Networks and Spatial Economics, 1: 2001 233–265

Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future

SRINIVAS PEETA
School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA

ATHANASIOS K. ZILIASKOPOULOS
Department of Civil Engineering, Northwestern University, Evanston, IL 60208, USA

Abstract

Dynamic Traffic Assignment (DTA) has evolved substantially since the pioneering work of Merchant and Nemhauser. Numerous formulations and solutions approaches have been introduced ranging from mathematical programming, to variational inequality, optimal control, and simulation-based. The aim of this special issue is to document the main existing DTA approaches for future reference. This opening paper will summarize the current understanding of DTA, review the existing literature, make the connection to the approaches presented in this special issue, and attempt to hypothesize about the future.

Keywords: Dynamic traffic assignment, real-time deployment, planning

1. Introduction

Dynamic Traffic Assignment (DTA), though still in a state of flux, has evolved substantially since the seminal work of Merchant and Nemhauser (1978a, 1978b). There is currently heightened interest in DTA, particularly in the development of approaches that can be deployed for large-scale real-time and planning applications. In addition, researchers have become increasingly aware that the theory of DTA is still relatively undeveloped, which necessitates new approaches that account for challenges from the application domains as well as for the fundamental questions related to tractability and realism. Agencies and practitioners are also increasingly realizing the potential of DTA to address longstanding problems with the unrealistic assumptions of existing static planning methods, as well as the potential of DTA to both evaluate Intelligent Transportation Systems (ITS) technologies, as well as be the main operational engine for deployment.

DTA refers to a broad spectrum of problems, each corresponding to different sets of decision variables and underlying behavioral and system assumptions, and possessing varying data requirements and capabilities in terms of representing the traffic system or control actions. One common feature of these models is that they depart from the standard static assignment assumptions to deal with time-varying flows. Another feature shared by these models is that none presently provides a universal solution for general networks. Perhaps the one aspect that fosters unanimity among researchers is that the general DTA problem is inherently characterized by ill-behaved system properties that are imposed by the need to adequately represent traffic realism and human behavior. This is further exacerbated by the time-dependency and randomness in system inputs. A fundamental consequence of this reality is that a theoretical guarantee of properties such as existence, uniqueness, and stability can be tenable only through compromises in depicting traffic theoretic phenomena and potentially restrictive assumptions on driver behavior. Viewed from the complementary perspective, an ability to adequately capture traffic dynamics and driver behavioral tendencies precludes the guarantee of the standard mathematical properties. This inherent complexity of DTA has spawned a clear dichotomy of approaches that range from the the analytical to the simulation-based. A fundamental practical consequence of the theoretical intractability is the focus of DTA researchers on developing deployable solution procedures that seek close-to-optimal solutions with a clear understanding that claims of uniqueness and/or global optimality are neither essential nor particularly meaningful in the real-world. This has manifested as the development of mostly heuristic implementation procedures that seek effectiveness, robustness, and deployment efficiency. Another outcome is the notion of commensurability of features among different approaches so that trade-offs among desirable features allow different degrees of responsiveness to different DTA problems given the broad scope of objectives and functional needs addressed under the general umbrella of DTA.

A reassuring practical aspect of the general DTA problem is that its mathematical intractability is not an all-encompassing barrier to the real-world utility of the associated solution approaches. Substantial research over the past decade suggests that effective and efficient solutions can be obtained for several realistic scenarios characterized by inconsequential mild violations and/or infrequent physical manifestations of the mathematically intractable aspects. There are several actual situations where ill-behaved problem characteristics do not arise, and even if they do, it is with minimal practical temporal and spatial consequences vis-à-vis the modeling assumptions. Hence, different approaches may address different functional needs with different degrees of robustness, precluding the notion of sweeping generalizations on mathematical intractability by focusing on pathological scenarios, especially if they are rare in practice. The consensus is that a deployable DTA approach should adequately represent traffic realism in the context of the problem objective.

This paper is organized as follows. The next section reviews past efforts in DTA by classifying the major approaches and putting them in perspective with the current understanding of DTA. Section 3 discusses some future research issues and directions, currently in their incipient stages, that are motivated by application domains or problem characteristics. Section 4 presents some concluding comments.

2. Review of the Past

This section reviews past DTA literature by classifying the various approaches into four broad methodological groups: mathematical programming, optimal control, variational inequality, and simulation-based. Of these, the first three groups are further labeled as analytical approaches. As in the static case (for comprehensive coverage, see Sheffi (1985) and Patricksson (1994)), most formulations tend to focus on the user equilibrium (UE) and system optimal (SO) objectives, or some variants of them. While a vast body of literature has been developed in this area over the past two decades given the wide range of problems addressed under the DTA label, this section focuses on some of the efforts that highlight the basic problem dimensions. Peeta (1994) provides a review of the literature on mathematical programming and optimal control based DTA formulations, with particular focus on real-time information provision. Ran and Boyce (1994) discuss various optimal control models. Recent efforts that focus on variational inequality based formulations are discussed in Ran and Boyce (1996a) and Chen (1999). Mahmassani et al. (1998a) review various simulation-based DTA models.

2.1. Mathematical programming formulations

Mathematical programming DTA models formulate the problem in a discretized time-setting. Merchant and Nemhauser (1978a, 1978b) represent the first attempt to formulate the DTA problem as a mathematical program. The formulation was limited to the deterministic, fixed-demand, single-destination, single-commodity, SO case. The model (hereafter referred to as the M–N model) uses a link exit function to propagate traffic and a static link performance function to represent the travel cost as a function of link volume. It results in a flow-based, discrete time, non-convex non-linear programming formulation. The model is shown to provide a proper generalization of the conventional static SO assignment problem, and the global solution is obtained by solving a piecewise linear version of the model. Ho (1980) shows that such a global optimum can be obtained by solving a sequence of at most N+1 linear programs, where N is the number of periods.

Carey (1986) proved that the M–N model satisfies the linear independence constraint qualification since the proposed exit function is continuously differentiable, validating the optimality analysis in Merchant–Nemhauser (1978a, 1978b). Carey (1987) reformulates the Merchant–Nemhauser problem as a well-behaved convex nonlinear program through the manipulation of exit functions, which offers mathematical and algorithmic advantages over the original formulation. The basic formulation is very similar to the M–N model. For the formulation to be convex, an exit function is used to bound the link outflow rather than being the outflow itself as in the M–N model. Since it is convex, standard mathematical programming software can be used to solve the problem. However, the paper indicates that special structures such as the staircase structure in the constraints should be explored to develop algorithms that are more efficient. Extensions of the basic model are introduced to handle multiple destinations and commodities. However, the resulting formulations remain problematic because of non-convexity issues arising from a "first-in, first-out" (FIFO) requirement. These difficulties appear to be inherent in all mathematical programming approaches to the time-dependent assignment problem, for both the UE and SO cases. Multiple destinations require the models to explicitly satisfy a first-in, first-out requirement that is essential from a traffic realism viewpoint. FIFO has two dimensions, physical and algorithmic. Since individual vehicles may at times violate FIFO in actual traffic networks (overtaking), it is satisfied only in an average sense from a traffic propagation standpoint.

The problematic FIFO violation implied here is algorithm-induced, one where some commodity (for example, traffic between one origin-destination (O-D) pair) physically jumps over another to reduce system costs. This is inconsistent with traffic realism. The FIFO requirement is easily satisfied in single-destination formulations and for certain special network structures. In general networks, this requirement would introduce additional constraints that yield a non-convex constraint set, destroying many of the nice properties of the formulation, and severely increasing the computational requirements of any eventual solution algorithm (Carey, 1992).

Another well-known phenomenon related to the SO mathematical programming models that impinges on traffic realism is the "holding-back" of vehicles on links. It arises when link exit constraints are satisfied as strict inequalities. In a traffic network, it may often be advantageous to favor certain traffic streams or movements over others to minimize system-wide travel delays (e.g. holding back traffic at the minor approach of an intersection in favor of the major approach). Unless otherwise specified, the solution of a SO assignment formulation may entail holding of traffic on one path in favor of traffic on other paths for some significant amount of time at points where the paths overlap or intersect. In other words, vehicles may be artificially delayed on a link for a time that exceeds what may be considered "fair" or "reasonable". Such a solution is probably not acceptable socially nor realistic operationally. From a modeling standpoint this would entail explicit constraints to preclude holding-back of traffic. Carey and Subrahmanian (2000) illustrate some aspects of the FIFO and holding-back issues.

Janson (1991a, 1991b) represents one of the earliest attempts to model the UE DTA problem as a mathematical program. One feature of his approach is that it seeks an equilibrium described in terms of experienced path travel times, instead of the instantaneous travel times assumed in several prior studies. Non-linear mixed integer constraints are proposed in the formulation to ensure temporal continuity of O-D flows, though they may be violated in the solution procedure specified which is a straightforward extension of the well-known incremental assignment heuristic for static formulations. The properties of this procedure are not sufficiently well-established, and it may lead to possible unrealistic traffic behavior. In addition, it relies on static link performance functions for traffic modeling.

Birge and Ho (1993) extend the M–N problem to the stochastic case by relaxing the assumption that O-D desires are known for the entire planning horizon. They develop a multistage stochastic mathematical programming formulation that is non-linear and non-convex as in the deterministic case. The model assumes a finite number of scenarios of random variable realizations, where a scenario is defined as a possible combination of past O-D desires in every period. However, the formulation assumes current assignment decisions to be independent of future O-D desires. The resulting model determines the O-D assignment pattern that minimizes the expected piecewise linear convex costs over several time periods through a sequence of linear optimizations.

Ziliaskopoulos (2000) introduces a linear programming formulation for the single destination SO DTA problem based on the cell transmission model (Daganzo, 1994) for traffic propagation. It circumvents the need for link performance functions as the flow propagates according to the cell transmission model, thereby being more sensitive to traffic realities. While not an operational model for real-world applications, it provides some insights on the DTA problem properties. Carey and Subrahmanian (2000) also propose a linear single destination SO DTA formulation, again with the intent to derive insights on the model properties. As stated earlier, they study the FIFO and holding-back issues in the context of mathematical programming formulations.

The substantial research in mathematical programming based DTA approaches highlights its current limitations for developing deployable models for general networks. A persistent issue is the need to trade-off mathematical tractability with traffic realism. This is primarily manifest in terms of the inadequate modeling ability and/or inconvenient constraints arising vis-à-vis the representation of traffic dynamics. One such aspect is the non-convex constraints needed to explicitly address the FIFO property. While there is a whole body of optimization literature that deals with non-convex programming, the non-convexity in the DTA context leads to a loss of analytical and computational tractability for deployment in general networks. In addition, in general, mathematical programming DTA formulations tend to have difficulties related to: (i) the use of link performance and/or link exit functions, (ii) holding-back of traffic, (iii) efficient solutions for real-time deployment in large-scale traffic networks, and (iv) a clear understanding of solution properties for realistic problem scenarios.

2.2. Optimal control formulations

In constrained optimal control theory DTA formulations, the O-D trip rates are assumed as known continuous functions of time, and the link flows are sought as continuous functions of time. The constraints are analogous to those for the mathematical programming formulations, but are defined in a continuous-time setting. This results in a continuous-time optimal control formulation, rather than a discrete-time mathematical program.

Friesz et al. (1989) discuss link-based optimal control formulations for both the SO and UE objectives for the single destination case. The models assume that adjustments from one system state to another may occur concurrently as the network conditions change; that is, the routing decisions are made based on current network conditions, but can be continuously modified as conditions change. The SO model is a temporal extension of the static SO model, and proves that at the optimal solution the instantaneous flow marginal costs on the used paths for an O-D pair are identical and less than or equal to the ones on the unused paths. It should be noted that a discrete-time version of this SO formulation would be identical to the M–N model. They also propose a time-dependent generalization of Beckmann's equivalent optimization problem (Beckmann et al., 1956) for static UE traffic assignment in the form of an optimal control problem by the equilibration of instantaneous user path costs. The models use exit functions to propagate traffic, and link performance functions to determine travel costs. Key issues include the lack of meaningful link performance and exit functions, and the difficulties in developing efficient solution algorithms. Also, the formulations treat the link inflow as a control variable, though the outflow is a function. This is problematic as the non-linearity of the exit function makes it difficult to establish the generalization of Wardrop's first principle (Wardrop, 1952) to the time-dependent case for a network with multiple origins and destinations. Wie (1990) extends the UE model to include elastic time-varying travel demand, which leads to the implicit consideration of departure time choices. Wie also enumerates several limitations of this approach.

Ran and Shimazaki (1989a) use the optimal control approach to develop a link-based SO model for an urban transportation network with multiple origins and destinations. They also define exit flow to be a function, precluding the generalization of optimality conditions. They use linear exit functions and quadratic link performance functions so as to reduce the computational burden for a time-space decomposition solution procedure that can only handle very small network problems. In addition to the unrealistic modeling of congestion, they do not consider the FIFO issue that arises for multiple destinations. Ran and Shimazaki (1989b) present an optimal control theory based instantaneous UE DTA model. The flows exiting links are treated as a set of control variables rather than functions to circumvent the generalization issue. No efficient algorithms are available to solve these formulations.

Ran et al. (1993) use the optimal control approach to obtain a convex model for the instantaneous UE DTA problem by defining link inflows and outflows to be control variables. They recognize the inability of the usual cost functions to account for dynamic queuing and congestion costs, and propose splitting the link travel cost into moving and queuing components. However, these functions are assumed to be non-negative, increasing and differentiable, and hence may not reflect traffic realism. Also, no specific instances of the functions are provided or tested. Boyce et al. (1995) discuss a methodology to solve the discretized version of the problem using the Frank–Wolfe algorithm and an expanded time-space network representation. However, they do not implement the procedure or illustrate it through examples. Also, the use of static link performance functions is a limitation of this model.

Though an attractive approach for describing dynamic systems, an optimal control type of DTA formulation suffers from many limitations, such as the lack of explicit constraints to ensure FIFO and preclude holding of vehicles at nodes, the inadequate and possibly unrealistic modeling of traffic congestion, and more crucially, the lack of a solution procedure for general networks. In addition, the substantive interpretation of the UE formulations based on instantaneous travel times requires rather strong and possibly unrealistic assumptions about the nature of the behavioral processes underlying the particular equilibrium conditions postulated by the researchers. Because of the limitations of optimal control theory in the DTA context and the advantages offered by variational inequality (VI), analytical DTA models have in recent years migrated towards the VI approach, discussed hereafter.

2.3. Variational inequality formulations

Variational inequality provides a general formulation platform for several classes of problems in the DTA context such as optimization, fixed point, and complementarity. It fosters a unified mechanism to address equilibrium and equivalent optimization problems. Also, its mathematical properties such as uniqueness can be illustrated in a simple manner. Dafermos (1980) introduced the VI approach in the static traffic equilibrium context.

Nagurney (1998) provides a comprehensive summary of VI and addresses various equilibrium problems. VI circumvents analytical tractability issues arising in constrained optimization formulations due to asymmetric link interactions. In that sense, it can handle more realistic traffic scenarios. Also, extensions and sensitivity analyses can be conveniently performed. While it is more general, and better equipped than the other analytical approaches to address several aspects of DTA problems, the broader limitations associated with analytical models, discussed earlier, remain.

Friesz et al. (1993) formulate a continuous time VI model to solve for the departure time/route choice by equilibrating the experienced travel times. The model uses link performance functions, penalty functions for early/late arrivals, travel demands, desired arrival times, and all possible paths between origins and destinations. The path cost is a combination of the travel cost determined by the link performance function and the penalty for early/late arrival caused by traveling along the path. The formulation incorporates relatively more realism in terms of traveler behavior, but has unresolved issues. For instance, there is no proof of solution existence or uniqueness. Also, since the formulation is a continuous-time infinite-dimensional VI problem, its solution requires solving a complex system of simultaneous integral equations and there is no efficient algorithm for this purpose.

Wie et al. (1995) introduce a discretized VI formulation for the simultaneous route-departure equilibrium problem to enable computational tractability, and propose a heuristic algorithm to approximately solve it. They show solution existence under certain regularity conditions, and use exit flow functions instead of exit time functions used in the Friesz et al. (1993) formulation. The use of exit flow functions raises the usual traffic flow realism issues. Since the formulation is path-based, and complete path enumeration is computationally burdensome, there is a need for an efficient method to identify a subset of relevant paths. However, that may entail some limiting assumptions on the traffic modeling (Ran and Boyce, 1996b).

To circumvent the problems with path-based VI models, Ran and Boyce (1996b) propose a link-based discretized VI formulation with fixed departure times. Akin to Friesz et al. (1993), they also equilibrate the experienced travel times. They include a queuing delay component to partly alleviate the traffic realism issues arising in the context of analytical models. However, the capacity and oversaturation constraints significantly increase the computational burden, leading to computational feasibility issues for realistic networks. Also, path-based formulations are inherently convenient in the context of route guidance. Ran et al. (1996) extend the proposed model to a link-based VI model for the simultaneous departure time/route choice problem.

Chen and Hsueh (1998) also propose a link-based VI formulation for the UE DTA problem. They show that without loss of generality, travel time on a link can be represented as a function of link inflow only (instead of function of link inflow, exit flow, and number of vehicles on the link). Through this simplification, they illustrate that the Jacobian matrix of the travel time function is asymmetric, and that any dynamic travel choice problem with such travel time function characteristics does not have an equivalent optimization program. A solution algorithm based on nested diagonalization procedure is proposed; however, it is still prohibitively expensive to implement on real networks.

The VI approach is more general than the other analytical approaches discussed earlier, providing greater analytical flexibility and convenience in addressing various DTA problems. It has been used to illustrate, with relative ease, the notion of experienced travel times for the so-called "simultaneous" and "ideal" UE DTA problems. Also, it highlights the inability of the mathematical programming approach in addressing scenarios with asymmetric Jacobian matrices for the travel cost functions. However, VI approaches are more computationally intensive than optimization models, raising issues of computational tractability for real-time deployment. These issues are further magnified for path-based VI formulations requiring complete path enumeration. Also, despite capabilities to better represent link interactions, traffic realism issues arising in the context analytical models persist.

2.4. Simulation-based models

Simulation-based DTA models use a traffic simulator to replicate the complex traffic flow dynamics, critical for developing meaningful operational strategies for real-time deployment. The terminology "simulation-based models" may be a misnomer. This is because the mathematical abstraction of the problem is a typical analytical formulation, mostly of the mathematical programming variety in the current literature. However, the critical constraints that describe the traffic flow propagation and the spatio-temporal interactions, such as the link-path incidence relationships, flow conservation, and vehicular movements, are addressed through simulation instead of analytical evaluation while solving the problem. This is because analytical representations of traffic flow that adequately replicate traffic theoretic relationships and yield well-behaved mathematical formulations are currently unavailable. Hence, the term "simulation-based" primarily connotes the solution methodology rather than the problem formulation. A key issue with simulation-based models is that theoretical insights cannot be analytically derived as the complex traffic interactions are modeled using simulation. On the other hand, due to the inherently ill-behaved nature of the DTA problem, notions of convergence and uniqueness of the associated solution may not be particularly meaningful from a practical standpoint. In addition, due to their better fidelity vis-à-vis realistic traffic modeling, simulation-based models have gained greater acceptability in the context of real-world deployment (Mahmassani et al., 1998b; Ben-Akiva et al., 1998).

In addition to the use of a simulator in a descriptive mode to determine the traffic flow propagation, most existing simulation-based models also use it as part of the search process to determine the optimal solution. Labeled the predictive-iterative mode, the simulator is used in each iteration to project the future traffic conditions as part of the direction-finding mechanism for the search process. Given the substantial computational burden associated with the use of a simulator, the choice of granularity (macroscopic, microscopic or mesoscopic) has significant implications for the real-time computational tractability of simulation-based models.

Mahmassani and Peeta (1992, 1993, 1995), and Peeta and Mahmassani (1995a), develop DTA models that use a mesoscopic traffic simulator, DYNASMART

(Jayakrishnan et al., 1994), as part of an iterative algorithm to solve for the SO and UE solutions for O-D demand with fixed departure times. Labeled as deterministic DTA models in the context of the assumptions on information availability for Advanced Traveler Information Systems (ATIS) operations, the problem assumes complete *a priori* knowledge of O-D demands for the entire planning horizon of interest. Both descriptive and normative objectives are considered. Also, akin to the analytical formulations, the models assume identical users, manifest as a single class of users in terms of information availability, information supply strategy, and driver response to the information provided. The adopted simulation logic combines a microscopic level of representation of individual tripmakers and drivers, with a macroscopic description of some of the interactions taking place in the traffic stream. This enables an acceptable solution accuracy level at a fraction of the computational cost required for a microscopic representation of traffic maneuvers. The use of a traffic simulator circumvents the traffic realism issues of analytical formulations and provides a solution procedure for general networks. Mahmassani et al. (1993) extend the single user class models to the more realistic multiple user classes scenarios where multiple user classes in terms of information availability, information supply strategy, and driver response to the information provided, are assumed. However, as with the analytical formulations, they are inadequate operationally for providing optimal real-time path information and/or instructions to network users under ATIS in response to unpredicted variations in network conditions. In addition to the conceptual and algorithmic aspects of the models, the real-time and large-scale nature of these problems make computational issues associated with the implementation of the models an integral part of the problem. In particular, the development of algorithmic procedures must reflect the issue of computational efficiency. As these deterministic DTA solution methodologies incorporate a simulator as part of the iterative search process, they are not feasible for real-time deployment without further modifications vis-à-vis implementation.

Ghali and Smith (1992a) propose a single user class formulation for the deterministic SO DTA problem in which congestion arises exclusively at specified bottlenecks modeled as deterministic queues. A simulation-based solution procedure is proposed, whereby vehicles are routed individually on paths determined using link marginal travel costs. Although the approach does not ensure system optimality, and has limitations due to certain assumptions on queuing, it addresses several traffic modeling issues that limit the realism and validity of analytical formulations. Ghali and Smith (1992b) discuss different levels of computing approximate marginal travel times, in addition to the computation of the global marginal travel times. However, their methodology for the computation of marginal travel times involves the brute force simulation of alternative scenarios, which is neither feasible for real-time application in the ATIS context, nor efficient in large-scale real-world networks. Ghali and Smith (1992c), and Smith (1994), address the SO and UE DTA problems and implement their solution procedures using the CONTRAM (Leonard et al., 1989; Taylor, 1990, 1996) simulation model. Ghali and Smith (1993) construct examples to illustrate the possibility of non-convexity and non-differentiability in DTA problems, which preclude the guarantee of convergence and of obtaining globally optimal solutions for general networks.

With the objective of generating a real-time deployment capability, Peeta and Mahmassani (1995b) develop rolling horizon DTA models to explicitly incorporate real-time variations in network conditions and foster computational efficiency to enable real-time tractability. The rolling horizon approach provides a practical method for addressing problems which ideally require future demand information for the entire planning horizon, a defining characteristic of DTA problems (Peeta, 1994). Its advantage is the ability to use currently available information, and near-term forecasts (of the order of 5 to 15 minutes) with some degree of reliability, to solve a problem in quasi real-time while preserving the effectiveness of the computational procedure in determining "good" control strategies. From an operational perspective, these information requirements are more realistic compared to the perfect knowledge assumptions of deterministic models. Since it is stage-based, the rolling horizon approach ensures that unpredicted variations in on-line traffic conditions can be adequately accounted for in subsequent stages. However, if the actual O-D desires in a stage are significantly different from the forecasts, the solution is sub-optimal. Also, despite being a stage-based approach and more efficient than deterministic DTA solution procedures, it can be computationally intensive in a centralized architecture.

Ben-Akiva et al. (1997a, 1997b) propose DynaMIT as a dynamic traffic assignment system to estimate and predict in real-time current and future traffic conditions. It consists of a demand and a supply simulator that interact to generate UE route guidance under the rolling horizon framework. No underlying formulation is proposed. The demand simulator estimates and predicts O-D demand using a Kalman Filtering methodology. It considers both historical information and the driver response to information. The supply simulator is used to determine the flow pattern based on the demand. It is a mesoscopic traffic simulator, where vehicles are moved in packets and links are divided into segments that include a moving part and a queuing part to model traffic flow.

Ziliaskopoulos and Waller (2000) introduce an internet-based GIS system that integrates data and models into one framework. The simulation-based DTA model in this system uses RouteSim, which is a mesoscopic model based on cell transmission (Daganzo, 1994) for traffic propagation, as the traffic simulator. The model captures many realities on the network, such as traffic signals, by using time-dependent cell capacities and saturation flow rates.

Simulation-based DTA models address, with relative ease, several modeling issues that are troublesome in analytical formulations. The use of a simulation model that incorporates traffic theoretic relationships to model traffic flow circumvents the limitations of analytical link performance and exit functions in replicating dynamic traffic phenomena and obtains flows consistent with those relationships. A traffic simulator also captures the complex vehicle interactions, thereby evaluating the non-linear objective function satisfactorily compared to idealized cost functions. In addition, a simulation model can implicitly satisfy the FIFO constraint and circumvent the problem due to holding-back of vehicles. All these factors, and the ability to keep track of paths of individual vehicles in the ATIS context, represent the advantages of simulation-based approaches in the DTA context. Hence, while analytical models have focused primarily on deriving theoretical insights, simulation-based models have concentrated on enabling practical deployment for realistic networks. Therefore, issues such as multiple user classes, information availability, driver response characteristics, deployable solution procedures, and deployment aspects such as consistency, robustness, calibration, and computational efficiency, have been the focus of simulation-based models. The deployment aspects are discussed further in Section 3. Simulation is an especially convenient option to realistically capture the complex interactions among multiple user classes that arise in actual traffic networks. Given the traffic modeling issues that arise for analytical models in the single user class context itself, introducing the notion of multiple user classes substantially adds to their intractability.

The key limitation of simulation-based DTA models, in general, is the inability to derive the associated mathematical properties. As discussed earlier, this is not a significant issue for general traffic networks. This is because the DTA problem for general traffic networks is inherently ill-behaved and rather intractable, restricting the ability of analytical models as well to analyzing problems with simplified assumptions that reduce realism. However, an ability to analyze the system properties even under simplified assumptions can be insightful in generating future directions to address the problems. Another limitation of using a simulator is in the deployment context. The computational burden associated with the use of a simulator as part of an iterative mechanism to project the future can be operationally restrictive. Hence, many recent simulation-based deployment frameworks, discussed further in Section 3, trade-off solution accuracy with computational efficiency.

3. Current and Future Directions: Challenges and Opportunities

The review of past DTA efforts provides pointers to the challenges and opportunities in the general DTA domain, both in terms of research directions and practical applications. DTA appeals to a broad audience of researchers and practitioners, and promises the potential for a wide range of applications. Recent trends suggest that it can impact a broad range of problems in transportation operations, planning, and engineering. Each, in turn, introduces a unique set of challenges as well as opportunities for DTA research. This section briefly discusses some of the current and/or future research directions and potential challenges. It also identifies some of the opportunities they present for advancing the DTA field. The challenges and opportunities are grouped into three categories: two in the application domain (real-time deployment and planning) and one related to the fundamental issues.

3.1. Real-time deployment

Till recently, the primary focus of research efforts in the DTA arena has been the development of methodological and algorithmic constructs that address the fundamental concepts, architecture, and logic, representing mostly an off-line developmental perspective. Current efforts, while continuing to address the fundamental issues, are increasingly focusing on real-time operational issues, to develop and refine capabilities for the efficient and effective real-time deployment of these methodologies and algorithms. These issues are: (i) computational tractability, (ii) robustness of the solution methodologies and stability of the associated solutions given the inherent system randomness, (iii) fault tolerance and system reliability, (iv) operational consistency and model calibration/validation, and (v) demand estimation and prediction. The associated objective is a practical one: how to implement these algorithms and strategies to achieve a feasible degree of practical effectiveness in large-scale traffic networks? Currently, the research on most of these issues is rather sparse.

While perfect *a priori* knowledge (on O-D demand and incidents) and system-wide coordination are highly desirable, they remain the obstacles to real-time operability in general networks because they introduce unrealistic expectations and computational intractability, respectively. This motivates the development of methodologies that are both robust and computationally efficient on-line.

3.1.1. Deployment frameworks: computational tractability The real-time deployment of DTA requires computationally tractable solution procedures for general networks. Real-time computational tractability can be addressed at different levels. At the computing hardware level, it strongly depends on the environment used (sequential, parallel, distributed, etc.) and the technological progress of computational processing and communication capabilities. At the computing software level, it depends on the efficient algorithmic coding and the overall system integration architecture used, for example the CORBA based implementation of a DTA system (Hawas et al., 1997). At the level of algorithmic logic, it depends on the control architecture (e.g., centralized, decentralized, or hybrid). As stated earlier, here, the deployment frameworks related to algorithmic logic are addressed. Most of these frameworks either circumvent the use of a simulator on-line or use it in a computationally efficient mode that avoids a predictive-iterative process.

To address the real-time computational burden, Hawas and Mahmassani (1995) propose a non-cooperative reactive decentralized architecture where spatially distributed controllers make routing decisions independent of each other. Local control rules using artificial intelligence techniques and currently available partial traffic information in the local area are used to determine user routes within the control territory. An attractive feature of the approach is the flexibility in defining the territorial size of each controller based on its processing capabilities, thereby circumventing issues of computational burden. Another significant aspect is that the approach does not require any O-D demand predictions unlike centralized frameworks. However, since controllers act independently based only on local rules, there is no coordination between them. If the proportion of inter-territory vehicles is large, the solution under this scheme deviates substantially from the optimal solution (Hawas and Mahmassani, 1995).

Hawas (1995) extends the above approach to develop a cooperative scheme that enables exchanging non-local information between neighboring controllers. While the decentralized architecture is more robust under incident situations because the local rule heuristics are more responsive to current network conditions, there can be substantial degradation in performance under non-incident conditions compared to the benchmark centralized deterministic DTA solution (Hawas and Mahmassani, 1997). Since the approach is reactive, it does not exploit available historical data, especially on time-dependent O-D demand and incidents.

Pavlis and Papageorgiou (1999) develop a reactive decentralized feedback control strategy based DTA model for meshed networks. The essential components of the strategy are simple decentralized control laws of the bang-bang, P, or PI type that could be designed based on trial-and-error. It reacts solely to the real-time measurements. The controllers independently calculate splitting rates for each path based on the traffic information on alternative paths. The control strategy is to equalize the instantaneous path travel times of alternative paths for each O-D pair. Since it is decentralized and circumvents forecasts of future traffic conditions, significant computational efficiency is achieved to promote real-time tractability. However, the associated framework is limited to specific network topologies and is opaque to systematic correctable errors in the prediction process.

Peeta and Zhou (1999a, 1999b) propose a hybrid predictive-reactive framework that combines off-line and real-time strategies to solve the real-time DTA problem. It ensures that the computationally intensive components are executed off-line while circumventing the need for very accurate real-time O-D demand forecast models. The primary concept of this approach involves the development of an *a priori* solution using an iterative centralized deterministic DTA-based mechanism off-line that exploits the available historical data to develop a robust initial solution for real-time use. The historical O-D demand and incident distributions are used to incorporate randomness. The *a priori* solution is based on the computationally intensive Monte Carlo simulation of various realizations. The initial solution is updated on-line based on the unfolding conditions through efficient real-time reactive strategies obtained using the field traffic measurements. Hence, no O-D demand prediction is required in real-time.

Peeta and Yang (2000a) develop an efficient dynamical systems based search process to solve the DTA problem in real-time. It is used in conjunction with a simulator in the descriptive mode as part of a feedback control strategy that circumvents the need for real-time O-D demand predictions and uses the field traffic measurements. A truncated decision horizon, called the assignment interval, is used to implement the strategy over the planning horizon of interest, akin to a rolling horizon scheme. The control strategy is extended (Yang, 2001) to incorporate an anticipatory capability that uses historical traffic data within a hybrid framework. Here as well, the simulator is used in a descriptive mode for a truncated horizon, ensuring real-time tractability.

3.1.2. Consistency checking The operational consistency problem, representing the potential divergence of the predicted system state from the actual conditions unfolding on-line, arises because of several factors that can significantly influence the performance of dynamic traffic networks. They include: (i) incorrect prediction of the time-dependent origin-destination (O-D) trip demand, (ii) unpredicted incidents, (iii) incorrect path predictions, (iv) incorrect traffic modeling, (v) incorrect assumptions on driver behavior and/or response to information provided, (vi) incorrect assumptions on system related parameters, (vii) noise and/or sparsity in measurements, and (viii) failure of the ATIS system components. Ensuring consistency is critical to the effectiveness of any DTA procedure. Additionally, from the perspective of traffic networks with advanced information systems, not all network users are likely to be equipped with in-vehicle route guidance systems for two-way communication with traffic control centers. Hence, it is imperative to robustly predict the travel actions and decisions of unequipped users so as to provide more accurate routing information to equipped users, with the broader objective of enhancing system performance. In a control context, the consistency aspect raises issues of system observability and controllability. A related problem that needs to be addressed in an operational context is the calibration of the various static and dynamic modeling and system parameters for the specific network. This aspect is also related to the model validation issue.

Mahmassani et al. (1998a) propose an integrated real-time and off-line adjustment module for consistency using a proportional-integral-derivative (PID) feedback control strategy. They develop a framework for real-time monitoring that reacts to any observed on-line deviations in traffic conditions. The real-time corrections are not necessarily the most accurate adjustments because some information could be missing. Thus, the system is further updated off-line where full information on past conditions is used to perform optimal adjustments to address observed real-time errors. The real-time scheme is used to correct the parameters of the travel time function and the flow propagation equations using real-time data on traffic measures such as average speed, inflow, and outflow. However, the associated adjustment process does not address the underlying fundamental phenomena that cause the inconsistencies, and the procedure can also be computationally expensive on-line. The approach is extended to an online traffic monitoring system in Ziliaskopoulos et al. (1998).

Peeta and Bulusu (1999) propose a framework for ensuring operational consistency vis-à-vis real-time dynamic traffic assignment in networks with ATIS. Formulated within a stage-based rolling horizon framework, the model first solves a deterministic DTA problem to predict the system state for the near future while optimizing certain system-wide objectives for the controller, and later seeks consistency between the predicted system state and the actual conditions unfolding on-line. The approach ensures that future state predictions and path assignments are consistent with the current actual system state rather than a presumed estimate of it, which synergistically reduces the propagation of consistency errors over time. The consistency problem is formulated as a constrained least squares model. It is under-determined, rank deficient, and potentially ill-conditioned for general networks. In addition, it lacks well-behaved properties and has a fixed-point element, characteristics inherited from the DTA problem. It is solved using generalized singular value decomposition based orthogonal transformations. However, the approach is computationally intensive.

3.1.3. Robustness: incorporating randomness The issue of on-line robustness of the DTA solution procedure arises primarily because of inherent stochastic system inputs and/or factors in general traffic networks. As discussed earlier, there are several sources of randomness that significantly influence the performance of dynamic traffic networks. They include the randomness in demand, supply conditions (for example, incidents), and the fractions of users belonging to the different user classes. Solution procedures that explicitly incorporate uncertainties so as to minimize or limit the deterioration in system performance are referred to as robust. The associated problem is typically labeled as the stochastic DTA problem.

Peeta and Zhou (1999a, 1999b) explicitly model the inherent stochasticity in O-D demand and/or network supply conditions. Thereby, the O-D desires and/or non-recurrent congestion characteristics are treated as random variables with known distributions (based on a historical database updated on-line or on a day-to-day basis). The associated real-time stochastic DTA problem is solved using the hybrid solution framework discussed in Section 3.1.1, which addresses the computationally intensive procedures off-line to reduce the processing time on-line. It involves the determination of an off-line solution that serves as a "good" initial solution for real-time application, to which adjustments are made on-line as warranted by unfolding traffic conditions. A robust solution is defined here as one that has minimal deviation, in terms of the expected system-wide travel time, from the corresponding deterministic DTA solution, on average, across the range of likely O-D demand matrices. The approach models only the time-dependent O-D demand explicitly as random variables while determining the off-line solution. While incidents can also be treated as random variables in this framework, they can be more robustly managed on-line using reactive strategies (Hawas and Mahmassani, 1997; Peeta and Zhou, 1999b). Also, the marginal effect of the occurrence or non-occurrence of a predicted incident on the degradation of the robustness of the off-line solution is, in general, relatively much higher compared to that of an O-D desire, as individual incidents can impact network performance significantly.

3.1.4. Stability Stability of the DTA solution is an important operational issue for the control of dynamic traffic networks. This is because inappropriate assignment proportions may lead to increased unpredictability and volatility, and/or catastrophic consequences for the system. There are several key factors that can affect the stability of the traffic system. They include incidents, randomness in time-dependent demand, driver behavior, information provision, driver response to information, user class fractions in the traffic stream, and external traffic controls. Conceptually, the notion of stability implies that all solutions are bounded and converge to the time-dependent desirable states. The practical implication is that a stable solution minimizes or limits the deterioration of system performance.

There is a substantial body of literature that addresses stability in the context of static traffic assignment. Yang (2001) provides a detailed summary of the various approaches. Smith (1979, 1984) addresses the stability of traffic network equilibrium for the static assignment problem. A dynamical system is used to model the route choice behavior. The Lyapunov function approach is used to study the stability of the equilibrium solutions. Horowitz (1984) proposes three models for the route choice decision-making process in a two-link network based on three weighted average measures. He defines the network equilibrium to be stable if the equilibrium point is unique and the convergence of link volumes to the equilibrium state from arbitrary initial points is guaranteed. Friesz et al. (1994) apply tatonnement adjustment processes from classical microeconomic equilibrium models to predict day-to-day changes in response to changes in demand. They analyze the behavior of day-to-day trajectories from disequilibrium under complete or incomplete information provision and discuss the stability properties. Cantarella and Cascetta (1995) consider the stability of stochastic equilibrium for general networks. Regularity conditions are proposed to ensure the existence and uniqueness of a stationary probability distribution of system states. Zhang and Nagurney (1995, 1996), and Nagurney and Zhang (1996, 1997), introduce the projected dynamical system concept to study the route choice adjustment process in elastic and fixed demand networks. The stationary point of such a dynamical system coincides with the UE flow pattern. They propose two distinct approaches: the monotonicity approach to analyze global stability and the regularity approach for local stability analysis. Watling (1999) extended Horowitz's results (1984) to general networks. A dynamical adjustment process is proposed for studying the stability of the general asymmetric stochastic equilibrium assignment problem. The stability analysis concentrates on the linear approximation of the original non-linear model.

Peeta and Yang (2000a, 2000b) analyze stability in dynamic traffic networks using a dynamical systems approach and propose stable real-time deployable route guidance control strategy models for different assignment principles. They show that the Lyapunov functions for the SO and UE objectives are their corresponding objective functions under DTA. The problem is formulated as a non-linear dynamical system. Incorporation of time-dependence enables the consideration of time variance in the O-D demand and link travel times. Rather than seeking the time-dependent desirable stable states themselves, the approach moves the system towards these states based on the current or predicted future network conditions (Yang, 2001). This is because the desirable states themselves may never be reached in practice given the complexities inherent to dynamic traffic networks.

3.1.5. Demand estimation and prediction As discussed in Section 2, the assumptions on the amount of information available on the O-D trip demand lead to different DTA formulations based on information availability. In this context, the real-time estimation and prediction of dynamic O-D demand has attracted a lot of research attention. Chang and Tao (1999) classify the associated models into two broad categories: DTA based and non-DTA based. Cremer and Keller (1987), Wu and Chang (1996), Sherali et al. (1997), Ashok and Ben-Akiva (2000), and Hu et al. (2001) belong to the non-DTA based category, while Cascetta et al. (1993) and Chang and Tao (1999) are representative of DTA based O-D demand models.

Cremer and Keller (1987) link the inflows into the network and the outflows from it through split parameters. The problem is then transformed into one where an estimate for the split parameters is sought. The formulation assumes that the travel time between the various origins and destinations is shorter than the sampling interval, an unrealistic assumption for general networks. Four different approaches are proposed for estimating the time-dependent splits. The first one uses the cross-correlation matrices to estimate the splits. This method cannot guarantee that the conditions on the error distribution of the splits are satisfied. The second method includes those conditions in the problem formulation as constraints and minimizes the squared error between the estimated exit flows and the observed values. These two methods seek the time-dependent splits in one step, while the other two approaches proposed are recursive and result in shorter computational times.

Wu and Chang (1996) propose a formulation that uses traffic flows through screenlines in addition to link volumes to estimate the dynamic O-D trip matrices. They consider the time-dependent travel times between O-D pairs, thereby relaxing the assumption in Cremer and Keller (1987) on travel times. The system observability is enhanced through the consideration of these two aspects. However, as no traffic assignment component is included in the formulation, the path choices are approximated using a logit model. The use of a logit model implies independence among paths, a restrictive assumption. Sherali et al. (1997) develop solution algorithms based on dynamic link volumes. The two formulations addressed are optimization problems: one is a constrained least-squares problem, the other attempts to minimize the sum of the absolute deviation between the estimated and observed flows. The solution algorithms perform better than the Kalman filter based recursive approach in Cremer and Keller (1987). However, static travel times are used and no assignment component is included.

Cascetta et al. (1993) and Chang and Tao (1999) incorporate a network loading process into the O-D demand estimation problem. Cascetta et al. (1993) propose a dynamic network loading approach that uses the link-path incidence matrix to describe how traffic propagates. However, it uses average travel times to compute this matrix and is, hence, not a robust assignment model. Two estimators are proposed for the demand matrix. One simultaneously estimates the O-D demand matrices for all time intervals; the other estimates an O-D demand matrix for each interval sequentially, using the current one as an initial estimate for the next time interval. Chang and Tao (1999) extend the screenline approach (Wu and Chang, 1996) to decompose the network into smaller sub-networks. A two-stage O-D demand estimation framework is proposed for a large network based on the network decomposition. A non-DTA based approach is used to estimate an initial O-D demand matrix. A DTA model assigns this initial demand matrix to the network and calibrates the demand matrix using the resulting link volumes.

Ashok and Ben-Akiva (2000) present two state-space models for the real-time estimation and prediction of time-dependent O-D demand. Instead of defining the state-vector as the O-D flows themselves, the first model focuses on the deviations in O-D demand. The second model defines the state-vector as the deviations of departure rates from each origin and the shares headed to each destination. Preliminary test results indicate that such formulations make the real-time estimation process computationally tractable. The second model yields better predictions with some loss of accuracy in the filtered estimates.

Hu et al. (2001) propose an adaptive Kalman Filtering algorithm for the dynamic estimation of O-D demands using the time-varying link traffic count information. The study does not address the problem of estimation of O-D demand matrices in networks with multiple routes, but can be extended to that scenario. A key aspect of the adaptive estimator is the use of a traffic simulator to predict the time-varying travel times used to compute time-varying assignment fractions.

Estimating and predicting the dynamic O-D demand from link volumes (and possibly, other information such as the historical O-D demand matrix) is a relatively under-explored research topic, with no approach currently capable of fully dealing with general networks. Models that have a DTA component can robustly address the various time-dependent network issues; however they can entail significant computational burden.

3.1.6. Error and fault tolerance The real-time deployment of DTA under ATIS involves the transmission of field data to the Traffic Control Center (TCC) for real-time processing. To enable reliable and uninterrupted operation, these real-time systems should be fault tolerant to critical hardware failure modes such as malfunctioning detectors and failed transmission/communications links. Transportation systems possess significant random elements, such as human behavior, weather conditions, demand variability, and incidents, which complicate fault detection in vehicular traffic networks.

Anastassopoulos (2000) proposes a Fourier transform based fault tolerant framework for a deployable DTA control architecture where both data faults (erroneous data) and incidents are treated as abnormalities in the monitored network. The approach first detects an abnormality and then distinguishes data faults from incidents. Data faults are corrected using a Fourier transform based data correction heuristic. The approach uses data directly without any predictive modeling, circumventing likely modeling errors and enabling adaptability to future demand/supply changes. It also predicts near-term traffic conditions efficiently.

3.2. Planning

3.2.1. Estimating and forecasting time-dependent demand Probably the single most challenging obstacle to overcome, before deploying DTA for planning applications, is that of estimating and predicting the time-dependent origin-destination demand. Most researchers assume that such a time-dependent O-D matrix is readily available. However, the current practice for demand modeling does not provide methodologies for estimating time-dependent O-D demand. While substantial work has been done in this area, the issue is far from closed. Planning agencies currently spend significant resources to obtain a 24-hour trip table; but no resources, to the best of our knowledge, are invested for collecting time-dependent data. Demand modeling in the traditional four-step procedure uses data from travel surveys to calibrate demand models and forecast future demand. Trip generation usually takes the form of either regression or category analysis, both of which use socio-economic characteristics of the travelers in a zone as independent variables. Trip distribution mainly relies on gravity models (which can be derived through entropy maximization) to estimate the linking of trip ends predicted by the trip generation model and to form an O-D matrix. Discrete choice models are then applied to compute the mode splits that estimate the percentage of travelers using each mode. The results from these steps are the O-D matrices for each travel mode. While this process has been questioned as arbitrary and various combined models have been suggested, the four-step process tends to be the one mostly adopted by practitioners.

Surprisingly, the problem of estimating the temporal distribution of demand has been addressed by only a few studies. The approaches appearing in this special issue discuss demand estimation, though the attention is mostly directed towards real-time applications (Mahmassani, 2001; Ben-Akiva, et al. 2001). Demand estimation and prediction are typically treated outside the simulation based DTA modules aiming to provide robust estimates of historical time-dependent demand. None of the papers, however, presents compelling evidence that such an approach would work for an actual network. Mahmassani (2001) was the first to adopt the rolling-horizon approach for real-time DTA applications, which can account for demand uncertainties by updating the flow patterns produced by those demands as they become known. Ben-Akiva et al. (2001) recognize the need for historical time-dependent data, but they mostly focus on short-term predictions for real-time applications. The Kalman filtering approach, however, aims to use all existing information to produce real-time O-D estimates, and could lead to the development of meaningful O-D matrices even for planning applications. A fairly elaborate discussion of the Ben-Akiva et al. approach based on a Kalman filtering framework is included in the paper. Carey (2001) introduces an elastic demand model in a similar manner. Specifically, for each O-D pair and time period the demand is treated as a price elastic variable, where travel demands may be subject to costs or penalties incurred if traffic arrives at the destination earlier or later than some desired times.

Another promising approach is suggested, where dynamic demand data are generated from two sets of information: the first set is the static demand data and the second includes a set of behavioral rules that describe the travelers' choice of departure time. A process is proposed for calibrating the set of behavioral rules with increasingly detailed and disaggregate data. The synthetic dynamic demand module is then incorporated into an iterative process with DTA within a simulation based model. Li (2001) suggested a time-dependent trip-distribution approach that extends the static trip generation and distribution gravity models. While it stops short of proposing a framework that would produce time-dependent demand data, it makes an important contribution by identifying temporal elements in existing data sets that could certainly be used to further enhance the temporal demand estimates.

While the time-varying nature of the demand in DTA models presents insurmountable difficulties, it also presents some unprecedented opportunities, namely accounting for the departure time choice and the uncertain nature of demand. The opportunity to account for people desiring to arrive at a certain time (fixed arrival time demand) is dealt with extensively in Friesz et al. (2001). Many of the references in the paper by Friesz provide promising formulations and state necessary and sufficient conditions. Li et al. (1999) introduced a linear programming formulation that can solve reasonable size networks accounting for fixed arrival time as well as mixed demand.

Waller (2000), Waller and Ziliaskopoulos (1998) and Ziliaskopoulos and Waller (2000) proposed approaches for accounting for demand uncertainty. Note that by relaxing the determinism of the demand, one can potentially afford to make more mistakes in estimating it, since the solutions of stochastic DTA models are by definition more robust. The specifics of these approaches are briefly discussed in section 3.3.3.

In summary, the estimation of time-dependent demand is a seemingly intractable problem that will occupy researchers in the years to come. However, accounting for additional complexities such as demand uncertainty and fixed arrival time users, coupled with data from a real-time system, may provide opportunities to overcome some of the difficulties in estimating time-dependent demand.

3.2.2. Modeling multiple user classes The simulation based approaches are convenient for capturing different user classes, as well as other realities of street networks. All approaches in this special issue deal with the issue of accounting for many user classes, while recognizing that the issue is complex and many assumptions need to be made.

Mahmassani et al. (2001), Boyce et al. (2001) and Ben-Akiva et al. (2001) report on approaches that account for many types of users. However, a key question that will occupy researchers for the years to come is: what are the fundamentally different user classes? Mahmassani et al. (2001) focus on user classes in terms of information availability and supply strategies (UE, SO, following local rules, etc.). Ben-Akiva et al. (2001) classify drivers according to the information level they receive (fixed-route vehicles, unguided drivers, drivers receiving descriptive or prescriptive information). Boyce et al. (2001) deal with the problem in a more abstract fashion, by considering classes of users, where each class is associated with a disutility or generalized cost function. The ideal UE DTA route choice conditions are defined for each class on the basis of travel disutilities instead of travel time only. Implicitly, Friesz et al. (2001) and Li et al. (1999) distinguish drivers based on whether they are departure time or arrival time based. Finally, Li (2001) and Li et al. (1999) classify users for planning applications by trip purpose and propose an analytical approach to model the interaction between fixed arrival time (FAT) and fixed departure time (FDT) based users.

The range of the classes modeled indicates the complexity and significance of the problem. At the disaggregate level, every user can be viewed as representative of a unique class; however, this would be impractical to describe in a model. A classification based on the type of on-board device, information compliance level, trip purpose and vehicle type should provide reasonable realism for modeling purposes. There is a great deal that needs to be done to account for all these classes; the difficulties are mostly related to representing user behavior and equilibrium conditions.

3.2.3. Modeling many modes: person assignment Very little has been done on assigning routes to people instead of drivers. As information technology matures in the transportation arena, ultimately, mode split will be more of a daily or real-time decision, instead of a long term planning decision assumed by the third planning step. The main difficulties for implementing a simulation based person-assignment approach would be representing the many modes of transportation and their perceived costs, defining equilibrium conditions, modeling the mode choice and computing paths on intermodal/multimodal networks. The intermodal optimum path algorithm by Ziliaskopoulos and Wardell (2000) provides a promising framework to both represent and compute paths that account for both highway and transit modes. Constructing analytical dynamic assignment approaches on networks with many modes would be even more challenging, especially since some of the fixed schedule lines need integer variables to be properly captured, potentially leading to intractable formulations.

3.2.4. Accounting for network controls and other realities Capturing network realities will be an ongoing research endeavor. While, in recent years, the computational power of commercially available machines has doubled approximately every other year, the need to capture more network realities will always stretch the hardware capabilities to the limit. Currently, the simulation based approaches (Mahmassani, 2001; Ben-Akiva et al., 2001; and Ziliaskopoulos and Waller, 2000) seem to be able to account for signal control, ramp metering, variable message signs, intersection movement delays, detector logic, different vehicle types and their interactions (trucks, buses), as well as information provision technologies. As DTA moves closer to deployment, it will become more critical to correctly account for control and the associated responses of drivers. It may not be sufficient to simply obtain the approximate impact of a traffic signal timing plan; a very precise representation may be required. This requires either microscopic or mesoscopic simulation models with time steps of a few seconds. Otherwise, the evolution of the queues behind traffic signals and ramp meters cannot be appropriately captured. However, the use of a very small time step (about 2 seconds) to capture the evolution of traffic on a fairly uniform long freeway segment is wasteful and unnecessary. Using multiple simulation time intervals is a promising alternative to finer simulation, though the traffic model should provide for such a capability. An extensive discussion on the models used to propagate traffic is included in the next section.

Another issue related to the realism of the network representation is that of capturing the interactions between some classes of vehicles and the infrastructure; such as modeling a truck negotiating a right turn at an intersection with inadequate geometry. For example, a stated preference survey by Peeta et al. (2000a) on driver response to variable message sign based messages suggests perceptible differences in the responses of truck drivers compared to those of non-truck drivers. This is because not all alternative diversion routes are viewed as feasible by truck drivers. Existing models only tentatively address this issue, especially on how infrastructure influences vehicle routing behavior. In a mixed traffic case, even computing the likely routes of vehicles is a difficult problem, as the travel delay experienced by a non-truck vehicle at an intersection with inadequate geometry will depend on whether a truck is present or not.

Finally, capturing the impact of information technology devices and the response of drivers to information can only be guessed at this point. Different devices and service providers enter the market every year; some have technology that needs to be simulated such as electronic toll collection readers, transit signal pre-emption technologies, real-time traffic adaptive responsive control (RT-TRAC) controllers and data collected from wireless phones.

Capturing some of the above realities with analytical closed form models will probably remain an open problem for a long time. Striving however to address these realities using any type of approach will be increasingly relevant as DTA is adopted by operating agencies.

3.2.5. Simultaneously optimizing controls and routing As discussed above, representing controls is an important issue that adds to the realism of the approach, but it is considerably simpler than actually optimizing the controls. To the best of our knowledge, no approach can currently simultaneously optimize routes and controls. The simulation DTA models can fairly easily represent them but they do not provide an appropriate structure to actually optimize them. The analytical approaches would become intractable even if all signalized intersections were treated in isolation. Again, this will possibly be an open issue for a very long time, which of course should not discourage researchers from striving to address it. Including signal optimization, capacity analysis and evaluation capabilities in a simulation assignment framework that can do user-specified hybrid macroscopic, mesoscopic and microscopic simulation and optimize routes will probably be the model that can finally integrate transportation planning, engineering and operations modeling and software. The benefits from the enhanced communication and database integration among these traditionally fragmented fields will be immense.

3.3. Other fundamental modeling issues

3.3.1. Describing equilibrium and behavioral issues Defining equilibrium on networks with time-dependent demand has consumed significant research energies, leading to various definitions and proofs. Most of them are simple extrapolations of Wardrop's static conditions. User equilibrium (UE) conditions mean that drivers follow time dependent least travel time paths, while system optimum (SO) conditions result from drivers following time dependent least marginal travel time paths. The simulation based approaches are flexible to perform UE, SO or multiple user classes assignment, though the equilibrium conditions are only heuristically approximated. Mahmassani (2001) assumes that users follow some user optimal routing behavior that could be influenced by information, leading to various classes of users. The boundedly rational drivers are probably the most flexible class that has the potential to capture a wide range of behavioral patterns. Ben-Akiva et al. (2001) assume user optimal behavior which implies that users cannot find a route they would prefer compared to the one they chose based on the provided information. Analytical approaches can guarantee equilibrium (UE or SO) and they can typically prove theoretically that a given formulation meets such conditions. Friesz et al. (1993) introduced the path integral equilibrium model, which is the first to generalize the Wardropian principle of equilibrium and account for the simultaneous route and departure-time choice on general networks. The generalization of the Wardropian principle leads to a variational inequalities (VI) formulation for the problem, where the dynamic equilibrium is expressed as a function of the integrals of path delays.

The question of whether equilibrium actually takes place or is a mathematical construct is a very old issue and probably precedes even the definition of traffic equilibrium itself. The context of time-dependent networks provides opportunities to relax the restrictive steady-state equilibrium assumptions and model phenomena of evolving disequilibria (Friesz et al., 1994). Probably the only experimental evidence of user decision-making behavior is an experiment, that involved 100 travelers over a 24-day period, performed by Chang and Mahmassani (1988). Two heuristic rules were proposed and calibrated for the adjustment of departure time. Friesz et al. (1994) addressed Mahmassani and Chang's design theoretically by introducing a tatonnement model for modeling transition of disequilibria from one state to another. The model led to the same conclusions as in Chang and Mahmassani (1988), that a day-to-day adjustment process could lead to an equilibrium state. The fact that there are always changes in supply, demand and traffic propagation, in combination with the stochasticity of all the involved parameters, makes the notion of equilibrium highly questionable.

3.3.2. Modeling traffic flow propagation Modeling traffic flow propagation is one of the fundamental DTA issues that concerned most researchers in the field, in their effort to balance realism and computational tractability. Merchant and Nemhauser (1978a, 1978b) were the first to realize that link performance functions (similar to the static BPR relationships) are not appropriate for dynamic networks, and they devised what they called exit functions to regulate the link outflow. Other formulations have been proposed since, including link performance functions used by some simulation based models, or approaches based on the cell transmission model and microscopic models. Next, we briefly discuss the main existing approaches to properly capture flow propagation on the links and nodes, and speculate on future developments.

Exit function Merchant and Nemhauser (1978a, 1978b) suggested the use of an exit function to capture congestion on a link. The exit function determines the outflow from a link given the number of vehicles on it. To express this function symbolically, let $x_a(t)$ denote the volume of traffic on arc a at time t, $u_a(t)$ the rate at which traffic enters arc a, and $g_a(x_a(t))$ the rate at which traffic exits the link. The M–N model, as stated earlier, is non-convex due to the exit functions considered. Carey (1986) showed that the model satisfied constraint qualifications, which ensured that the optimality conditions would hold at an optimum. Carey (1987) also showed that if the exit function constraint was re-written as an inequality, then the model is convex and, if certain weak restrictions hold, then in any solution of the model these inequalities will be binding, that is, satisfied as equalities. As discussed in Friesz et al.'s (2001) paper in this focus issue, the M–N flow balance equation based on the exit function can be written as follows:

$$\frac{dx_a}{dt} = u_a - g_a(x_a) \text{ for all arcs.}$$

Since first introduced, exit functions have been used extensively as a means to move traffic in DTA models (Carey, 1987; Friesz et al., 1989; Wie et al., 1995).

Since an exit function computes the outflow as a function of the number of vehicles on the link, it implicitly assumes that changes in density propagate instantaneously across the link. However, in a congested network, this shortcoming may not be a serious problem since links are not empty at either the beginning or the end of the time period of interest. The purpose of the model is to reflect the aggregate behavior of the vehicles on the network, not to track individual vehicle's movement. Exit flow functions have been criticized as difficult to specify and measure. In addition, they typically violate the first-in, first-out (FIFO) condition on the link as demonstrated in Carey (1986, 1987). This has led to the abandonment of exit functions for capturing traffic propagation on links. In the Friesz et al. (2001) paper, there is an extensive discussion of extensions to exit functions that can provide reasonably consistent propagation such as the exit time functions and their inverses. They also provide a way of expressing exit-time function based model of arc dynamics to obtain an alternative formulation involving constrained differential equations, state-dependent lags, and arc entrance and exit flows that are control variables rather than operators.

Link performance function Link performance functions are used in Janson (1991a), Ran et al. (1996), and Chen and Hsueh (1998) to decide the time-dependent link-path matrix. Most link performance functional forms appear to be straightforward temporal extensions of static BPR functions, which suffer from the well-documented drawbacks of realism and consistency. These are further exacerbated by the need to capture dynamics; i.e., to account for the spatio-temporal distribution of traffic on the link while propagating traffic. Some simulation models (Ben-Akiva et al., 2001) use similar functions but only on part of the link. Specifically, they divide links into moving and queuing sections; the BPR type function is applied only to the moving part which is characterized by more uniform conditions than the rest of the link. Mahmassani (2001) moves vehicles on links or sub-links according to a modified Greenshield type speed-density relationship, noting that other traffic stream models could also be incorporated based on field investigation.

Link performance functions tend to be convex, which makes them convenient for closed form formulations, but they suffer from serious drawbacks of realism, mainly because they cannot capture the link dynamics in propagating traffic. Dividing the link into small sub-sections can approximate the link queue evolution but it is not clear how accurate this representation is. Field calibration experiments will ultimately determine the suitability of these models.

The cell transmission model The cell transmission model was proposed by Daganzo (1994) to simulate the evolution of traffic on a single highway link. It was shown that if the relationship between traffic flow (q) and density (k) is of the form:

$$q = \min\{vk, q_{max}, w(k_j - k)\}, \quad \text{for } 0 \leq k \leq k_j$$

then Lighthill and Whitham's (1955) and Richards' (1956) equations for a single highway link can be approximated by a set of difference equations, with current conditions (the state of the system) being updated at every time interval. In the relationship above, v, q_{max}, w, and k_j are constants that denote the free flow speed, the maximum flow (or capacity), the backward propagation speed, and the jam density, respectively. The model discretizes the time period of interest (assignment period) into small intervals; based on that, it divides every link of the street network into small homogeneous segments, called cells, so that the length of each cell is equal to the distance traveled by the free-flow moving vehicles in one time interval.

Ziliaskopoulos and Lee (1997) extended the model to represent freeway and arterial street networks. Ziliaskopoulos (2000) suggested that the traffic flow propagation relationships could be captured by a set of linear constraints that results in linear programming DTA models. The main advantages of the cell transmission model are its simplicity, versatility and linearity; its major drawbacks are the determinism of the traffic propagation on highly congested links and the precision errors introduced when the discretization time interval is smaller than 2 seconds. While both concerns can be addressed, they could result in less tractable formulations. Embedding the cell transmission model in analytical formulations, however, presents unprecedented advantages, since it provides superior realism in a linear program that allows the vast existing literature on linear programming to be used to better understand and compute DTA. It can also provide for extensions to formulate network design and stochastic programming formulations as discussed in the following section.

In conclusion, it needs to be emphasized again that all the representations above are simply models and as such abstractions of reality; the ultimate model suitability will be determined when enough data is available to calibrate and evaluate them with real data.

3.3.3. Demand uncertainty DTA models typically assume known point estimates of time-dependent O-D demand. The main objective of the DTA research is to relax the assumption of time-invariance of demand. The next natural question is whether the assumption of the determinism of demand can be relaxed. In general, for both planning and operational applications, it is unclear what action should be taken or what the cost is, when the forecast demand is not realized. It becomes necessary, therefore, to develop a systematic way of accounting for demand uncertainty and develop a more robust solution, which is more likely to withstand extreme events. One such possibility is to run the deterministic model with a large number of randomly generated demand patterns, and to infer some rules and principles from the results. Such an approach has been proposed by Peeta and Zhou (1999b) to produce robust DTA solutions. The only drawback with this approach appears to be the computational requirements of a single DTA run and the enormous amount of scenarios that need to run to produce meaningful results and inferences. However, as discussed in Peeta and Zhou (1999a), the computational aspect can be addressed by using a hybrid deployment framework consisting of off-line and real-time components. Chance Constrained Programming (CCP) introduced by Waller and Ziliaskopoulos (1998) is an alternative approach. The CCP model is introduced as an extension of a deterministic linear programming model. The insights from the analytical model suggest that the uncertain demand needs to be inflated depending on its variance, which can also be applied to simulation based approaches. Waller (2000) formulates a two-stage stochastic programming approach with recourse for planning DTA approaches, with promising results. However, the issue remains open and would probably be one of the most active areas in DTA and traffic control/optimization in the years to come, as stochasticity seems to be the rule rather than the exception in traffic systems. In fact, one of the emerging questions is whether accounting for the stochastic nature of traffic systems is more important than accounting for dynamics. This is an issue that will probably be debated by the research community in the years to come.

3.3.4. User behavior Mahmassani (2001) reports extensively on the effort to account for behavioral elements. One of the principal features of this approach's interface with activity based behavioral models is its explicit representation of individual trip-making decisions, particularly for path selection decisions, both at the trip origin and en-route. Behavioral rules governing route-choice decisions are incorporated, including the special case in which drivers are assumed to follow specific route guidance instructions. Experimental evidence presented by Chang and Mahmassani (1988) and Mahmassani (2001) suggests that commuter route choice behavior exhibits a boundedly-rational character. This means that drivers look for gains only outside a threshold, within which the results are satisfying and sufficing for them. This can be translated to the following route switching model (Mahmassani and Jayakrishnan, 1991):

$$\delta_j(k) = \begin{cases} 1 & \text{if} \quad TTC_j(k) - TTB_j(k) > \max(\eta_j \cdot TTC_j(k), \tau_j) \\ 0 & \text{otherwise} \end{cases}$$

where $\delta_j(k)$ is a binary indicator variable equal to 1 when user j switches from the current path to the best alternate, and 0 if the current path is maintained; $TTC_j(k)$ and $TTB_j(k)$ are the trip times along the current path and along the best path from node k to the destination on current path, respectively; η_j is a relative indifference threshold, and τ_j is an absolute minimum travel time improvement needed for a switch. The threshold level may reflect perceptual factors, preferential indifference, or persistence and aversion to switching. The quantity η_j governs users' responses to the supplied information and their propensity to switch. The minimum improvement τ_j is currently taken to be identical across users according to user-defined values. Further discussion on this issue can be found in Mahmassani's (2001) paper in this special issue. Very little seems to have been done, besides the above-mentioned work, in this important field.

3.3.5. Path processing Path processing lies at the core of DTA development; routes need to be computed for every modeled user class accounting for intersection movement delays, general link cost functions, many modes of transportation, fixed arrival times and other network realities. The work by Ziliaskopoulos and Mahmassani (1993, 1994, 1996) addresses many of these issues, but given that in the existing operational models path processing seems to be a key computational bottleneck, more needs to be done. Ziliaskopoulos et al. (1997) proposed a parallel implementation that can potentially reduce the required computational time, but it requires rather elaborate hardware settings which discourages researchers from applying it. The evolution of the Internet and information technology state-of-practice presents many opportunities for distribution and parallelism not available only a few years ago. More research needs to be done to take advantage of these technologies. For example, Beowulf clusters (Peeta et al., 2000b) represent an economical and flexible high-performance computing environment in this context. Another area related to path processing is that of computing routes when the link travel times are not only time-dependent but also stochastic, which necessitates on-line shortest path approaches.

3.3.6. Tractable analytical approaches, FIFO violations and efficient decomposition approaches While there are various analytical approaches and solution algorithms, as discussed in the review section of this paper, there is no analytical formulation that can efficiently solve actual size networks. In other words, there is no equivalent to the Frank-Wolfe solution algorithm we have for static approaches that could conveniently decompose on of the analytical approaches to some sort of a path processing algorithm. Waller (2000) suggested a decomposition approach for the cell transmission based linear programming formulation that solves min-cost flow sub-problems on a modified network, but the approach is yet to be applied on actual size networks. In addition the analytical approaches seem to still suffer from the limitations of single destination (unless the first-in, first-out condition on the links is violated), especially for user optimum behavior. For system optimum solutions satisfying the FIFO condition may not be as critical, especially if they simply aim to provide a lower bound on the system performance. Constraining for non-FIFO conditions would at best provide a tighter bound which makes the solution questionable. It is the belief of the authors that the FIFO violation for SO solutions is the temporal equivalent of unreasonable paths produced by static approaches; the SO approach should hold traffic at various places in the network, since this provides a better system optimum solution.

3.3.7. Departure time choice modeling or FAT/FDT According to Li (2001), a recent travel survey in the Puget Sound region in the United States revealed that almost 70% of the trips during the morning commute are work or school related trips that tend to fix their arrival rather than their departure time (see Table 1). This presents challenges for researchers to develop models that account for this class of users, but also opportunities, since arrival time based demand is arguably easier to estimate than departure time based one. Accounting for fixed arrival time based demand leads to the simultaneous route choice and departure time assignment problem. The existing approaches in the literature for this problem can be grouped into empirical approaches, which are based on experimental or survey data, and analytical ones, which are typically limited to formulations with no efficient solution algorithms.

Chang and Mahmassani (1988), Mahmassani and Herman (1984), Friesz et al. (1993), De Palma et al. (1983), Hendrickson and Kocur (1981), and Friesz et al. (1994) dealt with the simultaneous route and departure-time choice on urban networks. Equilibrium conditions for simultaneous departure time and route were first defined by Friesz et al. (1993), and can be loosely stated as follows:

> *For any $h = (h_{rs}^{k\beta}(t) : k \in P, \beta \in B)$, the pair (h,μ) is designated as a simultaneous route-departure time equilibrium, if and only if the following two conditions are satisfied:*
>
> $$h_{rs}^{k\beta}(t) > 0 \Rightarrow T_{rs\beta}^{k}(t,h) + \phi[t + T_{rs\beta}^{k}(t,h)] = \mu_{rs\beta}(h),\ all\, r,\ s,\ \beta.$$
>
> $$T_{rs\beta}^{k}(t,h) + \phi[t + T_{rs\beta}^{k}(t,h)] \geq \mu_{rs\beta}(h),\ all\, r,\ s,\ \beta.$$

Where, $h_{rs}^{k\beta}(t)$ is the flow on path k that connects nodes r and s at time t and arrival time β measured at the entrance of the first link on the path, $\phi[t + T_{rs\beta}^{k}(t,h)]$ is a cost function of

Table 1. Temporal distribution of work and school trips

	Time of day				
	0:01–5:00	5:01–9:00	9:01–14:00	14:01–19:00	19:01–24:00
Work and school trips	49	2036	1115	493	55
Total trips	154	2998	4491	5986	1867
Percentage of work and school trips	31.5%	67.9%	24.8%	8.2%	2.9%

early or late schedule delay ($\phi[\beta] = 0$) and $\mu_{rs\beta}(h) = min\{T^k_{rs\beta}(t, h) + \phi[t + T^k_{rs\beta}(t, h)]\}$ over all paths k in P that connect r and s. These conditions simply state that at equilibrium a user cannot switch either departure time and/or path and reduce his/her personal gain. Based on this definition Friesz et al. (1993) proposed a variational inequality formulation and proved that it is equivalent to the definition stated above. Ziliaskopoulos and Rao (1999) introduced a simulation based fixed arrival based heuristic solution methodology that approximates the above equilibrium condition. A similar heuristic solution aimed to emulate day-to-day dynamics and individual drivers' decisions as discussed in Chang and Mahmassani (1988).

Finally, Li et al. (1999) and Li (2001) introduced a linear programming (LP) formulation for the simultaneous departure time and route choice problem under SO conditions for networks with multiple origins and destinations. The LP is formulated using the cell transmission model. The necessary and sufficient conditions for SO DTA fixed arrival time were proved to be the following:

> *The necessary and sufficient condition for SO FAT-DTA is that all used paths to a node at a certain arrival time from an origin have cost equal to the marginal cost of an additional unit of demand at that node and time interval from that origin, while all unused paths have cost higher than or equal to the marginal cost.*

Extensions to the model that account for penalties associated with schedule delays, and for simultaneous loading of arrival time based and departure time based demand were also proposed.

Despite the efforts outlined above, estimation of the departure time choices is not well understood. It is potentially one of the main advantages of DTA over steady state models, where departure time is meaningless. In closing this section, we venture to state that the potential of DTA to account for this important decision variable makes it not only superior to static approaches, but also necessary for planning applications. Ignoring the departure time when the common sense and experimental data suggest that it is the main decision variable most users are likely to consider adjusting would be unfortunate, especially since time-dependent demand for this user class can potentially be estimated more easily than for other user classes. If the arguably obsolete four planning steps were to be followed, estimating departure time would have been the fifth planning step.

4. Closing Comments

Dynamic traffic assignment has evolved rapidly over the past two decades, fueled by the needs of applications domains ranging from real-time traffic operations to long-term planning. Characterized by inherent mathematical intractability and challenging complexities, it has nevertheless spawned a vast body of literature that encompasses a broad gamut of problems with different underlying assumptions and functional objectives. While focused research has led to rapid strides in the understanding of the problem characteristics, it has also highlighted the difficulties involved in developing a universally applicable approach for general networks in the operational domain. For example, a mathematically tractable analytical model that is adequately sensitive to traffic realism vis-à-vis real-time operations, is still elusive. Consequently, current research still focuses on fundamental problem characteristics while venturing into deployment issues that are motivated by current and future operational needs.

Researchers in the next few years will debate and research the issue of applicability of DTA for planning applications. Early attempts were consumed with comparatively evaluating static and dynamic models applied to the same network and total demand; this is a futile effort, since the ground truth will never be known to actually compare the outcome of the models with. Debating on whether dynamic models are better than static ones is hardly the issue; dynamic models are obviously superior, since they relax more assumptions and capture more realities than the static approaches. Dynamic models are simply the natural evolution in the transportation field that like any other new effort suffers from early development shortcomings; it is only a matter of time before they are improved and ultimately adopted by the industry. Academics and researchers should look ahead and be concerned with some of the fundamental questions posed above, especially those related to the realism of assumptions and mathematical tractability.

References

Anastassopoulos, I. (2000). "Fault Tolerance and Incident Detection Using Fourier Transforms." M.S. Thesis, Purdue University.

Ashok, K. and M.E. Ben-Akiva. (2000). "Alternative Approaches for Real-time Estimation and Prediction of Time-dependent Origin-Destination Flows." *Transportation Science* 34(1), 21–36.

Beckmann, M.J., C.B. McGuire, and C.B. Winston. (1956). *Studies in the Economics of Transportation*. Connecticut: Yale University Press.

Ben-Akiva, M.E., Bierlaire, M., Bottom, J., Koutsopoulos, H.N. and Mishalani, R. (1997a). "Development of a Route Guidance Generation System for Real-Time Application." *Proceedings of the 8th International Federation of Automatic Control Symposium on Transportation Systems*, Chania, Greece, pp. 433–438.

Ben-Akiva, M.E., Koutsopoulos, H.N., Mishalani, R. and Yang, Q. (1997b). "Simulation Laboratory for Evaluating Dynamic Traffic Management Systems." *ASCE Journal of Transportation Engineering* 123(4), 283–289.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H.N., and Mishalani, R. (1998). "DynaMIT: A Simulation-based System for Traffic Prediction and Guidance Generation." *Proceedings of the 3rd Triennial Symposium on Transportation Systems*, San Juan, Puerto Rico.

Birge, J.R. and J.K. Ho. (1993). "Optimal Flows in Stochastic Dynamic Networks with Congestion." *Operations Research* 41(1), 203–216.

Boyce, D.E., B. Ran, and L.J. LeBlanc. (1995). "Solving an Instantaneous Dynamic User-Optimal Route Choice Model." *Transportation Science* 29, 128–142.

Cantarella, G.E. and E. Cascetta. (1995). "Dynamic Processes and Equilibrium in Transportation Networks: Towards a Unifying Theory." *Transportation Science* 29, 305–329.

Carey, M. (1986). "A Constraint Qualification for a Dynamic Traffic Assignment Model." *Transportation Science* 20, 55–58.

Carey, M. (1987). "Optimal Time Varying Flows On Congested Networks." *Operations Research* 35(1), 58–69.

Carey, M. (1992). "Nonconvexity of the Dynamic Assignment Problem." *Transportation Research* 26B(2), 127–133.

Carey, M. and E. Subrahmanian. (2000). "An Approach to Modeling Time-Varying Flows on Congested Networks." *Transportation Research* 34B(3), 157–183.

Cascetta, E., D. Inaudi, and G. Marquis. (1993). “Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts.” *Transportation Science* 27, 363–373.

Chang, G.-L. and H.S. Mahmassani. (1988). “Travel Time Prediction and Departure Time Adjustment Behavior Dynamics in A Congested Traffic System.” *Transportation Research* 22B(3), 217–232.

Chang, G.L. and X.D. Tao. (1999). “An Integrated Model for Estimating Time-Varying Network Origin-Destination Distributions.” *Transportation Research* 33A(5), 381–399.

Chen, H.-K. (1999). *Dynamic Travel Choice Models*. Berlin: Springer-Verlag.

Chen, H.-K. and C.F. Hsueh. (1998). “A Model and An Algorithm for the Dynamic User-Optimal Route Choice Problem.” *Transportation Research* 32B(3), 219–234.

Cremer, M. and H. Keller. (1987). “A New Class of Dynamic Methods for the Identification of Origin-Destination Flows.” *Transportation Research* 21B(2), 117–132.

Dafermos, S. (1980). “Traffic Equilibrium and Variational Inequalities.” *Transportation Science* 14, 42–54.

Daganzo, C. (1994). “Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory.” *Transportation Research* 28B(4), 269–287.

De Palma, A., Ben-Akiva, M., Lefevre, C. and Litinas, N. (1983). “Stochastic Equilibrium Model of Peak Period Traffic Congestion.” *Transportation Science* 17(4), 430–453.

De Palma, A. and F. Marchai. (2000) “Dynamic Traffic Analysis with Static Data”: Some Guidelines from the Application to Paris, TRB Preprint. *Transportation Research Record*.

Friesz, T.L., Bernstein, D., Mehta, N.J., Tobin, R.L. and Ganjalizadeh, S. (1989). “Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem.” *Operations Research* 37(6) 893–901.

Friesz, T.L., T.L., Bernstein, D., Smith, T.E., Tobin, R.L. and Wie, B.W. (1993). “A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem.” *Operations Research* 41(1), 179–191.

Friesz, T.L., Bernstein, D., Mehta, N.J., Tobin, R.L., and Ganjalizadeh, S. (1994). “Day-to-Day Dynamic Network Disequilibrium and Idealized Traveler Information Systems.” *Operations Research* 42(6), 1120–1136.

Ghali, M.O. and M.J. Smith. (1992a). “A Dynamic Traffic Assignment Model.” Presented at the *71st TRB Annual Meeting*, Washington D.C.

Ghali, M.O. and M.J. Smith. (1992b). “Dynamic Road Pricing and Dynamic Traffic Assignment for Congested Large Networks.” Presented at the *39th North American Regional Science Association International Conference*, Chicago.

Ghali, M.O. and M.J. Smith. (1992c). “Optimal Dynamic Traffic Assignment of a Congested City Network.” *Proceedings of the Second International Capri Seminar on Urban Traffic Networks*, Capri, Italy.

Ghali, M.O. and M.J. Smith. (1993). “Traffic Assignment, Traffic Control and Road Pricing.” *Proceedings of the Twelfth International Symposium on Transportation and Traffic Theory*, Daganzo (ed.), University of California, Berkeley, California, pp. 147–170.

Hawas, Y.E. (1995). “A Decentralized Architecture and Local Search Procedures for Real-Time Route Guidance in Congested Vehicular Traffic Networks.” Ph.D. Dissertation, The University of Texas at Austin.

Hawas, Y.E. and H.S. Mahmassani. (1995). “A Decentralized Scheme for Real-time Route Guidance in Vehicular Traffic Networks.” *Proceedings of the Second World Congress on Intelligent Transport Systems*, Yokohama, Japan.

Hawas, Y.E. and H.S. Mahmassani. (1997). “Comparative Analysis of Robustness of Centralized and Distributed Network Route Control Systems in Incident Situations.” *Transportation Research Record* 1537, 83–90.

Hawas, Y.E., Mahmassani, H. S., Chang, G.L., Taylor, R., Peeta, S. and Ziliaskopoulos, A. (1997). “Development of DYNASMART-X Software for Real-Time Dynamic Traffic Assignment.” *Technical Report ST067-85-TASK-E*, Center for Transportation Research, The University of Texas at Austin.

Hendrickson, C. and G. Kocur. (1981). “Schedule Delay and Departure Time Decisions in a Deterministic Model.” *Transportation Science* 15, 62–77.

Ho, J.K. (1980). “A Successive Linear Optimization Approach to the Dynamic Traffic Assignment Problem.” *Transportation Science* 14, 295–305.

Horowitz, J.L. (1984). “The Stability of Stochastic Equilibrium in a Two-link Transportation Network.” *Transportation Research* 18B, 12–28.

Hu, S.-R., Madant, S. M., Krogmeier, J.V. and Peeta, S. (2001). "Estimation of Dynamic Assignment Matrices and O-D Demands using Adaptive Kalman Filtering", to appear in the *Intelligent Transportation Systems Journal*.

Janson, B.N. (1991a). "Convergent Algorithm for Dynamic Traffic Assignment." *Transportation Research Record* 1328, 69–80.

Janson, B.N. (1991b). "Dynamic Traffic Assignment for Urban Road Networks." *Transportation Research* 25B(2/3), 143–161.

Jayakrishnan, R., H.S. Mahmassani, and T.Y. Hu. (1994). "An Evaluation Tool for Advanced Traffic Information and Management Systems in Urban Networks." *Transportation Research* 2C(3), 129–147.

Leonard, D., P. Gower, and N.B. Taylor. (1989). "CONTRAM: Structure of the Model." TRRL *Research Report* RR 178, Transport and Road Research Laboratory, Crowthorne, United Kingdom.

Li, Y. (2001). "Development of Dynamic Traffic Assignment Models for Planning Applications." Ph.D. Dissertation, Northwestern University, Evanston, IL.

Li, I.Y., A.K. Ziliaskopoulos, and S.T. Waller. (1999). "Linear Programming Formulations for System Optimum Dynamic Traffic Assignment with Arrival Time Based and Departure Time Based Demands." *Transportation Research Record* 1832, 121–142.

Lighthill, M.J. and J.B. Whitham. (1955). "On Kinematic Waves, I: Flow Modeling in Long Rivers; II: A Theory of Traffic Flow on Long Crowded Roads." *Proceedings of Royal Society* A229, 281–345.

Mahmassani, H.S. and G.L. Chang. (1985). "Dynamic Aspects of Departure-Time Choice Behavior in A Commuting System: Theoretical Framework and Experience Analysis." *Transportation Research Record* 1037, 88–101.

Mahmassani, H.S. and R. Herman. (1984). "Dynamic User Equilibrium Departure Time and Route Choice on Idealized Traffic Arterials." *Transportation Science* 18(4), 362–384.

Mahmassani, H.S. and S. Peeta. (1992). "System Optimal Dynamic Assignment for Electronic Route Guidance in a Congested Traffic Network." *Proceedings of the Second International Capri Seminar on Urban Traffic Networks*, Capri, Italy, 1992.

Mahmassani, H.S. and S. Peeta. (1993). "Network Performance Under System Optimal and User Equilibrium Dynamic Assignments: Implications for ATIS." *Transportation Research Record* 1408, 83–93.

Mahmassani, H.S. and S. Peeta. (1995). "System Optimal Dynamic Assignment for Electronic Route Guidance in a Congested Traffic Network." In Gartner and Improta (eds.), *Urban Traffic Networks: Dynamic Flow Modeling and Control*, Berlin: Springer-Verlag, pp. 2–27.

Mahmassani, Peeta, S., Hu, T., and Ziliaskopoulos, A. (1993). "Algorithm for Dynamic Route Guidance in Congested Networks with Multiple User Information Availability Groups." *Proceedings of the 26th International Symposium on Automotive Technology and Automation*, Aachen, Germany, pp. 273–280.

Mahmassani, H.S., Hawas, Y.E., Abdelghany, K., Abdelfatah, A., Chiu, Y.-C., Kang, Y., Chang, G.L., Peeta, S., Taylor, R. and Ziliaskopoulos, A. (1998a). "DYNASMART-X; Volume II: Analytical and Algorithmic Aspects." *Technical Report ST067-85-Volume II*, Center for Transportation Research, The University of Texas at Austin.

Mahmassani, H.S., Chiu, Y.-C., Chang, G.L., Peeta, S. and Ziliaskopoulos, A. (1998b). "Off-line Laboratory Test Results for the Dynasmart-X Real-Time Dynamic Traffic Assignment System." *Technical Report ST067-85-TASK-G*, Center for Transportation Research, The University of Texas at Austin.

Merchant, D.K. and G.L. Nemhauser. (1978a). "A Model and an Algorithm for the Dynamic Traffic Assignment Problems." *Transportation Science* 12(3), August 1978, 183–199.

Merchant, D.K. and G.L. Nemhauser. (1978b). "Optimality Conditions for a Dynamic Traffic Assignment Model." *Transportation Science* 12(3), August 1978, 200–207.

Nagurney, A. (1998). *Network Economics: A Variational Inequality Approach*. Boston: Kluwer Academic Publishers.

Nagurney, A. and D. Zhang. (1996). *Projected Dynamical Systems and Variational Inequalities with Applications*. Boston: Kluwer Academic Publishers.

Nagurney, A. and D. Zhang. (1997). "Projected Dynamical Systems in the Formulation, Stability Analysis, and Computation of Fixed-demand Traffic Network Equilibria." *Transportation Science* 31, 147–158.

Patriksson, M. (1994). *The Traffic Assignment Problem: Models and Methods*. Utrecht, The Netherlands: VSP.

Pavlis, Y. and M. Papageorgiou. (1999). "Simple Decentralized Feedback Strategies for Route Guidance in Traffic Networks." *Transportation Science* 33, 264–278.

Peeta, S. (1994). "System Optimal Dynamic Traffic Assignment in Congested Networks with Advanced Information Systems." Ph.D. Dissertation, The University of Texas at Austin.

Peeta, S. and S. Bulusu. (1999). "A Generalized Singular Value Decomposition Approach for Consistent Online Dynamic Traffic Assignment." *Transportation Research Record* 1667, 77–87.

Peeta, S. and H.S. Mahmassani. (1995a). "System Optimal and User Equilibrium Time-Dependent Traffic Assignment in Congested Networks." *Annals of Operations Research* 60, 81–113.

Peeta, S. and H.S. Mahmassani. (1995b). "Multiple user Classes Real-time Traffic Assignment for On-line operations: A Rolling Horizon Solution Framework." *Transportation Research* 3C(2), 83–98.

Peeta, S. and T.-H. Yang. (2000a). "Stability Issues for Dynamic Traffic Assignment." *Proceedings of the 9th International Federation of Automatic Control Symposium on Control in Transportation Systems*, Schnieder and Becker (eds.), Braunschweig, Germany, pp. 507–511.

Peeta, S. and T.-H. Yang. (2000b). "Stability of Large-scale Dynamic Traffic Networks under On-line Control Strategies." *Proceedings of the 6th International Conference on Applications of Advanced Technologies in Transportation Engineering*, Singapore.

Peeta, S. and C. Zhou. (1999a). "Robustness of the Off-line A Priori Stochastic Dynamic Traffic Assignment Solution for On-Line Operations." *Transportation Research* 7C, 281–303.

Peeta, S. and C. Zhou. (1999b). "On-line Dynamic Update Heuristics for Robust Route Guidance." In Adamski, Rudnicki and Zak (eds.), *Proceedings of the International Conference on Modeling and Management in Transportation*, Cracow, Poland, 2, 119–126.

Peeta, S., J.L. Ramos, and R. Pasupathy. (2000a). "Content of Variable Message Signs and On-line Driver Behavior." *Transportation Research Record* 1725, 102–108.

Peeta, S., R. Pasupathy, and S. Gedela. (2000b). "An Internet Based On-line Architecture for Real-Time Traffic Systems Control." *Proceedings of the 9th International Federation of Automatic Control Symposium on Control in Transportation Systems*, Schnieder and Becker (eds.), Braunschweig, Germany, 314–319.

Ran, B. and D. Boyce. (1994). *Dynamic Urban Transportation Network Models, Lecture Notes in Economics and Mathematical Systems* 417. Berlin: Springer-Verlag.

Ran, B. and D. Boyce. (1996a). *Modeling Dynamic Transportation Networks*. Berlin: Springer-Verlag.

Ran, B. and D.E. Boyce. (1996b). "A Link-based Variational Inequality Formulation of Ideal Dynamic User-Optimal Route Choice Problem." *Transportation Research* 4C(1), 1–12.

Ran, B. and T. Shimazaki. (1989a). "A General Model and Algorithm for the Dynamic Traffic Assignment Problems." *Proceedings of the Fifth World Conference on Transport Research*, Yokohoma, Japan.

Ran, B. and T. Shimazaki. (1989b). "Dynamic User Equilibrium Traffic Assignment for Congested Transportation Networks." Presented at the Fifth World Conference on Transport Research, Yokohoma, Japan.

Ran, B., D.E. Boyce, and L.J. LeBlanc. (1993). "A New Class of Instantaneous Dynamic User-Optimal Traffic Assignment Models." *Operations Research* 41(1), 1993.

Ran, B., R.W. Hall, and D.E. Boyce. (1996). "A Link-based Variational Inequality Model for Dynamic Departure Time/Route Choice." *Transportation Research* 30B(1), 31–46.

Richards, P.I. (1956). "Shockwaves on the Highway." *Operations Research* 4, 42–51.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. NJ: Prentice-Hall.

Sherali, H.D., N. Arora, and A.G. Hobeika. (1997). "Parameter Optimization Methods for Estimating Dynamic Origin-Destination Trip-Tables." *Transportation Research* 31B(2), 141–157.

Smith, M.J. (1979). "Existence, Uniqueness, and Stability of Traffic Equilibria." *Transportation Research* 13B, 295–304.

Smith, M.J. (1984). "The Stability of a Dynamic Model of Traffic Assignment – An Application of a Method of Lyapunov." *Transportation Science* 18, 245–252.

Smith, M.J. (1994). "A New Dynamic Traffic Model and the Existence and Calculation of Dynamic User Equilibria on Congested Capacity-constrained Road Networks." *Transportation Research* 27B, 49–63.

Taylor, N.B. (1990). "CONTRAM 5: An Enhanced Traffic Assignment Model." TRRL *Research Report* RR 249, Transport and Road Research Laboratory, Crowthorne, United Kingdom.

Taylor, N.B. (1996). "CONTRAM Modeling in the Real-time Context – A Review and Prognosis." TRRL *Project Report* PR/TT/088/96, Transport and Road Research Laboratory, Crowthorne, United Kingdom.

Waller, S.T. (2000). "Optimization and Control of Stochastic Dynamic Transportation Systems: Formulations, Solution Methodologies, and Computational Experience." Ph.D. Dissertation, Northwestern University, Evanston, IL.

Waller, S.T. and A.K. Ziliaskopoulos. (1998). "A Chance-Constrained Based Stochastic Dynamic Traffic Assignment Model: Analysis, Formulation and a Solution Algorithm." *TRISTAN III Proceedings*, San Juan, Puerto Rico.

Wardrop, J.G. (1952). "Some Theoretical Aspects of Road Traffic Research." *Proceedings, Institution of Civil Engineers* II(1), 325–378.

Watling, D. (1999). "Stability of the Stochastic Equilibrium Assignment Problem: a Dynamical Systems Approach." *Transportation Research* 33B, 281–312.

Wie, B.W. (1990). "Dynamic Analysis of User Optimized Network Flows With Elastic Travel Demand." Presented at the 70th TRB Annual Meeting, Washington, D.C., 1991.

Wie, B.W., R.L. Tobin, T.L. Friesz, and D. Bernstein. (1995). "A Discrete Time, Nested Cost Operator Approach to the Dynamic Network User Equilibrium Problem." *Transportation Science* 29(1), 79–92.

Wu, J.F. and G.L. Chang. (1996). "Estimation of Time-Varying Origin-Destination Distributions with Dynamic Screenline Flows." *Transportation Research* 30B(4), 277–290.

Yang, T.-H. (2001). "Deployable Stable Traffic Assignment Models for Control in Dynamic Traffic Networks: A Dynamical Systems Approach." Ph.D. Dissertation, Purdue University.

Zhang, D. and A. Nagurney. (1995). "On the Stability of Projected Dynamical Systems." *Journal of Optimization Theory and Applications* 85, 97–124.

Zhang, D. and A. Nagurney. (1996). "On the Local and Global Stability of a Travel Route Choice Adjustment Process." *Transportation Research* 30B, 245–262.

Ziliaskopoulos, A.K. (2000). "A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem." *Transportation Science* 34(1), 37–49.

Ziliaskopoulos, A.K. and S. Lee. (1997). "A Cell Transmission Based Assignment-Simulation Model for Integrated Freeway/Surface Street Systems." *Transportation Research Record* 1701, 12–23.

Ziliaskopoulos, A.K. and H.S. Mahmassani. (1993). "A Time-Dependent Shortest Path Algorithm for Real-Time Intelligent Vehicle/Highway Systems." *Transportation Research Record* 1408, 94–104.

Ziliaskopoulos, A.K. and H.S. Mahmassani. (1994). Technical Commentary on the Paper Entitled "The Fastest Path in Time-Dependent Networks for Intelligent Vehicle/Highway Systems," by D.E. Kaufman and R.L. Smith, in *IVHS Journal* 1, 91–95.

Ziliaskopoulos, A.K. and H.S. Mahmassani. (1996). "On Finding Least Time Paths Considering Delays for Intersection Movements." *Transportation Research* 30B(5), 359–367.

Ziliaskopoulos, A.K. and L. Rao. (1999). "A Simultaneous Route and Departure Time Choice Equilibrium Model on Dynamic Networks." *International Transactions of Operational Research* 6(1), 21–37.

Ziliaskopoulos, A.K. and S.T. Waller. (2000). "An Internet-based Geographical Information System that Integrates Data, Models and Users for Transportation Applications." *Transportation Research* 8C(1), 427–444.

Ziliaskopoulos, A.K. and W.W. Wardell. (2000). "An Intermodal Optimum Path Algorithm for Dynamic Multimodal Networks." *European Journal of Operational Research* 125, 486–502.

Ziliaskopoulos, A.K., D. Kotzinos, and H.S. Mahmassani. (1997). "Design and Implementation of Parallel Time-Dependent Shortest Path Algorithms for Real-Time Intelligent Transportation Systems." *Transportation Research* 5C(2), 95–107.

Ziliaskopoulos, A.K., D. Doan, and H.S. Mahmassani. (1998). "An On-line Monitoring System for Real-Time Traffic Management Applications." *Transportation Research Record* 1678, 142–150.