RMS: Removing Barriers to Analyze the Availability and Surge Pricing of Ridesharing Services

Hassan Ali Khan
hakhan@ncsu.edu
North Carolina State University
Raleigh, NC, USA

Hassan Iqbal
hiqbal@ncsu.edu
North Carolina State University
Raleigh, NC, USA

Muhammad Shahzad
mshahza@ncsu.edu
North Carolina State University
Raleigh, NC, USA

Guoliang Jin
guoliang_jin@ncsu.edu
North Carolina State University
Raleigh, NC, USA

Abstract

Ridesharing services do not make data of their availability (supply, utilization, idle time, and idle distance) and surge pricing publicly available. It limits the opportunities to study the spatiotemporal trends of the availability and surge pricing of these services. Only a few research studies conducted in North America analyzed these features for only Uber and Lyft. Despite the interesting observations, the results of prior works are not generalizable or reproducible because: *i)* the datasets collected in previous publications are spatiotemporally sensitive, *i.e.,* previous works do not represent the current availability and surge pricing of ridesharing services in different parts of the world; and *ii)* the analyses presented in previous works are limited in scope (in terms of countries and ridesharing services they studied). Hence, prior works are not generally applicable to ridesharing services operating in different countries.

This paper addresses the issue of ridesharing-data unavailability by presenting Ridesharing Measurement Suite (`RMS`). `RMS` removes the barrier of entry for analyzing the availability and surge pricing of ridesharing services for ridesharing users, researchers from various scientific domains, and regulators. `RMS` continuously collects the data of the availability and surge pricing of ridesharing services. It exposes real-time data of these services through *i)* graphical user interfaces and *ii)* public APIs to assist various stakeholders of these services and simplify the data collection and analysis process for future ridesharing research studies. To signify the utility of `RMS`, we deployed `RMS` to collect and analyze the availability and surge pricing data of 10 ridesharing services operating in nine countries for eight weeks in pre and during pandemic periods. Using the data collected and analyzed by `RMS`, we identify that previous articles miscalculated the utilization of ridesharing services as they did not count in the vehicles driving in multiple categories of the same service. We observe that during COVID-19, the supply of ridesharing services decreased by 54%, utilization of available vehicles increased by 6%, and a 5× increase in the surge frequency of services. We also find that surge occurs in a small geographical region, and its intensity reduces by 50% in about 0.5 miles away from the location of a surge. We present several other interesting observations on ridesharing services' availability and surge pricing.

CCS Concepts

• **Information systems** → **Web services**; • **General and reference** → *Measurement*; Evaluation; • **Applied computing** → Electronic commerce.

Keywords

Ridesharing services, Gig economy, Uber, Surge, COVID 19

ACM Reference Format:
Hassan Ali Khan, Hassan Iqbal, Muhammad Shahzad, and Guoliang Jin. 2022. RMS: Removing Barriers to Analyze the Availability and Surge Pricing of Ridesharing Services. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3491102.3517464

1 Introduction

Ridesharing services like Uber and Lyft are famous examples of the gig economy and have completed billions of rides around the world [19, 43]. These services offer shorter trip time and lower trip fare compared to other transportation modes [24, 41], create opportunities for people belonging to social minority groups to make a living from already available resources [9], help to reduce the vehicle traffic [8], and reduce carbon emissions [6].

Given the immense impact of ridesharing services on the current socioeconomic fabric, it is essential to make the real-time information of *availability* and *surge pricing* of these services available to various shareholders of ridesharing services, including users, researchers, and city regulatory authorities. The term *availability* in this paper encompasses three aspects of ridesharing services, which includes: *i) supply*, *i.e.*, the number of vehicles driving for that ridesharing service over any given period of time; *ii) utilization*, *i.e.*, the number of vehicles currently in supply that the passengers book over the given period of time; and *iii) idle driving time and distance*, *i.e.*, the time spent and distance traveled by the drivers of ridesharing service while looking for passengers. Regarding surge pricing, ridesharing services dynamically increase the prices of their

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517464

trips during times of low supply or excessive utilization. This concept of a dynamic increase in trip fares amid supply and utilization imbalance is referred to as *surge pricing*.

Publicizing the real-time information of availability and surge pricing can benefit multiple stakeholders of ridesharing services. Drivers of ridesharing services will be informed about the exact high supply and utilization zones at any time of the day, and passengers of such services will be able to see if they can get shorter pickup times or lower trip prices in neighboring locations. More importantly, researchers from various scientific domains (*e.g.*, economy, machine learning, environmental sciences, and logistics management) can analyze the real-time data of the availability and surge pricing, combined with demographic and transportation datasets (*e.g.*, median household income, density population, average paying capacity, fuel prices, and road infrastructure) to guide regulators on making informed polices on reducing carbon emission, avoiding traffic congestion, and reducing pricing bias of ridesharing services. However, ridesharing companies do not share real-time supply, utilization, or surge pricing details with the public. As a result of the unavailability of the related real-time data, it has been difficult for city transport authorities and researchers to measure the impact of ridesharing companies on the transportation paradigm.

Only a few multidisciplinary scholarships analyzed the impact of ridesharing services in the past. Using the limited datasets of availability and surge pricing of certain ridesharing services, studies found that ridesharing services serve more trip requests and offer lower trip prices than traditional taxi services [3, 10, 18, 21, 39]. Regarding the availability bias of ridesharing services in different parts of North America, researchers argued that "whiter" and "richer" neighborhoods in North America observe higher supply and black riders observe longer wait times and higher trip cancellation rates compared to white riders [3, 25]. Ridesharing services multiply floating-point values in the form of 1.× (*surge multipliers*) with the base trip fares to increase the trip prices, which in turn can increase the supply and/or decrease the utilization by offering higher profits to drivers and removing the price-elastic passengers from the demand. Studies that analyze surge pricing found that higher surge multipliers force ridesharing vehicles to go on a "wild goose chase" to pick up distant customers and push the drivers of ridesharing services to drive for long working hours through weekends [5, 7, 32].

Despite their interesting insights, the findings of the related studies are neither reproducible nor generalizable for different socioeconomic regions in the current times because of three limitations: *First*, privately owned ridesharing services are not transparent. They do not disclose the granular details of their vehicle supply, ride utilization, and surge algorithms to the public [21, 34, 49]. The absence of public real-time data is the primary barrier to reproducing or furthering the multidisciplinary research to analyze ridesharing services' availability and surge pricing. *Second*, previous studies presented their findings based on the analyses of small and time-sensitive datasets because ridesharing services do not share their trip level data with the researchers. Authors of related studies relied on restrictive data collection methods, such as scraping outdated official datasets made public by ridesharing services, manual passenger surveys, booking and paying for hundreds of rides, and even driving for ridesharing services [3, 10, 15, 18, 21, 39]. *Third*, findings of prior studies focus only on Uber and Lyft in a few cities of North America. Uber and Lyft in North America are not representative of the usage of ridesharing services around the world because the socioeconomic conditions of North America are not representative of those of other countries in the world.

To address the limitations mentioned above and remove the barriers to understanding and analyzing ridesharing services, we propose an end-to-end open-source data feed system called Ridesharing Measurement Suite (`RMS`). It overcomes the problem of the unavailability of public data of ridesharing services by continuously collecting, analyzing, and exposing the real-time information of availability (*i.e.*, supply, utilization, idle time, and idle distance) and surge pricing of multiple ridesharing services. `RMS` continuously collects the data of the availability and surge pricing of ridesharing services using the web requests used in the official smartphone applications of those services. It then processes the collected raw data and exposes the analyzed datasets to the public. As tasks like deciphering and analyzing the HTTPS web traffic of real smartphone applications and querying the servers of ridesharing services are mainly in the realm of computer science experts, they are hard for researchers of other fields because of the lack of knowledge of operating systems, difficulty to comprehend network programming, and unfamiliarity with the internet security concepts. `RMS` essentially removes these barriers for researchers to analyze the availability and surge pricing of ridesharing services.

`RMS` is the first data feed tool for ridesharing services with two key properties, and it offers two types of user interfaces. The two properties include: *i)* ***Generic***, *i.e.*, any ridesharing service that has a smartphone app can be added in `RMS` for data collection, analysis, and presentation; and *ii)* ***Geolocation Oblivious***, *i.e.*, it is deployable in any city around the world where a ridesharing service operates. The two interfaces with which `RMS` exposes the real-time information of the availability and surge pricing of ridesharing services include: *i)* ***Graphical User Interface,*** which presents the real-time intensity of availability and surge pricing of the ridesharing services in the form of heatmaps on a website; and *ii)* ***Public APIs*** which presents the availability and surge pricing data of ridesharing services in different cities through public JSON APIs. Researchers can consume these APIs in various measurement, comparative, and ethnographic studies exploring the usage of ridesharing services.

Currently, we have configured `RMS` for 10 ridesharing services (*i.e.*, Cabify, Careem, Gett, Heetch, Juno, Lyft, Ola, Shebah, Taxify, and Uber) and deployed it in nine different countries (*i.e.*, South Africa, India, UAE, UK, Australia, Mexico, USA, Canada, and France) to continuously collect, process, and present the availability and surge pricing information of ridesharing services. To demonstrate the applications of `RMS`, we perform a large-scale measurement study to understand the spatiotemporal trends of availability and surge pricing of ridesharing services in pre and during COVID-19 time periods using the data collected and analyzed through `RMS`. The findings of the measurement study are based on **1.4 billion** web API responses from the web servers of the ridesharing services that we collected using `RMS` over four weeks in the pre-pandemic period (**P1:** 04/21/2019 - 05/20/2019) and over four weeks during COVID-19 (**P2:** 01/18/2021 - 02/15/2021).

In general, we conduct a use-case measurement study of `RMS` to cater for the following two shortcomings of related literature: *First*, previous studies overestimated the supply and miscalculated the utilization because they did not address the fact that the majority of ridesharing vehicles stay active in multiple ride categories of the same service simultaneously. For example, a vehicle can be active in both *economy* and *luxury* ride categories of Uber simultaneously. This shortcoming hints that previous articles may have significantly overestimated the supply and miscalculated the utilization of ridesharing services. *Second*, although the current pandemic has severely affected a range of gig-economy businesses, and as a result, ridesharing services have laid off more than 17% of their employees amid COVID-19 [2, 20], we do not find any published article that studies the impact of the COVID-19 on the usage of ridesharing services.

Specifically, we focus on answering five important questions regarding the usage of ridesharing services in our RMS use-case measurement study. *First*, how to infer the availability of ridesharing services in the presence of vehicles that are active in multiple ridesharing services concurrently? *Second*, does the percentage of overlapping vehicles between ridesharing services vary across different countries? *Third*, how do the availability of ridesharing services differ in experiment regions across the time of the day and regions in P1 & P2? *Fourth*, how do frequencies, radii, and lifespans of the surge in ridesharing services vary across cities in pre and during COVID-19 periods? And *fifth*, does the effect of surge on the availability of services vary across cities?

We investigate the collected datasets and make several key observations that are unnoticed in previous works. We report for the first time that up to 40% of ridesharing drivers stay active in multiple categories of a service concurrently. This finding hints that previous studies may have miscalculated the utilization of ridesharing services. During the pandemic, the supply of ridesharing services decreased by 54%, utilization of available vehicles increased by 6%, and surge frequency of services increased by 5×. We also observe that, during COVID 19, Uber lost its popularity of having maximum supply in three major cities: New York, Toronto, and Dubai. Furthermore, the surge multiplier value decreases by at least 50% about 0.5 miles away from the surging location.

Overall, we make the following contributions in this paper:

- We present RMS, a publicly available tool for real-time data collection and analysis of availability and surge pricing data of ridesharing services. It can help the HCI community to evaluate the impact of new design recommendations and conduct related studies in the domain of ridesharing services.
- As a use case of RMS, we present a large-scale exploratory study of the availability and surge pricing of the 10 popular ridesharing services in nine different countries. We further compare these metrics thoroughly across the ridesharing services and the countries.
- Using the data collected, analyzed, and presented by RMS, we are the first to extensively quantify the impact of COVID-19 on the availability and surge pricing of ridesharing services. We make the dataset and analysis scripts used in this study publicly available [28].

Next, we present the related work (§2), carefully designed methodologies for continuous data collection of availability and surge pricing of ridesharing services (§3), and data analysis methodology with a focus on data cleansing methods to remove the ambiguous record in our collected datasets (§4). We present the RMS tool that continuously collects, analyzes, and presents the availability and surge pricing information of ridesharing services (§5). We use the collected datasets in pre and during COVID-19 phases to present the analysis of the availability and surge pricing around the world (§6). We discuss the implications of this work (§7), the limitation and future directions (§8), and ethical concerns of this study and RMS (§9). Finally, we conclude the paper (§10). We also summarize all the terms identified in this study in the appendix (A) at the end of the paper.

2 Related Work

As a result of the unavailability of public data of availability and surge pricing of ridesharing services, prior works employed various methods to collect ridesharing services data and analyzed several aspects of these services. Based on measurement methodologies used in prior work, we classify them into five categories: 1) as a driver, 2) user surveys, 3) booking rides, 4) provider datasets, and 5) smartphone app logging. Below, we discuss the related papers and explain the limitations of each category.

As driver: Henao *et al.* [21] collected the supply and utilization datasets of ridesharing services by becoming a driver for Uber and Lyft. They served 416 rides (Lyft, UberX, LyftLine, and UberPool) and added a survey collecting passenger demographics and income status while analyzing vehicle miles traveled and travel behavior. Caulfield [18] presented an economist account, who became an Uber driver, on economic aspects of the gig economy. Although becoming a driver gives hands-on experience, it is not scalable for large-scale measurement studies because of human intervention.

User Surveys: Rayle *et al.* [39] conducted 380 user surveys to compare taxis and ridesharing services in San Francisco. They collected passenger information such as age, household income, education, and gender, and studied metrics such as trip distance, vehicle occupancy, and wait time. Surveys are effective in relating user responses to the demographics of the users. However, they are not scalable, and the types of survey questions limit their insights.

Booking Rides: Ge *et al.* [14] booked about 1500 rides on UberX, Lyft, and Flywheel, in Boston and Seattle to access the racial discrimination in ridesharing services. They analyzed passenger-specific data, *e.g.*, waiting times, travel times, drivers' cancellation rates, costs, and (where applicable) ratings awarded by drivers to the travelers, and found evidence of discrimination against African American passengers. Although booking rides gives precise information on passenger-centric aspects, this method is prohibitively expensive to scale to multiple services operating in different cities.

Provider Dataset: In [3, 10], authors used public datasets of taxi services and acquired Lyft and Uber data from the respective providers. Brown *et al.* [3] utilized public data on taxi service provided by the Department of Transportation and requested data from Lyft that corresponded to the time frame of taxi service data (3 months). They studied the geographic distribution, individual usage, and explored evidence of racial or gender discrimination on ridesharing and taxi services. However, as noted by the authors, provider data is not comparable to other ridesharing services, and it also lacks information on the demographics of users. Cramer *et* *al.* [10] used public data on taxis and acquired limited data from Uber to study capacity utilization rate for taxis and UberX in Los Angeles and Seattle. They merged the data for multiple ride categories of Uber, such as UberX, UberXL, and UberSelect, into a single category making the data coarse-grained, and their observations do not represent the availability of specific Uber categories.

Smartphone App Logging: One body of work collects data by emulating ridesharing applications [7, 25] or by logging events in these applications [15, 16]. Chen *et al.* [7] emulated the smartphone app of Uber to collect data that the app presents to its users. They installed 43 emulated instances of Uber application on different GPS coordinates to collect data in downtown San Francisco and Midtown Manhattan NYC. They studied metrics such as surge multipliers, estimated wait times, car supply, and passenger demand for all categories of Uber. App emulation is an effective method for continuous data collection without human intervention. However, Chen *et al.* [7] only studied Uber, collected data for only two cities of the US, and did not provide a public tool or public data set. In [25], Jiang *et al.* collected ride-level traces from Uber and Lyft in San Francisco for 40 days and New York City for 27 days. They compared the data of Uber and Lyft with taxis and studied spatial-temporal aspects of supply and demand. Guo *et al.* [15] used application event logs and public data of taxis and trained neural network for surge prediction. They found that the average surge multiplier of geolocation is related to the hour-of-day, the day-of-week, and the location itself.

In terms of measurement methodology, our work aligns with the app emulation approach [7, 25] in that `RMS` uses web traffic traces of smartphone applications. However, `RMS` surpasses previous works because: *i)* it is a *generic* platform capable of measuring different ridesharing services; *ii)* it is *geolocation oblivious* as it is not affected by the underlying city architecture and navigation systems, *i.e.*, it is deployable in any city with any urban design without modifying the presented algorithms; *iii)* it is live as it continuously provides real-time data on the configured ridesharing services; and *iv)* it is publicly available.

3 Data Collection Methodology

This section explains the data collection methods used in `RMS`.

3.1 Overall Design

`RMS` uses HTTP/S web requests used in the applications of ridesharing services (presented in Table 1) for data collection. We add web requests in the data collection scripts of RMS with six steps. *First*, we install an Android-5.0 emulator (NOX [36]) on a computer. *Second*, we install and execute Charles proxy server [46] on that same computer. An IP address is automatically assigned to the Charles web proxy server program whenever we execute it. *Third*, we set up the web proxy on the Android emulator by adding the Charles proxy server's IP address in the Android emulator's network settings. Charles proxy server, which now sits between the Android emulator and the internet, passes web requests of the applications installed on the Android emulator to the internet and receives the responses from web servers on their behalf. *Fourth*, we install and execute the selected ridesharing applications on the Android emulator and log applications' web traffic traces, *i.e.*, record all the web requests that applications make to exchange data with their web servers. Each HTTP/S web request consists of a request line, header parameters, and body content. *Fifth*, we find the specific web requests that `RMS` will use to collect the availability and surge pricing data by manually reading the logs of each application's web traces. *Last*, we add the recognized web requests in the `RMS` data collection requests repository. `RMS` will not require the emulator and web proxy set up to collect the data of ridesharing services after this step. The data collection scripts of `RMS` periodically call web requests from its repository to collect various ridesharing services' availability and surge information in different parts of the world.

Table 1: Cities (short forms), services (short codes), and categories. The economic categories of ridesharing services are mentioned in bold style. Services with * sign quit their operations in P2.

City	Service	Categories
Cape Town (CPT)	Uber (U)	**X**, Assist, XL, Black
	Taxify-Bolt (T)	**Bolt**, Plus
Delhi (DEL)	Uber (U)	Pool , **Go** , Premier, XL
	Ola Cabs (0)	Mini, **Prime**, Play, Lux
Dubai (DXB)	Uber (U)	**Black**, Lux, XL
	Careem (C)	Go, **Economy**
London (LDN)	Uber (U)	Pool, **X**, Assist, X+, XL, Black
	Gett (G)	**Taxi**
Melbourne (MEL)	Uber (U)	Pool, **X**, XL, Lux, Assist
	Taxify-Bolt (T) *	**Bolt**
	Shebah (S)	**Economy**
Mexico City (MEX)	Uber (U)	Pool, **X**, Assist, XL, Black,
	Taxify-Bolt (T) *	**Bolt**
	Cabify (CB)	**Economy**
New York City (NYC)	Uber (U)	Pool, **X**, Car Seat, WAV, XL, Black
	Lyft (L)	Shared, **Lyft**, XL , Lux, Lux XL
	Juno (J) *	**Bliss**, Lux, SUV
Paris(PAR)	Uber (U)	Pool, **X**, WAV, Premuim, Van, Green, Berline
	Taxify-Bolt (T)	**Bolt**, Berline
	Heetch (H)	**Economy**
Toronto (YTO)	Uber (U)	Pool, **X**, XL, Black, Black SUV, Assist, WAV
	Lyft (L)	Shared, **Lyft**, LX, Lux, Black, Black XL

To view the encrypted HTTPS traffic of Android emulator applications in simple text, we add a self-signed private SSL certificate on the local machine that has the Charles web proxy installed. We add a corresponding self-signed public SSL certificate on the Android emulator as a trusted certificate. However, some of the mobile applications use the *SSL pinning* technique as an additional security layer to stop app-server web communication in the presence of a self-signed public SSL certificate. To remove the SSL Pinning checks in the ridesharing applications, we decompile the smartphone applications to get their source codes, then comment the SSL pinning checks in the source codes of applications using *Frida Hooks* [26] and *SSLUnpinning* module of *Xposed* framework [50], recompile the instrumented source codes, and install the patched application on the Android emulator.

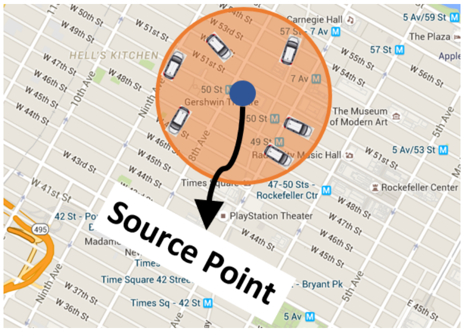

Figure 1. Illustration of data collection source point and nearby vehicles

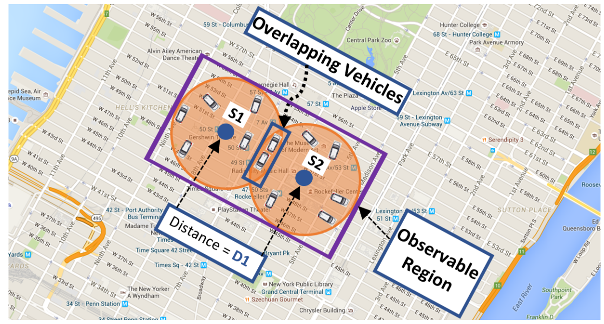

Figure 2. The distance D1 between two source points at times T1 in Midtown Manhattan NYC (with at least one overlapping car in the nearby vehicles datasets of both source points)

3.2 Blanketed Region

RMS collects the data of the availability and surge pricing of ridesharing services from the selected regions of cities that are mentioned in Table 1. We refer to the data collection region in each city as the *blanketed region*. To observe the *i)* maximum traffic of ridesharing vehicles and *ii)* effect of surge pricing in our pandemic effect use-case study, we collect the data of ridesharing services in the downtown region in every city (Table 1) in both P1 & P2. Blanketed regions of the cities in RMS are Sea Point in CPT-South Africa, Gole Market and Connaught Place in DEL-India, the vicinity of Dubai world trade center and Al Satwa community in DXB-UAE, the vicinity golden square in LDN-UK, the area of Melbourne Town Hall in MEL-AU, the city center in MEX-Mexico, Midtown Manhattan in NYC-USA, the vicinity of Cluny Museum in PAR-France, and the surroundings of the University of Toronto - St. George Campus in YTO-Canada. In each city, we keep blanketed region the same for all the available services of that city in both P1 & P2.

3.3 Nearby Vehicles

The smartphone applications of ridesharing services make HTTP/S web requests along with the user's geolocation in the request headers to their web servers to retrieve the real-time locations of vehicles that are near to the user (known as *nearby vehicles*) [22, 42]. The most common data values that we receive in response to nearby vehicles web requests by the servers of ridesharing services in our experiment are: *i)* a list of nearby vehicles in different ride categories (e.g., economy, luxury, etc.); *ii)* their vehicle ID and their geolocations; and *iii)* the timestamp of response.

RMS invokes the nearby vehicles web requests of each **service-city instance** (e.g., Uber-CPT and Lyft-YTO) from 50 different geolocations (also referred to as *source points*) within the blanketed region of each city. An illustration of a source point and nearby vehicles close to the geolocation of the source point is shown in Figure 1. The sampling rate of nearby vehicles request for every service-city instance is 10 requests/minute at every source point.

3.4 Placement of Source Points

Deciding on the appropriate geographical distance between the adjacent source points of data collection is a crucial step. If we collect data from locations too far apart, we will not observe all the available vehicles within the blanketed region.

To find the proper distance between source points, we collected the data of nearby vehicles of every service-city instance from two different source points on 03/22/2019 with the sampling frequency of 10 requests/minutes at each source point. For every service-city instance, we began with collecting the data of nearby vehicles with the distance between the two source points as 0.3 miles in the blanketed region in every hour. After that, we gradually decreased the distance between the source points by 100 feet until we found at least one overlapping economy category vehicle in the collected datasets of both the source points. Finally, we saved all the geographical distances between the source points with at least one overlapping vehicle. We used the minimum of all such distances to be the distance between the adjacent source points for each service in each city in P1. Figure 2 presents an example of finding the appropriate distance between data collection source points (S1 and S2) at time T1 of the day in Midtown Manhattan NYC, where D1 represents the least distance between source points with at least one overlapping vehicle in the collected datasets of S1 and S2.

We followed the same protocol on 01/16/2021 to calculate the appropriate distance between the adjacent source points of available service-city instances for P2. It helped by ensuring we did not miss any ridesharing vehicle within the blanketed region of every city in pre & during COVID-19 periods.

Although data collection source points of RMS are currently installed in each city's downtown, they can be installed in any geographical region and will not require editing the presented algorithms. We may observe the quantitative differences depending on the location of source points, but the placement of source points will not affect the data quality.

3.5 Observable Region

The *observable region* of any ridesharing service within a given time interval represents a polygon with its area as the product of distances between the horizontally and vertically farthest observed vehicles within that period. In other words, the *observable region* of any service-city instance represents the geographical area in which we can trace every available vehicle of that service. It is important to notice that as the number of source points stays fixed at 50, the *observable region* varies from time to time for each service-city instance as per the density of nearby vehicles (e.g., vehicles are dense during the rush hours but are sparse at 4 AM).

The purple-colored rectangle represents the observable region in Figure 2 in which we can trace every vehicle. Throughout the experiment, we find Uber's observable region as the smallest among the competing ridesharing services in each city. The averages of observable regions of Uber in our experiment are 7.1 mi^2 in P1 & 8.9 mi^2 in P2. In both P1 & P2, we observe a significantly higher number of Uber vehicles in NYC, YTO, and MEL than in other cities. Hence, the areas of observable regions of Uber in the mentioned cities are relatively smaller as compared to the other experiment cities, *i.e.*, NYC (P1: 5.58 P2: 7.37), YTO (P1: 5.16, P2: 7.74), and MEL (P1: 6.06, P2: 7.18).

4 Data Analysis Methodology

This section defines the data variables that RMS returns and discusses methods to clean and analyze the collected datasets.

4.1 Supply

In this paper, the *supply* of a service refers to the number of observed unique vehicles of that service within any given time interval. It corresponds to the total number of unique vehicles that show up in the nearby vehicles web requests of all the data collection source points of any service-city instance within a specific time window.

Out of 10 services that we study, we face issues identifying unique cars in the nearby vehicles datasets of three services, *i.e.*, Uber, Lyft, and Taxify. In the coming paragraphs, we explain the methodology RMS uses to identify unique cars for these three services. In Uber, in response to nearby vehicles web requests invoked at different source points, the same vehicles appear with different random encrypted IDs. For example, a same vehicle *(V1)* appears as *Car-1* at source point *S1* and as *Car-2* at source point S2. We find the same issue in Lyft's nearby vehicles datasets. This issue makes it difficult to count the total number of unique vehicles of these services. To classify multiple vehicle IDs as the same vehicle in the mentioned service, RMS uses the following algorithm. For each source point S_i in each city, it compares each vehicle ID's trajectory (a set of $\langle geolocation$ and $timestamp\rangle$ tuples) observed in S_i with the trajectories of vehicles observed in the source points that are adjacent to S_i. Then, if it finds overlapping parts in the trajectories of any two vehicles, it assumes that both IDs represent the same vehicle. Finally, RMS removes the false positives in our processed supply datasets by checking if two vehicles were simultaneously present at different geolocations, even if they have overlapping trajectory parts. This algorithm enables RMS to obtain the tight upper-bounds of the total number of vehicles for Uber and Lyft.

Taxify randomizes IDs of available vehicles after every minute, making it challenging to count the unique number of vehicles. By comparing the geolocation and bearing attributes of the first occurrence of a new ID with that of all the vehicles observed in all the source points right before the occurrence of the new ID, RMS identifies the new assigned ID of the vehicle. To obtain a realistic approximation of the total number of Taxify vehicles, we avoid adding any heuristic in RMS (except the distance between geolocation of "old" ID and the "new" ID to be less than 200 feet within a 6 seconds interval).

RMS also removes the vehicles from the supply datasets that remained online for less than 30 seconds. They account for only 12% of the total vehicles recorded in our experiment. Such vehicles were also observed in a previous study and referred to as short-lived vehicles [7]. A possible reason for observing short-lived vehicles is that the servers of ridesharing services in our experiment return only a specific number of vehicles close to a source point in their nearby vehicles web requests, *e.g.*, Uber servers return a maximum of eight nearby vehicles. It may be the case that short-lived vehicles are replaced with closer vehicles in the subsequent requests at the geolocation of a source point. We observe 87% of the short-lived vehicles in the datasets of source points close to the horizontal or vertical boundaries of blanketed regions. This observation makes it easy to understand that most short-lived vehicles were driving around or near to experiment areas briefly.

4.2 Utilization

For any service-city instance, vehicles from the supply that go offline within the observable region in a given time interval are referred to as *utilized* vehicles in this paper. We infer the utilization of ridesharing services by counting the number of vehicles that go offline within the observable region within a specific time window.

The inferred utilization in this paper is actually the *fulfilled* demand, *i.e.*, the number of vehicles that go offline, because none of the ridesharing services provide public data about the *quantity* demanded (the number of passengers that request rides). The presented utilization in our study represents the upper bound of actual demand since it is possible that some vehicles go offline not due to picking up the passengers but because the drivers are done for the time being and turn off the ridesharing applications. Next, we discuss our data cleaning approaches for utilization datasets.

A vehicle can go offline from the observable region for multiple reasons, *e.g.*, a driver turning off the rider application or going outside the experiment area. We cannot assess if a vehicle goes offline for the former reason. So, the utilization presented in the coming sections is an upper bound of the real utilization. To cater to the effect of the latter reason, RMS does not count those vehicles in utilization that go offline with the distance $\leq$ 500 feet from horizontal or vertical boundaries of the observable region.

While the vehicles of Uber, Lyft, and Taxify with transient random IDs cannot reappear in our observable regions after the completion of rides, the vehicles of services with persistent vehicle IDs (where a unique ID is observed for each vehicle in the entirety of our data collection periods) may reappear in the observable regions within a very short period. In such cases, based on previous findings [44], RMS treats $\geq$10 minutes gap between the two consecutive occurrences of the same vehicle ID as the completion of a ride.

Previous studies do not mention that the significant percentage of drivers in most of the ridesharing services drive in multiple ride categories concurrently (*e.g.*, the same vehicle being available in *UberX* and *UberXL* at the same time). There is no way to ascertain that a vehicle active in multiple categories goes offline with the trip request of which particular type. Counting such vehicles multiple

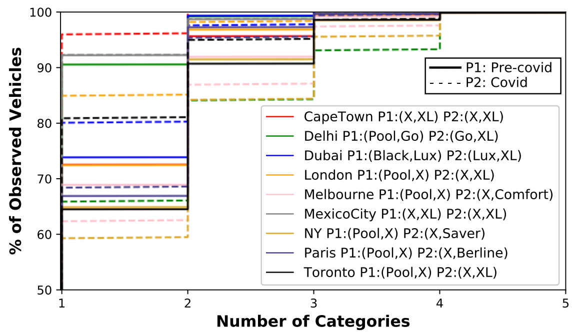

Figure 3. Distribution of Uber vehicles that are observed in one or more categories. A tuple with every city represents the most frequently repeated pair of categories in P1 & P2.

times in the utilization of available categories may result in miscalculating the utilization of ridesharing services. This finding hints that previous studies [7, 25] may have miscalculated the utilization of ridesharing services as they did not report the percentage of vehicles available in multiple categories of the same service.

Figure 3 shows the percentage of Uber vehicles active in one or more categories simultaneously during P1 & P2. We observe that 15%-40% of the total Uber vehicles in seven out of nine cities are active in multiple categories except for DEL and MEX in P1 and CPT and MEX in P2. The most frequently repeated pair of categories for vehicles driving in multiple categories are the two least expensive available ride options in each city in both P1 & P2. To find a close approximation to the true utilization of each category **C** in each service-city instance, `RMS` runs the following protocol: *i)* represent all the vehicles in category **C** with a set M; *ii)* represent vehicles in set M that are only available in category **C** and not in any other category with a set M' where $M' \subset M$; *iii)* count the number of vehicles in M' that go offline within the observable region, represented as n; and *iv)* scale the utilization count of category C as:

$$C_{Utilization} = n \times \frac{|M|}{|M'|}$$

Next, some drivers use multiple ridesharing services, *e.g.*, the same vehicle being active in both Uber and Lyft concurrently [12]. Upon receiving a trip request from any of the services, they usually turn off the other applications. Since there is no way to know which service offered a ride to such vehicles, they might be double-counted in the utilization for multiple ridesharing services. To identify the vehicles driving in multiple services, `RMS` first divides the observable region of each city into multiple 50 $feet^2$ blocks. Then, it observes whether the two vehicles available in different services follow the same trajectory of blocks during the same time window (with the maximum delta of ±6 seconds). On the positive outcome of the above-mentioned condition, `RMS` considers the vehicle available in multiple services to be the same.

Table 2 presents the average percentage of **daily** overlapping vehicles between pairs of all the available services in P1 & P2. Standard Deviation (*SD*) quantifies the spread of the number of overlapping vehicles between the services that are observed during each day of each phase (*i.e.*, P1 & P2). On average, we find 2% of the vehicles in P1 and 3.45% in P2 to be active in multiple ridesharing services. We notice ≥ 4% overlapping vehicles in some cities (NYC and YTO) while ≤ 1% in others (MEL and MEX). We also observe that Shebah (the only women ridesharing service) in MEL has < 1% overlapping vehicles with other ridesharing services. Overall, we find that the number of overlapping vehicles in multiple services during the pandemic is 1.7X greater than in 2019. Since the percentage of vehicles active in multiple ridesharing services concurrently is relatively small, `RMS` excludes them from the availability analysis.

Table 2: Percentages of daily shared economic category vehicles between the service-city instances. NA represents that either or both of the services stopped services in P2.

City	Services	Shared Vehicles	
		P1: Mean ± SD	**P2:** Mean ± SD
CPT	Uber - Taxify	2.89% ± 0.69%	3.19% ± 0.71%
DEL	Uber - Ola	2.81% ± 1.99%	4.32% ± 1.11%
DXB	Uber - Careem	3.45% ± 0.63%	5.26% ± 0.42%
LDN	Uber - Gett	3.28% ± 0.97%	3.93% ± 0.31%
MEL	Uber - Shebah	0.02% ± 0.01%	0.44% ± 0.02%
	Uber - Taxify	1.09% ± 0.45%	**NA** (T-MEL is not available in P2)
	Shebah - Taxify	0%	**NA**
MEX	Uber - Cabify	0.11% ± 0.06%	0.87% ± 0.47%
	Uber - Taxify	0.17% ± 0.09%	**NA**
	Cabify - Taxify	0.55% ± 0.37%	**NA**
NYC	Uber - Lyft	4.03% ± 0.72%	7.12% ± 1.51%
	Uber - Juno	3.69% ± 0.72%	**NA**
	Lyft - Juno	3.77% ± 0.61%	**NA**
PAR	Uber - Taxify	1.05% ± 0.3%	1.77% ± 0.62%
	Uber - Heetch	2.31% ± 0.92%	4.45% ± 0.39%
	Taxify - Heetch	0.06% ± 0.03%	0.81% ± 0.1%
YTO	Uber - Lyft	4.93% ± 1.44%	5.66% ± 1.23%

4.3 Idle Time and Idle Distance

For all the utilized vehicles that go offline in availability datasets of `RMS`, the time spent and distance traveled by those vehicles before going offline to serve a trip request are referred to as *idle time* and *idle distance*, respectively. To calculate idle time and distance, `RMS` records every epoch from the first appearance of a utilized vehicle till it goes offline and then calculates the distance traveled during those epochs using the geolocations returned by the servers of ridesharing service on each timestamp.

4.4 Surge

Ridesharing services apply surge in the form of a **1.X** multiplier on the base fare of their trips. They claim that increasing the trip cost with higher surge multipliers helps attract more drivers and reduces the services' demand in surging areas [38].

The web requests to fetch *Upfront Trip Price Estimate* from a **source** location to a **destination** location allows us to see the value of surge multiplier at the source location at any given time. We find the Upfront Trip Price Estimate web requests used in smartphone applications of ridesharing services in applications' web traces and execute them in `RMS`.

From each of the 50 data collection source points for each service-city instance, RMS periodically invokes the Upfront Trip Price Estimate web request of the respective service, with the geolocations of source points as the **source** and the airport of the respective city as the **destination** in the request, to collect the real-time values of surge multipliers of services at every source point. The sampling rate of the Upfront Trip Price Estimate web request for each service-city instance at each source point is 10 per minute.

Analysis of surge pricing includes the following five attributes in this paper: *i) Surge multiplier value, i.e.*, the floating-point value of surge intensity (in the form of **1.X**) which is multiplied with the base trip fare to increase the trip prices; *ii) Surge instance* represents the continuous-time window in which surge multiplier value starts to increase from 1.0 and drops back to 1.0; *iii) Surge frequency* represents the frequency of surge instances in every day of the experiment; *iv) Surge lifespan* represents the average length of surge instances in minutes; and *v) Surge radius* represents the average geographical radius (in miles) of surge instances.

5 Ridesharing Measurement Suite (RMS)

There are two primary components of RMS, i) data collection and analysis module and ii) data presentation module. Scripts of data collection and analysis module of RMS execute on a local computer connected to our university's internet, and the data presentation module is hosted on a shared internet hosting server. The data collection and analysis module of RMS periodically generates the code files of web pages for the presentation module and uploads them on the shared hosting server.

Currently, RMS collects, analyzes, and presents the data of 10 ridesharing services in nine cities, which makes the *working service-city instances set* of RMS. To add a new element in the working service-city instances set, we have to add a five-attributes tuple in that set. The attributes of each tuple are: *i)* the name of the ridesharing service; *ii)* the name of the city; *iii)* the bounding coordinates of the blanketed region (upper left, upper right, lower right, and lower left coordinates) in that city; *iv)* the official web requests for surge pricing and nearby vehicles; and *v)* the list of authentication tokens of that ridesharing service (a unique token that is assigned to each user on signing up) to avoid denial of service and prevent HTTP 429 (Too Many Requests) error responses. To collect the data of working service-city instances set, RMS *i)* parses the set; *ii)* finds the appropriate distance between data collection source points for each service-city instance; and *iii)* starts the data collection scripts for the nearby vehicles and surge pricing for each service-city instance.

Considering budget limitations, we decided to host the data presentation module of RMS on a shared hosting space that allows only a limited number of upload requests per day. Although the data collection and analysis scripts do not incur any significant delay in processing and presenting the live updates of availability and surge pricing of ridesharing services, to lower the bandwidth utilization of the shared hosting server, we decided on the update frequency of RMS as once every 15 minutes. After every 15 minutes of data collection, RMS invokes supply, utilization, and surge analysis modules to analyze the collected data sets by employing the methods presented in §4. After analyzing the collected datasets, RMS publishes the real-time information of supply, utilization, and

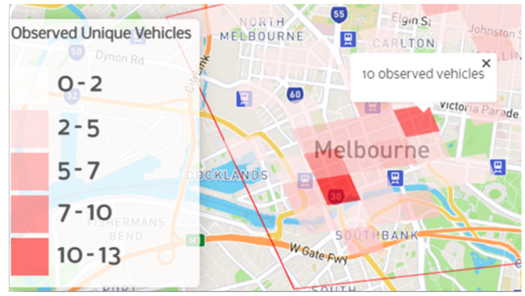

Figure 4. Supply heatmap of Uber-MEL in RMS on 02/27/2021

surge attributes (multiplier value, lifespan, and radius) of ridesharing services on two different interfaces on hosting space to serve the stakeholders of ridesharing services. It publishes the results *i)* on a Graphical User Interface [31], representing the intensity of the mentioned data variables using heatmaps in HTML5 and *ii)* with public JSON-based APIs [30] that researchers or regulators can use to query the data generated by RMS.

Before the publication of this paper, researchers from the Beijing Institute of Technology (China), which were part of the GEARS 2021 program [4] cohort at North Carolina State University, requested access to RMS APIs and used a set of scripts [29] to save the data returned by RMS APIs continuously. They collected the supply data of ridesharing services for 30 days. They used the collected datasets to train models for the supply prediction of ridesharing services at different times and days of the week. They presented their findings in a poster at the GEARS 2021 poster presentation demo.

Next, we describe how RMS collects, analyses, and presents the supply, utilization, and surge pricing data with examples.

***Supply*:** RMS continuously collects the data of nearby vehicles of each service-city instance from multiple source points with the frequency of 10 nearby vehicles requests per minute at each source point as explained in §3. After every 15 minutes, RMS invokes the supply cleaning module that disambiguates the collected nearby vehicles dataset of that interval using the data cleaning approaches for supply discussed in §4. After that, RMS invokes the supply analyzer module that divides the observation region of each service-city instance into a grid of 500 feet^2 blocks and records the number of unique economic category vehicles observed in each block.

An example of the supply heatmap of RMS for Uber-MEL is shown in Figure 4 that presents the number of unique Uber vehicles observed in different blocks of the experiment area in MEL on 02/27/2021 between 8 PM - 8:15 PM. During that interval, we observed the maximum supply at *King St* and *Lonsdale St.* The supply measurement variable of RMS can help the drivers to know the actual geographical distribution of ridesharing vehicles and avoid the high supply and competition zones.

***Utilization*:** To find the utilization of service-city instances in each 15 minutes interval, RMS invokes the utilization analysis module. For every 500 feet^2 block of the observable region, the utilization analysis module counts the vehicles that go offline in that block before the end of the interval and computes the idle time and distance of the utilized vehicles.

The utilization heatmap of RMS with idle time and distance, for Uber-MEL on 02/27/2021 between 8 PM - 8:15 PM, is shown in

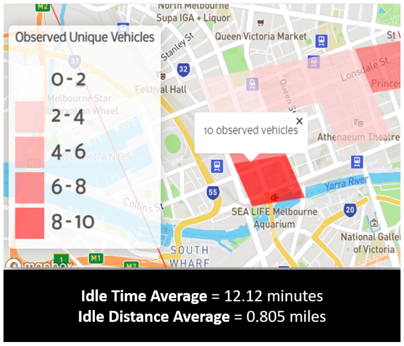

Figure 5. Utilization heatmap of Uber-MEL on 02/27/2021

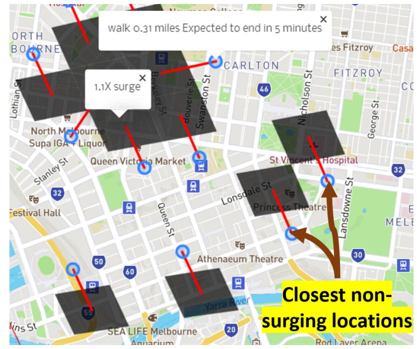

Figure 6. Surge heatmap of Uber-MEL in RMS on 02/27/2021

Figure 5. Like supply, we observe the maximum utilization of Uber at *King St* and *Lonsdale St*. Finding the high utilization areas in experiment cities can increase the chances for drivers of ridesharing services to get more rides.

Surges: RMS continuously collects the surge multipliers information at all the source points for each service-city instance after every six seconds interval. For every surge instance of any services observed at any source point, RMS saves the surge lifespan at that source point with respect to the time of the day to compute the expected surge lifespan based on the surge history at that location.

Figure 6 shows the surge heatmap with nearest non-surging locations (marked with blue circles) and the expected lifespans of surge (based on the previous history) for Uber-MEL at 8:10 PM on 02/27/2021. The surge module's output can help passengers know if they can avoid the surge by walking to a neighboring location or waiting for a few minutes. The presentation of the expected lifespan of a surge instance can also assist drivers in determining if they can reach the surging location before the end of the surge.

6 Measurement Study with RMS

In this section, we present a use-case measurement study of RMS, which quantifies the impact of COVID-19 on the spatiotemporal availability and surge pricing trends of ridesharing services. We analyze the data collected using RMS and present our observations for the availability and surge pricing of economic category vehicles of ridesharing services in our experiment (Table 1).

In the rest of this section, we represent Pearson correlation coefficient [1] as *Pearson's r* to measure the correlation across multiple timeseries of availability and surge pricing datasets. The value of *Pearson's r* ranges from -1 to +1, where 0 indicates that there is no correlation between the given variables while the values greater or less than 0 indicate the positive or negative correlation, *i.e.*, values in the given multiple timeseries grow in the same or opposite directions at the same time. The relationship between the given variables is considered strong (+ve or -ve) when $|r| \geq 0.7$. Furthermore, the notations like *Uber-CPT* represent *service-city* instances.

6.1 Supply

Previous studies found that Uber supply exhibits periodic daily trends, and Uber supply count peaks during the morning and evening rush hours and declines at night [7, 25]. However, it is unclear if the same supply patterns are observable for the ridesharing services across different regions and periods.

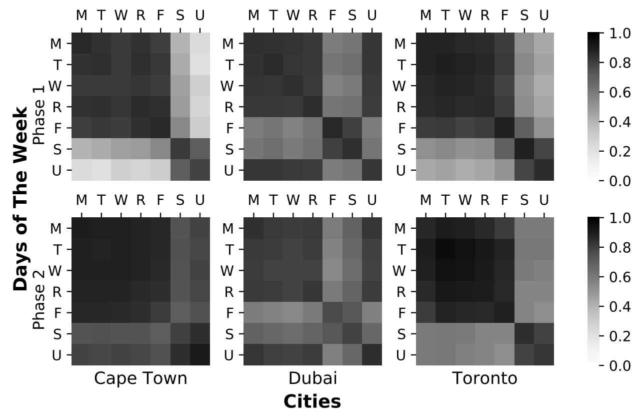

Figure 7. The average Pearson correlation coefficient between the supply of Uber across the days of the week in CPT, DXB, and YTO in P1 & P2

To investigate this, we first examine correlations between the daily supply timeseries. Second, we present the trends of daily supply timeseries by counting all the unique vehicles observed during each **five-minutes** interval of a day for each day of P1 & P2. Third, we compare supply counts and peak supply hours of the service-city instances in P1 & P2.

6.1.1 Temporal Correlation of Supply. To measure the correlation in the temporal behavior of the supply of ridesharing services, we compute the daily supply timeseries of each service-city instance in every five-minutes window. Each daily supply timeseries of service-city instances contains 288 integer values representing supply counts in every five-minutes window of the day.

Figure 7 shows the average values of *Pearson's r* between the daily timeseries of supply of Uber across the days of the week in CPT, DXB, and YTO in P1 & P2. The first observation from Figure 7 is that the timeseries of the supply counts of Uber over an entire day is strongly correlated across weekdays and weekends (*Pearson's* $r \geq 0.76$) but is weakly correlated between weekdays and weekends ($Pearson's\ r \leq 0.59$). This indicates that the timeseries of supply of ridesharing services appear to be different in the two parts of the week, both in P1 & P2 (the official weekend in DXB is Friday and Saturday). We also observe a strong correlation in the timeseries of supply within weekdays and weekends for all of the other service-city instances. We do not present the other services in Figure 7 because of the space limitation. We do not observe a negative correlation between the timeseries of supply for any service-city instance on different days of the week. A negative correlation indicates that we can observe a local maximum and local minimum supply count on two different days of the week at the same time of the day. Finally, the correlation coefficient between the supply of presented services across weekends and weekdays is higher in P2 than P1 because we do not observe a significant increase in the supply of services in the morning rush hours during weekdays in P2, possibly due to the closure of workplaces and colleges amid COVID-19.

6.1.2 Supply Over Time of the Day. Figure 8 shows the average number of vehicles per square mile of the observable regions of the available services that we observe in 24 hours on every day of the week in CPT, DXB, and YTO in P1 & P2. During P1, we observe the diurnal patterns during weekdays in the supply of ridesharing services in all the presented cities. The supply count increases at around 8 AM and decreases from midnight to 6 AM. If we compare the days of the week, we can see that there are two local peaks of supply count on weekdays, whereas, on weekends, there is only one per day, typically around 9 PM.

The evening supply count peaks of ridesharing services during weekdays in LDN, MEL, and MEX are significantly higher (~1.3X) than the morning supply count peaks of available services (figures not shown due to space limitation). We hypothesize that the well-pronounced dips between the two supply count peaks during weekdays in YTO are observed because of the business and school hours during weekdays as the blanketed region in YTO includes a university and some public offices. These observations show that the supply patterns of ridesharing services evolve differently during the day in different cities, and the finding regarding the supply of ridesharing services in one region cannot help model the supply patterns in other areas.

During P2, Figure 8 shows a significantly lower supply of ridesharing services. Overall, we observe 54% fewer ridesharing vehicles in P2 as compared to P1. Although P1 & P2 are conducted in different months and weather conditions, a few related articles that evaluated the impact of weather conditions on the supply and utilization of ridesharing suggest that weather does not significantly affect the ridership of ridesharing services [17, 45]. Our observations, coupled with the previous findings, render the pandemic a probable cause for the decrease in the supply of ridesharing services.

The services being affected most in P2 regarding supply counts are Uber-YTO, Ola-DEL, and Taxify-CPT. We record 60% fewer vehicles of these services in P2 than P1. We see the distinguishable morning and evening supply peaks of ridesharing services during pandemic weekdays in just four out of nine cities (YTO, DEL, MEX, and PAR). In two out of the nine experiment cities (NYC and MEL), we notice only one supply peak of ridesharing services per day at 11 AM in NYC and 8 PM in MEL. Except for the services in NYC, we

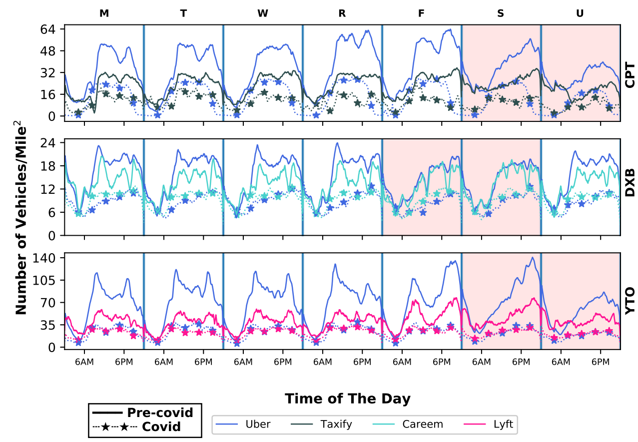

Figure 8. The average number of observed cars of available ridesharing services within a square mile region in CPT, DXB, and YTO (shaded areas represent the weekends) in P1 & P2

find the evening supply count of available services between 6 PM - 9 PM to be 1.31x higher than the morning supply count between 8 AM - 11 AM. Regarding the supply of services during weekends of P2, unlike P1, we do not find the supply to be noticeably lower on weekends than weekdays, except for services in NYC and MEL.

In summary, we see a significant shift in the supply count peaks of service-city instances during weekdays and weekends in the pandemic compared to P1. We also observe that the supply of ridesharing services in P2 is almost a half of P1. Our observations indicate that the supply of ridesharing services is time-sensitive, and the findings of related studies, conducted at one point of time, may not necessarily represent the future usage of these services.

6.1.3 The Average Supply of Ridesharing Services. This section analyzes the expected number of vehicles of each service-city instance that we can observe within an hourly interval. For this, we compute the daily timeseries of supply by counting all the unique vehicles observed during each **60-minute** interval of a day for each day of collected data. We find that the correlation between the timeseries of hourly supply is moderate to strong for each service-city instance for the same day of the week, with $Pearson's\ r > 0.69$ in P1 and $r > 0.73$ in P2. In other words, we observe an almost equal number of vehicles of each service-city instance in the corresponding hourly intervals on the same day of the week.

In both P1 & P2, the supply of Gett service in London is strongly correlated in an hourly interval ($r \geq 0.82$). Gett is also a metered taxi service [48] and may have supply regulations to make the service available in a uniform fashion across the time of the days. This effect was also mentioned in a previous study [25], in which the supply for taxi services maintained a similar pattern across the time of the days and exhibited less variance throughout the day.

The numbers of economic category vehicles that we observe per square mile in all the hourly intervals of P1 & P2 for each service-city instance are presented in Figure 9. We observe that the combined supply of Uber in all the experiment cities outnumbers the combined supply of other services in our experiment with a 1.8 to 1 ratio of vehicles in P1 and 1.5 to 1 ratio of vehicles in P2. In contrast to a previous work [25] that reported the supply of Uber to be 2.3x times greater than the supply of Lyft in NYC in 2018, we find the supply of Uber to be almost 1.12x greater than the supply of Lyft in NYC in P1. The difference between the supply of Uber and Lyft in NYC is almost negligible in P2.

During COVID-19, we observe 58% fewer vehicles of Uber than before the pandemic. Interestingly, in DXB, NYC, and YTO, we do not see Uber as the service with the maximum supply anymore compared to the contemporary services in each city. As compared to P1, overall, we record 54% fewer ridesharing vehicles in P2 with the maximum 67% drop in the supply of Uber in LDN. A potential reason for this observation is that British drivers filed a lawsuit pursuing to be classified as employed workers of Uber in P2, which in turn caused a decrease in the number of drivers for Uber. Further, we observe the maximum number of ridesharing vehicles (in an hourly interval within a square mile region) in NYC and YTO in P1 & P2. We observe the least number of unique ridesharing vehicles in an hourly interval in DXB during P1 and LDN during P2.

In an official Uber report [35], it is mentioned that the majority of the Uber drivers stay active on the Uber app for less than 2 hours a day. In this case, drivers will want to know the peak supply hours of ridesharing services to avoid competition while they are on the road. Although the information of peak supply hours of the services is not public, using the supply data, we examine the hours of the day in which we observe the maximum number of ridesharing vehicles in experiment cities. The top shaded parts of Figure 9 represent the most frequently repeated two *Peak Hours* of supply for every service-city instance during weekdays and weekends of P1 & P2. Green and red markers highlight the supply counts during the peak hours of weekdays and weekends, respectively. We observe the majority (*60%*) of presented peak supply hours of ridesharing services during weekdays between 8 AM - 12 PM in P1. Due to low supply during the mornings of weekdays in P2, we observe only 30% of the supply peak hours during the same interval. Regarding weekends in our experiment, almost 62% of the supply peak hours are observed between 8 PM - 1 AM in P1, and 65% of them are recorded between 3 PM - 7 PM in P2. On average, we find the supply of ridesharing services during the peak hours of weekdays to be 1.13x higher than the supply during weekends peak hours in P1, except for Uber in MEL. Unlike P1, we observe that the supply of ridesharing services is almost equal during the peak hours of weekdays and weekends in P2. In summary, we can see Uber losing its status as the most popular ridesharing service in terms of supply count in DXB, NYC, and YTO in P2.

Finally, we investigate whether we observe a similar number of ridesharing vehicles across time of the day for the same day of the week at a particular location (referred to as *spatial correlation*). The examination of spatial correlation of supply of services across time can help us in predicting the supply hotspots in experiment cities. For each service-city instance, we start by dividing the experiment region into a grid of multiple 200 $feet^2$ blocks. Then, we create the supply timeseries for every block by recording the number of observed unique vehicles in that block during each 60-minute interval for every day of data collection. Last, we compute the Pearson correlation coefficient between the timeseries of every block for

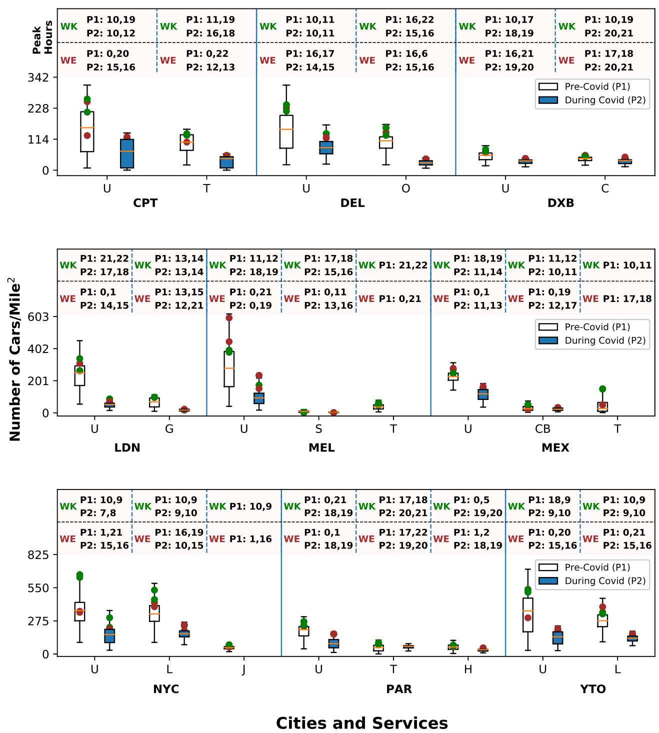

Figure 9. Number of economic category cars per mile^2 for all the service-city instances in hourly intervals. Green and brown markers represent supply in weekdays (WK) and weekends (WE) peak supply hours. Top shaded panels represent the most repeated peak hours in weekdays and weekends in P1 & P2.

the same days of the week in P1 & P2. We find the spatial correlation of service-city instances extremely weak, with the correlation coefficient $r < 0.13$ in P1 and $r < 0.17$ in P2. In other words, we find a strong positive correlation between the timeseries of supply counts for each service-city instance in its experiment region with respect to the time of the day and the day of the week. In contrast, we do not see a similar number of ridesharing vehicles at the same time for the same days of the week at a particular geolocation.

The key takeaway of this result is that it is difficult for drivers of ridesharing services to know the supply in real-time, considering that: *i)* the ridesharing services do not make the real-time supply information public and *ii)* the observation, in our experiment, that supply hotspots of ridesharing services are not predictable.

`RMS` can graphically present the supply data of ridesharing services. It shows the real-time number of unique vehicles of ridesharing services observed in different neighborhoods of cities. It also offers REST APIs to expose the information of real-time supply in JSON format.

6.2 Utilization

This section compares the daily utilization of ridesharing services across time within and across different regions. Figure 10 presents

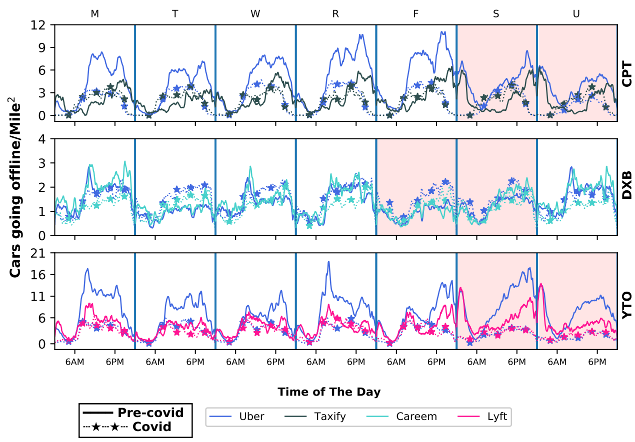

Figure 10. Utilization of ridesharing services throughout the week in three cities where the utilization is the number of economy category vehicles of available services going offline per square mile

the daily trends of utilization counts (the number of cars going offline/mi^2), averaged over five-minute windows, for every day of the week in CPT, DXB, and YTO for P1 & P2.

We make three key observations regarding temporal trends of utilization. *First*, excluding Taxify-CPT, we observe a strong correlation coefficient (*Pearson* $r > 0.704$) between daily supply and utilization timeseries of studied service-city instances in both P1 & P2. This explains why we observe local utilization peaks corresponding to supply peaks of the presented services. For Taxify-CPT in P1, we observe its utilization peak around evenings, with only one local peak in daily utilization observed throughout the week. This suggests that the majority of Taxify passengers may prefer to use the city bus service in the morning as we see a local peak of utilization for Taxify in CPT after 7 PM six days of the week (the city bus service shuts down at 7 PM [13]).

Second, in P1, we observe the utilization of ridesharing services to be 1.6x higher on weekend nights as compared to weekday nights (8 PM - 12 AM). We observe the maximum utilization during the evenings (6 PM - 10 PM) of the first day of the weekend in eight out of the nine experiment cities (except in DXB). In P2, we do not observe a significant increase in the utilization during weekend nights compared to weekday nights, probably because of country-wide lock-down orders since early 2020.

Third, the utilization patterns of ridesharing services in the experiment cities during both weekdays and weekends in P1 are not similar to those in P2. For example, during weekdays in five out of the nine cities (*i.e.*, CPT, DEL, DXB, MEL, and NYC), the utilization of available services is observed to be 11% higher in the morning rush hours as compared to the evening rush hours in P1. While in P2, we find the utilization of available services higher during the morning rush hours than the evening rush on weekdays in just NYC and YTO. As another example, in P1, the average utilization count of weekdays is ~16% higher than the average utilization count of weekends in seven out of nine cities except for MEX and NYC. Whereas in P2, we observe that the average utilization count of ridesharing services is only 3-5% higher during weekdays than weekends in the experiment cities, except for CPT and DEL. We observe a little higher weekend utilization of ridesharing services in CPT and DEL than their weekday utilization.

In summary, we observe 39% fewer inferred trips in our utilization datasets in P2 than P1. Unlike P1, the diurnal patterns in the utilization of ridesharing services are not significantly evident in P2 because of the low supply and utilization of such services during the pandemic.

Next, we study the average hourly utilization percentage (*AHU%*) of available vehicles of the ridesharing services in our experiment. We examine the utilization% ($\frac{utilization}{supply} \times 100$) of all of the service-city instances in every one-hour interval of P1 & P2. As a result of 28 days of data collection in each phase, we have 672 hourly utilization percentage values for available service-city instances in P1 & P2, respectively. Figure 11 shows the values of the AHU% of ridesharing services in the experiment cities in P1 & P2. Overall, we see 1.33x higher AHU% of ridesharing services in P2 (~28.1%±8%) compared to P1 (~21.7%±5%). In both phases of our experiment, we consistently observe the *AHU%* of ridesharing services in DXB and MEX to be the maximum and minimum, respectively, among the experiment cities. Regarding the services that are present in multiple cities, we find the *AHU%* of Uber, across all available cities, to be the highest as 27%. Interestingly, we see Shebah in MEL to exhibit the least *AHU%* of 12% in P1, but the second-highest *AHU%* of 36% in P2. Excluding Taxify in PAR and CPT, we do not find any service-city instance to exhibit lower *AHU%* of available ridesharing vehicles in P2 compared to P1.

In summary, we observe a close competition of regional ridesharing services in the AHU% with Uber, *i.e.*, we see Careem-DXB, Gett-LDN, Shebah-MEL, Cabify-MEX, and Lyft-YTO to have comparable or even better AHU% than Uber in the respective cities.

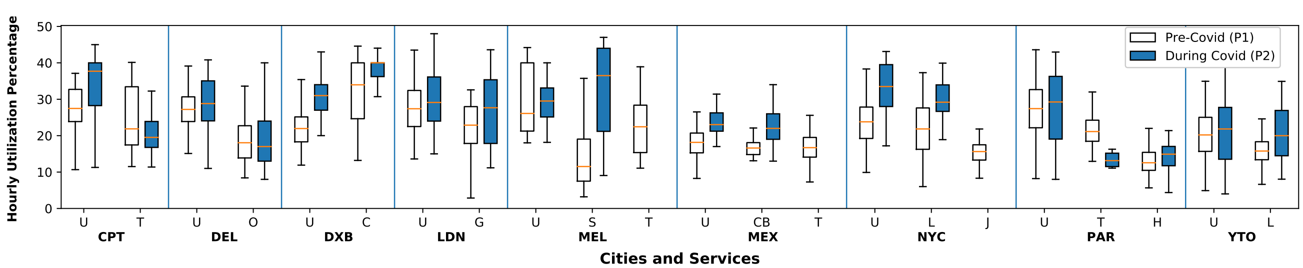

Figure 11. The average hourly average utilization percentage (*AHU%*) of services-city instances in P1 and P2 of our experiment.

6.3 Idle Time and Idle Distance

To find the time spent and distance traveled by the utilized vehicles (aka, *Idle time* and *Idle distance*, respectively) of ridesharing services before going offline, RMS records the geolocation of each utilized vehicle on every epoch from its first appearance until it goes offline.

Figure 12 shows the heatmaps of the idle distances and their corresponding idle times for the utilized ridesharing vehicles for all service-city instances in P1 & P2. The percentages in the red, green, blue, and black windows along Y-axes show the fraction of utilized vehicles observed with idle time between 0 - 5, 5 - 10, 10 - 15, and 15 - 20 minutes, respectively. The same windows along X-axes show the fraction of utilized vehicles observed with the idle distance between 0 - 1, 1 - 2, 2 - 3, and 3 - 4 miles, respectively. Crediting to the high demand of available vehicles in P2, we find the average idle time and distance of ridesharing services of P2 (7 minutes and 0.88 miles) to be lower than that of P1 (9 minutes and 1.77 miles). We find that, on average, the drivers of services in YTO wait for the shortest time (~5 minutes) compared to other service-city instances to get the ride requests. In contrast, the maximum average idle time of the ridesharing vehicles is observed in MEX as 12 minutes in P1 & P2.

Some interesting observations in Figure 12 are: *i)* the percentage of vehicles that traveled the idle distance of ≥2 miles is <25% in P1 and <10% in P2, which explains why we observe asymmetrical heatmaps of idle distances in both phases; *ii)* we observe that the percentages of vehicles that were idle for 15 to 20 minutes in P1 & P2 are similar but vary significantly in other windows of idle time; and *iii)* the distribution of the percentage of utilized vehicles with respect to the amount of idle distance in P1 & P2 is similar to a concave-up parabola. In other words, in our experiment, the percentage of the utilized vehicles decreases exponentially with an increase in the idle distance value.

Presenting the real-time utilization hotspots, average idle time, and idle distance can help the drivers of ridesharing services receive ride requests faster and be aware of the distance they might have to travel to get the rides. To serve this purpose, RMS provides GUI and respective REST APIs to present the real-time high-demand zones, the average idle time, and the average idle distance of available ridesharing services, which has been explained in §5.

6.4 Surge Analysis

Ridesharing services increase trip prices using surge multipliers, in the form of 1.**X**, to match the supply of drivers to the utilization of riders at any given time. The value of a surge multiplier can increase or go back to 1 depending on the values of supply and utilization during the surge. In our experiment, we observe eight services in P1 and seven services in P2 use surge multipliers. Ridesharing services in our experiment employ *dynamic surge pricing algorithms*, in which the surge multiplier value and time of the day of the surge are not pre-determined, except for Shebah in MEL, which applies a constant surge multiplier of 1.1x daily from 6 am to 8 am and 5 pm to 11 pm (both in P1 & P2). Also, we do not see Lyft in NYC and YTO to apply surge during COVID-19. In this section, we study the surge frequency for available service-city instances, attributes of surge instances (lifespans and radii), and evaluate surge multipliers' effect on ridesharing services' availability attributes.

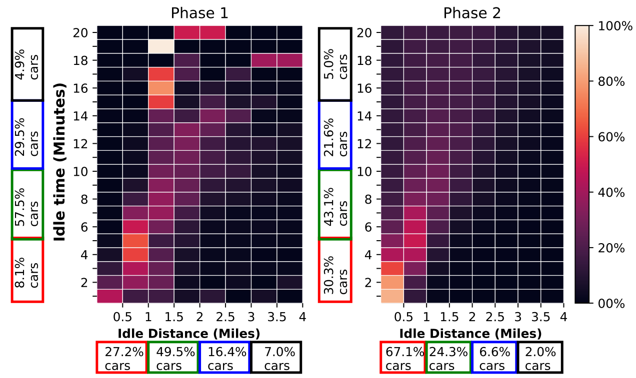

Figure 12. Heatmaps of idle distances and their corresponding idle times of utilized vehicles in P1 & P2. The percentages in red, green, blue, and black windows along Y-axes show the fraction of vehicles observed with idle time between 0 - 5, 5 - 10, 10 - 15, and 15 - 20 minutes, respectively. The percentages along X-axes show the fraction of utilized vehicles observed with the idle distance between 0 - 1, 1 - 2, 2 - 3, and 3 - 4 miles.

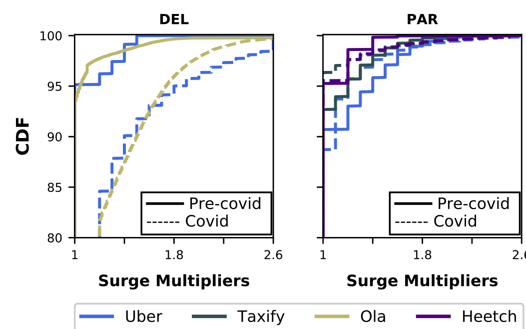

Figure 13. Surge multiplier values CDF of services in DEL and PAR.

6.4.1 Surge Frequency. Considering the space limitation, in Figure 13, we present the CDF of surge multipliers of available services in two cities with the highest surge frequencies in our experiment, *i.e.*, DEL and PAR. The results of surge frequency distribution are based on the responses of 20 million web requests for each service-city instance in each data collection phase.

We make three key observations regarding surge frequency distribution in our experiment. *First*, we see a similar distribution of surge multipliers in most of the ridesharing services, *i.e.*, a staircase pattern in the distribution of surge multipliers with gradually decreasing height of each step in most of the service-city instances (except for Ola in DEL and Cabify in MEX, for which we received the value of surge multipliers in the form of continuous floating-point numbers). This observation signifies that the smallest surge multiplier value greater than 1 is repeated the most in our experiment. *Second*, the frequency of surge multiplier >1 is 5x more in P2 than P1. We notice the value of surge multiplier >1 for 2.5% of total

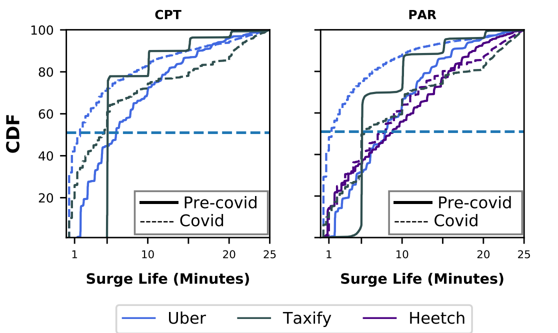

Figure 14. CDF of surge lifespan of services in CPT and PAR

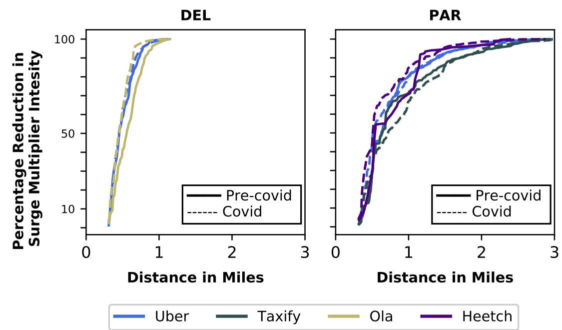

Figure 15. Distance (in miles) to be travelled to reduce the surge multiplier values of services in DEL and PAR

surge web requests in P1 and 12.33% in P2. This observation is not surprising considering that we recorded lower supply and higher utilization of available vehicles in P2 than P1. *Third*, we observe that cities in our experiment have drastically different surge characteristics across the data collection phases. We notice the average surge multiplier values as 1.33x in P1 and 1.41x in P2. Regarding the surge frequency in different cities, we find that the services in PAR and DEL surge most frequently both in P1 (7.3%) and P2 (22%). The city with the minimum surge frequency in our collected dataset is MEL in P1 and YTO in P2. Overall, we observe an increased value of the average surge multiplier and surge frequency in P2 in most of the cities, except for YTO, DXB, and LDN, where the surge frequency of available services is ≤3%.

6.4.2 Surge Lifespan and Surge Radius. To study the lifespan of surges in our dataset, we record the continuous length of time for which the multiplier is >1.

Figure 14 shows the distribution of lifespan of surge instances for services in CPT and PAR, the cities with the maximum expected surge lifespan, in P1 & P2. For this result, we consider the surge instances with a maximum lifespan of 25 minutes. 94% of the total surge instances in our dataset last for less than 25 minutes. We observe that the surge lifespan distribution of most of the service-city instances is similar to a concave down function, *i.e.*, the surge instances with a longer lifespan (>15 minutes) are rare in our experiment, except for Cabify in MEX and Taxify in CPT, MEL, MEX, and PAR in P1. We see a stair-case pattern in the lifespan of the surge instances for the mentioned services, with the minimum lifespan to be five minutes in P1. This shows that one could expect the surge multiplier value to update or end five minutes from the start of the surge for these two services in P1.

We see that the percentage of surge instances with a lifespan of ≤5 minutes is almost doubled in P2 compared to P1. In P1, 1/3 of the surge instances (with at least 45% in CPT, 35% in DEL, 53% in DXB, 18% in LDN, 33% in MEL, 25% in MEX, 35% in NYC, 33% in PAR, and 25% in YTO) end within five minutes for all available ridesharing services in the experiment cities. In P2, 2/3 of the surge instances end within the same time window (with at least 63% in CPT, 65% in DEL, 85% in DXB, 78% in LDN, 87% in MEL, 66% in MEX, 31% in NYC, 64% in PAR, and 78% in YTO). Regarding the service-city instance with the minimum and maximum expected lifespan of the surge, in P1, we find Lyft-NYC to have the minimum expected surge lifespan of three minutes and Cabify-MEX to have the maximum of 11 minutes. In P2, we observe services in DXB to have the minimum expected surge lifespan of two minutes, while Heetch-PAR exhibits a maximum of nine minutes.

We also observe that the lifespans of surges for each service-city instance are strongly spatiotemporally correlated in each data collection phase. In other words, the lifespan of a surge instance at a particular geolocation in a weekday or weekend, within any 15-minute interval of a day, will be close to the average of the lifespans of surge instances that were observed at the same geolocation in the last four weeks within the respective 15-minute interval of weekdays or weekends.

Next, we compute the radius of each surge instance in our dataset. For a surge instance at a source point (S_I), we record the surge multiplier values at every source point within the distance ≤ three miles from S_I to find the nearest non-surging location. We present the surge radii distribution of service in DEL and PAR (cities with service that exhibit the smallest and largest average value of surge radius) for P1 & P2 in Figure 15.

Overall, we do not find any significant difference in the surge radius of service-city instances in P1 & P2. Nor do we find any evidence for radii of surge instances to be pre-determined for any service-city instances in P1 & P2, except for Shebah-MEL. Further, we find that the surge multiplier value reduces by at least 50% if we move 0.5 miles away from the surging location. We also observe that for a half of the surge instances in our experiment, the nearest non-surging location is 0.84 miles from the surging location. Last, less than 1.5 miles of distance is expected to be traveled to avoid 100% of the surge instances for the available services in two out of the nine experiment cities, *i.e.*, DEL and YTO in our experiment.

The three major takeaways of the results related to surge lifespan and radius are: *i*) the ridesharing services update their surge algorithms regularly because we do not see the same staircase pattern in the lifespan of surge instances in P2 as compared to P1 and the previous publication [7]; *ii*) except for a few services, like Heetch-PAR and Cabify-MEX, the majority of surge instances in our experiment are short-lived, and we can avoid those surge instances by waiting for a few minutes, as 50% of the total surge instances ended in ≤ 8 minutes in P1 and ≤ 3 minutes in P2; and *iii*) ridesharing customers can also walk for 8-10 minutes in the *"right"* direction to get the surge multiplier frequency halved.

Based on the observations that the lifespans of surge instances are spatiotemporally correlated and walking in the "right" direction can help in avoiding the surge, RMS presents the real-time surge multiplier value, expected life of surge (based on the history of surge lifespans), and the closest non-surging location for available services. More details have been presented in §5.

6.4.3 Surge Impact on Availability. Ridesharing services claim that surge causes increased supply and reduced utilization. To verify this claim, we examine the supply and utilization of each ridesharing service in two intervals of time for each surge instance. For every data collection source point of every service-city instance with surge multiplier value > 1, we find the four bounding coordinates (upper left, upper right, lower left, and lower right) with surge multiplier values equal to 1. Next, let's say a surge instance starts at T_0 and ends at T_N where T_0 and T_N are minutes of the day from 0 to 1439, and $Surge_{Length}$ = T_N - T_0 + 1 is the total life of surge in minutes, where T_0 - $Surge_{Length} \geq 0$. Then, we calculate the total observed number of vehicles (supply) and the vehicles that went offline (utilization) within the four bounding coordinates of the surged area in two-time windows of $Before_{Surge}$ = T_0 - $Surge_{Length}$ to T_0 and $During_{Surge}$ = T_0 to T_N.

Our surge impact analysis validates the claim that most surge instances of ridesharing services can help increase supply and reduce utilization in the surging areas. We observe a 9% increase in supply volume of ridesharing services in 77% of the total recorded surge instances. However, out of the total vehicles that moved in the surging areas to increase the supply, only 9% in P1 and 6% in P2 are new vehicles with the first epoch recorded after the start of the surge instances. It implies that during the surge instances, the existing supply of ridesharing services gets redistributed geographically to the places where surge multipliers are greater than 1. That may result in long wait times and less supply in the areas where the drivers are moving away.

We also observe that the majority of the surge instances result in decreased utilization in experiment cities, except for DXB, where we see 59% of the surge instances to be ineffective in reducing the utilization. On the other hand, ridesharing services that are the most responsive to the surge are Uber-DEL, Ola-DEL, and Heetch-PAR, with the reduced utilization in ≥90% of the recorded surge instances. Overall, we observe a 5.5% decrease in the utilization counts of services in 71% of the total recorded surge instances.

Last, we observe that surge instances between 6 AM - 9 AM in P1 and 1 PM - 3 PM in P2 are the most effective in increasing the supply of ridesharing services. However, the most effective surge instances, in which we observe the decrease in utilization counts, are between 5 AM - 7 AM in P1 and between 4 AM - 5 AM in P2. The least effective time window for decreased utilization is 8 PM - 10 PM in both phases, with only 44% of surge instances resulting in decreased utilization.

To summarize: *i)* although it appears that surge serves its purpose, it does not help increase the total supply of the ridesharing service; instead, it redistributes the existing supply geographically and *ii)* being at the surging location between 8 PM - 10 PM can be more beneficial for the drivers of ridesharing services. We do not observe a drop in utilization of ridesharing services for the majority of the surge instances that occurred during that period.

7 Discussion

This section summarizes the implications of this work for quantitative and qualitative studies related to ridesharing services.

For Quantitative Studies: Quantitative ridesharing studies measure different aspects of availability and surge pricing and combine them with other demographic and transportation datasets (*e.g.*, median household income, density population, average paying capacity, fuel prices, and road infrastructure) to achieve a variety of different objectives. Examples of such objectives include quantifying the impact of the pandemics on the usage of such services, developing spatiotemporal supply, utilization, and surge prediction models, guiding regulators on making informed policies on reducing carbon emissions, avoiding traffic congestion, and reducing the pricing bias of ridesharing services.

As a use-case of the utility of RMS, we use the data collected and analyzed through it in the pre and during the COVID-19 periods to quantify the impact of COVID-19 on the availability and surge pricing of ridesharing services. Observations from this use-case measurement study reveal that during COVID-19, the supply of ridesharing services decreased by 54%, utilization of available vehicles increased by 6%, and surge frequency of services increased by 5×. Our observations further reveal that during the pandemic, Uber lost its popularity of having the maximum supply in three major cities: New York, Toronto, and Dubai.

Although we do not measure the impact of all attributes specific to each region that may affect the usage of ridesharing services, future work can leverage RMS to collect real-time data of the availability and surge pricing to investigate the correlation of different factors (*e.g.*, weather, national holidays, and public transport) with the usage of ridesharing services.

For Qualitative Studies: HCI researchers can use RMS to evaluate the outcomes of related qualitative studies (*e.g.*, [11, 27]) that suggest new ridesharing application designs for specific users (*e.g.*, the groups with low income, low digital literacy, and physical disabilities) by comparing the availability and surge pricing of ridesharing services in that region before and after introducing the interventions. These collaborations will help the HCI community understand if the new interventions prove to assist in bringing positive changes in the usage of ridesharing services, *e.g.*, increasing the utilization or reducing the idle times and distance.

Some HCI researchers work on diagnosing socioeconomic, gender, racial, and ethical biases in the provisioning or usage of commercial applications using applications' usage data (*e.g.*, [23, 33, 37, 47]). However, the unavailability of the usage data of ridesharing services makes it impossible to further the understanding of the provision and usages biases of these services in different neighborhoods. RMS can help remove the data unavailability barrier in analyzing the availability and surge pricing biases of ridesharing services.

8 Limitations and Future Work

This work has four limitations and open problems arising from making the mobility data of ridesharing services public, which we also see as opportunities for future work.

The *first* limitation is that we choose a small geographical area in each experiment city. Web servers of ridesharing services use the *rate-limiting* mechanism to restrict users from making too many web requests within a short period. Web servers identify each user through a unique OAuth-based authentication token added by the client in request headers. Once web servers activate rate-limiting, users receive **HTTP 429** (too Many Requests response status code) with **Retry-After** header value (in seconds) indicating how long to wait before making a new request. We were able to execute at the maximum six web requests within one second using the same authentication token. In simpler words, we can use the same user account token to get nearby vehicles or surge values from six different geolocations concurrently in our experiment without experiencing rate-limiting from the web servers of ridesharing services. To avoid **HTTP 429** errors while executing several nearby vehicles and surge web requests, RMS needs several unique user authentication tokens for each ridesharing service involved in our experiment. As per the standard registration process on social applications, we needed unique phone numbers to sign-up/register on ridesharing services. We opted to use an online messaging service [40] to receive OTPs (One Time Passwords) on unique phone numbers while signing up for ridesharing services. We used a total of 180 phone numbers for our experiment because of budget constraints. As a result of just 180 unique user authentication tokens of each ridesharing service involved in this experiment, RMS can only cover a small part of each city's downtown with 50 data collection source points for each service-city instance. Although RMS is deployed only in a limited region of nine different cities, with enough user authentication tokens and computational resources, RMS can track the supply, utilization, and surge pricing of ridesharing services in different neighborhoods of any city without modifying any presented algorithms in this paper.

Second, the utilization computed by RMS using the method explained in §4.2 may represent the upper bound of the actual utilization of ridesharing services since no ridesharing service provides the information about the real utilization. Some vehicles may go offline not to pick up passengers but because the drivers are done for the time being and turn off the ridesharing applications, and RMS can consider such vehicles as utilized vehicles.

Third, RMS presents the availability and surge pricing of just economic category vehicles of ridesharing services. Comparative analysis of all the vehicle categories of ridesharing can provide further insight into the usage of ridesharing services.

Fourth, we are currently hosting RMS on a shared web server and using our institute's computational resources and internet connections to collect and analyze the availability and surge pricing data. To keep the upload bandwidth utilization of the shared web server low, we set RMS update frequency to be once every 15 minutes. With more computational resources and allocated internet bandwidth, we can speed up the update frequency of RMS without modifying the presented algorithms.

In the future, we plan to undertake three tasks: *first*, we will study the spatiotemporal trends of availability and surge pricing trends of multiple ridesharing services in different neighborhoods of the same cities (*e.g.*, downtown and suburban areas) to compare the provision and usage of multiple vehicle categories in those regions; *second*, we plan to collaborate with HCI researchers (*e.g.*, [11, 27]) and use RMS to evaluate the impact of qualitative studies that suggest new ridesharing application designs for specific users (*e.g.*, groups with low income, low digital literacy, and physical disabilities) by comparing the availability and surge pricing of ridesharing services in that region before and after introducing the interventions; and *third*, we will also work on improving the user experience of RMS for different types of users by conducting usability workshops with ridesharing users. We will be focusing primarily on understanding the expectations of non-expert users from an algorithmic platform that will guide the design improvements of RMS.

9 Ethical Considerations

As this work involves data collection from the official web servers of ridesharing services, we have been careful to collect data ethically. *First*, we do not retain any personal information about the driver or passengers of ridesharing services, nor did we book any ride for our experiment. *Second*, we keep the impact of this experiment to be negligible on the ridesharing services' infrastructure; RMS scripts act as a few actual instances of the respective applications considering. The sum of downloads of all the discussed ridesharing services on Android Google Play, and Apple App Store is in hundreds of millions. *Next*, we also obtained approval for this study from the Institutional Review Board of our institute. *Last*, we thoroughly read the Terms of Conditions (ToC) documents of studied ridesharing services to ensure that we do not violate any of the terms while collecting the data. We do not find any term that may prevent us from scraping the data from their applications.

10 Conclusions

The resistance of ridesharing services in making the information of their availability (supply, utilization, idle time, and idle distance) and surge pricing public creates a barrier for researchers to investigate the usage of these services across different regions. This paper presents Ridesharing Measurement Suite (RMS), a data feed tool that removes entry barriers for analyzing ridesharing services' availability and surge pricing. RMS continuously crawls and analyzes the web traffic of smartphone applications of ridesharing services. It exposes real-time data on the availability and surges pricing of ridesharing services through a graphical user interface and a set of public APIs.

Using RMS, various stakeholders of ridesharing services can acquire real-time information on the availability and surge pricing of ridesharing services in different regions. It can also simplify the data collection process for future multidisciplinary ridesharing research studies. To highlight the utility of RMS, we use RMS to conduct a large-scale measurement study on the availability and surge pricing by collecting data for 10 popular ridesharing services in nine countries for eight weeks in two time periods: before COVID-19 and during COVID-19. Using the data collected and analyzed by RMS, we make several important observations regarding ridesharing services' availability and surge pricing, which are left unnoticed in the related literature. We believe that this work will further spark motivation in the research community to explore the dynamics of ridesharing services in different regions.

References

[1] Jacob Benesty. 2019. *Pearson Correlation Coefficient.* Wikipedia. Retrieved (August 28, 2021) from https://en.wikipedia.org/wiki/Pearson_correlation_coefficient
[2] Shannon Bond. 2020. *Lyft Lays Off Nearly 1,000 Employees As Virus Saps Demand For Rides.* NPR. Retrieved (August 28, 2021) from https:

//www.npr.org/sections/coronavirus-live-updates/2020/04/29/847904818/lyft-lays-off-nearly-1-000-employees-as-virus-saps-demand-for-rides
[3] Anne Elizabeth Brown. 2018. *Ridehail revolution: Ridehail travel and equity in Los Angeles.* Ph.D. Dissertation. UCLA.
[4] Michael Bustle. 2022. GEARS in North Carolina - Global Training Initiative. https://gti.ncsu.edu/gears/. (Accessed on 02/19/2022).
[5] Juan Camilo Castillo, Dan Knoepfle, and Glen Weyl. 2017. Surge Pricing Solves the Wild Goose Chase. In *Proceedings of the 2017 ACM Conference on Economics and Computation* (Cambridge, Massachusetts, USA) *(EC '17)*. Association for Computing Machinery, New York, NY, USA, 241–242. https://doi.org/10.1145/3033274.3085098
[6] Brian Caulfield. 2009. Estimating the environmental benefits of ride-sharing: A case study of Dublin. *Transportation Research Part D: Transport and Environment* 14, 7 (2009), 527–531.
[7] Le Chen, Alan Mislove, and Christo Wilson. 2015. Peeking Beneath the Hood of Uber. In *Proceedings of the 2015 Internet Measurement Conference* (Tokyo, Japan) *(IMC '15)*. Association for Computing Machinery, New York, NY, USA, 495–508. https://doi.org/10.1145/2815675.2815681
[8] Blerim Cici, Athina Markopoulou, Enrique Frias-Martinez, and Nikolaos Laoutaris. 2014. Assessing the Potential of Ride-Sharing Using Mobile and Social Data: A Tale of Four Cities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 201–211. https://doi.org/10.1145/2632048.2632055
[9] Bryna Corcoran. 2019. *Lyft National Statistics 2019.* Lyft. Retrieved (August 28, 2021) from https://www.lyftimpact.com/stats/national
[10] Judd Cramer and Alan B Krueger. 2016. Disruptive change in the taxi business: The case of Uber. *American Economic Review* 106, 5 (2016), 177–82.
[11] Tawanna R. Dillahunt, Vaishnav Kameswaran, Linfeng Li, and Tanya Rosenblat. 2017. Uncovering the Values and Constraints of Real-Time Ridesharing for Low-Resource Populations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2757–2769. https://doi.org/10.1145/3025453.3025470
[12] Erica Fink. 2014. *Uber threatens drivers: Do not work for Lyft.* CNN. Retrieved August 26, 2021 from https://ride.guru/lounge/p/what-is-the-average-trip-distance-for-an-uber-or-lyft-ride
[13] Tasneem France. 2019. *Capetown Bus Service Schedules.* MyCiti. Retrieved (August 28, 2021) from www.myciti.org.za
[14] Yanbo Ge, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf. 2016. *Racial and gender discrimination in transportation network companies.* Technical Report. National Bureau of Economic Research.
[15] Suiming Guo, Chao Chen, Jingyuan Wang, Yaxiao Liu, Ke Xu, and Dah Ming Chiu. 2018. Dynamic Price Prediction in Ride-on-Demand Service with Multi-Source Urban Data. In *Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (New York, NY, USA) *(MobiQuitous '18)*. Association for Computing Machinery, New York, NY, USA, 412–421. https://doi.org/10.1145/3286978.3286992
[16] Suiming Guo, Chao Chen, Jingyuan Wang, Yaxiao Liu, Ke Xu, and Dah Ming Chiu. 2020. Fine-grained dynamic price prediction in ride-on-demand services: Models and evaluations. *Mobile Networks and Applications* 25, 2 (2020), 505–520.
[17] Snigdha Gutha. 2017. *Analyzing the Effect of Weather on Uber Ridership.* SAS. https://support.sas.com/resources/papers/proceedings17/1260-2017.pdf (Accessed on 07/15/2021).
[18] Steve Hawk. 2018. *What an Economist Is Learning by Driving for Uber.* Stanford. https://www.gsb.stanford.edu/insights/what-economist-learning-driving-uber
[19] Andrew Hawkins. 2018. *Lyft hits a major milestone of 1 billion trips.* The Verge. Retrieved (August 28, 2021) from https://www.theverge.com/2018/9/18/17871514/lyft-milestone-one-billion-trips-rides
[20] Andrew J. Hawkins. 2020. *Uber lays off 3,000 more employees in latest round of COVID-19-inspired cuts.* The Verge. Retrieved (August 28, 2021) from https://www.theverge.com/2020/5/18/21262337/uber-layoff-3000-employees-covid-19-coronavirus
[21] Alejandro Henao. 2017. Impacts of Ridesourcing-Lyft and Uber-on Transportation Including VMT. *Mode Replacement, Parking, and Travel Behavior* 7, 3 (2017), 201–211.
[22] Alex Hern. 2017. *Uber denies misleading passengers with 'phantom cars' in its app.* Business Insider. Retrieved (August 28, 2021) from https://www.businessinsider.com/uber-denies-misleading-passengers-with-phantom-cars-in-its-app-2015-7
[23] Charles G. Hill, Maren Haag, Alannah Oleson, Chris Mendez, Nicola Marsden, Anita Sarma, and Margaret Burnett. 2017. Gender-Inclusiveness Personas vs. Stereotyping: Can We Have It Both Ways?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6658–6671. https://doi.org/10.1145/3025453.3025609
[24] Hyeonjun Hwang. 2016. Measuring the Benefits of Ridesharing Services to Urban Travelers. *growth* 29, 2 (2016), 2.
[25] Shan Jiang, Le Chen, Alan Mislove, and Christo Wilson. 2018. On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 863–872. https://doi.org/10.1145/3178876.3186134
[26] Anto Joseph. 2018. *frida-android-hooks.* Frida Hooks. Retrieved (August 28, 2021) from https://github.com/antojoseph/frida-android-hooks
[27] Joseph Kasera, Jacki O'Neill, and Nicola J Bidwell. 2016. Sociality, Tempo & Flow: Learning from Namibian Ridesharing. In *Proceedings of the First African Conference on Human Computer Interaction.* ACM New York, USA, New York, USA, 36–47.
[28] Hassan Ali Khan. 2021. *Graph plotting scripts used in this paper.* NC State University. Retrieved (August 28, 2021) from https://github.com/rmsstudydata/graphScripts
[29] Hassan Ali Khan. 2021. *Program to consume the APIs of RMS.* NC State University. Retrieved (August 28, 2021) from https://github.com/rmsstudydata/graphScripts/tree/main/RMS-Tool%20Data%20Collection%20Scripts/ML%20Scripts
[30] Hassan Ali Khan. 2021. *Ridesharing Measurement Tool - **APIs**.* NC State University. Retrieved (August 28, 2021) from https://rmstool.com/api
[31] Hassan Ali Khan. 2021. *Ridesharing Measurement Tool - **GUI**.* NC State University. Retrieved (August 28, 2021) from https://rmstool.com
[32] Brenton J Malin and Curry Chandler. 2016. Free to work anxiously: Splintering precarity among drivers for Uber and Lyft. *Communication, Culture & Critique* 10, 2 (2016), 382–400.
[33] Lena Mamykina, Elizabeth D Mynatt, and David R Kaufman. 2006. Investigating health management practices of individuals with diabetes. In *Proceedings of the SIGCHI conference on Human Factors in computing systems.* ACM New York, NY, USA, =New York, NY, USA, 927–936.
[34] A MARGHALL. 2018. Dying to know Uber's secrets, data-hungry cities get creative. *Wired* 3, 09 (2018), 2018.
[35] Lawrence Mishel. 2019. *Uber and the labor market.* Uber. Retrieved (August 28, 2021) from https://www.epi.org/publication/uber-and-the-labor-market-uber-drivers-compensation-wages-and-the-scale-of-uber-and-the-gig-economy
[36] Big Nox. 2021. *NoxPlayer - Free Android Emulator on PC and Mac.* Nox. (Accessed on 09/09/2021).
[37] Anna Offenwanger, Alan John Milligan, Minsuk Chang, Julia Bullard, and Dongwook Yoon. 2021. Diagnosing bias in the gender representation of HCI research participants: how it happens and where we are. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* ACM New York, NY, USA, New York, NY, USA, 1–18.
[38] Ben Popper. 2014. *Introduction to Bayesian Statistics.* The Verge. Retrieved (August 28, 2021) from https://www.theverge.com/2014/2/26/5445210/in-san-diego-uber-kept-drivers-off-the-road-to-encourage-surge
[39] Lisa Rayle, Susan A Shaheen, Nelson Chan, Danielle Dai, and Robert Cervero. 2014. *App-based, on-demand ride services: comparing taxi and ridesourcing trips and user characteristics in San Francisco.* Technical Report. Citeseer.
[40] Charles S. 2021. *Bulk Mobile number use for Phone verification by receive SMS online, help verify bulk accounts.* Get SMS Code. (Accessed on 12/16/2021).
[41] Sara Silverstein. 2014. *A recent study shows that ride sharing is cheaper than taxicab in 21 largest cities in the United States.* Business Insider. (Accessed on 12/16/2021).
[42] Benjamin Snyder. 2017. *Uber denies there are 'phantom' cars in its app.* Fortune. Retrieved (August 28, 2021) from https://fortune.com/2015/07/30/uber-phantom-cars/
[43] Sudeep Srivastav. 2020. *Uber Statistics (2019-2020 Updated).* Appinventiv. Retrieved (August 28, 2021) from https://appinventiv.com/blog/uber-statistics/#A27
[44] Ippei Takahashi. 2018. *What is the average trip distance for an Uber or Lyft Ride?* Ride Guru. Retrieved (August 28, 2021) from https://ride.guru/lounge/p/what-is-the-average-trip-distance-for-an-uber-or-lyft-ride
[45] Ye Tian Tian, Jian Sun Sun, He Dong Dong, and Guoyang Qin Qin. 2020. Quantifying the Impact of Rainfall on Taxi Hailing and Operation. *Journal of Advanced Transportation* 14, 7 (2020), 201–211.
[46] Karl von Randow. 2019. *Charles Proxy.* Charles. Retrieved (August 28, 2021) from https://www.charlesproxy.com
[47] Mihaela Vorvoreanu, Lingyi Zhang, Yun-Han Huang, Claudia Hilderbrand, Zoe Steine-Hanson, and Margaret Burnett. 2019. From Gender Biases to Gender-Inclusive Design: An Empirical Investigation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300283
[48] Itay Wein. 2020. *How much does Gett Taxi cost?* Gett. https://gett.com/uk/faq/user-faqs/getting-started/how-much-does-gett-taxi-cost
[49] Tim Wheatcroft. 2019. *Certify Ride Hailing Report: Uber and Lyft Achieve Record Popularity.* Boston News. Retrieved (August 28, 2021) from https://www.boston.com/news/business/2016/06/28/uber-data-boston-wants
[50] Hamid Zaeri. 2016. *SSLUnpinning Xposed.* Xposed Framework. Retrieved (August 28, 2021) from https://github.com/ac-pm/SSLUnpinning_Xposed

A Terminology

Following are the definitions of terms used in this paper.

Term	**Definition**
Gig Economy	A labour market characterized by the prevalence of short-term contracts or freelance work as opposed to permanent jobs.
Supply	The number of vehicles driving for a ridesharing service over any given period.
Utilization	The number of vehicles currently in supply that passengers book over the given period.
Hourly Utilization Percentage (HU%)	Utilization percentage in an hour.
Idle Time and Idle Distance	The time spent and distance traveled by the ridesharing service drivers while looking for passengers.
Availability	The term availability in this paper encompasses three aspects of ridesharing services, which include: *i)* supply; *ii)* utilization; and *iii)* idle time and idle distance.
Surge	Ridesharing services dynamically increase the prices of their trips during times of low supply or excessive utilization. This concept of a dynamic increase in trip fares because of supply and utilization imbalance is referred to as surge (dynamic) pricing.
Surge Multiplier	Surge multiplier value, *i.e.*, the floating-point value of the surge intensity variable (in the form of 1.x), multiplied with the base trip fare to increase the trip price.
Surge Instance	The surge instance represents the continuous-time window in which the surge multiplier value increases from 1.0 and drops back to 1.0.
Surge Frequency	Surge frequency represents the daily number of surge instances.
Surge Lifespan	Surge lifespan represents the length of surge instances in minutes.
Nearby Vehicles	Vehicles shown on a map in a ridesharing application close to users are referred to as nearby vehicles in this paper.
Source Points	Fifty different geolocations in a city from where `RMS` collects the data of the availability and surge pricing of ridesharing applications. Each data collection geolocation is referred to as a source point in this paper.
Blanketed Region	We refer to the data collection region in each city, where source points are installed, as the blanketed region of that city.
Observable Region	The observable region of any ridesharing service within a given time interval represents a polygon with its area as the product of distances between the horizontally and vertically farthest observed vehicles within that period.
Service-City Instances	A specific ridesharing service operating in a particular city, *e.g.*, Uber-Melbourne.
`RMS`	Ridesharing Measurement Suite: The tool presented in this paper which continuously collects, analyzes, and publishes the information of availability and surge pricing of ridesharing applications.
Working Service-City Instances Set of `RMS`	A set that represents all the service-city instances analyzed by `RMS`.