. Furthermore, when analyzed spatially, patterns in degree centrality can reveal topological inequities and support decisions aimed at improving network uniformity and access equity.

- **Betweenness centrality:** Betweenness centrality is a key topological metric that captures the extent to which a transit stop functions as a bridge or intermediary within the overall network structure. In public transportation analysis, this measure is particularly important for identifying strategic transfer points that facilitate connectivity between, otherwise disconnected or weakly linked areas of the network [91].

Conceptually, betweenness centrality evaluates the frequency with which a given node (i.e., transit stop) lies along the shortest paths between all other pairs of nodes in the network. A stop with high betweenness centrality is likely to serve as a critical juncture, indicating that a substantial volume of passenger flow may depend on that location for transfers between routes or corridors. Such nodes are structurally significant and often represent potential choke points, where congestion, delays, or infrastructure failure can disproportionately affect network-wide performance. Formally, the betweenness centrality $C_B(v)$ for a node v is computed using the formulation provided in Equation (2):

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{2}$$

σ_{st}represents the total number of shortest paths between nodes v and t, while $\sigma_{st}(v)$ denotes the count of these paths passing through v.

The inclusion of betweenness centrality in this study is methodologically motivated by its ability to reveal functional importance beyond local connectivity. While degree centrality assesses immediate neighborhood integration, betweenness centrality highlights the global navigational significance of each stop. In practical terms, this metric supports the identification of nodes that require infrastructural reinforcement, targeted scheduling interventions, or redundancy planning to mitigate systemic vulnerabilities.

- Closeness centrality

Closeness centrality is a global network metric that measures how efficiently a transit stop is positioned relative to all other stops in the network [92]. It reflects the average topological proximity of a node to all others, capturing the potential of that stop to facilitate quick and direct access across the transit system.

A stop with high closeness centrality is interpreted as being centrally located in the network topology, implying that it can reach other parts of the network with minimal transfers or travel time. Such nodes often serve as effective origins or destinations for journeys, and their spatial optimization is key to improving overall accessibility and service efficiency. The closeness centrality $C_c(v)$ of a node v is mathematically defined in Equation (3):

$$C_c(v) = \frac{n-1}{\sum_{u \neq v} d(v,u)} \tag{3}$$

where n is the total number of nodes in the network, and $d(v,u)$ denotes the length of the shortest path between nodes v and u.

Closeness centrality was selected for this study because it offers a direct measure of network accessibility from a system-wide perspective. Unlike degree or betweenness, which focus on local connectivity or transit flow through a node, closeness centrality provides insight into how quickly a passenger can navigate from one stop to the rest of the network. This makes it particularly useful for evaluating the spatial efficiency of the network and for identifying stops that enhance overall system cohesion.

3.2.2. Service Frequency Analysis

Service frequency is a critical temporal indicator of transit system performance, reflecting the availability and regularity of public transportation across different periods of the day [93–95]. In this study, we analyzed service frequency to capture the temporal dynamics of bus operations in Johor Bahru and Penang, focusing on variations during pre-peak, peak, and post-peak hours. Such differentiation allows for a nuanced understanding of whether service provision aligns with fluctuating commuter demand.

To characterize these temporal variations, we employ a piecewise exponential model, which defines service intensity as a function of the hour of the day, denoted by h . The formulation of the model is presented in Equation (4):

$$N(h) = \begin{cases} a_1 e^{b1h}, & h < 7\ (pre - peak\ period) \\ a_2, & 7 \leq h \geq 19\ (steady\ peak\ period) \\ a_3 e^{-b_2 h}, & h > 19\ (post - peak\ period) \end{cases} \tag{4}$$

In this formulation, $N(h)$ represents the number of transit trips scheduled per hour at a given time h. The constants a_1, a_2 and a_3 correspond to service intensity levels during the respective periods: rising, stable, and declining. The parameters $b1$ and b_2 capture the rate of exponential growth or decay in service density before and after peak hours, respectively. These parameters were estimated empirically from GTFS *stop_times.txt* data using nonlinear regression techniques, ensuring data-driven parameterization.

This approach was selected due to its ability to represent real-world temporal service characteristics with parsimony and analytical clarity. Unlike static headway averages, which obscure intra-day variability, the piecewise exponential model provides a continuous and interpretable profile of service fluctuations. It is particularly effective in revealing under-served time windows, enabling policymakers and operators to optimize schedules with respect to commuter flow dynamics. Furthermore, integrating this temporal analysis into our broader analytical approach enhances the capacity of the study to evaluate both spatial and temporal equity in service provision.

3.2.3. Stop Connectivity and Transfer Points

Stop connectivity is a critical dimension in evaluating the structural and functional performance of a transit network. It reflects how well a given transit stop is integrated into the broader system through its connections to multiple transit routes. High stop connectivity is typically associated with increased route choice flexibility, reduced transfer times, and improved network efficiency, especially in multimodal or high-frequency service contexts [96].

To quantify this, we introduce the Stop Connectivity Index, which measures the number of distinct routes that serve a given stop. This index provides a localized yet meaningful indication of potential transfer at each node. Mathematically, the connectivity *Cs* of a stop s is expressed as shown in Equation (5):

$$C_s = \sum_{i=1}^{N_s} \delta(r_i) \tag{5}$$

where N_s denotes the total number of transit routes in the system, and $\delta(r_i)$ is an indicator function that returns 1 if route r_i includes stop s, and 0 otherwise. This formulation ensures a binary inclusion count that is both intuitive and scalable, making it suitable for cross-network comparative diagnostics.

In addition to route-based connectivity, understanding the spatial distribution of stops and their clustering patterns is essential for identifying service concentration and potential redundancy. To this end, we evaluate Stop Connectivity Distribution using a hexagonal binning approach, a method particularly effective in spatial network analysis due to its minimal edge effects and consistent area coverage. The density function $f(x,y)$, which describes the spatial intensity of stops within a defined hexagonal unit $\mathrm{Hex}(x,y)$, is computed as shown in Equation (6):

$$f(x,y) = \sum_{i=1}^{N} I[(x_i,y_i) \in Hex(x,y)] \tag{6}$$

Here, (x_i, y_i) are the coordinates of stop i, and I is an indicator function that equals 1 if the stop falls within the spatial boundaries of the hexagon centered at (x, y), and 0 otherwise. The resulting surface identifies high-density clusters—which may represent efficient hubs or over-saturated zones—and sparsely served areas, often corresponding to underserved or transit-poor neighborhoods.

3.2.4. Spatial Distribution of Transit Stops and Kernel Density Estimation (KDE)

Understanding the spatial distribution of transit stops is central to evaluating the geographic accessibility and equity of public transportation systems. Uneven stop placement can result in service oversaturation in central areas and under-provision in peripheral or low-income regions, thereby contributing to spatial injustice in transit access [97–99]. To address this, we apply KDE, a non-parametric technique used to generate a continuous surface representation of point densities—here, the density of transit stops—across the urban landscape. The spatial density at a location (x,y) is computed using the bivariate kernel function presented in Equation (7):

$$f(x,y) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{x-x_i}{h}, \frac{y-y_i}{h}\right) \tag{7}$$

where $f(x,y)$ denotes the estimated density at coordinates (x,y), n is the total number of transit stops, h is the bandwidth parameter controlling the degree of spatial smoothing, and K is the kernel function. In this study, we use a Gaussian kernel due to its smooth decay properties and its established utility in spatial accessibility studies.

Bandwidth h is a critical parameter in KDE as it governs the scale at which density variations are detected. A smaller bandwidth captures fine-grained spatial heterogeneity but may introduce noise, while a larger bandwidth smooths the surface at the risk of masking local variation [100]. In our analysis, the bandwidth h for kernel density estimation was selected using Silverman's Rule of Thumb [89], a well-established heuristic that provides a bias-variance tradeoff optimal under Gaussian assumptions. Although the distribution of stop densities may deviate from strict normality, this method offers a computationally efficient and reproducible approach for cross-city comparisons. The consistent use of this rule ensures interpretability while mitigating overfitting, especially in cities where stop data is sparse or noisy.

For comparative purposes, and to isolate the influence of stop distribution along the x-dimension alone, a univariate form of the KDE is expressed in Equation (8):

$$f(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \tag{8}$$

The KDE-based density surface enables the identification of transit-rich zones, service deserts, and redundant spatial clusters, which may correspond to inefficient resource allocation or unplanned stop proliferation [101]. By visualizing stop intensity as a continuous surface, this method provides a more nuanced representation than conventional buffer-based accessibility models.

KDE has been chosen in this study for its methodological transparency, statistical robustness, and capacity to integrate seamlessly with raster-based spatial analyses. It allows planners and policymakers to pinpoint underserved regions, prioritize infrastructure upgrades, and evaluate the alignment between population density and transit provision. Moreover, KDE output forms an essential input for overlay analyses with POIs and population rasters, thereby enriching the overall evaluation of transit equity and spatial performance.

3.2.5. Route Overlapping Analysis and Spatial Route Representation

Evaluating route redundancy is essential for optimizing transit resource allocation, minimizing inefficiencies, and ensuring comprehensive yet non-duplicative service coverage [102]. We apply the JSI as a quantitative measure of spatial and operational overlap between pairs of transit routes [16].

The JSI evaluates the similarity between two sets by comparing their intersection relative to their union. In the context of transit analysis, this metric captures the proportion of common stops shared by two routes. A higher JSI value indicates greater overlap, suggesting potential service duplication, while a lower value reflects route differentiation and spatial diversification. Formally, for two routes r_1 and r_2 the overlap $O(r_1,r_2)$ is defined in Equation (9):

$$O(r_1,r_2) = \frac{|S(r_1) \cap S(r_2)|}{|S(r_1) \cup S(r_2)|} \tag{9}$$

where $S(r_1)$ and $S(r_2)$ denote the sets of stops serviced by routes r_1 and r_2, respectively. The numerator represents the number of shared stops between the two routes, and the denominator indicates the total number of unique stops across both routes.

The JSI was selected for this study due to its simplicity, interpretability, and broad acceptance in both transportation planning and spatial analytics. It allows for systematic pairwise comparisons of all routes, enabling planners to identify high-overlap corridors that may benefit from consolidation, staggered scheduling, or mode diversification.

To support spatial visualization and geospatial clustering of redundant routes, we calculate a centroid-based spatial representation of each transit route. This is derived by averaging the geographic coordinates (latitude and longitude) of all stops associated with a given route, yielding a representative spatial point for each route, as shown in Equation (10):

$$P(r) = \left(\frac{\sum_{i=1}^{n} lat_i}{n}, \frac{\sum_{i=1}^{n} lon_i}{n}\right) \tag{10}$$

Here, $P(r)$ denotes the centroid coordinates of route r, while lat_i and lon_i are the latitude and longitude of the $i-th$ stop, and n is the total number of stops on the route. This spatial abstraction enables route-to-route proximity analysis and supports the integration of JSI values into geovisual dashboards, cluster maps, or network heatmaps.

Together, the JSI and route centroid modeling provide a multi-scalar framework for assessing spatial efficiency and operational redundancy in urban transit systems. This dual analysis is especially relevant in high-density urban environments, where overlapping services may reflect inefficient scheduling or uncoordinated planning across parallel corridors. By identifying such redundancies, transit authorities can improve route rationalization and equity in service distribution.

3.3. Pedestrian Accessibility and First-Mile Connectivity Assessment

Ensuring equitable access to public transportation systems requires not only robust network coverage but also walkable access from origin points such as homes, schools, and commercial facilities. We employ a geospatial methodology to assess pedestrian accessibility and first-mile connectivity, using cost surfaces, friction-based transitions, and travel time estimations.

To represent the spatial impedance faced by pedestrians moving across the urban terrain, we utilize a global walking friction surface raster [80]. Each raster cell encodes a friction value f_{ij}, which reflects the relative difficulty or cost of traversing the terrain from cell i to cell j. These values incorporate land cover, slopes, and infrastructure barriers. The transition weight, W_{ij}, quantifies how easily a pedestrian can move between adjacent raster cells. As shown in Equation (11), the cost is modeled as the inverse of the friction value.

$$W_{ij} = \frac{1}{f_{ij}} \tag{11}$$

This formulation ensures that higher friction values (e.g., rough terrain or lack of footpaths) result in lower transition weights, indicating reduced walkability. This raster-based impedance model forms the foundation of the cumulative accessibility computation.

3.3.1. Pedestrian Accessibility

Using the transition weights, a cumulative cost surface is computed to represent the least-cost walking paths from transit stops to all other grid cells within the study area [103,104]. This is achieved via graph traversal algorithms such as Dijkstra's algorithm or cost-distance functions implemented within GIS [105–107]. The total accumulated cost $C(x,y)$ for reaching any cell (x,y) is defined in Equation (12):

$$C(x,y) = min \sum_{i=1}^{n} (d_i, W_i) \tag{12}$$

Here, d_i represents the Euclidean or geodesic distance between adjacent raster cells in the traversal path, and W_i is the transition weight at each step. This surface provides a spatial representation of walkability-adjusted accessibility, highlighting areas with high pedestrian impedance and identifying potential barriers to first-mile access. It also supports overlay analysis with population density or socioeconomic vulnerability indicators to diagnose spatial equity in access.

3.3.2. Walking Time Estimation from POIs to Transit Stops

In addition to the raster-based cost modeling, we perform a direct estimation of walking time from the centroids of selected POIs including schools, hospitals, commercial centers, and residential zones to the nearest transit stop. This metric, though simpler, offers an intuitive representation of accessibility and is calculated in Equation (13):

$$T_i = \frac{D_{ij}}{S} \tag{13}$$

The estimated walking time from a point of interest i to the nearest transit stop j, denoted as T_i is calculated by dividing the straight-line distance D_{ij} by the assumed walking speed S. In this formulation, D_{ij} represents the Euclidean distance between the POI and the corresponding transit stop, while S is set to a constant value of 5 kilometers per hour, in alignment with internationally recognized standards for average pedestrian travel speed in urban environments [108,109]. This provides a standardized basis for comparing first-mile accessibility across diverse geographic and infrastructural contexts.

The combined use of raster-based cumulative cost surfaces and point-to-stop time estimations allows for both high-resolution spatial diagnostics and intuitive urban planning metrics. While the former captures the complexity of real-world pedestrian movement under varying land conditions, the latter provides easily interpretable indicators suitable for stakeholder communication and policy benchmarking. This dual approach ensures that accessibility is evaluated not only from a geometric proximity perspective but also in terms of practical navigability, thus enhancing the study's relevance for sustainable transport planning, transit equity audits, and inclusive infrastructure design.

4. Results

This section presents the results of a comparative transit analysis conducted for two Malaysian urban contexts: Johor Bahru and Penang. These cities were selected due to their distinct spatial forms, socioeconomic roles, and transit service challenges, offering a meaningful contrast for evaluating public bus system performance.

Johor Bahru, situated at the southern tip of Peninsular Malaysia, serves as a major economic and cross-border commuting hub adjacent to Singapore. The city experiences high transit demand, particularly along key corridors such as the Johor Causeway, yet continues to face chronic issues related to congestion, service fragmentation, and inadequate last-mile connectivity [110]. Penang, encompassing both the island and mainland regions, is known for its heritage significance and tourism-driven economy. The state's transit system must contend with spatial constraints, inconsistent service frequency, and limited modal integration, particularly in areas such as George Town and across the Penang Bridge [111]. Together, these cases provide a basis for evaluating spatial accessibility, service frequency, and network efficiency using GTFS data. Their geographic extents are illustrated in Figure 4, which highlights the urban boundaries and transit corridors analyzed in this study.

4.1. Transit Network Coverage

4.1.1. Transit Network Coverage: myBAS Johor

The degree centrality analysis revealed that transit stops such as Larkin Sentral, Taman Dato Onn, and Pusat Khidmat Pelanggan TNB@Johor Bahru function as major local hubs due to their high number of direct connections. As shown in Figure 5(a), these stops exhibit high node degrees, suggesting a radial structure in central Johor Bahru where multiple routes converge. This characteristic supports efficient intra-urban travel but may also introduce pressure on these nodes during peak hours.

In terms of betweenness centrality, which reflects the potential of a node to act as a bridge within the shortest paths across the network, Larkin Sentral and Bandar Baru Uda emerge as strategic transfer points. These findings, shown in Figure 5(b), indicate a strong dependence on a small number of nodes for inter-route passenger transitions. Their functional significance, however, also implies a high susceptibility to systemic disruption in case of service failures, underscoring the need for resilience planning and route redundancy.

The closeness centrality analysis, depicted in Figure 5(c), measures the ease with which each node can access the entire network. Stops such as Opp Dewan Jkr Johor Bahru, Larkin Sentral, and Stadium Larkin scored highest, indicating their efficient positioning for minimizing travel times across the urban core. This spatial efficiency is particularly valuable in optimizing route planning and enhancing equitable accessibility. A composite scatter plot, presented in Figure 5(d), further

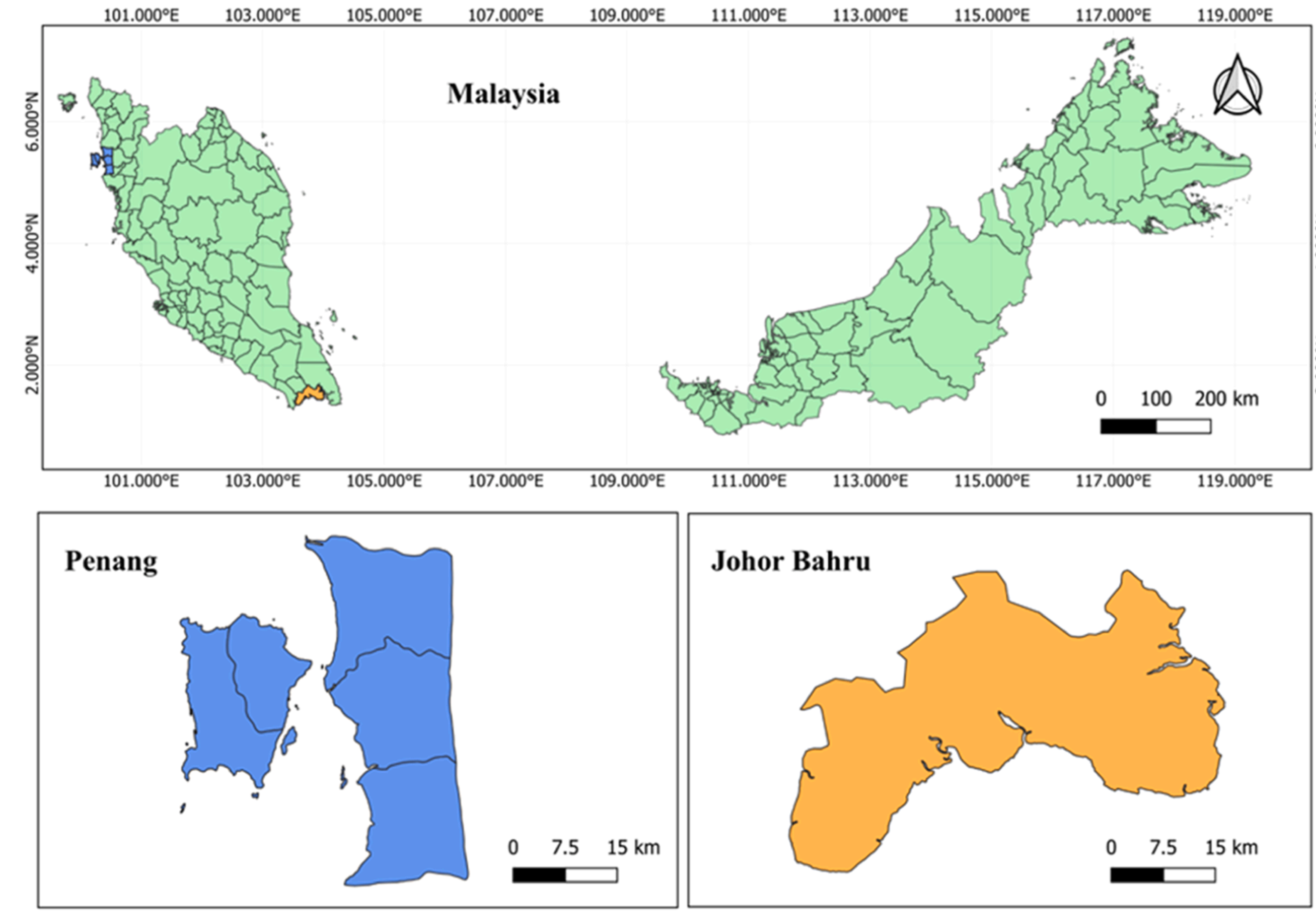


Figure 4. Geographical representation of study area: Johor Bahru and Penang.

elucidates the interrelationship between degree and betweenness centrality. Interestingly, the plot demonstrates a weak correlation between the two metrics: several stops with modest degree centrality nonetheless exhibit high betweenness centrality. This divergence indicates that even peripheral or moderately connected stops can serve critical roles as routing pivots highlighting the multidimensional nature of nodal importance within transit systems.

Overall, the centrality analysis reveals that myBAS Johor exhibits a hybrid network structure, combining features of both centralized (hub-and-spoke) and decentralized (mesh) typologies. Larkin Sentral, in particular, functions as the network's principal node across all three centrality dimensions. However, its overwhelming importance also flags it as a single point of failure. As such, strategic interventions should prioritize increasing route redundancy, fortifying infrastructure at high-betweenness stops, and expanding peripheral stop services to strengthen regional inclusivity and reduce systemic vulnerability.

4.1.2. Transit Network Coverage: Rapid Bus Penang

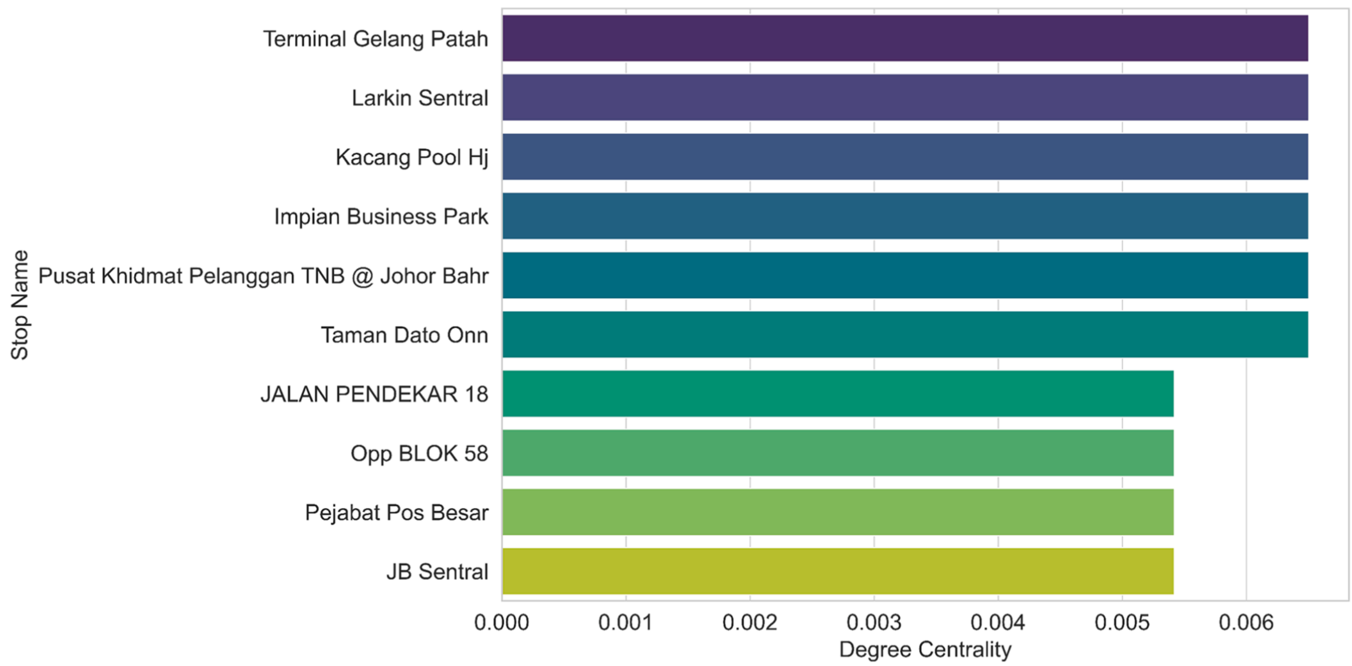
The Rapid Bus Penang transit network presents a distinct spatial and functional topology shaped by its dense urban core and strategic inter-regional linkages. Through the same triad of centrality indices, the network's efficiency, robustness, and vulnerability were critically assessed to inform optimization strategies.

The degree centrality map, presented in Figure 6(a), identifies Terminal Komtar and (M2) Megamall – Pacific as the most connected transit nodes, serving as key urban anchors. Their high degree values suggest that these locations support a diverse array of direct connections, enhancing route flexibility and minimizing the need for inter-route transfers. Such structural centrality is essential in high-density corridors where transit frequency and transfer minimization are paramount for efficient passenger flow. Analysis of betweenness centrality, visualized in Figure 6(b), designates (M2) Megamall – Pacific, (M1) Penang Sentral, and Queensbay Mall as crucial transit pivots. These nodes mediate significant proportions of route interconnectivity, acting as transfer points for cross-regional mobility. However, this reliance introduces fragility: congestion or disruptions at these stops could significantly affect network-wide performance. Accordingly, transport planners should consider introducing buffer nodes and alternative routing schemes to diffuse transfer dependency and enhance fault tolerance.

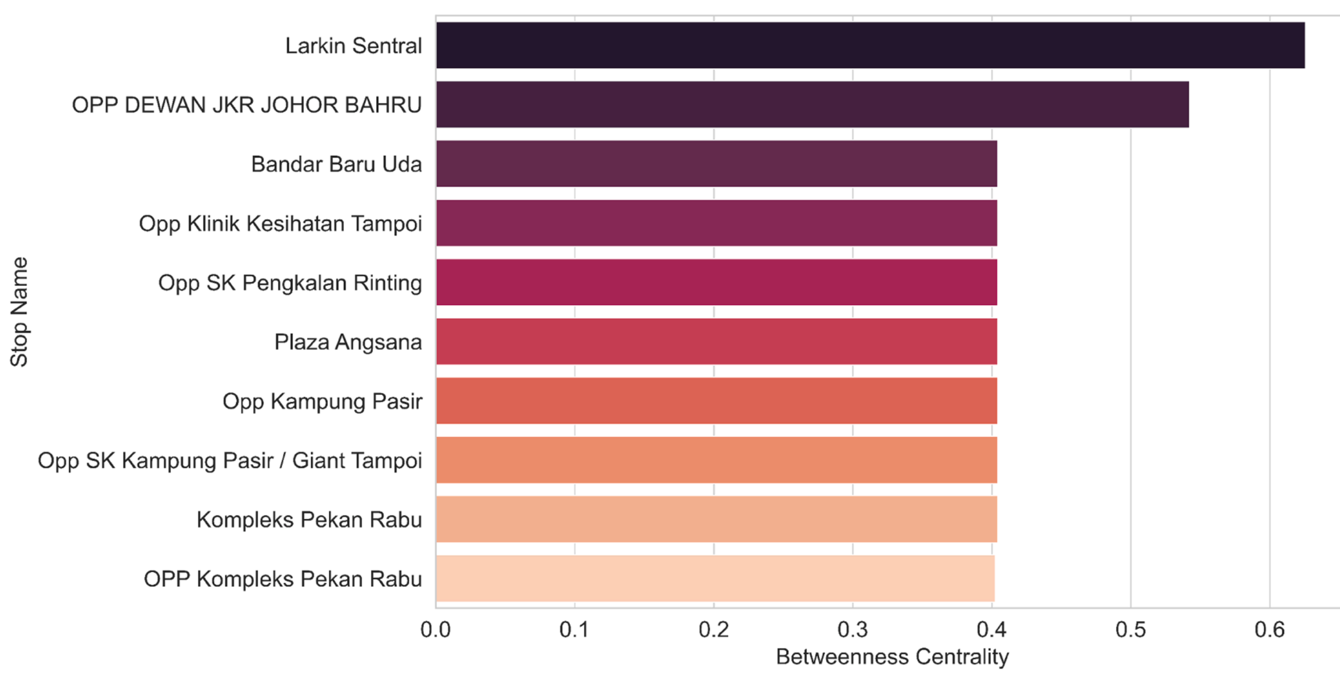
The closeness centrality distribution, as shown in Figure 6(c), highlights the accessibility advantages of Masjid Jamek Kampung Setol and (M2) Megamall – Pacific, indicating their capacity to serve as efficient access points within the network. These findings suggest that improving pedestrian infrastructure and service frequency at these stops could substantially reduce travel times and increase rider satisfaction, particularly in transit-dependent communities.

The combined metric analysis is further elucidated in Figure 6(d), where a scatter plot compares betweenness and degree centrality. Similar to Johor Bahru, the Rapid Penang network exhibits a non-linear association between these indices, emphasizing that strategic value in transit systems is not solely determined by connectivity volume. For instance, (M1) Penang Sentral, despite having moderate degree centrality, plays an outsized role in interzonal connectivity due to its spatial placement and network routing configuration.

These insights highlight the structural monocentric nature of the Rapid Bus Penang network, concentrated heavily in the Georgetown–Bayan Lepas corridor. While such centralization promotes



(a)

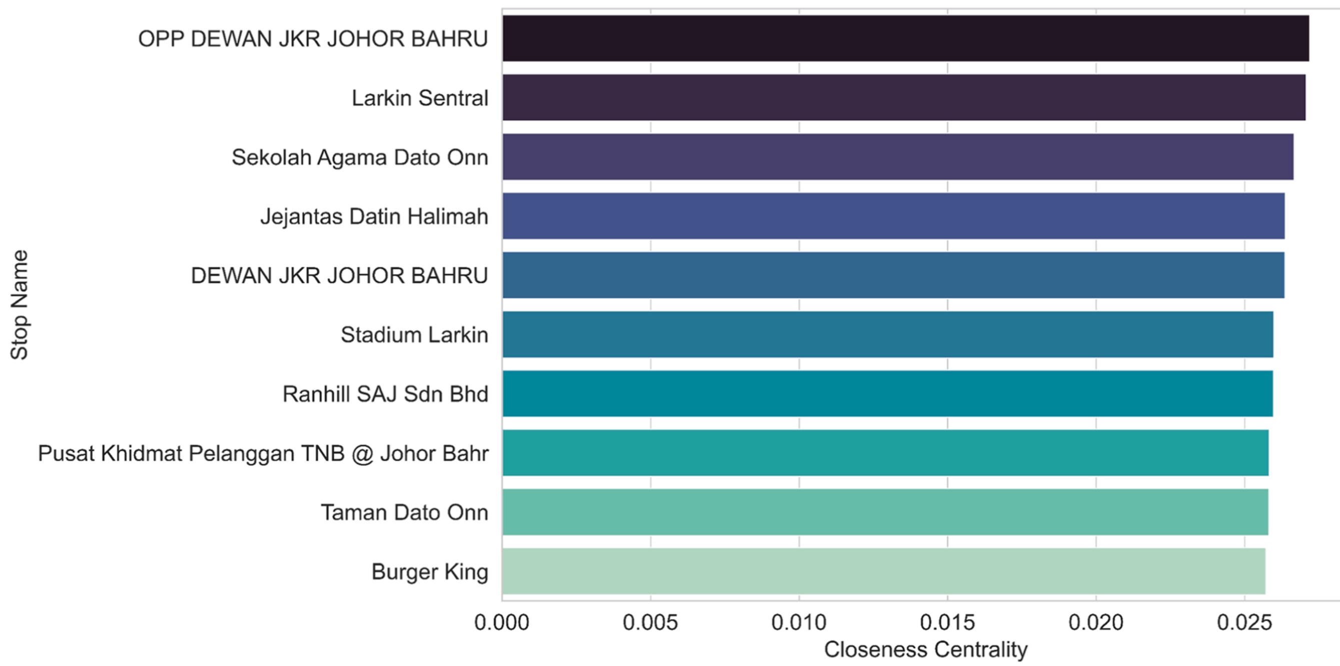


(b)

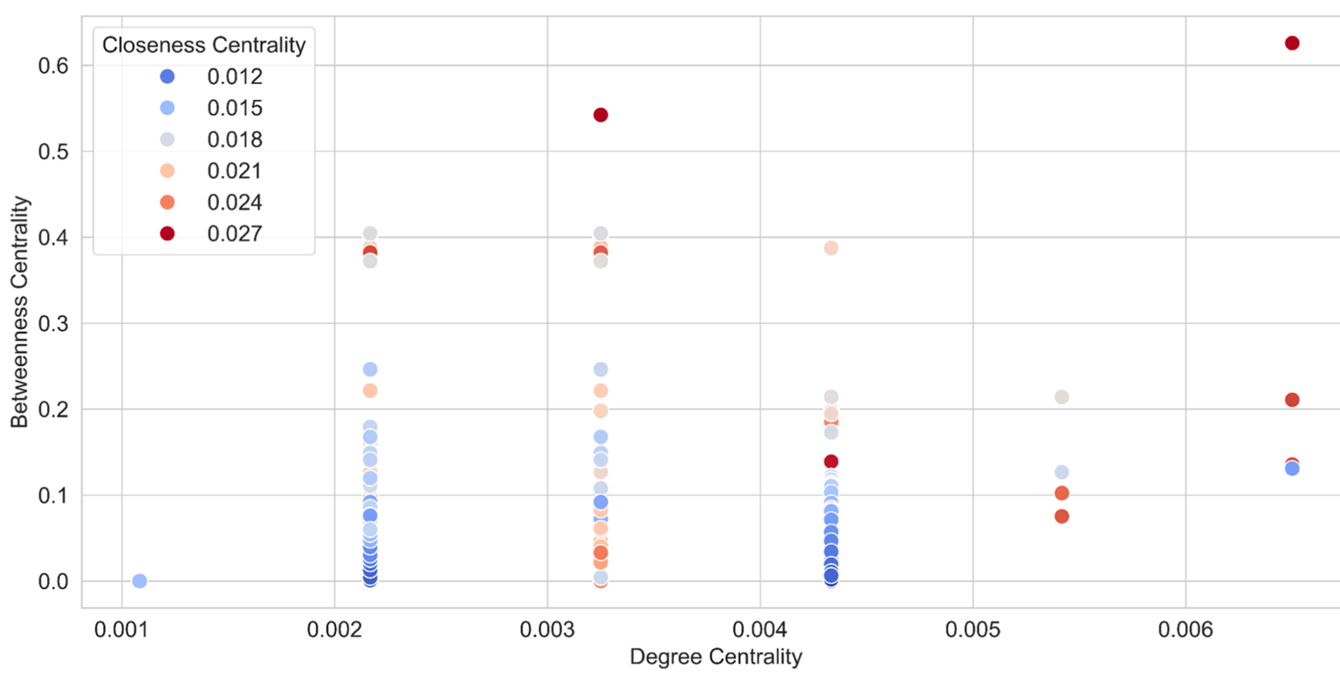
Figure 5. Centrality-based characterization of the myBAS Johor Bus network. (a) Degree centrality, (b) Betweenness centrality, (c) Closeness centrality and (d) Scatter plot between degree and betweenness centrality.

efficiency in the urban core, it raises equity concerns for outlying areas. To counterbalance this, a decentralized expansion strategy featuring capacity upgrades in peripheral nodes, dynamic route rebalancing, and multimodal integration should be considered.

A comparative synthesis of the two networks myBAS Johor and Rapid Bus Penang reveals both convergence and divergence in structural



(c)

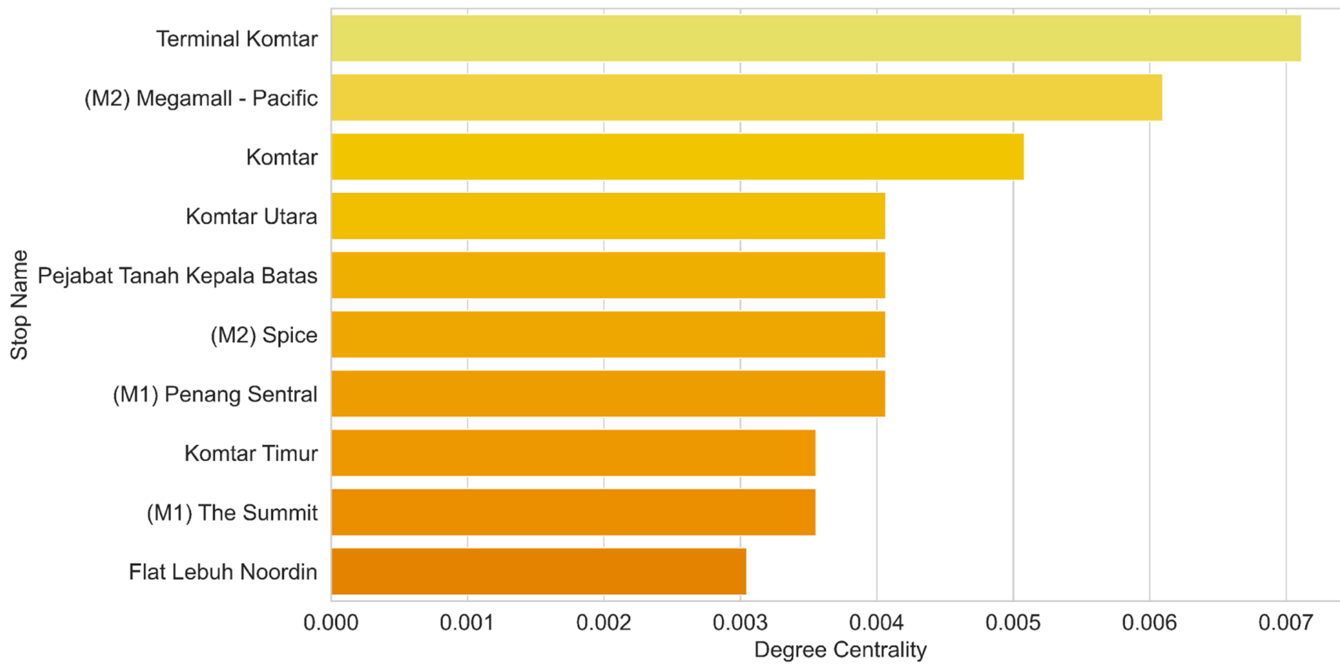


(d)

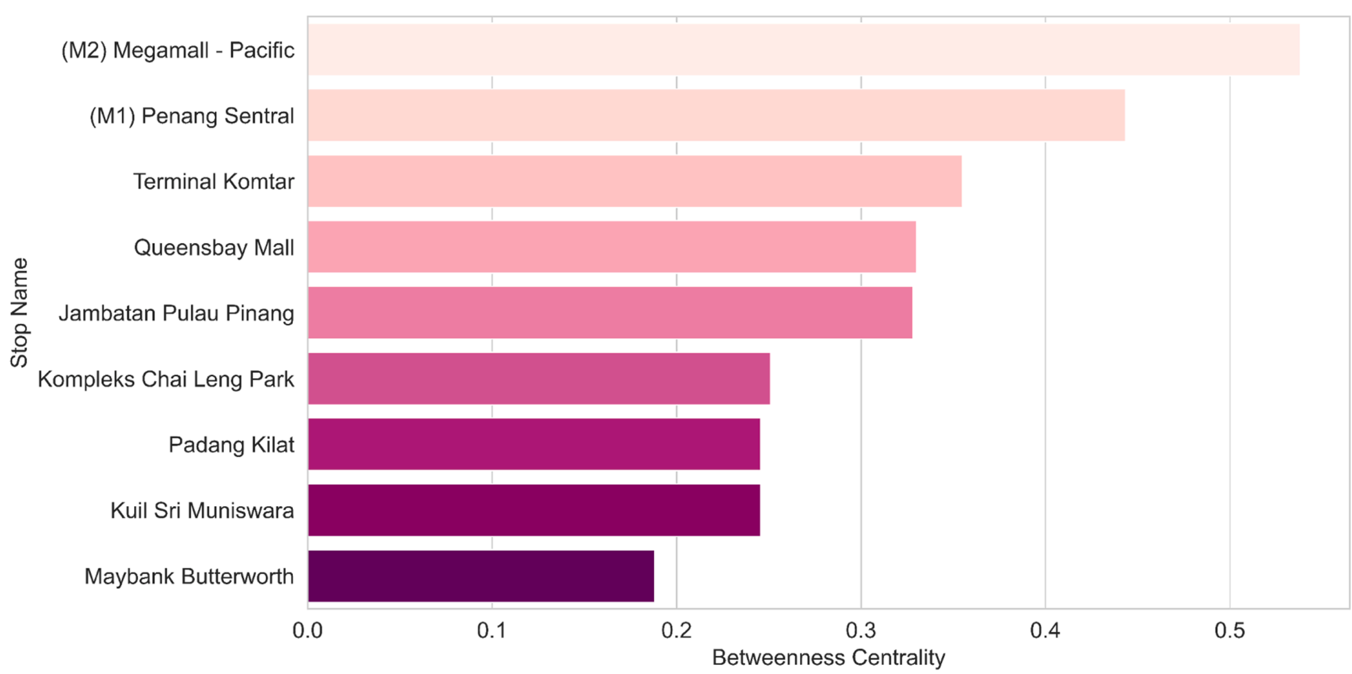
Figure 5. (*continued*).

characteristics. While both systems are anchored by high-centrality nodes Larkin Sentral in Johor Bahru and Terminal Komtar in Penang their underlying topological configurations exhibit notable differences. Johor Bahru exhibits a polycentric pattern with several high-closeness peripheral stops contributing to accessibility. In contrast, Penang's network is markedly monocentric, centralizing mobility around a few strategic corridors.

These distinctions call for differentiated intervention strategies. Johor's planners may focus on reinforcing secondary hubs and enhancing resilience at transfer nodes, whereas Penang could benefit



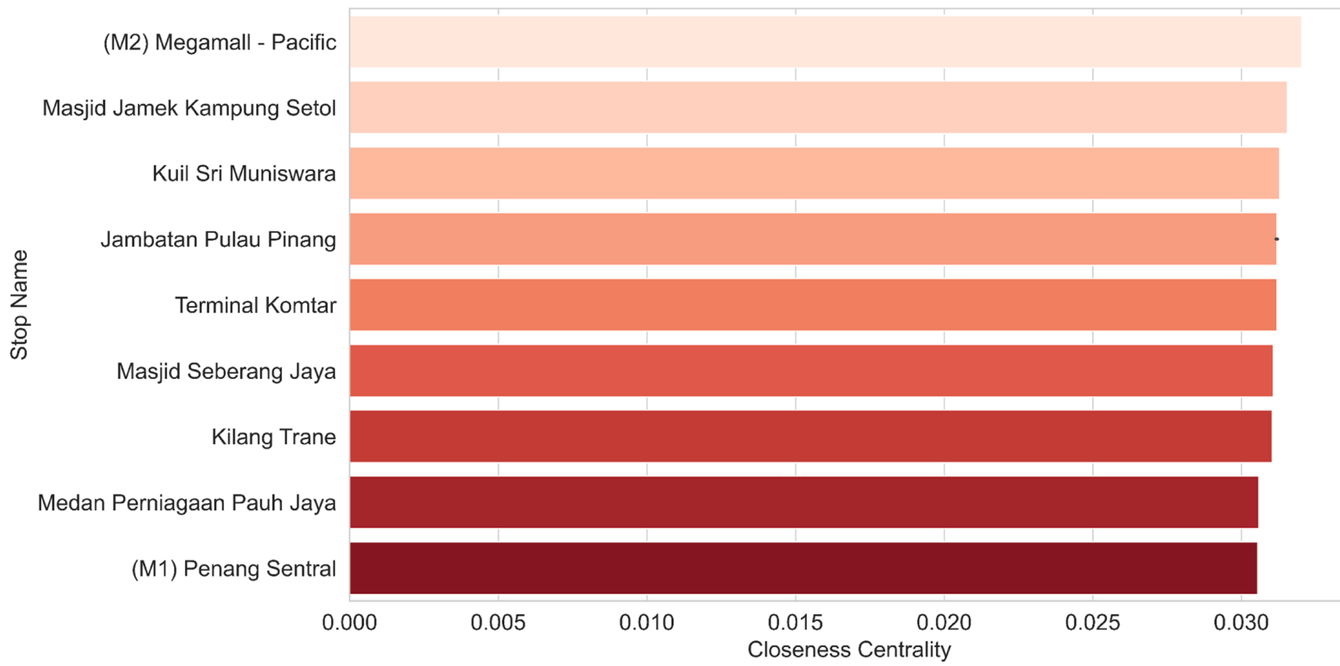
(a)



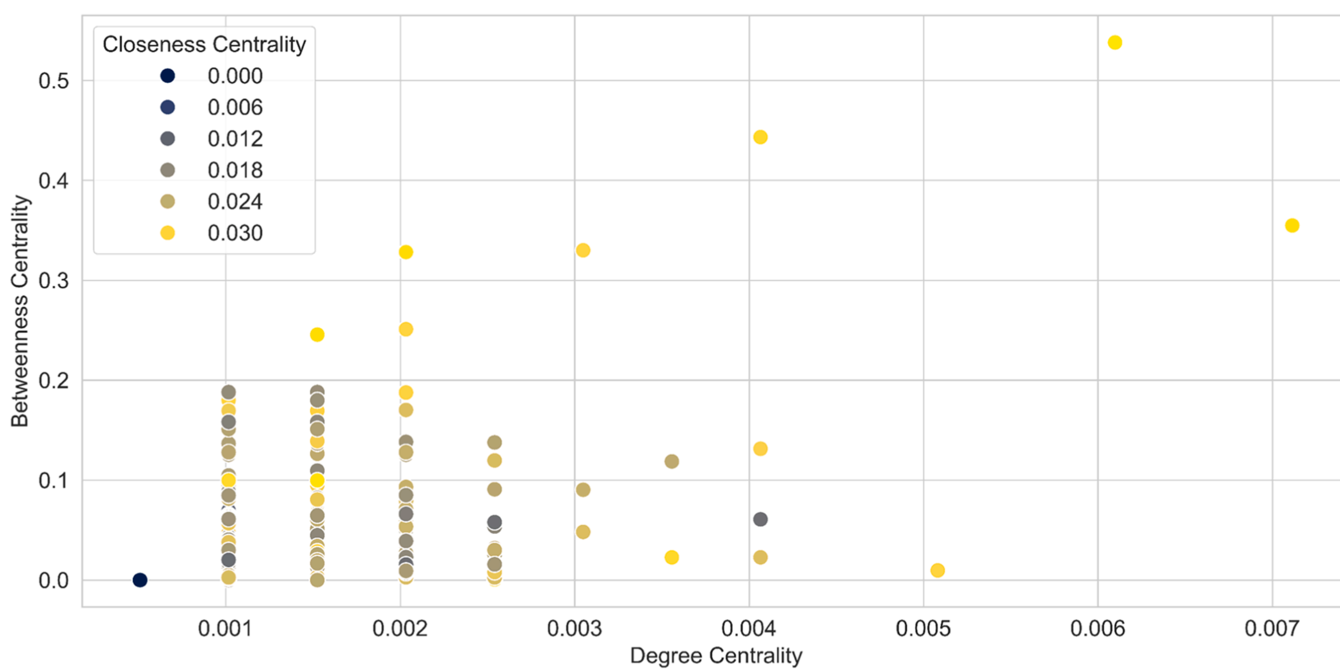
(b)

Figure 6. Centrality-based characterization of the Rapid Bus Penang bus network. (a) Degree centrality, (b) Betweenness centrality, (c) Closeness centrality and (d) Scatter plot between degree and betweenness centrality.

from deliberate decentralization and peripheral service extension. Both cities, however, should consider integrating dynamic GTFS-RT datasets in future analyses to capture operational variability and facilitate adaptive transit planning a shift from descriptive diagnostics to predictive system design.



(c)

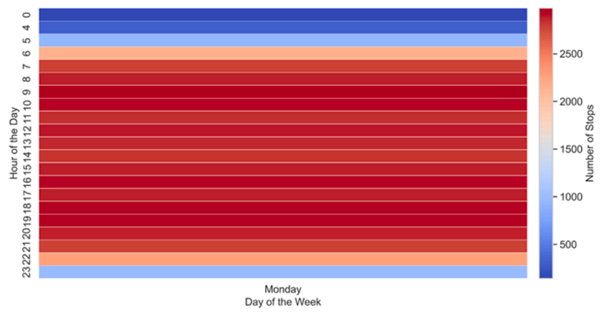


(d)

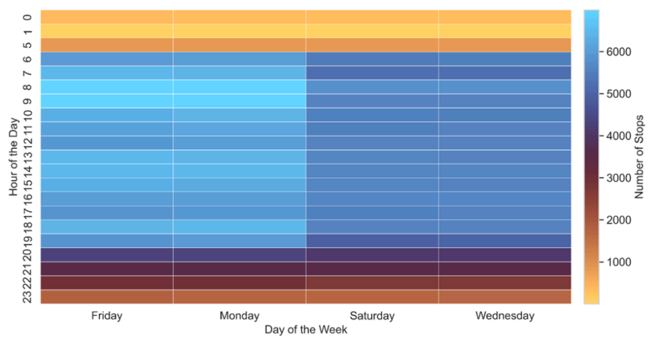
Figure 6. (*continued*).

4.2. Service Frequency Analysis

An analysis of service frequency provides temporal insights into the operational dynamics of the myBAS Johor and Rapid Bus Penang networks. By examining intra-day and inter-day patterns, this study evaluates the adequacy of transit coverage relative to urban mobility needs, with particular attention to peak-hour provision and off-peak service gaps. Heatmaps presented in Figure 7 visualize service intensities across hourly intervals and selected days of the week, enabling a granular assessment of public transport availability.



(a)



(b)

Figure 7. Temporal distribution of service frequency for public bus networks in Johor Bahru and Penang. (a) myBAS Johor (b) Rapid Bus Penang.

The service frequency heatmap for myBAS Johor, depicted in Figure 7(a), illustrates a structured temporal pattern concentrated within standard operational hours. The results indicate that service availability peaks consistently between 07:00 and 23:00, aligning with commuter demand driven by work, education, and general urban activity. This period is characterized by a dominant presence of warm colors reds and oranges that reflect a robust and steady transit supply. In contrast, early-morning (00:00–05:00) and late-night (post-23:00) intervals display cool tones, indicating a marked decline in service frequency. This suggests a structural gap in round-the-clock transit coverage, particularly affecting night-shift workers, airport travelers, and individuals reliant on late-night mobility options. Despite the strength of weekday operations, the near-absence of extended-hour service highlights a temporal inequity in the system's accessibility.

In Figure 7(b), It shows service variability across Friday through Monday, offering a nuanced view of weekday versus weekend dynamics. As with the Johor analysis, the vertical axis represents operational hours while the horizontal axis shows specific days. The color gradient, ranging from yellow and light blue (high frequency) to deep blue and purple (low frequency), delineates temporal intensity. Weekday service patterns mirror the Johor system with pronounced peaks during morning (07:00–09:00) and evening (17:00–19:00) rush hours. These peaks reflect conventional commuting behavior and are indicative of the network's alignment with daily workforce needs. However, what distinguishes the Penang network is a more diffuse and sustained frequency on weekends, particularly during evening hours between 19:00 and 23:00. This suggests a shift toward supporting leisure and recreational mobility, consistent with Penang's tourism-oriented and mixed-use urban character

Nevertheless, similar to myBAS Johor, transit service from midnight to early morning remains sparse, suggesting a common challenge across both systems in addressing the mobility needs of off-peak and late-night users. The service frequency analysis indicates that both transit systems myBAS Johor and Rapid Bus Penang are strategically structured to accommodate peak-hour demand, particularly on weekdays, thereby fulfilling their core function of facilitating commuter mobility. Nonetheless, temporal equity remains a significant planning challenge, particularly for users traveling during off-peak or unconventional hours.

The myBAS Johor network demonstrates consistent weekday operations, suggesting a degree of operational stability. However, its limited service adaptability during weekends and evening periods constrains its temporal inclusivity. In contrast, Rapid Bus Penang exhibits a comparatively more flexible temporal configuration, especially during weekends, reflecting its responsiveness to recreational and leisure-oriented mobility patterns. Despite this relative adaptability, both systems exhibit notable service gaps during late-night and early-morning hours—periods often critical for shift workers, service employees, and other transit-dependent populations.

From a transport planning and policy perspective, these temporal deficiencies represent systemic limitations in accessibility and inclusivity. Addressing these gaps necessitates targeted interventions. One key strategy involves extending late-night services along major corridors, particularly those linking employment hubs and entertainment districts, where mobility demands frequently extend beyond conventional hours. Complementarily, the deployment of demand-responsive transit (DRT) services during low-demand periods offers a flexible, cost-efficient mechanism to expand coverage without straining fixed-route resources [112].

In areas characterized by dynamic land-use or high tourist activity, especially during weekend evenings, optimizing service frequencies to reflect real-world temporal demand profiles would further enhance system responsiveness. Moreover, the integration of GTFS-RT data into service planning can support adaptive scheduling algorithms, allowing agencies to dynamically respond to ridership fluctuations and emerging temporal usage trends in near real-time.

Through the implementation of these empirically grounded strategies, both myBAS Johor and Rapid Bus Penang can improve temporal service equity, bolster off-peak mobility, and reinforce public transportation as a viable, inclusive mode for diverse user groups—irrespective of time of travel.

4.3. Stop-Level Network Structure

4.3.1. Top Transfer Points

The analysis of stop connectivity and transfer node hierarchies offers valuable perspectives on the network structure and functional capacity of the myBAS Johor and Rapid Bus Penang systems, informing both planning efficiency and service integration. By quantifying the number of unique routes converging at each transit hub, this analysis identifies the critical nodes that underpin passenger flow, support multimodal transitions, and influence system-wide efficiency and resilience.

In the myBAS Johor network, the spatial distribution and functional roles of transfer points are illustrated in Figure 8(a), which ranks the top ten hubs based on the number of intersecting routes. JB Sentral emerges as the most dominant node, servicing ten unique routes and operating as the central interchange for both urban and regional connectivity. This centrality aligns with its geographic placement within Johor Bahru's transit-oriented core, reinforcing its strategic importance in facilitating high-volume passenger transfers.

Other prominent hubs such as Pejabat Pos Besar and Wisma Peladang, with nine and eight intersecting routes respectively, form secondary convergence points that expand the network's reach and operational flexibility. Further down the hierarchy, Kompleks Pekan Rabu and Bandar Baru Uda with six to seven intersecting routes serve as vital peripheral nodes, bridging feeder and local routes to the main transit spine. While this nodal concentration supports route interconnectivity and travel efficiency, it also introduces risks of congestion and over-reliance. Hubs like JB Sentral are particularly susceptible to service bottlenecks during peak hours, which could compromise both user experience and operational stability.

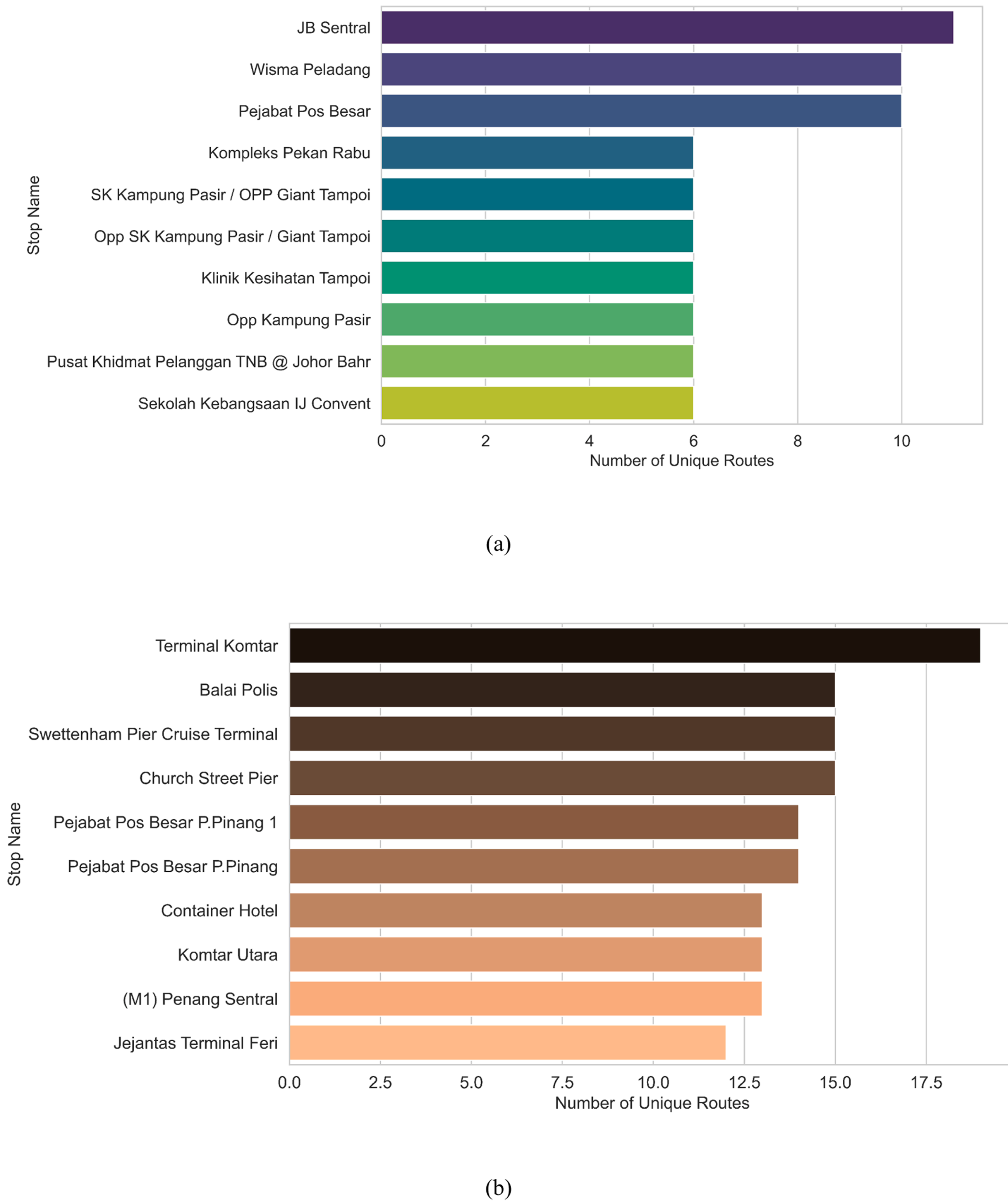


Figure 8. Top transfer points in the public bus networks of Johor Bahru and Penang, ranked by the number of unique intersecting routes. (a) myBAS Johor (b) Rapid Bus Penang.

Addressing these concerns requires targeted investment in infrastructure upgrades and capacity enhancements. Expanding terminal space, improving transfer amenities, and deploying intelligent passenger flow management systems at high-demand hubs would mitigate congestion risks. Simultaneously, bolstering the role of secondary hubs through increased service allocation and improved accessibility can foster a more balanced and resilient network topology.

A similar structural configuration is evident in Rapid Bus Penang, as depicted in Figure 8(b). Here, Terminal Komtar unequivocally functions as the principal transit hub, supporting more than seventeen unique routes. Its central role within Penang's mobility framework is underscored by its geographical location in George Town's dense commercial district, where multimodal integration particularly with pedestrian and tourism infrastructure is paramount.

Additional key transfer nodes such as Balai Polis, Church Street Pier, and Swettenham Pier Cruise Terminal, each accommodating between ten and twelve routes, form a robust second tier of strategic hubs. These stops connect urban, peri-urban, and coastal corridors, enabling dynamic passenger flows across different functional zones. Of particular interest is the pairing of Pejabat Pos Besar P. Pinang and Pejabat Pos Besar P. Pinang 1, which illustrates a deliberate planning approach whereby spatially proximate stops are designed to operate in tandem. This strategy distributes passenger loads, reduces congestion at singular points, and promotes smoother intermodal transitions.

Furthermore, secondary hubs such as Container Hotel, Komtar Utara, and Jejantas Terminal Feri, supporting six to eight intersecting routes, play a pivotal role in maintaining connectivity across lesser-served urban segments. These nodes facilitate localized access and reinforce the horizontal expansion of the network, ensuring transit services extend beyond the urban core to support inclusive and equitable mobility.

The combined findings from Johor and Penang emphasize the dual importance of highly connected primary hubs and strategically integrated secondary nodes. Infrastructure investment in major hubs particularly in terminal design, real-time passenger information systems, and multimodal access is crucial for accommodating growing ridership and improving service fluidity. At the same time, enhancing the operational profile of secondary transfer points can decentralize passenger flows and reduce over-dependence on a limited number of interchanges.

In addition to structural interventions, soft infrastructure improvements such as wayfinding systems, sheltered waiting areas, and pedestrian-friendly environments are essential for maximizing the usability of transfer nodes. Incorporating first-mile and last-mile solutions such as bicycle lanes, shared mobility services, or dedicated pedestrian corridors will further increase the functional accessibility of these hubs.

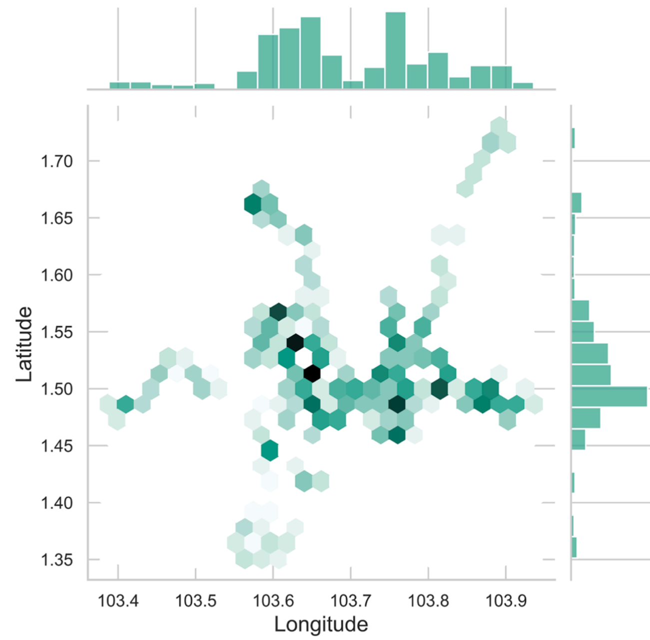
In Sum, an optimized hierarchy of transfer points contributes not only to operational efficiency and travel time reduction, but also to network resilience and urban livability. By adopting a holistic approach to transfer hub development, the myBAS Johor and Rapid Bus Penang networks can enhance system performance, support sustainable growth, and promote equitable access for diverse user groups.

4.3.2. Stop Connectivity Distribution

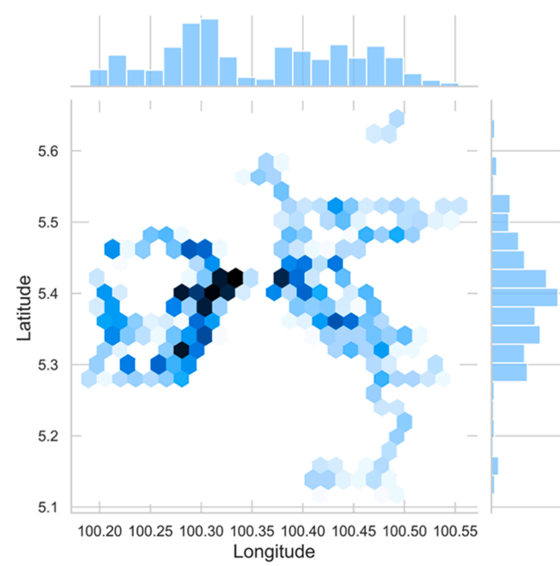
The spatial configuration of transit stops is a foundational element in understanding how well public transport networks serve their urban and regional landscapes. Through spatial analysis of stop distribution, transit agencies can identify clusters of connectivity, geographic gaps, and areas with latent demand. We used hexbin density plots and marginal histograms to assess the spatial dispersion of stops in the myBAS Johor and Rapid Bus Penang networks, thereby informing potential strategies for network optimization and spatial equity enhancement.

In myBAS Johor, the spatial distribution of stops, shown in Figure 9 (a), reveals a marked concentration around JB Sentral and its surrounding urban core. Darker hexagonal bins indicate areas of high stop density, corresponding with zones of intensified economic activity, population density, and established infrastructure. The accompanying marginal density plots along the longitudinal and latitudinal axes confirm this central clustering, reflecting the gravitational pull of central business districts in shaping transit demand.

Conversely, the peripheral and semi-rural zones exhibit significantly lower stop densities, visualized by lighter hexagons and sparse marginal densities. These underrepresented areas signal potential service deserts, where low transit availability may constrain accessibility for transit-dependent populations. Addressing these spatial disparities may involve targeted route extensions, infill stops, or the deployment of demand-responsive services tailored to low-density environments.



(a)



(b)

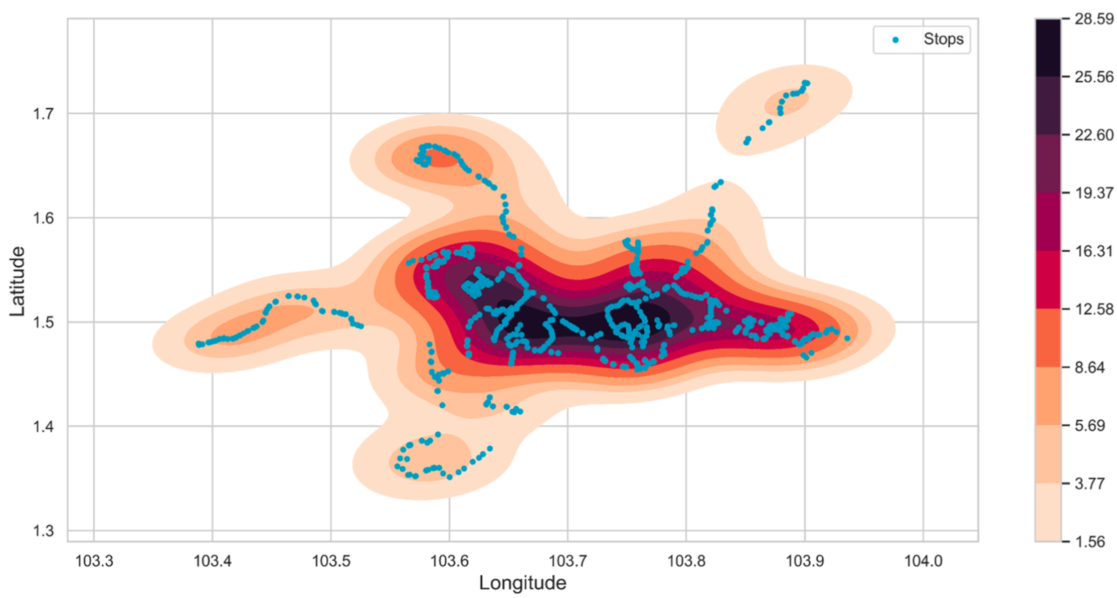
Figure 9. Connectivity distribution of (a) myBAS Johor and (b) Rapid Bus Penang.

Moreover, some high-density clusters exhibit limited route diversity, suggesting that stop saturation alone does not equate to optimal connectivity. In such cases, network performance could be enhanced through route diversification, intermodal integration, and improved headway reliability.

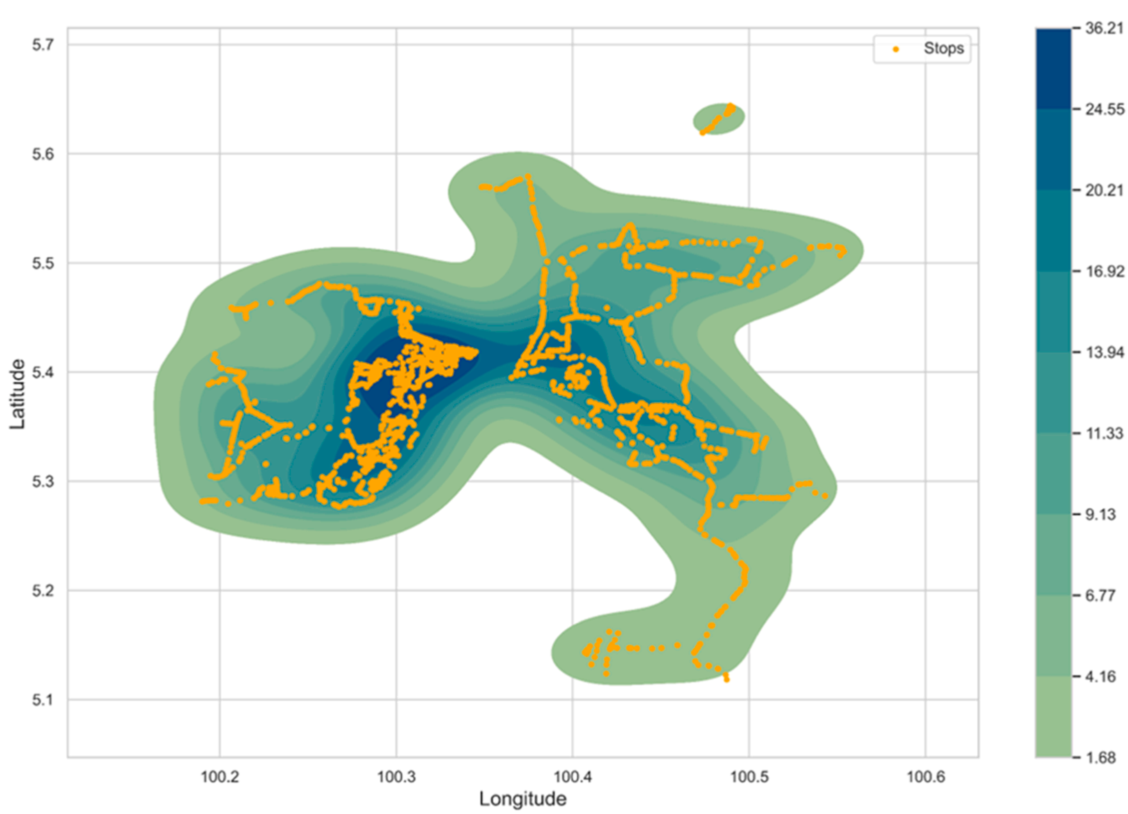
The Rapid Bus Penang network exhibits a comparable spatial profile, as shown in Figure 9(b). Transit stops are densely clustered in and around George Town, with Terminal Komtar emerging as a dominant connectivity node. Darker hexbin clusters reflect the high concentration of stops in core commercial and mixed-use districts. Meanwhile, areas on the urban periphery including parts of Balik Pulau and northern Penang Island—display sparse stop density, suggesting a similar issue of spatial inequity.

The marginal histograms reinforce this urban-rural disparity, illustrating an asymmetric distribution of service coverage. While this pattern may reflect rational resource allocation based on population distribution and demand elasticity, it also highlights the challenge of ensuring universal accessibility. In particular, rapidly urbanizing fringe areas may require proactive planning to avoid future transit deficiencies. Additionally, dense transit zones such as George Town may experience operational inefficiencies due to stop redundancy or congestion, necessitating interventions such as stop consolidation, streamlined route planning, or timed-transfer scheduling.

Overall, these spatial analyses yield important planning



(a)



(b)

Figure 10. Spatial distribution and density of transit stops for (a) myBAS Johor and (b) Rapid Bus Penang.

implications. Both systems demonstrate a radial structure, with high-density central nodes tapering into under-served peripheral zones. Although such a structure is efficient for unidirectional peak flows, it may inadequately support reverse commuting, cross-town trips, or multimodal continuity. Enhancing connectivity in underserved areas not only improves geographic coverage but also promotes social equity by reducing the mobility burden on vulnerable populations.

To improve network resilience and rider experience, planners should prioritize infrastructure investment in peripheral growth areas, integrate pedestrian and first-mile/last-mile connectivity, and ensure that high-demand corridors are supported by high-frequency, reliable service. Moreover, integrating multimodal planning combining fixed-route bus systems with feeder services, bike infrastructure, and walkable access can extend the effective reach of the network without requiring extensive physical expansion.

Overall, leveraging these spatial insights allows transit planners to move beyond reactive network maintenance toward proactive, data-informed planning. A well-distributed and optimally connected stop network is essential not only for operational efficiency but also for achieving broader goals of sustainable urban mobility and inclusive access.

4.3.3. Spatial Distribution of Transit Stops

The spatial distribution of transit stops serves as a vital indicator of a public transport network's coverage, accessibility, and operational efficiency. By applying KDE heatmaps integrated with GTFS static data, this section provides a spatially continuous perspective on stop density patterns across myBAS Johor and Rapid Bus Penang, identifying transit-rich zones, underserved regions, and potential misalignments in stop placement. The overlaid scatter plots further contextualize individual stop locations within these density landscapes, offering planners a dual-scale view of macro-level coverage and micro-level placement.

In the myBAS Johor network, the KDE heatmap in Figure 10(a) demonstrates a clear spatial asymmetry in transit accessibility. High stop density, illustrated by warmer tones such as yellow and green, is primarily concentrated in central Johor Bahru, which encompasses key commercial corridors, government institutions, and multimodal nodes such as JB Sentral. This concentration reflects a demand-responsive pattern driven by population agglomeration and concentrated economic activity. Red scatter points representing individual stops reinforce this clustering, highlighting both well-served corridors and zones of potential redundancy where closely spaced stops may impede operational speed without significant gains in accessibility.

In contrast, cooler tones ranging from light blue to purple mark the peripheral suburban and rural regions, where transit coverage diminishes significantly. These areas often lack continuous service corridors and are punctuated by sparse or isolated stops. A few scattered clusters in northern Johor appear disconnected from the central network, suggesting inefficiencies in route integration and potential overextension without corresponding feeder support. To address these spatial gaps, strategic interventions such as the deployment of feeder services, demand-responsive transit (DRT), or route reconfiguration may enhance connectivity while maintaining cost-effectiveness.

A comparable distributional structure is observed in Rapid Bus Penang, as shown in Figure 10(b). The KDE heatmap, superimposed with dark turquoise scatter points, reveals that central George Town retains the highest stop density within the system. This core zone exhibits short inter-stop distances and strong multimodal linkages, consistent with the city's mixed-use urban fabric and high pedestrian activity. These areas, shaded in deep blue to purple, underscore the role of high-frequency transit corridors in sustaining compact, transit-oriented development patterns.

Peripheral zones including parts of Balik Pulau, Batu Maung, and rural segments of Seberang Perai exhibit much lighter hues on the density gradient. Although some of these regions contain isolated stops, their weak integration into the network raises concerns about spatial equity and transit reliability. The presence of stops in low-density catchments, without complementary routing or frequency support, may indicate underperforming segments or legacy routing decisions that are misaligned with current demand. These inefficiencies could be mitigated through route rationalization, reallocation of operational resources, or introduction of scalable alternatives such as shared mobility integration.

Across both cities, the KDE analysis emphasizes the need for a balanced spatial approach to transit planning. In urban cores, where stop density is already high, emphasis should shift toward service optimization, stop consolidation where appropriate, and infrastructure upgrades—such as real-time information systems and pedestrian enhancements—to manage volume and maintain service quality. In contrast, rural and peri-urban areas require attention to coverage expansion, through either extending fixed routes or employing flexible service models suited to low-demand environments.

4.3.4. Distribution of Stop Visits per Route

Analyzing the distribution of stop visits across individual routes provides critical insights into patterns of passenger activity, service utilization, and operational balance within a public transit network. These distributions not only reflect demand intensity but also reveal underlying issues related to service redundancy, route configuration, and accessibility disparities. By visualizing these distributions for myBAS Johor and Rapid Bus Penang, this section identifies high-demand corridors, underutilized segments, and strategic intervention points that inform network-level optimization.

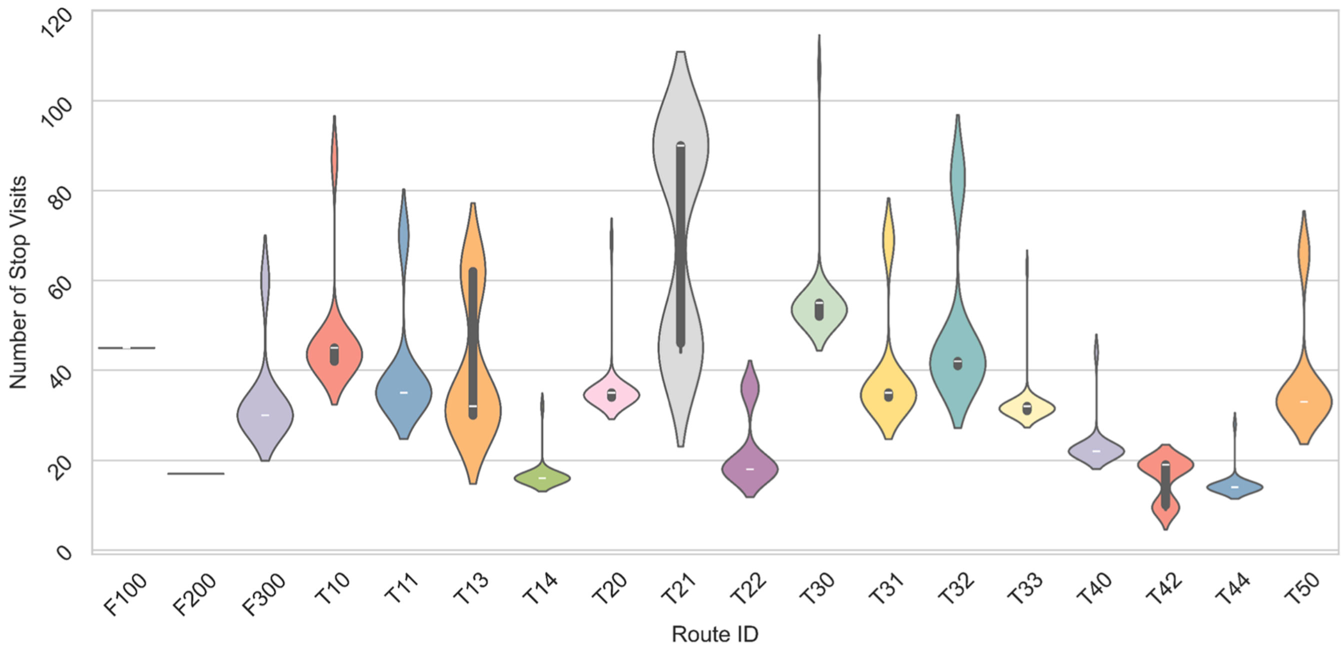
In the case of myBAS Johor, the distribution of stop visits is illustrated in Figure 11(a) using density plots for selected routes. The data reveal substantial variability across routes, with some exhibiting broad distributions such as T21 and T32 indicating heterogeneity in stop-level demand. These broad patterns suggest the presence of a small number of stops experiencing disproportionately high usage, likely corresponding to central business districts, educational institutions, or key intermodal terminals. In contrast, other stops on these routes show minimal activity, possibly due to lower residential densities, inadequate pedestrian access, or competition from alternative transport modes.

More uniform distributions, as observed in routes like T14 and T44, reflect consistent stop utilization. These routes are typically associated with homogeneous land use or evenly distributed demand across their catchment areas, facilitating predictable scheduling and resource efficiency. Notably, outliers and long tails, particularly evident in routes such as T13, point to anomalous usage patterns, where certain stops may cater to landmark-driven or event-based demand. In such cases, fixed-route services may be inefficient, and a hybrid model combining core service with flexible scheduling during peak hours may be more appropriate.

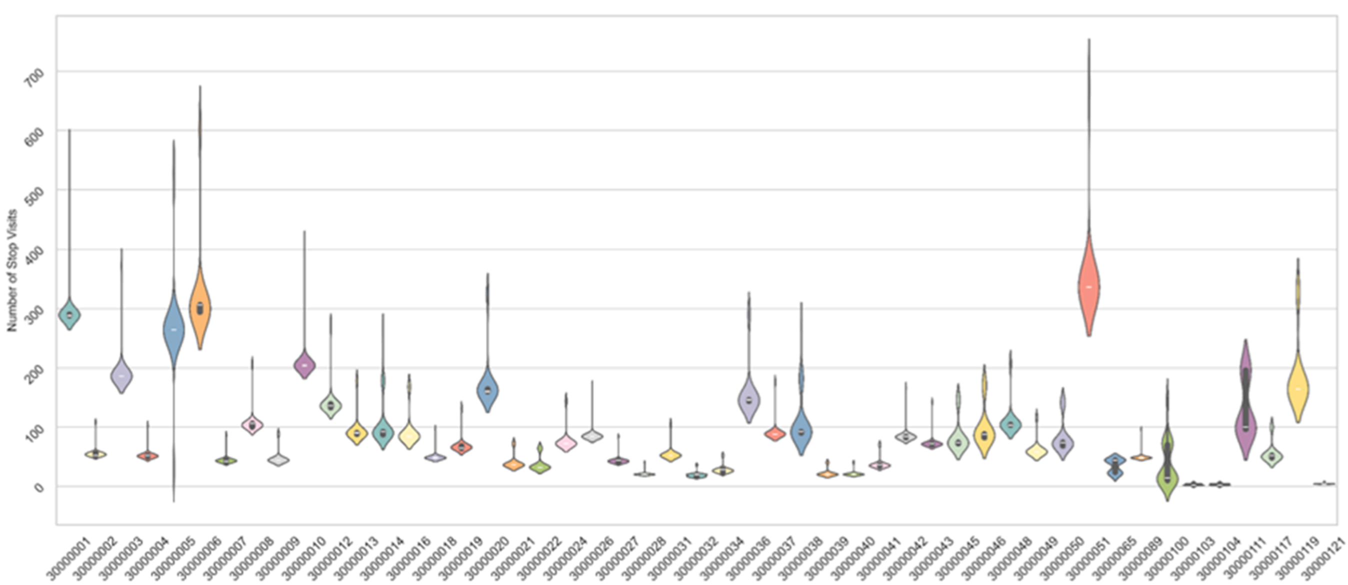
The presence of multimodal distributions, where multiple peaks are observed within a single route, reveals temporal or spatial variability in usage. These patterns may stem from the interaction of weekday commuter traffic and weekend leisure demand, or from transitions between urban and suburban environments. To address this complexity, service planners could implement differentiated headways, express variants, or dynamic frequency adjustments based on temporal demand profiling.

Turning to Rapid Bus Penang, Figure 11(b) presents violin plots capturing the density and spread of stop visit frequencies across multiple routes. Routes such as 30000005 and 30000035 exhibit elongated violin shapes, denoting considerable variance and highlighting the coexistence of highly frequented and sparsely used stops. These disparities are often linked to the dual function of such routes in connecting dense urban centers with peripheral areas, where stop-level demand fluctuates significantly.

In contrast, routes like 30000007 and 30000016 demonstrate compact, symmetric distributions, suggesting uniform demand profiles that support standard scheduling protocols. The internal consistency of



(a)



(b)

Figure 11. Distribution of stop visit frequencies for (a) myBAS Johor and (b) Rapid Bus Penang.

these routes presents opportunities for service enhancement through increased frequency, extended service hours, or the deployment of higher-capacity vehicles to accommodate latent demand.

Outliers embedded within the violin plots further highlight the inequity in stop-level utilization. High-traffic stops may require infrastructure investments, such as expanded waiting areas, real-time information systems, and enhanced pedestrian facilities, to accommodate growing volumes. Conversely, persistently underutilized stops might warrant closer scrutiny regarding their location, surrounding land use, or accessibility. Solutions may involve rerouting, better wayfinding, or targeted marketing efforts to raise awareness and improve ridership.

The findings from both cities reinforce the utility of stop-level analysis as a foundation for data-driven service design. Routes characterized by high variability demand flexible operational strategies, including the implementation of express services, skip-stop models, or time-based service differentiation. Stable-demand routes provide a more predictable environment for incremental improvements in capacity and reliability. Ultimately, aligning service supply with spatial and temporal demand enhances network efficiency, supports sustainable operations, and fosters a more user-centered transit experience.

4.4. Route Overlapping

Understanding the extent and structure of route overlaps within a public transit network is essential for evaluating operational efficiency, resource allocation, and service equity. Overlapping routes can signal either coordinated high-frequency service in high-demand corridors or unintended redundancy that leads to inefficiency. Conversely, areas with minimal overlap may reflect coverage gaps or under-resourced segments. We utilized network-based visualizations to assess route overlaps in myBAS Johor and Rapid Bus Penang, facilitating targeted recommendations for network optimization.

In the case of myBAS Johor, route overlap is visualized through a route-level network graph presented in Figure 12(a). Here, each node corresponds to a transit route, while edges represent shared stops between routes. The size of each node indicates the number of overlapping connections it maintains, while edge thickness is proportional to the Jaccard similarity coefficient, providing a quantitative measure of route redundancy.

Notably, routes such as T30, T31, and T40 form densely interconnected clusters, indicating substantial overlap in the central urban core. While this configuration suggests enhanced service frequency in high-demand zones such as Johor Bahru's commercial and administrative districts it may also imply inefficiencies where multiple routes serve similar passenger flows without sufficient differentiation. This phenomenon often leads to over-servicing along certain corridors while simultaneously limiting service in other areas.

On the other hand, peripheral routes such as F300 and T42 display minimal connectivity with other services, highlighting their roles in extending coverage to low-density or spatially isolated regions. These routes are vital for ensuring network reach and inclusivity, particularly for populations in transit-dependent suburban and semi-rural zones. The spatial positioning of the nodes, determined by the geometric centroids of their stop distributions, reinforces this contrast: central clusters serve densely built-up areas, whereas peripheral nodes extend to urban edges with minimal redundancy.

Addressing these dynamics requires a dual approach to route optimization. In areas of high overlap, targeted interventions—such as route consolidation, coordinated headways, or differentiated service patterns (e.g., limited-stop or express services)—can mitigate redundancy and reallocate capacity more effectively. Conversely, routes with low overlap but high geographic importance should be safeguarded and potentially reinforced through increased frequency, improved feeder integration, and investment in supporting infrastructure.

For the Rapid Bus Penang network, route overlap is presented in Figure 12(b) through a stop-based node-link diagram. In this figure, transit stops are represented as circles, where the size of each circle indicates the number of routes serving that location. Lines connecting the circles denote route intersections and shared service corridors. The core area of George Town emerges as a high-density hub, with multiple routes converging at key nodes such as Terminal Komtar, reflecting a planned effort to facilitate transfers and concentrate service where demand is greatest.

However, excessive convergence at central stops introduces challenges related to congestion, dwell time delays, and scheduling inefficiencies—particularly during peak commuting hours. Redundant routing can lead to diminishing returns, where additional services no longer contribute meaningfully to capacity or reliability but instead overburden limited infrastructure. In contrast, peripheral zones—particularly on the island's northern and western flanks, exhibit smaller, disconnected nodes with limited route overlap. These low-connection areas suggest limited-service availability and highlight ongoing challenges in achieving geographic equity in transit provision.

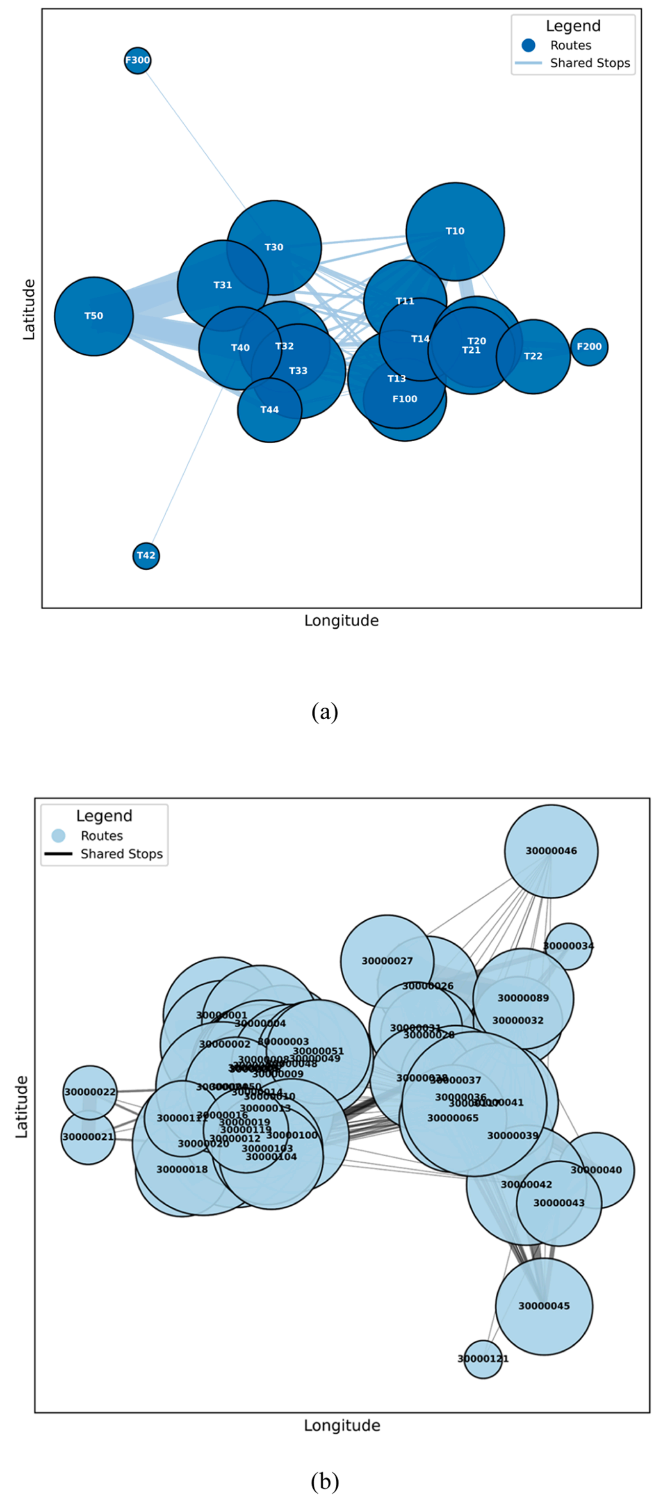


Figure 12. Route overlapping visualization for (a) myBAS Johor and (b) Rapid Bus Penang.

Strategically managing route overlap in Penang requires a rebalancing of resources. In urban cores, rationalization of overlapping services—through stop spacing adjustment, coordinated timetables, or restructured route geometries—can alleviate crowding and reduce operational strain. Simultaneously, improving connectivity in under-served districts through route expansion, express extensions, or mobility-on-demand services would extend the functional reach of the network and enhance regional accessibility.

Across both systems, the findings reinforce the importance of balancing frequency with spatial equity. High-overlap corridors should be preserved where justified by demand but optimized to avoid duplication. Meanwhile, areas with minimal overlap must receive proportionate attention to ensure that the benefits of urban transit extend beyond central nodes. Data-informed planning tools such as route clustering algorithms, accessibility audits, and demand forecasting models—can support this rebalancing by identifying where marginal gains in coverage or efficiency can yield the highest system-wide benefits.

4.5. Pedestrian Accessibility Assessment

Pedestrian accessibility to transit stops often referred to as the "first-mile" problem is a key determinant of public transport usability and inclusivity. We assessed the ease with which individuals can access transit infrastructure from major urban amenities, such as residential clusters, commercial zones, healthcare facilities, and educational institutions. By applying an Accumulated Cost Surface (ACS) methodology, calibrated using GTFS static data and a uniform walking speed of 5 km/h, this study analyzed spatial variations in pedestrian accessibility for myBAS Johor and Rapid Bus Penang.

4.5.1. Public Transit Stop Accessibility

In the myBAS Johor network, pedestrian accessibility is visualized in Figure 13(a) using travel time intervals to the nearest transit stop, highlighting spatial disparities in walkable access. The spatial outputs are categorized into five discrete walking time zones: less than 5 minutes (yellow), 10 minutes (orange), 15 minutes (pink), 30 minutes (purple), and more than 30 minutes (blue). This classification framework enables the identification of core-accessibility zones as well as areas of marginal access.

The analysis reveals that central Johor Bahru demonstrates high levels of accessibility, with most areas falling within the 5–10-minute walking range. These zones are typically characterized by mixed land use, higher population density, and concentrated commercial activity, which justifies the closely spaced and frequently serviced transit stops. However, a sharp gradient of declining accessibility emerges in the suburban and peri-urban fringes, where walking times to the nearest transit stop often exceed 15–30 minutes. These peripheral zones, shaded in purple and blue, signal limited transit penetration and highlight a spatial mismatch between residential development and transit infrastructure planning.

A parallel analysis for the Rapid Bus Penang network, illustrated in Figure 13(b), reveals a broadly similar pattern. Centralized areas particularly within George Town and the Bayan Lepas corridor exhibit robust accessibility, with most POIs falling within a 10-minute walking range. These zones benefit from dense transit stop deployment and well-integrated pedestrian networks. Conversely, the accessibility map of rural and low-density regions—particularly in Balik Pulau and parts of Seberang Perai—reveals significant travel time barriers, with isochrone bands extending beyond 30 minutes. These findings point to structural transit inequities, wherein core zones are served by overlapping transit routes while peripheral areas remain spatially isolated.

This disparity in walking-time accessibility highlights the uneven first-mile connectivity landscape across both systems. In urban centers, high-accessibility zones may still benefit from infrastructure enhancements, such as sheltered walkways, tactile paving for universal access, and real-time wayfinding systems. Meanwhile, in under-served areas, where walk times exceed 15 minutes, targeted interventions are critical. These may include the deployment of feeder services, new stop placements, or mobility-on-demand platforms to bridge the last-mile gap.

Importantly, the accessibility differentials observed in both cities are not merely operational inefficiencies; they constitute a broader equity concern, with direct implications for mobility-impaired individuals, elderly residents, and economically disadvantaged populations. Enhancing walkable access to transit thus represents a pivotal strategy for promoting inclusive, sustainable urban transport.

4.5.2. Spatial Accessibility Modelling to Key Urban Points of Interest

Equitable access to essential urban amenities, including healthcare facilities, educational institutions, commercial centers, and residential zones, constitutes a core dimension of transport justice and urban livability. The ability of individuals to reach these essential destinations efficiently reflects the performance of the public transport system not only in terms of operational design but also in terms of social inclusion. Pedestrian accessibility to key POIs is analyzed using a Pedestrian Accessibility Model, which estimates cumulative walking travel time from bus stops to destination points. The travel cost surfaces are visualized through isochrone maps with five travel-time intervals ($\leq$15 minutes, $\leq$30 minutes, $\leq$45 minutes, $\leq$60 minutes, and >60 minutes), offering a multi-scalar view of service coverage in myBAS Johor and Rapid Bus Penang networks.

- **Hospital Accessibility:** Access to healthcare is critical for both emergency response and routine health services, making hospital accessibility a high-priority metric in transit system evaluation. Using accumulated walking time surfaces, hospital access zones were delineated and color-coded (yellow to purple), as illustrated in Figure 14.

In Figure 14(a), central areas of Johor Bahru exhibit the most favorable accessibility conditions, with large swaths of territory falling within the 15-minute isochrone band. This reflects a confluence of dense transit coverage, higher urban compactness, and the spatial co-location of healthcare institutions with transport infrastructure. However, accessibility declines rapidly in peripheral and semi-urban zones. Outer districts such as Skudai and Tebrau show walking times in excess of 45 to 60 minutes, revealing critical access deficiencies for vulnerable populations, including the elderly, those without private vehicles, and low-income residents.

Figure 14(b) illustrates a similar spatial structure in Penang. George Town and the Bayan Lepas corridor exhibit high accessibility, while areas such as Balik Pulau and inland Seberang Perai demonstrate longer walking times to hospitals. These findings point to entrenched spatial disparities in transit-enabled healthcare access. Addressing these deficiencies requires targeted planning interventions, including bus route extensions, feeder services to healthcare clusters, and pedestrian infrastructure improvements such as walkways and safety crossings.

- **School Accessibility:** Spatial accessibility to schools is a critical aspect of equitable urban development, particularly for children from transit-dependent households. In Johor Bahru, Figure 15(a) shows that most schools in the urban core are within a 15–30-minute walking distance from the nearest transit stop. However, schools in suburban areas such as Ulu Tiram, Kempas, and Senai fall into the 45-minute or longer access bands. These patterns reflect limited stop density, fragmented transit routing, and poor-quality pedestrian infrastructure.

Figure 15(b) demonstrates a comparable pattern in Penang, where schools located in central urban districts exhibit strong accessibility, whereas those situated in peri-urban zones remain significantly under-served by the existing transit network. These spatial barriers affect school attendance reliability and long-term educational equity.

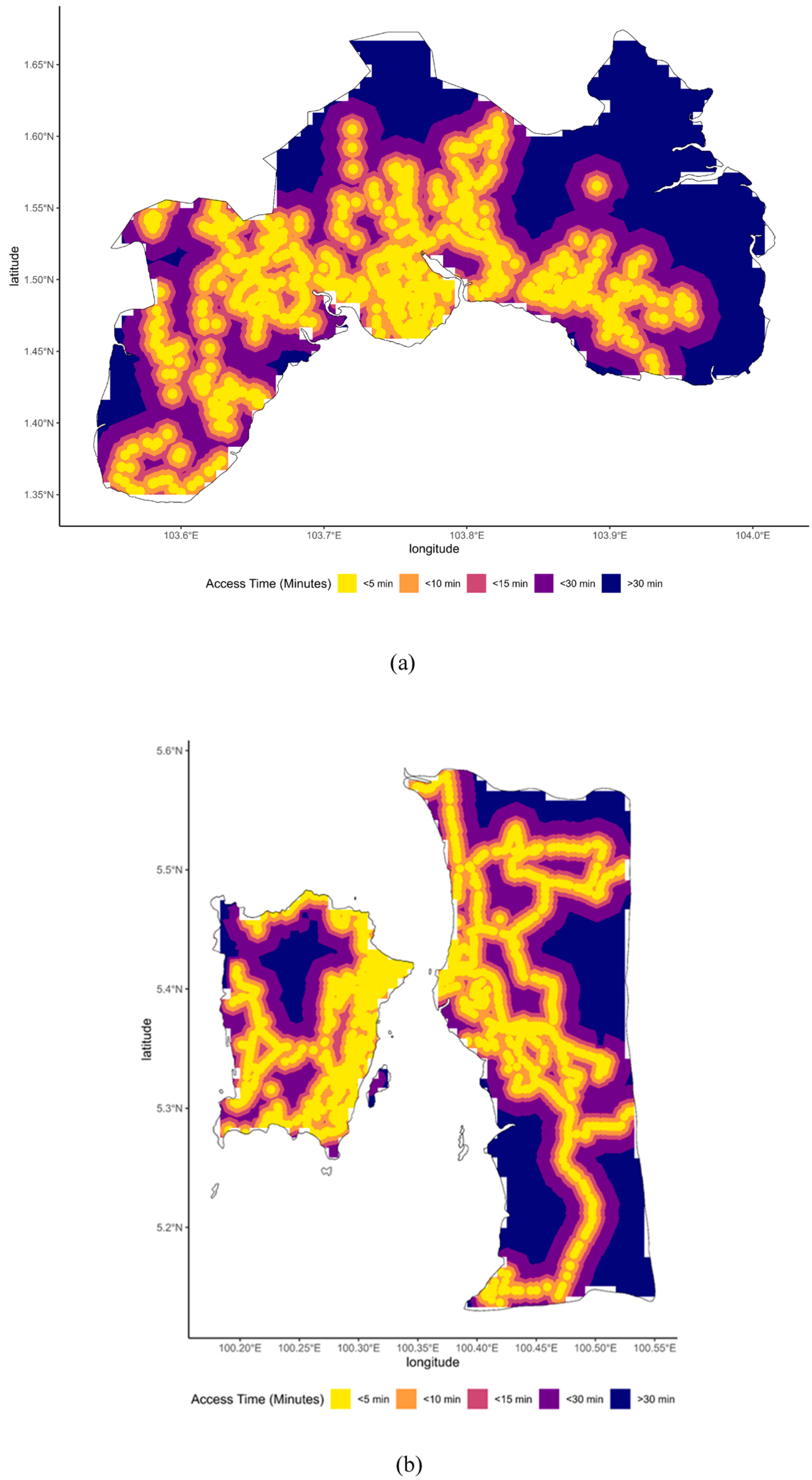
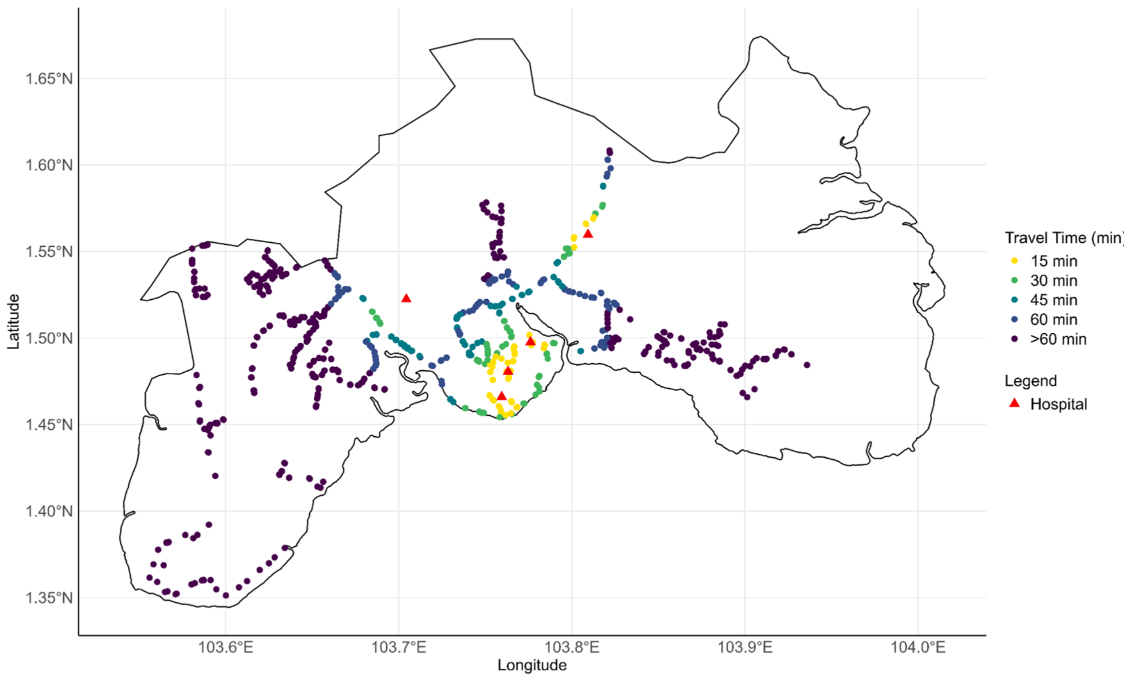
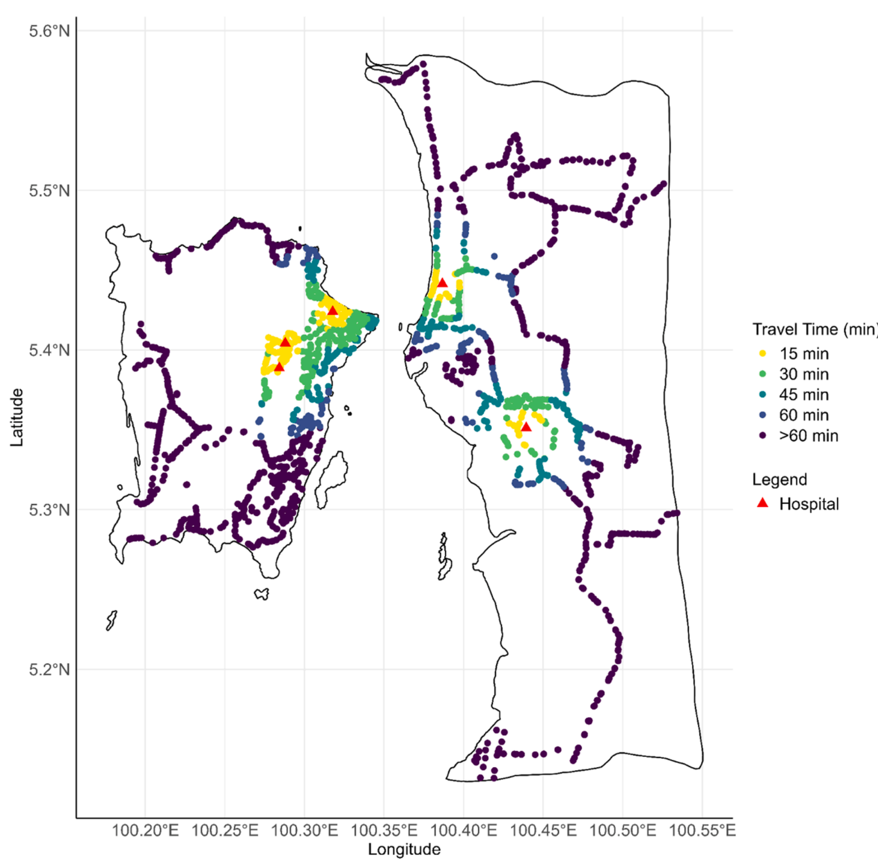


Figure 13. Accessibility zones to public bus stops (a) myBAS Johor and (b) Rapid Bus Penang based on pedestrian travel time.

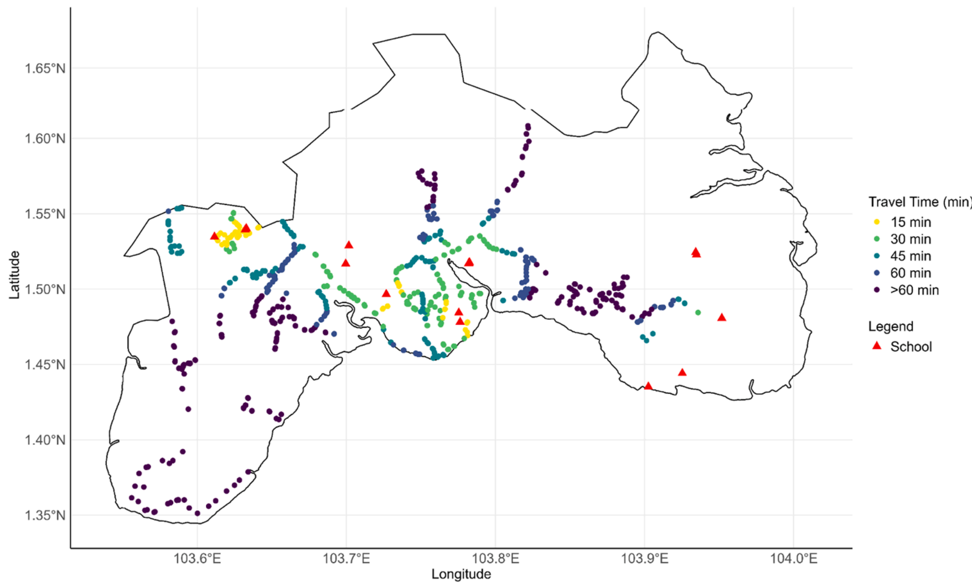


(a)

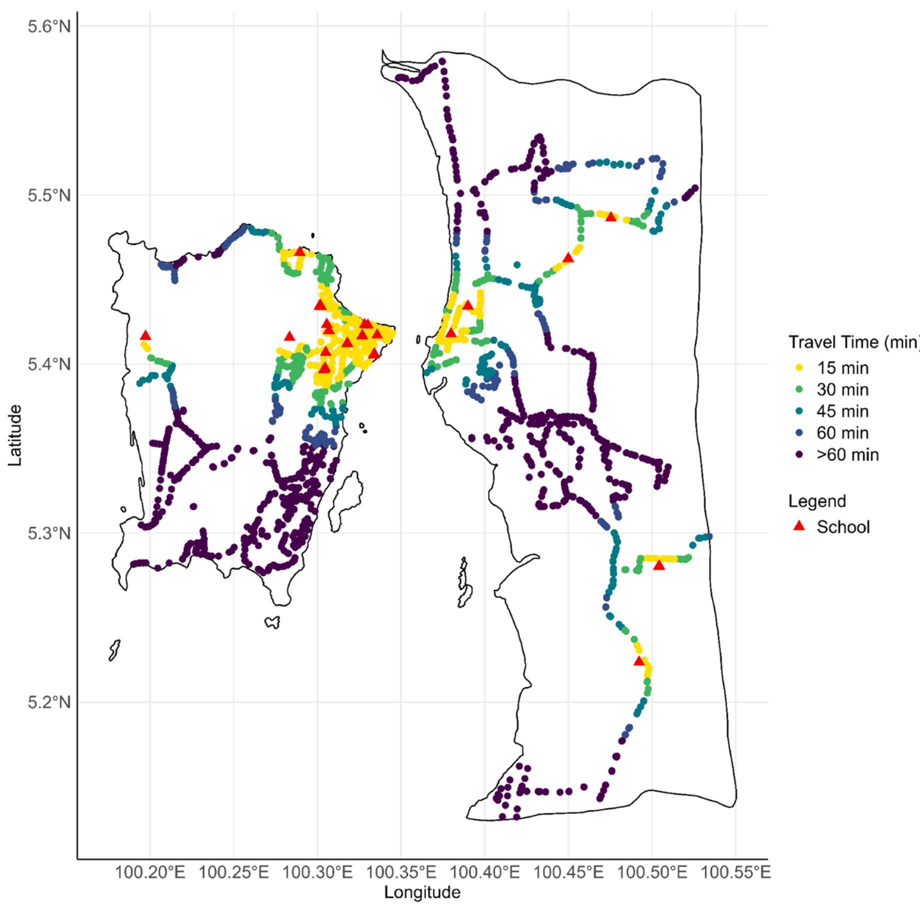


(b)

Figure 14. Pedestrian Travel time to the nearest hospital from public bus stops (a) myBAS Johor and (b) Rapid Bus Penang.



(a)



(b)

Figure 15. Pedestrian Travel time from schools to the nearest bus stops (a) myBAS Johor and (b) Rapid Bus Penang.

Addressing them will require school-focused transit strategies, including peak-period service integration, realignment of stops closer to school gates, and universal pedestrian design standards.

Mall Accessibility: Ensuring equitable access to commercial and retail hubs is vital for fostering economic participation, especially among transit-dependent populations whose livelihoods and daily activities are closely tied to public transport availability. Figure 16(a) indicates strong connectivity to central retail hubs such as City Square, KSL City, and AEON Tebrau, where travel times are generally under 30 minutes. However, accessibility declines markedly in peripheral commercial clusters, where low-density land use and poor pedestrian connectivity produce functional transport deserts.

Figure 16(b) shows that George Town in Penang enjoys high accessibility to major shopping destinations like Gurney Plaza and Prangin Mall. In contrast, commercial centers in more peripheral zones such as Design Village in Batu Kawan and Sunshine Square in Bayan Baru are significantly less accessible due to long transfer distances and inadequate last-mile infrastructure. These findings highlight the need for better integration of transit planning with commercial development policy, including strategic stop placement, route extension to emerging retail nodes, and the enhancement of pedestrian access facilities.

- ***Residential Area Accessibility:*** Accessibility to residential areas is a key component of inclusive urban mobility, especially for populations that rely on public transport for daily activities.

Figure 17(a) shows that central areas of Johor Bahru exhibit high levels of accessibility to residential zones, with most areas falling within the yellow and light green bands, indicating walking times under 30 minutes. However, peripheral neighborhoods particularly in southern and western sectors are characterized by significantly lower accessibility. These zones appear in blue and purple, reflecting walking times exceeding 45 to 60 minutes and revealing gaps in stop coverage and pedestrian connectivity. Figure 17(b) presents a spatial distribution in Penang. While the urban core offers good access to residential zones, outer areas such as those in Seberang Perai face longer walking distances and limited transit integration. These disparities highlight systemic deficiencies in linking transit infrastructure with housing in low-density or expanding urban zones.

To improve residential accessibility, targeted strategies are needed, including the extension of bus routes to underserved areas, increases in service frequency, and improvements in pedestrian infrastructure—such as safe crossings, sidewalks, and lighting. Such interventions are essential for ensuring equitable access to housing and fostering spatial inclusion across urban regions.

The spatial modeling of POI accessibility across hospitals, schools, malls, residential area in both myBAS Johor and Rapid Bus Penang reveals consistent structural inequities in pedestrian access to essential services. High-accessibility zones are concentrated in central, transit-rich districts, whereas outlying areas remain underserved. These disparities are not merely operational inefficiencies but reflect deeper spatial-justice concerns in transit planning, where mobility opportunities—and by extension, access to education, health, and economic life—are unevenly distributed across space.

5. Discussion

The analysis of myBAS Johor and Rapid Bus Penang presents a nuanced understanding of the operational, structural, and spatial dynamics underpinning Malaysia's urban bus transit systems. By integrating methods from network science, spatial accessibility modeling, and urban systems theory, this study illuminates the multifaceted nature of transit performance across both core and peripheral urban contexts. The findings offer critical insight into how centrality-driven efficiency in high-demand corridors coexists with and are often counterbalanced by inequities in access, service fragility in low-density zones, and route-level inefficiencies. This discussion unpacks these interconnected challenges, framing them within broader debates on transit equity, resilience, and sustainable urban mobility. It further proposes a set of policy, design, and research responses capable of addressing systemic vulnerabilities and supporting a more inclusive and adaptive public transport paradigm in Malaysia.

5.1. Network Dynamics and Centrality Structures

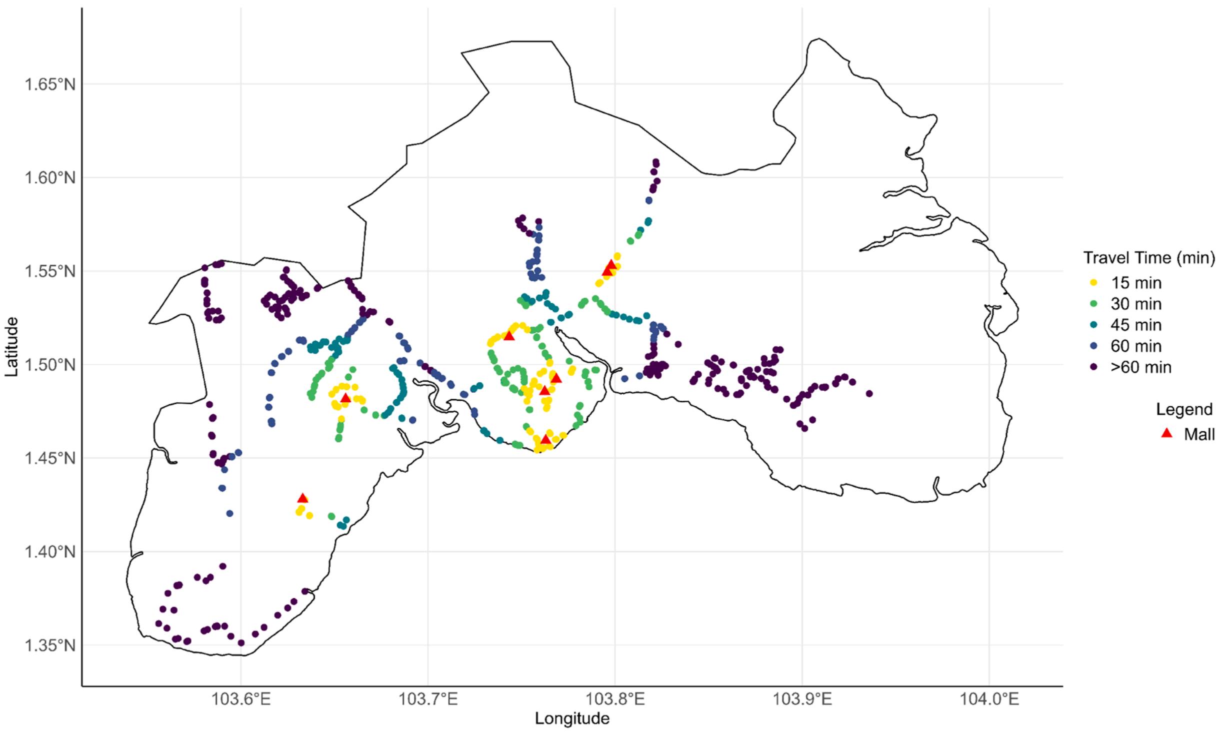
The centrality analysis focusing on degree, betweenness, and closeness confirms that both systems rely heavily on a limited set of strategically positioned transfer hubs. In Johor Bahru, Larkin Sentral consistently ranks highest across all centrality measures, functioning as a super-node that facilitates both local and intercity movement. Similarly, Terminal Komtar in Penang emerges as a critical convergence point for multiple high-frequency routes. These hubs exhibit high betweenness centrality, indicating their pivotal role in mediating passenger flow and inter-route connectivity. However, such topological centralization, while efficient for throughput in core areas, introduces systemic fragility: any disruption at these high-load nodes due to maintenance, congestion, or external shocks could have cascading effects across the network.

Moreover, the structural topology of both networks can be characterized as semi-hierarchical. While urban cores exhibit a hub-and-spoke configuration typified by centralized transfer points and radial route geometries outlying areas reveal more fragmented or linear structures. This hybrid topology does support flexible passenger movement in high-demand corridors and facilitates high-frequency circulation in central business districts. However, the lack of network redundancy in suburban and peri-urban zones limits adaptability. In practical terms, this means that outer zones often rely on a single or limited set of routes, with minimal options for rerouting or alternative access in the event of delays or interruptions. The resulting network design is efficient in terms of coverage compactness but lacks resilience, rendering it susceptible to localized disruptions. This pattern is consistent with other rapidly urbanizing contexts in Southeast Asia, where legacy transit systems evolve around economic cores and respond slowly to decentralized growth [113,114]. In both Johor and Penang, route convergence is overly concentrated within a small set of transfer nodes, while peripheral areas exhibit high impedance values indicating low nodal centrality and limited interconnectivity. This spatial imbalance raises important questions about network modularity and the capacity of the system to absorb demand shocks or reroute flows during service disruptions.

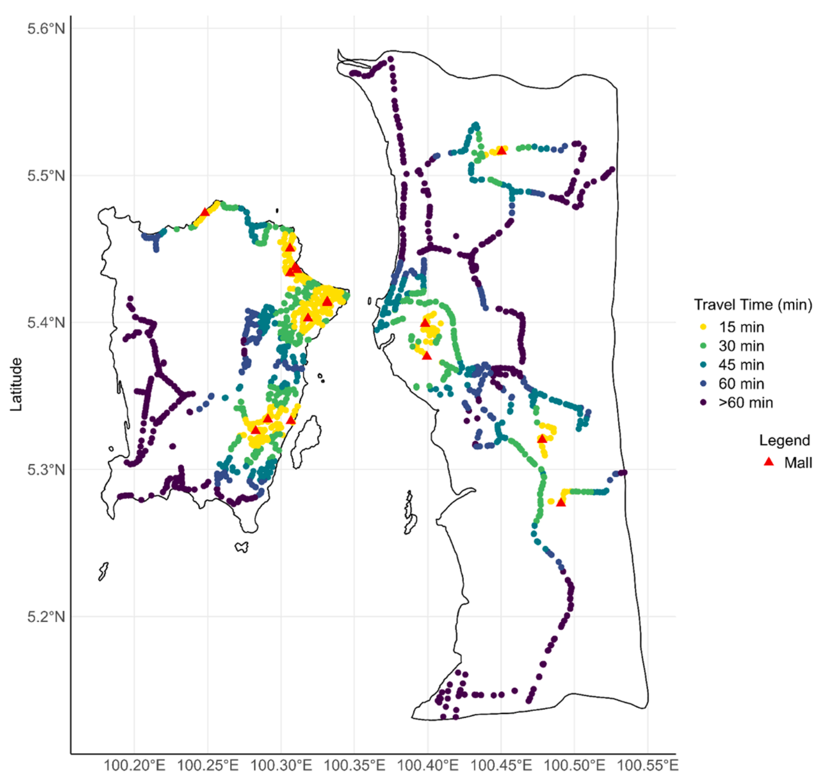
Additionally, the current node-weighted configurations optimized largely for ridership concentration and transfer efficiency do not adequately incorporate equity-based parameters such as population dispersion, mobility need, or access to critical services. This creates a structural misalignment between service availability and demographic geography, where transit infrastructure favors economic intensity over social inclusion [115–117]. The travel time surface models further validate this finding by showing that access to hospitals, schools, and commercial zones deteriorates sharply beyond the inner transit rings, revealing systemic spatial inequities masked by apparent network efficiency.

To strengthen the adaptive capacity of these networks, planning must move toward a paradigm of resilient decentralization. This involves redistributing transfer functionality across a broader set of secondary hubs, enhancing cross-town route connections, and investing in micro-transfer points in growth areas. Moreover, introducing modular redundancy whereby parallel or intersecting services offer backup functionality without full route duplication could significantly improve service continuity in peripheral districts.

In sum, while myBAS Johor and Rapid Bus Penang demonstrate high structural efficiency in facilitating intra-urban mobility within their respective cores, their reliance on centralized transfer architectures exposes both systems to strategic vulnerability. A shift toward distributed

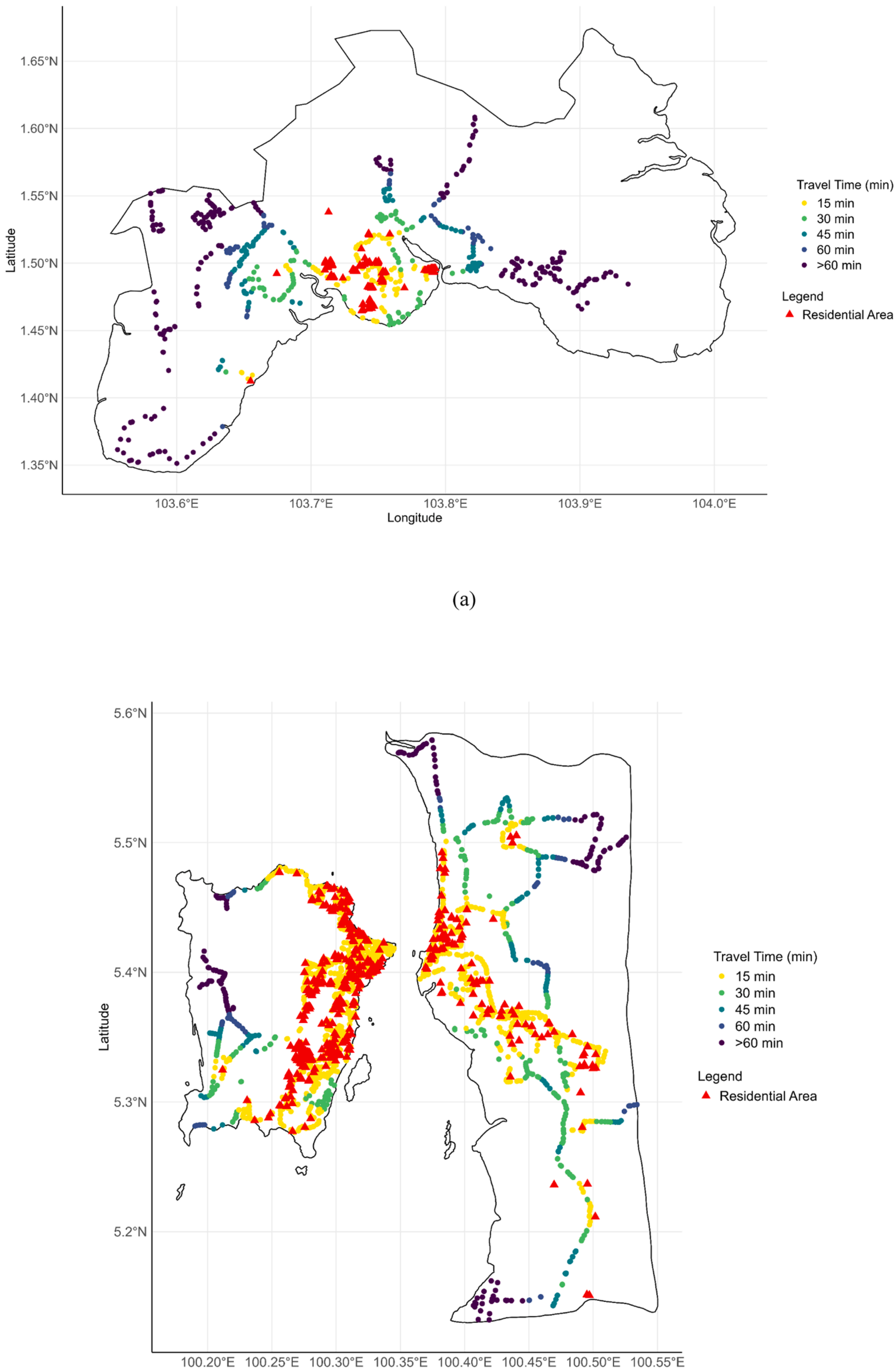


(a)



(b)

Figure 16. Travel time from mall to the nearest bus stops (a) myBAS Johor and (b) Rapid Bus Penang.



(a)

(b)

Figure 17. Pedestrian Travel time from residential area to bus stops in (a) myBAS Johor and (b) Rapid Bus Penang.

resilience, spatial redundancy, and modular connectivity will be essential to sustain performance under increasing demand and shifting urban morphology. This evolution requires not only technical reconfiguration but also institutional support for inclusive planning and infrastructure investment that anticipates rather than reacts to urban expansion.

5.2. Route Redundancy and Service Inefficiencies

The transit network analysis conducted in this study exposes significant complexities and inefficiencies in both the structural and temporal dimensions of service delivery across myBAS Johor and Rapid Bus Penang. While both systems perform well in terms of peak-hour coverage and central corridor saturation, their operational configurations often fall short when evaluated through the lens of systemic balance, fleet optimization, and inclusive accessibility.

Temporal analysis reveals that the two networks offer adequate service frequency during standard commuting hours, particularly between 07:00 and 19:00, aligning with traditional employment and educational schedules. However, temporal gaps become apparent during late-night and early-morning periods (00:00–05:00), where service frequency significantly diminishes or ceases altogether. These temporal blind spots disproportionately impact transit-dependent populations, including shift workers, informal sector laborers, night-time economy participants, and low-income commuters who lack private alternatives. The absence of off-peak services thus reinforces socio-temporal exclusion, undermining the principle of equitable access to urban opportunities.

Structurally, route overlap patterns offer insights into how spatial and operational redundancy can lead to suboptimal system performance. In Johor Bahru, high-density corridors such as those served by routes T30, T31, and T40 are characterized by extensive overlap. While these overlaps may appear beneficial in increasing frequency and reducing wait times, their marginal utility declines sharply in the absence of demand-responsive coordination. Without intelligent scheduling, real-time fleet management, or stop-level load balancing, such overlaps can result in vehicle bunching, fleet congestion, and inefficient fuel and labor utilization. In practical terms, this translates to unnecessary operational expenditure with limited user benefit, especially when service intervals are not differentiated by time-of-day or demand intensity.

In contrast, low-overlap routes such as F300 and T42, which serve lower-density or peri-urban zones, face a different set of challenges. While these routes extend spatial coverage and contribute to geographic equity, they often operate in isolation, without integration into feeder loops or modal transfer systems. The lack of connectivity to trunk routes or regional hubs reduces their effective utility, limiting the possibility of seamless transfers and thereby diminishing their role in supporting complete end-to-end travel chains. This underlines the importance of network cohesion, not merely geographic coverage, in determining system accessibility and efficiency.

The situation in Penang reflects similar structural dualities. Core urban areas, especially around George Town, exhibit high route density and significant overlap, which again raises concerns regarding resource cannibalization and redundant routing. Meanwhile, peripheral districts such as Balik Pulau remain poorly connected, both in terms of route coverage and service frequency. The resultant pattern is one of spatial and operational asymmetry: while central districts may be over-served, peripheral areas remain marginal in planning priorities, exacerbating spatial disparities in mobility access.

These findings highlight an inherent conflict within Malaysian transit governance between market-oriented service rationales and the equity-driven mandate for inclusive, universally accessible mobility. Current practices tend to over-emphasize financial metrics such as farebox recovery and cost-per-kilometer, often at the expense of social accessibility objectives. This policy imbalance manifests in the persistence of service gaps in low-density regions and the inefficient duplication of services in high-density corridors [118].

To address these issues, future operational strategies must adopt a differentiated service model, one that distinguishes between high-volume urban corridors requiring dense, frequent services and low-density regions better served through flexible, demand-responsive modalities. Route planning must also evolve toward network layering, whereby primary trunk routes serve major corridors, secondary routes provide cross-town linkages, and feeder services integrate peripheral zones into the broader network. Additionally, the use of real-time analytics, AVL systems, and demand forecasting algorithms can help optimize route scheduling and reduce operational waste [119].

Overall, achieving network efficiency in a socially inclusive manner requires not just technical redesign but also institutional recalibration, where service equity is institutionalized as a core performance criterion alongside economic efficiency. Such a shift would represent a meaningful step toward the development of transit systems that are not only operationally sound but also structurally just and strategically adaptive in the face of ongoing urban transformation.

5.3. Spatial Inequities in Transit Access

The spatial analyses conducted in this study, particularly through kernel density estimations, stop-level utilization profiling, and isochrone-based accessibility modelling, reveal a marked geography of inequality embedded within Malaysia's urban bus systems. While transit services in central districts of Johor Bahru and George Town exhibit compact, integrated configurations characterized by dense stop placement and multimodal convergence this spatial coherence rapidly dissipates at the city peripheries. In areas such as southern Tebrau and western Penang Island, walking distances to the nearest stop regularly exceed nationally acceptable thresholds, and the time required to reach key destinations is disproportionately higher.

First-mile access remains the most critical barrier in these peripheral zones. The average pedestrian travel time to the nearest bus stop in urban districts is generally below 10 minutes, reflecting well-structured local connectivity. However, in suburban and exurban neighborhoods, walking distances may reach or exceed 1.5–2.0 kilometers, translating into travel times of 25–35 minutes, particularly for children, elderly individuals, or those with limited mobility. This spatial fragmentation is compounded by low route frequency, limited transfer options, and in many cases, the absence of shelter or wayfinding infrastructure. Overall, these factors significantly reduce the practical usability of public transport for peripheral residents, even when nominal service exists.

The results from the accumulated cost surface models provide further clarity on this urban-suburban divide. Isochrone maps demonstrate that access to essential public services, including hospitals, educational institutions, and retail centers declines steeply as one moves beyond the inner urban ring. In centrally located districts, these facilities are generally accessible within 15 to 30 minutes via a single trip or a short transfer. Conversely, in peripheral regions, the required journey often exceeds 45 minutes, particularly during non-peak periods or in areas with low stop connectivity and route overlap. This disparity is not simply a matter of convenience but represents a deeper spatial-structural inequity, wherein one's physical location effectively constrains access to opportunity and social mobility.

In the Malaysian context, this issue is exacerbated by historically uneven development patterns, land-use fragmentation, and the prioritization of core-centric planning models. Urban infrastructure investment has largely followed economic density, leaving peripheral districts reliant on ad hoc expansions or low-frequency services that do not meet the demands of daily life [73,120].

Addressing these inequities cannot be achieved through service expansion alone. Simply extending existing routes into peripheral areas risks perpetuating low-ridership inefficiencies without tackling the structural misalignment between network design and settlement morphology. What is required is a strategic reconceptualization of transit equity as a fundamental principle in transport planning—one that foregrounds accessibility, affordability, and usability for all populations, regardless of geographic location [115,121].

This reconceptualization must include the development of equity-sensitive planning metrics, such as travel time to essential services for vulnerable groups, integration of demographic risk indices into service allocation models, and the identification of first-mile barriers using pedestrian impedance models. Furthermore, funding frameworks should move beyond farebox recovery and vehicle utilization, instead incorporating indicators of spatial justice, health equity, and educational access. Integrating such an approach within Malaysian transit governance would also align with the objectives of Sustainable Development Goal 11.2, which calls for "safe, affordable, accessible, and sustainable transport systems for all, with special attention to the needs of those in vulnerable situations." The empirical evidence presented here demonstrates that peripheral exclusion is not an incidental byproduct of urbanization but a predictable consequence of current planning paradigms—one that can and must be addressed through targeted, data-driven reform.

5.4. Policy and Operational Implications

The empirical insights derived from this research carry significant implications for urban transit policy, operational strategy, and equitable network design. As Malaysian cities continue to evolve under the pressures of demographic shifts, spatial decentralization, and infrastructure aging, there is a growing imperative to translate analytical findings into concrete, policy-relevant interventions that are both technically sound and socially inclusive [77].

The study's centrality analyses reveal the acute dependency of network performance on a limited number of high-throughput nodes such as Larkin Sentral in Johor and Terminal Komtar in Penang. These hubs, while facilitating efficient transfers, also represent potential bottlenecks and single points of failure. Enhancing infrastructure capacity at these locations through investments in platform expansion, real-time signage, crowd control mechanisms, and express route overlays is essential for maintaining throughput resilience, particularly during peak demand periods.

In low-density and peripheral areas, the study highlights inefficiencies in traditional fixed-route service models. Routes such as F300 and T42 exhibit low connectivity and sparse utilization, rendering them economically and operationally suboptimal. A shift toward demand-responsive transit (DRT) models in these areas can provide a scalable alternative. By leveraging mobile booking platforms and artificial intelligence for dynamic scheduling and routing, transit agencies can extend coverage without incurring the inefficiencies associated with underutilized fixed services [122,123].

Furthermore, the route overlap analysis identified cases of excessive redundancy in central corridors, notably along routes such as T30, T31, and T40. These findings support the case for network rationalization through hierarchical route structuring, where primary corridors are served by high-capacity routes and supported by secondary and tertiary feeder services. This model can enhance operational efficiency, reduce fleet congestion, and provide clearer service logic to end users.

The accessibility modeling component particularly the isochrone-based travel time analysis to key destinations, demonstrates the urgency of strengthening first-mile/last-mile connectivity. Peripheral zones often face walking distances and travel times that exceed acceptable thresholds, creating significant barriers to transit adoption. Investment in multimodal integration such as bicycle infrastructure, pedestrian pathways, and park-and-ride facilities is necessary to ensure full network usability, especially for marginalized or transit-dependent populations [124].

Finally, the POI-specific accessibility audits emphasize a persistent equity gap. Access to essential urban amenities including hospitals, educational institutions, and retail centers is heavily skewed in favor of central districts. As a result, transport infrastructure planning must move beyond efficiency metrics to incorporate equity-based criteria. Allocating resources based on spatial disadvantages, demographic vulnerability, and social need can ensure that transit systems fulfill not only their logistical function but also their role in fostering inclusive urban development.

To consolidate these linkages between empirical evidence and applied reforms, Table 5 presents a synthesized mapping the core findings of this study to specific policy and design interventions. This table serves as a practical tool for planners, policymakers, and researchers aiming to translate data-driven insights into sustainable transit strategies.

By embedding these principles into transit policy and design frameworks, Malaysian urban transport systems can advance toward a model that is not merely functional, but transformative one capable of enhancing accessibility, reducing systemic disparities, and supporting inclusive, sustainable urban development.

Table 5
Policy and Design Implications Informed by Empirical Evidence.

Policy/Design Focus	Empirical Basis from Study	Actionable Outcome
Enhancing Infrastructure at High-Demand Transit Hubs	Centrality analysis identified Larkin Sentral and Terminal Komtar as critical but vulnerable nodes	Allocate capital investment for expanded platforms, digital signage, crowd control, and dedicated peak-hour express services
Deploying Demand-Responsive Transit in Low-Density Areas	Peripheral routes (e.g., F300, T42) exhibited low frequency and limited overlap	Pilot mobile app-based DRT in underserved zones, using AI to match real-time demand and optimize routing
Implementing Route Hierarchies and Network Rationalization	Overlapping routes (e.g., T30, T31, T40) in dense corridors led to redundancy and resource inefficiency	Redesign network with primary, secondary, and feeder routes to balance efficiency with coverage and reduce corridor saturation
Improving First-Mile/ Last-Mile Accessibility	Accessibility models showed significant travel-time disparities between central and peripheral regions	Develop pedestrian zones, bike infrastructure, and multimodal hubs to close first-mile/last-mile gaps
Equity-Centered Infrastructure Planning	POI-based accessibility findings highlighted inequitable access to hospitals, schools, and commercial centers	Prioritize transport investments in underserved regions as part of social inclusion mandates, not just financial viability

5.5. Future Research Agenda for Sustainable Transit

Ensuring the future of public transportation in Malaysia aligns with the national vision of sustainability requires a deliberate, research-informed approach to system innovation. The current findings reveal substantial progress in network structuring, centrality optimization, and accessibility modeling. However, they also expose several systemic gaps particularly in temporal reliability, spatial equity, and adaptive infrastructure that demand a targeted research agenda capable of advancing both scholarly understanding and policy impact.

Given Malaysia's rapid urbanization, shifting travel behavior, and increasing exposure to climate-related risks, future transit research must focus on building resilience, promoting inclusivity, and enabling system adaptability. Such research should not only enhance the technical sophistication of transport modeling but also support Malaysia's commitment to the United Nations Sustainable Development Goals (SDGs), particularly those related to sustainable cities (SDG 11) [125], climate action (SDG 13), infrastructure innovation (SDG 9), and reduced inequalities (SDG 10) [120].

To guide this effort, Table 6 presents a structured research agenda

Table 6
Future Research Directions for Sustainable Transit in Malaysia with SDG Alignment.

Research Area	Research Gap	Contribution to Sustainable Transport & SDG Alignment
Resilience Modelling of Central Transit Hubs	Current network assessments are static; lack of disruption modeling for centrality-critical nodes limits operational resilience planning under scenarios such as system failures, natural disasters, or demand surges.	Enhances network robustness (SDG 11.2) by identifying vulnerability points and improving preparedness for operational shocks, ensuring service continuity and urban system stability.
Optimization of Demand-Responsive Transit (DRT) Systems	Limited implementation of real-time optimization using AI/ML in Malaysia restricts the scalability of DRT systems, particularly for bridging service gaps in low-density, underserved communities.	Promotes equitable, cost-effective access to transport in rural/ peripheral zones (SDG 9.1, SDG 11.2), expanding mobility opportunities through scalable technology solutions.
Granular Accessibility Audits to Critical POIs	Most accessibility studies apply general metrics; there is a lack of granular, POI-specific audits incorporating socio-demographic indicators, travel time equity, and walkability to essential services.	Strengthens inclusion by targeting underserved areas with tailored interventions (SDG 10.2, SDG 11.3), improving access to health, education, and employment nodes.
Evaluation of Transit-Oriented Development (TOD) in Secondary Cities	Malaysian TOD literature primarily focuses on Kuala Lumpur; systematic evaluations of land-use-transport integration in smaller or emerging cities are underrepresented and necessary for spatial equity.	Supports compact, low-carbon urban development (SDG 11.3, SDG 13.2) by aligning housing, transit, and services in secondary growth corridors and mitigating urban sprawl.
Integrated Simulation of Multimodal Urban Mobility Systems	Simulation of multimodal interactions between buses, rail systems, and non-motorized transport (NMT) is scarce; such models are essential for evaluating integrated fare systems, routing efficiency, and emissions reduction.	Enables energy-efficient, multimodal integration (SDG 7.3, SDG 11.2) through data-driven coordination across transport modes, reducing emissions and promoting system interoperability.
Passenger Behavior Analysis	Behavioral impacts of GTFS-RT data on travel decisions are not well understood; there is a need for empirical studies using mobile apps or A/B testing to assess service reliability and responsiveness.	Improves real-time adaptability of public transport (SDG 9.5, SDG 11.2) by enhancing user information systems and service flexibility, leading to greater rider confidence and efficiency.
Designing Climate-Resilient Public Transit Infrastructure	Transit infrastructure in flood-prone or heat-vulnerable zones lacks adaptive design; climate-informed planning tools are required to mitigate risks and ensure long-term system operability and safety.	Ensures long-term climate resilience (SDG 13.1, SDG 11.) by designing transit infrastructure that can withstand environmental hazards and protect vulnerable urban populations.
Real-Time Transit Modelling using GTFS-RT	Existing research in Malaysia on transit models lack real-time operational data integration, limiting responsiveness to dynamic demand, delays, or service disruptions.	Supports adaptive transit planning (SDG 11.2) by enabling data-driven decisions based on live service conditions, enhancing reliability, efficiency, and user satisfaction.

identifying key domains, the specific methodological and policy gaps they address, and their alignment with SDG targets. This synthesis provides a roadmap for scholars, planners, and decision-makers seeking to leverage empirical evidence for transformative improvements in urban mobility systems.

This proposed research framework not only addresses Malaysia's most pressing transit challenges but also contributes to global discourses on sustainable urban development. Each direction offers both technical feasibility and policy relevance, providing a strong basis for securing institutional support and interdisciplinary collaboration [126–129]. The integration of GTFS-RT presents a promising direction for future research aimed at advancing real-time transit modeling and enhancing operational responsiveness. Leveraging dynamic service data can support the development of adaptive planning frameworks and data-informed interventions responsive to fluctuating demand and service disruption ns [130,131]. Such an approach aligns with the objectives of SDG 11.2, offering pathways to improve the reliability, efficiency, and equity of public transport systems, particularly in dynamic urban environments [132–135]. Furthermore, by embedding SDG alignment within the formulation of research questions, scholars can ensure that their work supports both academic advancement and international development benchmarks.

In sum, a research-informed transit agenda one that is technically robust, equity-centered, and globally aligned represents a vital next step in the evolution of Malaysia's urban mobility systems. It offers a pathway toward systems that are not only efficient and scalable, but also socially inclusive, environmentally resilient, and grounded in evidence-based governance.

5.6. *Research Contributions and National Relevance*

Although GTFS-based methodologies are widely utilized in global transit research, their application within the Malaysian context remains exceptionally nascent. The official release of Malaysia's first GTFS datasets in November 2023 represents a critical inflection point in the country's transition toward data-driven urban mobility planning. Prior to this, transit system evaluations predominantly relied on static route maps, manual field surveys, and consultant-led reports—methods constrained by limited computational capacity and a lack of integrated geospatial analytics.

Existing Malaysian transit literature has largely focused on user satisfaction, travel behavior, and service perception, with minimal engagement in network structure, spatial accessibility modeling, or performance diagnostics. This study directly addresses that methodological void by presenting one of the first empirical applications of GTFS data for a multi-dimensional, systems-level analysis of public bus operations in Malaysia.

The contribution of this study is twofold. First, it introduces a comprehensive analytical orientation centered on transit network structure and spatial accessibility—dimensions that are notably absent in prevailing Malaysian research. Second, it operationalizes a suite of established international methodologies—including network science, accessibility modeling, and route efficiency analysis—within a cohesive, reproducible, and open-source workflow. This approach not only enhances transparency and methodological rigor but also provides practical tools for planners, government agencies, and researchers engaged in transit modernization efforts.

Beyond its methodological advancements, the study delivers policy- and design-relevant insights grounded in empirical evidence. These insights are translated into actionable interventions for network rationalization, demand-responsive transit deployment, infrastructure investment, and equity-centered planning. Furthermore, the research establishes a scalable foundation for future inquiry into sustainable urban mobility, aligning closely with Malaysia's commitments to the SDGs, particularly SDG 11.2 on inclusive and sustainable transport systems.

To contextualize these contributions, Table 7 summarizes the study's thematic areas of impact and their relevance to the evolving institutional and operational landscape of Malaysian transit planning.

6. Conclusion and Limitations

This study presents a comprehensive, data-driven evaluation of Malaysia's public bus transit systems, focusing on the operational and spatial dynamics of myBAS Johor and Rapid Bus Penang. Through the integration of network centrality analysis, route overlap diagnostics, spatial accessibility modeling, and isochrone-based evaluations to critical destinations, this research contributes new empirical evidence and integrated methodological approach for understanding the strengths and weaknesses of contemporary urban transit networks in a rapidly urbanizing Southeast Asian context.

Findings from the network analysis highlight the presence of core-centric connectivity patterns, with high betweenness centrality values concentrated at nodes such as Larkin Sentral and Megamall Pacific, which function as dominant interchange points within their respective systems. While this enhances throughput efficiency in high-demand corridors, it also exposes the system to disproportionate vulnerability in the event of service interruptions at these key hubs. The observed semi-hierarchical topology offers flexibility in urban cores but lacks redundancy and adaptability in suburban and peri-urban segments, where network sparsity and low service frequency limit equitable access.

Temporal assessments reveal that both systems maintain strong service availability during peak commuting hours but exhibit significant gaps during early-morning and late-night periods, particularly affecting shift workers and populations dependent on public transport outside the standard operating window. Spatial distribution analyses confirm that urban districts benefit from high stop densities and short average walking distances, whereas peripheral regions such as southern Tebrau and Balik Pulau face first-mile barriers, extended travel times, and diminished access to essential services like hospitals, schools, and commercial centers.

Furthermore, route overlap analyses highlight inefficiencies in resource deployment. In high-density corridors, excessive overlaps such as that seen among routes T30, T31, and T40 results in marginal frequency gains while increasing operational costs and underutilizing fleet potential. Conversely, low-overlap routes like F300 and T42 extend spatial coverage to underserved areas but often lack feeder integration or adequate frequency, limiting their usability for multimodal or transfer-based trips.

Table 7
Research Contributions to Malaysian Transit Planning and Policy.

Contribution Area	Relevance to Malaysian Context
Operationalizing GTFS for Malaysian Transit Planning	GTFS data became available nationally only in November 2023; this study is among the first to apply it systematically for service evaluation and spatial equity analysis.
Introducing Computational Methods in Local Contexts	Malaysian transit research has largely lacked data-driven approaches. This study integrates network analysis and geospatial modeling to fill this critical methodological gap.
Informing Evidence-Based Public Transport Policy	Technical results are translated into actionable recommendations for service redesign, route optimization, and infrastructure investment, fostering planning reform.
Establishing an Integrated Methodological Approach	The study develops replicable, open-source workflows using tools such as Python, R, GTFS Validator, and QGIS, supporting institutional uptake and scalability across cities.
Supporting Malaysia's SDG 11.2 Commitments	By diagnosing spatial inequities and proposing targeted interventions, the research contributes to national objectives on sustainable and inclusive urban transport.

This research contributes to literature by advancing a multi-dimensional assessment approach that captures structural complexity, temporal responsiveness, and equity in transit access. By combining network science, spatial analytics, and accessibility modeling, the study provides actionable insights for transit planners and policy stakeholders aiming to transition toward resilient, inclusive, and performance-optimized mobility systems.

While this study provides robust empirical insights, several limitations should be acknowledged. First, the analysis is based exclusively on GTFS-static datasets, which capture scheduled transit services but do not reflect real-time operational dynamics such as delays, missed trips, or service rerouting due to congestion or infrastructure failures. This temporal limitation restricts the ability to assess service reliability or real-world passenger experience.

To mitigate data reliability concerns, all GTFS datasets used in this study were systematically validated using the Canonical GTFS Schedule Validator. The validation process confirmed structural completeness and compliance with specification standards, with only non-critical warnings such as minor formatting inconsistencies and missing optional fields detected. These were addressed through post-validation refinement using custom Python scripts, which corrected time formats, filled missing entries, and standardized spatial coordinates. While these measures enhance the internal consistency of the data, the absence of real-time service information remains a constraint in fully capturing dynamic transit performance.

Second, the temporal granularity of the available data limits the capacity to assess seasonal, event-driven, or day-to-day demand fluctuations. A more dynamic and longitudinal data architecture incorporating GTFS-RT, AVL systems, or smartcard transaction data—would enable a deeper exploration of temporal elasticity and operational variability. Third, the study is limited by the absence of socioeconomic and demographic integration. While spatial and structural metrics offer important insights, they do not reflect how accessibility and service equity vary across income groups, gender, age cohorts, or other vulnerable populations. Future research should incorporate household travel surveys, census data, or demographic overlays to support a more equity-centered analysis of transit performance. Fourth, this study focuses exclusively on bus-based public transportation, without accounting for intermodal interactions with commuter rail, informal transit services (e.g., vans or minibuses), or emerging micro-mobility options such as bike-sharing and e-scooters. A system-level simulation of multimodal urban mobility including mode interconnectivity and fare integration would offer a more holistic understanding of travel behavior and network efficiency.

Finally, the study does not include cost-benefit or economic feasibility analysis of the proposed interventions. While the recommendations are grounded in spatial and operational analysis, the absence of financial modeling may limit their immediate policy uptake. Future work could incorporate cost-effectiveness assessment tools or multi-criteria decision analysis (MCDA) to evaluate trade-offs and implementation feasibility under budgetary constraints.

Despite these limitations, the present study offers a replicable and policy-relevant method for assessing and improving urban transit systems in Malaysia and other emerging metropolitan regions. It affirms the critical importance of aligning network design with social inclusion, operational resilience, and data-informed planning. Moving forward, transit reform efforts must prioritize not only throughput and coverage but also systemic adaptability and spatial equity, ensuring that public transportation serves as an enabling infrastructure for inclusive urban development. The insights generated here provide a valuable foundation for further research, institutional reform, and long-term policy innovation aimed at achieving Malaysia's sustainable mobility and climate-resilience goals.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author(s) utilized GPT-4 to assist with grammar correction, language refinement, and overall clarity of expression. Following the use of this AI tool, the author(s) thoroughly reviewed and edited the content to ensure accuracy, integrity, and alignment with the intended scholarly rigor. The author(s) take full responsibility for the final content of the published article.

CRediT authorship contribution statement

Mahbub Hassan: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hridoy Deb Mahin:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Ferdoushi Ahmed:** Writing – review & editing, Validation. **Md. Maruf Hassan:** Writing – review & editing, Validation. **Atikur Rahaman:** Writing – review & editing, Validation, Investigation, Funding acquisition. **Masuk Abdullah:** Writing – review & editing, Validation, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research did not receive any external funding.

Acknowledgments

This research was supported by the University of Debrecen Program for Scientific Publication. The authors would like to express their appreciation to the data.gov.my developer team for providing access to their open data sources and APIs, which facilitated the data collection and analysis in this study.

Data availability

The datasets generated and analyzed during this study are publicly available at data.gov.my under the Creative Commons Attribution 4.0 International License (CC BY 4.0). Additionally, the code and cleaned, validated datasets will be provided by the corresponding author upon reasonable request.

References

[1] M. Buchanan, The benefits of public transport, Nature Physics 15 (9) (Sep. 2019), https://doi.org/10.1038/s41567-019-0656-8.

[2] J.A.K. Alhassan, P. Anciaes, Public transport investments as generators of economic and social activity, Journal of Transport & Health 41 (Mar. 2025) 101989, https://doi.org/10.1016/j.jth.2025.101989.

[3] A.K. Yudha, N. Leelawat, J. Tang, A systematic review and bibliometric analysis of the impacts of COVID-19 on economy and mobility from the geospatial data perspective, Results in Engineering 26 (Jun. 2025) 105282, https://doi.org/10.1016/j.rineng.2025.105282.

[4] S. Kaewunruen, J. Sresakoolchai, H. Sun, Causal analysis of bus travel time reliability in Birmingham, UK, Results in Engineering 12 (Dec. 2021) 100280, https://doi.org/10.1016/j.rineng.2021.100280.

[5] R. Xiao, et al., Urban expansion in highly populous East Asian megacities during 1990–2020: Tokyo, Seoul, Beijing, and Shanghai, Ecological Informatics 83 (Nov. 2024) 102843, https://doi.org/10.1016/j.ecoinf.2024.102843.

[6] "The Growth of Asian Cities.". [Online]. Available: https://knowledge.csc.gov.sg/world-cities-summit-issue/the-growth-of-asian-cities/. Accessed: August, 2024.

[7] M. Seehamart, S. Pueboobpaphan, R. Pueboobpaphan, Actual walking distance vs. Perceived walking distance: Relationship and impact on determining acceptable walking distances to stations of the Bangkok Metro, Results in Engineering 25 (Mar. 2025) 104088, https://doi.org/10.1016/j.rineng.2025.104088.

[8] P. Miner, B.M. Smith, A. Jani, G. McNeill, A. Gathorne-Hardy, Car harm: A global review of automobility's harm to people and the environment, Journal of Transport Geography 115 (Feb. 2024) 103817, https://doi.org/10.1016/j.jtrangeo.2024.103817.

[9] M.I.M. Masirin, A.M. Salin, A. Zainorabidin, D. Martin, N. Samsuddin, Review on Malaysian Rail Transit Operation and Management System: Issues and Solution in Integration, in: IOP Conference Series: Materials Science and Engineering 226, Aug. 2017 012029, https://doi.org/10.1088/1757-899x/226/1/012029.

[10] A. Al Nafees, M. Hassan, A. Paul, S.S. Shraban, H.Deb Mahin, Public Transport Ridership Forecasting Using Machine Learning: A Case Study of Rapid Bus KL in Malaysia, in: 2024 27th International Conference on Computer and Information Technology (ICCIT), Dec. 2024, pp. 3366–3371, https://doi.org/10.1109/iccit64611.2024.11021879.

[11] Malaysia Government Open Data, "Public Transportation Dashboard,". [Online]. Available: https://data.gov.my/dashboard/public-transportation. Accessed: August, 2024.

[12] S.N. Wahab, M.I. Hamzah, N.M. Suki, Y.S. Chong, C.P. Kua, Unveiling passenger satisfaction in rail transit through a consumption values perspective, Multimodal Transportation 4 (1) (Mar. 2025) 100196, https://doi.org/10.1016/j.multra.2025.100196.

[13] D. Arias, et al., Using GTFS to Calculate Travel Time Savings Potential of Bus Preferential Treatments, Transportation Research Record: Journal of the Transportation Research Board 2675 (9) (Jun. 2021) 1643–1654, https://doi.org/10.1177/03611981211009226.

[14] Z. Wan, Y. Mahajan, B.W. Kang, T.J. Moore, J.-H. Cho, A Survey on centrality metrics and their Network Resilience Analysis, IEEE Access 9 (2021) 104773–104819, https://doi.org/10.1109/access.2021.3094196.

[15] J. Delso, B. Martín, E. Ortega, A new procedure using network analysis and kernel density estimations to evaluate the effect of urban configurations on pedestrian mobility. The case study of Vitoria –Gasteiz, Journal of Transport Geography 67 (Feb. 2018) 61–72, https://doi.org/10.1016/j.jtrangeo.2018.02.001.

[16] V. Verma, R.K. Aggarwal, A comparative analysis of similarity measures akin to the Jaccard index in collaborative recommendations: empirical and theoretical perspective, Social Network Analysis and Mining 10 (1) (Jun. 2020), https://doi.org/10.1007/s13278-020-00660-9.

[17] I. Herzog, Spatial Analysis Based On Cost Functions, Archaeological Spatial Analysis (Jan. 2020) 333–358, https://doi.org/10.4324/9781351243858-18.

[18] J. Wu, et al., A GTFS data acquisition and processing framework and its application to train delay prediction, International Journal of Transportation Science and Technology 12 (1) (Mar. 2023) 201–216, https://doi.org/10.1016/j.ijtst.2022.01.005.

[19] P. Prommaharaj, S. Phithakkitnukoon, M.G. Demissie, L. Kattan, C. Ratti, Visualizing public transit system operation with GTFS data: A case study of Calgary, Canada, Heliyon 6 (4) (2020) e03729, https://doi.org/10.1016/j.heliyon.2020.e03729. Apr.

[20] S. Para, et al., G2Viz: an online tool for visualizing and analyzing a public transit system from GTFS data, Public Transport 16 (3) (May 2024) 893–928, https://doi.org/10.1007/s12469-024-00362-x.

[21] A. Antrim, S.J. Barbeau, The many uses of GTFS data—opening the door to transit and multimodal applications, Location-Aware Information Systems Laboratory, University of South Florida, 2013 [Online]. Available: http://www.locationaware.usf.edu/wp-content/uploads/2010/02/The-Many-Uses-of-GTFS-Data-%E2%80%93-ITS-America-submission-abbreviated.pdf.

[22] K. Kaeoruean, S. Phithakkitnukoon, M.G. Demissie, L. Kattan, C. Ratti, Analysis of demand–supply gaps in public transit systems based on census and GTFS data: a case study of Calgary, Canada, Public Transport 12 (3) (Sep. 2020) 483–516, https://doi.org/10.1007/s12469-020-00252-y.

[23] P. Fortin, C. Morency, M. Trépanier, Innovative GTFS Data Application for Transit Network Analysis Using a Graph-Oriented Method, Journal of Public Transportation 19 (4) (Dec. 2016) 18–37, https://doi.org/10.5038/2375-0901.19.4.2.

[24] B. Goldstein, L. Dyson, A. Nemani, Beyond Transparency: Open Data and the Future of Civic Innovation, Code for America Press, San Francisco, CA, USA, 2013 [Online]. Available: https://books.google.com.bd/books?id=uzbxngEACAAJ.

[25] Z. Aemmer, A. Ranjbari, D. MacKenzie, Measurement and classification of transit delays using GTFS-RT data, Public Transport 14 (2) (Feb. 2022) 263–285, https://doi.org/10.1007/s12469-022-00291-7.

[26] P. Bellini, S. Bilotta, E. Collini, M. Fanfani, P. Nesi, Data Sources and Models for Integrated Mobility and Transport Solutions, Sensors 24 (2) (Jan. 2024) 441, https://doi.org/10.3390/s24020441.

[27] "Best Practices - General Transit Feed Specification." Accessed: May 25, 2025. [Online]. Available: https://gtfs.org/documentation/schedule/schedule-best-practices/.

[28] S.J. Barbeau, Quality control—lessons learned from the deployment and evaluation of GTFS-realtime feeds, in: presented at the *97th Annual Meeting of the Transportation Research Board (TRB)*, Washington, DC, USA, 2024. Jan. 7–11, 2018. Sponsored by TRB Committee AP000 Public Transportation Group. [Online]. Available: https://trid.trb.org/View/1496848. Accessed: Aug.

[29] T. Elliott, T. Lumley, Modelling the travel time of transit vehicles in real-time through a GTFS-based road network using GPS vehicle locations, Australian & New Zealand Journal of Statistics 62 (2) (Jun. 2020) 153–167, https://doi.org/10.1111/anzs.12294.

[30] N. Wessel, J. Allen, S. Farber, Constructing a routable retrospective transit timetable from a real-time vehicle location feed and GTFS, Journal of Transport

Geography 62 (Jun. 2017) 92–97, https://doi.org/10.1016/j.jtrangeo.2017.04.012.
[31] E. Chondrodima, H. Georgiou, N. Pelekis, Y. Theodoridis, Particle swarm optimization and RBF neural networks for public transport arrival time prediction using GTFS data, International Journal of Information Management Data Insights 2 (2) (Nov. 2022) 100086, https://doi.org/10.1016/j.jjimei.2022.100086.
[32] A. Aziz, J. Mohamad, Public Transport Planning: Local Bus Service Integration and Improvements In Penang, Malaysia, Planning Malaysia 18 (13) (Jul. 2020), https://doi.org/10.21837/pm.v18i13.784.
[33] M.N. Ustadi, N.A.M. Shopi, A Study towards the Efficiency of Public Transportation Hub Characteristics: A Case Study of Northern Region, Peninsular Malaysia, Procedia Economics and Finance 35 (2016) 612–621, https://doi.org/10.1016/S2212-5671(16)00075-7.
[34] J.C. Wong, Use of the General Transit Feed Specification (GTFS) in Transit Performance Measurement, Sch. of Civil and Environmental Eng. and Sch. of City and Regional Planning, Georgia Inst. of Technol, Atlanta, GA, USA, Dec. 2013. M. S. thesis[Online]. Available: https://nctspm.gatech.edu/sites/default/files/u60/WONG-THESIS-2013.pdf.
[35] P. Carleton, S. Hoover, B. Fields, M. Barnes, J.D. Porter, GTFS-Ride: Unifying Standard for Fixed-Route Ridership Data, Transportation Research Record: Journal of the Transportation Research Board 2673 (12) (Jul. 2019) 173–181, https://doi.org/10.1177/0361198119851748.
[36] E. Eros, S. Mehndiratta, C. Zegras, K. Webb, M.C. Ochoa, Applying the General Transit Feed Specification to the Global South, Transportation Research Record: Journal of the Transportation Research Board 2442 (1) (Jan. 2014) 44–52, https://doi.org/10.3141/2442-06.
[37] S. Jiang, Y. Sun, W. Wong, Y. Xu, X. Zhao, Real-Time Urban Traffic Monitoring Using Transit Buses as Probes, Transportation Research Record: Journal of the Transportation Research Board 2679 (2) (Jul. 2024) 219–236, https://doi.org/10.1177/03611981241260708.
[38] S. K. B. V, R. Fedujwar, A. Agarwal, Travel Time Variability of Bus Routes in Delhi Using Real-Time GTFS Data, in: 2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS), Jan. 2024, pp. 210–215, https://doi.org/10.1109/comsnets59351.2024.10427234.
[39] C. Bannur, C. Bhat, G. Goutham, H.R. Mamatha, General Transit Feed Specification Assisted Effective Traffic Congestion Prediction Using Decision Trees and Recurrent Neural Networks, in: 2022 IEEE 1st International Conference on Data, Decision and Systems (ICDDS), Dec. 2022, pp. 1–6, https://doi.org/10.1109/icdds56399.2022.10037408.
[40] J.Q. Zhou, J. Jackson, P. Stravitz, G.J. Barlow, P. Koonce, Addressing Data Latency in GTFS (General Transit Feed Specification) Realtime to Improve Transit Signal Priority, Transportation Research Record: Journal of the Transportation Research Board 2679 (1) (Jul. 2024) 1329–1341, https://doi.org/10.1177/03611981241255027.
[41] J.-W. Sim, G.-H. Cho, Dynamic public transit accessibility analysis using actual transit operation data, International Journal of Urban Sciences (May 2025) 1–23, https://doi.org/10.1080/12265934.2025.2503727.
[42] P. Gaudette, R. Chapleau, T. Spurr, Bus Network Microsimulation with General Transit Feed Specification and Tap-in-Only Smart Card Data, Transportation Research Record: Journal of the Transportation Research Board 2544 (1) (Jan. 2016) 71–80, https://doi.org/10.3141/2544-09.
[43] A. Vijaya, B.L. Gudissa, L.S. Angreani, H. Wicaksono, Predictive Analysis of Public Transportation Delays Using Machine Learning Models on GTFS Data, in: 2024 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Dec. 2024, pp. 450–454, https://doi.org/10.1109/ieem62345.2024.10857111.
[44] E. Chondrodima, H. Georgiou, N. Pelekis, Y. Theodoridis, Public Transport Arrival Time Prediction Based on GTFS Data, Machine Learning, Optimization, and Data Science (2022) 481–495, https://doi.org/10.1007/978-3-030-95470-3_36.
[45] Z. Aemmer, S. Sørbø, A. Clemente, M. Ruocco, Generalization strategies for improving bus travel time prediction across networks, Journal of Urban Management 13 (3) (Sep. 2024) 372–385, https://doi.org/10.1016/j.jum.2024.05.002.
[46] P.P. Lopes, G. Gramaglia, D. Bacciu, H.T. Marques-Neto, Towards Forecasting Bus Arrival Thorough A Model Based On GNN+LSTM Using GTFS and Real-time Data, in: Proceedings of the 4th International Conference on AI-ML Systems, Oct. 2024, pp. 1–9, https://doi.org/10.1145/3703412.3703417.
[47] L. Iovino, P.T. Nguyen, A.D. Salle, F. Gallo, M. Flammini, Unavailable Transit Feed Specification: Making It Available With Recurrent Neural Networks, IEEE Transactions on Intelligent Transportation Systems 22 (4) (2021) 2111–2122, https://doi.org/10.1109/TITS.2021.3053373.
[48] R. Javanmard, L. Liu, J.A. Long, J. Lee, Using Realtime GTFS to generate easy-to-use transit accessibility measures under travel time uncertainty, Travel Behaviour and Society 41 (Oct. 2025) 101054, https://doi.org/10.1016/j.tbs.2025.101054.
[49] Y. Li, W."David" Fan, Modeling and Evaluating Public Transit Equity and Accessibility by Integrating General Transit Feed Specification Data: Case Study of the City of Charlotte, Journal of Transportation Engineering, Part A: Systems 146 (10) (Oct. 2020), https://doi.org/10.1061/jtepbs.0000426.
[50] J. Wu, L. Zhou, C. Cai, F. Dong, J. Shen, G. Sun, Towards a General Prediction System for the Primary Delay in Urban Railways, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019, pp. 3482–3487, https://doi.org/10.1109/ITSC.2019.8916868.
[51] R. Ceccato, F. Deflorio, M. Diana, M. Pirra, Measure of urban accessibility provided by transport services in Turin: a traveller perspective through a mobility survey, Transportation Research Procedia 45 (2020) 301–308, https://doi.org/10.1016/j.trpro.2020.03.020.
[52] A. Vágner, Route planning on GTFS using Neo4j, Annales Mathematicae et Informaticae (2021), https://doi.org/10.33039/ami.2021.07.001. Accepted manuscript.
[53] N. Wessel, M.J. Widener, Discovering the space–time dimensions of schedule padding and delay from GTFS and real-time transit data, Journal of Geographical Systems 19 (1) (Dec. 2016) 93–107, https://doi.org/10.1007/s10109-016-0244-8.
[54] E. Boudabous, M. Karaa, L. Sboui, J. Montecinos, O. Alam, Analyzing Public Transit Schedule Deviations: A Case Study on Montreal Using Real-Time Data, in: 2024 IEEE 27th International Symposium on Real-Time Distributed Computing (ISORC), May 2024, pp. 1–6, https://doi.org/10.1109/isorc61049.2024.10551354.
[55] J. Vuurstaek, G. Cich, L. Knapen, A.-U.-H. Yasar, T. Bellemans, D. Janssens, GTFS Bus Stop Mapping to the OSM Network, Procedia Computer Science 109 (2017) 50–58, https://doi.org/10.1016/j.procs.2017.05.294.
[56] J. Godfrid, P. Radnic, A. Vaisman, E. Zimányi, Analyzing public transport in the city of Buenos Aires with MobilityDB, Public Transport 14 (2) (Mar. 2022) 287–321, https://doi.org/10.1007/s12469-022-00290-8.
[57] S. Devunuri, L. Lehe, TransitGPT: a generative AI-based framework for interacting with GTFS data using large language models, Public Transport (Jun. 2025), https://doi.org/10.1007/s12469-025-00395-w.
[58] S. Devunuri, S. Qiam, L.J. Lehe, ChatGPT for GTFS: benchmarking LLMs on GTFS semantics... and retrieval, Public Transport 16 (2) (Apr. 2024) 333–357, https://doi.org/10.1007/s12469-024-00354-x.
[59] J. P. Bazzo Vieira, R. H. M. Pereira, and P. R. Andrade, "Estimating public transport emissions from General Transit Feed Specification data," Nov. 2022, doi: https://doi.org/10.31219/osf.io/8m2cy.
[60] A. Nishino, A. Kodaka, M. Nakajima, N. Kohtake, A Model for Calculating the Spatial Coverage of Audible Disaster Warnings Using GTFS Realtime Data, Sustainability 13 (23) (Dec. 2021) 13471, https://doi.org/10.3390/su132313471.
[61] S. Hoover, A. Kost, J.David Porter, Using General Transit Feed Specification-Ride to Identify Riders' Use of Transfer Opportunities, Transportation Research Record: Journal of the Transportation Research Board 2676 (6) (Feb. 2022) 197–206, https://doi.org/10.1177/03611981211073311.
[62] K. Dick, A. Hasan, J. Dergham, A.J. Clarke, H. Khalil, G. Wainer, Systematic Analysis of Public Transit Data Availability in Canada, in: 2022 IEEE International Conference on Big Data (Big Data), Dec. 2022, pp. 4901–4908, https://doi.org/10.1109/bigdata55660.2022.10021073.
[63] H. Wang, B.P.Y. Loo, The public transport disadvantaged in a highly transit-oriented city: An analytical framework, key challenges and opportunities, J Transp Geogr 120 (2024), https://doi.org/10.1016/j.jtrangeo.2024.103983.
[64] A. Karner, Assessing public transit service equity using route-level accessibility measures and public data, Journal of Transport Geography 67 (Feb. 2018) 24–32, https://doi.org/10.1016/j.jtrangeo.2018.01.005.
[65] S. Senegas, A. Furno, P. Palmas, J. Simon, Framework for the Analysis and Enhancement of the Accessibility of Large-Scale Urban Transit Networks: A Data-Driven Study for the City of Lyon, France, Transp Res Rec 2678 (2) (Feb. 2024) 35–47, https://doi.org/10.1177/03611981231172939.
[66] S.P. Borgatti, M.G. Everett, A Graph-theoretic perspective on centrality, Soc Networks 28 (4) (2006) 466–484, https://doi.org/10.1016/j.socnet.2005.11.005.
[67] N. Wessel and S. Farber, "On the Accuracy of Schedule-Based GTFS for Measuring Accessibility," 2018, doi: https://doi.org/10.31235/osf.io/hzgpd.
[68] R. Passmore, K. Watkins, R. Guensler, Assessing Bike-Transit Accessibility, Transp Res Rec (Oct. 2024), https://doi.org/10.1177/03611981241234902;WGROUP:STRING:PUBLICATION.
[69] R.H.M. Pereira, P.R. Andrade, J.P.B. Vieira, Exploring the time geography of public transport networks with the gtfs2gps package, Journal of Geographical Systems 25 (3) (Dec. 2022) 453–466, https://doi.org/10.1007/s10109-022-00400-x.
[70] J.Y. Yap, N. Omar, I. Ismail, A Study of Traffic Congestion Influenced by the Pattern of Land Use, IOP Conference Series: Earth and Environmental Science 1022 (1) (May 2022) 012035, https://doi.org/10.1088/1755-1315/1022/1/012035.
[71] S. Bachok, M.Mohamed Osman, U.A. Khalid, M. Ibrahim, Commuters' Perceptions on Rail Based Public Transport Services: A Case Study of Ktm Komuter In Kuala Lumpur City, Malaysia, Planning Malaysia Journal 11 (3) (Nov. 2013), https://doi.org/10.21837/pmjournal.v11.i3.110.
[72] J. Tomas, A.N.A. Rahman, The Effectiveness of Public Transport System, in: Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering, 2022, pp. 247–256, https://doi.org/10.1007/978-981-16-7597-3_20.
[73] I. Yusoff, B.-K. Ng, S.A. Azizan, Towards sustainable transport policy framework: A rail-based transit system in Klang Valley, Malaysia, PLOS ONE 16 (3) (Mar. 2021) e0248519, https://doi.org/10.1371/journal.pone.0248519.
[74] N.H. Ramlan, M.M. Osman, N.S. Rabe, A. Azizan, N.A. Azmi, S. Amiruddin, Assessing the Current Implementation of Compact And Mixed-Use Development within Public Rail Transit Stations in Malaysia, Planning Malaysia 19 (May 2021), https://doi.org/10.21837/pm.v19i15.934.
[75] S. Bachok, Z. Ponrahono, M.M. Osman, S. Jaafar, M. Ibrahim, M. Mohamed, Apreliminary Study of Sustainable Transport Indicators in Malaysia: The Case Study of Klang Valley Public Transportation, Procedia Environmental Sciences 28 (2015) 464–473, https://doi.org/10.1016/j.proenv.2015.07.056.

[76] R. Ding, N. Ujang, H. bin Hamid, J. Wu, Complex Network Theory Applied to the Growth of Kuala Lumpur's Public Urban Rail Transit Network, PLOS ONE 10 (10) (Oct. 2015) e0139961, https://doi.org/10.1371/journal.pone.0139961.
[77] J. Kuys, G. Melles, A. Al Mahmud, S. Thompson-Whiteside, B. Kuys, Human Centred Design Considerations for the Development of Sustainable Public Transportation in Malaysia, Applied Sciences 12 (23) (Dec. 2022) 12493, https://doi.org/10.3390/app122312493.
[78] M.N. Borhan, A.N.H. Ibrahim, D. Syamsunur, R.A. Rahmat, Why Public Bus is a Less Attractive Mode of Transport: A Case Study of Putrajaya, Malaysia, Periodica Polytechnica Transportation Engineering 47 (1) (Oct. 2017) 82–90, https://doi.org/10.3311/pptr.9228.
[79] OpenStreetMap contributors, "OpenStreetMap," *OpenStreetMap*, [Online]. Available: https://www.openstreetmap.org/#map=7/23.721/90.351. Accessed: Aug. 2024.
[80] Malaria Atlas Project, Global Friction Surface 2019, Google Earth Engine Data Catalog (2019) [Online].Available: https://developers.google.com/earthengine/datasets/catalog/Oxford_MAP_friction_surface_2019. Accessed: Aug. 2024.
[81] D.J. Weiss, et al., A global map of travel time to cities to assess inequalities in accessibility in 2015, Nature 553 (7688) (Jan. 2018) 333–336, https://doi.org/10.1038/nature25181.
[82] WorldPop, Malaysia 2020 Population Density (1km resolution, UN-adjusted), WorldPop, University of Southampton (2020) [Online]. Available: https://hub.worldpop.org/geodata/summary?id=47354. Accessed: Aug. 2024.
[83] MobilityData, Validation Rules and Metadata - Canonical GTFS Schedule Validator, GTFS Validator (2024) [Online]. Available: https://gtfs-validator.mobilitydata.org/rules.html. Accessed: Aug.
[84] S. Lee, J. Jeon, K. Cho, J. Im, Does Intercity Transportation Accessibility Matter? Its Effects on Regional Network Centrality in South Korea, Land 14 (4) (Apr. 2025) 873, https://doi.org/10.3390/land14040873.
[85] M. Stępniak, S. Goliszek, Spatio-Temporal Variation of Accessibility by Public Transport—The Equity Perspective, The Rise of Big Spatial Data (Oct. 2016) 241–261, https://doi.org/10.1007/978-3-319-45123-7_18.
[86] J. Gelb, P. Apparicio, Temporal Network Kernel Density Estimation, Geogr Anal 56 (1) (Jan. 2024) 62–78, https://doi.org/10.1111/GEAN.12368.
[87] E. Chevallier, E. Kalunga, J. Angulo, Kernel Density Estimation on Spaces of Gaussian Distributions and Symmetric Positive Definite Matrices, SIAM Journal on Imaging Sciences 10 (1) (Jan. 2017) 191–215, https://doi.org/10.1137/15m1053566.
[88] J.M. Leiva-Murillo, A. Artés-Rodríguez, Algorithms for maximum-likelihood bandwidth selection in kernel density estimators, Pattern Recognition Letters 33 (13) (Oct. 2012) 1717–1724, https://doi.org/10.1016/j.patrec.2012.06.006.
[89] H. Läuter, B.W. Silverman, Density Estimation for Statistics and Data Analysis. Biometrical Journal, Chapman & Hall, London – New York, 1986, pp. 876–877, https://doi.org/10.1002/bimj.4710300745, 30Jan. 1988. [Online]. Available:.
[90] P. Jia, J. Liu, C. Huang, L. Liu, C. Xu, An improvement method for degree and its extending centralities in directed networks, Physica A: Statistical Mechanics and its Applications 532 (Oct. 2019) 121891, https://doi.org/10.1016/j.physa.2019.121891.
[91] Q. Zhang, R. Deng, K. Ding, M. Li, Structural analysis and the sum of nodes' betweenness centrality in complex networks, Chaos, Solitons & Fractals 185 (Aug. 2024) 115158, https://doi.org/10.1016/j.chaos.2024.115158.
[92] W. Wang, C.Y. Tang, Distributed estimation of closeness centrality, in: 2015 54th IEEE Conference on Decision and Control (CDC), Dec. 2015, pp. 4860–4865, https://doi.org/10.1109/cdc.2015.7402978.
[93] C. Frei, H.S. Mahmassani, Riding More Frequently, Transportation Research Record: Journal of the Transportation Research Board 2350 (1) (Jan. 2013) 65–71, https://doi.org/10.3141/2350-08.
[94] İ.Ö. Verbas, C. Frei, H.S. Mahmassani, R. Chan, Stretching resources: sensitivity of optimal bus frequency allocation to stop-level demand elasticities, Public Transport 7 (1) (Jan. 2014) 1–20, https://doi.org/10.1007/s12469-013-0084-6.
[95] M. Saberi, A. Zockaie, W. Fang, A. El-Geneidy, Definition and Properties of Alternative Bus Service Reliability Measures at the Stop Level, Journal of Public Transportation 16 (1) (Mar. 2013) 97–122, https://doi.org/10.5038/2375-0901.16.1.6.
[96] C. Li and Z. Huang, "Measuring Stop Connectivity within a Transit Network," CICTP 2012, pp. 1577–1586, Jul. 2012, doi: https://doi.org/10.1061/9780784412442.161.
[97] T. Munkhbayar, Z. Dashdorj, H.-H. Cho, J.-W. Lee, T.-K. Kang, E. Altangerel, Assessing Urban Activity and Accessibility in the 20 min City Concept, Electronics 14 (8) (Apr. 2025) 1693, https://doi.org/10.3390/electronics14081693.
[98] S.J. Fatima, M.A. Javed, S.R. Ahmed, Assessing the Serviceability Status of Existing Transport Network by Using Geospatial Techniques in Metropolitan City Lahore, Pakistan Journal of Scientific & Industrial Research Series A: Physical Sciences 63 (3) (Nov. 2020) 201–208, https://doi.org/10.52763/pjsir.phys.sci.63.3.2020.201.208.
[99] M.M. Hasan, G.M.T. Islam, P.S.S. Matsumoto, M. Novak, K.R. Rahaman, Employing Geographic Information Systems in Analyzing Pedestrian Accessibility to Public Bus Stops in Halifax, Journal of Urban Planning and Development 149 (3) (Sep. 2023), https://doi.org/10.1061/jupddm.upeng-4268.
[100] S. Węglarczyk, Kernel density estimation and its application, in: ITM Web of Conferences 23, 2018 00037, https://doi.org/10.1051/itmconf/20182300037.
[101] G. Brandes, C. Sieg, M. Sander, R. Henze, Driving Domain Classification Based on Kernel Density Estimation of Urban Land Use and Road Network Scaling Models, Urban Science 8 (2) (May 2024) 48, https://doi.org/10.3390/urbansci8020048.
[102] X. Xu, A. Chen, S. Jansuwan, K. Heaslip, C. Yang, Modeling Transportation Network Redundancy, Transportation Research Procedia 9 (2015) 283–302, https://doi.org/10.1016/j.trpro.2015.07.016.
[103] Q. Tang, W. Dou, An Effective Method for Computing the Least-Cost Path Using a Multi-Resolution Raster Cost Surface Model, ISPRS International Journal of Geo-Information 12 (7) (Jul. 2023) 287, https://doi.org/10.3390/ijgi12070287.
[104] T.R. Etherington, E. Penelope Holland, Least-cost path length versus accumulated-cost as connectivity measures, Landscape Ecology 28 (7) (Mar. 2013) 1223–1229, https://doi.org/10.1007/s10980-013-9880-2.
[105] H. Lu, W. Zheng, T. Wang, S. Wang, An effective cost distance calculation based on raster data model improved algorithm, in: Proceedings of 2011 International Conference on Computer Science and Network Technology, Dec. 2011, pp. 2214–2218, https://doi.org/10.1109/iccsnt.2011.6182416.
[106] A. Sadeghi-Niaraki, K. Kim, M. Varshosaz, Multi-criteria decision-based model for road network process, Int J Environ Res 4 (4) (2010) 573–582, https://doi.org/10.22059/ijer.2010.243.
[107] R. Rodríguez-Puente, M.S. Lazo-Cortés, Algorithm for shortest path search in Geographic Information Systems by using reduced graphs, SpringerPlus 2 (1) (Jul. 2013), https://doi.org/10.1186/2193-1801-2-291.
[108] A.M.C.B. Silva, J.R.R. da Cunha, J.P.C. da Silva, Estimation of pedestrian walking speeds on footways, Proceedings of the Institution of Civil Engineers - Municipal Engineer 167 (1) (Mar. 2014) 32–43, https://doi.org/10.1680/muen.12.00048.
[109] O.B. Jensen, M. Martin, M. Löchtefeld, Pedestrians as floating life - On the reinvention of the pedestrian city, Emotion, Space and Society 41 (Nov. 2021) 100846, https://doi.org/10.1016/j.emospa.2021.100846.
[110] A. Rizzo, S. Khan, Johor Bahru's response to transnational and national influences in the emerging Straits Mega-City Region, Habitat International 40 (Oct. 2013) 154–162, https://doi.org/10.1016/j.habitatint.2013.03.003.
[111] A. Marzuki, M. Bagheri, A. Ahmad, T. Masron, M.F. Akhir, Examining transformations in coastal city landscapes: spatial patch analysis of sustainable tourism—a case study in Pahang, Malaysia, Landscape and Ecological Engineering 20 (4) (Jul. 2024) 513–545, https://doi.org/10.1007/s11355-024-00613-w.
[112] L. Buics, A. Prayogo, and B. E. Balassa, "The Role of Demand-Responsive Transport Systems in Sustainable Urban Mobility: A Systematic Literature Review and Stakeholder Analysis," SMTS 2024, p. 40, Nov. 2024, doi: https://doi.org/10.3390/engproc2024079040.
[113] A. Sharifi, Resilient urban forms: A review of literature on streets and street networks, Building and Environment 147 (Jan. 2019) 171–187, https://doi.org/10.1016/j.buildenv.2018.09.040.
[114] S. Hamnett and D. Forbes, Eds., "Planning Asian Cities," Mar. 2012, doi: https://doi.org/10.4324/9780203804247.
[115] M. Shafiq, M. Amorim, A. Couto, Assessment of integrated rail and bus transport network design: Equity and efficiency perspectives, Transportation Research Interdisciplinary Perspectives 31 (May 2025) 101450, https://doi.org/10.1016/j.trip.2025.101450.
[116] T. Zhang, S. Dong, Z. Zeng, J. Li, Quantifying multi-modal public transit accessibility for large metropolitan areas: a time-dependent reliability modeling approach, International Journal of Geographical Information Science 32 (8) (Apr. 2018) 1649–1676, https://doi.org/10.1080/13658816.2018.1459113.
[117] E. Flom, Incorporating an equity-based approach into transportation operations, Public Roads 87 (1) (2023) 23–27 [Online]. Available: https://trid.trb.org/View/2213668.
[118] Y.A. Abdullah, N.S. Khalid, F.N. Nawawi, M.A. Marzukhi, Z.Ahmad Zaki, O.L. H. Leh, Rail-Based Transit-Oriented Development (TOD) From the Public's Perspectives in the Klang Valley, Malaysia, in: IOP Conference Series: Earth and Environmental Science 1217, Jul. 2023 012025, https://doi.org/10.1088/1755-1315/1217/1/012025.
[119] H. Zakaria, D. Kamarudin, F. Azizul, M.F. Idrus, N.R.H.M. Haron, N.M. Aripin, Automatic Vehicle Location (AVL): Evaluation on the Punctuality Index of City Public Bus Service, Advances in Intelligent Manufacturing and Mechatronics (2023) 75–86, https://doi.org/10.1007/978-981-19-8703-8_7.
[120] O. Sánchez, K. Castañeda, S. Vidal-Méndez, D. Carrasco-Beltrán, N.E. Lozano-Ramírez, Exploring the influence of linear infrastructure projects 4.0 technologies to promote sustainable development in smart cities, Results in Engineering 23 (Sep. 2024) 102824, https://doi.org/10.1016/j.rineng.2024.102824.
[121] V. Noferesti, H. Mirzahossein, Leveraging machine learning to predict residential location choice: A comparative analysis, Results in Engineering 25 (Mar. 2025) 104214, https://doi.org/10.1016/j.rineng.2025.104214.
[122] R. Gomes, J. Pinho de Sousa, T.Galvão Dias, Sustainable Demand Responsive Transportation systems in a context of austerity: The case of a Portuguese city, Research in Transportation Economics 51 (Sep. 2015) 94–103, https://doi.org/10.1016/j.retrec.2015.07.011.
[123] S. Raveau, Demand-Responsive Transit, Evaluation Studies, International Encyclopedia of Transportation (2021) 423–427, https://doi.org/10.1016/b978-0-08-102671-7.10361-6.
[124] J.A. Mesa, F.A. Ortega, M.A. Pozo, R. Piedra-de-la-Cuadra, Assessing the effectiveness of park-and-ride facilities on multimodal networks in smart cities, Journal of the Operational Research Society 73 (3) (Jan. 2021) 576–586, https://doi.org/10.1080/01605682.2020.1854628.
[125] S. Zyoud, A.H. Zyoud, Advancing sustainable cities and communities with internet of things: Global insights, trends, and research priorities for SDG 11, Results in Engineering 26 (Jun. 2025) 104917, https://doi.org/10.1016/j.rineng.2025.104917.

[126] V. Mohamad Taghvaee, et al., Sustainable development goals: transportation, health and public policy, Review of Economics and Political Science 8 (2) (Apr. 2023) 134–161, https://doi.org/10.1108/reps-12-2019-0168.
[127] G. Tiwari, C. Phillip, Development of public transport systems in small cities: A roadmap for achieving sustainable development goal indicator 11.2, IATSS Research 45 (1) (Apr. 2021) 31–38, https://doi.org/10.1016/j.iatssr.2021.02.002.
[128] V.M. Taghvaee, M. Nodehi, R.M. Saber, M. Mohebi, Sustainable development goals and transportation modes: Analyzing sustainability pillars of environment, health, and economy, World Development Sustainability 1 (2022) 100018, https://doi.org/10.1016/j.wds.2022.100018.
[129] A. Kwilinski, O. Lyulyov, T. Pimonenko, Environmental Sustainability within Attaining Sustainable Development Goals: The Role of Digitalization and the Transport Sector, Sustainability 15 (14) (Jul. 2023) 11282, https://doi.org/10.3390/su151411282.
[130] M. Hassan, H.D. Mahin, A. Al Nafees, A. Paul, S.S. Shraban, Big data applications in intelligent transport systems: a bibliometric analysis and review, Discover Civil Engineering 2 (1) (Mar. 2025), https://doi.org/10.1007/s44290-025-00205-z.
[131] M. Hassan, A. Al Nafees, S.S. Shraban, A. Paul, H.D. Mahin, Application of machine learning in intelligent transport systems: a comprehensive review and bibliometric analysis, Discover Civil Engineering 2 (1) (May 2025), https://doi.org/10.1007/s44290-025-00256-2.
[132] G. Ramkumar, S. Kannan, V. Mohanavel, S. Karthikeyan, A. Titus, The future of green mobility: A review exploring renewable energy systems integration in electric vehicles, Results in Engineering 27 (Sep. 2025) 105647, https://doi.org/10.1016/j.rineng.2025.105647.
[133] C. Banyong, et al., A machine learning comparison of transportation mode changes from high-speed railway promotion in Thailand, Results in Engineering 24 (Dec. 2024) 103110, https://doi.org/10.1016/j.rineng.2024.103110.
[134] P.M. Donald, D.N. Salamida, P.A. Panga, M.M. Mayengo, Modeling and control of global warming driven by transportation-induced carbon dioxide emissions through green economy investments, Results in Engineering 26 (Jun. 2025) 105048, https://doi.org/10.1016/j.rineng.2025.105048.
[135] P.K. Senecal, F. Leach, Diversity in transportation: Why a mix of propulsion technologies is the way forward for the future fleet, Results in Engineering 4 (Dec. 2019) 100060, https://doi.org/10.1016/j.rineng.2019.100060.