Transportation Research Part D 100 (2021) 103044

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd

Mining dockless bikeshare data for insights into cyclist behavior and preferences: Evidence from the Boston region

Bita Sadeghinasr[a,*], Armin Akhavan[a], Peter G. Furth[a], Steven R. Gehrke[b], Qi Wang[a], Timothy G. Reardon[c]

[a] Northeastern University, Department of Civil and Environmental Engineering, 360 Huntington Ave., Boston, MA 02115, USA
[b] Northern Arizona University, Department of Geography, Planning, and Recreation, Flagstaff, AZ 86001, USA
[c] Metropolitan Area Planning Council, 60 Temple Pl, Boston, MA 02111, USA

ARTICLE INFO

Keywords:
Bicycling
dockless bikeshare
Level of traffic stress
Contraflow
Bicycle routing, detour

ABSTRACT

Emerging micromobility services provide not only new transportation options, but also valuable sources of data for researchers and planners seeking to understand traveler behavior and preferences and to improve transportation networks. Dockless systems for shared bicycles and electric scooters have recently begun operating in many American cities, providing data not only on trip start and end, but also on the route followed. In this study, we analyze data from more than 110,000 dockless bikeshare trips in Boston's suburbs. About 15% of trip ends were within 100 m of a transit station, indicating that bikeshare serves many more functions than access to transit. More than 40% ended in a town or village center, suggesting that bikeshare may support the local economy. From examining GPS traces, it was clear that – contrary to the assumptions of common bike routing algorithms – bike riders often use sidewalks, other footways and soft-surfaced paths, parking lots, and driveways, and 36% ride the wrong way on one-way streets. An examination of the streets with high contraflow volumes indicates two common profiles. One is streets whose travel channel is so narrow that either the bike or vehicle has to pull into a parking lane to let the other pass; those streets typically have low traffic speeds and very low traffic volumes. The other is streets whose travel channel is wide enough for a bike and car to pass without either yielding to the other; they typically have greater vehicular speed and volume, but still only one lane. After classifying streets by level of traffic stress, we find that only 7 percent of trips use exclusively low-stress links, and that about 40 percent of bike-miles are ridden on low-stress links. These low percentages are consistent with an underdeveloped bike network that forces riders to use through streets for all except ultra-local trips; it may also reflect a prevalence of sidewalk riding on major arterials.

1. Introduction

Recent years have witnessed a proliferation of new forms of mobility in U.S. cities and around the world. Carsharing, ride-hailing, bikeshare, scooters, and other services provide not only new options for travelers, but also new sources of data that can help

* Corresponding author.
E-mail addresses: sadeghinasr.b@northeastern.edu (B. Sadeghinasr), arminakhavan@northeastern.edu (A. Akhavan), p.furth@northeastern.edu (P.G. Furth), steven.gehrke@nau.edu (S.R. Gehrke), q.wang@northeastern.edu (Q. Wang), treardon@mapc.org (T.G. Reardon).

https://doi.org/10.1016/j.trd.2021.103044

municipalities and state agencies make better transportation investments and policy choices. Previously, location data collected from sources such as smart phones (Barbosa et al., 2018; Sadeghinasr et al., 2019; Akhavan et al., 2019a; 2019b), call detail records (Gonzalez et al., 2008), social media (Wang et al., 2018), and GPS receivers installed on vehicles (Hosseinzadeh et al., 2021a,b; Mahdinia et al., 2021; Mohammadnazar et al., 2021a, 2021b) have been used to study human mobility in urban settings.

Data from dockless bikeshare systems provides a valuable and complex source of information for transportation planning. Unlike earlier-generation bike share systems that rely on docks, dockless bikeshare requires GPS functionality so that the location of available vehicles can be provided to would-be renters (Gehrke et al., 2021a,b; Lazarus et al., 2020). That functionality, in addition to the flexibility of dockless operation, can make dockless bike systems a unique and valuable source of data. Modern dock-based bikeshare systems such as those in Washington, DC, New York, and Boston offer time-stamped origin–destination data by trip, affording valuable insights on origin–destination patterns, temporal utilization, and interaction between bikesharing and transit (O'Brien et al., 2014; Vogel et al., 2011; Gehrke & Welch, 2019). Still, with dock-based bikeshare, origin and destination information are limited to dock locations, whereas with dockless bikeshare, detailed spatial information on users' ultimate destination can be captured.

Even more importantly, with little incremental cost, a dockless bikeshare system can be configured to write GPS records at regular intervals while bikes are en route. The Lime dockless bikeshare system operated in Boston's suburbs in 2018 and 2019, the subject of this study, has just such a configuration, and therefore yields data on the routes cyclists followed as well as bike speed on different route segments.

The sample size of cyclist routes that data from a dockless system can provide is far greater than what has been available for earlier research on cyclist routes. Compared to the 110,000 rides used in this study, Pritchard et al. (2019) used data from113 participants to study one contraflow bike lane; Park and Akar (2019) studied detour levels using 1531 trips from volunteers; Ton et al. (2017) looked into cyclists' route-choice behavior and preferences using 3045 trips; and Crist et al. (2019) assessed route preferences and the level of low stress cycling connection between origins and destinations using 1038 trips.

The route-level data that dockless bikeshare makes possible offers unique insights into cyclists' mobility patterns and route choice behavior. In particular, these data shine light on three aspects of cyclists' route choices.

One is the extent to which cyclists prefer segments with low traffic stress and avoid segments with high traffic stress. Furth et al. (2016) developed a method for classifying streets by the levels of traffic stress they impose on cyclists, and postulate that only a small fraction of people will tolerate high-stress segments. Their classification system or one like it is used by communities across the country for bike network planning, including identifying links that are important for bicycle connectivity but are high stress (Moran et al., 2018; Semler et al., 2017; Gehrke et al., 2020). With route traces, one can observe the extent to which riders follow a longer route from origin to destination in order to avoid high-stress road segments. Several researchers with access to bicycle route data have studied the degree to which cyclists choose routes with low traffic stress, sometimes with mixed results (Broach et al., 2012; Wang et al., 2016; Crist et al., 2018). The Lime dockless bikeshare data affords another such opportunity, with a large sample size and a population that is not dominated by experienced cyclists.

A second is the extent to which cyclists use links that are not typically included in the network cyclists are assumed to use in common routing apps, including informal paths, sidewalks, parking lots, and private driveways.

A third is the extent to which cyclists ride contraflow, that is, the wrong way on one-way streets. Contraflow riding tends to be under-studied because it is nearly invisible – often dispersed over many segments and occurring mostly on local streets, where counts are rarely made; dockless bikeshare data offers a unique view of what happens on those streets. Putta and Furth (in press) estimated demand for contraflow riding on streets in the four central municipalities of the Boston metro area by looking at shortest low-stress paths from every home to every job, assuming that contraflow is allowed on local one-way streets. They found contraflow is critical to low-stress bike network connectivity, increasing the fraction of jobs that people can reach from home using only low stress streets and paths from 1.2% to 8.7%. They also found that a substantial part of this connectivity benefit derives from a small number of links that provide attractive route connections for more than just local travelers.

Because those results are based on a theoretical shortest-path analysis, they still leave unanswered important questions. Are people actually willing to ride contraflow on one-way streets? And, by looking at the streets that attract a lot of contraflow riding, can we determine what type of street makes people feel that it's safe to ride contraflow? The route traces provided by the dockless bikeshare data provide a unique and valuable resource for answering these questions.

While trip and fleet statistics for dockless systems are generally available through vendor dashboards or periodic reporting, we are aware of no other publications using trip-level route data on this scale to study cyclist travel patterns. While some rider or link-level data have been made available on occasion by providers of fitness apps and devices (e.g, Strava, Fitbit), the population using those products are more likely to be regular, confident cyclists, and unrepresentative of the "interested but concerned" cyclist type described by Geller (2016). In contrast, the majority of Boston-area Lime Bike users are not dedicated, regular cyclists; according to a Lime-administered rider survey (First Miles: Examining 18 Months of Dockless Bikeshare in Metro Boston), 53% of Lime Bike users last rode their personal bike, if they even have one, more than one month prior to being surveyed. These results suggest that many of the behaviors and preferences of bikeshare users may be reasonably representative of new or casual bike riders – the kind of people that policy-makers aim to attract to bicycling – making results applicable for bicycle transportation planning and policy in general.

1.1. Background: dock-based and dockless bikeshare in the Boston region

In the Boston region, dock-based bikeshare, originally named Hubway and later changed to Bluebikes, launched in July 2011 in the City of Boston and subsequently expanded to serve Cambridge, Somerville, and Brookline; in late 2019, Everett also joined the Blue Bikes system. The regional expansion of the system was facilitated by the Metropolitan Area Planning Council (MAPC), the regional planning agency which conducted the procurement process and contracting between the vendor (formerly Alta, now Motivate) and the participating municipalities, public agencies, and sponsors.

Following a pilot program in the town of Malden that began in September 2017, dockless bikes were formally introduced to the region in April 2018. Lime, a transportation-rental company based in the United States, was selected as the vendor through a competitive procurement issued by 15 cities and towns and again facilitated by MAPC (a second vendor, Spin, was also selected but withdrew before actually deploying any bikes) (Harmon, 2018). Due to conditions established by the contract with Motivate, all of the Bluebikes municipalities were prohibited from participating in the regional dockless bikeshare contract, and Lime is prohibited from operating its system in those municipalities. (The lone exception was Everett, which for a short period in 2019 was in both the Lime and Bluebikes system.) Lime vehicles may be ridden into the Bluebike service area but may not be rented once locked and must instead be collected by Lime. Lime charged $1 per 30 min for riding the manual bikes, $1 to unlock and $0.15 per minute to ride their pedal-assist e-bikes, or $29.99 per month for the LimePrime membership.

At the system launch in 2018, Lime's fleet consisted of pedal-powered bikes and a very small number of electric-assist bicycles. In September 2018, Lime began to transition the entire fleet to electric-assist, removing most of the pedal-powered bikes over the winter of 2018–19. At the end of 2019, Lime stopped renting bicycles altogether and refocused its business on scooters. This study deals with Lime bike trips, both pedal-powered and electric-assist.

Fig. 1 shows the municipalities in the Lime service area with their rate of trip making per capita (resident population plus total employment) during the one-year period from October 2018 to September 2019. Malden, with 450 trip-origins per 1000 residents and employees, ranked first followed by Everett, Arlington, and Winthrop. Average population density of the Lime region is 5396 people per square-mile, with communities ranging from a population density of 1,024 (Bedford) to 18, 225 (Chelsea) residents per square mile.

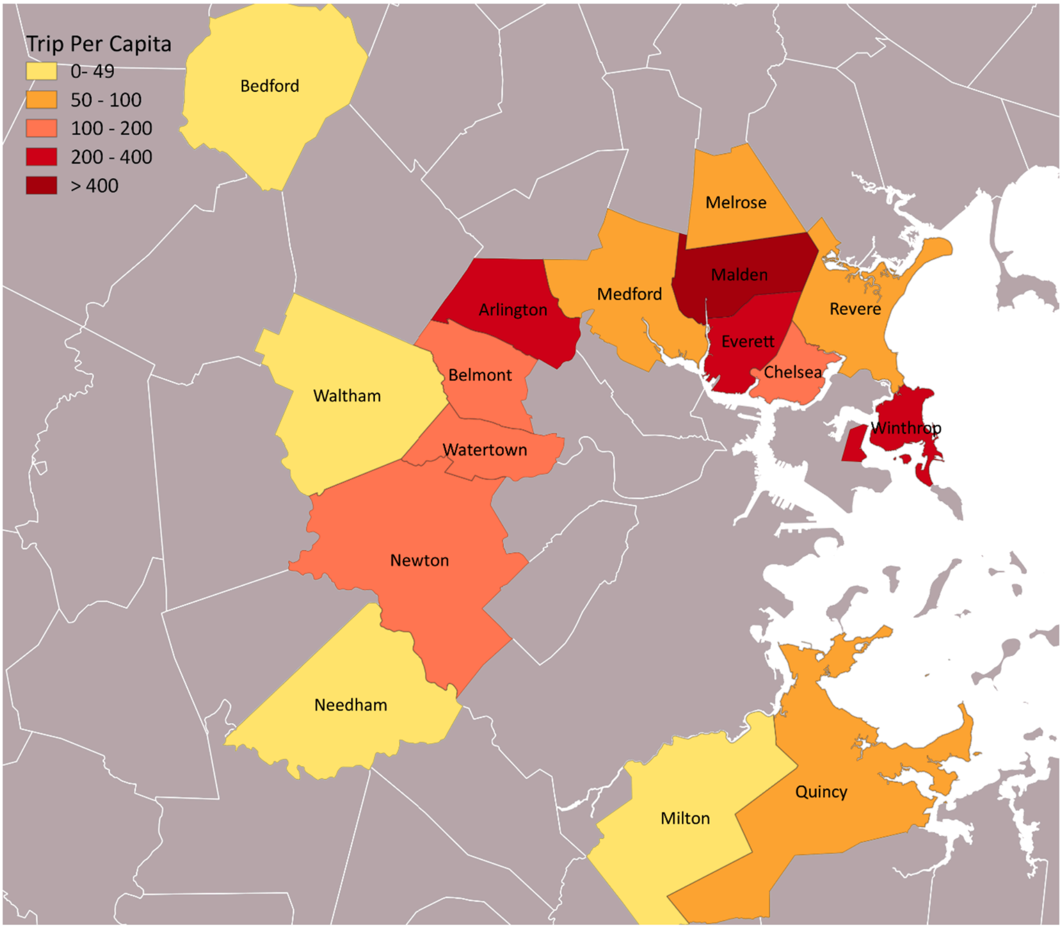


Fig. 1. Lime bike trip making per capita across the service area.

2. Methods

2.1. Data source

The regional Memorandum of Understanding governing the Lime system stipulates access for MAPC to trip level data for the system. Data of 397,996 dockless bikeshare trips were captured over an 18-month period from April 1, 2018, through September 30, 2019. Data were accessible through the Lime API and formatted according to the Mobility Data Specification (MDS), a nationally-recognized data standard now maintained by the Open Mobility Foundation (https://github.com/openmobilityfoundation/mobility-data-specification). Fields common to every record belonging to a trip include provider ID, provider name, device ID, GPS accuracy, vehicle ID, vehicle type, propulsion type, start time, end time, trip ID, trip distance, and trip duration. Trip distance is believed to be the actual on-road distance of the trip, as determined by Lime based on GPS data (though as described below, many trips lack distance information.)

GPS data is held in three additional fields: latitude, longitude, and a timestamp. Records were made at variable intervals, sometimes as short as 2 s but more often in the neighborhood of 30–40 s.

2.2. Data cleaning and map matching

Data errors arose from malfunctioning GPS units, memory issues, user error (e.g., failing to lock a bike at the end of a trip) or indecision (e.g., unlocked and quickly re-locked a bike), vehicle engagement problems, or vehicle disrepair. This required filtering out trips with obvious data errors. Table 1 summarizes the filters applied and how many trips were disqualified in each successive filter.

Of the 397,996 bicycle trips, 25.2% were filtered out because they lacked intermediate waypoints. Another 14.2% were filtered out because they were unreasonably short (less than 1 min or 100 m). Many of these could be failed attempts at unlocking bikes or rejection of a non-working bike. A small fraction of trips (0.4%) was filtered out because they were longer than 5 h or 20 km, or had GPS coordinates outside the Greater Boston area (0.3%).

Many of the trips had successive records with identical coordinates, usually followed by a distant point whose time was shortly after the last record at the repeated point. Rather than discarding the trip because of the impossibly high speed implied by the "jump" to a new point, records that repeated an earlier location were discarded, assuming instead that the GPS unit had been unable to update its coordinates. After this correction, another 23.6% of trips were filtered out because their remaining GPS records implied a speed greater than 12 m/s (27 mph, 43 km/h) between successive waypoints, a speed few people can attain on a racing bike, and virtually impossible to attain on a bike-share bike, even on the steepest downhills of metro Boston. The distribution of speed by link has a 98th percentile of about 6.6 m/s. A frequently used boundary for outliers is the 75th percentile plus 3 times the inner quartile, which for this distribution is about 9 m/s. Therefore, trips with segments speed exceeding 12 m/s likely suffer from data error.

An additional 7.8% of the trips were filtered out because they had gaps between waypoints greater than 200 m (straight-line). Most street segments (blocks) in the service area are shorter than 200 m, and for the network analysis that followed, it was important to the agency that bikes were positively detected on or very near the street segments to which they were matched rather than imputed based on shortest path routing.

Finally, waypoints were snapped to the OpenStreetMap (OSM) road network. In the Lime service area, OSM is well provisioned with bicycling and walking paths as well as with driveways in large parking lots. In some communities in the Lime service area, sidewalks are typically coded as distinct features, while in others they are not. Map-matching was done using Graphhopper (Ramm, 2017), an open source routing and navigation engine written in Java and using OpenStreetMap network data. For map-matching, it uses the Viterbi algorithm, a dynamic programming algorithm that finds the best path through a Hidden Markov Model lattice (Newson & Krumm, 2009), finding the most likely sequence of snapped points by maximizing the product of measurement probabilities (likelihood that a GPS record Zt would be observed if the vehicle were actually on road segmentr) and transition probabilities (probability of a vehicle moving between the candidate road-matches at two consecutive times Zt and Zt_{+1}).

Our first attempt at map matching resulted in many highly circuitous and unlikely trips. Closer examination of the original GPS waypoints revealed that many trips used parking lot driveways, sidewalks, and footpaths, and many traveled the wrong way on one-way streets, all of which are forbidden by the standard cyclist profile in Graphhopper. Also, in OSM, bike paths that are not hard-surfaced are coded as footpaths. Therefore, a new routing profile was created that allows the use of such facilities and, essentially,

Table 1
Number of trips disqualified in each successive filter.

All bike trips	397,996
1. Trips missing waypoints	100,402
2. Trips shorter than 1 min or 100 m	56,494
3. Trips longer than 5 h or 20 km	1786
4. Trips partially or entirely out the study area	1384
5. Trips with a link with a speed higher >12 m/s	93,897
6. Trips with a jump >200 m	30,984
7. Trips that couldn't be matched to the street network	3,061
Remaining trips	110,024

routes bikes as if they were pedestrians, but will prioritize bike facilities over other facility types. With this change, matched routes appeared to be reasonable for almost all trips, and only 0.8% of the original set of trips had to be filtered out because of route matching error.

Of the different filters applied, only one that excluded more than 1 percent of trips was discretionary, which is the maximum gap length of 200 m. Because this filter removed only 7.8% of trips, and because a more relaxed filter would have retained only some of those trips, it is unlikely that a more relaxed filter would have altered by much the results derived from the trips that were retained.

In the following analyses that depend on the routing data, only trips that passed all of the filters are included (n = 110,024). However, figures and distributions that provide information on the number of rides are from all trips except those outside the great Boston area, or too short or too long (n = 338,332).

2.3. Identifying contraflow travel

Contraflow riding was detected by comparing the direction of travel between successive waypoints in the GPS trace to the allowed direction of travel on one-way road segments. In OSM, segments tagged as one-way may be simple one-way streets, but they can also be unidirectional carriageways that are part of divided road. Network analysis was used to distinguish the two.

2.4. Classifying streets by level of traffic stress

To analyze riders' use and preference for streets with different degrees of traffic stress, all network links were classified by level of traffic stress (LTS) (Furth et al., 2016). There are four levels of traffic stress: LTS 1 is meant to be suitable for children and timid adults, LTS 2 involves limited traffic interactions that most adults can tolerate and corresponds to the population stratum "interested but concerned" in Geller's taxonomy (Geller, 2016; Dill & McNeil, 2013), LTS 3 involves a greater amount of interaction with traffic, though still at a level that "enthused and confident" cyclists will tolerate, and LTS 4 involves traffic interactions and danger that only the "strong and fearless" will tolerate. LTS 1 and 2 are considered "low-stress."

The classification criteria used roughly follow Furth's LTS version 2.0 (Furth, 2017; Furth et al., 2018), adapted as follows to the data available from OpenStreetMap.

- LTS 1: bike paths (as coded in OSM), streets with physically separated bike paths, and streets classified as residential.
- LTS 11 (a subset of LTS 1): Sidewalks (where coded in OSM as distinct network features), other footpaths including soft-surfaced multiuse paths, driveways in parking lots, and other driveways outside the public right of way.
- LTS 2: streets with conventional bike lanes having no more than one lane per direction and a speed limit of no more than 30 mph.
- LTS 3: streets with conventional bike lanes that don't meet LTS 2 qualifications, and non-residential streets with no more than one lane per direction and speed limit no more than 30 mph.
- LTS 4: non-residential streets without bike lanes that fail to meet LTS 3 qualifications.

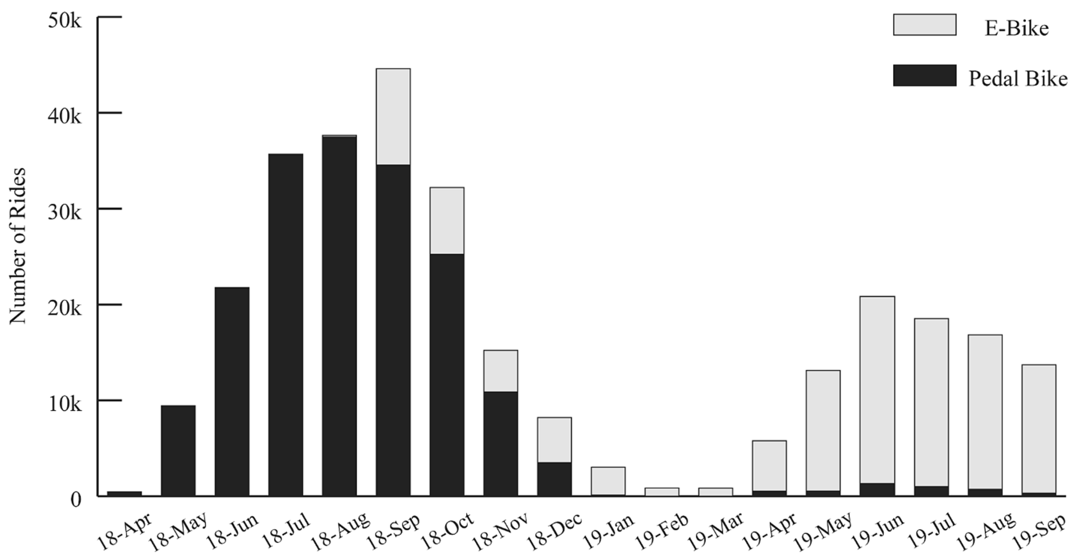


Fig. 2. Monthly dockless bikeshare ridership by propulsion type.

3. Findings

3.1. Temporal trends, trip length, and means of propulsion

The ridership trend across the first 1.5 years of the system's life is shown in Fig. 2. Ridership by propulsion type (pedal powered and electric-assist) largely reflects Lime's changing provision of bicycle types. The significantly lower ridership in the second year may reflect a reduction in the fleet size as Lime transitioned to electric-assist bikes. Municipal staff reported that there were fewer bikes available in 2019, but unfortunately Lime was not able to provide firm statistics on the number of vehicles deployed over the study period. Lower ridership in 2019 may also reflect the higher price charged for electric assist bikes. The very low ridership over the winter probably also reflects supply more than demand, as pedal bikes were removed and not replaced until spring with electric-assist bikes.

The median trip length trend is shown in Fig. 3. Median trip length in 2019, when most of the fleet was electric-assist bikes, is about 2 km or 1.25 mi, while in 2018, when most of the fleet was pedal-powered bikes, it is about 1.6 km, or 1 mi. These low trip lengths suggest a substantial degree of trip substitution from walking and ride hailing. The difference between the two years suggests that people consider the practical range of bikeshare to be greater with electric-assist bikes, which are both faster and demand less exertion.

Interestingly, while median trip length was greater in 2019, median trip lengths within a given year were about the same for the two types of propulsion. Because the fleet was dominated by pedal powered bikes in 2018 and by electric-assist bikes in 2019, it may be that in 2018 people were expecting to get a pedal-powered bike and determined their trip length accordingly, even if they got an electric-assist bike, and vice versa in 2019. Comparing September 2018 with September 2019 (the only year-over-year comparison in which the fleet had both pedal-powered and electric assist bikes), median trip length for the two propulsion types combined was 30% greater in 2019.

The trip length distribution by season and day type is shown in Fig. 4. 75 percent of Lime trips were between a half mile and two miles in length, comparable to dockless e-bikeshare users in San Francisco (Lazarus et al., 2019). Quarters 2 and 3 (warmer weather, more daylight) had nearly the same shape distribution as quarters 4 and 1 (in Fig. 4, with similar quarters grouped to reduce clutter), but trip lengths a little longer in quarters 2 and 3. Weekend trips are on average 0.13 miles longer than weekday trips regardless of the time of year.

Ridership varies with time of the day (Fig. 5). On weekdays, ridership is greatest between noon and 9 p.m. with a relatively flat peak around 5 p.m. A small a.m. peak can also be observed, but it has less ridership than any hour in the afternoon. On weekends, there is a relatively flat peak from 1 to 6 p.m. On both weekdays and weekends, late night use is strong; not until after 1 a.m. does it drop to less than 20% of peak ridership. Many of those late-night bikeshare trips may be substitutes for taxi and other ride-hailing trips.

We sus

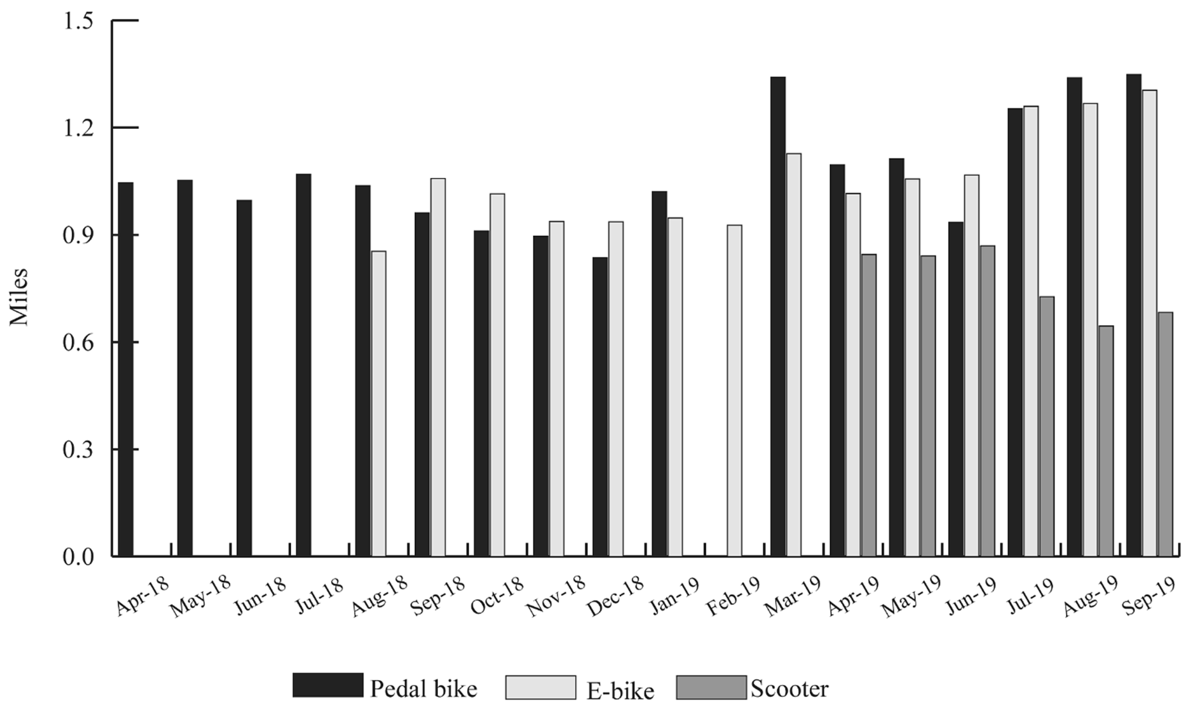


Fig. 3. Median trip length by month and vehicle / propulsion type.

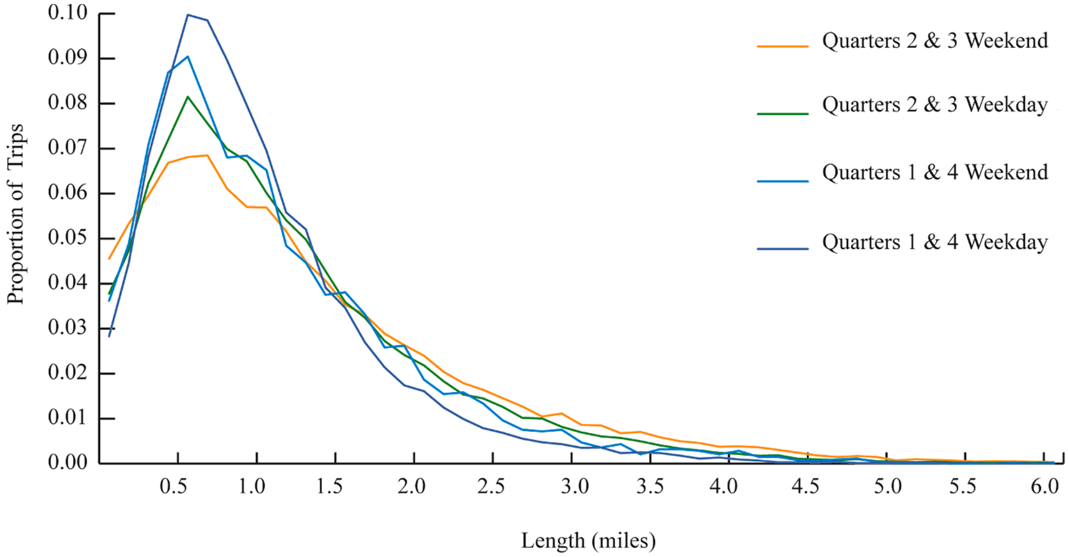


Fig. 4. Trip length distribution by quarter and weekday / weekend.

3.2. *Origins, destinations, and transit connections*

Around 15 percent of Lime bike rides across the region began or ended within 100 m of a rapid transit, commuter rail, or major bus route stop, showing that bikeshare is used for much more than just "last mile" access to transit. Of trips beginning in Arlington, Bedford, Melrose, and Needham, a substantial share of trips (9 to 14 percent) end at a transit station located in a neighboring municipality, indicating a need for regional coordination.

City, town, and village centers accounted for 45% of trip origins and 40% of trip destinations, suggesting that bikeshare may be supportive of local businesses.

Residential areas accounted for 24% of trip origins and 27% of trip destinations. This imbalance between neighborhoods (where destinations exceed origins) and business districts (where origins exceed destinations) required Lime to semi-regularly "rebalance" by removing bikes from places with excess destinations to places with excess origins. While information on rebalancing operations was not available, it is possible from the data to determine locations with excess destinations and excess origins. An interactive map provided in an online report shows these locations using 250 m × 250 m grid cells; one can see there that the excess origins cells tend to be at town centers and commercial main streets, while the excess origins cells are largely in outlying residential areas (Metropolitan Area Planning Council, 2019).

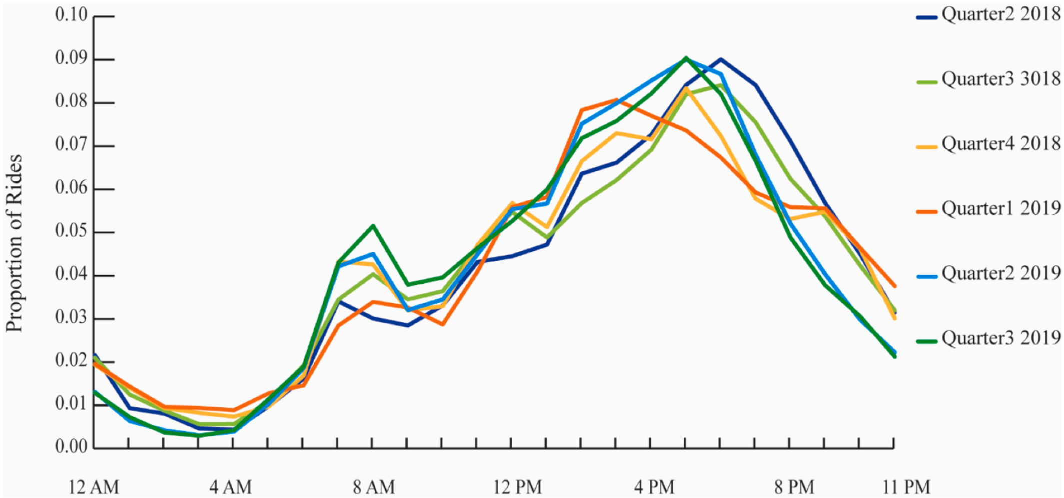
3.3. *Contraflow riding*

Contraflow riding – that is, riding in the wrong direction on a unidirectional link – can occur on simple one-way streets as well as on unidirectional carriageways that are part of a divided, two-way road. One might expect to find very little contraflow riding on divided roads, yet we found that of all Lime bikeshare trips with matched routes, 23% included at least one segment in which the cyclist rode the wrong way on a divided road. A review of such links with more than 100 contraflow trips found that they were all multilane arterial roads or multilane traffic circles lacking bike lanes and bike paths, but with sidewalks. Therefore, we suspect that most of this apparent contraflow riding actually represents riding on sidewalks, suggesting a need for protected bicycling facilities on these roads.

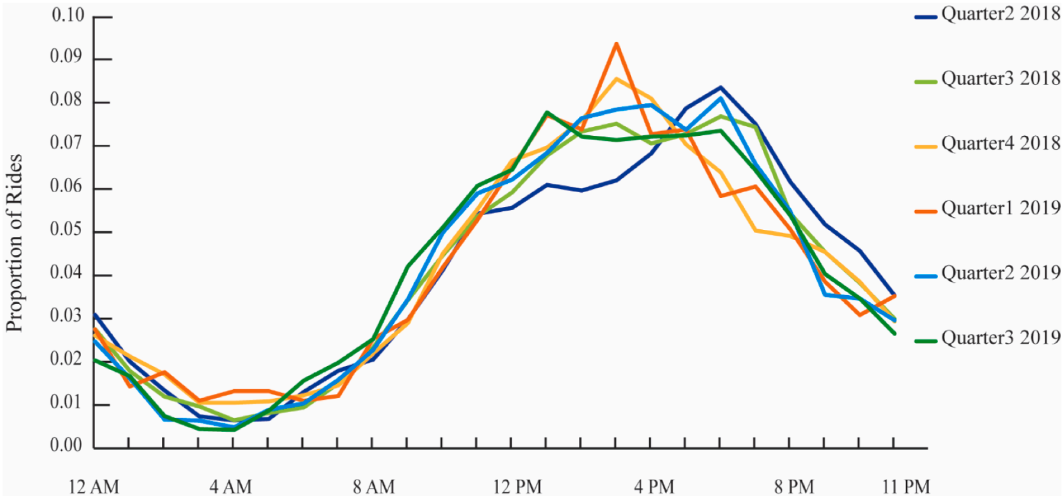
Looking only at contraflow riding on simple one-way streets (i.e. streets that are not part of a two-way divided road; also excludes private driveways), 36% of Lime bike users rode contraflow for at least part of their trip. It should be noted that many of the municipalities in the Lime service area have a lot of one-way streets; in some of those municipalities, the percentage of riders with a contraflow segment is almost 50%. Over the entire Lime service area, contraflow on simple one-way streets accounts for 6.2% of bike-miles ridden.

One would expect that frequent contraflow use occurs where two conditions are met: cyclists consider the road safe for contraflow riding and it offers a shorter route to reach one's destination. Fig. 6 maps the simple one-way road segments with high contraflow volumes in Malden, Everett, and Chelsea, the three communities with most contraflow Lime riding. Street segments with at least 100 contraflow bike trips were examined to determine what makes them attractive for contraflow from both the comfort / safety perspective and the network / routing perspective.

The streets with high contraflow use are nearly all single-lane roads (and for the exceptions, we suspect most cyclists use the sidewalk). They can be divided into two profiles, a narrow "single channel" profile and a broader "dual channel" profile. In the single channel profile, there is parking on both sides and the opening between the parking lanes is too narrow for a car and bike to pass one another at speed, and so one of the parties must pull over to let the other pass, an operating mode called "courtesy operation." (Motor



a. Weekdays



b. Weekends

Fig. 5. Ridership by hour of the day.

vehicles commonly use this operating mode on narrow two-way streets called "courtesy streets" or "yield streets", with cars pulling over to let another car pass (NACTO, 2013)). High-contraflow-use streets with the single channel profile all have low traffic volumes and low traffic speeds. Speeds are low partly due to road narrowness and parking friction and partly because, due to their low volumes, they have a Stop sign at almost every intersection. Parking lane occupancy on these streets, as viewed using Google StreetView, appears to be below 50%, leaving plenty of gaps into which a bicycle or vehicle can pull to allow the other party to pass.

Streets with a dual channel profile (a car and opposite direction bike can pass one another without slowing or pulling over) and high contraflow use are mostly 24 ft wide, with parking on only one side. After allowing 8 ft for a parking lane and 5 ft (the standard width of a bike lane) for a contraflow cyclist, that leaves drivers an 11 ft channel – the standard width of a vehicle lane on an arterial roadway, and thus wide enough to pass unimpeded. Moreover, in nearly all of the cases examined, parking is on the right side, meaning that contraflow cyclists ride along the curb, with no danger of being "doored" by a parked car, increasing the sense of safety. Streets with this profile have both greater traffic volumes (inferred from the street's position in the road network) and speed (inferred from both the greater channel width and lack of Stop signs where they met other local streets) than high contraflow use streets with the single channel profile. However, the lack of needed interaction with traffic, apart from following the "keep to the right" convention, makes the traffic stress low in spite of higher traffic volume and speed. Indeed, some of the high contraflow use streets with the dual channel profile are classified as collector streets ("tertiary" in OSM).

What makes these high-contraflow use streets attractive for contraflow from a network perspective? We found that some of the

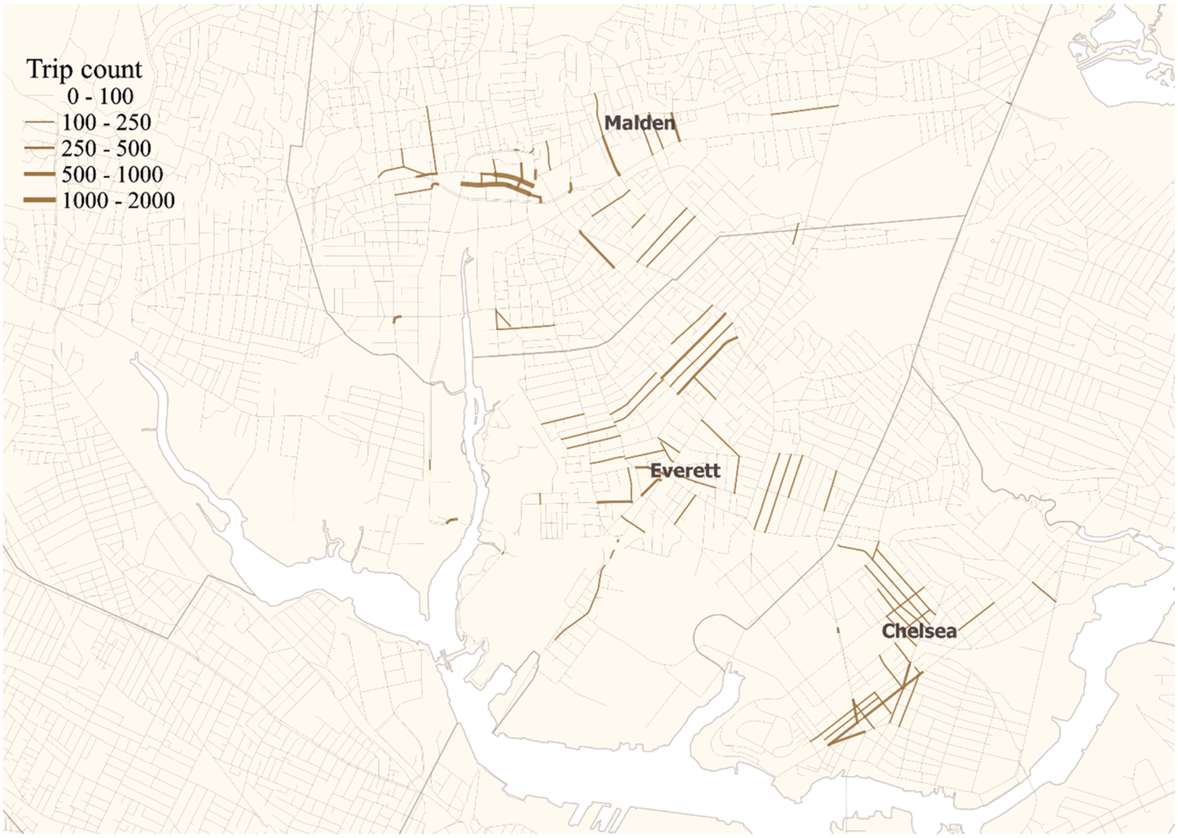


Fig. 6. Street segments with at least 100 contraflow bike trips.

streets with high contraflow use are purely local streets which are *not* good shortcuts to anywhere except to abutting homes, and that they tend to have equal bike volumes in both directions. This reveals that when leaving or approaching one's home, most cyclists choose the direction that most directly leads to their destination – as would a pedestrian – without regard for one-way restrictions.

However, most of the streets with high contraflow use have a position in the road network that makes them an attractive direct route for non-local travel. Three patterns for such streets can be identified. Some are streets that afford a connection where the road network is sparse due to a linear barrier such as a stream or rail corridor; for example, if a cyclist can't use Malden's Franklin Street, it is a considerable detour to the next street that crosses a (former) rail corridor. Several are one-way streets that are a continuation of a two-way street, such as Everett's Norwood Street and Glendale Street. A third group is quiet streets parallel to a higher stress through road, such as Malden's Exchange Street, Boylston Street, and Upham Street. In those cases, riding contraflow allows cyclists to use a lower stress route.

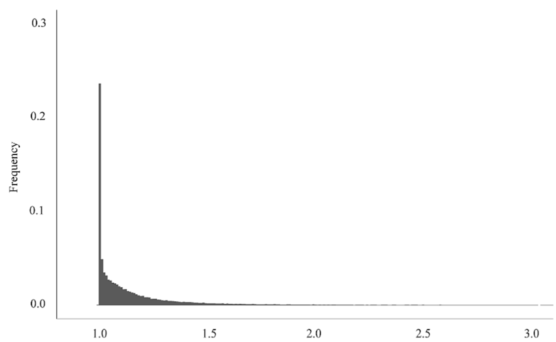


Fig. 7. Distribution of detour factor (trip length / shortest path length).

3.4. Route directness and detour

Previous research has shown that cyclists have a strong preference for a direct route. Researchers have found that people are willing to ride up to about 20 percent further to use a route that is more pleasant or less stressful (Broach et al., 2012; Khatri et al., 2016). Lime's pricing scheme, in which riders are charged per minute, adds a further incentive to use the shortest route. Fig. 7 shows the distribution of detour factor, defined as the ratio of trip length to the shortest path length. For this purpose, the shortest path was determined by Graphhopper using the same algorithm (profile) and OSM network to which the GPS traces had been snapped i.e., allowing travel in parking lots, on footpaths, and in both directions on one-way streets.

The majority of trips are direct: 58% have a detour factor less than 1.10, and 72% have a detour factor less than 1.20. The detour factor distribution has a long tail because it includes trips that are tours with nearly the same start and end location and trips with long deviations that reflect recreation or exploration rather than seeking a utilitarian route to a destination. It is not likely that high detour factor trips represent round trip errands such as shopping because locking a bike ends a trip (as Lime defines it), and because the per-minute user fee gives riders an incentive to end the trip and then start a new one after their shopping is done.

3.5. Use of streets with different levels of traffic stress

Fig. 8 maps Lime bicycle volumes across the service area by level of traffic stress, with volume indicated by line thickness and LTS by color. The thick, green lines that stand out represent popular bike paths. The large number of long red and purple lines, not as thick, represent high stress roads that many cyclists are using, presumably because in many corridors, they are the only practical route to get to one's destination. Only a few blue lines stand out, reflecting the relatively low number of streets in the service area with LTS 2 (two-lane streets with bike lanes).

Fig. 9 shows the distribution of trips by the proportion of the trip's distance that is on low-stress segments, i.e., LTS 1, 2, and 11. The fraction of trips whose length is 100% low-stress is only 7%. While this low figure could reflect the absence of an aversion for high-stress streets, it more likely reflects a property of urban areas with an underdeveloped bike network, discovered by Furth et al. (2016), that even though the street network may have many miles of low-stress local streets, they tend *not* to connect in a coherent network because road networks are deliberately designed to prevent through traffic from using local streets. This tendency is especially strong where, as in greater Boston, the road network is not a rectangular grid. That forces cyclists, like motorists, to use collector and arterial roads for all but the shortest of trips. These results are consistent with a study of the core communities of the Boston metro area (Boston,

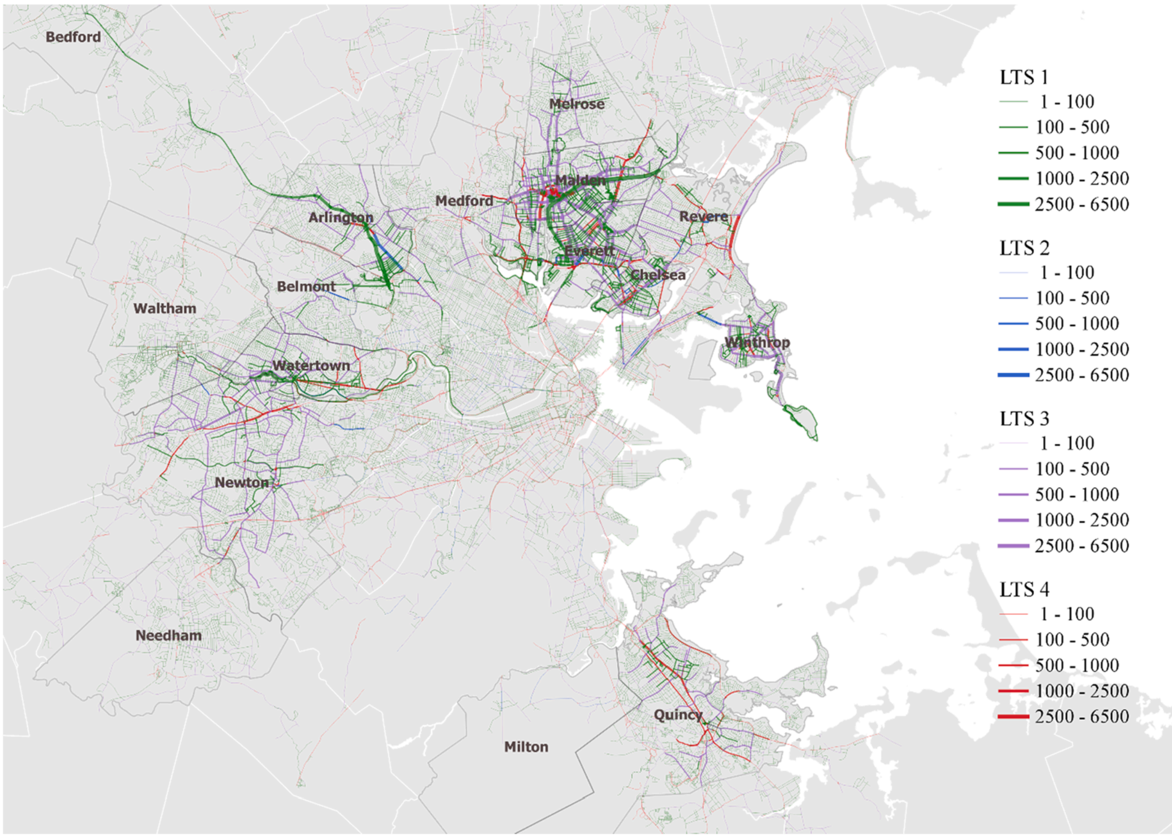


Fig. 8. Lime bike volume across the service area by level of traffic stress.

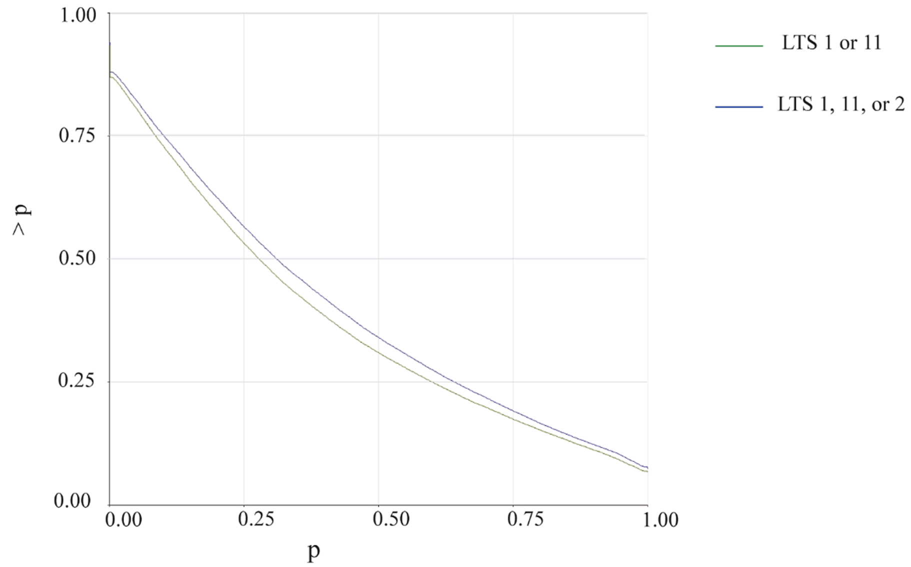


Fig. 9. Distribution of trips by proportion of trip length found on low-stress links.

Brookline, Cambridge, and Somerville) which found that, even with contraflow allowed on all local streets, a low-stress connection was available for only 8.7% of home-to-job location pairs (Putta and Furth, in press).

Another reason for the high use of high-stress links may be that in the Lime service area, riding on sidewalks is common, and by doing so one can ride on something coded as a high-stress link without actually being in high-stress traffic. Recall that for many roads with sidewalks, the sidewalks are not coded in OSM as distinct features. In Massachusetts, sidewalk bike riding is legal except in business districts.

Fig. 10 shows the proportion of bicycle-miles ridden by level of traffic stress. For trips with level of detour less than 1.25, suggestive of a utilitarian trip purpose, 35 to 40 percent of bicycle-miles are on low-stress links. Where level of detour exceeds 2, suggestive of recreational travel, around 60% of the miles ridden are low stress. The distribution of bicycle-miles by LTS varies little between periods of the week, though we found a little more use of high-stress segments off-peak and on weekends. This may be because roads have less traffic and are therefore less stressful during those periods compared to weekday peak hours.

Comparing the route cyclists actually took against the shortest path reveals some evidence of preference for low stress routes. This analysis is limited to trips whose detour factor is less than 1.25, since the concept of shortest path is meaningless for a loop trip or a trip with a deliberate deviation. Average trip length is 1768 m (1.10 mi), which can be decomposed into a shortest path distance of 1657 m and an incremental distance of 111 m. Fig. 11 shows the distribution of this incremental distance by level of traffic stress. The incremental LTS 4 distance is negative, meaning that compared to the shortest path LTS distribution, LTS 4 distance declines. And while LTS 3 distance increases, that increase is less than the proportionate increase in LTS 1 and 2. Thus, by deviating from the shortest path, people are avoiding some riding on very high stress roads and are increasing the proportion of their trip on low-stress roads.

4. Conclusion

Dockless bikeshare is a relatively new mode of micromobility about which little was previously known regarding travel patterns. In addition, the GPS tracking feature of dockless bikeshare makes it possible to track routes, offering insights into cyclist behavior and preferences that may apply to cyclists in general.

Only 15% of dockless bikeshare trips begin or end at a transit station, showing that bikeshare serves many more functions than simply a first- or last-mile transit access function. More than 40% of trips begin or end at a town or village center, which may indicate bikeshare's value in connecting people to local businesses. More people end trips in residential areas than started them, while the reverse is true in commercial centers, resulting in an imbalance that the bikeshare system operator has to correct by periodically moving bikes.

Cyclists have a strong tendency toward using the most direct route. Cyclists often use sidewalks, other footways, and parking lot driveways, and 36% rode the wrong way on one-way streets; in short, they tend to route themselves as would a pedestrian. Other than people making touring trips, most cyclists deviate little, on average, from the shortest path. Where they deviate, it results in less riding on very high stress links and more on low stress links.

Contraflow turns out to be very common. 36% of trips included a contraflow segment on a simple one-way street, accounting for 6% of all bike-miles. One-way streets with high contraflow volumes are nearly all single-lane streets. They include low volume, low speed residential streets with parking on both sides and a travel channel so narrow that when a contraflow cyclist meets an oncoming vehicle, one of them must pull over to let the other pass; they also include streets with greater traffic speed and volume (though still only a

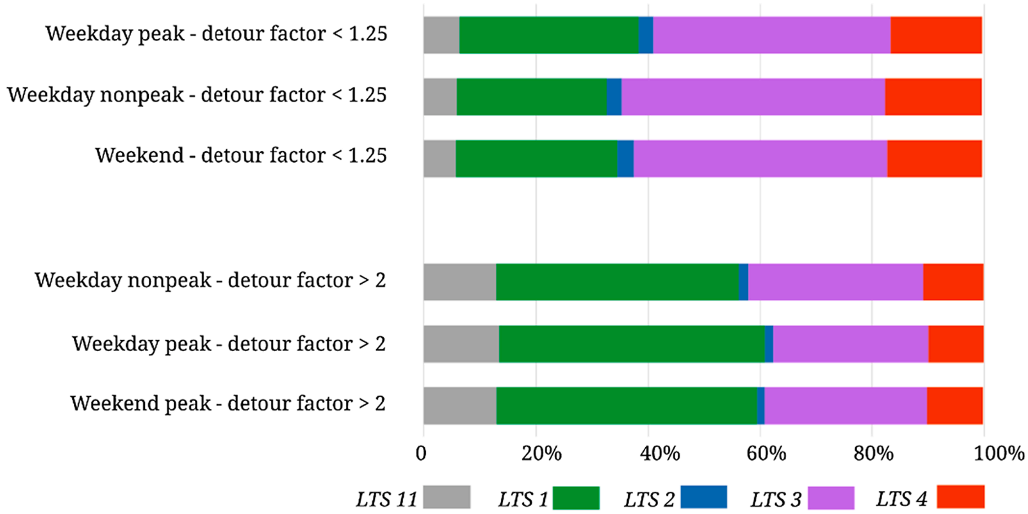


Fig. 10. Percentage of bicycle-miles ridden by level of traffic stress, by time of the week, for low-detour and high-detour riders.

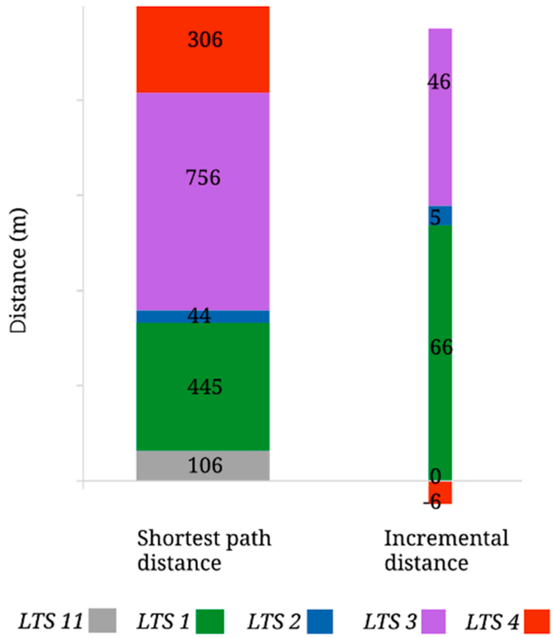


Fig. 11. Distribution of shortest path and incremental distance ridden by level of traffic stress.

single travel lane), but with a travel channel wide enough for a bike and car to get past one another without interaction.

The data show that most trips entail travel on high-stress roadways; only a small fraction of trips were completed using low-traffic-stress links only. This reflects three factors. One is the underdeveloped nature of the bike network, in which low stress links do not connect to form a coherent and direct network. A second factor is cyclists' strong preference for direct routes; it is not sufficient to provide low-stress connections if they involve a considerable degree of detour. A third factor is sidewalk use, through which a person can ride on a link classified as high stress yet actually be physically separated from traffic. Sidewalk use in the Lime service area is something known well anecdotally and can also be inferred from the data in the form of cyclists riding the wrong way on one side of a divided road.

While these results apply directly to bikeshare users, it is reasonable to believe that they reflect the behavior and preferences of a large population of cyclists and people considering becoming cyclists. From the public policy perspective, in keeping with government goals to improve bike safety and to attract more people to bicycling because of its health, environmental, economic, and equity benefits, several lessons can be drawn. One is a need to provide separated bike lanes on arterial roads, since they are often the only practical route to get from here to there. A second is a need to provide a dense network of direct and connected low stress routes in all directions as opposed to, say, routes in only a select set of corridors or a sparse network that would entail substantial detours. A third is to address the clear desire of cyclists to ride contraflow on one-way streets. Finally, a fourth is to consider promoting electric-assist bikes as a way of increasing the range for comfortable bicycling.

5. Limitations

Our study is not without limitations. While the number of trips analyzed is large, still a lot of trips had data errors or missing information about route followed and were therefore not used in studying contraflow riding, route directness, and the use of streets with different stress levels.

There are several reasons for these losses of data which are known as GPS device errors. The relevant errors to this study include the "PDOP mask" which is defined as the relationship between the error in user position and the error in satellite position limit. GPS data collection is suspended when PDOP rises above a limit. The collected dataset could also suffer from other GPS device limitations including receiver timing synchronization error, riding through an underpass, multipath effects due to tall buildings, and atmosphere refraction. To ensure better data quality, municipalities should require that bikeshare vendors supply sufficient documentation and details on software limitations.

CRediT authorship contribution statement

Bita Sadeghinasr: Investigation, Software, Formal analysis, Writing – original draft, Visualization. **Armin Akhavan:** Investigation, Software, Visualization, Formal analysis. **Peter G. Furth:** Conceptualization, Methodology, Writing – original draft, Visualization, Supervision. **Steven R. Gehrke:** Conceptualization, Writing – review & editing, Supervision. **Qi Ryan Wang:** Conceptualization, Supervision, Writing – review & editing. **Timothy G. Reardon:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgement

None.

References

Akhavan, A., Phillips, N.E., Du, J., Chen, J., Sadeghinasr, B., Wang, Q.i., 2019a. Accessibility inequality in Houston. IEEE Sens. Lett. 3 (1), 1–4.

Akhavan, A., Gately, C., Gehrke, S.R., Hydrick, G., Guerrero, J., Reardon T.G., and Sadeghinasr, B., 2019. First Miles: Examining 18 Months of Dockless Bikeshare in Metro Boston. [Online] Available at: https://storymaps.arcgis.com/stories/f9c8e9cddc444dd7a47a678158fd3580.

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C.R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J.J., Simini, F., Tomasini, M., 2018. Human mobility: Models and applications. Phys. Rep. 734, 1–74.

Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transport. Res. A: Pol. Pract. 46 (10), 1730–1740.

Crist, K., Schipperijn, J., Ryan, S., Appleyard, B., Godbole, S., Kerr, J., 2019. Fear factor: level of traffic stress and GPS assessed cycling routes. J. Transport. Technol. 09 (01), 14–30.

Dill, J., McNeil, N., 2013. Four types of cyclists? Examination of typology for better understanding of bicycling behavior and potential. Transp. Res. Rec. 2387 (1), 129–138.

Furth, P.G., 2017. Level of traffic stress. [Online] Available at: http://www.northeastern.edu/peter.furth/criteria-for-level-of-traffic-stress/.

Furth, P.G., Mekuria, M.C., Nixon, H., 2016. Network connectivity for low-stress bicycling. Transp. Res. Rec. 2587 (1), 41–49.

Furth, P.G., Putta, T.V., Moser, P., 2018. Measuring low-stress connectivity in terms of bike-accessible jobs and potential bike-to-work trips. J. Transp. Land Use 11 (1), 815–831.

Gehrke, S.R., Welch, T.F., 2019. A bikeshare station area typology to forecast the station-level ridership of system expansion. J. Transp. Land Use 12 (1), 221–235.

Gehrke, S.R., Akhavan, A., Furth, P.G., Wang, Q.i., Reardon, T.G., 2020. A cycling-focused accessibility tool to support regional bike network connectivity. Transport. Res. D: Transp. Environ. 85, 102388. https://doi.org/10.1016/j.trd.2020.102388.

Gehrke, S.R., Sadeghinasr, B., Wang, Q.i., Reardon, T.G., 2021a. Patterns and predictors of dockless bikeshare trip generation and duration in Boston's suburbs. Case Stud. Transp. Pol. 9 (2), 756–766.

Gehrke, S.R., Russo, B.J., Sadeghinasr, B., Riffle, K.R., Smaglik, E.J., Reardon, T.G., 2021b. Spatial interactions of shared e-scooter trip generation and vulnerable road user crash frequency. J. Transport. Saf. Secur. 1–17. https://doi.org/10.1080/19439962.2021.1971813.

Geller, R., 2016. Four Types of Cyclists. Portland Bureau of Transportation, Portland, OR.

Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.L., 2008. Understanding individual human mobility patterns. Nature 453 (7196), 779–782.

Hosseinzadeh, A., Algomaiah, M., Kluger, R., Li, Z., 2021a. Spatial analysis of shared e-scooter trips. J. Transp. Geogr. 92, 103016. https://doi.org/10.1016/j.jtrangeo.2021.103016.

Hosseinzadeh, A., Algomaiah, M., Kluger, R., Li, Z., 2021b. E-scooters and sustainability: investigating the relationship between the density of E-scooter trips and characteristics of sustainable urban development. Sustain. Cities Soc. 66, 102624. https://doi.org/10.1016/j.scs.2020.102624.

Khatri, R., Cherry, C.R., Nambisan, S.S., Han, L.D., 2016. Modeling route choice of utilitarian bikeshare users with GPS data. Transp. Res. Rec. 2587 (1), 141–149.

Lazarus, J., Pourquier, J.C., Feng, F., Hammel, H., Shaheen, S., 2020. Micromobility evolution and expansion: understanding how docked and dockless bikesharing models complement and compete – a case study of San Francisco. J. Transp. Geogr. 84, 102620.

Mahdinia, I., Mohammadnazar, A., Arvin, R., Khattak, A.J., 2021. Integration of automated vehicles in mixed traffic: evaluating changes in performance of following human-driven vehicles. Accid. Anal. Prev. 152, 106006. https://doi.org/10.1016/j.aap.2021.106006.

Mohammadnazar, A., Arvin, R., Khattak, A.J., 2021a. Classifying travelers' driving style using basic safety messages generated by connected vehicles: application of unsupervised machine learning. Transport. Res. C: Emerg. Technol. 122, 102917. https://doi.org/10.1016/j.trc.2020.102917.

Mohammadnazar, A., Mahdinia, I., Ahmad, N., Khattak, A.J., Liu, J., 2021b. Understanding how relationships between crash frequency and correlates vary for multilane rural highways: Estimating geographically and temporally weighted regression models. Accid. Anal. Prev. 157 https://doi.org/10.1016/j.aap.2021.106146.

Moran, S.K., Tsay, W., Lawrence, S., Krykewycz, G.R., 2018. Lowering bicycle stress one link at a time: where should we invest in infrastructure? Transp. Res. Rec. 2672 (36), 33–41.

NACTO, 2013. Urban street design guide. Island Press/Center for Resource Economics, Washington, DC.

Metropolitan Area Planning Council, 2019. First Miles: Examining 18 Months of Dockless Bikeshare in Metro Boston. [Online] Available at: https://storymaps.arcgis.com/stories/f9c8e9cddc444dd7a47a678158fd3580.

Newson, P., Krumm, J., 2009. Hidden Markov map matching through noise and sparseness. In: Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems (pp. 336-343).

O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. J. Transp. Geogr. 34, 262–273.

Park, Y., Akar, G., 2019. Why do bicyclists take detours? A multilevel regression model using smartphone GPS data. J. Transp. Geogr. 74, 191–200.

Pritchard, R., Bucher, D., Frøyen, Y., 2019. Does new bicycle infrastructure result in new or rerouted bicyclists? A longitudinal GPS study in Oslo. J. Transp. Geogr. 77, 113–125.

Putta, V.V.K.T., Furth P.G., in press. One-way streets and bicycle contraflow: impact on connectivity and accessibility for low-stress bicycling. Transportation Research Record.

Ramm, F., 2017. Routing engines für OpenStreetMap, FOSSGIS, Passau (2017). Accessed 24 Mar 2017. 2017.

Sadeghinasr, B., Akhavan, A., Wang, Q., 2019. Estimating Commuting Patterns from High Resolution Phone GPS Data. In: Computing in Civil Engineering 2019: Data, Sensing, and Analytics. American Society of Civil Engineers, Reston, VA, pp. 9–16.

Semler, C., Sanders, M., Buck, D., Graham, J., Pochowski, A., Dock, S., 2017. Low-Stress bicycle network mapping: the district of Columbia's innovative approach to applying level of traffic stress. Transp. Res. Rec. 2662 (1), 31–40.

Ton, D., Cats, O., Duives, D., Hoogendoorn, S., 2017. How do people cycle in amsterdam, netherlands? Estimating cyclists' route choice determinants with gps data from an urban area. Transp. Res. Rec. 2662 (1), 75–82.

Vogel, P., Greiser, T., Mattfeld, D.C., 2011. Understanding bike-sharing systems using data mining: exploring activity patterns. Procedia-Soc. Behav. Sci. 20, 514–523.

Wang, H., Palm, M., Chen, C., Vogt, R., Wang, Y., 2016. Does bicycle network level of traffic stress (LTS) explain bicycle travel behavior? Mixed results from an Oregon case study. J. Transp. Geogr. 57, 8–18.

Wang, Q., Phillips, N.E., Small, M.L., Sampson, R.J., 2018. Urban mobility and neighborhood isolation in America's 50 largest cities. Proc. Natl. Acad. Sci. 115 (30), 7735–7740.