Estimating 24/7 origin-destination demand using high granularity multi-source traffic data

Wei Ma, Sean (Zhen) Qian
Civil and Environmental Engineering, Carnegie Mellon University

Abstract

Dynamic origin-destination (OD) demand is central to transportation system modeling and analysis. The dynamic OD demand estimation problem (DODE) has been studied for decades, most of which solve the DODE problem on a typical day or several typical hours. There is a lack of methods that estimate high-resolution dynamic OD demand for a sequence of many consecutive days over several years (referred to as 24/7 OD in this research). Having multi-year 24/7 OD demand would allow a better understanding of characteristics of dynamic OD demands and their evolution/trends over the past few years, a critical input for modeling transportation system evolution and reliability. This paper presents a data-driven framework that estimates day-to-day dynamic OD using high-granular traffic counts and speed data collected over many years. The proposed framework statistically clusters daily traffic data into typical traffic patterns using t-Distributed Stochastic Neighbor Embedding (t-SNE) and k-means methods. A stochastic projected gradient descent method is proposed to efficiently solve the multi-year 24/7 DODE problem.

Background

- Most of the existing DODE methods require a dynamic traffic loading (DNL) process to describe the traffic flow evolution and congestion spill-back endogenously. As the DNL process requires relatively high computational budget, it takes hours to estimate dynamic OD demand even for a single day.
- The majority of previous studies estimate OD demand for a single time period or a single day. The daily, weekly, monthly patterns of OD demand has not be taken into consideration of the OD estimation methods.

Contributions

- It proposes a framework for estimating multi-year 24/7 dynamic OD demand using high-granular traffic flow counts and speed data.
- It takes into account day-to-day features of flow patterns by defining and calibrating the dynamic assignment ratio (DAR) matrix using real-world data, which enables realistic representation and efficient computing of network traffic flow.
- It adopts t-SNE and k-means methods to cluster day-to-day traffic data into different traffic scenarios.
- It proposes a stochastic projected gradient descent method to solve the DODE problem. The proposed method is suitable for GPU computation, which enables efficiently estimating high-dimensional OD over many years.

Framework

$$x_a(t_2) = \sum_{rs\in K_q}\sum_{k\in K_{rs}} \delta_{rs}^{ka} f_{rs}^{k}\left(\tau_{rs}^{ka}(t_2)\right)$$

- For any link a, the flow at time t_2 is the sum of flow that goes through link a, and departs at time
- Assumptions
  - ❖ Traffic spread evenly on each TMC, within 30 seconds
  - ❖ First in first out (FIFO)
- Travelers make route choice at each time interval t

$$f(t) = P(t)q(t)$$

- Dynamic Assignment Ratio (DAR)

$$\boxed{\rho_{rs}^{ka}(h_1,h_2) = \frac{\int_{t_1\in H_{h_1}\cap\tau_{rs}^{ka}(H_{h_2})} f_{rs}^{k}(t_1)dt_1}{\bar{f}_{rs}^{h_1 k}}}$$

- Minimize the gap between observed counts and estimated counts

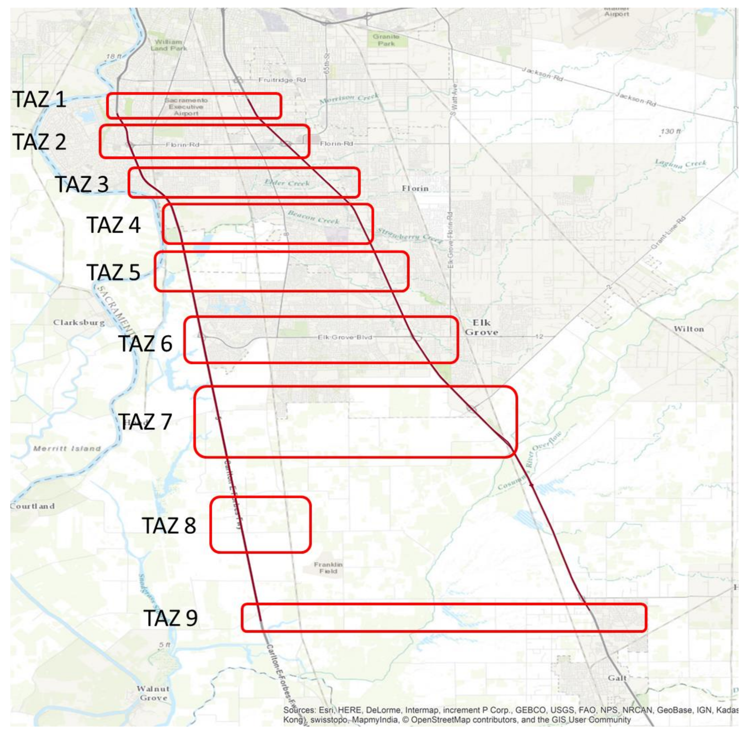
$$\min_{\{q_{rs}^{h_1}\}_{r,s,h_1}} \sum_{a\in A^o}\sum_{h_2=1}^{N}\left\|\tilde{x}_a^{h_2} - \sum_{rs\in K_q}\sum_{k\in K_{rs}}\sum_{h_1=1}^{N}\delta_{rs}^{ka}\rho_{rs}^{ka}(h_1,h_2)p_{rs}^{kh_1}\tilde{q}_{rs}^{h_1}\right\|_2^2$$

$$\text{s.t.}\quad \tilde{q}_{rs}^{h_1} \geq 0 \qquad \forall rs\in K_q, 1\leq h_1\leq N$$

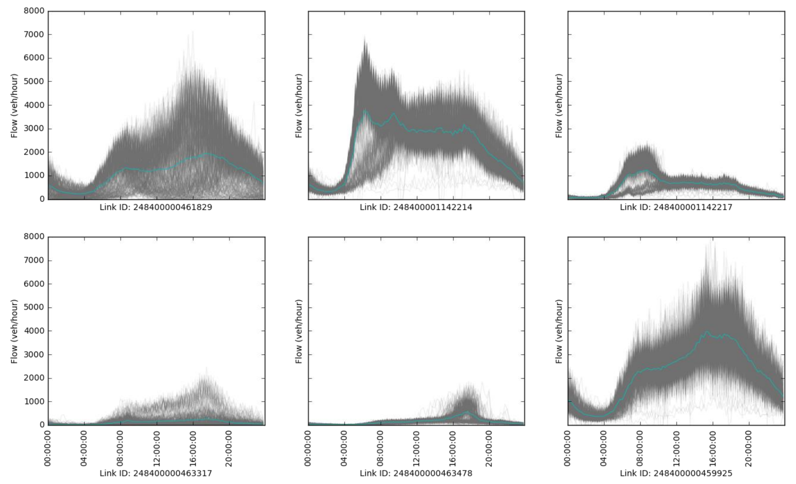
Case Study: I-5 and Hwy-99 towards Sacramento

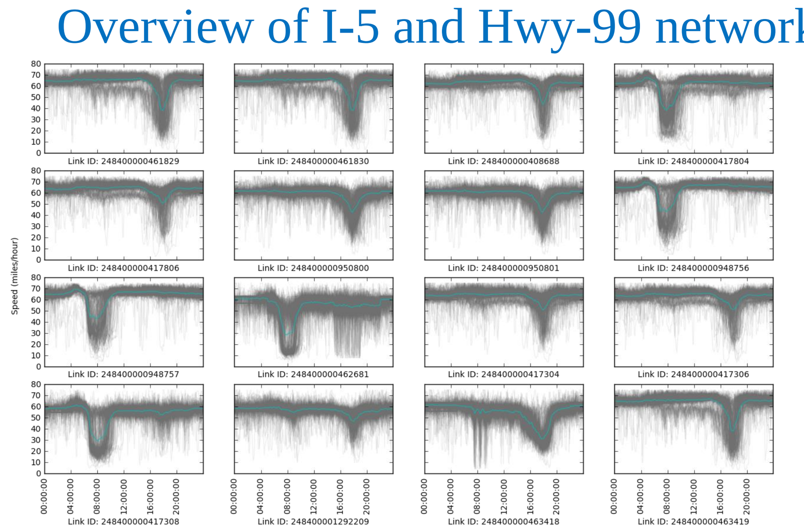
The network for I-5 and Hwy-99 is obtained from HERE. We extract the two highways to be our base network. The OD connectors are constructed based on the residence region and exits/entrances of two highways. We discompose the whole network into 9 traffic analysis zones (TAZs), and attach one origin and one destination to each TAZ.

- The flow counts raw data are obtained from Caltrans Performance Measurement System (PeMS).
- Traffic speed data were obtained from HERE. The traffic speed data are provided at the geographic level of Traffic Message Channel (TMC), one of the geo-reference protocols.

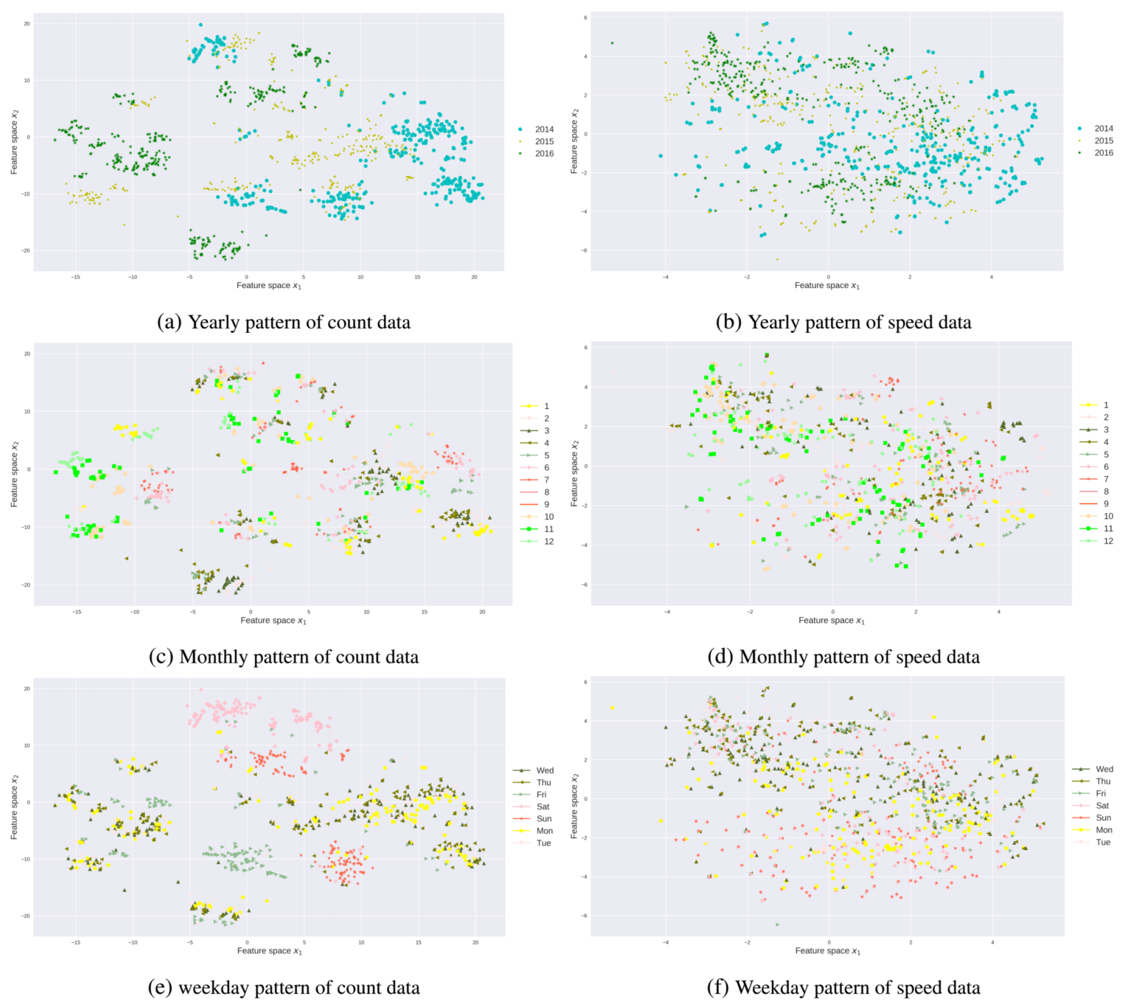


Overview of I-5 and Hwy-99 network

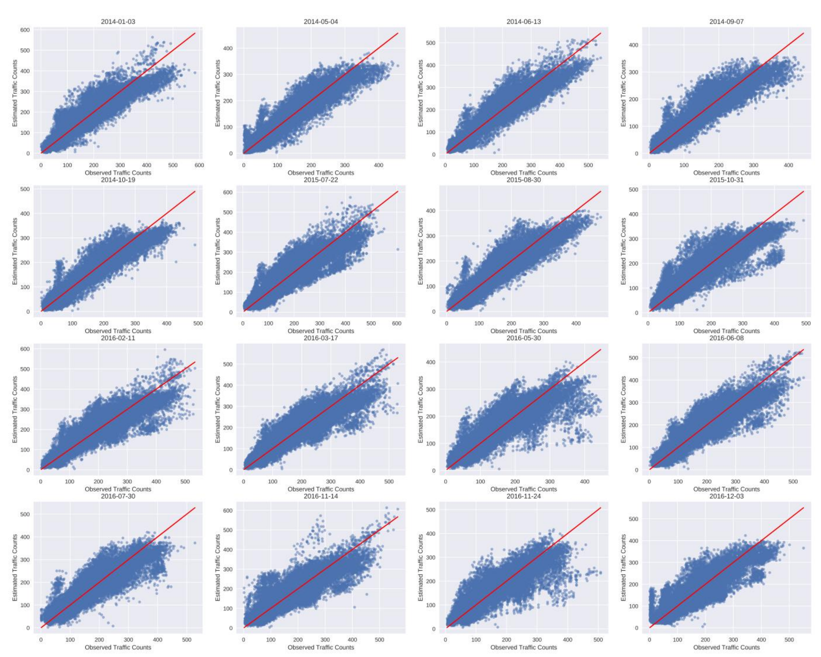




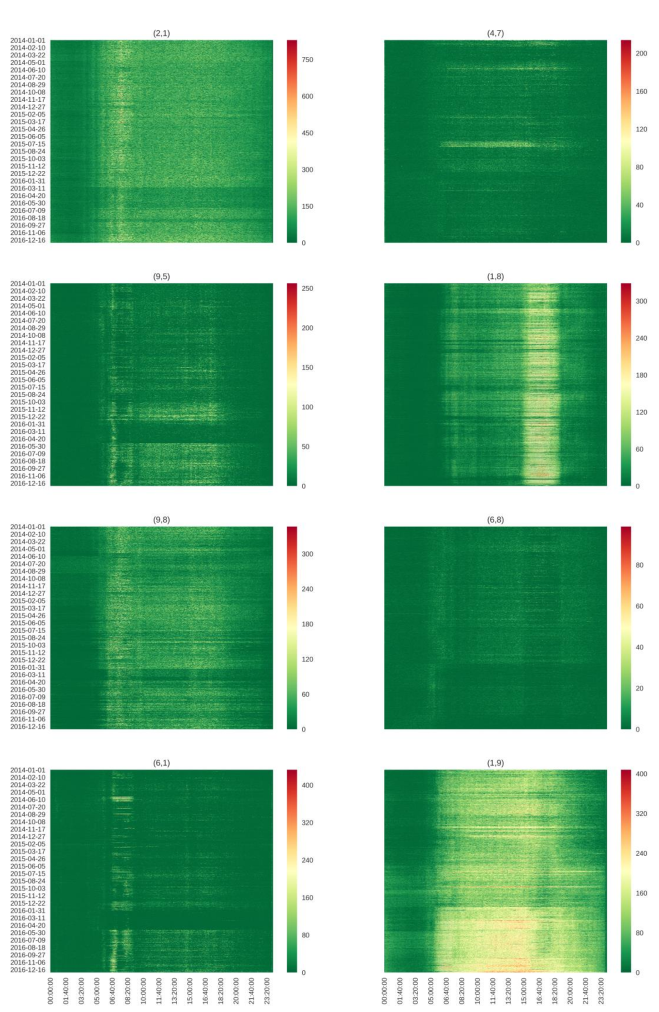
Overview of count and speed data



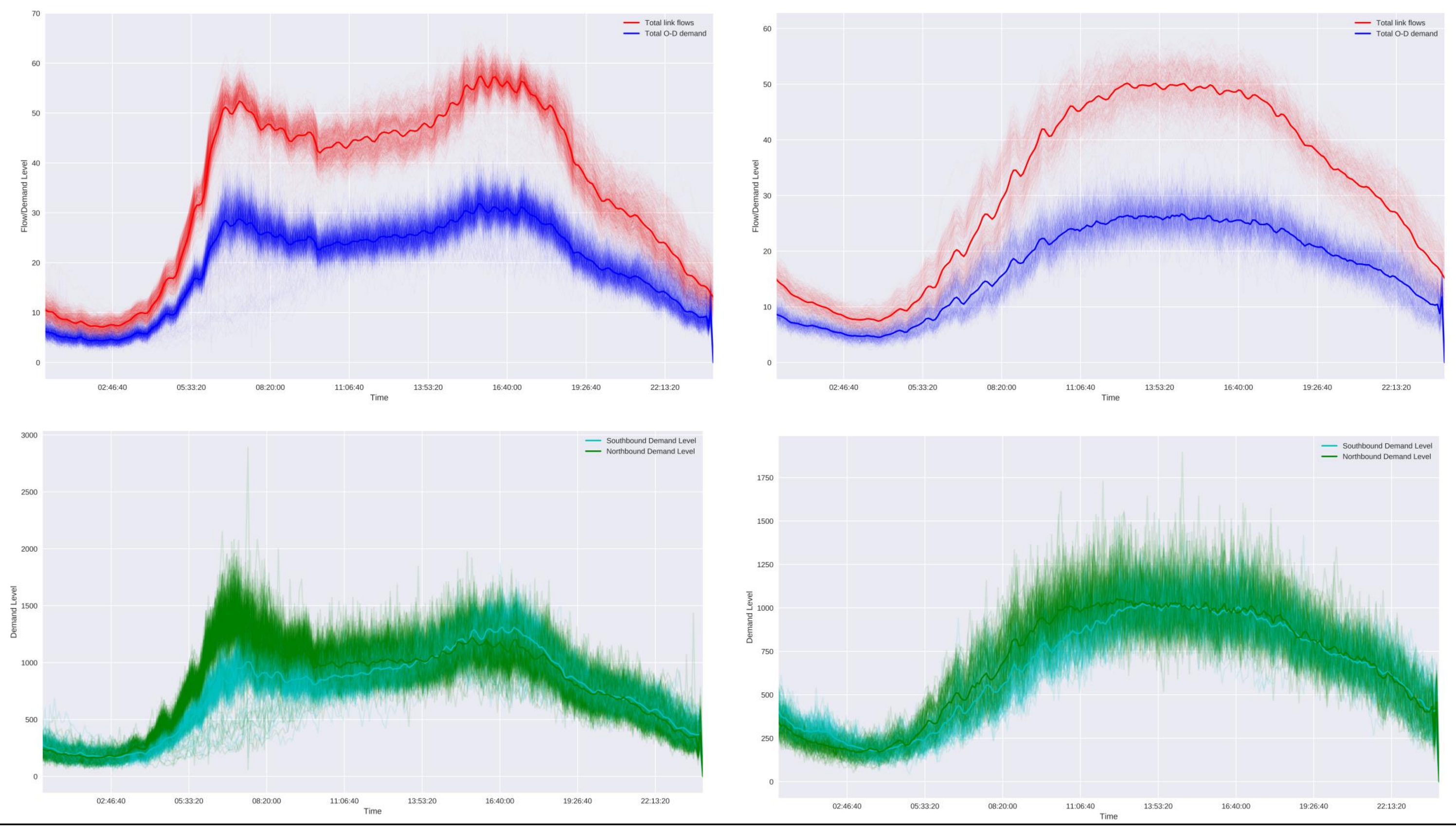
Daily, weekly, monthly patterns of traffic data



Goodness of fit



Traffic demand for eight randomly selected OD pairs



MAC Mobility Data Analytics Center
Carnegie Mellon University

A Division of Traffic 21 Reseearch Institute

traffic21
a transportation research institute of Carnegie Mellon University