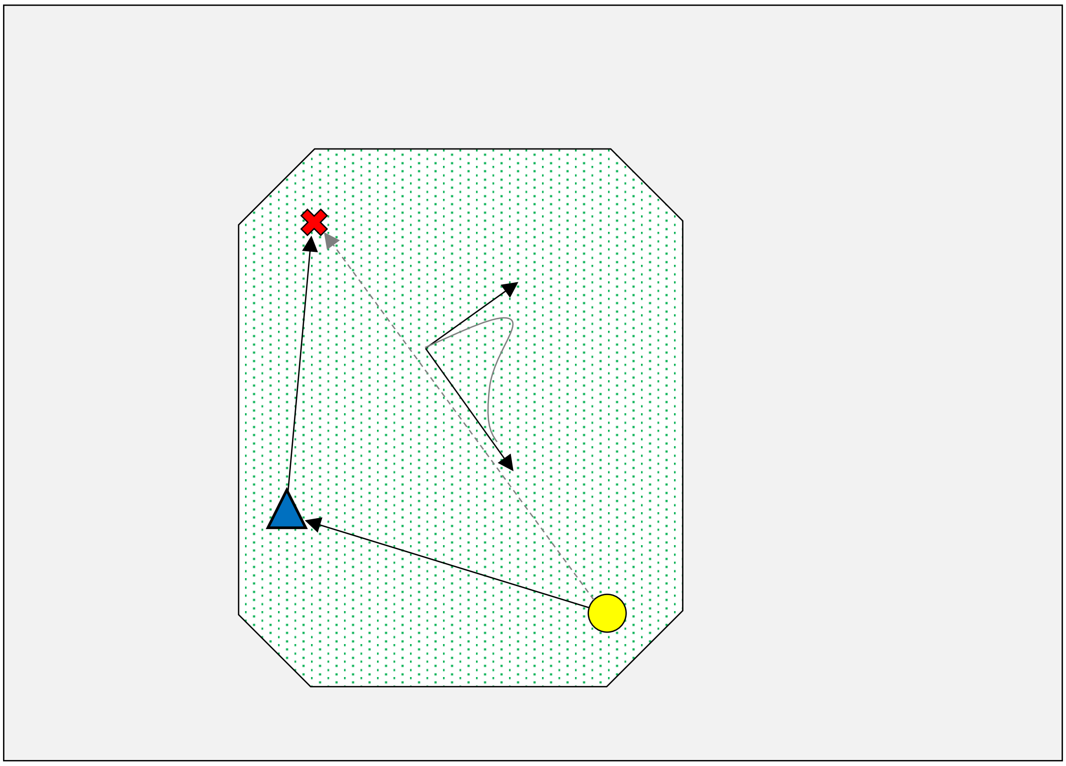


Estimating distances, passenger-vehicle matching and positioning for ridesourcing systems

Caio Vitor Beojone, Urban Transport Systems Laboratory (LUTS)
Nikolas Geroliminis, Urban Transport Systems Laboratory (LUTS)

Conference Paper STRC 2019

STRC | 19th Swiss Transport Research Conference
Monte Verità / Ascona, May 15-17, 2019

Estimating distances, passenger-vehicle matching and positioning for ridesourcing systems

Caio Vitor Beojone, Nikolas Geroliminis

Urban Transport Systems Laboratory (LUTS)
Station 18 CH-1015 Lausanne

T: +41 21 693 24 81
E: {caio.beojone@epfl.ch, nikolas.geroliminis@epfl.ch}

May 2019

Abstract

The motivation of this paper is to develop knowledge and tools to improve the probabilities to match rides and therefore increase the number of shared trips in a ridesourcing system. This paper aims to provide estimations for travel distances on a simplified rectangular urban traffic network to support the modeling of ridesourcing decision problems. Space is a continuous approximation of an urban street network (does not consider streets nor blocks). Thus, space measurement considers Manhattan distances. Estimations are mainly analytical and supported by simple Monte Carlo simulations. Idle vehicles, passengers and their destinations were considered uniformly distributed along the region. Estimated distances refer to travels such as closest vehicle from passenger, single-passenger trip distance, traveling between different regions. Other results evaluate the probabilities of matching passengers and partially busy vehicles (one passenger inside). Simulations considered different moving strategies for maximizing of minimizing the matching probabilities. Moving vehicles tend to concentrate in the center of the region, unless we explicitly order them to avoid the region centroid during their trips.

Keywords

Continuous approximation, Manhattan distances, statistics, urban transportation, ride-sharing, Monte Carlo simulations

1. Introduction

Ridesourcing (also known as Transportation Network Companies, ride-hailing, and e-hailing, for instance) has driven a lot of attention in recent years with the expansion of companies like Uber, Lift, and many others around the world (Jin et al., 2018). Due to its business model, this kind of services raised many concerns regarding labor laws, competitiveness, and safety of drivers and passengers (Rayle et al., 2016).

Ridesourcing sounds a promising direction to improve mobility, fighting car ownership. Moreover, most TNCs allow multiple people to share a ride, like UberPOOL. This shared mobility advance could ensure higher efficiency and sustainability (Mora et al., 2017). However, matching rides is still a challenge. Santi et al. (2014) and Mora et al. (2017) try to tackle this problem using shareability networks.

As part of the effort to model ridesourcing services, Stiglic et al. (2016) showed how a continuous space could help on modeling ride-sharing. Ansari et al. (2018) showed that, since its debut at Newell (1971), continuous approximations developed and matured as a powerful instrument for logistics and transportation problems. Therefore, relaxing discrete variables into continuous ones, like space, is a potential tool to create intuition on urban mobility and ridesourcing problems. Developing intuition about this can lead to relevant improvements (and/or mitigation of negative effects) in the expansion of ridesourcing services.

The motivation of this paper is to develop knowledge and tools to improve the probabilities to match rides and therefore increase the number of shared trips in a ridesourcing system. For this purpose, the urban grid space considered rather small block sizes compared to the area to simplify the calculations and estimations.

Therefore, this paper aims to provide estimations for travel distances on a simplified rectangular urban traffic network to support the modeling of ridesourcing decision problems. Besides traveling distances, this paper focus on the trajectory, and on the matching problem between passengers and vehicles. Estimations are mainly analytical and supported by simple Monte Carlo simulations.

This paper is structured as follows. Section 2 presents a brief literature review on ridesourcing, its problems, and improvements on the quality of service. Section 3 has the estimations on travel distances, matching probabilities, and positioning for ridesourcing vehicles. Section 4 presents some final considerations and directions for further research.

2. Literature Review

Urban efficiency and sustainability are affected by the expansion of ridesourcing services around the world. Jin et al. (2018) presented a recent review of ridesourcing operations. Although it is unclear the impacts on congestion in city centers and reductions on energy consumption and emissions, ridesourcing has shown to be capable of improving economic efficiency (Jin et al., 2018). Hall et al. (2018) evaluated the relationship between ridesourcing and public transportation competition in metropolitan areas in the U.S. Contreras and Paz (2018) used a linear regression analysis to measure the effects of ridesourcing over the taxicab industry in Las Vegas, Nevada. Schwieterman and Smith (2018) used multiple regression analysis and found reductions on trip times between neighborhoods and customers' savings for sharing their trips. Vinayak et al. (2018) used a generalized heterogeneous data model to find social dependency effects of shared mobility service usage. Zha et al. (2016) investigated ridesourcing through an economic point-of-view, considering scenarios with different competitiveness and regulations. Nie (2017) analyzed GPS data from Shenzhen, China, and found that the loss of the taxicab industry to ridesourcing tends to stabilize, also, ridesourcing mildly worsened congestion.

Some recent works tried to improve matching strategies, relocation of empty vehicles, and route choices aiming to improve the quality of ridesourcing services. Santi et al. (2014) used shareability networks to match passengers. Mora et al. (2017) extended its use to match these passengers to vehicles and relocate empty vehicles. Stiglic et al. (2016) used a continuous space simulation to show improvements on carpooling with different driver and rider flexibilities. On dynamic vehicle routing literature, Berbeglia et al. (2010) reviewed dynamic pick-up and delivery problems, as dial-a-ride problems. Molenbruch et al. (2017) reviewed dial-a-ride problems, their solution methods and classified them in categories. Masmoudi et al. (2018) presented a dial-a-ride problem with battery swapping.

Note that this review did not embrace most of the ride-sharing and dial-a-ride literature. Those interested in such problems are encouraged to read Furuhata et al. (2013), the previously mentioned papers, and the references therein.

3. Problems

This section presents a set of problems and analytical or simulation-based results for them. Problems regard to the way a vehicle would move inside a continuous space and its statistical implications.

Suppose that a rectangle with defined by lengths a and b represents a simplified urban street network. Now suppose that a vehicle on a city must travel from point A to point B, as illustrated in Figure 1. Also, suppose that the path between points A and B is composed of a grid (like streets in a city) of perpendicular and parallel set of lines. Manhattan distances are required to model this situation. Streets are uniformly spaced and the distance between a street n and its neighbor $n-1$ is $\Delta x = |x_n - x_{n-1}|$ (for both directions – x and y). Finally, neighbor street segments are close to each other in comparison to the size of the entire city, thus $\Delta x \rightarrow 0 \therefore \Delta x \equiv dx$.

Figure 1: Illustration of the continuous urban area represented in a rectangle.

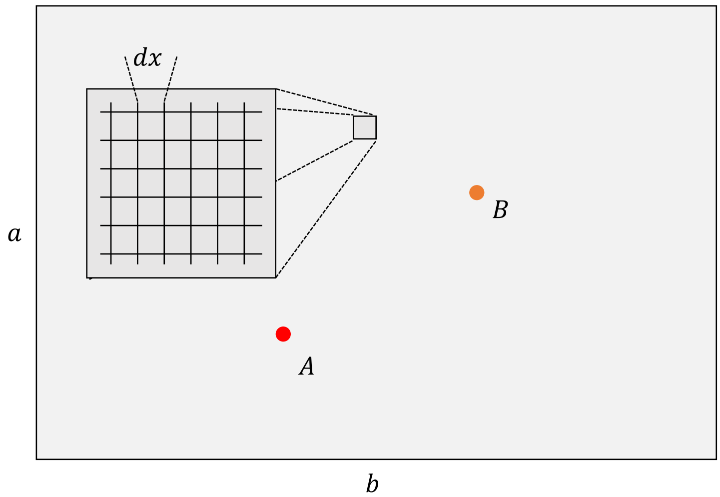
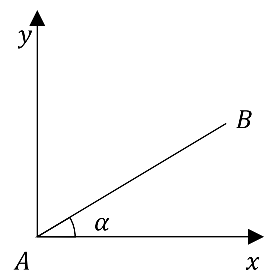


Figure 2 shows in a simple Cartesian plane the space between points A and B. If we consider that the vehicle travels as close as possible to the straight line that connects points A and B and remember that $\Delta x \to 0$, it is reasonable to assume that the angle α is going to remain constant during the trip. Therefore, the trip is the straight line connecting both points on average.

Figure 2: Cartesian representation of the angle between points A and B.



Considering that points A and B are unitary distant (in Manhattan distance), one can assume that the speed is a vector $\vec{v}$ and its decomposition follows the rule presented as the system of equations shown in Equation (1).

$$\begin{cases} |\vec{v}_x| + |\vec{v}_y| = 1 \\ \dfrac{|\vec{v}_y|}{|\vec{v}_x|} = \tan(\alpha) \end{cases} \tag{1}$$

Suppose that $0 \le \alpha \le \frac{\pi}{2}$. This allows us to find a single solution for the problem (Equation 2).

$$v_x = \frac{1}{\tan(\alpha) + 1}, \qquad v_y = \frac{\tan(\alpha)}{\tan(\alpha) + 1} \tag{2}$$

If we replace $\tan(\alpha)$ by D_y/D_x, where D_l is the total distance on axis l, we have Equation (3).

$$v_x = \frac{1}{\frac{D_y}{D_x} + 1} = \frac{1}{\frac{D_y + D_x}{D_x}} = \frac{D_x}{D_y + D_x} = \frac{D_x}{D}, \qquad v_y = \frac{D_y}{D_y + D_x} = \frac{D_y}{D} \tag{3}$$

For the general case ($0 \le \alpha < 2\pi$), $D_x = x_B - x_A$, $D_y = y_B - y_A$, and $D = \|B - A\|_1$. In the next subsections, we refer to this result as "straight as possible".

The latter shows that the estimations provided in this paper are only susceptible to errors when $\min(D_x, D_y)$ is smaller than the block size, i.e. when it might be necessary to go around the block to reach the destination.

3.1 Distances

In this section, we examine analytically at the distances that a vehicle might travel in a continuous rectangular space with random origins and destinations.

3.1.1 Distance between two uniformly random points in one dimension

If we assume that the origin and destination of a passenger are random and uniformly distributed across the region, one can define their positions in one dimension (said dimension represented by axis y, for instance) as $P_{\text{orig.}} \sim U(0, a)$ and $P_{\text{dest.}} \sim U(0, a)$. Therefore, each point is defined by the probability density function (pdf) on Equation (4).

$$f_{P_i}(p_i) = \begin{cases} \frac{1}{a}, & \text{if } 0 \leq p_{i1} \leq a \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Hence, the distance in the chosen dimension is a random variable denoted by Z, defined as $Z = |P_{\text{dest.}} - P_{\text{orig.}}|$. Therefore, using a convolution, the pdf is given by Equation (5).

$$f_Z(z) = \begin{cases} \frac{2}{a} - \frac{2z}{a^2}, & \text{if } 0 \leq z \leq a \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

3.1.2 Total distance between two uniformly random points in two dimensions

Based on previous results, the two-dimensional distance is a random variable denoted by W, defined as $W = Z_y + Z_x$ (the sum of the random distances in both directions). This result is possible because travels occur according to an urban street grid (Manhattan distances). Hence, a convolution can find the pdf for such variable (Equation 6).

$$f_W(w) = \begin{cases} -\frac{\frac{2w(-w^2 + 3bw)}{3} - \frac{2aw(6b - 3w)}{3}}{a^2b^2}, & \text{if } 0 \leq w < a \\ \frac{2(a + 3b - 3w)}{3b^2}, & \text{if } a \leq w \leq b \\ \frac{2(a + b - w)^3}{3a^2b^2}, & \text{if } b < w \leq a + b \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

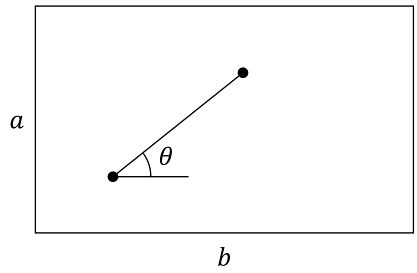
Gaboune et al. (1993) found similar results and their study included Euclidean and Chebychev distances. Equation (7) gives the average total distance.

$$E[W] = \frac{a+b}{3} \tag{7}$$

3.1.3 Angle between two uniformly random points

It is relevant to define an angle when traveling between two random points in a continuous space. It is particularly important when considering that the traveler will follow a straight line (as possible) to his destination. Figure 3 illustrates this situation.

Figure 3: Angle θ between two random points in a rectangular area.



If the random points are uniformly distributed across the region, an angle Θ is a random variable defined only in the first quadrant as seen in Equation (8).

$$\Theta = \arctan\left(\frac{|y_{\text{dest.}} - y_{\text{orig.}}|}{|x_{\text{dest.}} - x_{\text{orig.}}|}\right) \therefore 0 \leq \Theta \leq \frac{\pi}{2} \tag{8}$$

Considers $K = \tan\Theta$ as an auxiliary random variable. Equation (9) gives an alternative formulation for K.

$$K = \frac{Z_y}{Z_x} \therefore Z_y = Z_x K, \qquad k \geq 0 \tag{9}$$

The pdf of K is found through the quotient between two independent random variables. Thus, Equation (10) calculates the pdf of K.

$$f_K(k) = \begin{cases} \dfrac{b(2a-bk)}{3a^2} & \text{if } k \leq \dfrac{a}{b} \text{ and } k \geq 0 \\ -\dfrac{a(a-2bk)}{3b^2k^3}, & \text{if } k > \dfrac{a}{b} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

As we are interested in the random angle Θ, and not its tangent, Equation (11) shows the variable change process.

$$f_\Theta(\theta) = f_K(\tan(\theta)) \left|\frac{d\tan(\theta)}{d\theta}\right| \tag{11}$$

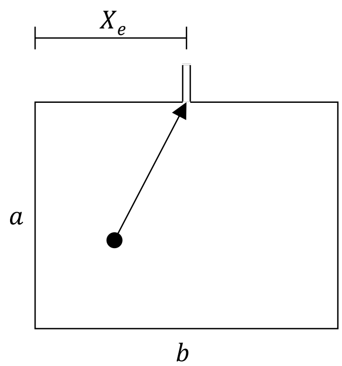
3.1.4 Distance to a neighbor region

Since a traveler may not restrict his trip to a single area, it is worth modeling the process in which one leaves a region to another. We define this movement as going from a random point in the direction to a specific point in the frontier of the current region.

One exit (general case)

The first case consists of only one exit to the neighbor region. Figure 4 illustrates the case for an exit at the top of the region (without loss of generality).

Figure 4: Schematic representation of the one exit problem.



Define distance D, a random variable defined as $D = |X_e - X_{\text{orig.}}| + |Y_e - Y_{\text{orig.}}|$, where X_e and Y_e are coordinates to the exit to the neighbor region. These coordinates are not random variables. $Y_e = a$, then we have Equation (12).

$$D = \underbrace{|X_e - X_{\text{orig.}}|}_{d_X} + \underbrace{|a - Y_{\text{orig.}}|}_{d_Y} \tag{12}$$

Without loss of generality, consider $0 \leq X_e \leq \frac{b}{2}$ (if $X_e > \frac{b}{2}$, it is enough to just mirror the solution). A convolution of both directions obtains the pdf of D. Note that the solution is for two different cases of X_e compared to a (Equations 13-15).

$$d_Y(y) = \begin{cases} \frac{1}{a}, & \text{if } 0 \leq y \leq a \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

$$d_X(x) = \begin{cases} \frac{2}{b}, & \text{if } 0 \leq x \leq X_e \\ \frac{1}{b}, & \text{if } X_e < x \leq b - X_e \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

$$f_D(d) = \begin{cases} \begin{cases} \frac{2d}{ab} & \text{if } 0 \leq d < X_e \\ \frac{X_e + d}{ab} & \text{if } X_e \leq d \leq a \\ \frac{X_e + 2a - d}{ab} & \text{if } a < d \leq a + X_e \\ \frac{1}{b}, & \text{if } a + X_e < d \leq b - X_e \\ \frac{a + b - X_e - d}{ab}, & \text{if } b - X_e < d \leq a + b - X_e \\ 0, & \text{otherwise} \end{cases}, & \text{if } a \geq X_e \\ \begin{cases} \frac{2d}{ab}, & \text{if } 0 \leq d < a \\ \frac{2}{b}, & \text{if } a \leq d \leq X_e \\ \frac{X_e + 2a - d}{ab}, & \text{if } X_e < d \leq a + X_e \\ \frac{1}{b}, & \text{if } a + X_e < d \leq b - X_e \\ \frac{a + b - X_e - d}{ab}, & \text{if } b - X_e < d \leq a + b - X_e \\ 0, & \text{otherwise} \end{cases} & \text{if } a \leq X_e \leq \frac{b}{2} \end{cases} \tag{15}$$

Multiple exits (general case)

The next case is a general form for multiple exits when all exits are in the same edge of the rectangular shaped region. Without loss of generality, a two-exit example illustrates the problem.

If we consider that a traveler will always choose the shortest path to the neighbor region, then the distance $D = \min(D_1, D_2)$, where D_1 and D_2 are the distances to each of the two options to reach the neighbor region, respectively. If both D_1 and D_2 are expressed as the sum of the distances in both directions (Manhattan distances), one would have Equation (16).

$$D_i = |X_{ei} - X| + |a - Y| = \Delta X_i + \Delta Y \tag{16}$$

By definition, $\Delta Y \sim U(0, a)$. In the other hand, Equation (17) has the pdfs for ΔX.

$$f_{\Delta X_i}(x) = \begin{cases} \frac{2}{b}, & \text{if } 0 \le x \le X_{ei} \\ \frac{1}{b}, & \text{if } X_{ei} < x \le b - X_{ei} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

The problem D is redefined as Equation (18).

$$D = \min(D_1, D_2) = \min(\Delta X_1, \Delta X_2) + \Delta Y \tag{18}$$

Finally, a cumulative density function of ΔX, in the form of $F_{\Delta X_i}(x)$, calculates the minimal distance in the dimension x (Equation 19).

$$X = \min(\Delta X_1, \Delta X_2) \therefore F_X(x) = F_{\Delta X_1}(x) + F_{\Delta X_2}(x) - F_{\Delta X_1}(x) F_{\Delta X_2}(x) \tag{19}$$

Equation (20) gives the expected traveled distance.

$$E[D] = E[X] + E[\Delta Y] = E[X] + \frac{a}{2} \tag{20}$$

Infinite exits (particular case of multiple exits)

Here, we present a particular case of the multiple exits case. Assume that the number of exits to the neighbor region is very high relative to the region width. It is the same as breaking the problem in infinite minor problems until $\Pr(D_X = 0) \to 1$. Thus, Equation (21) holds.

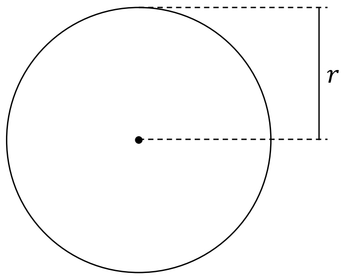
$$D = \Delta Y \therefore E[D] = E[\Delta Y] = \frac{a}{2} \tag{21}$$

3.1.5 Minimal distance to a passenger

Exceptionally in this section, we obtain results considering Euclidean distances first. Consider that a passenger arrives and there are n available vehicles within an acceptance radius r (see

Figure 5). The goal here is to find the distance between the passenger and the closest available vehicle, on average.

Figure 5: Simple problem representation with a passenger as the center of the circle.



One Available vehicle

Now consider that only one vehicle is available inside the circle and its position is uniformly random. It is critical to remark that the density of vehicles inside the circle is constant to find a probability density function for the distance to the center of the circle. For instance, consider $r' < \frac{r}{2}$ as an inner radius r' of a ring with width $\Delta r \to 0$. One should expect to find k vehicles within this ring; furthermore, if we double the inner radius to $2r'$ and keep the same width $\Delta r \to 0$, one should expect to find $2k$ vehicles within this second ring as both have the same density. Thus, Equation (22) describes the probability density function for the distance of a single vehicle to the center of the circle.

$$f_X(x) = \begin{cases} \frac{2x}{r}, & \text{if } 0 \leq x \leq r \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

Without loss of generality, we consider $r = 1$ for now on. Therefore, Equation (23) gives the average distance of a vehicle to the center of the circle. It is a well-known result, which one may find across many papers in the literature, like Geroliminis et al. (2011).

$$\mathrm{E}[X] = \int_0^r x f(x) dx = \int_0^1 2x^2 dx = \frac{2x^3}{3}\bigg|_0^1 = \frac{2}{3} \tag{23}$$

Multiple available vehicles

If more than one available vehicle is inside the circle, a random variable $Z = \min(X_1, X_2, \dots, X_n)$ can read the minimal distance to the center. As $X_1, X_2, \dots, X_n$ are independent and equally distributed, it is possible to find the distribution of Z, using Equation (24) (Larsen and Marx, 2012).

$$F_Z(z) = 1 - \left(1 - F_X(z)\right)^n \therefore F_Z(z) = 1 - (1 - z^2)^n \tag{24}$$

Thus, Equation (25) gives the pdf for Z.

$$f_Z(z) = F_Z'(z) = \begin{cases} 2nz(1 - z^2)^{n-1}, & \text{if } 0 \leq z \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

The expected distance to the center is seen in Equation (26).

$$\mathrm{E}[Z] = \int_{-\infty}^{\infty} z f_Z(z) dz = n\beta\left(\frac{3}{2}, n\right) = \prod_{i=1}^{n} \frac{2i}{2i + 1} \tag{26}$$

The proof for the previous equation is tedious. Those willing to show it should have a few things in mind: 1) $n \in \mathbb{N}_+$; 2) Write beta function using gamma functions; 3) the relationship between gamma function and factorials; and 4) Some particular values for Gamma functions.

Relationship with Manhattan distances

So far, we considered Euclidean distances for this problem. However, the same results apply to Manhattan distances as well. The reason for this is that the same property that defined the probability density function using rings within the circle applies to the use of Manhattan distances with a couple of adjustments. First, the region does not have a circular shape, but a diamond shape. Second, instead of radius, consider the Manhattan distance to the frontier of the diamond (which turns out to be constant). By doing so, the same property about double the inner radius apply for the diamond, and, therefore, all last results apply.

3.2 Passenger-vehicle matching

In this section, we look at the problem of a ride-sourcing system to match passengers, idle vehicles, and pooling vehicles (with already one passenger inside).

3.2.1 Matching passengers and idle vehicles within an acceptance range

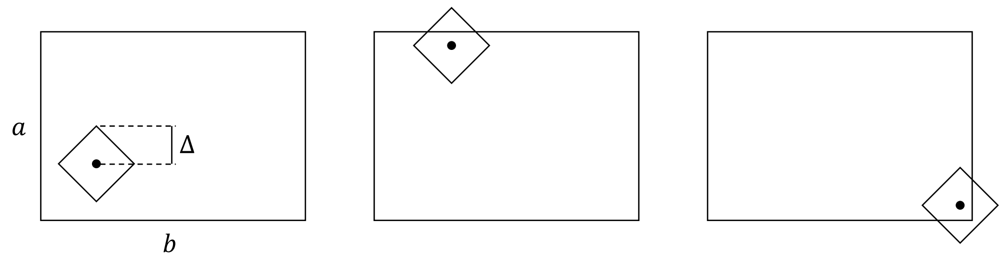
Consider that there is a region and an idle vehicle inside this region. New passengers arrive and idle vehicles must pick-up these passengers. However, passengers will only accept the ride if vehicles are capable of serving them within a certain amount of time, the system loses passengers, otherwise. The service provider needs to determine the fleet size so that it is capable of serving a certain percentage of the demand. One can define this problem using the following Equation (27).

$$\mathrm{Pr(Acceptance)} = \frac{\mathrm{E[Covered\ area]}}{\mathrm{Total\ area}} \tag{27}$$

Single vehicle in a rectangular area

Now consider a rectangular shaped region defined by lengths a (height) and b (width). Inside this region, there is an idle vehicle available to pick-up new passengers. Assume that new passengers arrive uniformly across the region and the traveling speed is approximately constant. Additionally, vehicles travel through a uniform network of streets, a taxicab geometry applies, therefore. Therefore, one can describe the shape of the acceptance region using a diamond shape with the parameter Δ. One can physically interpret the Δ as the maximum distance that a vehicle can travel to pick-up an arriving passenger within a waiting time limit. Figure 6 shows that, as the vehicle gets closer to the border, the covered area of this vehicle diminishes.

Figure 6: Rectangular region and covered areas: completely inside (left), close to one border (middle), and close to two borders (right).



As shown in Figure 6, the covered area depends on the position of the vehicle. Equation (28) calculates the covered area of a single vehicle as a function of its position (x, y), considering that $\left(\frac{b}{2}, \frac{a}{2}\right)$ is the region's center, and, so, its origin is at $(0,0)$. In summary, if the vehicle is close to a border, we must subtract the area, corresponding to a triangle, which is not covering the rectangle (is outside). Moreover, it shows the calculations separately for each direction and the unified (both dimensions) formula.

$$
\begin{aligned}
A_X(x) &= \begin{cases} 2\Delta^2 - (\Delta - x)^2, & \text{if } 0 \le x \le \Delta \\ 2\Delta^2, & \text{if } \Delta < x < b - \Delta \\ 2\Delta^2 - (b - x - \Delta)^2 & \text{if } b - \Delta \le x \le b \\ 0, & \text{otherwise} \end{cases} \\
A_Y(y) &= \begin{cases} 2\Delta^2 - (\Delta - y)^2, & \text{if } 0 \le y \le \Delta \\ 2\Delta^2, & \text{if } \Delta < y < a - \Delta \\ 2\Delta^2 - (a - y - \Delta)^2 & \text{if } a - \Delta \le y \le a \\ 0, & \text{otherwise} \end{cases} \\
A(x, y) &= \frac{A_X(x) A_Y(y)}{2\Delta^2}
\end{aligned} \tag{28}
$$

Finally, assuming a uniformly random position for the vehicle, Equation (29) gives the expected coverage area.

$$
\mathrm{E}[A] = \int_0^a \int_0^b \frac{A(x, y)}{ab} dx\, dy \tag{29}
$$

As the demand is also uniformly distributed, it is straightforward to update Equation (27) and obtain an estimate for the probability that a single vehicle covers an arriving passenger (Equation 30).

$$
\Pr(Acceptance) = \frac{\mathrm{E}[\text{Covered area}]}{\text{Total area}} = \frac{\mathrm{E}[A]}{ab} \tag{30}
$$

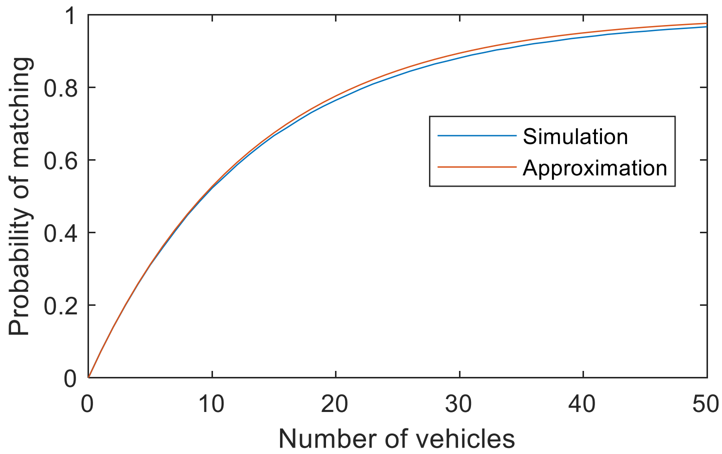
Multiple idle vehicles problem

Now consider that there are $n \ge 1$ idle vehicles in the same rectangular region. One can approximate the probability that at least one vehicle is close enough to an arriving passenger. The proposed approximation is the probability of at least one success on n Bernoulli trials (Equation 31). Where the probability of success is $p = \frac{E[A]}{ab}$ and the probability of failure is $q = 1 - p$.

$$
\Pr(D \le \Delta) = \sum_{j=1}^{n} \sum_{i=1}^{j} \binom{j}{i} p^i q^{j-i} \tag{31}
$$

Figure 7 presents a simple numerical result in comparison to a simulation.

Figure 7: Acceptance probability: Simulation *vs* Approximation.



One critical assumption behind this approximation is that vehicles do not cooperate. Otherwise, they would organize themselves like in a maximum expected coverage problem, therefore in a non-random way.

3.2.2 Matching passengers and partially busy vehicles

Suppose that a vehicle is taking a passenger p_1 to his/her destination through the shortest path possible. Additionally, suppose that inter-arrival times for passengers that match with p_1 follow an exponential distribution with a constant arrival rate λ (in arrivals per unit of time per unit of area – $arrivals \cdot h^{-1}km^{-2}$, for instance). Therefore, it is straightforward to assume that this is a pure-birth process from the current time to the time of arrival at the destination. Without loss of generality, assume that the number of matchings (arrivals) is zero at epoch $t = 0$. The expected time to reach the destination is t_f. The probability that at least one arrival ($\Pr(N > 0)$) will happen during the trip defines the likelihood to occur a matching before reaching the destination.

Therefore, Equation (32) gives the probability to remain with zero arrivals after Δt time units, where $0 < \Delta t \ll t_f$ holds, and $A(t)$ is the area of acceptance for arriving passengers that match p_1 at time t.

$$\Pr(N(t + \Delta t) = 0 | N(t) = 0) = P_0(t + \Delta t) = 1 - \lambda A(t)\Delta t \tag{32}$$

Hence, the following transformation, where $P_0(t)$ is the probability of occurring zero arrivals after t time units, holds:

$$P_0(t + \Delta t) = P_0(t)(1 - \lambda A(t)\Delta t)$$

$$P_0(t + \Delta t) = P_0(t) - \lambda A(t)\Delta t P_0(t)$$

$$P_0(t + \Delta t) - P_0(t) = -\lambda A(t)\Delta t P_0(t)$$

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda A(t) P_0(t)$$

On the limit $\Delta t \to 0$, we have Equation (33).

$$\lim_{\Delta t \to 0} \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = \frac{dP_0(t)}{dt} = -\lambda A(t) P_0(t) \tag{33}$$

Equation (34) is the result of integrating Equation (33) in the time interval $\left(0, t_f\right]$ to reach the destination.

$$P_0(t) = e^{-\lambda \int_0^{t_f} A(t)dt} \tag{34}$$

Finally, Equation (35) gives the probability of occurring at least one arrival that matches p_1.

$$\Pr(N > 0) = 1 - P_0(t) = 1 - e^{-\lambda \int_0^{t_f} A(t)dt} \tag{35}$$

If we cast some more light to the definition of matching, it is possible to reach a more suitable solution. Thus, consider that $A(t)$ is truly the intersection of two different areas. The first area A_w is the same as the diamond shape area seen in Figure 6, the area that the vehicle can reach obeying a maximum waiting time Δ. The second area A_{p_1} relates to a parameter Ω that describes the maximum distance that the vehicle can detour from the shortest path to the destination of p_1. For simplicity, consider that a vehicle can only carry two passengers at a time. Thus, Equation (36) describes this definition of $A(t)$.

$$A(t) = A_{p_1}(v, p_1, \Omega) \cap A_w(v, \Delta) \tag{36}$$

As seen in Figure 6, $A_w(v, \Delta)$ has a diamond shape defined as a function of the location of a vehicle v and the length $\Delta > 0$. However, $A_{p_1}(v, p_1, \Omega)$ has an octagonal shape due to the Manhattan distance metric. Figure 8 illustrates the calculation of A_{p_1} using parameter Ω, and the positions of v, and the destination of p_1. For simplicity, in the remaining of this section, we refer to $A_{p_1}(v, p_1, \Omega)$ and $A_w(v, \Delta)$ only as A_{p_1}, and A_w, respectively.

Figure 8: Illustration of area A_{p_1} calculation for the matching problem.

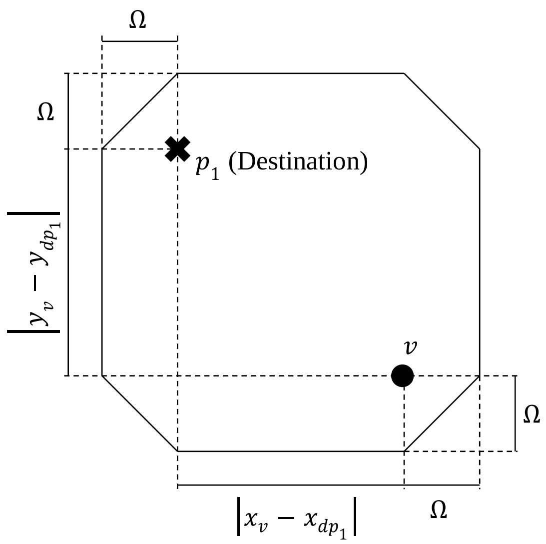
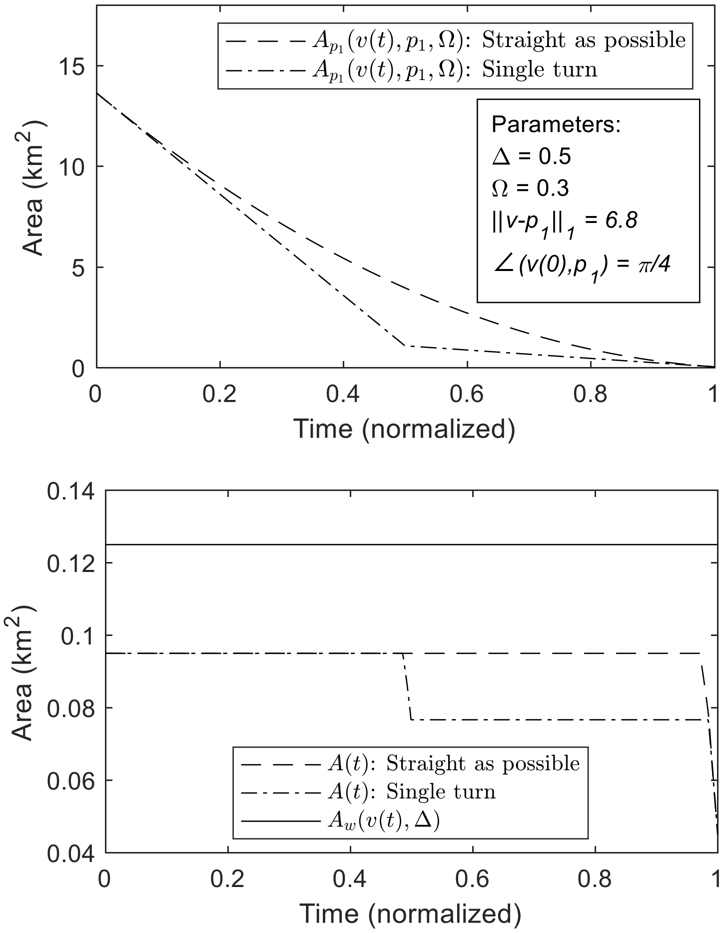


Figure 9 presents a numerical example of areas A_{p_1}, A_w, and $A(t)$ using different strategies to reach the destination. Note that $A(t) \leq \min(A_{p_1}, A_w)$. If $\Delta \leq \Omega$, then $A_w \subset A_{p_1}$, therefore $A(t) = A_w$.

Figure 9: Numerical results on matching areas.



a)

b)

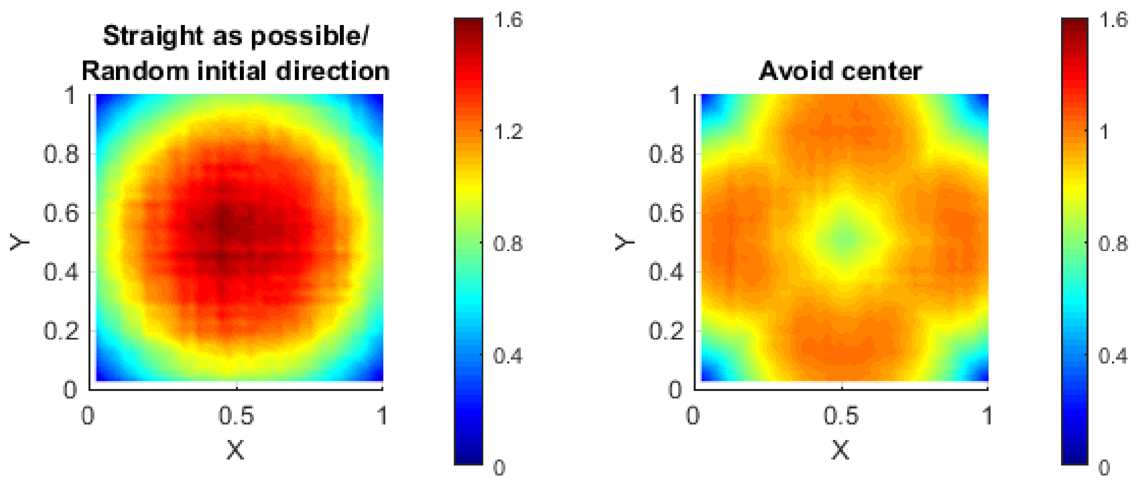
We would like to state that, although "straight as possible" strategy has a larger area of matching; it means a higher number of turns, which usually leads to smaller speeds (and longer travel times). Therefore, it is imperative to evaluate which movement strategy is more efficient. Further research could investigate how time influences the arrival rate of requests that match p_1. One way for this could combine the distributions for distances and angles from Section 3.1.

3.3 Positions

In this section, we analyze the positions of traveling vehicles. We used a simple discrete-event simulation to evaluate the positioning of traveling vehicles in a region. The simulator creates a random origin and destination for a vehicle. The vehicle travels at a constant speed and records positions at fixed time steps. Once the vehicle reaches the destination, it creates new random origin and destination. We plot a histogram of the positions and derive a pdf from it. However, we derive no analytical results.

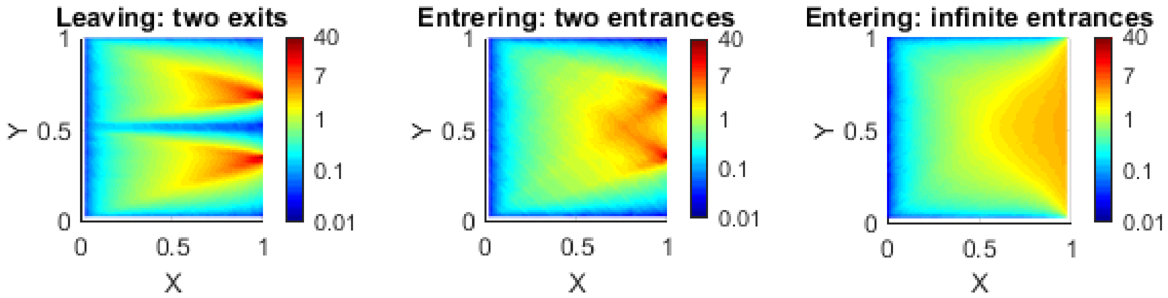
Figure 10 shows the pdfs for two different movement strategies. In the first strategy, the vehicle chooses its initial direction to travel randomly. However, the second strategy forces the vehicle to choose the initial direction such that it passes as far as possible from the center of the region (without increasing the total traveled distance). Note that forcing vehicles to avoid the center created a smoother pdf, suggesting that this would lead to a more homogeneously distributed congestion.

Figure 10: Numerical results on position density functions for intra-region trips.



For the second part of this analysis, we observed the simulation results for inter-region trips. Figure 11 presents the results of the simulation regarding three scenarios: 1) Leaving the region through two exits; 2) Entering a region through two entrances; and 3) Entering a region through infinite entrances. Note that the color map is on a log scale and that the entrances/exits are at the right side of each figure. It is interesting to observe that selecting the closest exit (Left figure) creates a valley between both exits (Supporting the findings from Section 3.1.4). Especially in the figures from the left and the middle, the densities close to the entrances/exits are far higher than in other regions, thus a potential bottleneck. Infinite entrances (figure on the right) distributed the densities.

Figure 11: Numerical results on position density functions for inter-region trips.



4. Final considerations

This paper aimed to provide estimations for travel distances, trajectories and matching on a simplified rectangular urban traffic network to support the modeling of ridesourcing decision problems. Among the most interesting findings, one could highlight three of them. First, Equation (26) shows that just increasing the number of vehicles is not an efficient strategy to lower waiting times if there is no cooperation among them. Second, Section 3.2.1 suggests that cooperation might play a critical role in maximizing the area that vehicles cover, directly influencing waiting times, as well. Third, trajectories influences the chances of matching a second passenger and congestion patterns, as shown in Figure 9 and Figure 10, respectively. Therefore, one willing to increase successful matchings should consider movement strategies.

These findings presented only initial evidence to support more elaborated analytical or simulation models on ridesourcing problems and urban mobility. Besides the continuous space assumption, several other limitations require further examination to provide more reliable and accurate insights. Namely, some of these limitations are: i) uniformly distributed origins and destination, opposing the behavior of a morning commute, for instance; ii) homogeneity in the region configuration, as large urban areas tend to face heterogeneous infrastructure and configuration; iii) exponentially distributed inter-arrival times.

Therefore, further research could focus on coping with the latter. Additionally, one could extend these simple simulations and analytical results for other insights on ridesourcing and urban mobility.

5. References

Ansari, S., M. Başdere, X. Li. and Y. Ouyang (2018) Advancements in continuous approximation models for logistics and transportation systems: 1996-2016, *Transportation Research Part B*, **107**, 229-252.

Berbeglia, G., J.F. Cordeau and G. Laporte (2010) Dynamic pickup and delivery problems, *European Journal of Operational Research*, **202**, 8-15.

Contreras, S.D. and A. Paz (2018) The effects of ride-hailing companies on the taxicab industry in Las Vegas, Nevada, *Transportation Research Part A*, **115**, 63-70.

Furuhata, M., M. Dessouky, F. Ordóñez, M.E. Brunet, X. Wang and S. Koenig (2013) Ridesharing: The state-of-art and future directions, *Transportation Research Part B*, **57**, 28-46.

Gaboune, B., G. Laporte and F. Soumis (1993) Expected Distances Between Two Uniformly Distributed Random Points in Rectangles and Rectangular Parallelepipeds, *Journal of the Operational Research Society*, **44** (5), 513-519.

Hall J.D., C. Palsson and J. Price (2018) Is Uber a substitute or complement for public transit?, *Journal of Urban Economics*, **108**, 36-50.

Jin S.T., H. Kong, R. Wu and D.Z. Sui (2018) Ridesourcing, the sharing economy, and the future of cities, *Cities*, **76**, 96-104.

Larsen, R.J. and M.L. Marx (2012) *An Introduction to Mathematical Statistics and Its Applications*, Prentice Hall, 5th Edition, 757p.

Masmoudi, M.A., M. Hosny, E. Demir, K.N. Genikomsakis and N. Cheikhrouhou (2018) The dial-a-ride problem with electric vehicles and battery swapping stations, *Transportation Research Part E*, **118**, 392-420.

Molenbruch, Y., K. Braekers and A. Caris (2017) Typology and literature review for dial-a-ride problems, Annals of Operations Research, 259, 295-325.

Mora, J.A., S. Samaranayake, A. Wallar, E. Frazzoli and D. Rus (2017) On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment, *Proceedings of the National Academy of Sciences of the United States of America*, **115** (3), 462-467.

Newell, G.F. (1971) Dispatching policies for a transportation route, *Transportation Science*, **5** (1), 91-105.

Nie Y. (2017) How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China, *Transportation Research Part C*, **79**, 242-256.

Rayle, L., D. Dai, N. Chan, R. Cervero and S. Shaheen (2016) Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Fancisco, *Transport Policy*, **45**, 168-178.

Santi P., G. Resta, M. Szell, S. Sobolevsky, S.H. Strogatz and C. Ratti (2014) Quantifying the benefits of vehicle pooling with shareability networks, *Proceedings of the National Academy of Sciences of the United States of America*, **111** (37), 13290-13294.

Schwieterman, J. and C.S. Smith (2018) Sharing the ride : A paired-trip analysis of UberPool and Chicago Transit Authority services in Chicago, Illinois, *Research in Transportation Economics*, **71**, 9-16.

Stiglic, M., N. Agatz, M. Savelsbergh and M. Gradisar (2016) Making dynamic ride-sharing work: The impact of driver and rider flexibility, *Transportation Research Part E*, **91**, 190-207.

Vinayak, P., F.F. Dias, S. Astroza, C.R. Bhat, R.M. Pendyala and V.M. Garikapati (2018) Accounting for multi-dimensional dependencies among decision-makers within a generalized model framework: An application to understanding shared mobility service usage levels, *Transport Policy*, **72**, 129-137.

Zha, L., Y. Yin and H. Yang (2016) Economic analysis of ride-sourcing markets, *Transportation Research Part C*, **71**, 249-266.