

21st International Symposium on Transportation and Traffic Theory, ISTTTT21 2015,
5-7 August 2015, Kobe, Japan

Cellpath: fusion of cellular and traffic sensor data for route flow estimation via convex optimization

Cathy Wu^{a,*}, Jérôme Thai^a, Steve Yadlowsky^a,
Alexei Pozdnoukhov^b, Alexandre Bayen^{a,b,c}

^aElectrical Engineering and Computer Sciences, UC Berkeley, 652 Sutardja Dai Hall, Berkeley, CA 94720, USA

^bCivil and Environmental Engineering, UC Berkeley, 109 McLaughlin Hall, Berkeley, CA 94720, USA

^cInstitute for Transportation Studies (ITS), UC Berkeley, 109 McLaughlin Hall, Berkeley, CA 94720, USA

Abstract

A new convex optimization framework is developed for the route flow estimation problem from the fusion of vehicle count and cellular network data. The issue of highly underdetermined link flow based methods in transportation networks is investigated, then solved using the proposed concept of *cellpaths* for cellular network data. With this data-driven approach, our proposed approach is versatile: it is compatible with other data sources, and it is model agnostic and thus compatible with user equilibrium, system-optimum, Stackelberg concepts, and other models. Using a dimensionality reduction scheme, we design a projected gradient algorithm suitable for the proposed route flow estimation problem. The algorithm solves a block isotonic regression problem in the projection step in linear time. The accuracy, computational efficiency, and versatility of the proposed approach are validated on the I-210 corridor near Los Angeles, where we achieve 90% route flow accuracy with 1033 traffic sensors and 1000 cellular towers covering a large network of highways and arterials with more than 20,000 links. In contrast to long-term land use planning applications, we demonstrate the first system to our knowledge that can produce route-level flow estimates suitable for short time horizon prediction and control applications in traffic management. Our system is open source and available for validation and extension.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Selection and peer-review under responsibility of Kobe University

Keywords: Route flow estimation; cellular network data; convex optimization; traffic assignment; simplex constraints; isotonic regression

1. Introduction

While there is a wealth of literature in transportation science that is aimed at modeling, computing, and estimating the movement of people in terms of *link* flows and *origin-destination (OD)* flows, there is relatively little work focused on *route* flow estimation. The route flow estimation problem is particularly important because route flow estimates can capture phenomena in traffic behavior that link flows and OD flows (also called OD demands) cannot. For instance, route flows would enable analysis and re-routing of commuters who would be most affected by a link closure. Additionally, route flows provides a rich state estimate of the network which may be used to compute link

* Corresponding author. Tel.: +1-510-642-3585; fax: +1-510-643-3955.

E-mail address: cathywu@eecs.berkeley.edu

flows, OD flows, turning ratios, etc., thereby providing backwards compatibility with past and ongoing work that builds upon those estimates.

Simultaneously accurate and efficient methods for estimating route flows are crucial for large scale urban network analysis and planning demands. However, the first step for many approaches to estimating route flow requires enumerating all feasible routes, which is an unreasonable task for many urban road networks because it may require exponential time to compute (Ford & Fulkerson, 1962, §1.2). Classically, the set of potential routes may be extracted from the induced equilibrium in network flow models. At the cost of restrictive assumptions, *deterministic user equilibrium* (UE) (Wardrop & Whitehead, 1952) permits the modeling of unique link flows and feasible route (or path) flows without requiring full route enumeration (Sheffi, 1985, §3.3), (Bell & Iida, 1997, §5.2)). The *stochastic user equilibrium* (SUE) (probit-based (Daganzo & Sheffi, 1977; Maher & Hughes, 1997) and logit-based (Fisk, 1980; Bell & Iida, 1997)) addresses some of the shortcomings of the UE by modeling imperfect knowledge of the network and variation in drivers' preferences, which makes the estimation of route flows possible (Bell *et al.*, 1997). However, frequent perturbations in traffic networks indicate that real-world transportation networks may not be in equilibrium (or only approximately so) (Hato *et al.*, 1999), so we develop a data-driven framework that focuses on effectively utilizing the large amount of data available for estimation in traffic networks. Indeed, in recent years, the growing number of mobile sensors in urban areas enables the use of probe vehicles for route inference from GPS traces (Hunter *et al.*, 2009; Rahmani & Koutsopoulos, 2013).

1.1. Traffic data sources

Traditional traffic sensing systems such as loop detectors embedded in the pavement and cameras provide accurate volume and speed estimates, but their placements are typically sparse and their information content is too coarse. Most importantly, they measure total counts of vehicles passing through a road segment without distinguishing between vehicles following different routes. In order to partially address the shortage of information on the routes followed by vehicles, other types of static sensors have been deployed on the road network: cameras that measure split ratios at different intersections (Veeraraghavan *et al.*, 2003) and plate scanning systems (Castillo *et al.*, 2008, 2010). However these systems require costly infrastructure and only provide highly localized traffic information. Meanwhile, given the large penetration of mobile phones among the driving population and the ubiquitous coverage of service providers in urban areas, mobile phones have become an increasingly popular source of location data for the transportation community. For example, dynamic probing by means of in-car GPS traces (Work *et al.*, 2008; Herrera *et al.*, 2009; Hunter *et al.*, 2009) is a promising technology for trajectory recovery and travel time estimation. However, due to the read-only nature of GPS signals, the low penetration rate of GPS-enabled devices running a dedicated sensing application currently limits the ability to accurately estimate traffic volumes, and it is also unlikely that such data would become available to public agencies (Patire *et al.*, 2013).

Cellular network data, in contrast to GPS traces, benefit from dedicated communication between mobile phones and cellular network network base stations, and the (coarse) location data are available directly from cellular communication network operators. Cellular network infrastructures record a variety of phone to cell communication events, such as *handovers* (HO), *location updates* (LU) and *call detail records* (CDR) (Volinsky *et al.*, 2011a,b), and this data has already been shown to be effective in studying urban environments (Candia *et al.*, 2008a; Jiang *et al.*, 2013; Toole *et al.*, 2012). Since typical cellular networks in urban areas include thousands of cells, HO/LU/CDR events are dense enough to be used effectively to estimate the route choice of agents without requiring any additional infrastructure. When an agent is moving, HOs transfer ongoing calls or data sessions from one cell to another without disconnecting the session, and LUs allow a mobile device to inform the cellular network when the device move from one location (or cell) to the next. CDRs (mainly used by service providers for billing purposes) contain timestamped summaries of the cell through which each data transmission came, and therefore contain abundant mobility traces for a majority of the population. Due to the granularity of sensing, these records alone are not sufficient for recovering agent routes precisely. The spatial resolution of CDR, HO, and LU data varies with the density of antennas and is roughly proportional to the daytime population density. In the present work, we use a standard localization approach when dealing with cellular data based on Voronoi tessellation, a simple model solely based on the locations of the cell towers (Baert & Seme, 2004; Candia *et al.*, 2008b).

Download English Version:

<https://daneshyari.com/en/article/1106921>

Download Persian Version:

<https://daneshyari.com/article/1106921>

[Daneshyari.com](https://daneshyari.com)