# Travel time estimation accuracy in developing regions

Daksh Shah, Aravinda Kumaran, Rijurekha Sen, Ponnurangam Kumaraguru

## Objectives

To investigate the quality of travel time estimates in the Indian capital city of Delhi and the National Capital Region (NCR).

## Introduction

We collected data about 610 Uber trips from 34 users. We empirically show the unpredictability of travel time estimates. It seems that the unpredictability leads to a whopping 28.4% of the requested trips being cancelled. Our empirical observations differ significantly from the high accuracies reported in travel time estimation literature.

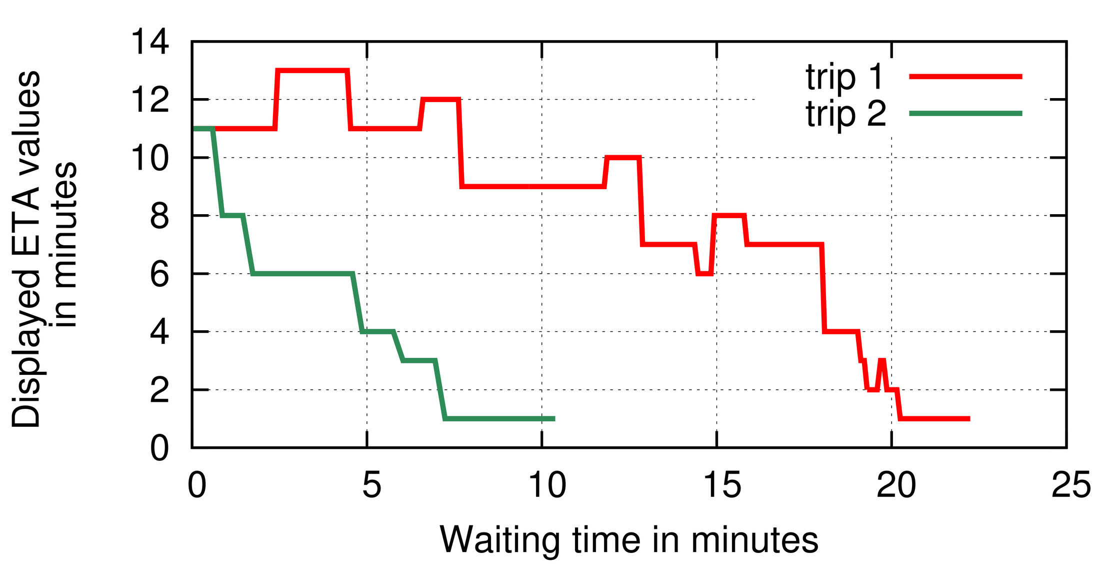


Figure 1: Unpredictable static values and jumps in ETA during waiting, for 11 mins initial ETA

Fig. 1 depicts two different trips with the same *t2_first* along the y-axis (11 minutes). The x-axis shows the actual waiting times. Ideally, both curves should monotonically decrease by one minute every minute and reach 0 after 11 minutes. Trip denoted in red has *t2* reaching 0 in almost 21 minutes, with many more *t2_stationary* and *t2_jumps* instances.

Fig. 3, CDF in red shows the difference between the ETA shown after booking (*t2_first*) and the actual waiting time. For less than $20^{th}$ percentile, the actual waiting times were less than or equal to the ETA. More common, however, is 5-10 minutes of differences (median to $80^{th}$ percentile), with the difference going to more than 20-25 minutes in the worst case.

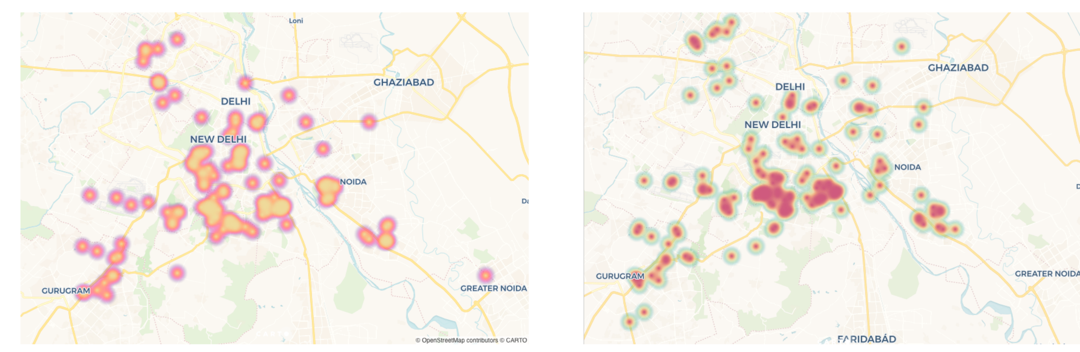

Figure 2: Heatmaps of source and destination of the trips made by our recruited participants, showing the significant geographical coverage by the participants

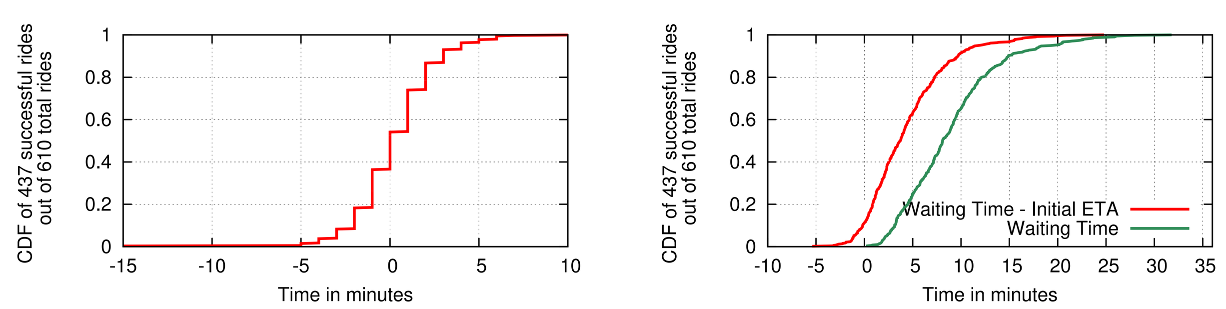


Figure 3: a) ETA difference before and after booking b) Waiting times and difference between initial ETA displayed after booking and actual waiting times

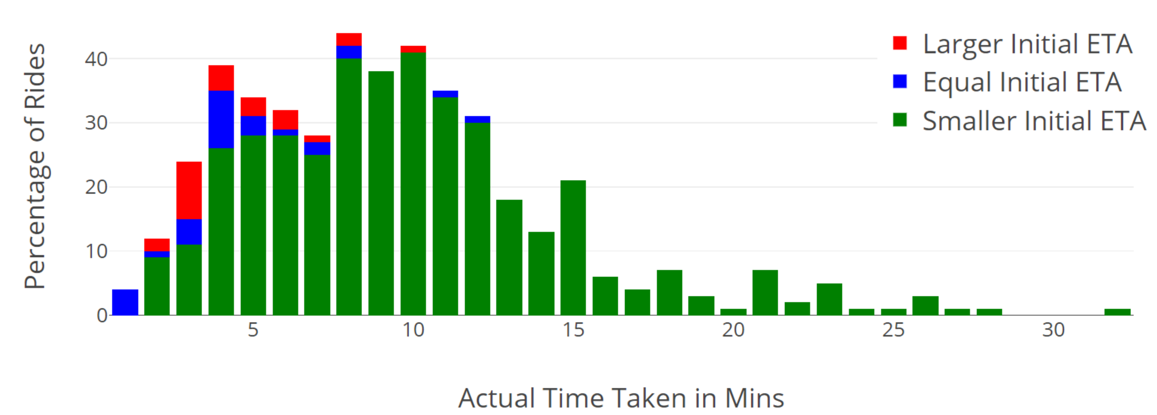


Figure 4: Initial ETA and actual waiting times mismatch

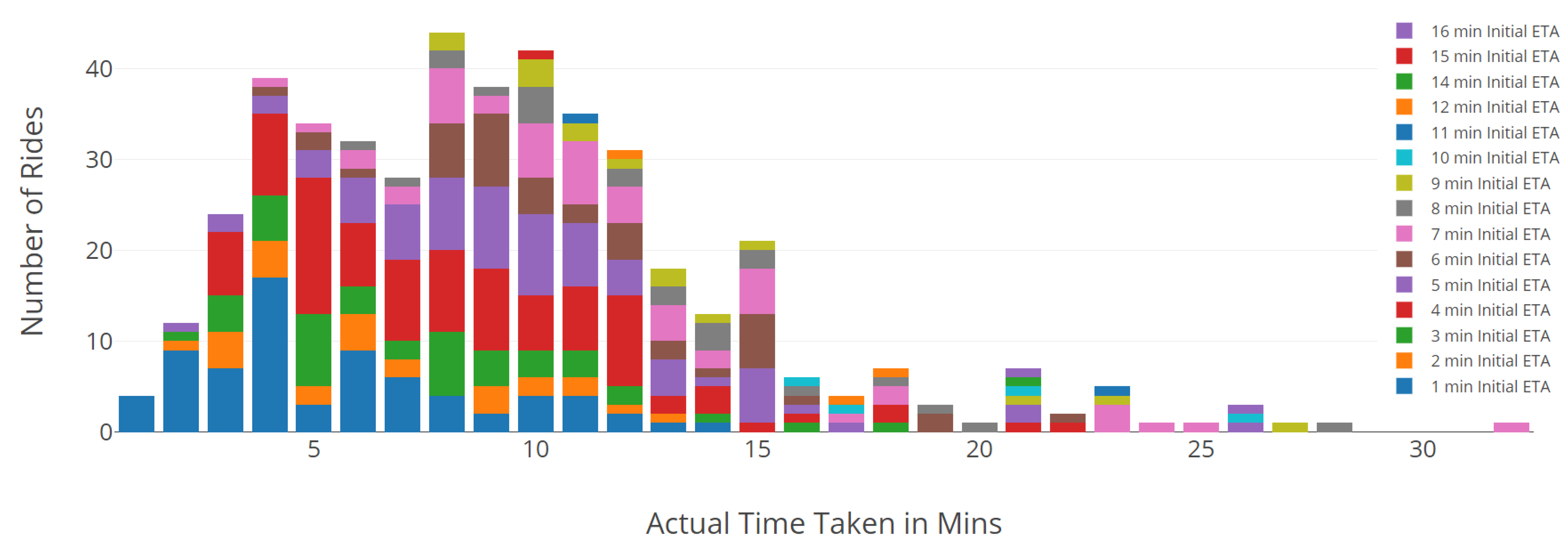


Figure 8: Wide range of initial ETA for the same actual waiting time

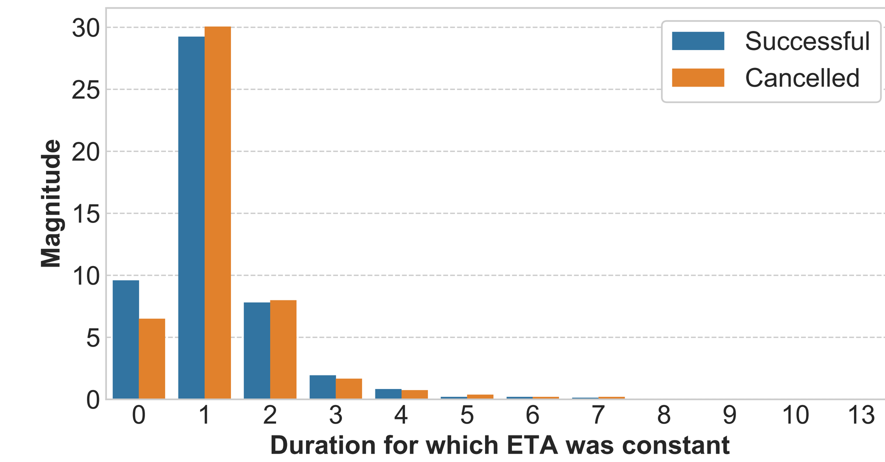


Figure 5: ETA showing constant value for more than a minute

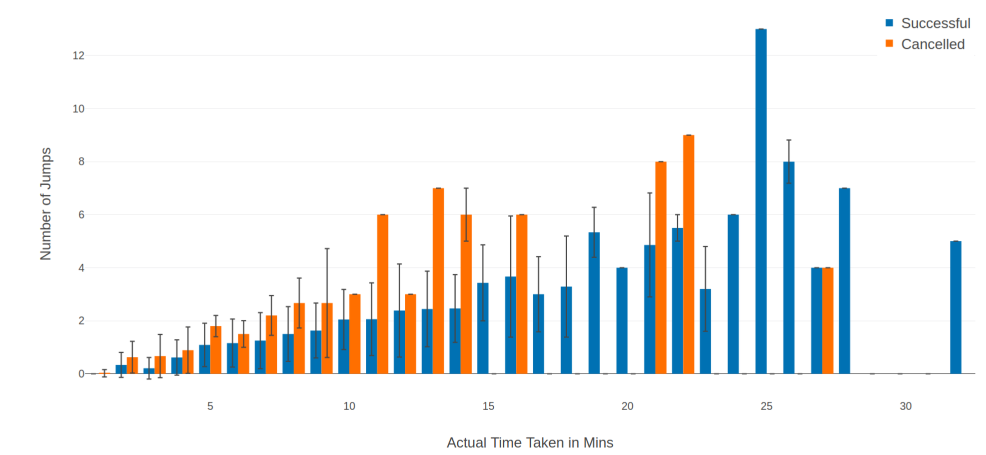

Figure 6: ETA showing upward jumps

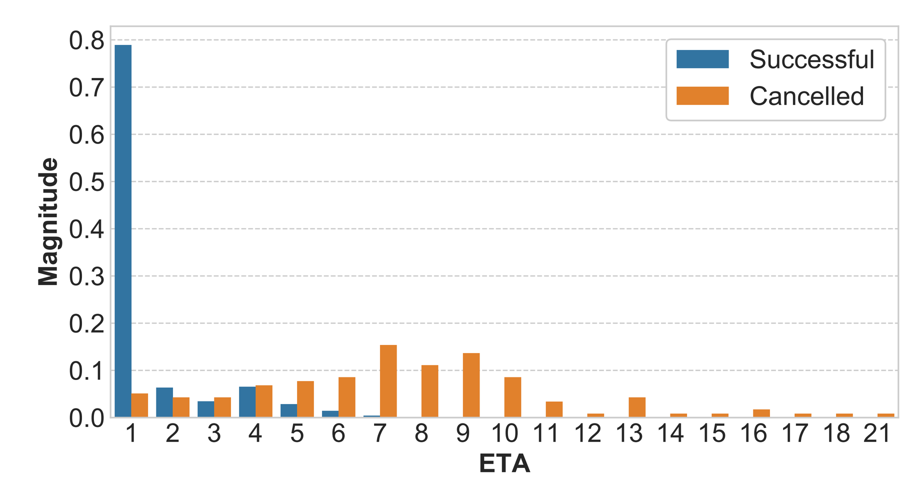


Figure 7: Final ETA for successful vs. cancelled trips

## Glossary

1. ***t1***: ETA shown when the app is opened and location is detected (same as the ETA returned by API).
2. ***t2***: After booking, a series of ETA values are displayed in the app, until the cab finally arrives or the trip is cancelled. This range of displayed ETA values as the passenger is waiting is referred to as t2
3. $\boldsymbol{t2_{first}}$: First ETA as shown in the app after the cab is booked.

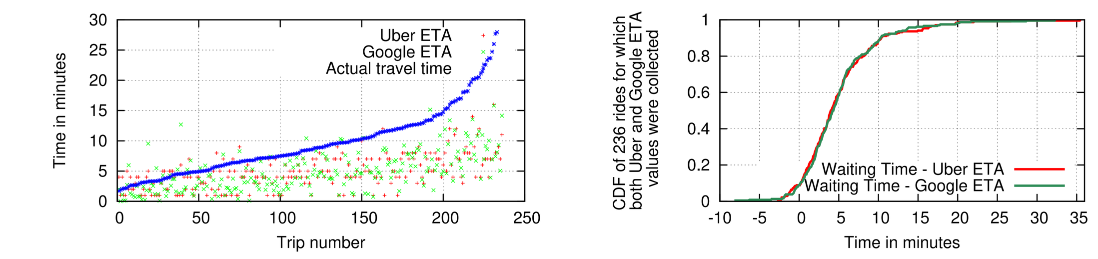

Figure 9: Google vs. Uber ETA values. Though the absolute values are different, the distribution of the differences with actual waiting times is very similar.

| Travel time estimation algorithms | Porto MAE (sec) | Shanghai MAE (sec) |
|---|---|---|
| RTTE [Rahmani et al., 2013] | 169.45 | 214.01 |
| PTTE [Wang et al., 2014] | 159.43 | 168.48 |
| SVR [Asif et al., 2014] | 241.41 | 424.12 |
| SAE [Lv et al., 2015] | 222.06 | 310.47 |
| spd-LSTM [Ma et al., 2015] | 217.37 | 302.45 |
| TEMP[Wang et al., 2016] | 193.61 | 248.70 |
| **Deeptravel[Zhang et al., 2018]** | **113.24** | **126.59** |

Table 1: Mean Absolute Error (MAE) of different travel time estimation algorithms as per State of the Art Papers.

## Possible Reasons for Mismatching Estimates

1. A lack of training data issue for developing countries
2. An algorithmic shortcoming that cannot capture the (lack of) historical patterns in developing region travel times
3. A conscious policy decision by Uber platform or Uber drivers, to mismatch the correctly predicted travel time estimates and increase cab cancellation fees?

## Conclusion

This paper identifies an important literature vs. practice gap in travel time estimation accuracy in developing regions using em- pirical data. In future, we will work on quantifying training data and algorithmic limits for this problem in developing regions, by generating large scale travel time datasets. It is necessary to bring more transparency to complex urban mobility services like Uber, and this discussion paper establishes this necessity.