

統計資料

一開始paper先簡介一些統計資料，第2頁下圖表是參與survey的統計，有性別、年齡等等。在第3頁上方分別是每次trip花費時間和一日有幾個trip的histogram。統計方式方式是每個taxi安裝GPS，定時間記錄taxi的坐標。

建模

做法把地圖切成方格狀（第4頁上圖），每個方格標定數字。他將每個trip從頭到尾經過的方格記成數列。以第4頁下圖左為例，駕駛經過的方格序列可以記成 $S = 2, 7, 12, 11, 16, \dots, 8$ 。

同時，每個格子會有一個「理論到目的地最少時間」，paper記做 t ，這數值先用Microsoft MapPoint計算出來，接，對於任何一個方格序列 S ，他統計序列中每進入一個新的格子之後 t 變化多少。

以 $S = 2, 7, 12, 11, 16, \dots, 8$ 為例，有2的 t 和7的 t 的差值、7的 t 和12的 t 的差值等，差值可以是正數可以是負數，記作 Δt 。

把這些 Δt 搜集起來得到第4頁的histogram。理論上如果駕駛規劃路徑能力不錯， Δt 應盡可能持續負值。

為了衡量駕駛可以做到「較佳的路徑規劃」的能力，他定義 p 值為 $\frac{\Delta t_{sample}}{\Delta t_{sample}}$ 。

目的地預測

給定一個方格路徑序列 S ，代表目前已經行走的路徑，計算目的地為 c_i 的機率，寫為 $p(c_i|S)$ 。

然後套用aBaye's rule

$$p(c_i|S) = \frac{p(S|c_i)p(c_i)}{\sum_{j=1}^{N_c} p(S|c_j)p(c_j)}$$

N_c 是全部的格子數目

paper中假設 $p(c_i) = \frac{1}{N_c}$ ，以及（抄原文避免失原意）

$$p(S|c_i) = \prod_{j=2}^{N_s} \begin{cases} p & \text{if } s_j \text{ is closer to } c_j \text{ than any previous cell in } S \\ (1-p) & \text{otherwise} \end{cases}$$

p 值利用上面的 Δt 建模方法計算。