Accessibility Analysis of East Asian Metro Systems

Hanyu Cheng*

[1]Delft University of Technology, department of Transport & Planning at the faculty of Civil Engineering and Geosciences, Delft, the Netherlands

ABSTRACT

Accessibility reflects the ease with which different individuals can overcome travel impedance to reach spatially-distributed opportunities. Since accessibility of public transport networks is jointly determined by network topology and service attributes, this study applies access graphs derived from time-weighted L-space and frequency-weighted P-space graphs as the standardized framework. For previously underresearched East Asia region, a new dataset is compiled from open-source data for 61 systems and access graphs are constructed over increasing generalized travel-time budgets, and reachability and equity indicators at critical time points are benchmarked against systems in other world regions. Finally, based on the temporal evolution of average degree, East Asian and metro networks worldwide are classified into four clusters using k-means clustering. The findings show that East Asian metro systems vary widely in sizes. Influenced by the accessibility growth pattern, most medium- to mega-sized networks follow logistic (S-shaped) curves of degree growth consistent with a core–branch structure, whereas smaller or systems with degraded service have irregular degree growth curves. In terms of performance, large systems show greater spatial disparity and less uniform service quality, resulting in lower equity and reachability at 30 minutes. Regardless of size, East Asian metros tend to underperform in the medium and late stages of accessibility growth due to early decay in degree-growth rate and limited reachability at time of maximum degree growth rate. When controlling for size, East Asian metros remain less reachable than European systems and experience earlier growth-rate decay than both European and North American counterparts. Based on the timing of the peak and subsequent decay in degree growth rate, four distinctive accessibility growth patterns are identified across the worldwide systems. Clustering analysis further reveals that large East Asian networks experience an earlier decline in accessibility growth, similar to networks lacking direct cross-line connections, whereas most medium and small networks sustain growth longer on par with those in other regions. The presence of networks of similar size in different clusters suggests that improving service attributes can enhance overall accessibility performance.

Under the supervision of Prof. dr. O. (Oded) Cats and Dr. Dipl.-Ing. R. (Rajat) Verma

Keywords: Accessibility, Metro networks, Network science

1 Introduction

Accessibility is the ease with which individuals can overcome distance barrier and reach spatially distributed opportunities via transport systems[1]. Accordingly, it should be the primary performance objective of public transport networks (PTNs)[2]. Recent studies have shown that the degree to which a PTN facilitates accessibility depends jointly on service attributes like frequency and the topology. Luo et al. formulate a generalized travel cost metric consisting of in-vehicle time, waiting time, and transfer penalty costs on network graphs and reveal that including service properties reveals stronger spatial disparities in accessibility and larger networks show higher average impedance due mainly to longer in-vehicle times[3]. Luo et al. further demonstrate that weighted and service-space topological indicators correlate strongly with observed passenger flows, supporting the view that accessibility patterns are reflected in network topology[4]. Jin et al. (2017) use GIS-based travel-time analysis to assess regional accessibility in East Asian high-speed rail networks, showing that network expansion can enhance accessibility around core cities but may also result in higher spatial disparities in certain parts of the region[5].

While the incorporation of accessibility into topology analysis is relatively recent, the study of PTNs' topology using graph theory from network science is long-established[6]. Its mature phase emerge around 2010 with the works of von Ferber et al. and Derrible & Kennedy. Von Ferber et al.[7] analyse worldwide PTNs in standardized graph representations and reported statistical patterns including power-law (scale-free) degree behaviour and small-world properties. Derrible and Kennedy[8–10] systematically apply graph-theory principles to metro networks, likewise identifying small-world and scale-free features, and suggest using state, form and structure to quantify network complexity, connectivity and directness. To preserve key features while capturing essential topological characteristics, L-space and P-space are most commonly used in afore-mentioned studies. L-space models the infrastructure graph where nodes are stations and links connect consecutive stations along a line[7], whereas

P-space models the service graph where nodes are still stations but links represent transfer-free connections between any pair of stations[7]. In the context of metro networks, other important structural characteristics have also been identified. Roth et al.[11] reveal a universal structural pattern in which sophisticated and developed metro systems tend to converge toward a dense, highly interconnected core with radiating branches.

Given the extensive body of topological research on PTNs including metro networks and the fact that metros, with their high capacity and exclusive right-of-way operations, can substantially reduce travel impedance[12], this study focuses on metro systems. Building on evidence from other rail-bound systems showing that both service attributes and network topology influences accessibility[3–5], and considering the recurring structural features of metro networks such as scale-free degree distributions, small-world connectivity, and convergence toward a core–branch layout[8–11], this study examines how metro system accessibility can be measured by integrating service attributes into network topology and whether similar accessibility patterns can be observed across networks. In this study, accessibility refers exclusively to the travel impedance involved in reaching destinations across the network, while the concept of metro follows the general and broad definition established in the UITP report[12]. The structure of this study is illustrated in the following figure 1.

To provide a generalized and comparable framework for accessibility study, this study adopts the **access graph**, a recently developed graph-based method proposed by Sfiligoj et al.[13]. The construction of access graph further discussed in chapter 4.2 and the accessibility of metro networks is quantified using seven concise indicators derived from the access graph, which collectively evaluate accessibility from three perspectives: (i) Critical time: t_M and $t_{M,\mathrm{rel}}$, where t_M is the time budget at which the average-degree growth rate is maximal, and $\delta t_M = t_M/t_{\max}$ is its relative form with respect to the maximum generalized travel time; (ii) Reachability: $\overline{D}_M$, normalized average degree at t_M and $\overline{D}_{30}$, normalized average degree at 30 minutes; (iii) Equity: G_M and G_{30} (Gini coefficients of the degree distribution at t_M and 30 minutes). Additionally, the size of all metro networks is denoted by the number of stations, N, which has a systematic influence on topological complexity and accessibility performance[2,3,10,14].

In the access graph at time budget t_b, the degree of a station is the number of other stations reachable from it within t_b. 30 minutes threshold is chosen for $\overline{D}_{30}$ and G_{30} as it aligns with typical daily travel-time budget per direction[15,16] and has been applied in studies of 30-minute reachable areas in East Asian PTNs[2,17,18]. At $t_{\max}$, all station pairs are connected and the access graph becomes complete ($\bar{D}(t_{\max}) = N-1$). Since the growth curve of the average degree in access graphs follows an S-shape in many worldwide metro systems[13], the most informative moment occurs when each additional unit of travel time yields the largest marginal gain in reachability. Beyond t_M, further increments in travel time investment become less efficient for passengers as it no longer yield the same amount of newly reachable stations. Additionally, since people's daily travel time budgets are roughly stable across cities[16], t_M can serve as an absolute indicator of when the network becomes generally accessible, with smaller t_M indicating earlier broad reach. Since larger networks tend to naturally have higher travel impedance[3], the normalized indicator $t_{M,\mathrm{rel}} = \delta t_M = t_M/t_{\max}$ is used to factor out the size effect, where smaller δt_M indicates earlier occurrence of the peak growth in average degree relative to the system's full extent. $\overline{D}_M$, $\overline{D}_{30}$ characterize the state of interconnection of station pairs and path diversity and can be used to track the pattern of accessibility gain in the access graph[10,13]. Unevenness in station degrees between the downtown core and suburban branches has been observed[11,14] and accordingly, the Gini coefficients G_M (at t_M) and G_{30} (at $t_b = 30$ min) are used as they are computationally simple, applicable on region level and provide diagnostics towards critical area to accurately evaluate inequality in the degree distribution[11,14,19]. Accordingly, the Gini coefficients G_M (at t_M) and G_{30} (at $t_b = 30$ min) are adopted for accurately assessing inequality in the degree distribution, as they are computationally simple, applicable at the urban scale, and provide diagnostic insights into critical areas[19]. Together, these indicators enable consistent and robust comparisons across cities and regions.

In terms of regional scope, this study focuses on East Asian metro systems, as previous accessibility analyses have primarily concentrated on Europe and North America[13], where most metro systems are completed early and their development has since then plateaued[11,12]. In contrast, East Asian cities possess favourable socio-economic conditions for metro development[20] and have experienced rapid, large-scale network expansion since the 2010s[12]. As of 2024, the region concentrates a substantial share of the world's metro infrastructure, with China, Japan, and South Korea together accounting for seven of the ten longest networks and 52.25% of global metro length[12]. Therefore, the region warrants a regional-level analysis and a comparative study with the rest of the world. However, existing accessibility research on East Asian metro networks remains limited in scope, often focusing on individual cities. For instance, the OECD/ITF (2023) report[18] finds that accessibility of Seoul's metro network measured by door-to-door time and cumulative opportunities is highest in the city centre and declines sharply toward suburban areas due to lower service frequencies and network densities. Due to the limited number of metro networks being studied, no prior large-scale comparative study has placed all major East Asian metro networks within an international context alongside other regions. Although in the broader context of PTNs, Wu et al.[2] show that compact urban form and high-frequency PT services in Chinese cities yield high accessibility levels comparable to those in Europe, whereas North American networks lag behind.

The previous lack of research on the accessibility of East Asian metro networks can largely be attributed to the unavailability

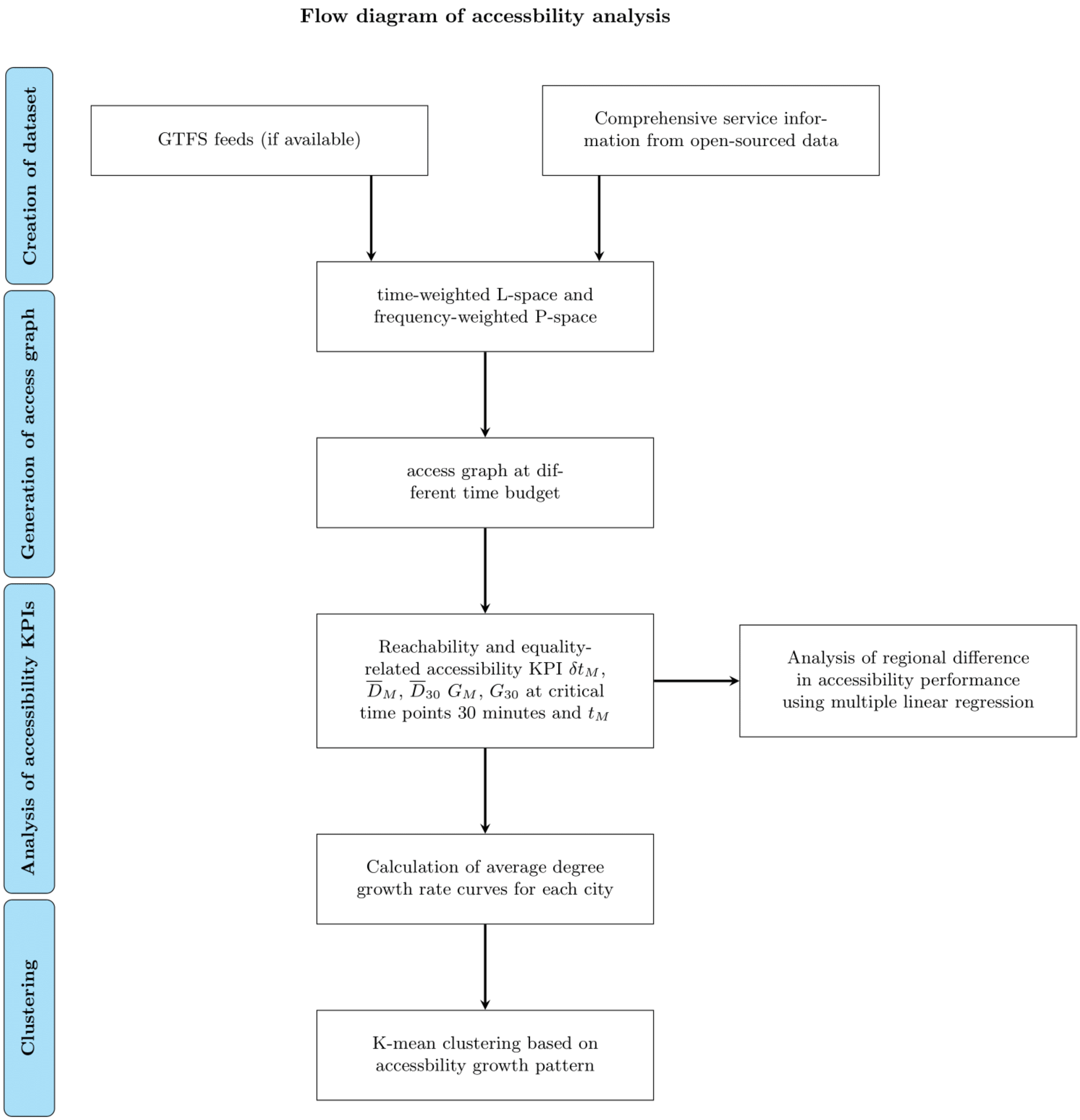


Figure 1. The flow chart of accessibility analysis of metro networks

of GTFS data in East Asia[2,17,21,22]. Nevertheless, recent studies have demonstrated that open sourced data like localized map service can serve as GTFS-equivalent sources to overcome this limitation[2,17]. Following such solution, this study constructs a comprehensive dataset of metro networks in major East Asian cities using open-source data[23,24], and systematically evaluates their accessibility performance at the regional level and in comparison with networks from other geographical regions. By broadening the research scope and introducing region-level generalization in accessibility analysis from a topological perspective, this study aims to address the following research question: **How do East Asian metro systems perform in terms of accessibility, and how can they be characterized when considered alongside networks from other geographical regions?**

2 Results

2.1 Selection and categorization of East Asian metro networks

To analyse metro networks, the selection of rail-bound systems follows the definition in UITP report. A line is treated as metro if it primarily serves an urban area (with possible extensions to adjacent municipalities), operates on an exclusive right-of-way that is fully grade-separated, and delivers at least medium capacity at high frequency[12]. Some metropolitan commuter rail lines in Seoul, Busan, or Tokyo may satisfy these criteria in practice. In this study, such lines are **not** classified as metro unless they are operated as an integral part of the metro system, i.e., (i) directly operated by the metro operator (ii) use metro-standard rolling stock and (iii) provide through-running or tightly coordinated services with connecting metro lines.

The numerical values of the accessibility indicators and the number of stations N for the East Asia region are summarized in Table 1, and their geographical locations and t_{max} value of metro networks are shown in Figure 2. Across the 62 East Asian systems, N ranges from 15 (Dongguan, Taizhou, Macau) to 417 (Beijing). The mean and median are 120.26 and 91.00, respectively, indicating a right-skewed distribution dominated by a few very large systems.

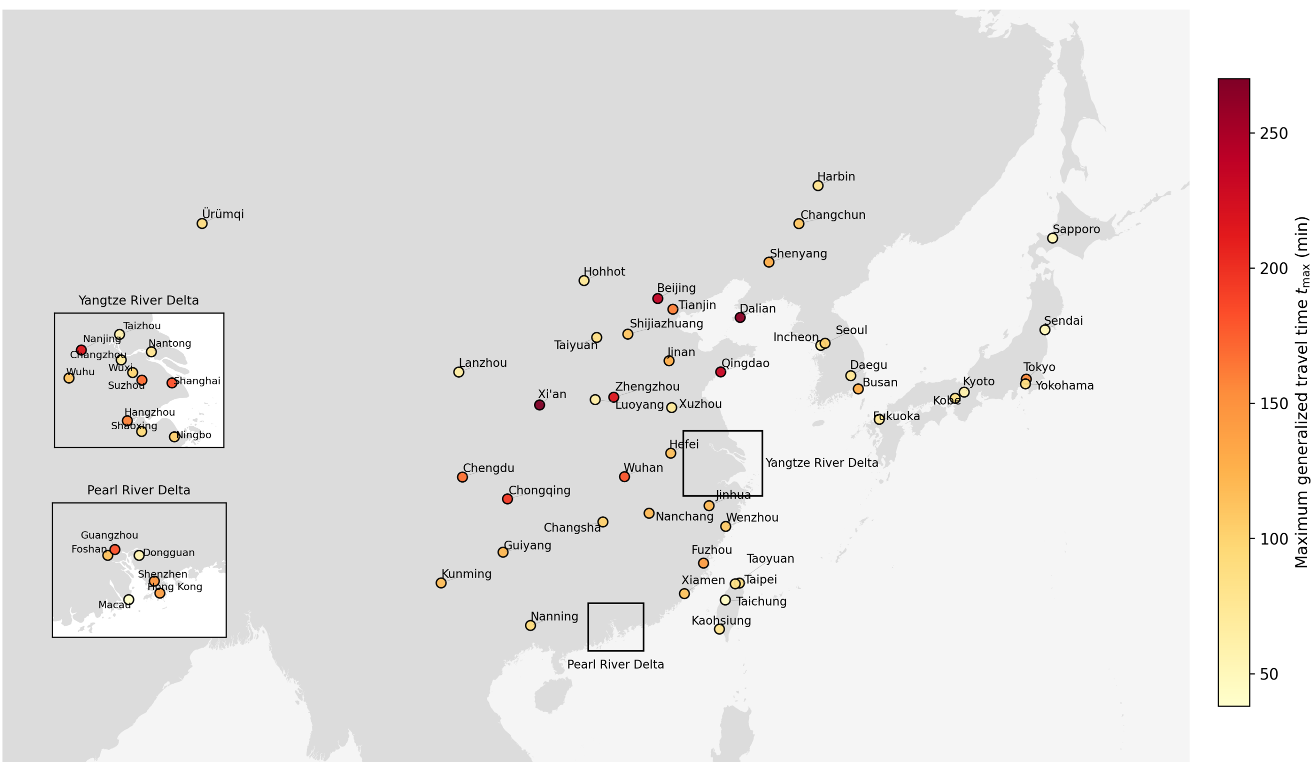


Figure 2. East Asian networks included in the study. Each dot marks a city and its filled colour indicates the maximum generalized travel time t_{max} (minutes). Insets enlarge the urban agglomerations in Yangtze River Delta and Pearl River Delta for legibility.

On rail-bound public transport networks (PTNs), larger systems have been previously found to exhibit greater travel impedance[3]. Therefore, it is necessary to classify metro systems by size. However, previous research do not classify worldwide metro networks by size, but instead analyse how topological properties evolve continuously with increasing network scale, denoted by N[8–11]. A recent review categorizes Chinese metro systems by the number of stations (1–50, 51–100, 101–200, 201–300, and > 300)[25], noting that this measure correlates with socio-economic factors, administrative hierarchy, and system length. This study extends this classification to other East Asian cities, as they share broadly similar socio-economic conditions and development contexts for expansion of metro networks[20]. Following this approach, the 62 East Asian metro networks analysed in this study are grouped by N into five categories: micro ($1 \leq N \leq 50$; 24 cities), small ($51 \leq N \leq 100$; 13), medium ($101 \leq N \leq 200$; 10), large ($201 \leq N \leq 300$; 10), and mega ($N > 300$; 5).

Table 1. Values of accessibility indicators for East Asian metro networks (alphabetical by city; floats with 2 decimals). N: number of nodes in the network; $t_{\max}$: maximum generalised travel time (min); t_M: time of maximum average degree growth (min); $\delta t_M = t_M / t_{\max}$; $\overline{D}_M$: average degree at t_M; $\overline{D}_{30}$: average degree at $t_b = 30$ min; G_M: Gini coefficient at t_M; G_{30}: Gini coefficient at $t_b = 30$ min.

City	N	$t_{\max}$	t_M	δt_M	$\overline{D}_M$	$\overline{D}_{30}$	G_M	G_{30}
Beijing	417	238	73	0.31	0.40	0.03	0.31	0.28
Busan	127	128	41	0.32	0.39	0.18	0.25	0.20
Changchun	111	110	41	0.37	0.50	0.22	0.19	0.19
Changsha	141	102	47	0.46	0.52	0.16	0.25	0.30
Changzhou	43	70	35	0.50	0.51	0.39	0.19	0.19
Chengdu	324	166	51	0.31	0.48	0.11	0.26	0.37
Chongqing	270	194	49	0.25	0.38	0.11	0.31	0.38
Daegu	91	76	43	0.57	0.70	0.32	0.16	0.24
Dalian	100	264	49	0.19	0.35	0.14	0.27	0.29
Dongguan	15	56	17	0.30	0.10	0.51	0.21	0.14
Foshan	64	114	43	0.38	0.48	0.25	0.21	0.21
Fukuoka	36	74	11	0.15	0.13	0.45	0.17	0.16
Fuzhou	93	142	47	0.33	0.39	0.13	0.22	0.23
Guangzhou	309	182	51	0.28	0.37	0.10	0.32	0.40
Guiyang	93	120	45	0.38	0.43	0.17	0.22	0.20
Hangzhou	258	158	45	0.29	0.36	0.12	0.29	0.32
Harbin	73	78	39	0.50	0.68	0.36	0.13	0.17
Hefei	172	114	43	0.38	0.35	0.14	0.30	0.28
Hohhot	43	74	43	0.58	0.66	0.37	0.15	0.17
Hong Kong	97	140	43	0.31	0.51	0.24	0.25	0.35
Incheon	68	68	31	0.46	0.52	0.46	0.21	0.22
Jinan	46	128	43	0.34	0.45	0.25	0.16	0.19
Jinhua	30	118	17	0.14	0.09	0.27	0.17	0.13
Kaohsiung	38	80	13	0.16	0.12	0.41	0.16	0.16
Kunming	103	116	47	0.41	0.49	0.19	0.21	0.25
Kyoto	31	60	9	0.15	0.06	0.53	0.12	0.15
Kobe	26	84	21	0.25	0.20	0.37	0.26	0.19
Lanzhou	27	66	11	0.17	0.07	0.50	0.13	0.17
Luoyang	33	64	11	0.17	0.06	0.42	0.09	0.16
Macau	15	38	11	0.29	0.20	0.84	0.28	0.08
Nanchang	113	118	41	0.35	0.42	0.21	0.28	0.31
Nanjing	225	216	55	0.26	0.43	0.10	0.30	0.36
Nanning	93	92	41	0.45	0.54	0.23	0.16	0.19
Nantong	43	78	29	0.37	0.36	0.37	0.25	0.25
Ningbo	137	102	37	0.36	0.46	0.26	0.25	0.30
Qingdao	177	232	35	0.15	0.20	0.12	0.36	0.36
Sapporo	46	52	33	0.64	0.69	0.52	0.16	0.21
Sendai	29	50	11	0.22	0.13	0.59	0.25	0.15
Seoul	296	102	35	0.34	0.35	0.20	0.28	0.31
Shanghai	411	182	51	0.28	0.36	0.07	0.34	0.44
Shaoxing	41	94	17	0.18	0.04	0.25	0.13	0.19

Continued on next page

City	N	t_{max}	t_M	δt_M	$\overline{D}_M$	$\overline{D}_{30}$	G_M	G_{30}
Shenyang	134	128	41	0.32	0.39	0.17	0.28	0.27
Shenzhen	325	146	35	0.24	0.27	0.17	0.35	0.38
Shijiazhuang	61	110	35	0.32	0.37	0.27	0.18	0.18
Suzhou	237	168	43	0.26	0.41	0.14	0.27	0.28
Taichung	18	44	11	0.25	0.14	0.77	0.14	0.10
Taipei	118	88	33	0.38	0.46	0.34	0.25	0.29
Taiyuan	46	86	39	0.45	0.55	0.36	0.17	0.17
Taizhou	15	60	13	0.22	0.10	0.54	0.17	0.12
Taoyuan	22	84	25	0.30	0.26	0.34	0.10	0.09
Tianjin	223	160	49	0.31	0.45	0.11	0.28	0.30
Tokyo	251	154	51	0.33	0.33	0.06	0.33	0.30
Wenzhou	36	106	13	0.12	0.05	0.24	0.16	0.16
Wuhan	272	180	55	0.31	0.46	0.10	0.27	0.34
Wuhu	35	112	17	0.15	0.06	0.22	0.16	0.13
Wuxi	91	98	39	0.40	0.45	0.23	0.25	0.28
Xiamen	72	112	39	0.35	0.46	0.26	0.24	0.26
Xian	238	270	41	0.15	0.32	0.14	0.32	0.34
Xuzhou	57	72	45	0.62	0.70	0.28	0.16	0.22
Yokohama	40	86	5	0.06	0.05	0.51	0.20	0.13
Zhengzhou	238	216	45	0.21	0.31	0.08	0.32	0.33
Ürümqi	22	88	17	0.19	0.08	0.44	0.09	0.16

2.2 Accessibility growth pattern of East Asian metro networks

The readability of access graphs is constrained by dense edges at higher time budgets. Therefore, for cross-city comparison at the regional level, this study summarizes the evolution of overall accessibility on metro networks with heatmaps of degree distributions overlaid with the average degree, as shown in Figure 3. As noted in chapter 1, degree measures interconnection, while the degree distribution provides a distributional perspective on equity. Summarizing both across the full time-budget range, Figure 3 enables systematic and streamlined cross-city comparison, complements and forms mathematical basis for accessibility indicators, and reveals how accessibility expands within each network. In Figure 3, each subplot is a heatmap of node–degree distributions, with the time budget t_b increasing in 2 minute steps. In every subplot, the x-axis denotes the time budget (min) and the y-axis denotes the degree in the access graph at the corresponding time. Heatmap cells encode the percentage of nodes in each degree bin, and the average degree $\overline{D}(t_b)$ is indicated with red markers. If common accessibility growth patterns are observed in topologically similar networks, the influence of topology on accessibility can then be examined by referring to the L-space of the networks.

As shown in Figure 3, many cities follow a logistic S-shaped growth pattern in the average degree of their access graphs but exceptions are also common. To define the core precisely, study follows Roth et al.[11] and selects the k-core with $k = 2$, where all one-degree nodes are iteratively removed until none remain. The remaining subgraph is the core, and the removed set forms the branches. Denoting the numbers of nodes and edges in the full graph by N and M, and in the core/branches by N_C, M_C and N_B, M_B, the fraction of branch is $\beta = N_B/N$. Core connectivity is summarized by the core's average degree $k_{\text{core}} = 2M_C/N_C$ and by the fraction of degree-2 core nodes $f_2 = |\{n \in N_C : \deg_C(n) = 2\}|/N_C$. A lower f_2 indicates greater interconnection, since degree-2 nodes typically have one line passing through and do not offer transfer opportunities.

Within category 1, 70% (21/30) of cities have $N > 100$ stations and converge toward a stable core–branch structure, while the remaining 30% (9/30) are smaller systems with N between 64 and 97. The branch share β ranges from 0.24 to 0.60 (mean = median = 0.42); k_{core} ranges from 2.12 to 2.59 (mean 2.37, median 2.36); and f_2 ranges from 0.69 to 0.90 (mean 0.77, median 0.77). This is consistent with the long-time limits reported by Roth et al. (2012)[11], who observed that for mature systems ($N \gtrsim 100$) the branch share stabilizes near $\beta \in [0.35, 0.55]$ (with many around 45%), the core mean degree around $k_{\text{core}} \approx 2.3$–$2.5$, and $f_2 \gtrsim 0.6$. It is worth noting that most category 1 systems fall squarely within these convergence ranges, with a slightly wider lower bound of β and a higher upper bound of k_{core} as a result of the rapid expansion of East Asian metro networks after the 2010s[12]. However, two mild outliers are observed for the branch share β and the core mean degree k_{core}, notably Busan ($\beta = 0.60$, $k_{\text{core}} = 2.12$) and Harbin ($\beta = 0.58$, $k_{\text{core}} = 2.15$). Their weaker topological connectivity is effectively compensated by high service frequency and short in-vehicle times along major corridors, resulting in overall accessibility levels and average degree growth pattern comparable to those of networks with denser cores.

In contrast, networks in category 2 generally exhibit a higher branch share, with β ranging from 0.45 to 0.65 (mean 0.55,

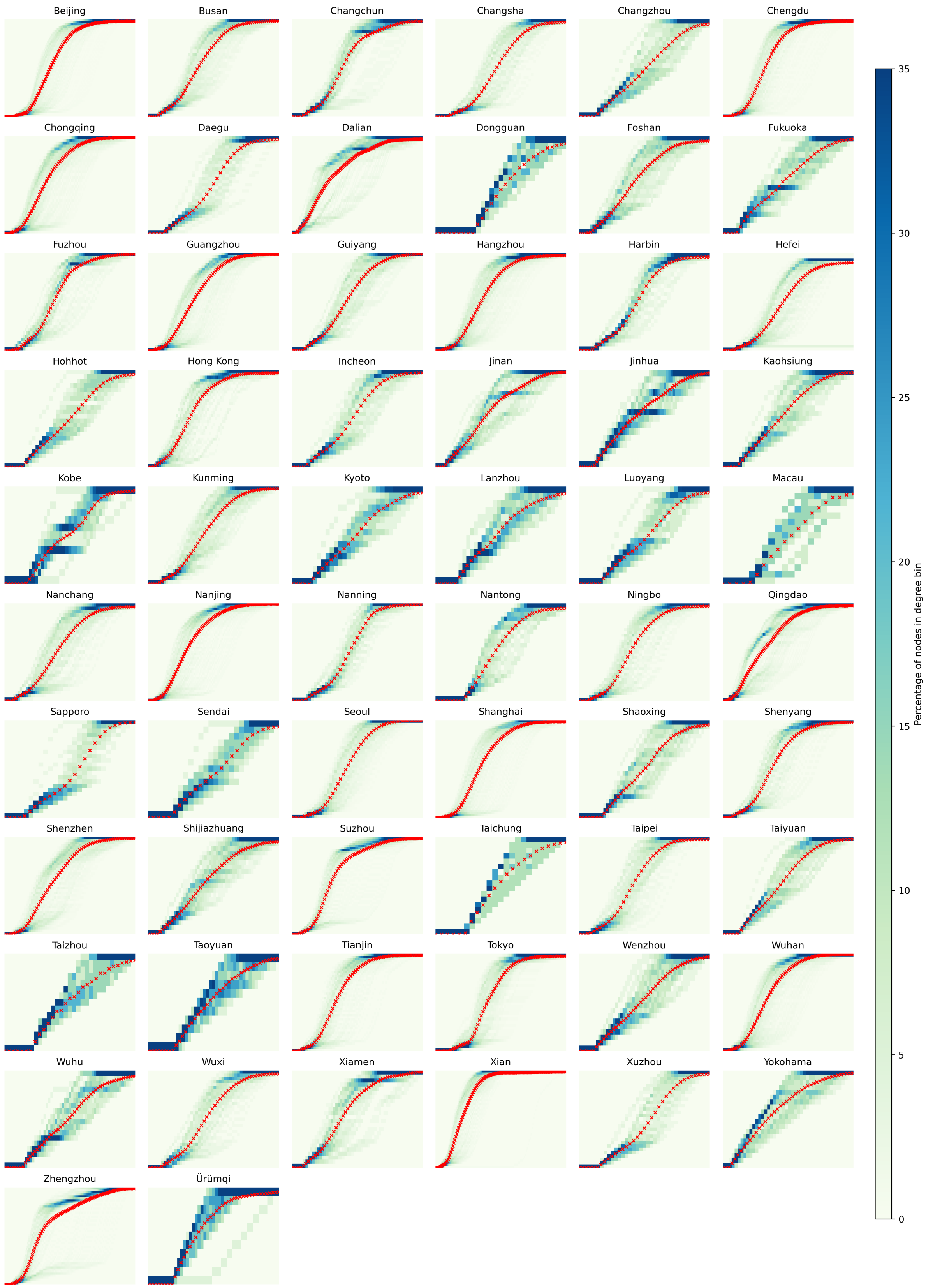


Figure 3. Degree distributions and average degree growth curve of each East Asian's access graph with varying t_b. In each subplot, y-axis indicates the degree and x-axis indicates time budget.

median 0.55), a less dense and less transferable core with k_{core} ranging from 2.12 to 2.31 (mean 2.22, median 2.21), and f_2 ranging from 0.79 to 0.90 (mean 0.84, median 0.85). Since k_{core} does not converge to the mature-network range $k_{\text{core}} \approx 2.3$–2.5 and the mean branch share lies at the upper bound of the convergence interval $\beta \in [0.35, 0.55]$, these systems can be interpreted as less developed networks compared with those in category 1. Their structural deficiencies are not compensated by service attributes, resulting in a slower pace of accessibility growth and less regular S-shaped growth curves.

While the regularity of the S-shaped curves differs between category 1 and category 2, both still follow the general pattern of accessibility growth. In both cases, as indicated by high f_2 values, increases in the average degree depend heavily on transfer stations within the core that connect multiple lines. In the early stage, most connected station pairs are non-transferable nodes located on separate lines, and cross-line connections are largely limited to stations adjacent to transfer hubs. As a result, the growth rate remains slow, with connections forming only within isolated subgraphs. As the time budget increases, more non-transferable stations within the core begin to reach transfer stations, allowing previously isolated subgraphs to coalesce through these transfer stations. This coalescence results in a rapid phase of accessibility growth, marked by a surge in cross-line connectivity and culminating in a highly interconnected core. The inflection point of the S-curve typically occurs at this stage. Beyond the inflection point, most remaining isolated subgraphs are located outside the core and on suburban branches and require substantially longer travel times to become connected to the core. As a result, average degree growth in the late stage is primarily driven by the gradual integration of these remote subgraphs with the core, leading to a deceleration in accessibility growth. The interconnection of stations on each line's terminal areas results in the eventual completion of the S-curve.

Category 3 consists mainly of micro-sized metro networks with the exceptions of Dalian, Qingdao, and Zhengzhou, which are substantially larger than the rest of the cities. The branch fraction of category 3 is highest among all three categories with β mean $= 0.777$, median $= 0.770$. It is worth noting that single-line or simple cross-shaped metro networks without discernible core appear only in this category and are effectively one-line equivalents with $\beta = 1$. Given the pronounced branching effects, topology has limited influence on accessibility compared with service attributes like in-vehicle time and frequency. Consequently, as shown in Figure 3, the degree-growth curves and degree distribution corresponding to different time budget are irregular and non-uniform because service attributes vary substantially across cities. The networks of Dalian ($\beta = 0.60$, $k_{\text{core}} = 2.13$, $f_2 = 0.89$) and Qingdao ($\beta = 0.46$, $k_{\text{core}} = 2.23$, $f_2 = 0.85$) have cores with low transferability and are dominated by suburban lines with long in-vehicle travel times and thus their degree-growth curves differ from networks with similar size. For Zhengzhou, the irregularity is mainly caused by station pairs involving the Zhengxu line (Zhengzhou–Xuchang), a low-frequency intercity metro lines that reduces the percentage of high-degree stations after the inflection point t_M and consequently dampens the late-stage ascent of the S-curve.

In sum, metro networks in cities with a distinct core–branch structure generally follow a clear S-shaped growth curve. By contrast, systems with a less developed core tend to exhibit less regular S-shaped patterns. Networks with very high branch fractions or low-frequency suburban lines are strongly influenced by service attributes and therefore deviate from the typical S-shaped accessibility growth. Overall, the results demonstrate that the degree to which a metro network exhibits logistic accessibility growth depends jointly on its topological maturity and operational characteristics. With the accessibility growth pattern being investigated, the following chapters 2.3 discuss the values of the accessibility indicators derived from access graph to evaluate the overall accessibility performance of East Asian metro networks.

Table 2. Classification East Asian metro networks' growth curves of average degree. Networks with clearly logistic (S-shaped) growth patterns are in Category 1; less regular S-trends in Category 2; and irregular, noisy, or non-logistic curves in Category 3. Core–branch indicators by city, grouped by visual growth-curve category. β is the branch share; $k_{\text{core}} = 2M_c/N_c$ is the mean degree of the 2-core; f_2 is the share of degree-2 nodes in the 2-core (all values rounded to 2 decimals).

Category	City	β	k_{core}	f_2
1 Clearly Logistic				
	Beijing	0.31	2.59	0.69
	Busan	0.60	2.12	0.90
	Changchun	0.24	2.17	0.89
	Changsha	0.52	2.35	0.78
	Chengdu	0.31	2.43	0.78
	Chongqing	0.38	2.39	0.74
	Foshan	0.42	2.31	0.79
	Fuzhou	0.38	2.17	0.86
	Guangzhou	0.32	2.51	0.69
	Guiyang	0.53	2.30	0.78

Continued on next page

Category	City	β	k_{core}	f_2
	Hangzhou	0.32	2.50	0.71
	Harbin	0.58	2.15	0.87
	Hefei	0.40	2.36	0.77
	Hong Kong	0.40	2.37	0.75
	Incheon	0.49	2.31	0.78
	Kunming	0.44	2.28	0.80
	Nanchang	0.48	2.35	0.78
	Nanjing	0.33	2.49	0.70
	Nanning	0.45	2.28	0.82
	Ningbo	0.47	2.32	0.79
	Seoul	0.41	2.47	0.72
	Shanghai	0.34	2.53	0.69
	Shenyang	0.40	2.36	0.77
	Shenzhen	0.36	2.47	0.71
	Taipei	0.43	2.36	0.76
	Tianjin	0.38	2.40	0.74
	Tokyo	0.42	2.43	0.73
	Wuhan	0.37	2.44	0.74
	Wuxi	0.48	2.30	0.79
	Xiamen	0.51	2.28	0.81
2 Moderately Logistic				
	Changzhou	0.52	2.27	0.83
	Daegu	0.58	2.18	0.87
	Nantong	0.65	2.12	0.90
	Shijiazhuang	0.65	2.15	0.87
	Suzhou	0.45	2.31	0.79
	Taiyuan	0.60	2.17	0.87
	Wuhu	0.46	2.24	0.83
	Xi'an	0.49	2.31	0.79
3 Non-Logistic / Irregular				
	Dalian	0.60	2.13	0.89
	Dongguan	1.00	–	0.00
	Fukuoka	0.72	2.08	0.91
	Hohhot	0.65	2.14	0.88
	Jinan	0.43	2.30	0.80
	Jinhua	0.82	2.04	0.94
	Kaohsiung	0.66	2.12	0.90
	Kobe	0.54	2.22	0.83
	Kyoto	0.67	2.11	0.90
	Lanzhou	0.74	2.00	1.00
	Luoyang	0.65	2.12	0.90
	Macau	1.00	–	–
	Qingdao	0.46	2.23	0.85
	Sapporo	1.00	–	–
	Sendai	1.00	–	–
	Shaoxing	1.00	–	–
	Taichung	1.00	–	–
	Taizhou	1.00	–	–
	Taoyuan	1.00	–	–
	Wenzhou	1.00	–	–
	Xuzhou	0.84	2.00	1.00
	Yokohama	1.00	–	–
	Zhengzhou	0.41	2.51	0.72

Continued on next page

Category	City	β	k_{core}	f_2
	Ürümqi	1.00	–	–

2.3 Accessibility performance of East Asian metro networks

In line with chapter 1, accessibility indicators derived from the access graph systematically summarize the performance of each metro network and enable cross-city comparisons at the regional level. With 30 minutes serving as the t_M, representing the typical travel-time budget for passengers and the threshold at which the network becomes generally accessible, G_{30}, G_M, D_{30}, and D_M, together with the heatmap in figure 3, evaluate reachability and equity at critical time points. In addition, given the wide range of network sizes and previous studies reporting both higher travel impedance in large rail-based PTNs[3] and centre–suburb accessibility gaps[18], a spearman correlation matrix is used to examine the interdependence among the accessibility indicators and multiple linear regression is performed to assess the influence of N. This study summarizes the accessibility performance of East Asian metro networks as follows:

First, for metro networks of at least medium size, the accessibility performance in the late stage is heavily constrained by suburban lines, where the quality of service deteriorates compared with the city centre. Most East Asian networks have experienced major extensions since the 2010s. By comparing typical large systems ($N > 100$), such as Seoul, Shanghai, and Beijing, with their layouts in 2012[11], it can be observed that the core defined in chapter 2.2 remains largely unchanged, while the overall network has become more spatially extensive due to the addition of long suburban lines. Both the multiple linear regression and the spearman correlation show a show a strong positive relationship between N and $t_{\max}$ ($R^2 = 0.58$, $\beta_{\log N} = 44.68$; $\rho(N, t_{\max}) \approx 0.80$). The extended tail before saturation, as shown in Figure 3, is a common feature for networks with N>100. By tracing the station pairs that become connected before reaching $t_{\max}$ and referring to the accessibility growth pattern discussed in chapter 2.2, it can be observed that stations located near the terminals of individual lines only become connected at time budgets corresponding to the tail of the accessibility growth curve. Compared with the rest of the station pairs, station pairs connected at the very late stage have a much larger share of the travel budget to be spent on suburban segments characterized by longer in-vehicle times and lower service frequencies and tend to have higher transfer penalty and waiting time, thereby yielding strongly diminishing returns in accessibility. In contrast, micro- and small- sized systems are largely free from this constraint, as their lines mainly serve the urban areas where service quality remains consistent across the entire network.

Second, regardless of network size, a common issue among East Asian metro networks is the early decay of the degree-growth rate and moderate reachability at t_M, which causes passengers to experience greater travel impedance during the medium and late stages of accessibility growth. When size factor is isolated, the δt_M of East Asian metro networks has a median and mean of 0.31 and IQR of 0.22–0.38. By size class, δt_M shows limited between-class variation as suggested by the low correlation between N and δt_M ($R^2 = 0.01$, spearman $\rho(N, \delta t_M) \approx 0.20$). If we construct a normalized time axis indicating the ratio of time budget to the maximum generalized travel time, $\tau = t_b/t_{\max}$, and define three stages of average-degree growth as the initial stage ($0 < \tau < 0.3$), the medium stage ($0.33 < \tau < 0.66$), and the late stage ($0.66 < \tau < 1$), the vast majority of systems experience maximum growth in average degree during the initial stage and the early medium stage. The decay of growth rate in micro and small-sized networks can be attributed to lack of direct connection between terminal areas and more frequent detours in later stage. For larger networks such decay can be attributed to earlier interconnection of core. The median and IQR of $\overline{D}_{30}$ (0.38 & 0.20–0.46) indicate that, at t_M, over half of the stations remain unreachable from any other stations. Consequently, for passengers still travel on networks in medium to late stage, travel impedance greatly increases as a large number of stations remain unconnected and spatially dispersed, and the rate of degree growth over time slows down. By absolute time value, a similar issue also exists. 89% of East Asian metro networks have t_M values below 60 minutes, corresponding to the typical maximum daily travel time that a person can afford[16]. In micro- and small-sized networks, 60 minutes is close to $t_{\max}$ and suggests that a maximum daily trip typically involves long-distance travel between stations in terminal areas. In larger networks, the accessibility pattern indicates that such trips usually correspond to journeys outside the core area. Therefore, long daily trip close to typical maximum time budget are also subject to greater travel impedance in late-stage.

Third, the core–branch structure, which is commonly observed in networks of at least medium size, provides only marginal improvements in reachability at t_M but has an overall negative influence on reachability within 30 minutes and on network equality. As suggested by the relationships between N and $\overline{D}_{30}$ ($R^2 = 0.71$, $\beta \log N = -0.16$; $\rho(N, \overline{D}_{30}) \approx -0.88$), G_{30} ($R^2 = 0.81$, $\beta_{\log N} = 0.08$; $\rho(N, G_{30}) \approx 0.91$), and G_M ($R^2 = 0.58$, $\beta_{\log N} = 0.06$; $\rho(N, G_M) \approx 0.76$), larger systems tend to exhibit declining short-term reachability and increasing inequality in accessibility. The higher level of inequality is also suggested by a more scattered distribution of node degrees around t_M on the figure 3 heatmap. It has been analysed in chapter 2.2 that cities with a typical core–branch structure usually experience faster interconnection among station pairs within the core than among those involving stations on the branches. As the results, stations located on branch and further away from core will have lower degree compared to stations inside the core before the final saturation of average degree growth, resulting in the

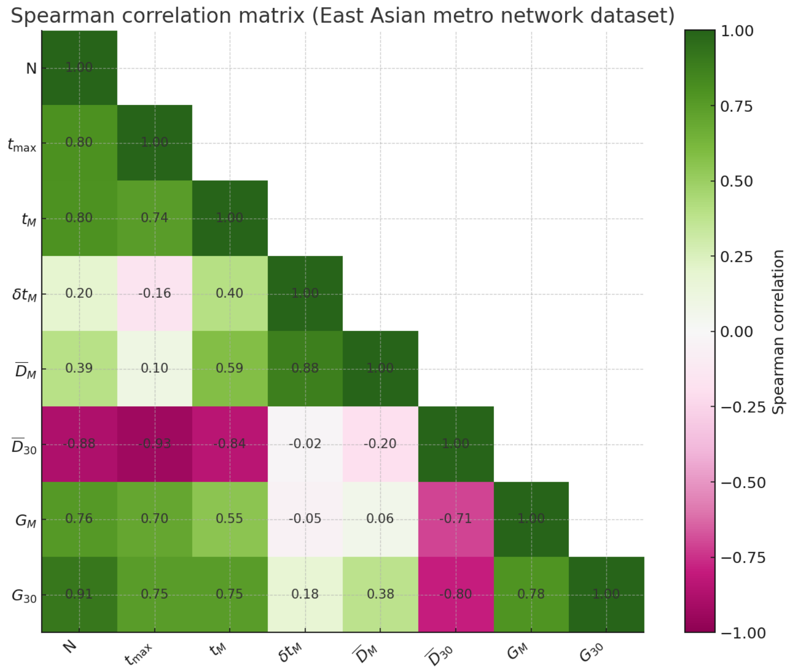


Figure 4. Spearman correlation matrix for accessibility indicators of East Asian metro networks

inequality and reduce the overall reachability. It is noteworthy that inequality, compared to reachability, is less sensitive to further size increases, because in large and mega networks the branch share β converges below 0.55 and the share of stations on branches no longer changes sharply. By contrast, greater spatial disparity in larger networks raises the generalized travel costs for branch stations to connect to the core, and hence reachability is affected more than equity. The creation of a core has only a trivial improvement in reachability at t_M, as suggested by $(R^2 = 0.19,\ \beta \log N = 0.06;\ \rho(N, \overline{D}_M) \approx 0.39)$, because a dense core provides strong connectivity for station pairs between lines. By comparison, micro- and small-sized networks have higher branch fractions and lower structural heterogeneity. Their equity and reachability are less affected by topology and depend more on service attributes. Smaller networks generally operate with limited spatial coverage and more uniform service quality, and therefore they are more reachable at low time budgets and exhibit greater equity.

In sum, East Asian metro systems on general underperform in the medium to late stage of accessibility growth as average degree growth decays early and reachability remains modest at t_M. Reaching unconnected stations after t_M will introduce greater travel impedance due to decreasing growth rate of average degree and higher generalized travel cost, and resulting in poorer overall accessibility performance. Micro- and small-sized networks have greater structural homogeneity, more uniform service quality, and limited spatial coverage. Their accessibility performance is therefore primarily determined by service attributes rather than topology, resulting in higher equality and reachability at 30 minutes compared to networks that are at least medium-sized. Conversely, the form of core–branch structure in medium- to mega-sized networks diminishes overall equality and short-term reachability, offering only marginal improvements in D_M due to higher cross-line connectivity. Additionally, for larger networks, the accessibility performance at late stage further declines as service quality on suburban extensions deteriorates.

2.4 Cross-Region comparison of accessibility performance

In 2.1 the result of accessibility indicator of and performance of East Asia metro networks on region level is discussed. However, the region's performance on accessibility relative to other regions remains unclear. For comparison, the study defines two other regions as North America and Europe. These regions are chosen because the historical development processes of their metro networks substantially different from those in East Asia[11]. Additionally, European public transport networks (PTNs) are found to lead in accessibility, while North American systems lag behind[2]. As recent network development in both regions has largely levelled off[12], their accessibility performance can be considered relatively stable and thus serve as a useful benchmark for East Asian networks. Following the regional split of the full dataset, the number of cities in each regions are 13 for North America, 62 for East Asia, and 31 for Europe. Sample-size considerations also guide this choice. The 2022 dataset includes a small number of systems from other world regions, including South America, South Asia, Middle East and Africa[26,27], but these regions are excluded since their small sample size could not support region-level generalization.

Table 3. Summary statistics (median, mean, and IQR) of key timing, connectivity, and equality indicators by network-size category (62 metro systems). Indicators showing weak correlation with N are excluded. Category groupings (1–3) follow the corrected classification in Table 2.

Category	t_{max} (min)			t_M (min)		
	Median	Mean	IQR (25–75%)	Median	Mean	IQR (25–75%)
Micro ($1 \leq N \leq 50$)	76	77	60–87	17	20	11–26
Small ($51 \leq N \leq 100$)	110	114	78–120	43	41	39–45
Medium ($101 \leq N \leq 200$)	115	124	104–126	41	41	38–42
Large ($201 \leq N \leq 300$)	174	182	158–210	47	47	44–50
Mega ($N > 300$)	182	183	166–182	51	52	51–51
All cities (62)	108	118	78–145	39	35	18–45

Category	$\overline{D}_{30}$			$\overline{D}_M$		
	Median	Mean	IQR (25–75%)	Median	Mean	IQR (25–75%)
Micro ($1 \leq N \leq 50$)	0.42	0.44	0.36–0.51	0.12	0.22	0.07–0.29
Small ($51 \leq N \leq 100$)	0.25	0.26	0.23–0.28	0.48	0.51	0.43–0.54
Medium ($101 \leq N \leq 200$)	0.18	0.20	0.16–0.22	0.44	0.42	0.39–0.48
Large ($201 \leq N \leq 300$)	0.11	0.12	0.10–0.14	0.37	0.38	0.34–0.42
Mega ($N > 300$)	0.10	0.10	0.07–0.11	0.37	0.38	0.36–0.40
All cities (62)	0.24	0.28	0.14–0.37	0.38	0.35	0.20–0.46

Category	G_{30}			G_M		
	Median	Mean	IQR (25–75%)	Median	Mean	IQR (25–75%)
Micro ($1 \leq N \leq 50$)	0.20	0.20	0.16–0.24	0.16	0.17	0.14–0.19
Small ($51 \leq N \leq 100$)	0.25	0.25	0.21–0.28	0.21	0.20	0.16–0.24
Medium ($101 \leq N \leq 200$)	0.30	0.29	0.27–0.31	0.25	0.26	0.25–0.28
Large ($201 \leq N \leq 300$)	0.33	0.33	0.30–0.35	0.30	0.30	0.28–0.32
Mega ($N > 300$)	0.38	0.37	0.37–0.40	0.32	0.32	0.31–0.34
All cities (62)	0.22	0.24	0.17–0.30	0.23	0.22	0.16–0.28

Table 4. Multiple linear models on $\log N$ for East Asian metro networks and the values of R^2, $\beta_{\log N}$, and $p_{\log N}$ (2 d.p.)

KPI	R^2	$\beta_{\log N}$	$p_{\log N}$
t_{max}	0.58	44.68	0.00
t_M	0.67	13.20	0.00
$\overline{D}_{30}$	0.71	−0.16	0.00
$\overline{D}_M$	0.19	0.09	0.00
G_{30}	0.81	0.08	0.00
G_M	0.58	0.06	0.00
δt_M	0.01	0.02	0.25

Table 5 compares the core accessibility indicators across regions. Overall, East Asia exhibits the lowest reachability and reaches its maximum degree growth latest on the absolute time axis. When normalized by network size, however, it shows the earliest decline in degree growth, with the peak occurring in the initial stage of accessibility growth. In terms of equity, East Asia records the highest inequality at t_M and the second-highest at 30 minutes, though cross-regional differences at t_M remain marginal. In line with Wu et al.[2], compact regional urban form and well-connected transport networks are associated with higher accessibility outcomes. European metro PTNs exemplify this pattern, achieving high levels of accessibility through dense metro networks primarily serving the urban areas. Meanwhile, as discussed in chapter 2.3, East Asian networks tend to introduce spatial and service disparities as they expand into suburban areas. Such outward expansion has consequently resulted in their accessibility gap relative to other regions.

Network size N systematically affects accessibility and large systems are more common in East Asia as reflected by higher station counts (mean $N = 120.26$, median 91) than in North America and Europe. Accordingly, multiple linear regression models are estimated for each accessibility indicator as a function of $\log N$, with regional dummy variables included and East Asia serving as the baseline category. Since allowing region-specific slopes adds only 1–4% to the explained variance and the joint interaction tests are not significant, the results in table 6 indicate a similar influence of size on accessibility across regions. Importantly, regional difference persist after controlling for size. At a given network size (N), Europe achieves higher reachability at both short time budget and maximum degree growth time point ($\overline{D}_{30}$: $\beta_{\text{Europe}} = +0.26$, $\overline{D}_M$: $\beta_{\text{Europe}} = +0.15$) and lower generalized travel-time threshold when networks become generally accessible (t_M: $\beta_{\text{Europe}} = -4.95$, $p = 0.02$) than East Asia, again reflecting the positive influence of compact urban form on accessibility. In contrast, North America displays higher t_M ($\beta_{\text{NA}} = +9.17$). This is consistent with the effect of sprawling development on North American PTNs[2]. However, since North America networks experience less expansion and have less spatial disparity, their reachability at t_M ($\overline{D}_M$: $\beta_{\text{NA}} = +0.16$) is greater than East Asian networks. Both Europe ($\beta_{\text{Europe}} = +0.08$, $p = 0.01$) and North America ($\beta_{\text{NA}} = +0.12$, $p < 0.001$) exhibit significantly higher δt_M than East Asia, indicating that their networks can sustain a longer period of degree growth, and passengers travelling during the medium and late stages of average-degree growth experience lower travel impedance.

In sum, whether the size effect is controlled for or not, East Asian metro networks, on average, underperform European systems in reachability and in reducing travel impedance during the late stage of network evolution. To a lesser extent, in terms of reachability, North American networks outperform East Asian systems of comparable size in accessibility at t_M ($\overline{D}_M$) and exhibit higher t_M as well as a later decay in the accessibility growth rate on normalized time axis. The three regions exhibit broadly similar levels of accessibility equity, which appears to be more sensitive to network size than to geographical differences.

Table 5. Per-region *Mean* and *Median* — Core indicators (2 d.p.)

Region	**Statistic**	**N**	$t_{\max}$	t_M	δt_M	$\overline{D}_M$	$\overline{D}_{30}$	G_M	G_{30}
North America	Mean	83	95	38.85	0.42	0.47	0.32	0.21	0.26
North America	Median	52	94	39.00	0.40	0.49	0.36	0.21	0.25
East Asia	Mean	120	118	34.61	0.31	0.35	0.28	0.22	0.24
East Asia	Median	91	108	39.00	0.31	0.38	0.24	0.23	0.22
Europe	Mean	77	69	25.06	0.39	0.47	0.60	0.21	0.18
Europe	Median	50	58	25.00	0.36	0.48	0.66	0.20	0.16

Table 6. Multiple linear regression models on $\log N$ with accessibility indicators for different geographical regions. p-values (2 d.p.)

Accessibility indicators	R^2	$\beta_{\log N}$	$p_{\log N}$	β_{Europe}	p_{Europe}	β_{NA}	p_{NA}
t_M	0.64	11.81	0.00	−4.95	0.02	9.17	0.00
$\overline{D}_{30}$	0.72	−0.16	0.00	0.26	0.00	−0.02	0.56
$\overline{D}_M$	0.23	0.07	0.00	0.15	0.00	0.16	0.00
G_{30}	0.73	0.09	0.00	−0.02	0.15	0.06	0.00
G_M	0.41	0.05	0.00	0.01	0.49	0.01	0.73
δt_M	0.11	0.01	0.45	0.08	0.01	0.12	0.00

2.5 Accessibility growth-rate profile clustering for cross-regional characterisation

In table 6 it can be seen that geographical regions does not always explain the difference in accessibility performance, as suggested by some high p-values. Prior studies shows that metro network structure is not random but follows recurring regimes (e.g.,core–branch) with measurable functional implications[8–10]. However, as Luo et al.[3] and the analysis in chapter 2.2 indicate, accessibility cannot be fully understood from topology alone. When service attributes such as frequency, waiting, and transfer times are incorporated, stronger spatial disparities emerge as central zones benefit from dense and frequent services, while peripheral areas experience higher travel impedance and weaker connectivity. To meaningfully characterize East Asian metros alongside systems from other regions, this study further classifies cities by the temporal shape of accessibility growth curve, which dynamically reflects travel impedance from the passengers' perspective. As stated in Chapter 2.3, a declining growth rate indicates rising travel impedance as reaching the remaining unconnected stations becomes progressively more time-consuming. When this decay occurs early, overall accessibility deteriorates, because late-stage accessibility is heavily constrained. Additionally, since the clustering is not constrained by geographical regions, it encompasses all East Asian, North American and European networks as well as networks that are previously excluded in chapter 2.4 due to insufficient sample size[26,27].

For each city i, the normalized average degree $\overline{D}_i(t_b)/(N_i - 1)$ is computed on a 2-min grid, interpolate to a normalized time axis $\tau \in [0,1]$ previously defined in chapter 2.3. The discrete growth rate, defined as $G_{\tau,i} = \Delta\overline{D}/\Delta\tau$, is used to construct a Pearson-correlation similarity matrix. The $G_{\tau,i}$ on normalized time axis for each city is shown in figure . Subsequently, spherical k-means clustering is applied, and based on the cosine-silhouette criterion, four distinct clusters are identified, and their detailed composition is reported in table 7. The growth-rate characteristics and common features of each cluster are as follows:

(1) Cluster 1 comprises networks with an early peak and a mid-stage decline in $G(\tau)$; their topology is mainly characterized by large systems with grid-like cores (e.g., Tokyo, Seoul, Paris) as well as networks with simple cores and radiating branches (e.g., Chicago, Rome).

(2) Cluster 2 includes networks with the fastest initial-stage rise and the earliest decline in $G(\tau)$; most early East Asian metro networks with a core–branch structure fall into this cluster, along with some networks expanding over long corridors (e.g., Dubai, Helsinki).

(3) Cluster 3 consists of networks exhibiting a balanced growth rate across the initial and medium stages; these are mainly single-line or cross-shaped systems with compact and simple topologies.

(4) Cluster 4 contains networks characterized by a late peak and subsequent decline in $G(\tau)$, typically featuring simple cores in quasi-circular shapes (e.g., Nanning), radiating branches (e.g., Stockholm, Washington), or other compact topologies (e.g., Warsaw, Taiyuan).

Consistent with the typology and service attributes, C3 leads on reachability at 30 minutes and equity (high $\overline{D}_{30}$, low G_{30}/G_M), C4 attains the highest reachability at t_M(high $\overline{D}_M$), C1 is intermediate, and C2 underperforms at short horizons with higher inequality. In the worldwide context, large East Asian systems predominately fall in C2 and are subject to an early decay in the growth of accessibility, similar to networks dominated by long corridors (e.g., Dubai, Helsinki) or those lacking direct cross-corridor links (e.g., Toronto), which struggle to achieve high short-horizon reachability. The core–branch structure also tends to raise inequality. However, large-sized networks with grid-like cores and higher, more homogeneous frequencies can partially avoid early decay as Tokyo, and Seoul fall into C1 and show patterns similar to Paris, which has been found to perform well in accessibility[13]. With the exception of Qingdao (whose lines are mainly suburban lines with long distance), medium-sized systems are generally less spatially dispersed, remain centred on the urban core, and often fall into C1. If service attributes deteriorate less across branches, some medium systems achieve stronger outcomes, as illustrated by Hefei and Changsha in C4. By contrast, smaller networks can benefit more from homogeneous frequencies across the network due to their compact size and often fall into C3 and C4, which perform well overall. This is consistent with Wu et al.[2], concentrating development within a compact urban core improves accessibility outcomes. However, limited size is not a guarantee of performance as topological deficiencies like long-corridor lines without transfer (e.g., Yokohama, Jinhua) or low-frequency branch operations (e.g., Urumqi) can still undermine accessibility, even in smaller networks.

The clustering results also show possible method to improve accessibility performance. As indicated by metro networks of Seoul and Tokyo, large networks that are C2-type can delay early saturation requires reducing suburban impedances, which can be achieved by shortening in-vehicle times on outer segments via reduced station dwell times and introducing an express overlay service, more even headways across branches. Infrastructure wise, grid-like core structure can also potentially increase the average degree growth period. Improvements in service attributes can also enhance the accessibility performance of medium-sized networks. Consequently, medium-sized networks of C1 type could draw lessons from those of C4 type. For

Table 7. Cluster groups of world metro networks ($k = 4$, silhouette(cosine)=0.4550). Networks located in East Asia are in **bold**.

Cluster	Cities
1	**Busan**, **Changchun**, Chicago, **Foshan**, **Fuzhou**, **Guiyang**, **Kunming**, Milan, **Nanchang**, **Nantong**, **Ningbo**, Paris, Rennes, Rome, Rotterdam, San Francisco, **Seoul**, **Shenyang**, **Shijiazhuang**, **Taipei**, **Tokyo**, Vancouver, Vienna, **Xiamen**
2	Athens, **Beijing**, Berlin, Bilbao, **Chengdu**, **Chongqing**, **Dalian**, Dubai, **Guangzhou**, **Hangzhou**, Helsinki, **Hong Kong**, **Jinan**, **Jinhua**, Lille, London, Madrid, **Nanjing**, Naples, New York, **Qingdao**, **Shanghai**, **Shenzhen**, **Suzhou**, **Tianjin**, Toronto, Turin, Valencia, **Wuhan**, **Xi'an**, **Yokohama**, **Zhengzhou**, **Ürümqi**
3	Atlanta, Baltimore, Budapest, **Changzhou**, Cleveland, Copenhagen, **Dongguan**, **Fukuoka**, Genoa, **Hohhot**, Hyderabad, **Kobe**, **Kochi**, **Kyoto**, **Lanzhou**, Los Angeles, **Luoyang**, **Macau**, Malaga, Marseille, Oslo, Philadelphia, Prague, **Sapporo**, **Sendai**, **Shaoxing**, **Taichung**, **Taizhou**, **Taoyuan**, **Wuhu**, **Xuzhou**
4	Amsterdam, Boston, Brussels, Buenos Aires, Cairo, **Changsha**, **Daegu**, **Harbin**, **Hefei**, **Incheon**, **Kaohsiung**, Lisbon, Lyon, Montreal, **Nanning**, Nuremberg, Santiago, Stockholm, **Taiyuan**, Toulouse, Warsaw, Washington, **Wenzhou**, **Wuxi**

Table 8. the mean and median of accessbility indicators for different clusters (2 d.p.)

Cluster	Statistic	N	$t_{\max}$	t_M	δt_M	$\overline{D}_M$	$\overline{D}_{30}$	G_M	G_{30}
1	Mean	113.33	105.67	37.92	0.36	0.44	0.30	0.24	0.25
1	Median	100.50	110.00	39.00	0.36	0.45	0.25	0.24	0.25
2	Mean	176.27	146.00	36.21	0.25	0.35	0.28	0.26	0.28
2	Median	177.00	140.00	41.00	0.26	0.37	0.23	0.27	0.30
3	Mean	33.71	62.52	22.81	0.38	0.34	0.52	0.17	0.15
3	Median	30.50	62.00	21.00	0.32	0.26	0.47	0.16	0.15
4	Mean	73.25	75.83	34.25	0.46	0.54	0.44	0.20	0.21
4	Median	67.50	77.00	34.00	0.46	0.55	0.38	0.21	0.20

micro- and small-sized systems, maintaining homogeneous service levels helps preserve the advantages of low spatial disparity inherent to simple network topologies.

3 Discussion

This study adopts a unified graph-based framework to first quantify the topological accessibility of 62 East Asian metro networks. By constructing time- and frequency-weighted access graphs and deriving indicators of reachability, and equity at critical time points, the analysis provides a comprehensive understanding of how topology and service jointly shape accessibility performance.

First, East Asian metros display a high degree of heterogeneity in both size and structural topology. Most large networks have core-branch structures and generally follow a logistic (S-shaped) pattern of accessibility growth consistent, whereas smaller networks without core-branch difference and networks with degraded service attributes deviate from this trend. This finding indicates that service attributes can magnify the effect of network topology on accessibility development.

Second, despite their scale and rapid spatial expansion in 2010s, East Asian metros networks generally tend to reach saturation of degree increase at initial stage of accessibility growth. The early decay in degree-growth rate suggests that newly added suburban branches in large networks hinder the integration of network structure and service, resulting in higher travel impedance and lower reachability at later stages. Large systems also show greater spatial disparity and less uniform service quality, which results in weaker short-term reachability and equity compared with smaller or more compact systems.

Third, when network size is controlled for, regional differences persist. While their equity levels are comparable, East Asian networks are less reachable than European metros networks at critical time points. Lastly, clustering analysis reveals four recurring accessibility growth pattern that transcend geographical regions. Large East Asian networks tend to experience early

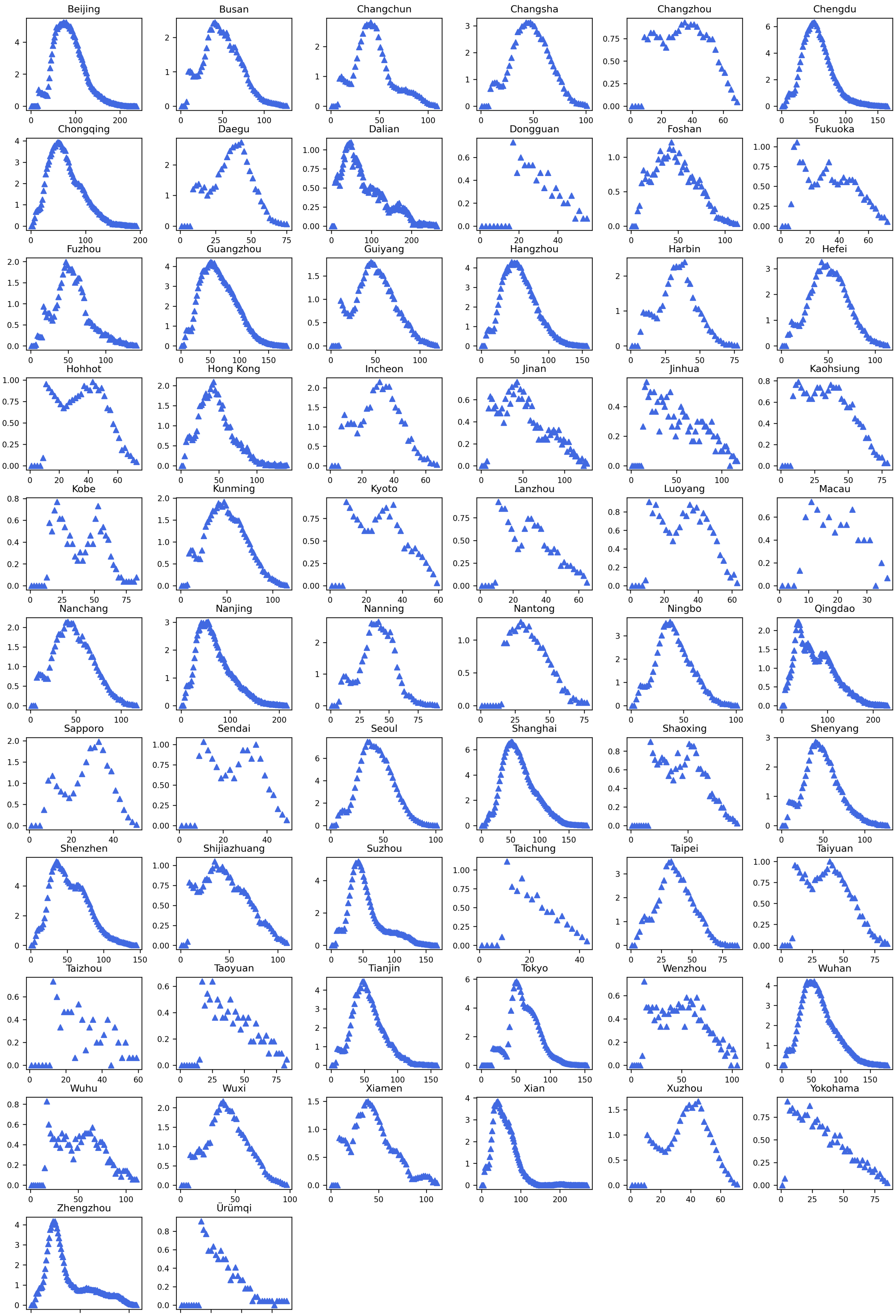


Figure 5. Rate of growth (first difference quotient) of the average node degree for East Asian networks. Degree growth rate graph of other worldwide metro networks can be found at Sfiligoj et al.[13]

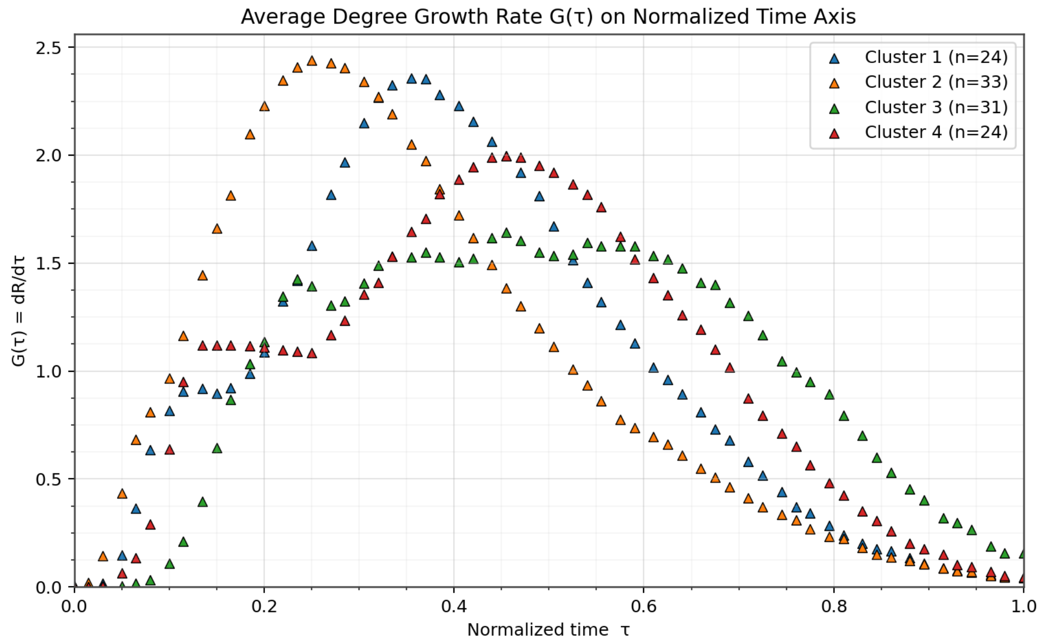


Figure 6. Accessibility growth-rate for 4 clusterson normalized time axis. Let $\tau = t_b/t_{\max}$ and $G(\tau_k) = \dfrac{d\overline{D}}{d\tau} = r_\tau(\tau_k) = \dfrac{\overline{D}(\tau_{k+1}) - \overline{D}(\tau_k)}{\Delta\tau}$, plotted as $y = G(\tau_k)$ versus $x = \tau_k$ ($k = 4$ clusters).).

decline in accessibility growth, similar to systems lacking direct inter-line connections, whereas medium and small networks sustain growth longer, on par with those in other regions. The coexistence of networks of similar size in different clusters highlights the role of service attributes in shaping accessibility.

In sum, the findings demonstrate that expanding network length alone does not ensure equitable or efficient accessibility. For future expansion of metro networks in East Asia regions as well as other geographical regions, strengthening cross-line integration, harmonising frequencies, and prioritising service quality in suburban corridors can deliver more balanced accessibility gains.

Although this study compiles comprehensive datasets of East Asian metro networks, several limitations remain. First, for Chinese cities, data were mainly obtained from local map services and calibrated using first and last train timetables. While this approach allows coverage across numerous systems, it is subject to greater temporal and spatial inaccuracy compared with GTFS datasets directly provided by operators. Second, due to data availability, some through services and express operations (for example, through-running services across multiple lines in Chongqing) are not represented in the current dataset. These services can significantly affect generalized travel times and network connectivity, and their omission may lead to a conservative estimate of accessibility performance. Third, the analysis excludes commuter rail following strict definition of metro systems. However, some East Asian cities like Tokyo provide cross-regional corridors that complement metro services. Consequently, the reported metro-only accessibility levels are likely to be the lower bound for such coordinated multi rail-bound PTNs[28]. Finally, the framework focuses on structural and service-based accessibility under normal conditions. Integrating demand, crowding, and timetable-driven dynamics following rail-bound PTN studies[4] would further enhance the accuracy of future accessibility assessments.

4 Method & Data

4.1 Data

For maximal coverage for East Asia region, the dataset ultimately manages to include 61 East Asian metro systems. Data sources differ by city group. Data for Seoul, Busan, Daegu, Incheon are primarily from `https://www.ktdb.go.kr` ; Taipei, Taoyuan, Taichung, Kaohsiung from `https://tdx.transportdata.tw/` ; and Tokyo, Yokohama, Kyoto, Kobe, Sapporo, Sendai, Fukuoka from `https://www.odpt.org/`. The remaining 47 systems use `https://lbs.amap.com/` as the primary source. While GTFS feeds are unavailable, station coordinates, service frequencies, and in-vehicle travel times are provided on afore-mentioned platform. Therefore, weighted L-space and P-space networks can be constructed. Due to the absence of obtainable open data, metro networks of Pyongyang, Daejeon, Gimpo, Gwangju, Osaka and Nagoya are excluded. The data for Kobe and metro networks in other regions is from a 2022 dataset[26,27].

4.2 Construction of access graph

The access graph is based on the L-space and P-space representation of the metro networks. The generalised travel time is computed with in-vehicle and waiting times, together with transfer costs. In the access graph G_A, the nodes denote the stations and have the same meaning as in L-space and P-space representation. The edge set of the access graph is generated based on the following rules[13]:

1. The L-space and P-space are first weighted by service frequencies f_{ij} (equals number of rides per hour) and in-vehicle time t^L_{ij}, respectively.
2. The edge weights of P-space is transformed into average waiting time t^{wait} in minutes, which equals as half of the headway $t^{\text{wait}}_{ij} = 30/f_{ij}$, with the assumption that passengers arrive randomly.
3. The in-vehicle time weights $t^{\text{in-veh}}_{ij}$ are assigned to P-space edges from the corresponding consecutive paths in L-space.
4. The distance matrix $D^{P,\text{wait}}$ is built based upon the waiting-time-weighted P-space representation.
5. The distance matrix of in-vehicle times $D^{P,\text{in-veh}}$ is obtaned by the in-vehicle time weights of the paths obtained in step 3 .
6. The distance matrix $D^{P,u}$ is determined by the corresponding path lengths in *unweighted* P-space. In $D^{P,u}$, $\left[D^{P,u}\right]_{ij} - 1$ equals the number of transfers along the path between i and j.
7. Construct the generalised travel-time matrix $D = D^{L,\text{in-veh}} + w_{\text{wait}} D^{P,\text{wait}} + w_{\text{transfer}}\left(D^{P,u} - J_N\right)$, where $D^{L,\text{in-veh}}$ is the in-vehicle time matrix in L-space, $D^{P,\text{wait}}$ is the waiting-time matrix in P-space, $D^{P,u}$ is the unpenalized transfer matrix in P-space, J_N is the all-ones matrix, and $w_{\text{wait}}, w_{\text{transfer}}$ are weights. The values of $w_{\text{wait}} = 2$ and $w_{\text{transfer}} = 5$ are determined from the literature and reflect the perception of time by typical passengers[29].

For any stop pair connecting stops i and j, entry $d_{ij} = [D]_{ij}$ of the generalised travel time matrix represents the generalised travel time between stops i and j. In the next step, the time budget t_b range between 0 and t_{max} is introduced. A binary of x_{ij} could represent the existence of edges between stops i and j in access graph and its value is obtained based on:

$$x_{ij} = \begin{cases} 1, & d_{ij} \le t_b, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Consequently, the access graph is time-budget dependent. For a given budget t_b, an undirected, unweighted edge (i, j) if j is reachable from i along a shortest path whose generalized travel time is $\le t_b$. Although derived from standard network representations via their travel-time weights, edges in the access graph encode reachability within t_b rather than physical adjacency or line co-membership. Accordingly, the node degree $D_i(t_b)$ equals the number of stops reachable from i within t_b, serving as a stop-level accessibility metric that is equivalent to a gravity measure.

With the calculated accessibility indicators,the natural logarithm of system size is used for the region-wide comparison

$$\text{KPI}_i = \alpha + \beta_{\ln N} \ln(N_i) + \varepsilon_i,$$

where $\ln(\cdot)$ denotes the natural logarithm (base e). Coefficients therefore represent changes in the indicator for proportional changes in the number of stations; for example, a doubling of network size corresponds to $\Delta \ln N = \ln(2) \approx 0.693$.

4.3 K-means clustering

Across all cities, the access-graph average degree exhibits logistic-like growth with the time budget and converges to its maximum. For city i, let $R_i(t)$ denote the normalized average degree at budget t, obtained by dividing degrees by $N_i - 1$. To enable comparison across networks with different sizes and maximum generalized travel times $t_{\max,i}$, time is reparameterized to the unit interval via $\tau = t/t_{\max,i} \in [0,1]$, and $R_i(\tau)$ is evaluated on a common grid $\{\tau_m\}_{m=1}^{M}$.

To capture the speed at which accessibility develops, this study approximates the growth-rate profile $G_i(\tau)$ by first differences of $R_i(\tau_m)$ and forms the vector $\mathbf{x}_i = (G_i(\tau_1), \dots, G_i(\tau_M))$. For each city pair (i, j), the Pearson correlation $\rho_{ij} = \text{corr}(\mathbf{x}_i, \mathbf{x}_j)$ is computed to obtain the similarity matrix $S = (\rho_{ij})$. Following standard practice in correlation-based clustering of temporal profiles, a Fisher z-transform, imputation for undefined entries, column-wise standardization, and ℓ_2-normalization of rows are applied to construct cosine-normalized feature vectors $\mathbf{r}_i$.

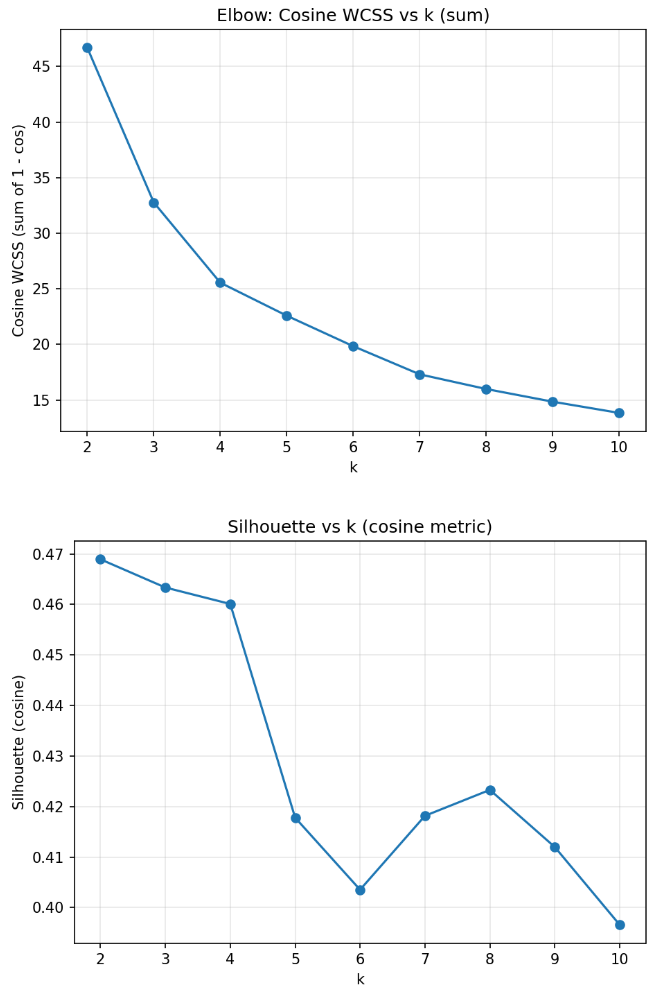


Figure 7. Cluster number selection using elbow and silhouette criteria on cosine-normalized features.

Spherical k-means clustering is chosen because it optimizes cosine similarity on the unit sphere, thereby capturing directional similarity more effectively than standard k-means[30]. It is then performed on $\{\mathbf{r}_i\}$ by minimizing within-cluster cosine dispersion.The number of clusters K is selected by combining elbow and cosine-silhouette criteria on the within-cluster sum of squares and pairwise distances (Figure 7). Both diagnostics indicate $K = 4$ as the largest value that maintains strong cohesion and separation while yielding balanced and interpretable groups.

References

1. Hansen, W. G. How accessibility shapes land use. *J. Am. Inst. Planners* **25**, 73–76, DOI: https://doi.org/10.1080/01944365908978307 (1959).

2. Wu, H. *et al.* Urban access across the globe: an international comparison of different transport modes. *npj Urban Sustain.* **1**, 16, DOI: https://doi.org/10.1038/s42949-021-00020-2 (2021).

3. Luo, D., Cats, O., van Lint, H. & Currie, G. Integrating network science and public transport accessibility analysis for comparative assessment. *J. Transp. Geogr.* **80**, 102505, DOI: https://doi.org/10.1016/j.jtrangeo.2019.102505 (2019).

4. Luo, D., Cats, O. & van Lint, H. Can passenger flow distribution be estimated solely based on network properties in public transport systems? *Transportation* **47**, 2757–2776, DOI: https://doi.org/10.1007/s11116-019-09990-w (2020).

5. Jin, F., Jiao, J., Qi, Y. & Wang, C. Evolution and geographic effects of high-speed rail in east asia: An accessibility approach. *J. Geogr. Sci.* **27**, 515–532, DOI: https://doi.org/10.1007/s11442-017-1390-8 (2017).

6. Lin, J. & Ban, Y. Complex network topology of transportation systems. *Transp. Rev.* **33**, 658–685, DOI: https://doi.org/10.1080/01441647.2013.848955 (2013).

7. von Ferber, C., Holovatch, T., Holovatch, Y. & Palchykov, V. Public transport networks: empirical analysis and modeling. *The Eur. Phys. J. B* **68**, 261–275, DOI: https://doi.org/10.1140/epjb/e2009-00090-x (2009).

8. Derrible, S. & Kennedy, C. Network analysis of world subway systems using updated graph theory. *Transp. Res. Rec. J. Transp. Res. Board* **2112**, 17–25, DOI: https://doi.org/10.3141/2112-03 (2009).

9. Derrible, S. & Kennedy, C. The complexity and robustness of metro networks. *Phys. A: Stat. Mech. its Appl.* **389**, 3678–3691, DOI: https://doi.org/10.1016/j.physa.2010.04.008 (2010).

10. Derrible, S. & Kennedy, C. Characterizing metro networks: state, form, and structure. *Transportation* **37**, 275–297, DOI: https://doi.org/10.1007/s11116-009-9227-7 (2010).

11. Roth, C., Kang, S. M., Batty, M. & Barthelemy, M. A long-time limit for world subway networks. *J. The Royal Soc. Interface* **9**, 2540–2550, DOI: https://doi.org/10.1098/rsif.2012.0259 (2012).

12. International Association of Public Transport (UITP). GLOBAL METRO FIGURES 2024. *www.uitp.org* https://www.uitp.org/wp-content/uploads/sites/7/2025/08/20250822_Global-Metro-Figures_Statistics-Brief_WEB.pdf (2025).

13. Šfiligoj, T., Peperko, A. & Cats, O. Access graph: a novel graph representation of public transport networks for accessibility analysis. *arXiv* https://doi.org/10.48550/arXiv.2507.08361 (2025).

14. Pei, A., Xiao, F., Yu, S. *et al.* Efficiency in the evolution of metro networks. *Sci. Reports* **12**, 8326, DOI: https://doi.org/10.1038/s41598-022-12053-3 (2022).

15. Mokhtarian, P. L. & Chen, C. Ttb or not ttb, that is the question: A review and analysis of the empirical literature on travel time (and money) budgets. *Transp. Res. Part A: Policy Pract.* **38**, 643–675, DOI: https://doi.org/10.1016/j.tra.2003.12.004 (2004).

16. Ahmed, A. & Stopher, P. Seventy minutes plus or minus 10 — a review of travel time budget studies. *Transp. Rev.* **34**, 607–625, DOI: https://doi.org/10.1080/01441647.2014.946460 (2014).

17. Liang, Y., Wang, D., Zhou, X., Hao, J. & Guo, Y. Assessing the impact of network and station accessibility on station-level rail transit ridership during peak and off-peak hours. *Transp. Res. Part A: Policy Pract.* **199**, 104574, DOI: https://doi.org/10.1016/j.tra.2025.104574 (2025).

18. OECD/International Transport Forum (ITF). Accessibility in the Seoul Metropolitan Area: Does Transport Serve All Equally? *OECD iLibrary* https://doi.org/10.1787/4ef85abc-en (2023).

19. Gori, S., Mannini, L. & Petrelli, M. Equity measures for the identification of public transport needs. *Case Stud. on Transp. Policy* **8**, 745–757, DOI: https://doi.org/10.1016/j.cstp.2020.06.006 (2020).

20. Chang, Z. & Phang, S.-Y. Urban rail transit ppps: Lessons from east asian cities. *Transp. Res. Part A: Policy Pract.* **105**, 106–122, DOI: https://doi.org/10.1016/j.tra.2017.08.015 (2017).

21. Vijlbrief, S. Including service information in a topological comparison of metro networks worldwide: A comparison of 51 metro networks worldwide using gtfs static data. *TU Delft Repository* https://resolver.tudelft.nl/uuid:5197db6d-1b39-413d-bf24-2ee13639e590 (2022).

22. Satomura, H., Yasutake, Y., Uchibayashi, T., Sueyoshi, C. & Inenaga, K. A support tool for converting community bus operation data to GTFS format. In *2024 International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC)*, 1–6, DOI: https://doi.org/10.1109/ITC-CSCC62988.2024.10628417 (Okinawa, Japan, 2024).

23. Cats, O. & Cheng, H. A curated data set of l-space representations for 61 east asian metro networks. version 1. *4TU.ResearchData* https://doi.org/10.4121/d665ec96-dd8b-4e6e-918d-d062a4e1ad42.v1 (2025).

24. Cheng, H. & Cats, O. A curated data set of p-space representations for 61 east asian metro networks. version 1. *4TU.ResearchData* https://doi.org/10.4121/ca8682da-5ab3-4bd4-af37-95517f20a0a5.v1 (2025).

25. Lin, D., Nelson, J. D., Beecroft, M. & Cui, J. An overview of recent developments in china's metro systems. *Tunn. Undergr. Space Technol.* **111**, 103783, DOI: https://doi.org/10.1016/j.tust.2020.103783 (2021).

26. Vijlbrief, S., Cats, O., Krishnakumari, P., van Cranenburgh, S. & Massobrio, R. A curated data set of l-space representations for 51 metro networks worldwide. *4TU.ResearchData* https://doi.org/10.4121/21316824.v1 (2022).

27. Vijlbrief, S., Cats, O., Krishnakumari, P., van Cranenburgh, S. & Massobrio, R. A curated data set of p-space representations for 51 metro networks worldwide. *4TU.ResearchData* https://doi.org/10.4121/21316950.v2 (2022).

28. Cao, Z., Asakura, Y. & Tan, Z. Coordination between node, place, and ridership: Comparing three transit operators in tokyo. *Transp. Res. Part D: Transp. Environ.* **87**, 102518, DOI: https://doi.org/10.1016/j.trd.2020.102518 (2020).

29. Yap, M., Wong, H. & Cats, O. Passenger valuation of interchanges in urban public transport. *J. Public Transp.* **26**, 100089, DOI: https://doi.org/10.1016/j.jpubtr.2024.100089 (2024).

30. Dhillon, I. S. & Modha, D. S. Concept decompositions for large sparse text data using clustering. *Mach. Learn.* **42**, 143–175, DOI: https://doi.org/10.1023/A:1007612920971 (2001).

Reflection report on research internship at Smart Public Transport Lab

Hanyu Cheng*

[1]Delft University of Technology, department of Transport & Planning at the faculty of Civil Engineering and Geosciences, Delft, the Netherlands

ABSTRACT

This report reflects on the extent to which my personal learning objectives were achieved during the research internship at the Smart Public Transport Lab, focusing on independent research, collaboration, and method development. It also discusses my contribution within the lab and identifies areas for further improvement. Finally, it reflects on the ethical and societal implications of accessibility research in public transport, emphasizing equity and inclusion.

Reflection Report

Prior to the research internship, I had already developed a strong and long-standing interest in rail-bound public transport and completed the course CIEQ6232 Public Transport Demand and Network Planning and Operation, taught by Prof. Oded Cats, one of the directors of the Smart Public Transport Lab (SPTL). The lab's mission is to advance the understanding of passenger transport systems through data science, simulation, and optimisation models that improve network performance, accessibility, and sustainability[1]. The course itself, along with the book *Light Rail Transit Systems: 61 Lessons in Sustainable Urban Development*[2] written by another director of the Smart Public Transport Lab, particularly interested me, as both offer perspectives on global rail transit systems including light rail, trams, and metros and I began to notice regional differences while reflecting on general guidance for future scheduled rail-bound systems.

During the course CIEQ6232, I worked with GTFS data of metro networks and identified the accessibility limitation of the Toronto metro networks based on the passenger flow distribution and travel times. Specifically for Toronto network. the limitations in accessibility arise from the absence of direct connections between certain stations. Concurrently, I observed that East Asian metro systems were largely missing from previous dataset, as many operators in the Asia-Pacific region do not provide publicly available GTFS feeds[3,4]. Drawing on my travel experience on systems such as the Chongqing Loop Line and Chengdu Line 4, where station-to-station accessibility declines noticeably once metro trains leave the underground sections in compact urban core and transition to suburban elevated sections, I began to question whether the large-scale and ambitious network expansions in East Asia metro systems during the 2010s[5]genuinely deliver improvement in accessibility compared with North American metro networks like Toronto. After completing the course, I approached Prof. Oded Cats, and we discussed to have a systematic accessibility analysis of East Asian metro networks that would also extend the coverage of existing datasets. During the research internship from September to November, I gained deeper insights into the academic study of accessibility in rail-bound systems and developed a clearer understanding of the role of the SPTL. The findings suggest that future metro operators could learn valuable lessons from East Asian metro networks, where ambitious infrastructure growth does not necessarily translate into improved accessibility.

Before the formal start of the internship, the main personal learning objectives were formulated as:

1. **Independent research competence:** Strengthen the capacity to design, accomplish, and document an academic research project on urban rail transit networks under high-level guidance and constructive feedback from supervisors".

2. **Collaboration and communication:** Enhance collaboration and scientific communication skills through supervision meetings and knowledge exchange within the Smart Public Transport Lab, including clear reporting of progress, challenges, and results.

3. **Technical and data-science skills:** Advance practical skills in handling and processing large-scale raw public transport datasets using Python and relevant libraries, and in implementing reproducible, code-based workflows.

4. **Methodological expertise in network science and graph-based accessibility analysis:** Acquire hands-on experience with methods related to network science and its application in public transport field (e.g. access graphs, L-space and

P-space representation) for the accessibility evaluation of rail-bound PT networks, and translate the result of methods into insights for future metro operators.

Upon completion of the internship, I am confident that all of my personal learning objectives have been fully accomplished. In retrospect, I strengthened my independent research competence by taking responsibility for the full workflow, which consists of defining the scope, compiling the dataset, calculating indicators, and iteratively improving methods and interpretation based on feedback from supervisors. In terms of technical and data-science skills, I became more proficient in using `pandas`, `numpy` and `networkx` to clean heterogeneous raw data, construct L-space and P-space graphs, and compute accessibility indicators in a systematic way. All analyses, figures, and tables were produced from reproducible, script-based workflows, documented in Jupyter notebooks to ensure that each result can be traced, verified, and updated. At the end of the internship, I also successfully contributed to broadening the scope of previous accessibility research at SPTL through the inclusion of East Asian metro networks[6,7]. This hands-on experience with mathematical modelling and Python-based analysis has also extended beyond the research internship itself, enabling me to contribute as a teaching assistant in CEGM1000 Modelling, Uncertainty and Data for Engineers at the civil engineering department. While network science is well established, its application to public transport requires domain-specific expertise learnt at Transport & Planning Department and SPTL tailored to rail-bound systems, including access graphs, degree-based indicators, core–branch structures, and clustering of accessibility growth profiles. Developing specialized skills enabled me to translate quantitative results into operational insights for future metro operators, such as demonstrating that improvement in the network sizes does not automatically lead to higher accessibility when service design and structural integration are not aligned.

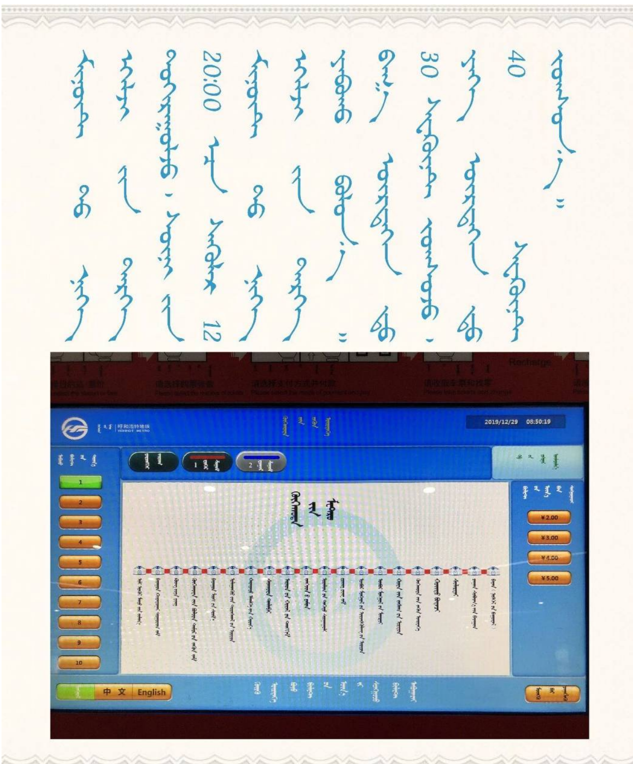
Lastly, my communication and collaboration skills were strengthened as several face-to-face interactions with my supervisors created more space for genuine dialogue rather than the more passive learning common in classroom contexts. During the internship, I actively formulated questions, presented intermediate results, and discussed possibility of usage of clustering approaches. Inspired by Alonso-González et al.[8], who use latent class clustering to characterise user attitudes towards MaaS, I proposed applying clustering to accessibility growth profiles to identify recurring regimes across metro systems. After discussion with my supervisors, this idea was integrated into the project and provided additional valuable insights, illustrating how active communication and collaboration directly improved the research output.

As the research internship has significantly enhanced my academic communication and research independence, I aim to become future traffic engineer or PhD researcher who can translate research outcomes into valuable insights for the transport industry. This experience also marked my first formal research in traffic engineering field after switching tracks from structural engineering, and I initially tended to focus too much on precise technical details, such as transfer station categorization algorithms based on distance and naming similarity. Over time, by consistently focusing on the main research questions, I learned to adopt a more balanced approach to data analysis. Beyond technical development, I faced challenges like extracting raw data from Korean and Japanese sources and representing Tokyo's multi-operator metro network, where I often relied on my linguistic and cultural knowledge to overcome translation barriers. My enthusiasm for rail-bound public transport sometimes led me to overwork, but the work–life balance culture at Transport & Planning Department taught me the value of maintaining a healthy balance. In sum, my key strengths and weaknesses during the research internship can be summarized as follows:

Strengths	Strong analytical and programming skills; perseverance and curiosity; cultural adaptability through international experience.
Weaknesses	Occasional perfectionism and a tendency to focus too much on technical details at the expense of communication or time management.
Opportunities	Pursuing a PhD in accessibility metrics; expanding interdisciplinary collaboration with policy experts.
Threats	Data quality limitations in open-source datasets; balancing academic ambition with personal well-being during intensive research phases.

The SPTL internship not only strengthened my technical expertise but also deepened my understanding of the societal responsibilities of transport researchers. Accessibility research inherently raises ethical questions about equity, inclusion, and fairness in the public transport networks. In my internship project on East Asian metro networks, I examined how differences in network structure and service quality result in disparities in reachability. The results indicated that larger systems often experience greater spatial inequality and reduced accessibility during the later stages of passengers' long journeys, revealing how technical design decisions can unintentionally introduce additional travel impedance for residents living outside urban cores. The inclusion value was also reflected in the regional scope of my research. Accessibility analysis in East Asia is scarce compared to studies focusing on Europe and North America. On the regional level, I went well beyond major and well-known East Asian cities like Shanghai, Beijing, Seoul, and Tokyo, and also included smaller or less internationally recognized systems. My own background as a member of the Tujia minority in southwest China further motivated the inclusion of cities such as

Hohhot (figure 1a) and Ürümqi where ethnic minorities constitute a significant share of the population. In the future, the Tibetan-speaking county in my home province of Sichuan also intends to introduce a rack-rail-based rail system (figure 1b), and the lessons learned from Hohhot and Ürümqi will be of vital importance. This reflected an ethical commitment to ensure that regional diversity and social inclusivity were represented in both regional and global benchmarking. This methodological decision was not only analytical but also moral as traffic engineering and students should use their knowledge on data-driven models to capture voices that are often overlooked in international research. Furthermore, I recognized that open-data methods carry societal responsibilities. Quantifying accessibility has policy implications as it can guide equitable investment, sustainable growth, and balanced regional development. My current and future role in accessibility study was therefore not limited to producing numerical results but also to ensuring that my analytical process adhered to values of transparency, reproducibility, and fairness in data use.



(a) Mongolian news celebrating the opening of the Hohhot metro system, which features Mongolian language support.

(b) Tibetan report on the planned rack rail-bound system in Ngawa Tibetan and Qiang Autonomous Prefecture (Sichuan).

Figure 1. The inclusion of rail-bound systems in regions with minority populations like (a) Inner Mongolia and (b) Ngawa Prefecture (Sichuan) fosters inclusion, equality, and the amplification of local voices

In summary, the research internship at the Smart Public Transport Lab deepened my technical expertise and ethical awareness as a student in traffic engineering. With the inclusion of East Asian metro systems in accessibility-related, I revealed how large systems may face greater spatial inequality and reduced accessibility in suburban areas. The experience strengthened my ability to work independently, communicate effectively, apply domain-specific knowledge, and balance analytical precision with social awareness. Motivated by my Tujia cultural background and by my study, travel, and living experiences across East Asia, I prioritized inclusion by ensuring that all cities with open-sourced metro data were represented in the study. Overall, the internship shaped me into someone who values both solid quantitative analysis and social responsibility in research and practice.

References

1. Delft University of Technology. Smart Public Transport Lab. *TU Delft* https://www.tudelft.nl/citg/over-faculteit/afdelingen/transport-planning/labs/smart-public-transport-lab (2025).

2. van der Bijl, R., van Oort, N. & Bukman, B. Chapter 6 – costs–benefits of light rail revised. *Light. Rail Transit Syst.* 139–154, DOI: https://doi.org/10.1016/B978-0-12-814784-9.00006-2 (2018).

3. Vijlbrief, S., Cats, O., Krishnakumari, P., van Cranenburgh, S. & Massobrio, R. A curated data set of l-space representations for 51 metro networks worldwide. *4TU.ResearchData* https://doi.org/10.4121/21316824.v1 (2022).

4. Vijlbrief, S., Cats, O., Krishnakumari, P., van Cranenburgh, S. & Massobrio, R. A curated data set of p-space representations for 51 metro networks worldwide. *4TU.ResearchData* https://doi.org/10.4121/21316950.v2 (2022).

5. International Association of Public Transport (UITP). GLOBAL METRO FIGURES 2024. *www.uitp.org* https://www.uitp.org/wp-content/uploads/sites/7/2025/08/20250822_Global-Metro-Figures_Statistics-Brief_WEB.pdf (2025).

6. Cheng, H. & Cats, O. A curated data set of p-space representations for 61 east asian metro networks. *4TU.ResearchData* https://doi.org/10.4121/ca8682da-5ab3-4bd4-af37-95517f20a0a5.v1 (2025).

7. Cats, O. & Cheng, H. A curated data set of l-space representations for 61 east asian metro networks. *4TU.ResearchData* https://doi.org/10.4121/d665ec96-dd8b-4e6e-918d-d062a4e1ad42.v1 (2025).

8. Alonso-González, M. J., Hoogendoorn-Lanser, S., van Oort, N., Cats, O. & Hoogendoorn, S. Drivers and barriers in adopting mobility as a service (maas) – a latent class cluster analysis of attitudes. *Transp. Res. Part A: Policy Pract.* **132**, 378–401, DOI: https://doi.org/10.1016/j.tra.2019.11.022 (2020).