Transportation Research Part B 105 (2017) 270–296

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

A decomposition approach to the static traffic assignment problem

Ehsan Jafari*, Venktesh Pandey, Stephen D. Boyles

Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, 1 University Station C1761, Austin, TX 78712, United States

ARTICLE INFO

Article history:
Received 7 September 2016
Revised 14 September 2017
Accepted 16 September 2017
Available online 25 September 2017

Keywords:
Traffic assignment
Spatial decomposition
Parallel schemes for traffic assignment
Statewide modeling
Equilibrium sensitivity analysis

ABSTRACT

This paper describes a spatial parallelization scheme for the static traffic assignment problem. In this scheme, which we term a decomposition approach to the static traffic assignment problem (DSTAP), the network is divided into smaller networks, and the algorithm alternates between equilibrating these networks as subproblems, and master iterations using a simplified version of the full network. The simplified network used for the master iterations is based on linearizations to the equilibrium solution for each subnetwork obtained using sensitivity analysis techniques. We prove that the DSTAP method converges to the equilibrium solution on the full network, and demonstrate computational savings of 35–70% on the Austin network. Natural applications of this method are statewide or national assignment problems, or cities with rivers or other geographic features where subnetworks can be easily defined.

1. Introduction

The traffic assignment problem (TAP) formulated by Beckmann et al. (1956) is used in transportation planning throughout the world, to predict drivers' route choice, and the resulting flows on roadway links (Patriksson, 1994). Owing to its elegant formulation as a convex program with an underlying network structure, this problem can be efficiently solved to high precision on city-scale networks using any number of modern algorithms (Dial, 2006; Bar-Gera, 2010; Gentile, 2014). However, as computational hardware and algorithms advance, attention shifts to more demanding applications of the traffic assignment problem. Such applications include bilevel programs whose solution often requires the solution of many TAP instances as subproblems, accounting for forecasting errors with Monte Carlo simulation of input parameters, and broadening the geographic scope of models to the statewide or national levels.

Parallel computing is a general technique for reducing the running time of algorithms, by identifying problem components which can be solved independently, and brought together at a later point in time. Many algorithms for TAP naturally lend themselves to parallelization (Chen and Meyer, 1988; Karakitsiou et al., 2004). For instance, the classic Frank–Wolfe algorithm can be parallelized by origin or destination when finding shortest paths and building the all-or-nothing link flow vector used in the search direction, and by link when determining the step size. The disaggregate simplicial decomposition method proposed by Larsson and Patriksson (1992) can be parallelized by origin–destination pair (Lotito, 2006; Karakitsiou et al., 2004). Similarly, while most bush-based algorithms have been specified (and convergence proved) with sequential

* Corresponding author.
E-mail addresses: ejafari@utexas.edu (E. Jafari), venktesh@utexas.edu (V. Pandey), sboyles@mail.utexas.edu (S.D. Boyles).

http://dx.doi.org/10.1016/j.trb.2017.09.011
0191-2615/

solution of the origin-specific subproblems, in principle these calculations can be parallelized as well. (A review on computation saving from these parallelization techniques is provided in the next section.)

This paper introduces a new way of parallelizing traffic assignment, using a spatial decomposition by geographic region rather than by origin. We term this algorithm a *decomposed static traffic assignment problem* (DSTAP) approach, in which the full network is divided into subnetworks. A *regional network* is also created as an abstraction of the full network. The DSTAP algorithm iterates between solving equilibrium on these subnetworks and on the regional network, with the demand across the boundaries of the subnetworks obtained from the regional network, and the structure of the regional network updated based on the subnetwork equilibria.

As is shown in this paper, the DSTAP algorithm converges to the same equilibrium solution as would be obtained for the full network. Numerical experiments on the Austin, TX regional network also show a substantial reduction in computation time (ranging from 35–70%) in comparison with solving the full network using gradient projection. Although we show correctness on general networks, DSTAP is most obviously suited for assignment problems on networks which naturally divide into subnetworks. Examples include statewide or nationwide models, where clearly-defined urban areas are connected by sparser rural regions, or in cities partitioned by rivers or other geographic features. In this paper we do not consider how best to partition a network into subnetworks, although this is a highly interesting problem for future research.

The remainder of this paper is organized as follows. Section 2 presents a review of current modeling methods for TAP, and research studies related to this modeling approach. Section 3 defines terminology related to the spatial decomposition scheme. An informal review of the DSTAP algorithm is provided in Section 4. Section 5 describes the proposed DSTAP approach in detail, and Section 6 provides a proof of its convergence to the full network equilibrium. Section 7 presents numerical results when applying DSTAP to a regional network from Austin, TX, and Section 8 concludes the paper.

2. Literature review

This section provides an overview of the existing literature in the following areas: modeling approaches to solve TAP on large scale networks; a review of network aggregation techniques and their applications in the field of transportation planning; and methods to parallelize the solution of TAP.

Many solution methods have been developed to solve the traffic assignment problem, which can be broadly classified into link-based methods, path-based methods, and bush-based methods. Link-based methods require less operational memory and work in the space of link flows to solve the optimization problem (Frank and Wolfe, 1956; LeBlanc et al., 1975; Mitradjieva and Lindberg, 2013). Path-based methods offer faster convergence compared to link based methods and act on the space of path flows; however, they have larger memory requirements (Jayakrishnan et al., 1994; Florian et al., 2009). Bush-based methods exploit the fact that the set of used paths from each origin at equilibrium forms an acyclic network (Bar-Gera, 2002; Dial, 2006; Nie, 2010; Bar-Gera, 2010; Gentile, 2014). Other recent advancements include $\epsilon-$optimal bush algorithm in Zheng and Peeta (2014) and adaptation of network simplex method to solve TAP for large scale networks proposed in Zheng (2015). These methods have improved the existing state-of-the-art of algorithms and are fast and memory efficient for networks with large scale. However, applying such algorithms to solve equilibrium on very large-scale networks may remain impractical. The computation time is also of concern when an application requires multiple runs of TAP. Network design problems, which are hierarchical optimization problems with user equilibrium constraints, are one such example. The algorithms to approximate solutions to network design problems which rely on performing sensitivity analysis, e.g. Josefsson and Patriksson (2007), require multiple runs of TAP. The Monte Carlo sampling techniques proposed in Duthie et al. (2011) to predict the influence of demand uncertainty and correlations on traffic predictions, also require multiple TAP runs, which necessitates the need of faster algorithms.

Network aggregation techniques are commonly employed when an approximate solution to TAP is desired, within reasonable computation framework. Current statewide planning models still rely on aggregation of the networks within cities, capturing only the major freeways and demand using the freeways. For instance, the Texas Statewide Analysis Model (SAM) captures the lower-level transportation system using centroid connectors which serve as an abstract but aggregate representation of urban transportation networks in different cities (Texas Department of Transportation, 2013). Many planning models also utilize the aggregation of traffic analysis zones (TAZ) to simplify the network at larger scale. Such techniques are employed by most of the statewide planning models in US which aggregate the zones and links in MPO models in the statewide network representation (Horowitz, 2006).

Several methods have been proposed to combine the zones and links in a network to form an aggregate network. Link extraction methods remove the unimportant links and nodes from the network (Haghani and Daskin, 1983), but as shown in Chan (1976), such extraction might lead to unpredictable flow patterns on network. Other researchers have proposed link abstraction methods where set of links and nodes between two nodes are replaced with a single aggregated link. Methods proposed to aggregate series and parallel links for purposes of sketch planning are one form of link abstraction (Boyce et al., 1985; Eash et al., 1983).

Researchers have proposed methods to abstract the links outside a subnetwork in form of an artificial link. Zhou et al. (2006) determine the OD matrices for subnetworks by using a virtual link to capture the behavior of travelers shifting away from the subnetwork. The split proportion for that link is determined using a proportional model based on travel times on paths inside and outside the subnetwork. The virtual link represents trips between boundary nodes that bypass the subnetwork (defined by paths for which at least 50% of its links, but not all, pass through the subnetwork), but