nature computational science

Article

https://doi.org/10.1038/s43588-022-00318-w

Demand-driven design of bicycle infrastructure networks for improved urban bikeability

Received: 13 January 2022

Accepted: 12 August 2022

Published online: 3 October 2022

Check for updates

Christoph Steinacker[1] ✉, David-Maximilian Storch[1], Marc Timme[1,2] and Malte Schröder[1] ✉

Cycling is crucial for sustainable urban transportation. Promoting cycling critically relies on sufficiently developed infrastructure; however, designing efficient bike path networks constitutes a complex problem that requires balancing multiple constraints. Here we propose a framework for generating efficient bike path networks, explicitly taking into account cyclists' demand distribution and route choices based on safety preferences. By reversing the network formation, we iteratively remove bike paths from an initially complete bike path network and continually update cyclists' route choices to create a sequence of networks adapted to the cycling demand. We illustrate the applicability of this demand-driven approach for two cities. A comparison of the resulting bike path networks with those created for homogenized demand enables us to quantify the importance of the demand distribution for network planning. The proposed framework may thus enable quantitative evaluation of the structure of current and planned cycling networks, and support the demand-driven design of efficient infrastructures.

Human mobility critically depends on the existing infrastructure underlying it[1,2]. The transition to more sustainable mobility in particular requires a sufficiently developed infrastructure to promote, for example, cycling over motorized mobility for short and medium-distance intra-urban trips[3,4]. During the COVID-19 pandemic, a number of cities such as Paris, New York and Bogotá pushed to open more street space to cyclists, expanded the size of side-walks, or blocked car traffic to enable social distancing[5]. Similarly, many cities have vouched to invest into cycling infrastructure[6,7] and are gaining increasing support among the population for these investments.

In general, designing suitable and efficient infrastructure networks constitutes an intricate problem as the networks are subject to multiple, often opposing technical, economic and social constraints[8–10]. Examples for efficient network structures can be found in various biological[11,12] and social networks[13,14], balancing resource and energy costs with efficient physical or information transport and robustness to failures. For mobility and infrastructure networks, examples of efficient topologies include the core–periphery structure of air travel networks, balancing the cost of direct flights with the inconvenience of transfers[11,15–17], and the emergent backbone structure in street networks of cities[18,19].

For the design of bike path networks, three major constraints include: (1) budget constraints, which limit the total length of bike paths due to, for example, construction or maintenance cost[15,20]; (2) bike path networks have to support the mobility demand and enable fast travel between frequented locations without large detours[15,16,18]; (3) bike path networks should also enable safe travel of cyclists along highly frequented routes[21]. Different implementations of bike path infrastructure weigh these constraints differently. For example, colored bike lanes on the street do not cost much but only slightly protect cyclists, whereas physically separated bike lanes greatly improve cyclists' safety but require more space and investment[22,23]. From a network structure

[1]Chair for Network Dynamics, Center for Advancing Electronics Dresden (cfaed) and Institute for Theoretical Physics, Technische Universität Dresden, Dresden, Germany. [2]Lakeside Labs, 9020 Klagenfurt, Austria. ✉e-mail: christoph.steinacker@tu-dresden.de; malte.schroeder@tu-dresden.de

https://doi.org/10.1038/s43588-022-00318-w

perspective, each of these aspects is simple to understand individually. For example, a connected bike path network in a city that minimizes the required budget is simply given by the minimum spanning tree of the underlying street network[17,19]. More sophisticated approaches to find connected network structures have recently been proposed. Such approaches, based on percolation processes[24–26], optimize the connectivity of the bike path network; however, connectivity alone is not sufficient to support demand, as routes along streets equipped with bike paths would likely be indirect and require large detours. Shortest path trees would optimally support the demand only from and to a single location. By contrast, direct routes between other locations would require cycling along streets without dedicated bike infrastructure and would not be as safe[27,28]. Finally, the safest and most convenient network for cyclists—in which a bike path exists along every street—would naturally exceed any reasonable budget constraints[20,29]. However, efficient bike path networks have to simultaneously adhere to all three constraints to enable both convenient and safe travel with feasible investment[30,31].

In this article we propose a framework for constructing a family of efficient bike path networks, all adapted to the given street network and demand distribution. The algorithm realizes inverted network growth: based on a simplified cyclist routing model and starting from a network fully equipped with bike paths, the algorithm generates a sequence of bike path networks by successively removing bike paths from the street segments with the least impact given the usage patterns in the current network. We observe that both convenience and safety of cycling in the network remain high, even if only a small part of the street network remains equipped with bike paths. Application of the algorithm to synthetic, homogeneous demand conditions enables us to quantify the importance of the cycling demand for the structure of the resulting bike path networks. The proposed framework is extendable to include different routing models and its applicability to different street networks and demand distributions may support planning of bike path networks that promote a desired cycling demand.

Results

Cyclist route choice model

The benefit of bike paths fundamentally depends on their usage and in turn on the routes of cyclists. We map cyclists' route choices to a shortest-path problem on a preference graph $G = (V, E)$ with $N = |V|$ number of nodes (intersections) and $M = |E|$ number of edges (street segments). We derive the preference graph G from the physical street network G^{street}. Both graphs share the same set of nodes V. Each edge $e_{ij} \in E$ in the cyclist preference graph represents a street segment that connects intersections $i, j \in V$ and is assigned a perceived distance

$$l_{ij} = l_{ij}^{\text{street}}\, p_{ij}. \quad (1)$$

Here, l_{ij}^{street} denotes the physical length of the corresponding street segment e_{ij}^{street} in the street network and $p_{ij} \in \{p_{ij}^{B}, p_{ij}^{0}\}$ is a penalty factor summarizing cyclists preferences against riding along the street segment e_{ij}. The set of street segments equipped with bike paths $E_B \subseteq E$ contains street segments $e_{ij} \in E_B$ without distance penalty $p_{ij}^{B} = 1$. Street segments that are not in this set $e_{ij} \notin E_B$ have penalty factors $p_{ij}^{0} > 1$. The value of these penalty factors p_{ij}^{0} may depend on different characteristics of the individual street segments, representing the perceived safety or convenience for cycling.

Adopting this perspective of a cyclist preference graph, we take cyclists to choose their route based on the shortest path $\Pi^{*}_{i\to j} = \operatorname{argmin}\left[L_{i\to j}(\Pi_{i\to j})\right]$ between their origin i and destination j, minimizing the perceived trip distance

$$L_{i\to j}(\Pi_{i\to j}) = \sum_{e\in\Pi_{i\to j}} l_e \quad (2)$$

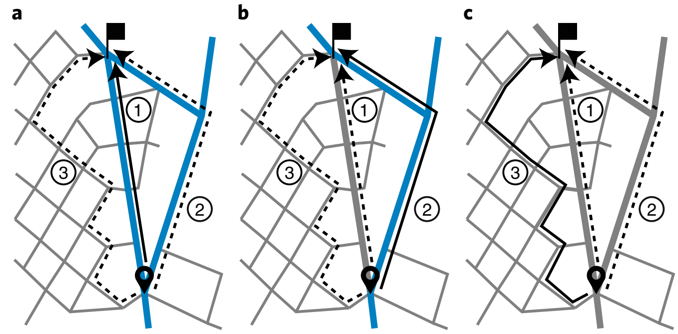

Fig. 1 | Cyclists' route choices balance speed and safety. a, If all major streets (thick edges) are outfitted with dedicated bike paths (thick blue lines), cyclists choose the most direct route (1, solid black arrow) from their origin (pin) to a destination (flag) over alternative paths (dashed arrows). **b**, If only some major streets are equipped with a bike path, cyclists avoid busy roads without a bike path (thick gray lines) and may prefer a short detour (2, solid black arrow). **c**, If none of the streets have dedicated bike infrastructure, cyclists balance the distance and safety of their route choices and may prefer long detours (3, solid black arrow) via low-traffic residential streets (thin gray lines) to more direct routes with high car traffic.

over their potential paths $\Pi_{i\to j}$. Effectively, cyclists choose the most direct path to keep the physical distance of their trip as short as possible but accept detours to avoid busy streets and use bike paths or low-traffic residential streets as alternative routes (Fig. 1).

This simplified route choice model enables efficient calculation of route choice decisions, particularly compared with more complex stochastic models[32,33]. To illustrate the concept, we focus here on the effect of the street type and take the penalties p_{ij}^{0} of a street segment to depend on the volume of car traffic on the respective segment, where higher penalties correspond with larger car traffic volumes (that is, a lower perceived safety or convenience; see Methods for details). In principle, the approach can be extended to include additional factors (see Supplementary Note 1 for a brief discussion) such as slopes, (left) turns or crossings by appropriately modifying the cyclist preference graph (for example, adding more edges with a penalty for left turns) as well as more complex route choice models or heterogeneous preferences among cyclists[27,28].

Network generation

We describe the bike path network of a city as a subgraph $G_B = (V, E_B) \subseteq G^{\text{street}}$ of the city's street network, in which each street segment $e_{ij} \in E$ may ($e_{ij} \in E_B$) or may not ($e_{ij} \notin E_B$) be equpped with a bike path. Even in this simple binary model, the number of possible bike path networks G_B scales exponentially with the number M of edges in the street network, as each street segment may or may not be equipped with a bike path (there are thus 2^M possible subgraphs). Testing all of these networks is impossible for real-world cities in reasonable time (see Supplementary Note 2 for a more detailed description of the underlying optimization problem). Recent approaches utilize forward network percolation models to construct bike path networks[24,26] or apply percolation models to a fixed cyclist flow[25] to find efficient networks.

Here we employ a complementary approach that follows the idea of pruning links from a network—as previously employed in network community-detection algorithms[34] and to study the structure of aviation networks[16]. Specifically, we create a sequence $\{G_B(M')\}_{M'}$ of bike path networks where $M' \in \{0, 1, \ldots, M\}$ street segments are outfitted with a bike path (see Fig. 2): we start from an optimal bike path network in which every street segment is equipped with a bike path ($E_B(M) = E$) such that there is no penalty for any street segment ($p_{ij} = p_{ij}^{B} = 1$ for all edges). We then compute the route choice decisions of the cyclists in their preference graph G (as described above) on the basis of their

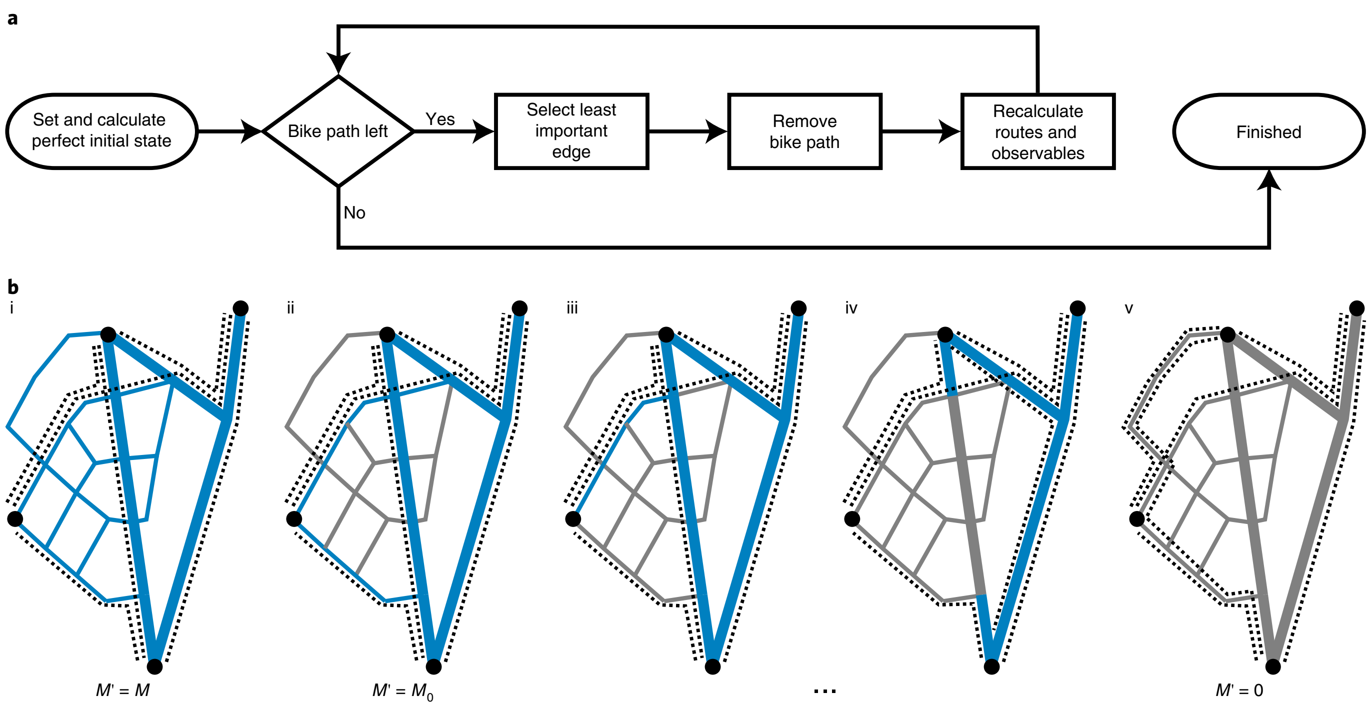

Fig. 2 | Constructing a sequence of efficient bike path networks. **a**, Block diagram of the algorithm. **b**, Illustration of bike path networks $G_B(M')$ with different numbers M' of street segments equipped with bike paths. Edges represent major busy streets (thick lines) or minor residential streets (thin lines), and whether the street is equipped with a bike path (blue) or not (gray). The black dotted lines indicate cyclists' route choices. (i) We start from a full bike path network $G_B(M) = G$, in which all M street segments of the network G are equipped with a bike path ($M' = M$). (ii) We first remove all of the bike paths that are not used by any cyclists, $n_{ij} = 0$, leaving us with the smallest subgraph $G_B(M_0)$ that still optimally serves the given demand. (iii, iv) We then remove the least important edges one by one, defined by the smallest product $p^0_{ij}\, n_{ij}$ of the penalty factor and the number of cyclists using the bike path, updating the cyclists' route choice and recording one network $G_B(M')$ for each number M' of bike paths. (v) The algorithm terminates with an empty bike path network $G_B(0)$ once all of the bike paths have been removed.

demand distribution $n_{i\to j}$, which denotes the number of cyclists traveling from nodes i to j. To construct the family of bike path networks, one by one we remove the least important bike path $e^*_{ij}(M')$ from the network, $E_B(M'-1) = E_B(M') \setminus \{e^*_{ij}(M')\}$, adjusting the penalty of that street segment from $p_{ij} = p^B_{ij} = 1$ to $p_{ij} = p^0_{ij} > 1$ in the cyclist preference graph G. We quantify the importance of a bike path $e_{ij} \in E_B(M')$ in the current state of the bike path network (with M' remaining bike paths) as the product $p^0_{ij}\, n_{ij}(M')$ of the penalty p^0_{ij} (if the street had no bike path) and the number of cyclists that using that street segment $n_{ij}(M')$. The product represents the graph-theoretical weighted betweenness centrality of the edge in the cyclist preference graph. This approach minimizes the negative impact of each removed bike path on the perceived distance of the cyclists in the current bike path network. After each change to the cyclist preference graph G, we update the route choice decisions of the cyclists, ensuring that the algorithm continually adapts to the cycling demand given the currently available set of bike paths $E_B(M')$. The process terminates with an empty bike path network $G_B(0) = (V, \varnothing)$ once all of the bike paths have been removed.

See Supplementary Note 3 for a discussion on the computational runtime of the network generation.

In contrast to iteratively adding bike paths to an initially empty graph and building on the suboptimal cycling routes in networks with few bike paths, this procedure creates bike path networks adjusted to ideal cycling conditions; for example, it keeps bike paths that may not be important in the perfect network if the cyclists start to use them more heavily as the other bike paths are removed (see Supplementary Note 4 for details).

Inputs to our algorithm are: (1) the street network G^{street}, (2) the penalty factors p^0_{ij} for each street segment not equipped with a bike path, (3) the demand distribution $n_{i\to j}$ and (4) the cyclists' route choice model. These parameters may either be as-is empirical values, or planned/desired ideal values (for example, describing the desired or predicted demand for cycling in a city). The latter application might be particularly relevant for planning bike path network extensions if urban quarters develop or are repurposed.

Application

We test the proposed algorithm using data from two German cities: Dresden and Hamburg (see Fig. 3). We take the street networks of both cities from OpenStreetMap (OSM)[35]—using the street classification as a proxy for their expected traffic load—as input data; we also take data from local bike-sharing services to model the cycling demand. We fix the penalty factors against those of physically protected bike path infrastructure based on the street type classification decoded in OSM (see the Methods for a detailed description of the data).

The two cities are representatives of two archetypes of local demand constellations: spatially homogeneous all-to-all demand and confined few-to-few demand. Bike-sharing usage patterns in Hamburg indicate a local demand structure that refers to the first archetype (see Fig. 3c,d), whereas corresponding data for Dresden hint at the latter archetype, which is reflected by the dominance of trips between the university and main train station (see Fig. 3a,b).

Algorithmic generation of bike path networks. We generate families of bike path networks $\{G_B(M')\}_{M'}$ for both cities. We chose a network in which all primary and secondary (P + S) street segments (as per their classification in OSM) are equipped with bike paths and compare this network to our generated network with the same total length

https://doi.org/10.1038/s43588-022-00318-w

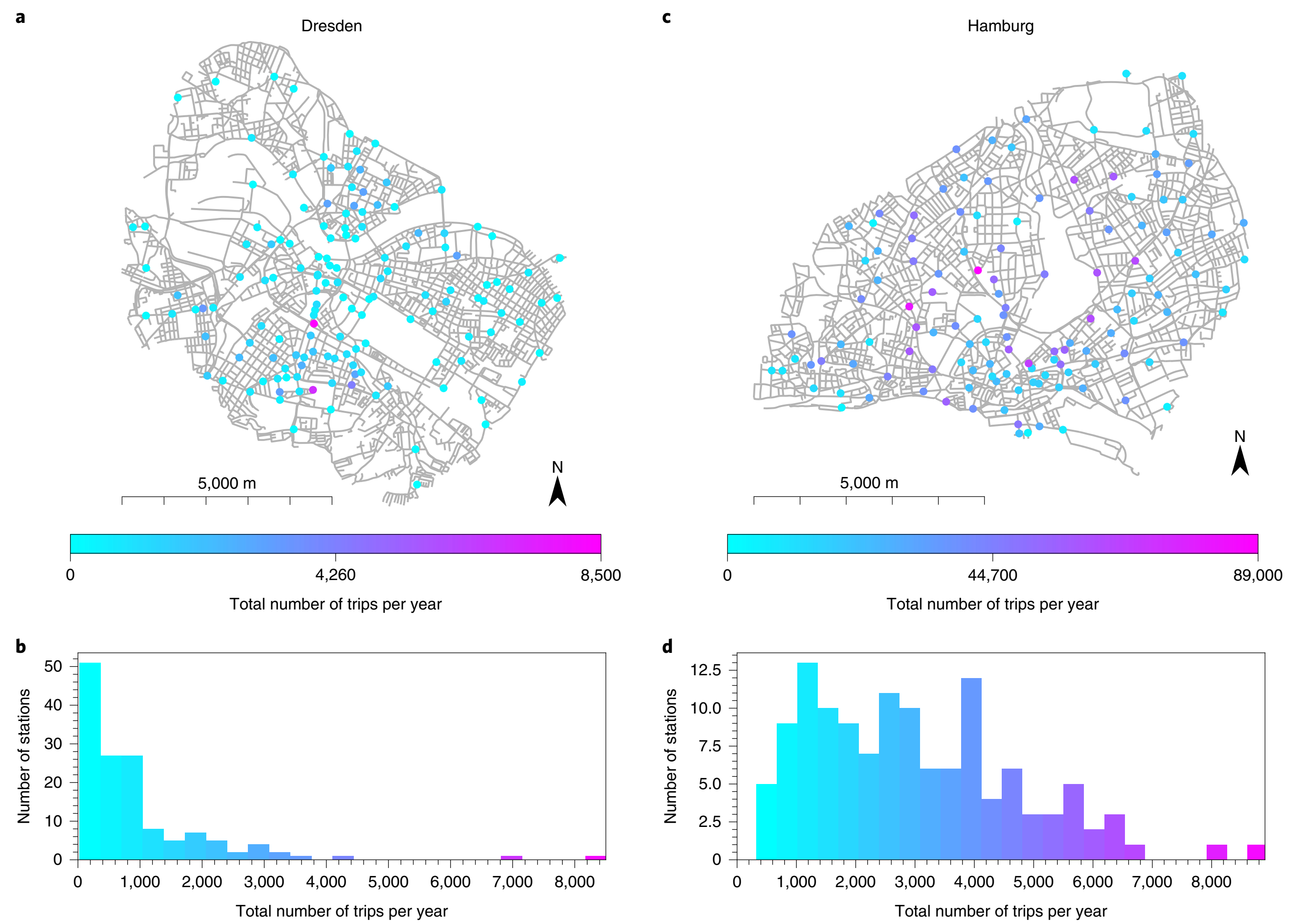

Fig. 3 | Street networks and bike-sharing demand in Dresden and Hamburg. **a**, Street network (gray lines), bike-sharing-station locations and station activity (colored circles) in Dresden between November 2017 and March 2020[44]. **b**, Distribution of station usage, measuring the combined number of in- and outgoing trips per station in Dresden. Bike-sharing usage is strongly heterogeneous and is dominated by two heavily used stations (pink) along the north–south axis between the central train station (center) and university campus (south). Station density reflects this usage pattern and is highest in the central city (north/center) and near the university campus (south). **c**, Street network, bike-sharing-station locations and station activity in Hamburg between January 2014 and May 2017[45]. **d**, Distribution of station usage, measuring the combined number of in- and outgoing trips per station in Hamburg. The station activity distribution is homogeneous across a broad spectrum of total number of trips. This homogeneous usage is also reflected in a more homogeneous distribution of bike-sharing stations, which is slightly denser only in the inner city (south).

$\Lambda(M') = \sum_{e\in E_B(M')} l_e^{\text{street}}$ of bike paths such that $\Lambda(M') = \Lambda_{\text{P+S}}$. Taking the installation and maintenance cost of the bike path network proportional to its length, we thus compare networks with the same budget. Due to some antiparallel one-way streets, our algorithm may place slightly more bike paths along other streets than effectively exist in the P + S network. As we assume bidirectional paths, our algorithm may equip only one of the antiparallel streets with a bike path, instead of both as in the P + S network.

Figure 4 illustrates both types of networks for Dresden and Hamburg. The network generated by our algorithm largely coincides with the primary and secondary roads due to the high penalty if bike paths are removed; however, we observe strong differences in the density of the bike path coverage. Especially in Dresden, the resulting bike path network is much denser along the central north–south axis of high station density and bike-sharing usage, indicating that our algorithm correctly adapts the network to the input demand conditions (Fig. 4c). The differences for Hamburg are smaller due to the comparatively homogeneous demand across the city, although our algorithm introduces bike path shortcuts through residential areas in cases of high demand or to connect stations to the bike path network.

To quantitatively compare the bike path families for both cities, we normalize the length of bike paths $\lambda = \Lambda(M')/\Lambda(M_0)$ with respect to the length $\Lambda(M_0)$ after removing all of the unused bike paths (see Fig. 2b). We define the total perceived distance of all trips in the cyclist preference graph as

$$\mathcal{L}(\lambda) = \sum_{i,j\in V} n_{i\to j} L_{i\to j}(\Pi^*_{i\to j}(\lambda), \lambda) \tag{3}$$

where $L_{i\to j}(\Pi, \lambda) = \sum_{e\in\Pi} l_e(\lambda)$ (compare with equation (2)). Here, $l_e(\lambda)$ denotes the effective length of the street segment e in the cyclist preference graph G, given a set of bike paths $E_B(M')$ with normalized length λ (that is, including penalties only for those streets from which we have removed the bike path); $\Pi^*_{i\to j}(\lambda)$ denotes the shortest path in this cyclist preference graph and thus the route chosen by cyclists going from i to j. To compare the total perceived distance across both cities, we measure the overall performance $b(\lambda)$ of the resulting network as the bikeability

$$b(\lambda) = \frac{\mathcal{L}(0) - \mathcal{L}(\lambda)}{\mathcal{L}(0) - \mathcal{L}(1)}, \tag{4}$$

https://doi.org/10.1038/s43588-022-00318-w

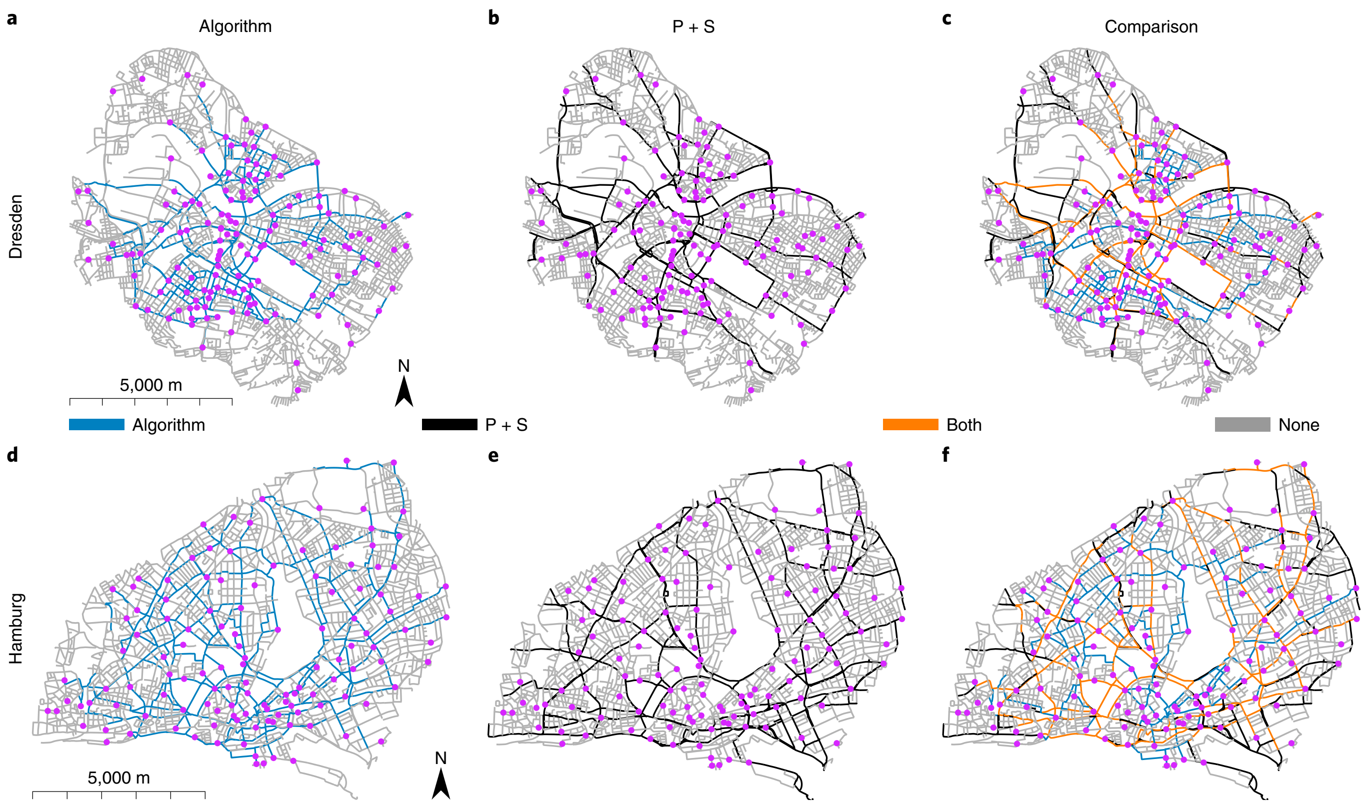

Fig. 4 | Demand-efficient bike path networks. **a**–**f**, Networks for Dresden (**a**–**c**) and Hamburg (**d**–**f**) with bike-sharing-station locations (purple, compare with Fig. 3). **a**,**d**, Bike path networks generated by the algorithm (blue) with the same total length as all of the primary and secondary streets, $\Lambda(M') = \Lambda_{P+S}$. **b**,**e**,Networks for the scenario in which only the primary and secondary streets (as per their OSM classification) are equipped with bike paths (black). **c**,**f**, Comparisons between both networks. The networks generated by the proposed algorithm largely coincide with the primary–secondary networks (orange edges in **c** and **f**) but more accurately reflect the input demand structure by also keeping highly used tertiary or residential streets equipped with bike paths and exhibiting a higher density of bike paths in high-demand areas.

where we again normalize the absolute values to the best- ($\lambda = 1$) and worst-case ($\lambda = 0$) scenarios; $b(0) = 0$ describes the network with no bike paths, whereas $b(1) = 1$ is the optimal network with bike paths along all of the shortest paths. We remark that our definition of bikeability differs from past measures[36] in that it quantifies the efficiency of the bike path network with respect to a specific demand distribution. The total difference between the physical trip length of cyclists and the direct shortest paths is only on the order of 10%, consistent with the empirical observations of cyclists' route choice behavior[27,28].

Figure 5 illustrates the bikeability across the generated sequence of bike path networks. Interestingly, a small fraction of bike paths with a small relative length of $\lambda > 0.1$ is sufficient to achieve more than 50% of the maximal bikeability in both cities. The larger area under the bikeability curve for Dresden compared with Hamburg is consistent with the differences in the demand structure between the two cities: we achieve a faster improvement in Dresden due to the more concentrated demand distribution, whereas we have to cover most of the city of Hamburg due to the more homogeneous demand. See Supplementary Fig. 3 for a brief overview of the bikeability of a further twelve cities.

A comparison with the bikeability of the primary–secondary bike path network with the same relative length λ_{P+S} of bike paths highlights the better adaptation to the demand structure in our algorithm. The bikeability is already high when all large roads are equipped with bike paths (about 0.87 for Dresden and 0.82 for Hamburg). Yet our algorithm manages to further increase this value to about 0.97 for Dresden and 0.95 for Hamburg (capturing more than 70% of the remaining potential of an optimal network $b(1) = 1$). Moreover, by adjusting the network to the route choice behavior, cyclists keep to streets equipped with bike paths for more than 89% of their total trip distance, compared with only about 60% in the primary–secondary network (see Fig. 5b,d). A negligible but non-zero fraction of the distance is cycled on tertiary and secondary streets without a bike path. See Supplementary Notes 5 and 4 for a comparison with static and dynamic forwards percolation approaches, respectively.

Impact of demand structure. We attribute the difference between the two cities in the above analysis to the structure of the bike-sharing demand distributions. To quantify the impact of the demand structure on our bike path network families and their bikeability curves, we compare the above results to synthetic bike path networks with homogenized demand. We create these homogeneous demand settings by first distributing demand equally between all stations and then distributing the stations as equidistantly as possible in the street network (see Fig. 6a and Methods).

Comparing the bikeability curves $b(\lambda)$ and $b_{\mathrm{hom}}(\lambda)$ in the empirical and the homogeneous demand settings, respectively, we find a comparatively large difference for Dresden and a much smaller difference for Hamburg (Fig. 6b,c). We quantify these differences by the area

$$\beta_{\mathrm{hom}} = \int_0^1 \left[b(\lambda) - b_{\mathrm{hom}}(\lambda)\right] \,\mathrm{d}\lambda, \qquad (5)$$

between the two bikeability curves, which describes the impact of the patterns—as well as the structure of the station and demand distribution—on the bikeability ($\beta_{\mathrm{hom}} \approx 0.015$ for Dresden, $\beta_{\mathrm{hom}} \approx 0.007$ for Hamburg; compare with Fig. 6). This confirms our above analysis that

https://doi.org/10.1038/s43588-022-00318-w

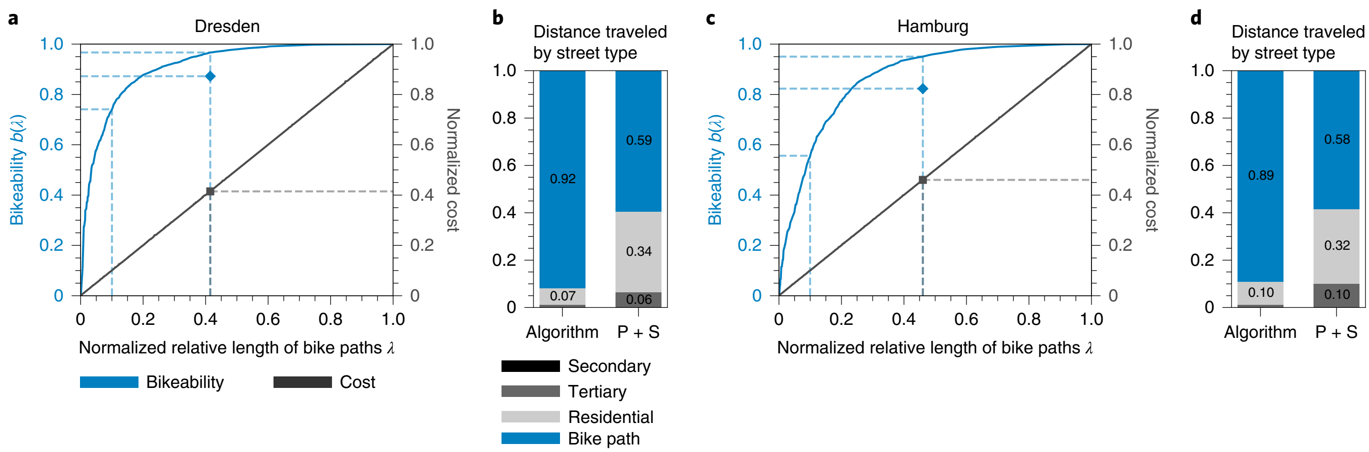

Fig. 5 | Demand-driven design of bike path networks improves bikeability. **a**,**c**, The resulting bikeability $b(\lambda)$ (blue, equation (4)) and relative cost of the network (gray) for Dresden (**a**) and Hamburg (**c**) as a function of the normalized length of the bike path network λ. **b**,**d**, Corresponding fraction of total distance traveled on streets with and without bike paths, in networks with bike paths with the same relative length λ_{P+S}, and the primary–secondary (P + S) comparison network. Adapting the network to the demand structure achieves over 89% of the distance traveled on bike paths (blue), also reducing the distance traveled on residential streets (light gray). For our algorithm, a negligible fraction of the total distance is cycled on tertiary (dark gray) and secondary (black) roads without a bike path (not visible in the bar chart).

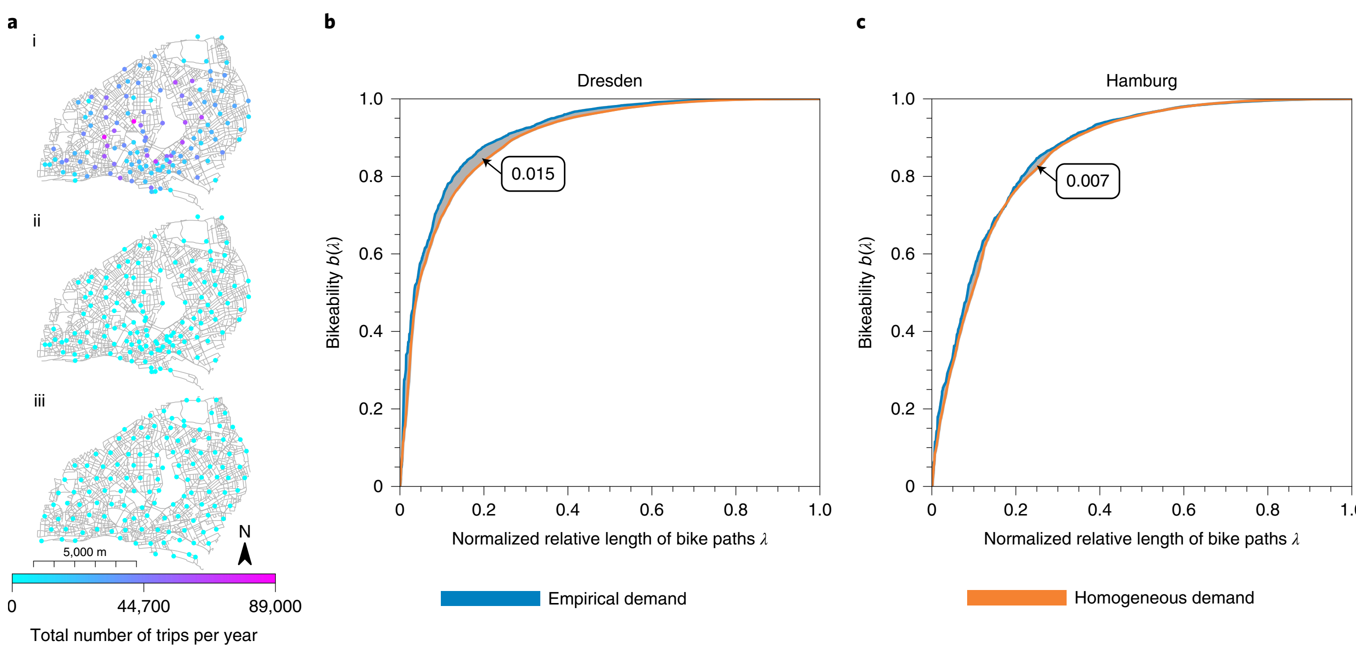

Fig. 6 | Bikeability differences quantify the importance of demand structures. **a**, Illustration of the demand homogenization for Hamburg. Starting from the empirical demand (i), we create uniform demand between all stations (ii) and then distribute the stations approximately equidistantly across the city (iii). **b**,**c**, Comparison of the bikeability $b(\lambda)$ (equation (4)) from the empirical (compare with Fig. 5) and homogenized demand data for Dresden (**b**) and Hamburg (**c**). The comparatively smaller area between the curves (shaded gray, equation (5)) for Hamburg suggests a more homogenous empirical demand, in line with our observations in Fig. 4.

the heterogeneous, centralized demand and station distribution in Dresden enhances the bikeability as fewer streets have to be equipped with bike paths to cover a large fraction of the total demand, whereas there is a much smaller effect in Hamburg. A similar approach may be used to quantitatively compare the impact that different street networks, or different types of cycling or desired demand distributions may have on the resulting bike path networks.

Discussion

Promoting cycling as a way to improve the sustainability of urban mobility is a complex problem. The attractiveness of cycling depends on the available infrastructure and safety of the trips, alternative mobility options, other mode choice decisions that influence the amount of car traffic, as well as aspects such as topography and weather[21].

Existing approaches to study bike path networks rely on various types of input data and focus on different properties; for example, improving connectivity of existing bike paths[24], purely structural network growth models[26] or standard forward percolation models based on static route choice data[25]. Compared with these more abstract percolation models, our adaptive inverse percolation framework trades computational speed for the explicit inclusion of cyclist demand. Compared with more detailed infrastructure models, we trade accuracy of the route choice model for the ability to adaptively adjust route choices as the network evolves.

Similar network pruning or inverse percolation techniques—as used in cascading failure models[37,38]—have previously been applied to understand the structure of transportation networks and their robustness[16,18]. The approach may therefore also find applications in designing and analyzing infrastructure and transportation networks beyond cycling. Combining the suggested approach with additional optimization steps, for example, by using successive addition and removal of individual bike paths or simulated annealing techniques[15] to explore the vicinity of networks constructed by our greedy approach, may further improve the quality of the resulting networks.

The proposed framework relies on several types of input, all with potential limitations in terms of data quality, modelling accuracy and interpretability of the results; however, the framework is highly adaptive and can easily be extended to overcome these challenges provided more detailed input data are available.

First, the quality and type of input data is critical for the resulting networks and their interpretation. Available street network data are mostly of high quality (see Methods) and even existing bicycle infrastructure may be included in the framework by preventing the removal of bike paths from specific street segments in the network; however, the resulting bike path networks have to be interpreted in the context of the demand input (compare our results for Dresden). For example, the empirical bike-sharing demand used to illustrate the framework may be strongly influenced by the currently (non-)existing infrastructure and the type of users of the service, thus skewing the generated networks to further improve already efficient parts of the network. At the same time, the framework is not limited to empirical demand data. Constructing efficient networks for desired and predicted demand may help guide extensions of bicycle infrastructure networks by suggesting efficient network structures[4,25]. Combining full mobility demand with a suitable mode choice model may even enable us to capture effects of induced additional demand when sufficient infrastructure is provided.

Second, the framework relies on a simplified route choice model to enable fast computation by mapping the route choice of cyclists to an effective shortest-path problem. We illustrated the framework with an effective network capturing only the effect of car traffic volume on cyclists route choice. A more detailed definition of the penalties, including additional deterrents such as slopes and (left) turns, may improve the accuracy of the route choice model. Furthermore, explicit safety considerations such as accident risk may be included in the penalties as proposed in ref. [39], enabling more accurate quantification of the actual safety in addition to the perceived safety. A more accurate representation of cyclist route choice behavior and heterogeneous preferences among cyclists, as assumed in common probabilistic route choice models, may be indirectly possible by considering multiple types of cyclists and creating a cyclist preference graph with appropriate penalties for each user type.

All of these potential extensions naturally require more details in the input data, such as information on traffic signals, street quality or expected driving behavior for car traffic across the street network. Recent contributions to the data-driven analysis and planning of cycling infrastructure and route choice as well as data collection methods are essential to establish a foundation of input data and ensure reliable results and predictions[24–26]. Although, eventually, first-hand on-site experience and detailed case-by-case modelling must determine the sensibility and feasibility of the suggested networks, with sufficiently accurate input data our framework may provide scenarios for efficient bicycle infrastructure networks, helping to guide planned extensions of existing infrastructure networks[4,25]. Our framework may thus complement current urban planning approaches[22] by helping to develop a more detailed understanding of the theoretical properties of efficient bike network structures across cities and offering a baseline of an efficient network to develop a more detailed long-term strategy to expand bicycle infrastructure.

Overall, the framework presented in this article may enable quantitative analysis of bike path networks with a large range of tools from network science by providing a way to quickly generate and compare families of efficient bike path networks in different settings and under different conditions.

Methods

Street networks

We download physical street networks for Hamburg and Dresden from OSM[35,40]. Although OSM data are crowdsourced, it is of high quality in developed countries, especially in Western Europe[41]. For other regions of the world, OSM is sometimes the only feasible source of data[42].

For both cities, we restrict ourselves to the area covered by the local bike-sharing schemes. We exclude city peripheries that are either sparsely populated with a low density of bike-sharing stations or a long distance from the city center. This results in a reasonably bikeable area of approximately 65 km^2 (Dresden) and 49 km^2 for Hamburg (Fig. 3).

We simplify these raw street networks by merging nodes within a radius of less than 35 m (for example, simplifying the detailed structure of intersections) and placing the resulting aggregated node at the centroid of their former positions[40]. Finally, we discard bike-inaccessible roads based on the OSM street classification hierarchy[43], in particular, edges labeled as: motorway, motorway_link, trunk, or trunk_link, which describe highways or high-speed motorways where cycling is not possible.

Based on the remaining nodes and edges (see Fig. 3), we create the physical street network G^{street} by assigning each street segment a length l_{ij}^{street} corresponding to its physical length in the OSM data.

Penalty factors

To fix the penalty factors in these networks without detailed information on traffic volume, we rely on the street type classification decoded in OSM as a proxy for the traffic load. Within the OSM category of car-accessible streets, we use OSM's classification hierarchy of: primary, secondary, tertiary or residential streets, replacing link street types with their corresponding normal street types, for example, primary_link with primary. Street segments not labeled with one of the four aforementioned classifications are assigned the residential status, as most other OSM classifications are reserved for small streets, for example, living_street. In case of ambiguity on the street segment length or street type, we always choose the first value in the list of the OSM data.

We assume that the expected traffic load and thus the corresponding penalty factor p_{ij}^0 monotonically increases from residential to primary roads. Based on this input, we construct the cyclist preference graph G with the perceived edge lengths $l_{ij} = p_{ij}\, l_{ij}^{\mathrm{street}}$ (equation (1)).

The penalty factors p_{ij}^0 quantify the trade-off between the distance a cyclist is willing to ride along a street with a bike path to avoid a specific street segment a without bike path. To illustrate our framework, we take penalty values representing the perceived distance compared with physically protected bike paths as they are the safest option for cyclists; however, the penalties can be adapted to represent other types of bike path implementations. We employ penalty factors as estimated in ref. [28]; the authors associated different penalties with different properties of the route and street segment by contrasting actually chosen bike routes with possible alternatives in a logit route choice model. We take the penalty factors representing accepted detours for streets with more than 30,000 vehicles per day, 20,000 to 30,000 vehicles per day, and 10,000 to 20,000 vehicles per day. Due to lack of detailed information on traffic volume, we directly match these penalties to the OSM street types: primary street segments are assigned a penalty $p_{ij}^0 = 7.0$, secondary $p_{ij}^0 = 2.4$, and tertiary $p_{ij}^0 = 1.4$. We add a smaller penalty factor $p_{ij}^0 = 1.1$ for residential street segments. These values serve as illustrative

https://doi.org/10.1038/s43588-022-00318-w

examples of the influence of car traffic and may vary depending on the location, trip purpose, bicycle infrastructure or for individual cyclists.

Bike-sharing demand data

We take demand data from local bike-sharing services in Dresden and Hamburg as demand input for the algorithm (Fig. 3). For both cities, we map the bike-sharing stations to their respective nearest node in the street network and obtain the origin–destination-resolved demand statistics by counting the overall number of trips $n_{i\to j}$ made throughout the entire observation period per pair (i, j) of stations. To quantify the total usage of each station, we compute the sum of bike rentals and returns at the station (see Fig. 3b,d). The two datasets represent two archetypes of local demand constellations: confined few-to-few demand in Dresden (Fig. 3a,b) and spatially homogeneous all-to-all demand in Hamburg (Fig. 3c,d).

Dresden. Dresden's local bike-sharing scheme operated on a combined station-based and free-floating mode during the observation period. Although bikes could be rented or returned on an as-needed basis within a pre-defined area in the inner city center, they needed to be rented from, and returned to, one of 159 stations outside of the free-floating zone[44].

To estimate the local bike demand, we use a proprietary dataset of approximately 440,000 bike-sharing trip records conducted by students of the local universities between November 2017 and March 2020 (except February and September 2018), which accounts for about 80% of all trips of the service[44]. The dataset contains, among other things, information on the trip origin and destination if the trip started or ended at a station, as well as the pickup and drop-off timestamps. The dataset does not contain positional information on trips conducted in free-floating mode.

To be able to fix the demand distribution to nodes in the street network, we exclude trips for which no origin or destination information is available (for example, trips starting or ending in the free-floating zones). Furthermore, we exclude nine stations that are distant from the city center and are thus not included in the core polygon illustrated in Fig. 3a, leaving approximately 163,000 trip records for the demand analysis. The remaining 150 stations are mapped to 142 nodes of the street network G (eight stations are mapped to the same node as another station, for example, two stations on two sides of a large street crossing).

Hamburg. In 2017, Hamburg's station-based bike-sharing scheme operated 206 stations, of which 129 were distributed in the core city (see Fig. 3b). Between January 2014 and May 2017 the service facilitated approximately 8.6 million rides for which detailed trip information is publicly available[45]. The data contains, among others, information on trip origin and destination station, pickup and drop-off timestamps, as well as user or bike-related information.

We again exclude trips where no origin or destination information is available as well as trips which start and end at the same station, leaving about 6.4 million trips in our region of interest (approximately 74% of all trips). After mapping the 129 stations to the street network G, we keep 127 unique locations (two stations are mapped to the same node as another station).

Homogenized demand

To quantify the impact of the demand distribution on the bikeability, we compare our results for the empirical demand distribution with results for homogenized demand distributions. We generate these randomized comparisons in three steps: (1) create approximately homogeneous demand and station distribution; (2) generate bike path ensembles for ten realizations of the demand and station distribution; (3) average the bikeability results from the different realizations.

We create these homogeneous demand settings by first distributing demand equally between all stations, setting $n_{i\to j}^{(\mathrm{hom})} = 1$ for all bike-sharing-station pairs i and j, where $i \neq j$. Second, we distribute the stations as equidistantly as possible in the street network. To achieve this station distribution, we create a triangular lattice in the polygon of the physical street network with a slightly higher total number of lattice points. We then map each lattice point to the closest node in the underlying street network G and delete excess points starting with those lattice points whose position in the triangular lattice is furthest from its corresponding node in G until we are left with the same number of nodes as station in our original data (Fig. 6a).

We create bike path networks as described above, compute the resulting bikeability and other measures, and average them over ten realizations of random homogeneous station distributions.

Data availability

The bike-sharing trip records used to estimate the cycling demand in Hamburg are publicly available from ref. [45]. The specific data used for Hamburg in this paper can be found at Zenodo with the code[46]. The bike-sharing trip records for Dresden are a proprietary asset of the Studierendenrat of the Technische Universität Dresden and nextbike GmbH, and cannot be made accessible by the authors. Source Data are provided with this paper.

Code availability

The code for our algorithm and a guide to reproducing the results is available through GitHub (https://github.com/PhysicsOfMobility/BikePathNet) under AGPL-3.0 license. The specific version of the package used to produce the results for this manuscript is available from Zenodo (ref. [46]).

References

1. Banister, D. The sustainable mobility paradigm. *Transp. Policy* **15**, 73–80 (2008).
2. Mazzoncini, R., Somaschini, C. & Longo, M. in *Green Planning for Cities and Communities: Novel Incisive Approaches to Sustainability* (ed. Dall'O', G.) 255–277 (Springer, 2020).
3. Buehler, R. & Pucher, J. Cycling to work in 90 large american cities: new evidence on the role of bike paths and lanes. *Transportation* **39**, 409–432 (2011).
4. Creutzig, F. et al. Urban infrastructure choices structure climate solutions. *Nat. Clim. Change* **6**, 1054–1056(2016).
5. Rhoads, D., Solé-Ribalta, A., González, M. C. & Borge-Holthoefer, J. A sustainable strategy for open streets in (post)pandemic cities. *Commun. Phys.* **4**, 183 (2021).
6. Schwedhelm, A., Li, W., Harms, L. & Adriazola-Steil, C. *Cycling During COVID-19* (World Resources Institute, 2020); https://www.wri.org/blog/2020/04/coronavirus-biking-critical-in-cities
7. Goetsch, H. & Quiros, T. P. *COVID-19 Creates New Momentum for Cycling and Walking. We Can't Let it Go to Waste!* (World Bank, 2020); https://blogs.worldbank.org/transport/covid-19-creates-new-momentum-cycling-and-walking-we-cant-let-it-go-waste
8. Jackson, M. O. *Social and Economic Networks* (Princeton Univ. Press, 2010).
9. Schröder, M., Nagler, J., Timme, M. & Witthaut, D. Hysteretic percolation from locally optimal individual decisions. *Phys. Rev. Lett.* **120**, 248302 (2018).
10. Aldous, D. & Barthelemy, M. Optimal geometry of transportation networks. *Phys. Rev. E* **99**, 052303 (2019).
11. Tero, A. et al. Rules for biologically inspired adaptive network design. *Science* **327**, 439–442 (2010).

12. Karschau, J. et al. Resilience of three-dimensional sinusoidal networks in liver tissue. *PLoS Comput. Biol.* **16**, (2020).
13. Kleinberg, J. M. Navigation in a small world. *Nature* **406**, 845 (2000).
14. Molkenthin, N., Schröder, M. & Timme, M. Adhesion-induced discontinuous transitions and classifying social networks. *Phys. Rev. Lett.* **121**, 138301 (2018).
15. Gastner, M. T. & Newman, M. E. J. Optimal design of spatial distribution networks. *Phys. Rev. E* **74**, 016117 (2006).
16. Verma, T., Russmann, F., Araújo, N. A., Nagler, J. & Herrmann, H. J. Emergence of core–peripheries in networks. *Nat. Commun.* **7**, 10441 (2016).
17. Barthélemy, M. & Flammini, A. Optimal traffic networks. *J. Stat. Mech. Theory Exp.* **2006**, L07002 (2006).
18. Barthélemy, M. Spatial networks. *Phys. Rep.* **499**, 1–101 (2011).
19. Scellato, S., Cardillo, A., Latora, V. & Porta, S. The backbone of a city. *Eur. Phys. J. B* **50**, 221–225 (2006).
20. Duthie, J. & Unnikrishnan, A. Optimization framework for bicycle network design. *J. Transp. Eng. A* **140**, 04014028 (2014).
21. Muñoz, B., Monzon, A. & Lois, D. Cycling habits and other psychological variables affecting commuting by bicycle in Madrid, Spain. *Transp. Res. Record* https://doi.org/10.3141/2382-01 (2013).
22. Rik de Groot, H. (ed) *Design Manual for Bicycle Traffic* Record no. 28 (Crow, 2016).
23. Bushell, M. A., Poole, B. W., Zegeer, C. V. & Rodriguez, D. A. *Costs for Pedestrian and Bicyclist Infrastructure Improvements* Techical Report (UNC Highway Safety Research Center, University of North Carolina, 2013); https://www.pedbikeinfo.org/cms/downloads/CountermeasureCosts_Report_Nov20131.pdf
24. Natera Orozco, L. G., Battiston, F., Iñiguez, G. & Szell, M. Data-driven strategies for optimal bicycle network growth. *R. Soc. Open Sci.* **7**, 201130 (2020).
25. Olmos, L. E. et al. A data science framework for planning the growth of bicycle infrastructures. *Transp. Res. C* **115**, 102640 (2020).
26. Szell, M., Mimar, S., Perlman, T., Ghoshal, G. & Sinatra, R. Growing urban bicycle networks. *Sci. Rep.* **12**, 6765 (2022).
27. Menghini, G., Carrasco, N., Schüssler, N. & Axhausen, K. Route choice of cyclists in Zurich. *Transp. Res. A* **44**, 754–765 (2010).
28. Broach, J., Dill, J. & Gliebe, J. Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transp. Res. A* **46**, 1730–1740 (2012).
29. Guerreiro, Td. C. M., Kirner Providelo, J., Pitombo, C. S., Antonio Rodrigues Ramos, R. & Rodrigues da Silva, A. N. Data-mining, GIS and multicriteria analysis in a comprehensive method for bicycle network planning and design. *Int. J. Sustain. Transp.* **12**, 179–191 (2018).
30. Banister, D. & Berechman, Y. Transport investment and the promotion of economic growth. *J. Transp. Geogr.* **9**, 209–218 (2001).
31. Buehler, R. & Dill, J. Bikeway networks: a review of effects on cycling. *Transp. Rev.* **36**, 9–27 (2016).
32. Daganzo, C. F. & Sheffi, Y. On stochastic models of traffic assignment. *Transp. Sci.* **11**, 253 (1977).
33. Storch, D.-M., Schröder, M. & Timme, M. Traffic flow splitting from crowdsourced digital route choice support. *J. Phys. Complexity* **1**, 035004 (2020).
34. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
35. OpenStreetMap contributors. *Street Networks Retrieved From openstreetmap.org* (OpenStreetMap, 2021); www.openstreetmap.org/copyright/en
36. Kellstedt, D. K. et al. A scoping review of bikeability assessment methods. *J. Commun. Health* **46**, 211–224 (2021).
37. Carmona, H. A., de Noronha, A. W. T., Moreira, A. A., Araújo, N. A. M. & Andrade, J. S. Cracking urban mobility. *Phys. Rev. Res.* **2**, 043132 (2020).
38. Schäfer, B., Witthaut, D., Timme, M. & Latora, V. Dynamically induced cascading failures in power grids. *Nat. Commun.* **9**, 1975 (2018).
39. Folco, P., Gauvin, L., Tizzoni, M. & Szell, M. Data-driven bicycle network planning for demand and safety. Preprint at https://arxiv.org/abs/2203.14619 (2022).
40. Boeing, G. OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **65**, 126–139 (2017).
41. Graser, A., Straub, M. & Dragaschnig, M. Towards an open source analysis toolbox for street network comparison: indicators, tools and results of a comparison of OSM and the official Austrian reference graph. *Trans. GIS* **18**, 510–526 (2014).
42. Quinn, S. & Bull, F. *Geospatial Information System Use in Public Organizations* 80–96 (Routledge, 2019).
43. OpenStreetMap contributors. *Key:Highway—OSM Wiki* (Wikipedia, 2022); https://wiki.openstreetmap.org/wiki/Key:highway
44. *Trip Data Dresden Universities* (NextBike GmbH, 2020).
45. *Buchungen Call a Bike* (Stand 05/2017) (CC BY 4.0) (Deutsche Bahn, accessed 14 April 2020); https://data.deutschebahn.com/dataset/data-call-a-bike
46. Steinacker, C. *PhysicsOfMobility/BikePathNet: Update for Supplementary Material for "Demand-driven Design of Bicycle Infrastructure Networks for Improved Urban Bikeability"* (Zenodo, 2022); https://doi.org/10.5281/zenodo.6975015

Acknowledgements

We thank S. Huber, S. Lißner and R. Gerike for stimulating discussions. The authors are grateful to the Centre for Information Services and High Performance Computing (ZIH) TU Dresden for providing facilities for high-throughput calculations. C.S. acknowledges support from the Deutsche Bundesstiftung Umwelt (DBU, German Federal Environmental Foundation). D.S. acknowledges support from the Studienstiftung des Deutschen Volkes. M.T. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Center for Advancing Electronics Dresden (cfaed). The project was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; project no. 493613373 to M.S). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

D.-M.S. and M.S. initiated the research supported by M.T. All of the authors conceived and planned the research. C.S. collected and analyzed the empirical data supported by D.-M.S. C.S. and M.S. designed the algorithm. C.S. wrote the code, performed and analyzed the simulations, supported by M.S. All of the authors contributed to interpreting the results and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43588-022-00318-w.

Correspondence and requests for materials should be addressed to Christoph Steinacker or Malte Schröder.

Peer review information *Nature Computational Science* thanks Nuno Araújo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Handling editor: Jie Pan, in collaboration with the *Nature Computational Science* team. Peer reviewer reports are available.