Applied Network Science

RESEARCH   **Open Access**

Topology dependence of on-demand ride-sharing

Debsankha Manik[1,2] and Nora Molkenthin[1,2*]

*Correspondence:
nora.molkenthin@ds.mpg.de
[1]Max Planck Institute for Dynamics
and Self-Organization (MPIDS),
37077 Göttingen, Germany
[2]Chair for Network Dynamics,
Institute for Theoretical Physics and
Center for Advancing Electronics
Dresden (cfaed), Technical
University of Dresden, 01069
Dresden, Germany

Abstract

Traffic is a challenge in rural and urban areas alike with negative effects ranging from congestion to air pollution. Ride-sharing poses an appealing alternative to personal cars, combining the traffic-reducing ride bundling of public transport with much of the flexibility and comfort of personal cars. Here we study the effects of the underlying street network topology on the viability of ride bundling analytically and in simulations. Using numerical and analytical approaches we find that system performance can be measured in the number of scheduled stops per vehicle. Its scaling with the request rate is approximately linear and the slope, that depends on the network topology, is a measure of the ease of ridesharing in that topology. This dependence is caused by the different growth of the route volume, which we compute analytically for the simplest networks served by a single vehicle.

Introduction

The increasing demand for mobility in modern urban, suburban and rural areas presents a wide range of ecological and logistic challenges. While urban areas struggle with traffic jams, air pollution and parking space shortages (Guerreiro 2018; NYC Department of Transportation 2018), rural areas are often unable to provide accessible and frequent public transport. The recent rise of the sharing economy (Belk 2014; Cohen and Kietzmann 2014; Kamargianni et al. 2016; Greenblatt and Shaheen 2015) has brought up ride-sharing as a possible answer to all of these problems. Ride-sharing poses an appealing alternative to personal cars, combining the traffic-reducing ride bundling of public transport with much of the flexibility and comfort of personal cars (Spieser et al. 2014; Zhang and Pavone 2016; Barbosa et al. 2018; Macharis and Keseru 2018; Vazifeh et al. 2018). Intelligent on-demand ride-sharing services are hoped to reduce the ecological footprint associated with individual mobility by dynamically bundling rides together, reducing the amount of vehicles necessary for the same number of rides (Tachet et al. 2017; Santi et al. 2014; Sorge et al. 2015; Sorge 2017).

However, the complex behaviour of such dynamic dial-a-ride problems (DARP) (Berbeglia et al. 2010) is not yet fully understood. Recent studies have examined the dynamical behaviour of specific ride-sharing strategies analytically (Herminghaus 2019;

Molkenthin et al. 2019) or in simulations (Alonso-Mora et al. 2017; Ma et al. 2013; Agatz et al. 2011; Horn 2002). However, the general scaling behaviour, or dependence on street network topology and request patterns are not currently understood. Such an understanding would be necessary to compare different dispatching strategies and network settings and make informed decisions about which dispatching strategy works best for a particular network.

Here we study the effects of the underlying street network topology on the viability of ride bundling analytically and in simulations in the low-density limit by studying the performance of a single vehicle. We find that for finite request rates and vehicles not restricted by capacity, there is always a quasi-stationary regime of operation, varying in waiting time and typical vehicle occupancy. We develop a probabilistic description of the steady-state route length, relying on *route-volume*, a topological characteristic of a network that we define, and use this to derive the equilibrium stop-list length. Based on this we show the scaling of the steady-state stop-list length with the dimensionless request density to be linear with a slope depending on the topology. The dependence of the route-volume on the stop-list length can be approximated explicitly for some simple topologies (ring and star) and numerically otherwise.

We apply this analysis to unweighted real-world street networks. The general layout of urban centers is predominantly grid-like in structure, whereas rural areas appear to be best described as interconnected rings with long stretches of unbranching streets. This leads to the surprising effect that, while the request density of cities tends to be better suited for ride-sharing, the topologies show the opposite trend with rural areas allowing easier bundling. This is particularly important as cities already have well functioning public transport options, which have proven to be impractical in less densely populated areas.

Model

We use a simplified model to study ride-sharing, in which the street network is represented as a simple graph with N nodes. Requests are placed according to a request pattern $P_{i,j,t}$ and serviced by a *single* vehicle, where a request is an ordered pair of a pick-up node i and a drop-off node j drawn from the request pattern at time t. The request pattern used throughout this manuscript is uniform across all node pairs and time to facilitate analytical treatment. Related studies have found, however, that such simplifications only have small impacts on the overall behaviour of the system (Molkenthin et al. 2019). New requests arrive according to a Poisson process with an average time Δt between requests to be included in the route according to a *dispatcher algorithm*. The capacity of the vehicle is presumed to be unlimited to allow for the analysis of a wide range of request rates. Introducing a finite capacity would lead to rejections and thus effectively an upper bound for the request rate, beyond which the system overloads. The dispatcher algorithm checks if the request's pick-up node can be inserted to the existing stop list without incurring any detour. If this is the case it checks if the drop-off node can also be inserted to the existing stop list without any detour, otherwise it is appended at the end of the stop list. If the pickup cannot be inserted with zero detour, then both the pick-up and the drop-off are appended right after each other at the end of the stop list.

We introduce the dimensionless request rate x in order to compare system properties across network topologies

$$x = \frac{2\langle l \rangle}{v\Delta t}, \tag{1}$$

where $\langle l \rangle$ is the average length of the requested ride and v is the bus speed.

A request rate of $x = 1$ means that the vehicle covers a distance of $2\langle l \rangle$ between two requests, where $2\langle l \rangle$ is the expected distance from the endpoint of the route to the new pick-up and from there to the new drop-off. Therefore even without ride-sharing, rides are on average completed within one Δt and even a taxi system would be able to serve them all, operating at maximum capacity. For $x > 1$ the taxi can no longer serve the system and waiting times diverge, in this case the ride-sharing system transports x times as many passengers as the taxi could.

To quantify passenger satisfaction in a ride-sharing system, we investigate the service time t_s, i.e. the time it takes from placing the request until being delivered at the requested drop-off location. The number of planned stops in the system, the stop list length n, on the other hand serves as a measure of performance from the perspective of the system as a whole.

Starting from an empty vehicle in a random position, we subsequently generate random requests with pick-up and drop-off nodes chosen uniformly randomly from the nodes of the graph, that are then included in the route by the dispatcher algorithm and served at constant velocity. This is repeated until the steady state is established and performed on a range of different network topologies (star, ring, grid, city layouts) and request rates ($0 < x < 40$). Results of the simulations on a ring with 10 nodes are shown in Fig. 1. The complete simulation code is available in (Manik 2020).
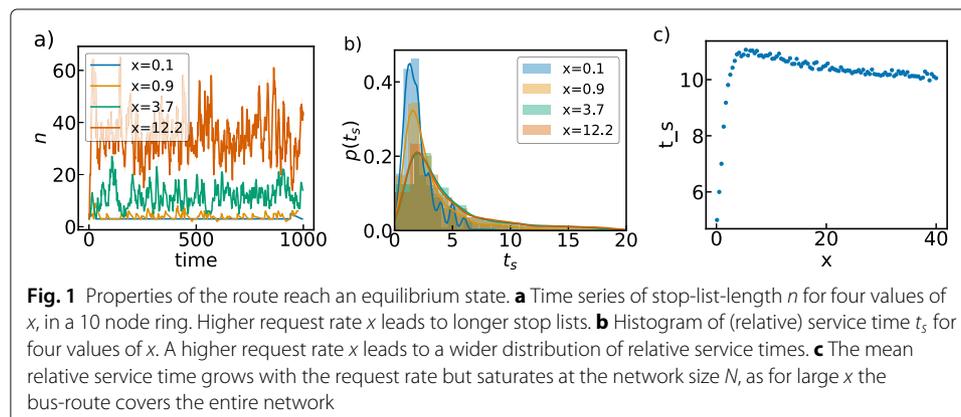
Analytics

We analytically derive approximations for the stop list length n in the steady state, by solving the evolution equation for the length of the planned route after r insertions L_r.

$$L_{r+1} = L_r + l^+(n) - v\Delta t, \tag{2}$$

Where $l^+(n)$ is the average added length per request, $v\Delta t$ is the distance driven in between requests and the discrete time parameter r counts the requests.

The planned route length reaches its equilibrium at

Fig. 1 Properties of the route reach an equilibrium state. **a** Time series of stop-list-length n for four values of x, in a 10 node ring. Higher request rate x leads to longer stop lists. **b** Histogram of (relative) service time t_s for four values of x. A higher request rate x leads to a wider distribution of relative service times. **c** The mean relative service time grows with the request rate but saturates at the network size N, as for large x the bus-route covers the entire network

$$l^+(n) = v\Delta t = \frac{2\langle l \rangle}{x}, \tag{3}$$

where we used the definition of x from Eq. 1.

In case of a taxi system the added length is independent of n at two times the average shortest path length in the network

$$l^+_{taxi} = 2\langle l \rangle, \tag{4}$$

as new segments are simply added to the end. If $x > 1$, the system no longer has an equilibrium as the route length keeps growing, if on the other hand $x < 1$, the taxi has time between subsequent rides, in which it stands still, lowering the average velocity.

In a ride-sharing system with a sensible dispatcher algorithm, however the length of the added segments depends on the planned route. In our model for example, as the current length of the route n grows, the probability that a new request's pick-up and/or drop-off being already included in the route increases, resulting in a smaller $l^+(n)$.

We take the added length to be the average over three possibilities:

a   Both, pick-up and drop-off node are already on the route.
b   The pick-up node is on the route but the drop-off node is not.
c   The pick-up node is not on the route.

In case a) no length is added to the route, in case b) the average added length is $\langle l \rangle$ and in case c) the route gets longer by an average of $2\langle l \rangle$. This means that

$$l^+(n) \approx P_b\langle l \rangle + 2P_c\langle l \rangle, \tag{5}$$

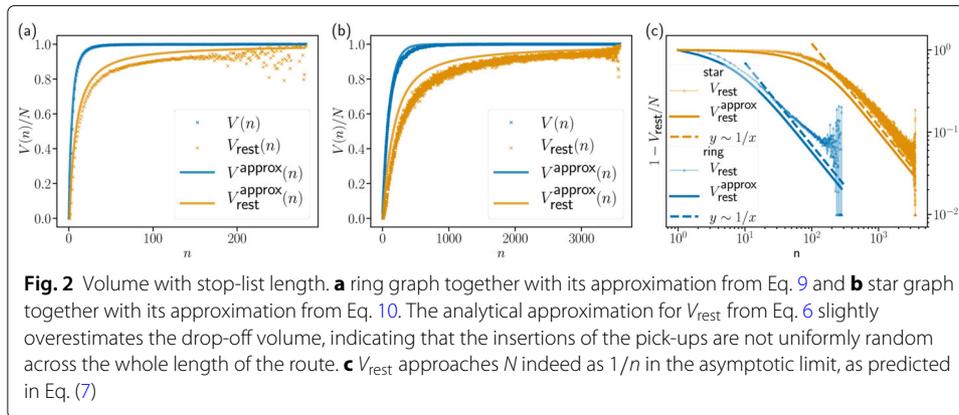where P_b and P_c are the probabilities of b and c respectively.

To evaluate the probabilities, we introduce *route volume V* as the number of nodes that can be reached within the stop list without a detour (illustrated in Fig. 7 in Appendix A). The probability for the requested pick-up node to be on the route depends on the expectation value $\langle V \rangle$ of the volume of the route. The volume depends on the length of the stop-list as well as the topology of the underlying network. As the requested drop-off point has to be on the route after the pick-up, there is a second relevant route volume V_{rest}, and its expectation value $\langle V_{\text{rest}} \rangle(n)$, the average volume of the route after the pick-up. Assuming the position of the pick-up is uniformly randomly located somewhere along the route, the fact that the insertion of the drop-off is always after the pick-up leads to (we employ a simplifying assumption that the pick-up is equally likely to be inserted at any position in the stop-list)

$$\langle V \rangle_{\text{rest}}(n) \approx \sum_{k=1}^{n} \langle V \rangle(n-k)/n = \sum_{k=0}^{n-1} \langle V \rangle(k)/n. \tag{6}$$

We see in Fig. 2a-b that V_{rest} so computed is slightly lower than the actual value. This is because in reality, the pick-up location is slightly more likely to be at the second half of the stop list than the first half. We demonstrate this in Fig. 9 in Appendix A.

Furthermore, we note that the function $\langle V \rangle(n)$ is always monotonously growing with n and asymptotically approaching N. We thus express $\langle V \rangle_{\text{rest}}(n)$ for large n as

$$\langle V \rangle_{\text{rest}}(n) = \frac{1}{n}\sum_{k=0}^{n-1}[N - (N - \langle V \rangle(k))] = N - \frac{1}{n}\sum_{k=0}^{n-1}[N - \langle V \rangle(k)] \approx N\left(1 - \frac{\alpha}{n}\right),$$

$$\tag{7}$$

Fig. 2 Volume with stop-list length. **a** ring graph together with its approximation from Eq. 9 and **b** star graph together with its approximation from Eq. 10. The analytical approximation for V_{rest} from Eq. 6 slightly overestimates the drop-off volume, indicating that the insertions of the pick-ups are not uniformly random across the whole length of the route. **c** V_{rest} approaches N indeed as $1/n$ in the asymptotic limit, as predicted in Eq. (7)

where $\alpha = \sum_{k=0}^{n-1} [1 - \langle V \rangle(k)/N]$, if this limit exists. Note that $\lim_{n \to \infty} 1 - \langle V \rangle(n)/N = 0$, and if in addition we know that $1 - \langle V \rangle(n)/N$ goes to 0 faster than $1/n$, then α is guaranteed to exist. In this case, α is a constant, depending only on the volume growth in a particular network, so $\langle V \rangle_{rest}$ approaches N with n^{-1}. We demonstrate in Fig. 2c that at least for rings and stars, this assumption holds.

Using this, we can express the probabilities for the three insertion types:
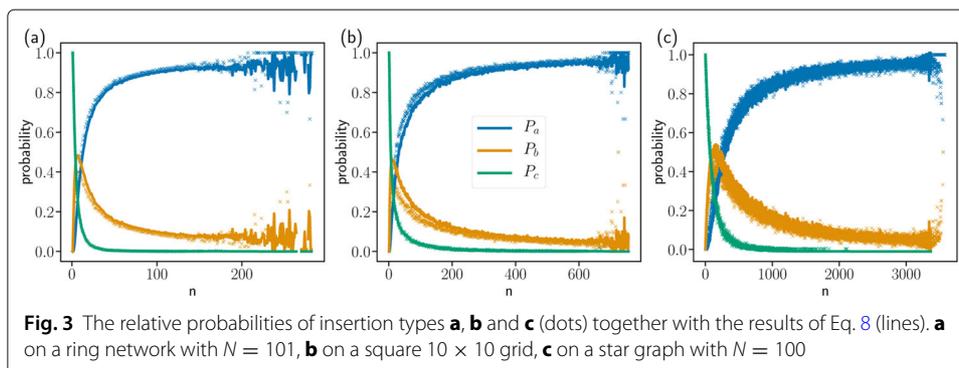
$$
\begin{aligned}
P_a &= \frac{\langle V \rangle(n)}{N} \frac{\langle V \rangle_{rest}(n)}{N} \\
P_b &= \frac{\langle V \rangle(n)}{N} \left(1 - \frac{\langle V \rangle_{rest}(n)}{N} \right) \\
P_c &= 1 - \frac{\langle V \rangle(n)}{N}.
\end{aligned}
\tag{8}
$$

This is shown for a number of different networks in Fig. 3.

In a ring of length N the expected route volume $\langle V \rangle_{ring}(n)$ for a stop list of length n is given by the recursive relation

$$
\langle V \rangle(n+1) = \begin{cases} \frac{N}{4} + \frac{\langle V \rangle(n)}{2} + \frac{\langle V \rangle(n)^2}{3N} & \text{if} \langle V \rangle(n) \le \frac{N}{2} \\ -\frac{N^2}{4\langle V \rangle(n)} + \frac{5N}{4} - \frac{2\langle V \rangle(n)}{3} + \frac{2\langle V \rangle(n)^2}{3N} & \text{if} \langle V \rangle(n) > \frac{N}{2}. \end{cases}
\tag{9}
$$

A detailed derivation (22) is given in the Appendix A. This approximation holds very well as shown in Fig. 2a.
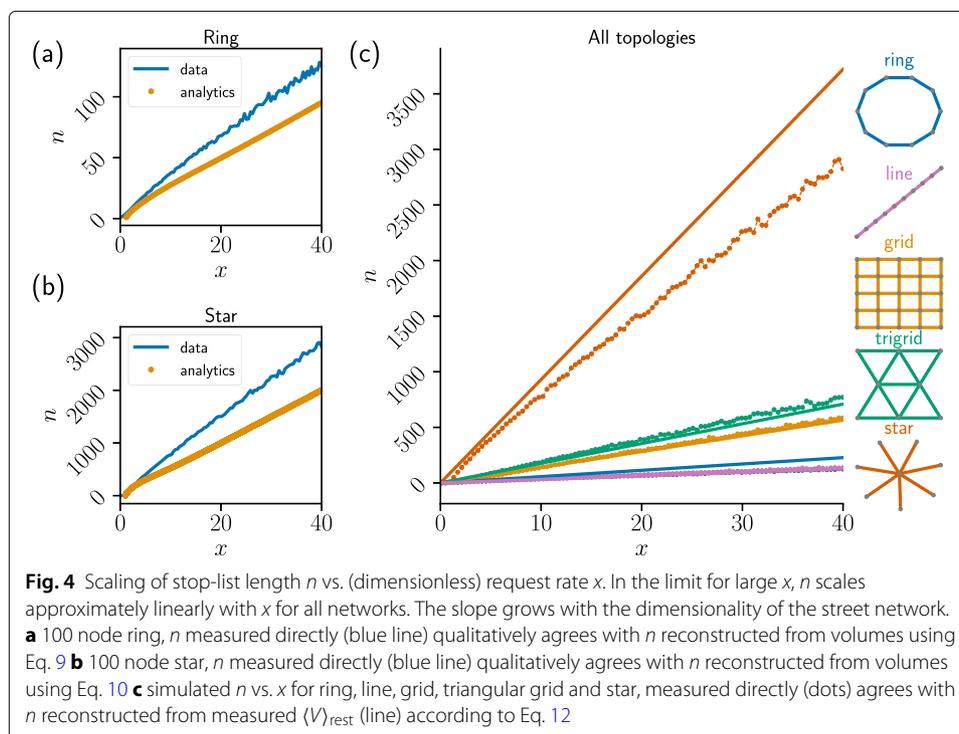

Fig. 3 The relative probabilities of insertion types **a**, **b** and **c** (dots) together with the results of Eq. 8 (lines). **a** on a ring network with $N = 101$, **b** on a square 10×10 grid, **c** on a star graph with $N = 100$

In a star with N nodes, the number of nodes on the route is approximately equal to the number of unique random draws. This is given by

$$\langle V \rangle_{star}(n) \approx N \left(1 - \left(\frac{N-1}{N} \right)^n \right). \tag{10}$$

This approximation does not account for the special role of the center point of the star, which is always on the route, as soon as the stop-list contains two or more nodes. Nonetheless it holds reasonably well, as shown in Fig. 2b.

Results

We have now gathered all necessary input for computing the expected steady-state stop-list length. Inserting the approximated volumes from Eq. 9 and Eq. 10 (or using volumes extracted from the simulation if no such approximation is available) into the probability functions from Eq. 8 and the approximation of the second volume from Eq. 6 to then insert into the steady state added length from Eq. 5 and solving for n, we find an approximately linear rise of the stop list length with the dimensionless request rate x (see Fig. 4a and b). While the general trend is captured well, the slopes of the analytical approximations underestimate the actual value. This is due to the overestimation of $\langle V \rangle_{rest}(n)$ by the assumption of uniformly random insertion of pick-ups in the stop-list, limitations of which assumption is demonstrated in Fig. 9 in Appendix A. The discrepancy therefore largely disappears when measured values are used for $\langle V \rangle_{rest}(n)$, as in Fig. 4c

Fig. 4 Scaling of stop-list length n vs. (dimensionless) request rate x. In the limit for large x, n scales approximately linearly with x for all networks. The slope grows with the dimensionality of the street network. **a** 100 node ring, n measured directly (blue line) qualitatively agrees with n reconstructed from volumes using Eq. 9 **b** 100 node star, n measured directly (blue line) qualitatively agrees with n reconstructed from volumes using Eq. 10 **c** simulated n vs. x for ring, line, grid, triangular grid and star, measured directly (dots) agrees with n reconstructed from measured $\langle V \rangle_{rest}$ (line) according to Eq. 12

Independently of the exact form of $\langle V \rangle (n)$ we can exploit the asymptotic behaviour of $\langle V \rangle (n)$ and set $\frac{\langle V \rangle}{N} \to 1$, leading to (using (3) and (5))

$$
\begin{aligned}
\frac{2\langle l \rangle}{x} = l^+(n) &\approx P_b \langle l \rangle + 2 P_c \langle l \rangle \\
&= \langle l \rangle \left[\frac{\langle V \rangle}{N} \left(1 - \frac{\langle V \rangle_{\text{rest}}}{N} \right) + 2 \left(1 - \frac{\langle V \rangle}{N} \right) \right] \\
&= \langle l \rangle \left(1 - \frac{\langle V \rangle_{\text{rest}}}{N} \right),
\end{aligned}
\tag{11}
$$

where we inserted the probabilities from Eq. 8 into Eq. 5 and set $\frac{\langle V \rangle}{N} \to 1$.

In this we use the expression for α from Eq.7 and solve for x, giving

$$
x \approx \frac{2n}{\alpha},
\tag{12}
$$

where $\alpha = \lim_{n \to \infty} \sum_{k=0}^{n-1} [1 - \langle V \rangle (k)/N]$ is computed from the analytical expressions of $\langle V \rangle (k)$ or directly from simulated $\langle V \rangle$ and $\langle V \rangle_{\text{rest}}$. In Fig. 4c the estimated results for α are inserted in Eq. 12 and plotted with the directly simulated $n(x)$. For each topology and request rate, 10000 requests were simulated, with the origin and destination of each request drawn uniformly randomly from the nodes (Manik 2020).
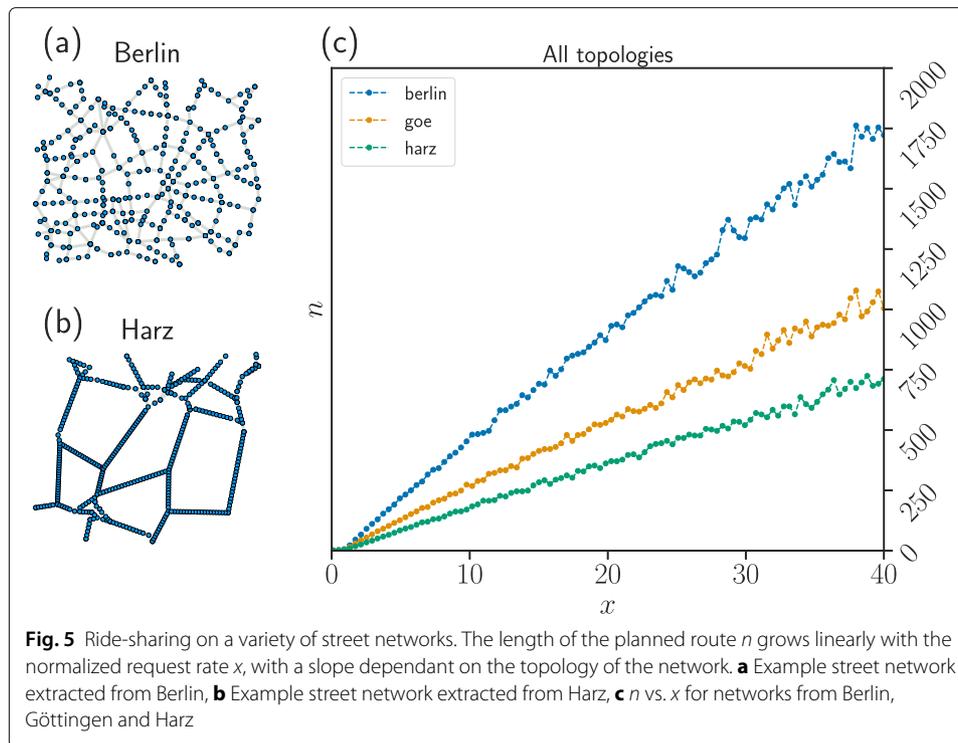
We find values of $\alpha \approx 11.4$ for the ring, $\alpha \approx 6.4$ for the line, $\alpha \approx 28.2$ for the grid, $\alpha \approx 35.4$ for the triangular grid and ≈ 186.0 for the star. The resulting lines capture the behaviour of the curves reasonably well.

The star graph has by far the steepest curve, indicating the worst layout for ride-sharing. This was expected as there is only one point that is on the way while all other nodes are detours. The grids perform slightly better, as there are multiple routes between any two points. The ring and line essentially represent the ride-sharing in an elevator, which works without much route adjustment by simply going up and down and collecting whomever is going in the current direction of the elevator.

This shows that ride sharing on a ring or line is very natural and typically possible, while almost no two distinct rides can be bundled on a graph with star topology (see Fig. 4).

We have applied the methods to real street networks to assess their respective ride-sharing feasibility and compare rural, urban and suburban areas. To this end we have extracted street networks from the freely available open dataset curated by the Open-StreetMap project (OpenStreetMap contributors 2020) using the open-source library OSMnx (Boeing 2017) and translated the weights into a corresponding number of equally long links, since our method is meant for unweighted graphs. First, if two nodes are less than 100 m apart, they are collapsed into one, then all links longer than 400 m were subdivided by adding additional nodes. This procedure ensures that link lengths are approximately uniform. This procedure is further illustrated in Appendix B and Fig. 11. In the resulting graph the distribution of edge lengths are rather sharply peaked around 400 m. For example, in case of the Berlin network in Fig. 5, edge lengths have mean 310.19 m.

These alterations leave the shortest path lengths in the networks largely unaffected. In each of Berlin, Göttingen and Harz, the average shortest path length remains within 20% of the original value (when comparing the original *weighted* average shortest path length in meters against the new *unweighted* shortest path length multiplied by the average link lengths in meters). In urban networks, the original street lengths are largely homogeneous, leading to few added intermediate nodes. In rural areas, on the other hand, street
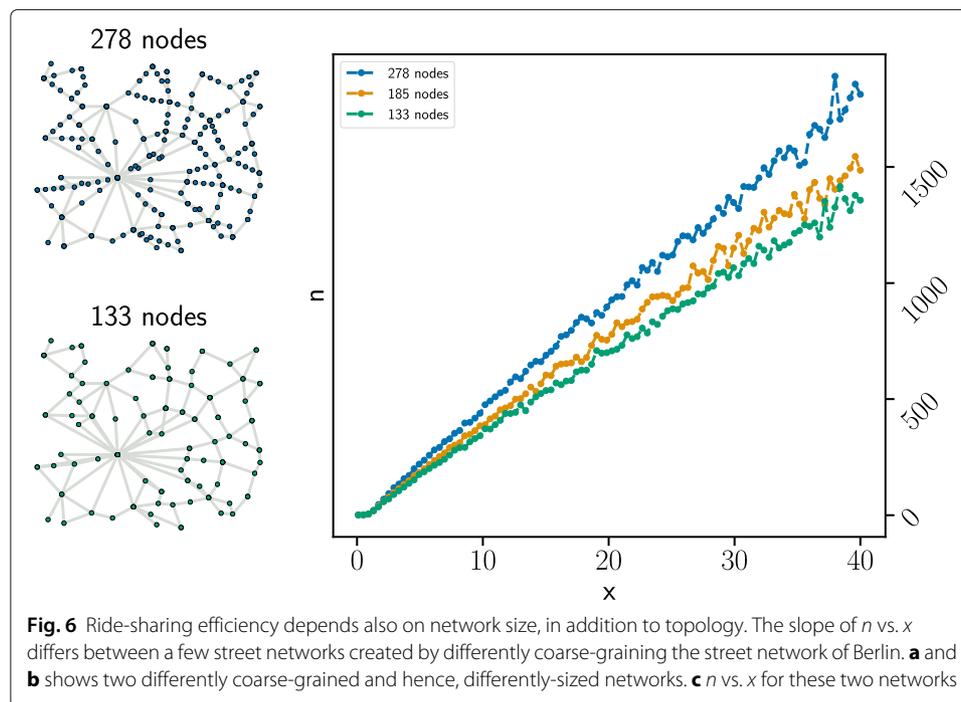
Fig. 5 Ride-sharing on a variety of street networks. The length of the planned route *n* grows linearly with the normalized request rate *x*, with a slope dependant on the topology of the network. **a** Example street network extracted from Berlin, **b** Example street network extracted from Harz, **c** *n* vs. *x* for networks from Berlin, Göttingen and Harz

lengths in between settlements are far larger than those within villages, leading to large numbers of added nodes. As a result, street networks in the city, as shown in Fig. 5a resemble a grid, whereas those in the countryside resemble a loose mesh.

In order to distinguish effects of network size N and the interaction structure on ride-sharing efficiency, we compare networks of equal size and link length. This is done by selecting geographical regions to include comparable numbers of nodes at the same coarse graining criteria: collapsing node pairs that are < 100 m distant, then introducing intermediate nodes along edges to achieve a target edge length of 400 m. By suitably choosing the geographical bounding boxes, we generate networks with ≈ 500 nodes and simulate 10000 requests in each case. The resulting stop-list length is plotted in Fig. 5 over the dimensionless request rate, assuming a uniformly random distribution of requests, as in the case of the artificial networks. We observe linear behaviour with the slope depending on the underlying network, just like we did for synthetic networks in Fig. 4.

In particular we find that the grid-like structure of the large city (Berlin) leads to a far steeper slope than the loose mesh of the rural area (Harz), or a single town surrounded by smaller villages (Göttingen), indicating that the rural topology may be more suitable for ride-sharing, when the request pattern is uniformly randomly distributed.

We furthermore explore the effect of network size, while keeping the network structure constant by comparing ride-sharing between different coarse-graining levels of the same region of Berlin: we collapse all node pairs < 200 meters distant, then introduce intermediary nodes to achieve target edge lengths of 200, 600 and 800 meters. Despite the normalization of the request rate with the average shortest path length, we find the slope to depend on the network size, as well as structure, as shown in Fig. 6. This means if the same geographical area is modelled as a sparser network (i.e. with fewer nodes), it shows higher ride-sharing efficiency than a denser network. This happens because in a sparser

Fig. 6 Ride-sharing efficiency depends also on network size, in addition to topology. The slope of n vs. x differs between a few street networks created by differently coarse-graining the street network of Berlin. **a** and **b** shows two differently coarse-grained and hence, differently-sized networks. **c** n vs. x for these two networks

network, it is more likely that the shortest path between the origin and destination of two different requests will have an overlap. Indeed, the same effect is also encountered in the optimum bus stop design problem: Fewer bus stops lead to higher efficiency (Moura et al. 2012; Schöbel et al. 2009; Nalawade et al. 2016), but this comes at the cost of less customer satisfaction since the average distance to cover for a customer to get to their pick-up location is increased. Various on-demand ride-pooling service providers as well choose to allow pick-up and drop-offs only at selected *virtual bus-stops* (i.e. bundled stop locations, meant to reduce the number of stops and amount of detour for the ride-sharing vehicles) instead of any node in the underlying street network (Pettersson 2019; Racah et al. 2017).

Conclusion and discussion

Here we have tackled the question how network topology affects the feasibility of ride sharing. For this, we have studied the steady states of a one-vehicle ride-sharing system in a range of simple homogeneous networks as well as real regional street network topologies using both analytical and numerical methods and analyzed the backlog (i.e. the stop-list) in the system depending on its load.

We find in Fig. 4 that, while similar networks also result in similar scaling behaviour of the stop-list length (i.e. ride-sharing is almost indistinguishable between the ring and the line or different grids of the same size), there are large differences between networks with different dimensions (i.e. ring (1D), grid (2D) and star (∞D)). Furthermore, there is a substantial impact of the number of nodes N on the ride-sharing predisposition.

We find these differences to be summarized by the cumulative route volume parameter α.

In the light of these findings we have compared a range of very different regional street networks ranging from urban centers to smaller towns and rural areas. Assigning them

the same request density and selecting bounding boxes, such that the networks have comparable distances and node numbers N, we find, that ride-sharing is topologically harder in urban areas, as their grid-like structure causes routes to be more likely to be distinct while the loose mesh, characteristic of rural areas topologically forces ride-sharing on long stretches of the connecting streets. We expect this effect to be even stronger with more realistic request patterns, which would remain largely homogeneous in urban centers, but be centered around settlements in rural areas, effectively lowering the number of active nodes.

This effect is, however, counteracted by the typically inconveniently low request rates for public service options in rural areas. To make use of the beneficial network structure, it would therefore be necessary to convince more customers of participating in shared flexible transport options.

The combination of nearby nodes (coarse graining) may improve ride-sharing and thus deliver convenient, efficient public transport. This may be realized in practice by offering the customers a choice of *virtual bus-stops* (Pettersson 2019; Racah et al. 2017). Cities on the other hand already have an inexpensive and efficient public transport in line-services. Further research is needed to determine how introducing line-services on the most frequented routes would affect or be combined with on-demand ride-sharing.

Outlook

As mentioned above, there seems to be a relationship between ride-sharing efficiency and network dimension. In future works it would be highly valuable to further explore this connection and establish a network measure, that predicts the slope directly from the network topology.

In order to directly apply the results of this research to realistic ride-sharing scenarios, the simplified demand pattern would need to be replaced by more realistic input. Equally one would likely work with several finite sized vehicles that allow small detours, rather than one infinitely large one making zero detour. In the choice of those parameters, intuitions gained from this analysis can be used to make educated guesses for such parameters.
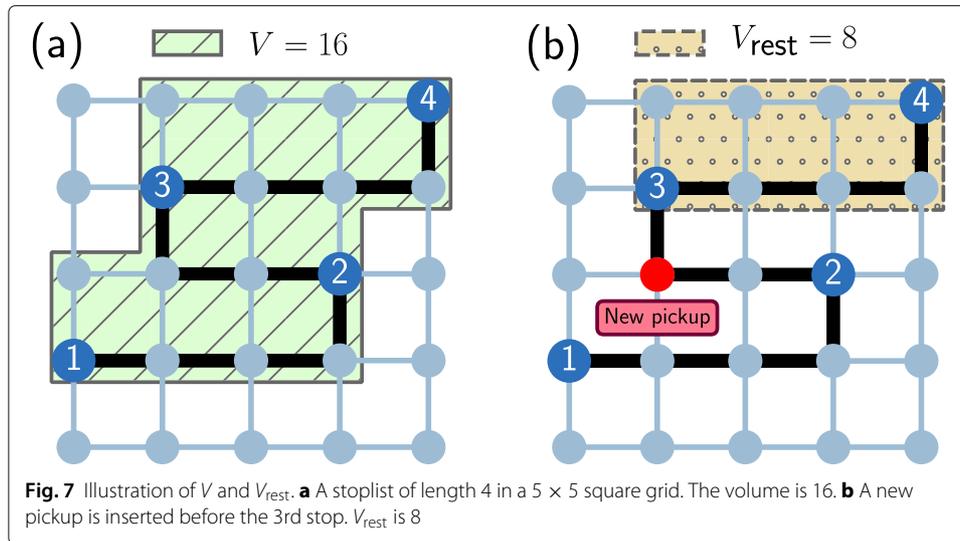
One question in particular to be addressed with such a more detailed model would be an analysis of the gains to be achieved through virtual bus stops in a trade off between efficiency and personal comfort.

Appendix A: Stoplist volume and insertion into a stoplist

We want to know what is the expected (i.e. ensemble average) *stoplist volume* $\langle V \rangle(n)$, when n stops are added to an initially empty stoplist according to the procedure described in "Model" section. The route volume V of a stoplist is the number of nodes that can be reached without taking any detour while traversing it. Given that a pickup point is inserted into the list, V_{rest} is the volume of the rest of the route after the pickup position. This is illustrated in Fig. 7 in Appendix A.

Analytical stoplist volume for rings

Consider an N node ring with nodes labelled $1, 2, \cdots, N$. Let a stoplist of length n be constructed as per "Model" section and V_n be its volume. Our goal is to find out the expectation value $E[V_n] := \langle V \rangle(n)$.

Fig. 7 Illustration of V and V_{rest}. **a** A stoplist of length 4 in a 5×5 square grid. The volume is 16. **b** A new pickup is inserted before the 3rd stop. V_{rest} is 8

We go to the continuum limit here for ease of analytical computations (i.e. we approximate the N node ring with a continuous ring of length N). Further, we do our calculations in phase variables *angular volume* $w_n \in [0, 2\pi]$, instead of in V_n, by defining

$$V_n := \frac{w_n}{2\pi} N. \tag{13}$$

We further assume without loss of generality that the stoplist (with angular volume w_n) has its extremities at 0 and w_n, effectively going to a rotated coordinate system.

We approach the problem by looking at the expectation value of the *increment* in volumes $\Delta_n(w_n, \zeta_n, \zeta_{n+1})$ when a new stop at ζ_{n+1} is added to an n-length stoplist with angular volume w_n whose last stop is at ζ_n:

$$w_{n+1} = w_n + \Delta_n(w_n, \zeta_n, \zeta_{n+1}) \tag{14}$$

$$E[w_{n+1}|w_n] = w_n + E[\Delta_n|w_n]. \tag{15}$$

If we know $E[\Delta_n|w_n]$, a recursive formula can be derived for $E[w_n]$.

The stoplist-generation process described in "Model" section necessarily means the function $\Delta_n(w_n, \zeta_n, \zeta_{n+1})$ will have the form described in Fig. 8.
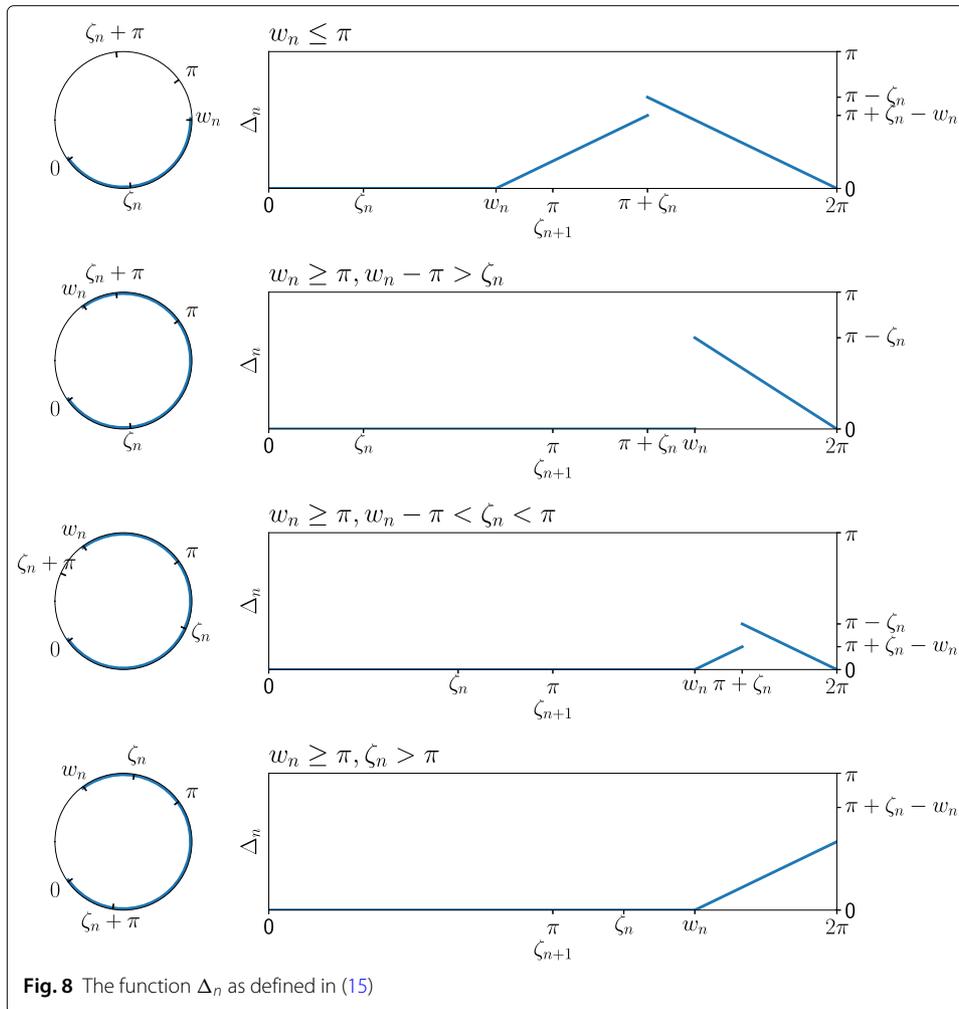
The main difficulty in computing $E[\Delta_n|w_n]$ lies in the fact that although ζ_{n+1} and w_n are independent, ζ_n and w_n are not independent (because ζ_n, the last point on the stoplist, influences the value of w_n). We therefore make out first simplifying assumption: Ignore this dependence and assume that ζ_n is uniformly randomly distributed in the interval $(0, w_n)$, i.e. the last point in the stoplist is equally likely to be found anywhere within the volume. Then we have

$$\zeta_n \sim U(0, w_n) \tag{16}$$

$$\zeta_{n+1} \sim U(0, 2\pi), \tag{17}$$

leading to

$$E[\Delta_n|w_n] = \int_0^{w_n} d\zeta_n \frac{1}{w_n} \int_0^{2\pi} d\zeta_{n+1} \frac{1}{2\pi} \Delta(w_n, \zeta_n, \zeta_{n+1}). \tag{18}$$

Fig. 8 The function Δ_n as defined in (15)

Plugging in the piecewise-linear functional form of Δ_n as described in Fig. 8 into Eq. (18) yields

$$E[\,\Delta_n|w_n] = \begin{cases} \frac{1}{2\pi}\left(\frac{w_n^2}{3} - w_n\pi + \pi^2\right) & \text{if} w_n \leq \pi \\ \frac{2\pi - w_n}{2\pi w_n}\left(-\frac{2w_n^2}{3} + 2\pi w_n - \pi^2\right) & \text{if} w_n > \pi, \end{cases} \tag{19}$$
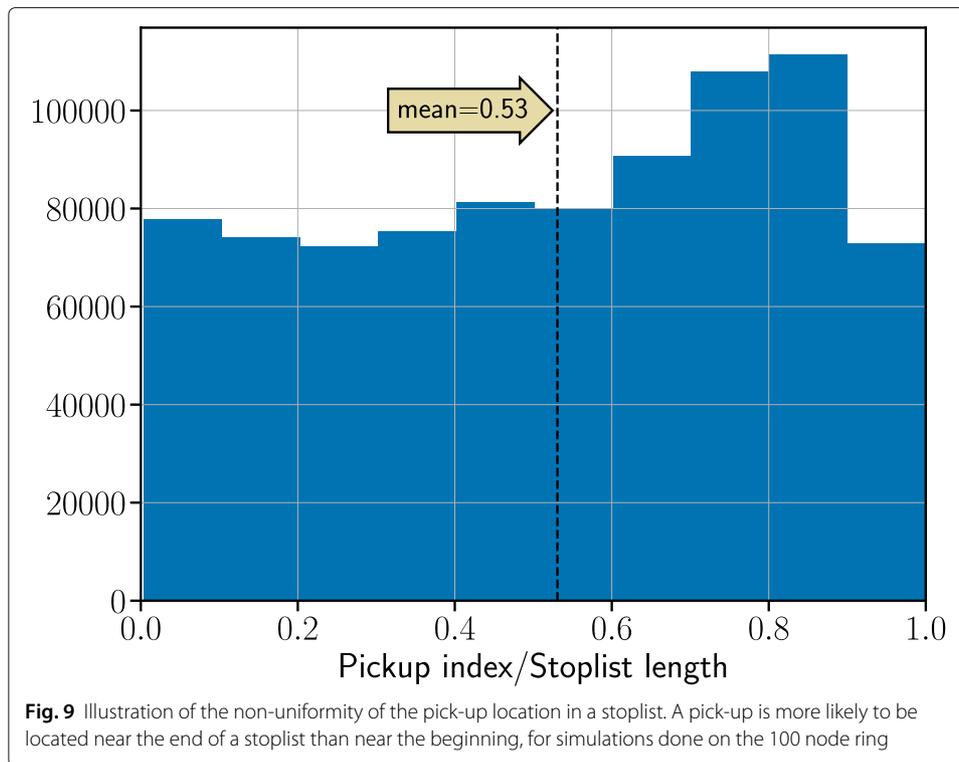
and using (15), we get

$$E[\,w_{n+1}|w_n] = \begin{cases} \frac{\pi}{2} + \frac{w_n}{2} + \frac{w_n^2}{6\pi} & \text{if} w_n \leq \pi \\ -\frac{\pi^2}{w_n} + \frac{5\pi}{2} - \frac{2w_n}{3} + \frac{w_n^2}{3\pi} & \text{if} w_n > \pi. \end{cases} \tag{20}$$

Now we can go back to the discrete V_n from the continuous w_n by using Eq. 13:

$$V(n) := E[\,\upsilon(n)], \tag{21}$$

$$\langle V\rangle(n+1) = \begin{cases} \frac{N}{4} + \frac{\langle V\rangle(n)}{2} + \frac{\langle V\rangle(n)^2}{3N} & \text{if} \langle V\rangle(n) \leq \frac{N}{2} \\ -\frac{N^2}{4\langle V\rangle(n)} + \frac{5N}{4} - \frac{2\langle V\rangle(n)}{3} + \frac{2\langle V\rangle(n)^2}{3N} & \text{if} \langle V\rangle(n) > \frac{N}{2}. \end{cases} \tag{22}$$

Fig. 9 Illustration of the non-uniformity of the pick-up location in a stoplist. A pick-up is more likely to be located near the end of a stoplist than near the beginning, for simulations done on the 100 node ring

Pick-up location in a stoplist

We see in Fig. 2a and b that V_{rest} computed based on the assumption that the pick-up is equally likely to be inserted after any stop in the stop-list, is slightly overestimating the actual value. In reality, however, the pick-up location is slightly more likely to be in the second half of the stop-list, as demonstrated in Fig. 9 in the case of the 100 node ring.

Appendix B: Coarse graining street network and homogenizing the edge lengths

Obtaining openStreetMap data

We developed our analytical results of "Analytics" section for unweighted graphs. In "Results" section, we applied the analytical insights to real world street networks, in addition to simple motif networks (Fig. 4). For this purpose, we used freely available and open data from OpenStreetMap project (OpenStreetMap contributors 2020). We used the open source python library OSMnx (Boeing 2017) to download OpenStreetMap networks for a number of geographical regions including Berlin, Göttingen and the Harz national park in Germany. In each case, we downloaded networks including all streets upto 'secondary' level (OpenStreetMap Wiki Contributors 2020).

Coarse graining

Since the networks so obtained were still quite large, we coarse grained them by merging any two nodes that were closer than 100 meters into a single node. This is illustrated in Fig. 10, panels (a) - (c).
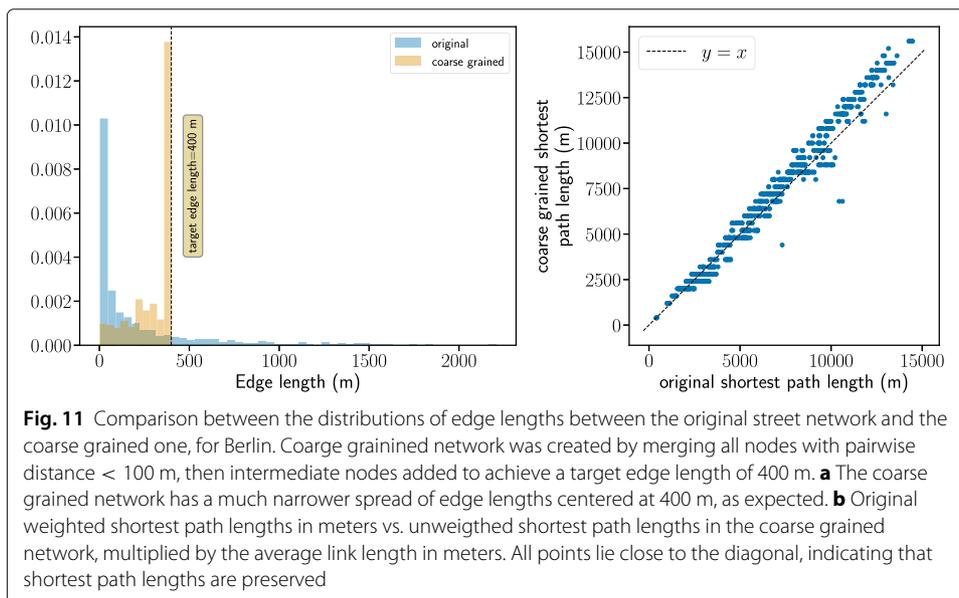
Fig. 10 Illustration of the coarse graining process. **a** The original network. **b** The three nodes in the center are smaller than a predefined buffer. **c** Hence they are merged into a single node. **d** In each of the edges, intermediary nodes are introduced so that most of the new edges have the same edge length. Note that at one end of the original edges, the new edge is slightly smaller than the target length

Converting to unweighted networks

At this stage we had a much smaller network, but the edges were uneven in length, especially between the different regions, making them not comparable in travel times, especially as our ride-sharing algorithm is based on unweighted networks. In order to get a representation of the network in the form of an unweighted network that still preserves the topological features, we introduced intermediary nodes in each of the edges so that the target edge length would be 400 meters. So, if an edge (u, v) had a length of 1000 meters, now two nodes w_1 and w_2 would be inserted between u and v so that the edges $(u, w_1), (w_1, w_2)$ and (w_2, v) have lengths 400, 400 and 200 meters respectively. This step is illustrated in Fig. 10, panels (c) - (d).

Properties of the coarse grained network

The resulting network so obtained (called the "coarse grained network") has similar average shortest path lengths as the original network. In case of Berlin, we find that the average shortest path length is 5859.6 meters for the original network and 5143.4 meters for the coarse grained one.

In Fig. 11, we demonstrate the difference in the distributions of the edge lengths between the original and the coarse grained network. The original network has a wide spread of edge lengths, as well as large number of very short edges. In the coarse grained

Fig. 11 Comparison between the distributions of edge lengths between the original street network and the coarse grained one, for Berlin. Coarge grainined network was created by merging all nodes with pairwise distance < 100 m, then intermediate nodes added to achieve a target edge length of 400 m. **a** The coarse grained network has a much narrower spread of edge lengths centered at 400 m, as expected. **b** Original weighted shortest path lengths in meters vs. unweigthed shortest path lengths in the coarse grained network, multiplied by the average link length in meters. All points lie close to the diagonal, indicating that shortest path lengths are preserved

network, most of the short edges are gone, and the vast majority of the edges have the desired target edge length of 400 (see Converting to unweighted networks). At the same time Fig. 11b demonstrates that distances between nodes are preserved by the coarse graining procedure.

References
Agatz N, Erera AL, Savelsbergh MW, Wang X (2011) Dynamic ride-sharing: A simulation study in metro atlanta. Procedia-Soc Behav Sci 17:532–550
Alonso-Mora J, Samaranayake S, Wallar A, Frazzoli E, Rus D (2017) On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc Natl Acad Sci 114:201611675. National Acad Sciences
Barbosa H, Barthelemy M, Ghoshal G, James CR, Lenormand M, Louail T, Menezes R, Ramasco JJ, Simini F, Tomasini M (2018) Human mobility: Models and applications. Phys Rep 734:1–74. Human mobility: Models and applications
Belk R (2014) You are what you can access: Sharing and collaborative consumption online. J Bus Res 67(8):1595–1600
Berbeglia G, Cordeau J-F, Laporte G (2010) Dynamic pickup and delivery problems. Eur J Oper Res 202(1):8–15
Boeing G (2017) Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput Environ Urban Syst 65:126–139
Cohen B, Kietzmann J (2014) Ride on! mobility business models for the sharing economy. Organ Environ 27(3):279–296
Guerreiro C (2018) Air quality in Europe - 2018 report. Publications Office of the European Union, Luxembourg. ISBN:978-92-9213-990-2
Greenblatt JB, Shaheen S (2015) Automated vehicles, on-demand mobility, and environmental impacts. Curr Sustain/Renew Energy Rep 2(3):74–81
Herminghaus S (2019) Mean field theory of demand responsive ride pooling systems. Transp Res A Policy Pract 119:15–28
Horn ME (2002) Fleet scheduling and dispatching for demand-responsive passenger services. Transp Res C Emerg Technol 10(1):35–63
Kamargianni M, Li W, Matyas M, Schafer A (2016) A critical review of new mobility services for urban transport. Transp Res Procedia 14:3294–3303
Ma S, Zheng Y, Wolfson O (2013) T-share: A large-scale dynamic taxi ridesharing service. In: 2013 IEEE 29th International Conference on Data Engineering (ICDE). IEEE, Brisbane. pp 410–421. https://doi.org/10.1109/ICDE.2013.6544843
Macharis C, Keseru I (2018) Rethinking mobility for a human city. Transp Rev 38(3):275–278. https://doi.org/10.1080/01441647.2018.1423612
Manik D (2020) Topology Dependence of Ridesharing: Simulations in Multiple Networks (Version 1.1). http://doi.org/10.5281/zenodo.3613791
Molkenthin N, Schröder M, Timme M (2019) Topological universality of on-demand ride-sharing efficiency. https://arxiv.org/abs/1908.05929
Moura JL, Alonso B, Ibeas Á., Ruisánchez FJ (2012) A two-stage urban bus stop location model. Netw Spat Econ 12(3):403–420
Nalawade DB, Nagne AD, Dhumal RK, Kale K (2016) Multilevel framework for optimizing bus stop spacing. IJRET Int J Res Eng Technol 5:298–304
NYC Department of Transportation (2018) Mobility Report. http://www.nyc.gov/html/dot/downloads/pdf/mobility-report-2018-screen-optimized.pdf. Accessed 02 Aug 2020
OpenStreetMap contributors (2020) OpenStreetMap Project. https://wiki.openstreetmap.org/wiki/About_OpenStreetMap. Accessed 02 Aug 2020

OpenStreetMap Wiki Contributors (2020) OpenStreetMap Wiki: Highway. https://wiki.openstreetmap.org/wiki/Key:highway#Roads. Accessed 02 Aug 2020

Pettersson F (2019) An international review of experiences from on-demand public transport services. K2 Work Pap. http://www.k2centrum.se/sites/default/files/fields/field_uppladdad_rapport/on-demand_pt.pdf

Racah Y, Sheba A, Shoval O, Ramot D, Marcovitch S (2017) Continuously updatable computer-generated routes with continuously configurable virtual bus stops for passenger ride-sharing of a fleet of ride-sharing vehicles and computer transportation systems and computer-implemented methods for use thereof. US Patent 9,562,785

Santi P, Resta G, Szell M, Sobolevsky S, Strogatz SH, Ratti C (2014) Quantifying the benefits of vehicle pooling with shareability networks. Proc Natl Acad Sci 111(37):13290–13294. https://doi.org/10.1073/pnas.1403657111

Schöbel A, Hamacher HW, Liebers A, Wagner D (2009) The continuous stop location problem in public transportation networks. Asia-Pac J Oper Res 26(01):13–30

Sorge A (2017) Towards a statistical physics of collective mobility and demand-driven transport. PhD thesis, Georg-August-Universität Göttingen

Sorge A, Manik D, Herminghaus S, Timme M (2015) Towards a unifying framework for demand-driven directed transport (d3t). In: Proceedings of the 2015 Winter Simulation Conference. IEEE Press, Huntington Beach. pp 2800–2811. ISBN:9781467397414

Spieser K, Treleaven K, Zhang R, Frazzoli E, Morton D, Pavone M (2014) Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in singapore. In: Meyer G, Beiker S (eds). Road Vehicle Automation. Lecture Notes in Mobility. Springer, Cham. pp 229–245. https://doi.org/10.1007/978-3-319-05990-7_20

Tachet R, Sagarra O, Santi P, Resta G, Szell M, Strogatz S, Ratti C (2017) Scaling law of urban ride sharing. Sci Rep 7:42868

Vazifeh MM, Santi P, Resta G, Strogatz SH, Ratti C (2018) Addressing the minimum fleet problem in on-demand urban mobility. Nature 557(7706):534

Zhang R, Pavone M (2016) Control of robotic mobility-on-demand systems: a queueing-theoretical perspective. Int J Robot Res 35(1-3):186–203

Publisher's Note