Transportation Research Part E 103 (2017) 261–285

Contents lists available at ScienceDirect

Transportation Research Part E

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

Optimal design of sustainable transit systems in congested urban networks: A macroscopic approach

Mahyar Amirgholy [a], Mehrdad Shahabi [b], H. Oliver Gao [a,*]

[a] School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853, United States
[b] Department of Civil and Coastal Engineering, University of Florida, Gainesville, FL 32611, United States

ARTICLE INFO

Article history:
Received 27 November 2016
Accepted 16 March 2017

Keywords:
Transportation system sustainability
Public transit systems
Macroscopic fundamental diagram
Continuum approximation
System optimization

ABSTRACT

Mass transit is a key component of a sustainable transportation system in urban networks. In this research, we propose a continuum approximation model to optimize the line spacing, stop spacing, headway, and fare of the transit system by minimizing a linear combination of (1) users generalized cost, (2) agency operating cost, and (3) external cost of the emission in the urban region. The design of the transit system can be optimized by minimizing the total cost of the transportation system in three different network allocation scenarios: (i) mixed network (Bus), (ii) dedicated lanes (BRT), and (iii) parallel networks (Metro).

1. Introduction

Rapid growth in urban travel demand in large cities has heightened the need for alternative modes of transport that are socially beneficial, economically justifiable, and environmentally friendly. In populated urban regions, a transit system can play a prominent role in improving the sustainability of transportation by alleviating the unpleasant impacts of congestion across the network. Providing a quality of service competitive with the automobile mode can convince a considerable portion of users to leave their private cars at home and use the public transit system for their trips in the network. A competitive transit system can improve user mobility and reduce the emission of air pollutants in the network by decreasing the number of circulating vehicles and increasing the average speed of traffic flow. However, the share taken by a transit system from the network travel demand largely depends on the quality of service it offers to users. Meanwhile, improvement in the service quality of the transit system may entail a significant rise in the operating cost of a system. Therefore, it is of great importance to strike a balance between the operating cost and service quality of a transit system by minimizing the social, economic, and environmental costs of the system in the design of a sustainable transit system.

The network-wide transit design problem is introduced in Holroyd (1967) that proposes an analytical model for optimal headway and line spacing in a grid transit network by minimizing the summation of operating cost and user travel time. The classic analytical model of a transit system has also been extended for different types of network structures in the literature. Byrne (1975) designs a radial system with optimal line spacing and headway using a polar coordinate system. The radial design of the transit network has been further elaborated by adding rings to the radial system in Vaughan (1986). Byrne and Vuchic (1972) and Wirasinghe et al. (1977) also analytically model the operation of a system of corridors with feeder-bus service in grid and radial transit networks. The optimal design of a hub-and-spoke system with minimal operat-

* Corresponding author.
E-mail address: hg55@cornell.edu (H.O. Gao).

http://dx.doi.org/10.1016/j.tre.2017.03.006
1366-5545/© 2017 Published by Elsevier Ltd.

ing cost is also discussed in Newell (1979). The optimal network structure, operating characteristics, and fare pattern of transit systems have been further elaborated in recent studies (Yu et al., 2015; Bagherian et al., 2016; Chebbi and Chaouachi, 2016; Huang et al., 2016).

The stochasticity in the demand and supply sides of a transit system has been studied in the literature as well (Chien et al., 2001; Daganzo and Smilowitz, 2004). To optimize the infrastructure of a transit system with elastic demand, Daganzo (2012) maximizes the social welfare of the system with multiple classes of users. Promoting the efficiency of a transit system by improving the mobility of the feeder service and land-use pattern around the trunk sections has been studied in Sivakumaran et al. (2014). Amirgholy and Gonzales (2016) also optimizes the operation of demand responsive transit systems with time-dependent demand by minimizing the summation of the generalized cost of users and the operating cost of the service for agency costs.

The role of the transit system (as an alternative to automobile mode) in reducing congestion in urban regions has been widely investigated in the literature. Daganzo (2010a) proposes a design for a transit service with a competitive level of accessibility to the automobile at a reasonable cost by optimizing the headway and structure of a hybrid hub-and-spoke system with two levels of coverage in the central and peripheral districts of a grid network with uniform demand distribution. The concept of the hybrid system has been generalized in Estrada et al. (2011) to design a high-performance transit system for the city of Barcelona. Nourbakhsh and Ouyang (2012) also proposes an alternative flexible-route transit system with hybrid structure for low demand areas. The hybrid model has been recently extended in Badia et al. (2014) for the ring-and-radial network structure using a polar coordinate system. Chen et al. (2015a) also adopts the hybrid concept to develop two continuum approximation models for grid and ring-and-radial structures in the transit network that can be solved numerically for real size problems. The long-term effects of designing a ring-and-radial network structure for light rail transit in urban regions have been studied in Saidi et al. (2016).

Recent studies have taken steps towards accounting for the interaction between automobiles and the transit system in bimodal urban networks. Guler and Cassidy (2012) shows that the delay that users experience in arterial networks can be significantly reduced when the available capacity of the network is optimally shared between the automobile and transit modes. The strategy of dedicating separate lanes to the transit system to improve user mobility in the network is also investigated in Gonzales and Daganzo (2012) by extending Vickrey's (1969) congestion theory to model the user equilibrium condition between the automobile and transit modes in a bottleneck over the morning peak period. To optimize the bimodal system, a dynamic pricing strategy is proposed for both automobile and transit modes that makes it possible to eliminate the delay in the bottleneck. In another paper, these authors employ the bottleneck model to compare the equilibrium condition between the automobile and transit modes over the morning and evening peak periods (Gonzales and Daganzo, 2013). Building upon these papers, Daganzo (2013) proposes a day-long pricing strategy for both automobile and transit modes that can optimize the system by minimizing the generalized cost of commuters. Tian et al. (2013) and Xiao et al. (2015) also adopt time-dependent credit schemes for managing congestion in a bi-modal system with heterogeneous user preferences.

The idea of using the macroscopic fundamental diagram (MFD) for optimal allocation of network capacity to different modes of transport is shaped in Gonzales et al. (2010) and developed in Gonzales and Daganzo (2012). In large urban regions (neighborhoods), the macroscopic relationship between the aggregated traffic variables of the network, i.e. speed, flow, and density, can be represented by the MFD of the region. The MFD model is analytically developed in Daganzo (2007) and Daganzo and Geroliminis (2008), and empirically measured for the city of Yokohama in Geroliminis and Daganzo (2008). Properties of the MFD largely depend on characteristics of the network and the distribution of demand in the region. Recent research on the macroscopic relationship between the traffic variables of the network shows that the scatter and the shape of the MFD can be significantly affected by the spatial distribution of congestion in the network. In fact, heterogeneity in the spatial distribution of congestion can slow down traffic flow in the network, and such effects can even be magnified by a rise in the vehicle density of the system (Mazloumian et al., 2010; Geroliminis and Sun, 2011; Gayah and Daganzo, 2011; Mahmassani et al., 2013). In real size urban networks where the distribution of congestion is naturally not homogenous, the region can be partitioned into a reasonable number of homogenous clusters in order to derive well-defined (low-scattered) MFDs for the subregions of the network (Ji and Geroliminis, 2012). The effects of heterogeneity in the spatial distribution of the congestion in the network on the scatter, shape, and accuracy of the MFD have also been examined using region-based and subregion-based MFD models (Ramezani et al., 2015; Xie et al., 2016). The macroscopic relationship between the traffic variables in real urban networks can be approximated by estimating a limited number of observable parameters (Laval and Castrillón, 2015). Alternatively, the MFD of the region and the average vehicle density of the network can be estimated using low penetrated probe data from the network (Gayah and Dixit, 2013; Ji et al., 2014; Leclercq et al., 2014; Nagle and Gayah, 2014; Du et al., 2015a). This information can be crucial for designing real-time traffic control strategies in urban networks (Keyvan-Ekbatani et al., 2012; Haddad and Geroliminis, 2012; Geroliminis et al., 2013; Haddad et al., 2013; Gayah et al., 2014; Haddad and Shraiber, 2014; Du et al., 2015b; Ramezani et al., 2015). In addition to applications for the design of real-time control strategies, the MFD model can be applied in evaluating the performance of the transportation system for planning and design purposes.

In this paper, we make a use of the MFD model to capture the effect of operating a transit system on the average speed and outflow of the urban network. The objective of the research is to design a sustainable transit system with a quality of service competitive with the automobile by minimizing the social, economic, and environmental cost of the transportation system. To cope with the complexity of the problem, we put the focus of the paper on aggregating the behavior of users and the performance of the system by ignoring the fine details of the problem to develop a continuum approximation model that