



OPEN Network accessibility as the emergence of cliques

Tina Šfiligoj¹✉, Aljoša Peperko^{2,3} & Oded Cats⁴

We propose a topological formulation of accessibility based on the notion of Access Graph, in which two nodes are connected if they are reachable within a given travel time. We trace the emergence and evolution of its subgraphs with imposed levels of connectedness, specifically maximal clique and k -cores. We propose two complementary sets of accessibility indicators, cumulative and threshold, based on integral measures of subgraph growth and times at emergence of k -cores, respectively. For a meaningful comparison of networks across different dimensions, we contrast the realised accessibility with that of an idealised network on the same set of nodes. The proposed measures offer a view of accessibility that extends beyond the commonly used node-averaged indicators. Empirical analysis of 42 metro networks worldwide demonstrates universal patterns of accessibility behaviour. We illustrate the practical application of this approach on a case study where we examine the accessibility impacts yielded by alternative infrastructure and service developments. Our results amount to the reconceptualisation of accessibility within the complex network framework.

Transport network design is a key determinant of travel accessibility. At the same time, accessibility is a key determinant of transport use¹. It is therefore paramount to understand how network structures influence transport accessibility. In particular, urban sustainability development goals call for improving the accessibility and therefore access to opportunities provided by public means of transport.

The classical definition of accessibility is the Hansen definition as travel impedance between spatially dispersed opportunities for activity². In contrast to other more technical measures, such as robustness or vulnerability, common across various fields of engineering³, accessibility is specific to the spatial movements of people, and thus to disciplines such as geography and transport science, and is often perceived as a “soft” measure. The Hansen definition conceptually captures these characteristics and, while it is qualitatively well-founded, it lacks quantitative rigour to guide formal analysis. Consequently, concrete operationalisation and the formulation of related indicators differ significantly across analyses, often relying on ad-hoc interpretations of the devised indicators. This dispersion has made it challenging to devise a unified and generalisable framework for accessibility analysis, leading to the recent recognition of the need to unify accessibility studies^{4–6}.

Accessibility is essentially about connectedness - the ability to reach various destinations within a certain time budget. A complex network perspective arguably offers a promising framework to establish a rigorous and unified analysis. Public transport systems can naturally be modelled as networks, and the methods of network science have been previously applied to study their topological structure^{7–9}. Past works have often focused on network science applications for studying the vulnerability and robustness of public transport networks^{10–14}.

Recently, a complex network approach has been gaining importance also in accessibility studies. Previous studies employ standard topological measures as accessibility indicators. In particular, topological measures such as average path length or network (in)efficiency at the global level^{15,16}, or metrics related to closeness centrality at the node level¹⁷ have been proposed as indicators of accessibility. All of the aforementioned proposed network-level measures are based on node-level averages. In addition, analyses often adopt a strictly topological approach, discarding the infrastructural and operational properties of public transport systems, such as in-vehicle and waiting times, which ultimately influence their use. These properties are direct outcomes of network design decisions such as line configuration, service frequencies and travel speeds. Moreover, there are several standard graph representations, reflecting either the underlying infrastructure or operational characteristics, since reconciling all aspects in a single representation is non-trivial.

In this work, we approach the problem from a complementary perspective, asking first how accessibility can be qualitatively redefined in terms of reachability within well-connected parts of the network. Specifically, our objective is to identify parts of the network where (almost) all nodes are reachable from one another within

¹Faculty of Maritime Studies and Transport, University of Ljubljana, Portorož, Slovenia. ²Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia. ³Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia. ⁴Department of Transport & Planning, Delft University of Technology, Delft, The Netherlands. ✉email: tina.sfiligoj@fpp.uni-lj.si

a given maximum allowed travel time. To this end, we adopt the notion of an access graph \mathcal{G}_A ¹⁸. Its node set represents the public transport stops, and its edge set is constructed by calculating the generalised travel time (*GTT*) matrix and connecting all pairs of nodes i and j for which $GTT(i, j)$ is within a certain cut-off time t_c . *GTT* is comprised of in-vehicle and waiting time, together with time-equivalent transfer penalty. The edge set of the access graph is thus t_c -dependent and \mathcal{G}_A gradually evolves from the empty graph at $t_c = 0$ to a complete graph when t_c reaches the maximum travel time in the network.

To formalise this view of accessibility, we examine several maximal subgraphs of the access graph and their evolution with increasing cut-off time. The strictest condition on connectedness is imposed by the maximal clique, where each pair of nodes shares an edge. A looser condition is that of a k -core, defined as a maximal subgraph where each node has a degree of at least k . Note that the access graphs and all its subgraphs used in our analysis are unweighted and undirected. These allow us to define reachability levels by setting predetermined values of k , for example observing the part of the network where at least $x\%$ of all the stops (of the original network) can be reached from each node. Specifically, we observe the $x = 25, 50, 75$ xk -cores in increasing order of stringency. Each such k -core emerges at a specific *GTT* threshold τ_k in the evolution of \mathcal{G}_A .

The analysis of (maximal) cliques in transport networks in the literature is hitherto relatively scarce. Some interesting works include¹⁹ where the authors consider urban statistical areas as nodes in the graph that are connected by commuting flows. The identification of maximal cliques in this case corresponds to clusters of employment nodes. In²⁰, contact networks of public transport passengers are constructed and cliques of passengers are identified based on similarity in journey and transfer patterns to identify the high-risk journeys in epidemic spreading. In addition,²¹ consider the topological characteristics of the Chinese airline network, focusing on the identification of cliques to identify densely connected parts of the network.

We note that the analysis of k -cores, especially k -core decomposition methods, is often used to examine node importance in networks (e.g.²²) or the hierarchical structure of networks (e.g.²³). k -core decomposition methods have been used in transport research, e.g. for studying network vulnerability²⁴, finding the backbone structure of the network²⁵, and airline network evolution²⁶. The methods used are based on the main core, i.e. the k -core with the largest k . In contrast to the existing analyses, in our approach we consider fixed values of k that reflect the desired levels of connectedness.

We trace the evolution of the selected subgraphs by progressively increasing the cut-off time. The process is visualised in Fig. 1c and d, where the access graph and the considered subgraphs are visualised for six increasing values of t_c . Specifically, the growth of the maximal clique as the most stringent condition on connectedness is illustrative (nodes and edges shown in red). We introduce two complementary sets of accessibility indicators related to the evolution observed for each network: cumulative accessibility indicators $S^{(H)}$ and threshold indicators $\tau^{(H)}$. Cumulative indicators are defined as integral under the curve $N^{(H)}(t_c)$, where $N^{(H)}$ is the dimension of the subgraph at t_c . Threshold indicators $\tau^{(H)}$ reflect a complementary perspective and measure the cut-off time at which the k -cores for fixed values of k first occur. Intuitively, the networks where the subgraphs grow fast early on in shorter time intervals can be regarded as more accessible. The cumulative and threshold indicators correspond to the notions of primal and dual accessibility, respectively²⁷.

To allow for the construction of indicators that enable meaningful comparison across networks, we contrast the concept of an idealised network on the set of nodes, each with given geographical coordinates, similar to²⁸. This allows us to incorporate geographic and topological scaling into the indicator construction. The introduced measures are then normalised to their idealised counterparts to allow for reliable accessibility indicators.

We perform the analysis for 42 metro networks worldwide and compare their accessibility levels. We illustrate our approach in Fig. 1 where we exemplify our analysis for the cases of Paris and London metro networks. The networks are similar in size and are among the largest networks in the dataset. Sub-figures (a) and (b) show the growth dimensions versus t_c for the studied subgraphs. The evolution of the access graphs and their maximal cliques and kx -cores is visualised in Fig. 1c, d. The curves for the maximal clique of the actual and idealised networks are shown in figures (e) and (f). Visually, the networks where the curves are closer to the idealised curves, corresponding to smaller gray areas in the plots, are more accessible. In this example, the London metro network has significantly higher accessibility than the Paris metro.

The accessibility indicators proposed in this study are compared to commonly used topological indicators, as well as to traditional indicators from transport geography. The implications of this approach for transport planners are illustrated on a case study of the Stockholm metro, where different developments of the network are simulated and the changes in accessibility levels are measured. We discuss the implications of the proposed framework for future accessibility studies.

Results

We construct access graphs of 42 metro networks worldwide and examine the evolution of their maximal subgraphs with progressively increasing cut-off time. For each network, we observe four characteristic subgraphs H : the maximal clique and xk -cores for $x \in \{25, 50, 75\}$. For accessibility measurement, we follow the same procedure for the idealised network to give a meaningful reference for actual accessibility levels. Sect. 2.1 presents the results for the idealised scenario and sets the scene for the results of the existing networks. The results for the actual networks and their comparison to their idealised counterparts are the subject of Sect. 2.2. Comparative assessment of the cities based on the proposed accessibility indicators is presented in Sect. 2.3. In Sect. 2.4, the proposed indicators are compared to the conventional accessibility indicators.

Idealised network

To be able to meaningfully assess the level of accessibility and compare networks of differing topological and geographical dimensions, a notion of an ideally accessible network is needed. In the absence of which, one compares to zero travel times between all origins and destinations which is not only unrealistic but also distortive

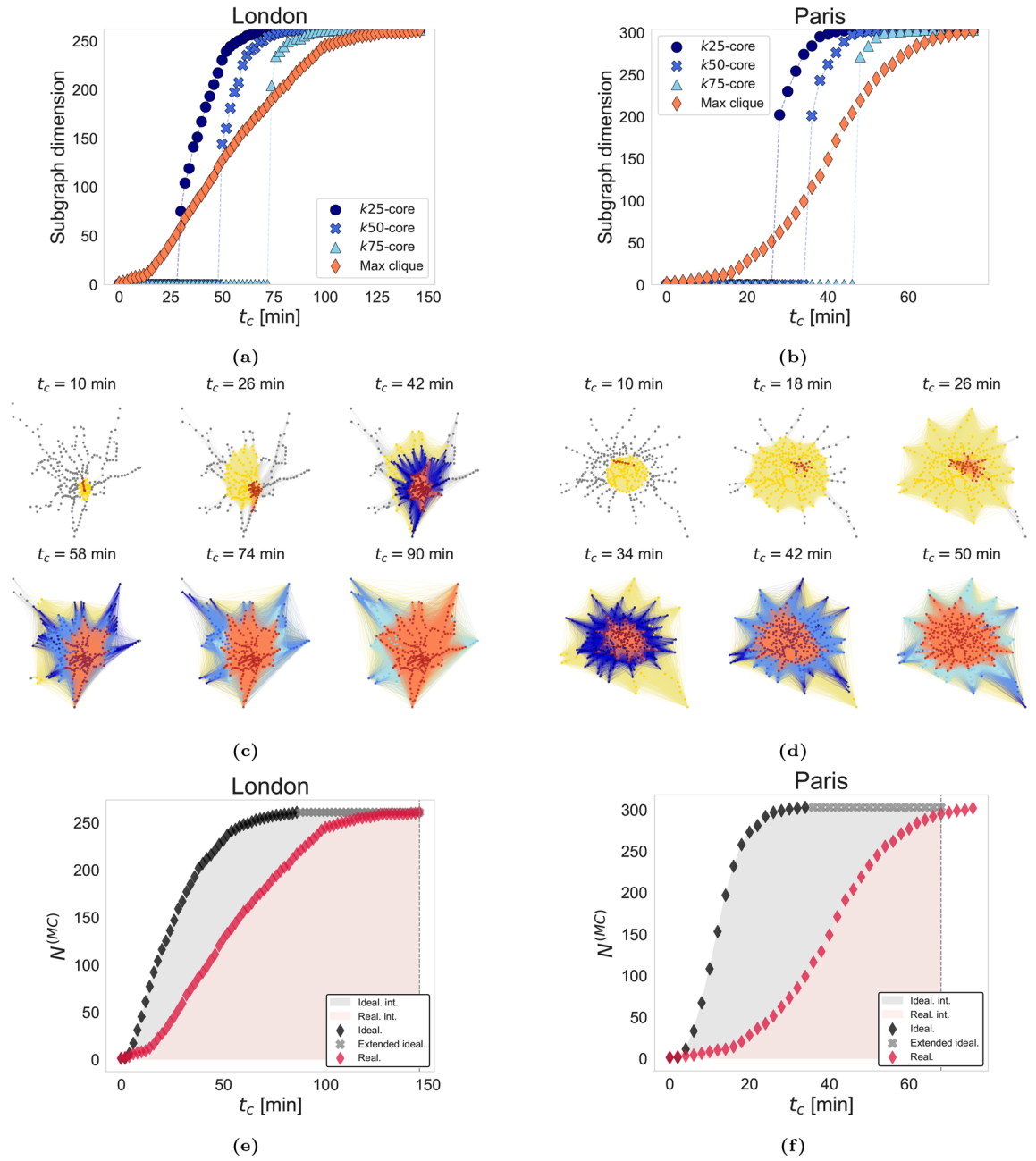


Fig. 1. Illustration of the subgraphs and related indicators for the London and Paris metro networks. Plots (a) and (b) show the dimension of the subgraph with respect to t_c . The red markers show the dimension of the maximal clique vs. cut-off time t_c . Dark, medium and light blue markers show the dimensions of the k_{25} , k_{50} , and k_{75} cores, respectively. Plots (c) and (d) show the evolution of the access graph (grey), k_{25} -core (dark blue), k_{50} -core (medium blue), k_{75} -core (light blue) and maximal clique (red), together with the idealised maximal clique (yellow) for London and Paris networks, respectively. The cumulative accessibility indicators used in the analysis are shown in plots (e) and (f) for the respective networks. Dimension of the idealised maximal clique is shown in black markers. We set the upper integral boundary to $\min(t_{max}, 2 \cdot t_{max}^{ideal})$, where t_{max} is the maximum travel time in the network and $2 \cdot t_{max}^{ideal}$ is twice the maximum idealised travel time. The minimum is taken, as in some networks $t_{max} < 2 \cdot t_{max}^{ideal}$. This happens in the London network, while the $2 \cdot t_{max}^{ideal}$ threshold is shown in (f) for Paris. The cumulative accessibility indicator represents the shaded red area relative to the shaded grey area. The latter represents the idealised cumulative accessibility against which the actual cumulative accessibility is measured. The values of selected indicators for both networks are: $S^{(MC)} = 0.76$, $S^{(k_{50})} = 0.77$, $\tau_{(k_{50})} = 0.34$ for London; and $S^{(MC)} = 0.53$, $S^{(k_{50})} = 0.57$, $\tau_{(k_{50})} = 0.50$ for Paris.

when comparing networks of different scales. Public transport networks are spatial networks, and the physical constraints of distance coupled with allowed travel speed pose an upper limit to possible accessibility levels. The problem is formally stated as: *given a set of N nodes, each with geographical coordinates (lat, lon), what is the theoretical upper limit of accessibility levels?*

The problem is nontrivial and even in the idealised case requires a set of assumptions. The maximum speed must be set so that velocity and acceleration are safe for passengers. Physical constraints also determine the maximum capacity, so a finite frequency must be set.

Even with unlimited resources, designing a realistic (i.e. satisfying the physical constraints) optimally accessible network requires several stages of planning, such as line planning and frequency setting, rendering this an NP-hard problem. We therefore approach the problem from a strictly theoretical perspective. We assume that each origin-destination pair can be travelled directly and in a straight line. We then set typical values of average speed and waiting time (proportional to inverse frequency) for all connections in all networks. The construction of the idealised network is described in more detail in the Methods Sect. 4.2. The maximum and mean Haversine distances, together with the corresponding idealised travel times, are shown in Table 1. Maximum distances vary by an order of magnitude, spanning from $\approx 7.5\text{km}$ (Marseille) to $\approx 75\text{km}$ (San Francisco), with the average maximum distance of $\approx 26\text{km}$ across all networks. Mean distances are highly linearly correlated with maximum distances ($r = 0.94$). Figure 2 shows the scatter plot of maximum Haversine distance ℓ^{max} and number of nodes N . The relationship is approximately linear with a few outliers that we comment on here. On one end, the most notable outlier is the San Francisco network, where 50 stops cover a maximum distance of more than 75 km. The Valencia network is another outlier in the same direction. On the other end, the New York and Paris networks cover a comparatively small geographic area relative to the number of nodes in the network. The difference stems from the metro systems covering different areas in different cities: in some cases only the urban area (e.g. Paris, whereas the RER extends further into parts of Ile-de-France), while in other cases they reach far into the outskirts (e.g. San Francisco and parts of the Bay area).

Comparison of real and idealised network behaviour

We construct the access graphs for the metro networks and their idealised counterparts. The growth of the examined subgraphs in both cases is shown separately for each city in Figure 3. Similar growth patterns are observed across all networks. The maximal clique dimension is a continuous monotonic function of t_c that increases from $N^{(MC)} = 1$ at $t_c = 0$ to $N^{(MC)} = N$ at $t_c = t^{max}$. Conversely, kx -core dimension $N^{(kx)}$ is 0 until a certain threshold time $t_c = \tau_{kx}$, and emerges with $N^{(kx)} \geq kx$ at τ_{kx} and monotonically increases after that. The difference between the growth of the actual and idealised cliques, as visually demonstrated by the relative difference of the areas under the violet (idealised) and orange (actual) curves, is examined in more detail in the next Sect. (2.3).

To understand the differences in maximal clique and k -core growth behaviour between realised and idealised networks, we examine possible sources of discrepancy. In the idealised network, all distances can be travelled directly. In real networks, this is - either due to network infrastructure or line planning - generally not the case, and the large majority of connections have a larger actual distance. The network indicator encapsulating this behaviour is circuitry, which measures the pair-wise discrepancy between the network distance $d_G(i, j)$ and the straight-line distance $d_E(i, j)$ in terms of their ratio (Eq. 2 in Methods)^{29,30}. The average circuitry over all node pairs is the corresponding global metric (Eq. 3 in Methods).

The average circuitry of the idealised network is $C = 1$ by design since it means that all connections are direct without requiring any detours. The average circuitry of the realised network is thus a measure of discrepancy from the ideal case. Since we only have data on travel times (and not on the exact shape of physical paths), we consider two measures of distance: in-vehicle time for infrastructural circuitry C^I , and waiting time for operational circuitry C^O . The values of C^I and C^O for all networks are given in Table 2. The relationship of infrastructural and operational circuitry to accessibility measures is presented and discussed in Sect. 2.3.

For all networks, the growth of the idealised network subgraphs consistently occurs significantly earlier in the time interval than that of the actual network. In addition, consistent patterns in the shapes of kx -core and maximal clique growth for each network point to universal behaviour of accessibility. A quantitative examination of the discussed behaviour will provide the means for comparative assessment of networks in terms of accessibility levels in the following section.

Accessibility measures

Cumulative accessibility indicators are defined as integrals under the subgraph dimension curves of the actual network, relative to the idealised network (see Sect. 4.4 for more detail). To relax the idealised accessibility assumption, the integrals are evaluated up to twice the maximum idealised travel time, $2t_I^{max}$. For some networks, this value is larger than the actual maximum travel time. Specifically, this happens for the Cairo, Copenhagen, Dubai, Hyderabad, Lille, London, Milan, Prague, San Francisco, Santiago, Valencia, Warsaw and the Washington DC networks. The upper bound T of the integral is thus determined as $T = \min(t^{max}, 2 \cdot t_I^{max})$. This is illustrated in Fig. 4. Similarly, the threshold indicators $\tau_{(kx)}$ are defined relative to T .

The rankings with the color-coded normalised indicator values for the full dataset are shown in Fig. 5. Note that higher values of the cumulative accessibility indicators mean higher accessibility, and vice versa for the threshold indicators. The five most accessible networks are London, San Francisco, Valencia, Bilbao and Dubai. Note that all of the networks, except Bilbao, fall in the $T = t_{max}$ category, i.e., have relatively low actual maximum travel times compared to the idealised case, and also in the idealised case the maximal clique growth has a distinctively slow convergence compared to most other networks (Figs. 3 and 4). At the lower end, Buenos Aires, Marseille, Philadelphia, Kobe and Oslo score the lowest on the proposed accessibility indicators. All five networks are characterised by a steep growth of the maximal clique dimension in the idealised case and a slow

City	N	ℓ^{max} [km]	$\bar{\ell}$ [km]	t_I^{max} [min]	\bar{t}_I [min]
Amsterdam	39	14.5	5.6	21.8	10.0
Athens	61	26.5	7.9	37.8	13.1
Atlanta	38	34.8	11.7	48.9	18.0
Berlin	174	29.3	8.2	41.6	13.5
Bilbao	42	18.9	7.1	27.7	11.9
Boston	52	26.2	7.3	37.5	12.2
Brussels	59	14.2	4.9	21.4	9.0
Budapest	48	11.9	4.1	18.4	7.9
Buenos Aires	78	11.3	4.4	17.6	8.4
Cairo	61	35.0	10.9	49.2	17.0
Chicago	137	39.4	11.5	55.0	17.8
Copenhagen	39	11.8	4.4	18.2	8.4
Dubai	53	43.9	16.5	61.0	24.5
Hyderabad	56	24.7	8.2	35.4	13.5
Kobe	26	16.7	6.6	24.7	11.4
Lille	60	18.1	7.0	26.6	11.9
Lisbon	50	10.9	4.3	17.1	8.2
London	261	61.7	14.3	84.7	21.6
Lyon	40	10.5	3.8	16.5	7.6
Madrid	240	36.3	10.0	50.8	15.8
Marseille	29	7.5	2.9	12.5	6.3
Milan	106	27.7	7.3	39.5	12.2
Montreal	67	17.5	6.5	25.8	11.1
Naples	28	13.9	4.1	21.1	8.0
New York	421	37.9	12.5	53.0	19.1
Nuremberg	49	16.1	5.0	23.9	9.2
Oslo	101	24.7	7.8	35.4	12.9
Paris	303	22.3	5.6	32.2	10.0
Philadelphia	50	16.8	6.1	24.9	10.6
Prague	58	21.3	6.7	30.9	11.5
Rome	73	24.7	7.6	35.4	12.6
Rotterdam	70	32.2	11.3	45.4	17.6
San Francisco	50	75.9	28.7	103.8	40.8
Santiago	119	29.2	8.8	41.5	14.2
Stockholm	101	22.4	8.4	32.4	13.7
Toronto	75	27.5	9.3	39.2	15.0
Toulouse	37	9.8	3.9	15.6	7.7
Valencia	95	60.3	14.9	82.8	22.4
Vancouver	52	29.9	11.3	42.3	17.5
Vienna	98	19.5	6.1	28.5	10.6
Warsaw	33	19.9	6.8	29.1	11.6
Washington	89	43.2	12.4	60.1	19.1

Table 1. Key statistics of geographical distances and idealised travel times for all cities. N : number of nodes; ℓ^{max} : maximum Haversine distance in the network; $\bar{\ell}$: average distance; t_I^{max} : maximum idealised travel time; \bar{t}_I : average idealised travel time.

growth in the actual network in the beginning of the interval, resulting in larger discrepancies between the existing network behaviour and that of its idealised version (Fig. 4).

We observe that for most of the networks, the rankings based on different indicators tend to be overall in agreement. This is also confirmed by the finding that all indicators are strongly correlated (Fig. 6) with positive correlations within each type of indicators - cumulative accessibility and threshold-based - and strong negative correlations between those, as expected. Notwithstanding, for a subset of networks, the rankings vary significantly between indicators. Among those are the Paris network, where kx -core indicators rank it significantly higher than the maximal clique indicator, whereas the pattern is reversed for the Copenhagen network. As can be seen in (Fig. 6), the correlations are the strongest for the maximal clique and $k50$ -core cumulative indicators ($r = 0.96$). Linear correlations are the highest for pairs of complementary indicators $S^{(xk)}$ and τ_{xk} at the same values of x ($r = -0.90$ for $x = 25$; $r = -0.96$ for $x = 50$; $r = -0.92$ for $x = 75$). The correlation strength remains

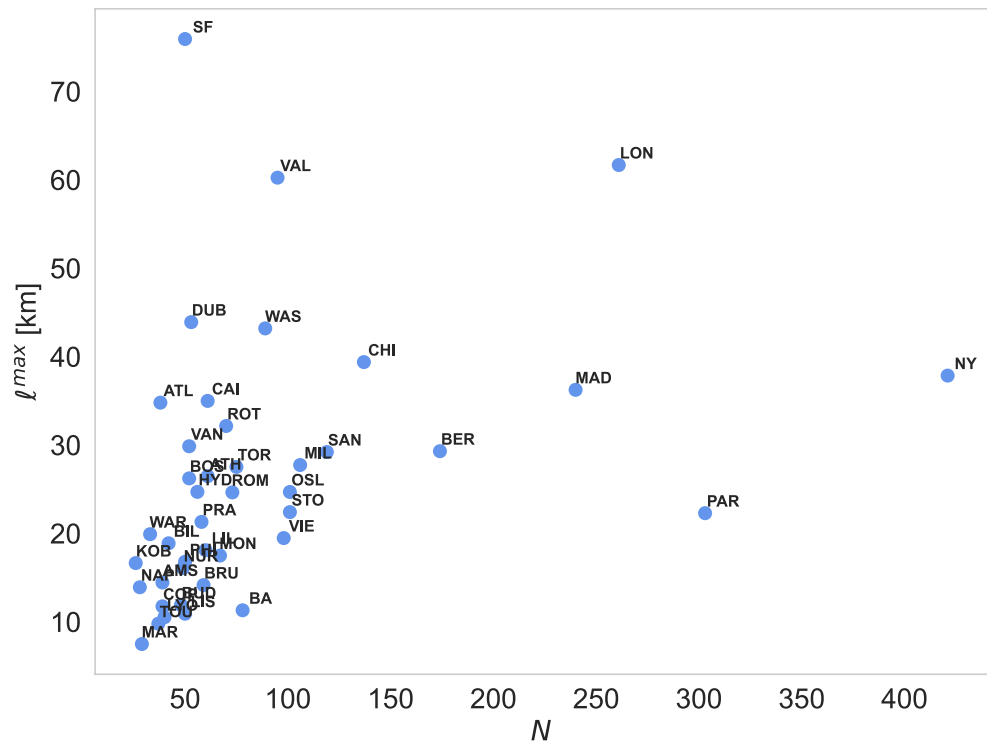


Fig. 2. Scatter plot of maximum Haversine distance ℓ^{max} vs. number of nodes N for all cities.

high across all combinations, but decreases for pairs of cumulative and threshold indicators with differing values of x (-0.82 to -0.89 for consecutive values of x and $r = -0.73 - r = -0.76$ for pairs with $x = 25$ and $x = 75$). The cross-correlations of the cumulative indicators at different values of x are the lowest for $S^{(k25)}$ and $S^{(k75)}$ ($r = 0.81$). Compared to $x = 50$, the two subgraphs for $x = 25$ and $x = 75$ are more sensitive to specific conditions at lower and upper boundaries of the time interval that might differ for each network. The $k50$ -based indicators are therefore deemed the most stable. The overall results suggest high regularity in maximal clique and k -core growth behaviour across all networks, pointing to universal patterns in network accessibility.

To further examine the relationship between accessibility measures and circuitry (Sect. 2.2), we observe their correlations (Fig. 7).

Correlations imply that the more circuit the network is the less accessible it is, in agreement with expectations, i.e. higher circuitry indicates lower accessibility. The correlations are stronger for infrastructural circuitry, implying the greater impact of infrastructure on accessibility levels compared to operational characteristics. Specifically, we note that the networks with the highest values of C^I - greatest circuitry - are Naples, Buenos Aires and Oslo (Table 2) are those that rank among the lowest few in Fig. 5. Conversely, the cities with the lowest values of C^I - San Francisco, Atlanta and Copenhagen, do not per-se, with the exception of San Francisco, correspond to the highest ranked cities. For operational circuitry, an interesting observation is that London has by far the lowest average waiting time and is also overall the most accessible network according to the indicators introduced here. On the other hand, San Francisco has one of the highest values of C^O . At the same time, it is an outlier in terms of infrastructural circuitry values with $C^I = 0.9$, i.e. less than one. In the construction of idealised networks, the speed on links is taken to be constant across all networks. San Francisco is an outlier in its geographical diameter (approx. 75 km) and the actual speeds on (long) rail segments can be significantly higher than this value.

The access graph and three of the subgraphs (maximal clique, $k25$ -core and $k50$ -core) at a characteristic value of $t_c = \tau_{k50}$ are visualised in Fig. 8. The networks are ranked and their subplot borders are color-coded based on the $S^{(MC)}$ indicator. The most accessible networks have maximal cliques of comparable dimensions to $k50$ -cores at the latter's emergence (compare also with the plots in Fig. 3). Moreover, the cliques are geographically relatively dense, while the spatial density is visibly low for the least accessible networks. Compare also the L-space layout with circuitry values in Table 2.

The geographical distribution of the cities included in the analysis is shown in Fig. 9.

Relation with conventional indicators

Topological indicators

Accessibility in general, and the topological indicators we consider in particular, rely on the identification of shortest paths. We therefore investigate their relation with commonly used topological indicators which are closely related to the notion of shortest paths: network efficiency E ; average shortest paths $\langle SP \rangle^L$ for in-vehicle time-weighted L-space representation, $\langle SP \rangle_w^P$ and $\langle SP \rangle^P$, for weighted and unweighted P-space representation,

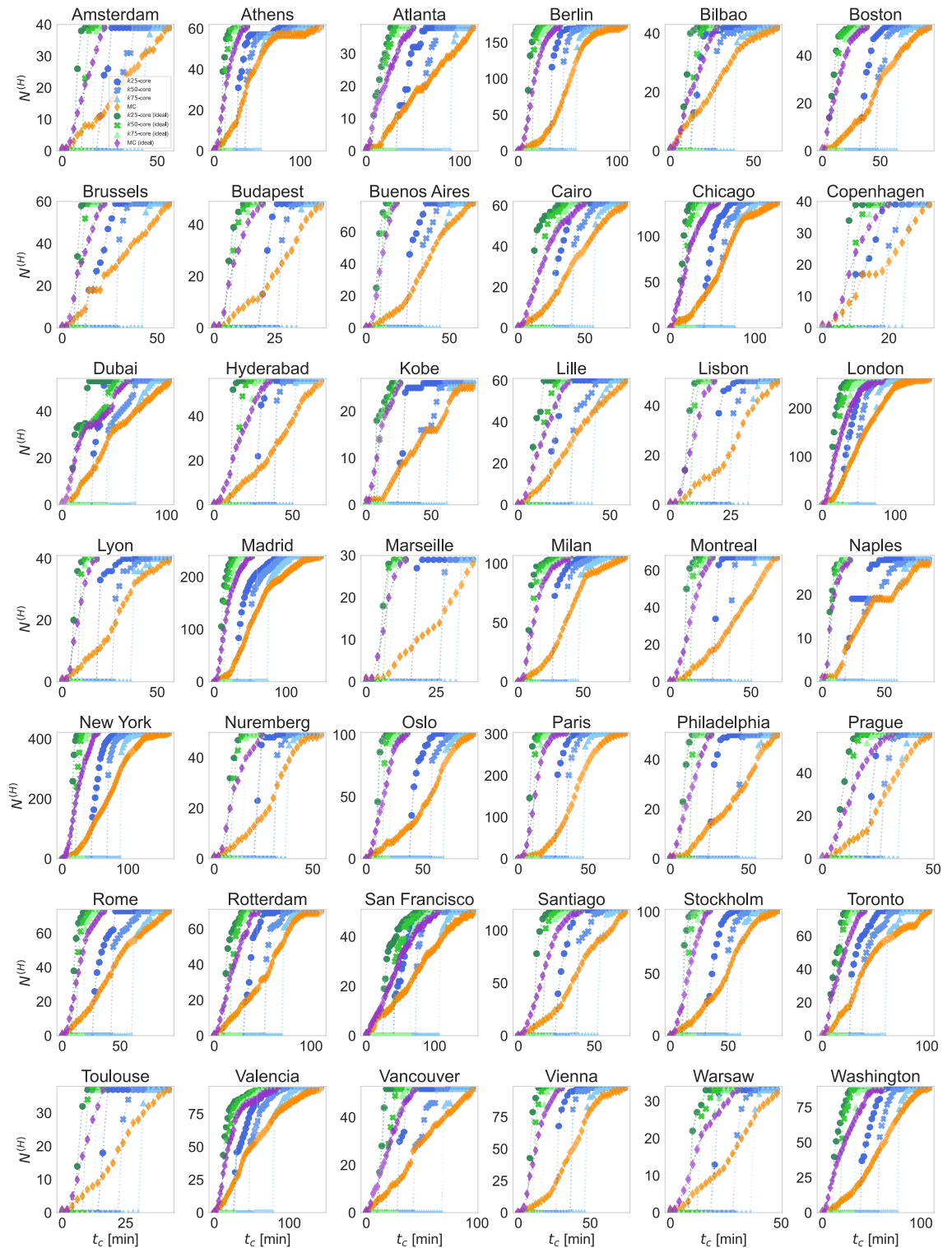


Fig. 3. Comparison of kx -cores and maximal cliques dimension growth vs. cut-off time t_c for the idealised and actual metro networks for all cities. Green (kx -cores) and violet (maximal clique) curves represent the idealised network, and blue and orange curves represent the actual network in each subplot.

respectively. Similarly, we consider the network diameter d in all three representations. All measures are normalised to their counterparts in the idealised network (see Sect. 4.5 for definitions).

The correlation matrix is shown in Fig. 10. The correlations are the highest for efficiency and weighted P-space indicators. Note that the definition of distance in the calculation of efficiency is *GTT*, Generalised Travel Time, i.e. the same metric that we use in the construction of the access graph.

City	C^I	C^O
Amsterdam	1.5	6.5
Athens	2.1	6.8
Atlanta	1.3	16.8
Berlin	2.1	5.6
Bilbao	1.9	2.8
Boston	2.3	9.2
Brussels	2.2	3.9
Budapest	2.4	3.5
Buenos Aires	2.7	4.4
Cairo	2.1	2.2
Chicago	2.4	9.0
Copenhagen	1.4	2.7
Dubai	1.5	5.1
Hyderabad	1.8	4.6
Kobe	2.5	11.2
Lille	2.0	1.9
Lisbon	1.8	3.2
London	2.0	1.7
Lyon	2.3	4.0
Madrid	2.4	6.5
Marseille	2.4	4.6
Milan	2.2	3.7
Montreal	1.8	6.0
Naples	3.5	8.4
New York	2.3	7.9
Nuremberg	2.1	4.1
Oslo	2.7	9.6
Paris	2.5	3.4
Philadelphia	2.0	8.1
Prague	1.6	3.4
Rome	2.3	5.7
Rotterdam	1.8	8.0
San Francisco	0.9	12.3
Santiago	1.5	5.6
Stockholm	2.3	6.7
Toronto	2.0	4.7
Toulouse	2.3	1.9
Valencia	2.1	4.6
Vancouver	1.7	5.4
Vienna	2.0	4.9
Warsaw	1.7	2.8
Washington	1.9	9.1

Table 2. Circuitry metrics across metro networks.

Similarly, correlations with average shortest paths are moderate, especially for the weighted representations. The correlations with shortest paths and diameter are negative since longer shortest paths reflect lower accessibility. The correlations are generally the highest for the weighted P-space, followed by weighted L-space and unweighted P-space, reflecting the influence of the service properties contained in the definition of GTT , used to construct the access graph (Eq. (4)). Insignificant to low correlations with network dimension N are expected due to the scaled construction of indicators using the notion of idealised accessibility.

Geographical indicators

Accessibility is a mature research area within transport geography. The classical indicators are based on the Hansen definition as the ease of travel between spatially dispersed opportunities for activity². A standard operationalisation of this notion is the cumulative opportunities measure of accessibility which measures the number of reachable opportunities within a given threshold time (definitions in Eq. 9 and Eq. 10 in Methods).

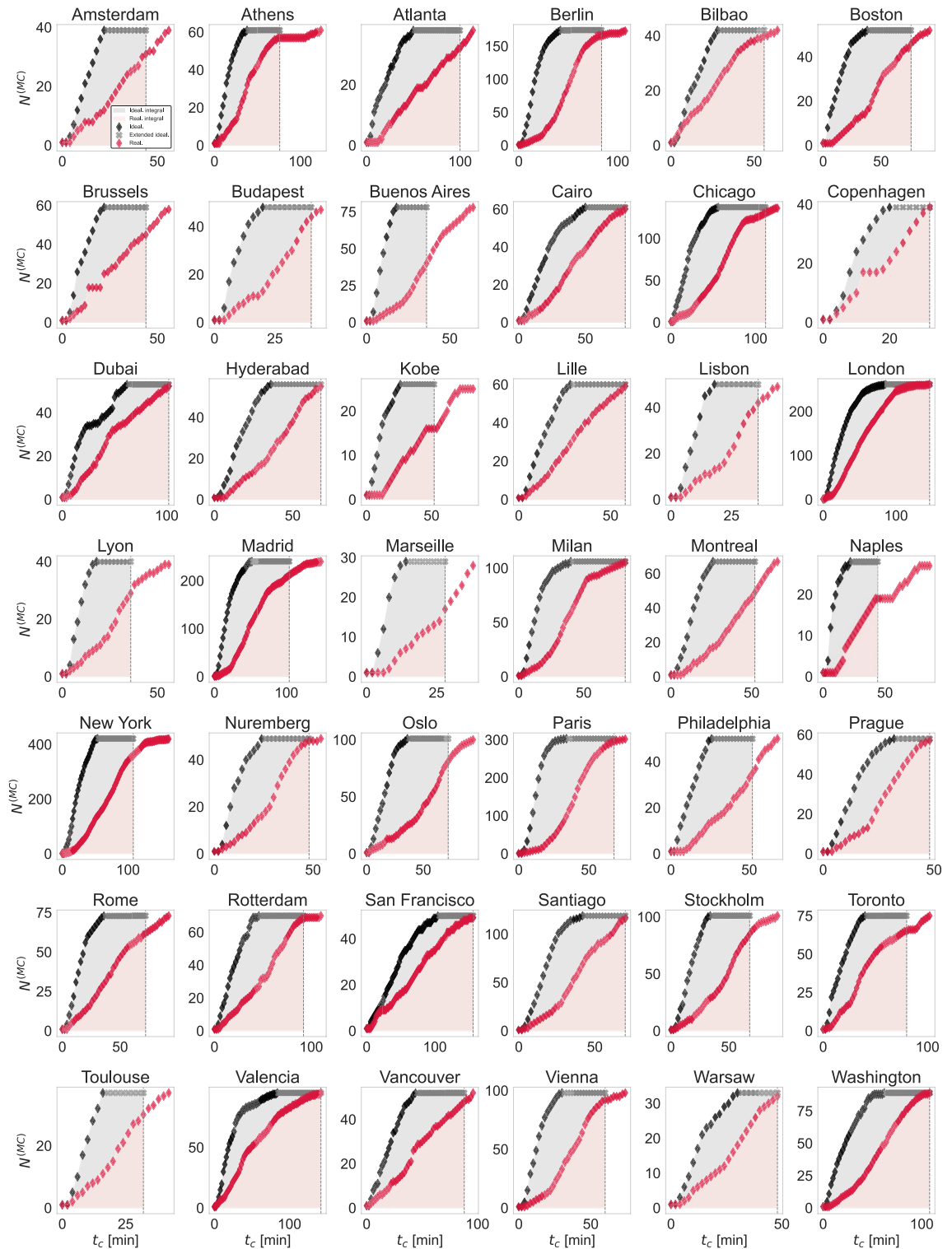


Fig. 4. Comparison of maximal clique growth for the ideal (black) and original (red) networks for each city. For indicator construction, we integrate up to twice the maximum idealised time (grey markers). If the maximum travel time in the original network is smaller than twice the idealised value, we integrate to maximum travel time (e.g. Cairo, Dubai, London). This is the case for 11 of 42 networks, i.e. about a quarter.

In the access graph, this corresponds exactly to node degree k_i at the threshold time t_c . To compare and contrast the introduced indicators to classical measures, we examine the average node degree in the access graph (i.e. the cumulative opportunities accessibility) for a few select values of t_c : $t_c = \{30, 45, 60\}$ minutes. Additionally, we consider a few values of t_c relative to the network-specific maximum GTT, t_{max} : $t_c = \{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}\}t_{max}$.

City	$S^{(MC)}$	$S^{(k25)}$	$S^{(k50)}$	$S^{(k75)}$	$\tau_{(k25)}$	$\tau_{(k50)}$	$\tau_{(k75)}$
London	1	3	1	1	2	2	1
San Francisco	2	2	3	3	4	6	8
Valencia	3	4	2	2	1	1	2
Bilbao	4	1	5	6	3	5	4
Dubai	5	6	4	5	6	3	11
Copenhagen	6	5	21	18	8	26	25
Cairo	7	9	7	7	15	14	14
Milan	8	11	6	4	14	4	3
Lille	9	7	8	9	7	12	13
Vancouver	10	8	9	23	10	9	20
Toronto	11	15	10	21	12	8	16
Athens	12	21	12	17	15	7	9
Warsaw	13	10	20	14	25	26	23
Prague	14	22	13	8	34	20	17
Washington	15	17	15	12	19	16	12
Chicago	16	23	19	10	17	15	6
Santiago	17	12	17	19	20	18	22
Rotterdam	18	13	22	22	12	17	17
Hyderabad	19	16	18	16	24	21	21
Nuremberg	20	25	27	15	28	25	15
Madrid	21	19	16	20	9	10	10
Rome	22	27	25	28	18	19	26
Berlin	23	18	14	13	21	13	7
Atlanta	24	20	24	31	11	23	31
Paris	25	24	11	11	22	11	4
Brussels	26	14	26	34	5	26	33
Vienna	27	29	23	24	30	22	17
Toulouse	28	26	32	32	31	32	32
Amsterdam	29	28	34	35	25	32	33
Budapest	30	36	31	26	33	30	27
Boston	31	33	29	27	29	24	28
New York	32	30	28	25	23	26	24
Lisbon	33	35	30	29	37	31	29
Stockholm	34	37	35	30	31	35	30
Lyon	35	38	36	37	37	36	37
Montreal	36	34	33	36	37	34	36
Naples	37	40	39	38	25	37	42
Oslo	38	41	37	33	40	39	35
Kobe	39	32	40	38	35	38	40
Philadelphia	40	31	38	38	35	40	38
Marseille	41	39	42	38	41	42	39
Buenos Aires	42	42	41	38	42	41	41

Fig. 5. Color-coded city rankings for the proposed indicators. The left-hand side shows the results for the four cumulative accessibility indicators $S^{(MC)}$, $S^{(k25)}$, $S^{(k50)}$, $S^{(k75)}$, and the right-hand side shows the three threshold indicators $\tau_{(k25)}$, $\tau_{(k50)}$, $\tau_{(k75)}$. For cumulative indicators, the city is ranked higher (i.e. has lower ranking score) in terms of accessibility for higher values of indicators. Conversely, lower values of time variables indicate higher accessibility. $\tau_{(kx)}$ values were multiplied by (-1) to achieve consistent ranking.

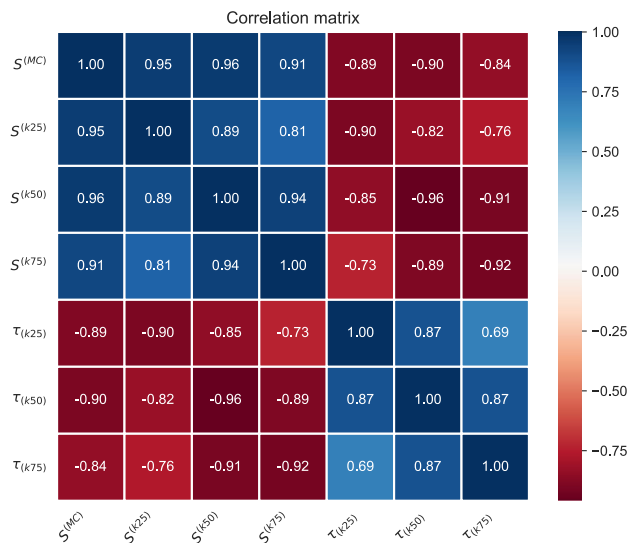


Fig. 6. Pearson correlation matrix of the indicators. $S^{(H)}$: cumulative accessibility indicators; $\tau_{(H)}$: threshold accessibility indicators. Blue and red shades represent positive and negative correlations, respectively.

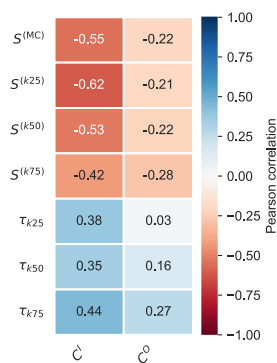


Fig. 7. Pearson correlation matrix of the indicators and circuitry measures. $S^{(H)}$: cumulative accessibility indicators; $\tau_{(H)}$: threshold accessibility indicators. C^I : infrastructural circuitry; C^O : operational circuitry. Blue and red shades represent positive and negative correlations, respectively.

The correlations of the introduced indicators to absolute-time values are low and negative (Fig. 11). This is in line with expectations since the geographical areas and thus the physical constraints on travel ability vary significantly between cities and the introduced indicators are designed to allow for meaningful comparisons across networks of different scales. The correlations with time points relative to maximum time are moderately positive which is in agreement with our formulation of accessibility as the one realised by the infrastructure network and the super-imposed service plans compared to the system’s geographical constraints.

We stress however that the indicators introduced in this work reflect the temporal evolution of the accessibility levels, and are thus qualitatively different from the cumulative opportunities definition. The cumulative indicators $S^{(H)}$ proposed in this work can be considered to represent “cumulative cumulative accessibility”, i.e., at a higher level of aggregation than cumulative accessibility in its traditional sense.

Illustration

For an illustration of the practical meaning of the introduced accessibility indicators, we investigate the impacts of different developments for the case of the Stockholm metro. We then compare accessibility measures of the modified networks with the original network. The Stockholm metro currently has $N = 101$ nodes and operates $L = 3$ lines and $R = 7$ routes (Fig. 12). It scores relatively low in accessibility rankings (30 to 37 out of 42 across all measures). It has large infrastructural circuitry as can be seen from its topology: a central core to the east, and cross-center radial routes spread out to the west, the nodes on which are geographically relatively close but topologically far apart.

For the sake of illustration, we consider the following three scenarios: (i) increasing (doubling) the frequencies on all lines in the network (refer to it as $2F$); (ii) introducing a new line (Line X, Fig. 13), and; (iii) combined previous two points ($2F + \text{Line X}$).

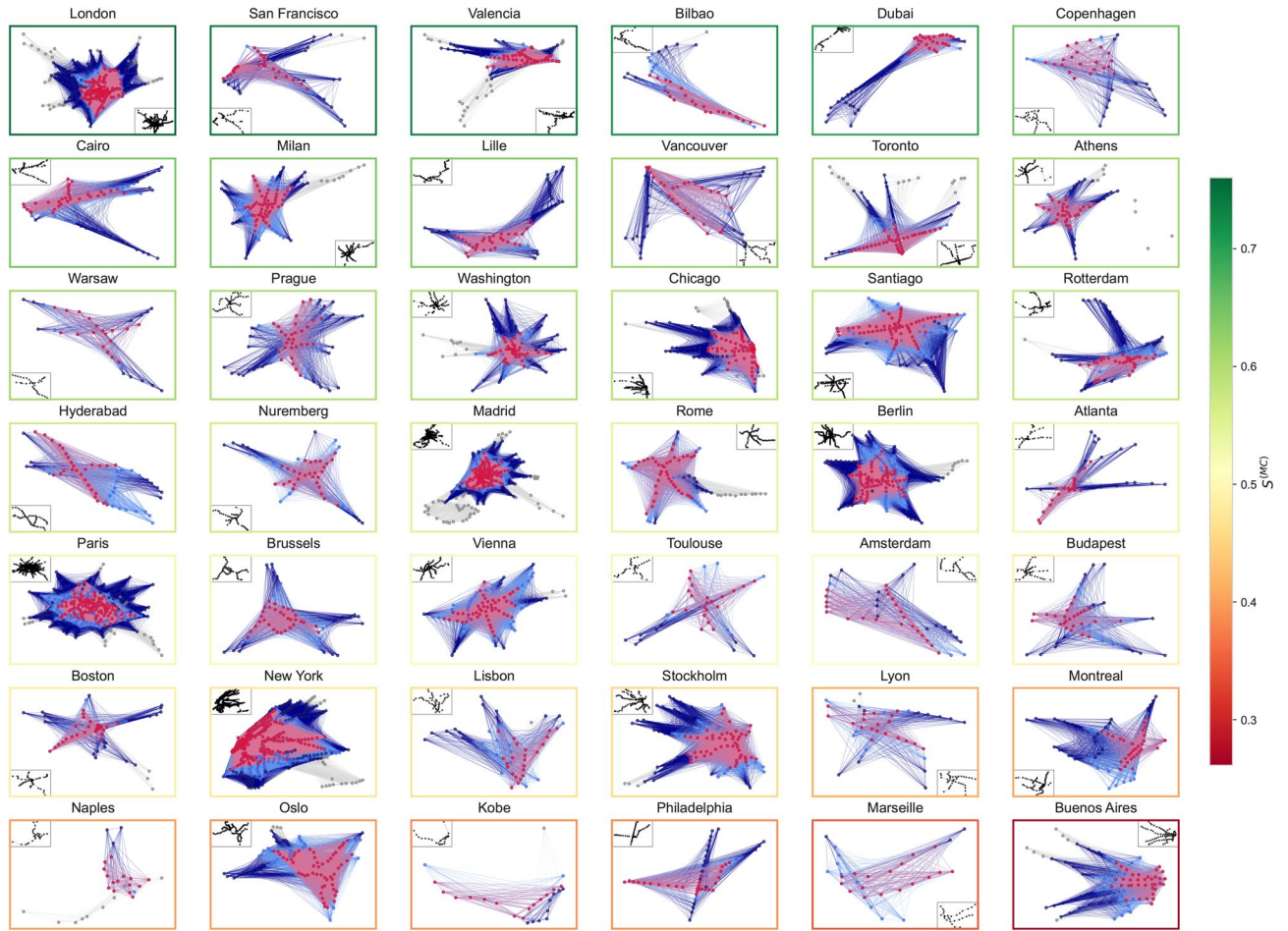


Fig. 8. Visualisation of subgraphs at $\tau_{k=50}$ for each metro network. Dark blue: k_{25} -core; light blue: k_{50} -core; red: maximal clique. L-space infrastructural representation is shown in a figure insert for each city. Cities are sorted based on their maximal clique cumulative accessibility indicator $S^{(MC)}$ and the subplot borders are color-coded based on the value of this indicator.

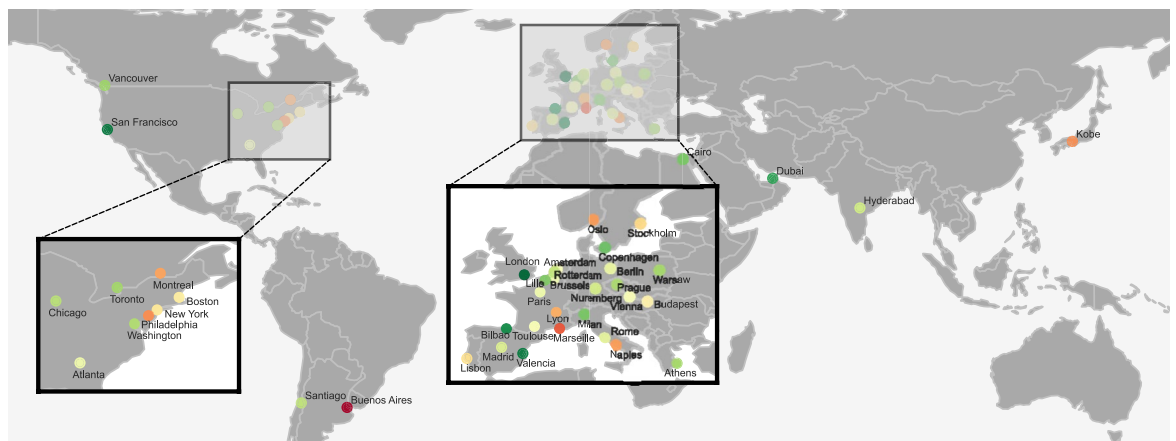


Figure 9. Geographical distribution of the studied networks with color-coded accessibility levels based on the $S^{(MC)}$ indicator. The ranking and the scale is the same as in Fig. 5. Two areas with the highest density of cities included in the analysis are zoomed in in the map.



Fig. 10. The Pearson correlation matrix for accessibility indicators: cumulative indicators $S^{(H)}$ and threshold indicators $\tau(xk)$. Blue and red shades represent positive and negative correlations, respectively.

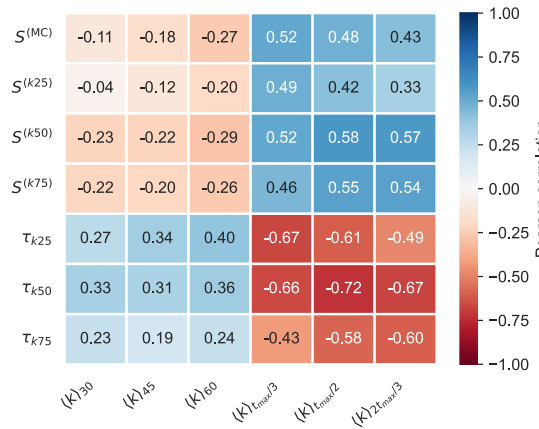


Fig. 11. The Pearson correlation matrix for accessibility indicators: cumulative indicators $S^{(H)}$ and threshold indicators $\tau(xk)$. Geographical indicators are the average degrees of the access graph at different time points: $t_c = \{30, 45, 60\}$ minutes and $t_c = \{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}\}t_{max}$. $(k)_{30}$ is the average degree at $t_c = 30$ min, and $(k)^{t_{max}/3}$ is the average degree at $t_c = \frac{1}{3}t_{max}$ (and similarly for others). Blue and red shades represent positive and negative correlations, respectively.

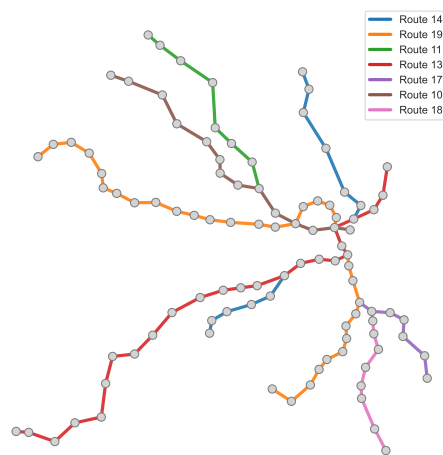


Fig. 12. Stockholm L-space with colour-coded routes.

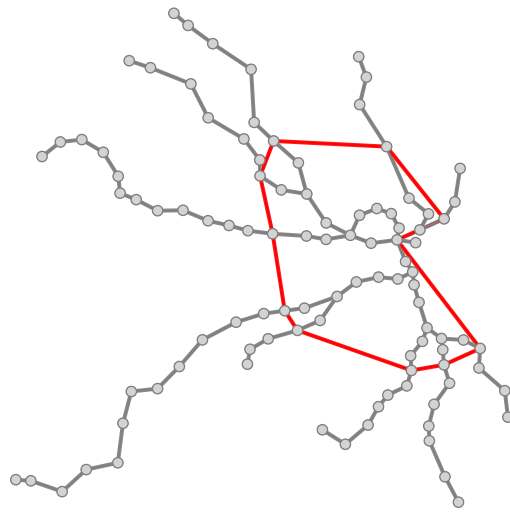


Fig. 13. Stockholm L-space with added “Line X” (shown in red).

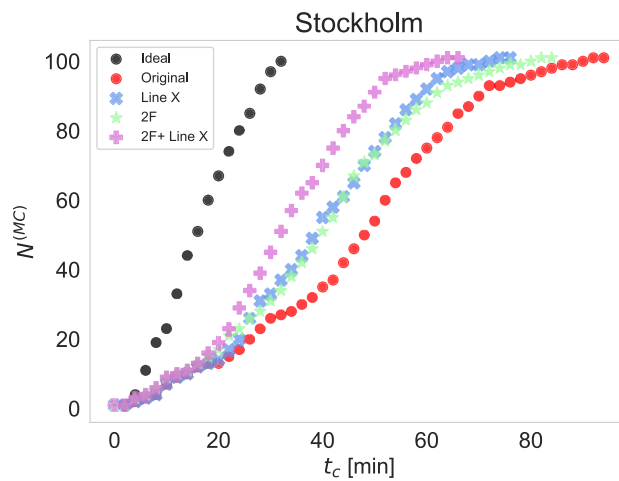


Fig. 14. Comparison of maximal clique growth with simulated improvements to the metro network.

Scenario	$S^{(MC)}$	$S^{(k25)}$	$S^{(k50)}$	$S^{(k75)}$	τ_{k25}	τ_{k50}	τ_{k75}
Original	0.44	0.54	0.31	0.11	0.48	0.76	0.94
Line X	0.57	0.66	0.51	0.33	0.39	0.61	0.79
2F	0.56	0.65	0.47	0.33	0.42	0.64	0.79
2F + Line X	0.69	0.76	0.67	0.58	0.33	0.49	0.64

Table 3. Values of indicators for different scenarios.

We perform the complete access graph analysis on the modified networks and compare the accessibility measures in each case with the values obtained for the original network. The comparison of maximal clique growth in all cases is shown in Fig. 14.

The improvements in the $S^{(MC)}$ indicator are comparable for the 2F and Line X scenarios, and both show considerable improvement in overall network connectedness for values of GTT above 20 minutes. This coincides with the time needed for the maximal clique “outskirts” to reach the nodes that are crossed by the new line. After that the improvement is rapid and the Line X scenario outperforms the 2F scenario considering the maximum GTT . The combination of both modifications to the network amounts to another significant improvement in accessibility levels. The values of all indicators in all scenarios are summarised in Table 3.

Discussion

This study presented a topological formulation of public transport accessibility by focusing on several characteristic subgraphs of the access graph. The proposed approach studies global accessibility levels beyond averaged node-level measures by tracing the emergence and evolution of well-connected regions of the network with increasing maximum allowed travel time. This framework amounts to a re-conceptualization of accessibility that offers several novel insights and opens promising ways for future analysis. Our approach is consistent with the popular notion of the X-minute city³¹ and thereby facilitates related investigations of accessibility using different travel modes and in relation to various amenities.

The maximal clique represents the most stringent condition on connectedness, and is at the same time the most interpretable measure. The high correlations between maximal clique and k -cores measures suggest that k -cores analysis suffices for the presented accessibility analysis, especially the $k50$ -core indicator. This is crucial when applying the analysis to larger networks due to tractability issues as finding maximal cliques is an NP -hard problem. In contrast, algorithms for finding k -cores have polynomial time complexity³². However, this correlation should first be established when applying the analysis to other modalities. Highly effective methods for finding cliques should be adopted for scalability to larger networks (e.g.³³).

Moderate to strong correlations with efficiency and shortest paths strengthen the relevance of the proposed indicators that are in line with current measures. The main strengths of our approach compared to traditional approaches is twofold: first, identifying the subgraphs offers a detailed view of accessibility going beyond simple averages. This offers an interpretable operationalisation of accessibility and allows for informed planning and optimisation of public transport networks. Second, by using the notion of the idealised network on the same set of (geographically positioned) nodes, it offers an approach that scales with the network geographic and topological size.

These findings warrant a rethinking of accessibility through the lens of complex networks. More importantly, and with broader implications, the observations support the need for a methodological shift: rather than adapting existing network metrics to transport concepts, relevant transport properties should guide the development of suitable graph-theoretic measures.

Next to these theoretical implications, the introduction of the access graph and its subgraph-based accessibility measures can serve transport planners in designing highly accessible networks. The proposed methodology allows for simulating changes in infrastructure (e.g. adding new stops or links) as well as service planning (e.g. introducing new lines and modifying service frequencies). The illustration of this approach on the Stockholm metro network offers a preliminary confirmation of its feasibility, and such simulations of alternative developments can guide decision making when planning public transport developments.

In the following, we outline the limitations of this study, and suggest venues for future research. We first note that in transport geography, accessibility is understood in relation to different amenities. In the present framework, we focus on the travel impedance component of accessibility, and simply consider the number of reachable nodes in the PT network as a proxy for number of opportunities for different activities. Connected to the previous point, actual land use or register data are not used in the model, and there is no estimate on the actual number of opportunities or their differentiation. This view of accessibility thus only addresses the efficiency of the existing system. Including the coverage, and number of different opportunities data in terms of node weights seems a promising direction for future work.

In terms of urban PT studies, to gain a more complete picture of accessibility, it will be necessary to consider other PT modes. In this work, we only consider metro systems that have a specific role in the urban PT system: compared to bus or tram networks, metro has high speed, large capacity and direct connections, and carries a significant percentage of passenger flows in multi-modal urban PT systems. However, due to its infrastructure, it has comparatively fewer stations and less spatial coverage than bus networks.

On the travel impedance side, the construction of the access graph is based on the cut-off function and is an undirected unweighted graph. Adding edge weights and using directed edges would give a more realistic view of accessibility. The maximal weighted clique problem would be interesting to apply (e.g.³⁴). In addition, a detailed analysis of the impacts of each component of GTT on accessibility levels will offer a deeper understanding of the interplay between PT network infrastructure and operation on access.

On the methodological side, the maximal clique is the strictest constraint and defines a single core part of the network. This is oblivious to any other well-connected parts of the network. A complementary perspective would be the analysis of all k -cliques, or a deeper analysis of network geometry and the forces shaping the network topology by examining smaller cliques and their interactions (similar to e.g.³⁵). Especially in larger cities, there might exist several smaller well-connected parts, which would also be interesting to see with respect to the idea of the X-minute city. Community detection is another promising alternative to predefined subgraphs.

Next to those mentioned above,

this framework opens several lines of further inquiry. We propose the identification of realistic maximally accessible networks, given the physical constraints as well as incorporating the cost of resources, emissions and energy consumption of operating the network, and accounting for (expected) passenger demand as a promising venue for future research. The consideration of metro systems within multi-modal public transport systems is another promising direction, and a similar analysis for multilayer networks should be performed to assess accessibility of integrated PT networks. In addition, an integrated framework for analysing accessibility and robustness will reveal the interplay between the two properties. Specifically, measuring the difference in accessibility levels in terms of the proposed subgraph indicators under failures on either nodes or links emerges as a promising direction.

Methods

Datasets and software

The access graphs are constructed using the curated datasets of L- and P-space representations of 51 metro networks^{36,37}. In the L-space representation (also known as the Infrastructure space,³⁸), there is an edge if the corresponding stops are consecutive stops on a route. In P-space (also known as the Service space,³⁸), each route is a clique, effectively counting the legs of a journey based on transfers. The edges in L- and P-space are weighted with in-vehicle and waiting times, respectively. Of the 51 networks included in the datasets, the 42 networks that have more than one operating line are considered for analysis. The networks included in our empirical analysis vary significantly in dimension, from $N = 26$ (Kobe) to $N = 421$ (New York).

The code is written in Python. The NetworkX library³⁹ is used for most of the analysis and the NetworKit library⁴⁰ is used for intensive clique computations.

Idealised network

For a reliable and meaningful comparison across networks of varying dimensions, we use the notion of the idealised network on the given set of nodes, each with geographical coordinates (lat, lon)²⁸. The idealised network is the complete graph on the given set of nodes that connects each pair of stops in the network using a straight line.

The idealised network is constructed as follows. First, we calculate the matrix of Haversine distances $[\ell]_{ij}$ between all pairs of nodes. The distances are divided by the average speed \bar{v} , and the average waiting time, equal to half the headway, is added to obtain the idealised generalised travel time GTT^{ideal} :

$$GTT_{ij}^{ideal} = \ell_{ij}/\bar{v} + t^{wait}, \quad (1)$$

We set fixed values of $\bar{v} = 45\text{km/h}$ and $t^{wait} = 5\text{ min}$ - corresponding to a frequency of 12 vehicles per hour - for all networks. The values are chosen based on the typical specifications of the real-world metro systems included in our dataset. The edge weight w_{ij} between the nodes i and j is thus the value of GTT_{ij}^{ideal} .

Comparison with realised networks

In the idealised network, connections of all pairs of nodes can be travelled directly by design. To quantify the deviations from the idealised case and measure the pair-wise discrepancy between the network distance $d_G(i, j)$ and the straight-line distance $d_E(i, j)$, we consider network circuitry, defined as:

$$C_{ij} = \frac{d_G(i, j)}{d_E(i, j)}. \quad (2)$$

The value of circuitry $C = 1$ means that the connection is direct and larger values of circuitry mean larger detours. The average circuitry of the network is the average over all pairs:

$$C = \frac{1}{N(N-1)} \sum_{i,j} C_{ij}. \quad (3)$$

Since the exact shapes of paths and the actual distances travelled are not given in our data, we consider two measures of distance and their corresponding circuitries:

- *Infrastructural circuitry* C^I , where the distances $d_G(i, j)$ and $d_E(i, j)$ represent in-vehicle times in the realised and idealises network, respectively.
- *Operational circuitry* C^O , where the distances $d_G(i, j)$ and $d_E(i, j)$ represent waiting times in the realised and idealises network, respectively.

Access graph construction

The access graph \mathcal{G}_A is constructed by calculating generalised travel times (GTT) between all pairs of nodes (stops) in the network and directly connecting the nodes that are reachable within a given cut-off time t_c . The edge set of \mathcal{G}_A is thus dependent on the external cut-off parameter t_c , gradually increased from 0 to maximum GTT. We use steps of 2 minutes in the analysis.

Generalised travel times are calculated as follows, similar to the methodology presented in⁴¹. The k -shortest paths, with $k = 3$, in the weighted L- and P-space representations are calculated between each pair of nodes i and j using Yen's algorithm⁴². For each of the paths obtained in a given graph representation, the path in the remaining graph representation is matched. For each path l of the resulting $2k$ paths, the generalised travel time is calculated as:

$$GTT_{ij}(l) = d_{ij}^L(l) + w^{wait} d_{ij}^P(l) + w^{transfer} d_{ij}^{Pu}(l) \quad (4)$$

where $d_{ij}^L(l)$ and $d_{ij}^P(l)$ are the weights of the L- and P-space shortest paths, respectively, and $d_{ij}^{Pu}(l)$ is the path length in unweighted P-space. $w^{wait} = 2$ is the waiting time weight and $w^{transfer} = 5\text{ min}$ is the transfer penalty. The values are determined from the literature and reflect average passengers' valuation of time⁴³. The generalised travel time between nodes i and j is then determined as the minimal GTT over the $2k$ paths:

$$d_{ij} = \min_l GTT_{ij}(l) \quad (5)$$

Topological indicators of accessibility

Topological indicators in this study are based on several subgraphs of the access graph:

- *Maximal clique*: a clique is a complete subgraph; maximal clique is the largest such subgraph.
- *k-core*: the largest subgraph where each node has a degree of at least k (within the subgraph). Note that for a given k , the dimension of the k -core might be zero. In the case of the access graph, a k -core for a given k will first occur at a sufficient cut-off time τ_k .

At each t_c -step, the k -cores for $k = \{\lceil 0.25N \rceil, \lceil 0.50N \rceil, \lceil 0.75N \rceil\}$ and maximal clique are computed. For $k = \lceil x\%N \rceil$, the respective subgraph is termed the kx -core. The motivation for this choice follows the goals of public transit service providers where accessibility should reach a certain level of coverage, e.g. 75%. The time-dependent growth of the subgraph H is examined, where $H \in \{\text{maximal clique}, k25\text{-core}, k50\text{-core}, k75\text{-core}\}$. The dimensions of the subgraphs are increasing functions of t_c in the case of access graphs. The dimension of a k -core for a given k becomes non-zero only when a sufficient number of edges are added, i.e. at a certain threshold value of t_c .

For accessibility indicator construction, the evolution of the subgraphs with increasing t_c is observed for both the actual and the idealised network. The value for the actual network is normalised to that of the idealised network. A set of cumulative accessibility indicators $S^{(H)}$ is defined as the integrals under the respective curves $N^{(H)}(t_c)$:

$$S^{(H)} = \frac{\int_0^T N^{(H)}(t_c) dt_c}{\int_0^T N^{(H^{ideal})}(t_c) dt_c} \quad (6)$$

The upper boundary T of the integral is defined as $T = \max(t^{max}, 2 \cdot t_{ideal}^{max})$, where t^{max} is the maximum GTT in the actual network and t_{ideal}^{max} the maximum travel time in its idealised counterpart.

A set of threshold indicators represents the values of the cut-off times τ_{kx} at which the kx -cores first appear, relative to T . By definition, $S^{(H)} \in [0, 1]$ and $\tau_{(xk)} \in [0, t^{max}/T]$.

Comparison with conventional topological indicators

Topological indicators

We investigate how the proposed accessibility indicators relate to standard topological measures. Alongside the number of nodes, the measures related to the shortest paths are chosen since they are closely related to the standard notion of accessibility.

- Average shortest paths (SP).
- Network diameter d : the longest of the shortest paths.
- Network efficiency E : averaged sum of inverse shortest paths d_{ij} between all pairs of nodes i and j :

$$E = \frac{\sum_{\{i,j \neq i\}} \frac{1}{d_{ij}}}{N(N-1)}. \quad (7)$$

All conventional measures are normalised to their idealised network counterparts. For network efficiency, this idea was proposed in²⁸ and elaborated in^{44,45} where the authors considered efficient spread of information in diverse networks, from social and communication to transport networks. The normalised efficiency is defined as:

$$\bar{E} = \frac{E}{E^{ideal}}, \quad (8)$$

where E^{ideal} is the efficiency of the complete weighted graph of the idealised network.

Geographical indicators

Traditional measures of accessibility within transport geography rely on several models, such as cumulative opportunities gravitational models^{46,47}. The cumulative opportunities definition of accessibility of node i is defined as:

$$A_i = \sum_{j=1}^N O_j f(c_{ij}), \quad (9)$$

where O_j is the number of opportunities at location j and c_{ij} is the generalised cost (impedance) of travel between nodes i and j and $f(c_{ij})$ is a function of the cost that can take many different definitions, that are generally decreasing functions of the cost (see e.g.⁶). The simplest, and often used, such function is the cut-off function, defined as:

$$c_{ij} = \begin{cases} 1; t_{ij} \leq t_c, \\ 0; \text{otherwise.} \end{cases} \quad (10)$$

Data availability

The data used to construct the access graphs analysed in this study are publicly available in the references [36, 37]. The code and the access graph data set generated for the purpose of this study are publicly available at Zenodo (<https://doi.org/10.5281/zenodo.18077788>).

Received: 25 July 2025; Accepted: 6 January 2026

Published online: 13 January 2026

References

1. Ingvardson, J. B. & Nielsen, O. A. How urban density, network topology and socio-economy influence public transport ridership: Empirical evidence from 48 European metropolitan areas. *J. Transp. Geogr.* **72**, 50–63 (2018).
2. Hansen, W. G. How accessibility shapes land use. *J. Am. Inst. Plann.* **25**(2), 73–76 (1959).
3. Artime, O. et al. Robustness and resilience of complex networks. *Nat. Rev. Phys.* **6**(2), 114–131 (2024).
4. Van Wee, B. Accessible accessibility research challenges. *J. Transp. Geogr.* **51**, 9–16 (2016).
5. Levinson, D. & Wu, H. Towards a general theory of access. *J. Transp. Land Use* **13**(1), 129–158 (2020).
6. Wu, H. & Levinson, D. Unifying access. *Transp. Res. Part D: Transp. Environ.* **83**, 102355 (2020).
7. von Ferber, C., Holovatch, T., Holovatch, Y. & Palchykov, V. Network harness: Metropolis public transport. *Physica A* **380**, 585–591 (2007).
8. Von Ferber, C., Holovatch, T., Holovatch, Y. & Palchykov, V. Public transport networks: Empirical analysis and modeling. *Europ. Phys. J. B* **68**, 261–275 (2009).
9. Derrible, S. & Kennedy, C. Applications of graph theory and network science to transit network design. *Transp. Rev.* **31**(4), 495–519 (2011).
10. Cats, O. The robustness value of public transport development plans. *J. Transp. Geogr.* **51**, 236–246 (2016).
11. Chopra, S. S., Dillon, T., Bilec, M. M. & Khanna, V. A network-based framework for assessing infrastructure resilience: A case study of the London metro system. *J. R. Soc. Interface* **13**(118), 20160113 (2016).
12. Pagani, A. et al. Resilience or robustness: Identifying topological vulnerabilities in rail networks. *Royal Soc. Open Sci.* **6**(2), 181301 (2019).
13. Cats, O. & Krishnakumari, P. Metropolitan rail network robustness. *Physica A* **549**, 124317 (2020).
14. Xu, Z. & Chopra, S. S. Interconnectedness enhances network resilience of multimodal public transportation systems for safe-to-fail urban mobility. *Nat. Commun.* **14**(1), 4291 (2023).
15. Dimitrov, S. D. & Ceder, A. A. A method of examining the structure and topological properties of public-transport networks. *Physica A* **451**, 373–387 (2016).
16. de Regt, R., von Ferber, C., Holovatch, Y. & Lebovka, M. Public transportation in great Britain viewed as a complex network. *Transportmetrica Transp. Sci.* **15**(2), 722–748 (2019).
17. 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 (2019).
18. Šiligoj, T., Peperko, A., & Cats, O. “Access graph: a novel graph representation of public transport networks for accessibility analysis,” [arXiv:2507.08361](https://arxiv.org/abs/2507.08361). (2025).
19. Zhu, G., Corcoran, J., Shyy, P., Pileggi, S. F. & Hunter, J. Analysing journey-to-work data using complex networks. *J. Transp. Geogr.* **66**, 65–79 (2018).
20. Hajdu, L., Bóta, A., Krész, M., Khani, A. & Gardner, L. M. Discovering the hidden community structure of public transportation networks. *Netw. Spat. Econ.* **20**(1), 209–231 (2020).
21. Yang, H., Le, M. & Wang, D. Airline network structure: Motifs and airports’ role in cliques. *Sustainability* **13**(17), 9573 (2021).
22. Kitsak, M. et al. Identification of influential spreaders in complex networks. *Nat. Phys.* **6**(11), 888–893 (2010).
23. Rajeh, S., Savonnet, M., Leclercq, E. & Cherif, H. Interplay between hierarchy and centrality in complex networks. *IEEE Access* **8**, 129717–129742 (2020).
24. Mouronte-López, M. L. Analysing the vulnerability of public transport networks. *J. Adv. Transp.* **2021**(1), 5513311 (2021).
25. Dai, L., Derudder, B. & Liu, X. Transport network backbone extraction: A comparison of techniques. *J. Transp. Geogr.* **69**, 271–281 (2018).
26. Wang, J., Mo, H. & Wang, F. Evolution of air transport network of China 1930–2012. *J. Transp. Geogr.* **40**, 145–158 (2014).
27. Cui, M. & Levinson, D. Primal and dual access. *Geogr. Anal.* **52**(3), 452–474 (2020).
28. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87**(19), 198701 (2001).
29. Barthélemy, M. Spatial networks. *Phys. Rep.* **499**(1–3), 1–101 (2011).
30. Huang, J. & Levinson, D. M. Circuitry in urban transit networks. *J. Transp. Geogr.* **48**, 145–153 (2015).
31. Logan, T. et al. The x-minute city: Measuring the 10, 15, 20-minute city and an evaluation of its use for sustainable urban design. *Cities* **131**, 103924 (2022).
32. Batagelj, V., & Zaversnik, M. “An o (m) algorithm for cores decomposition of networks,” [arXiv preprint cs/0310049](https://arxiv.org/abs/cs/0310049). (2003).
33. Jain, S., & Seshadhri, C. “The power of pivoting for exact clique counting. In *Proceedings of the 13th international conference on web search and data mining*, pp. 268–276, (2020).
34. Rozman, K., Ghysels, A., Janežič, D. & Konc, J. An exact algorithm to find a maximum weight clique in a weighted undirected graph. *Sci. Rep.* **14**(1), 9118 (2024).
35. Šuvakov, M., Andjelković, M. & Tadić, B. Hidden geometries in networks arising from cooperative self-assembly. *Sci. Rep.* **8**(1), 1987 (2018).
36. Vijlbrief, S., Cats, O., Krishnakumari, P., van, S., Cranenburgh, & R. Massobrio, A curated data set of l-space representations for 51 metro networks worldwide. (2022).
37. Vijlbrief, S., Cats, O., Krishnakumari, P., van, S., Cranenburgh, & Massobrio, R. “A curated data set of p-space representations for 51 metro networks worldwide,” (2022).
38. Luo, D., Cats, O. & van Lint, H. Can passenger flow distribution be estimated solely based on network properties in public transport systems?. *Transportation* **47**(6), 2757–2776 (2020).
39. Hagberg, A., Swart, P.J., & Schult, D.A. “Exploring network structure, dynamics, and function using networkx,” tech. rep., Los Alamos National Laboratory (LANL), Los Alamos, NM (United States), (2008).
40. Staudt, C. L., Sazonovs, A. & Meyerhenke, H. Networkit: A tool suite for large-scale complex network analysis. *Netw. Sci.* **4**(4), 508–530 (2016).
41. Massobrio, R. & Cats, O. Topological assessment of recoverability in public transport networks. *Commun. Phys.* **7**(1), 108 (2024).
42. Yen, J.Y. (1971) Finding the k shortest loopless paths in a network. *management Science* **17**(11) 712–716.
43. Yap, M., Wong, H. & Cats, O. Passenger valuation of interchanges in urban public transport. *J. Public Transp.* **26**, 100089 (2024).

44. Latora, V. & Marchiori, M. Economic small-world behavior in weighted networks. *Europ. Phys. J. B-Condensed Matter Complex Syst.* **32**, 249–263 (2003).
45. Bertagnolli, G., Gallotti, R. & De Domenico, M. Quantifying efficient information exchange in real network flows. *Commun. Phys.* **4**(1), 125 (2021).
46. El-Geneidy, A.M., & Levinson, D.M. Access to destinations: Development of accessibility measures (2006).
47. Kapatsila, B., Palacios, M. S., Gris , E. & El-Geneidy, A. Resolving the accessibility dilemma: Comparing cumulative and gravity-based measures of accessibility in eight canadian cities. *J. Transp. Geogr.* **107**, 103530 (2023).

Acknowledgements

The first and second author were partially supported by ARIS-Slovenian Agency for Research and Innovation (grants P2-0394 (A) and P1-0222, respectively, and grant SN-ZRD/22-27/0510 for both authors).

Author contributions

Conceptualisation: T.Š., O.C., A.P.; Methodology, Investigation: T.Š. and O.C.; Software, Formal analysis and Writing - Original Draft: T.Š.; Writing - Review and Editing: T.Š., O.C., A.P.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to T.Š.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

  The Author(s) 2026