Uber driver strategic behavior

September 26, 2016

1 Basic Model

Consider a set of n nodes as vertices of a graph $G = (V, E)$ defined over a geographical area. We are interested in creating an optimal contingency plan of trips, for a particular driver who starts his day at node $v_0 \in V$. We assume that the driver decides the maximum units of service time T that he would spend ferrying passengers at the beginning of the day. Let time t denote the number of time units left for a driver before the end of his pre-decided service time.

- For $i, j \in V, d_{ij}$ is the Euclidean distance between the two nodes. The costs to the driver in terms of gas and vehicle depreciation value in traversing this edge is c_{ij} and his earnings in ferrying a passenger from i to j is represented by p_{ij}. Time required to traverse the edge is denoted by τ_{ij}. The costs, the earnings and the time required to make this ride are all proportional to the distance travelled.

- Let $S^{(t)}_{n\times 1}$ be the surge multiplier vector over the nodes of the graph. Each entry $s^{(t)}_i$ denotes the surge multiplier in effect at node i at time t. When $s^{(t)}_i > 1$, the driver earning from ferrying a passenger from node i to j of the graph is $s^{(t)}_i \times p_{ij}$.

- The arrival process of passengers at node i at time t is a Poisson process with rate $\lambda^{(t)}_i$. Let vector $\Lambda^{(t)}_{n\times 1}$ denote the passenger arrival rates across all the vertices of G at time t. While some passengers may arrive simultaneously (compound Poisson process), we assume that this effect is negligible. Furthermore, we assume that if the driver is present at node i when a passenger arrives, he picks up the passenger with no further delay. As the inter-arrival times of passengers are exponentially distributed, the expected value of idle time for a driver waiting at node i is $1/\lambda^{(t)}_i$. We implicitly assume that the passenger arrival process is stationary for a finite time slice around t.

- The matrix $M^{(t)}_{n\times n}$ represents the transition probabilites between the nodes of the graph such that the fraction of passengers at node i whose destination is node j at time t is represented by entry $m^{(t)}_{ij}$, where $\forall t, m^{(t)}_{ij} \in \mathbb{R}_{\geq 0}, m^{(t)}_{ii} = 0$ and $\sum_j m^{(t)}_{ij} = 1$.

- H(.) denotes the Heaviside step function.

$$H(x) = \max \begin{cases} 0, & \text{if } x < 0, \\ 1, & \text{if } x \geq 0 \end{cases}$$

Table 1: Description of Notation

		Definition
d_{ij}	:	Distance between nodes i and j
c_{ij}	:	Driver cost for trip from node i to node j
p_{ij}	:	Driver earnings from passenger trip from node i to node j
τ_{ij}	:	Travel time from node i to node j
$s_i^{(t)}$	:	Surge multiplier active at node i at time t
$\lambda_i^{(t)}$	:	Poisson passenger arrival rate at node i at time t
$m_{ij}^{(t)}$	:	Fraction of passengers at node i destined to node j at time t
$OPT(t,i)$	:	Maximum expected earnings in remaining time units t, while at node i
$H(.)$	:	Heaviside function

We formulate a dynamic program to find the contingency plan to maximise the expected earnings of a driver in his service time T. Each entry $OPT(t,i)$ of the matrix OPT represents the maximum expected earnings in the remaining t units of service time when the driver is stationed at node i. The maximum expected revenue of the driver at the beginning of the day is thus $OPT(T, v_0)$.

- Initialization $\forall i, OPT(0,i) = 0$.
- While calculating each $OPT(t,i)$, our algorithm is faced with two possible choices.
  1. Keep waiting at node i for a passenger.
  2. Traverse to some other node j in search of a passenger.
- It chooses whichever choice maximises the expected earnings in the remaining time units t.

$$OPT(t,i) = \max \begin{cases} \sum\limits_{j \neq i} m_{ij}^{(t)} \left[\left(s_i^{(t)} p_{ij} - c_{ij} \right) + OPT(t - \tau_{ij}, j) \right], & \text{if } \frac{1}{\lambda_i^{(t)}} + \mathbf{E}[\tau_i | M] \leq t, \\ \max\limits_j \left[\left(OPT(t - \tau_{ij}, j) - c_{ij} \right) \times H(t - \tau_{ij}) \right], & \text{otherwise} \end{cases}$$

- While solving the above dynamic program, we maintain an output vector corresponding to a series of driver choices between waiting for a passenger at his current node i or taking an empty ride to some other node j.

Some issues:

1. What happens when the condition $\frac{1}{\lambda_i^{(t)}} + \mathbf{E}[\tau_i | M] \leq t$ is not satisfied, while at the same time $\forall j, H(t - \tau_{ij}) = 0$ i.e., there is no node j such that $t - \tau_{ij} > 0$? Should the driver stay at node i (first condition is probabilistic in nature) or should he simply call it a day?

2. If the condition $\frac{1}{\lambda_i^{(t)}} + \mathbf{E}[\tau_i|M] \leq t$ is satisfied, to calculate the quantity $OPT(t-\tau_{ij}, j)$, we will need future values of transition matrix $M^{(t-\tau_{ij})}$ and surge multiplier vector $S^{(t-\tau_{ij})}$. Either, we will have to assume that a driver has complete knowledge of how the system will behave during the day, right at the beginning of the day or make some assumption e.g., transition matrix and surge multiplier vector remain fixed at the initial values throughout the day, $\forall t, M^{(t)} = M^{(T)}, S^{(t)} = S^{(T)}$.