Integrating an Agent-Based Travel Behavior Model with Large-Scale Microscopic Traffic Simulation for Corridor-Level and Subarea Transportation Operations and Planning Applications

Lei Zhang\textsuperscript{1}; Gang-Len Chang, M.ASCE\textsuperscript{2}; Shanjiang Zhu\textsuperscript{3}; Chenfeng Xiong\textsuperscript{4}; Longyuan Du\textsuperscript{5}; Mostafa Mollanejad\textsuperscript{6}; Nathan Hopper\textsuperscript{7}; and Subrat Mahapatra\textsuperscript{8}

Abstract: Application of microscopic traffic simulation beyond the corridor level analysis is not widely seen in literature. This is partly because of the fact that a simulation model cannot capture behavior responses such as peak spreading. This study develops a framework that integrates agent-based travel behavior models with large-scale traffic simulation to capture the regional impacts of new development. The proposed model is then applied to the I-270/I-495/I-95 corridor in the north Washington, DC metropolitan area in a case study. Findings from this study reveal the potential of the proposed model to capture network dynamics and behavioral reactions. This framework also provides a valuable tool for the evaluation of new transportation infrastructure, such as the intercounty connector (ICC) corridor currently under construction, and its operation strategies. DOI: 10.1061/(ASCE)UP.1943-5444.0000139, © 2013 American Society of Civil Engineers.

CE Database subject headings: Traffic management; Simulation; Traffic models; Travel patterns.

Author keywords: Traffic simulation; Agent-based model; Departure time choice; Intercounty connector.

Introduction

Microscopic traffic simulation models exhibit strong advantages in capturing detailed traffic dynamics and have proved to be a valuable tool for evaluating corridor capacity expansion and traffic operation improvements. Their applications have recently been extended to address a broader range of transportation-related issues, including congestion management, multimodal corridor improvements, evacuation planning, land use, and economic development. However, a comprehensive analysis of many of these issues requires models that can consider various demand responses to these traffic management strategies, such as peak spreading, modal shifts, and traffic diversions at the corridor and regional levels, all of which are conventionally given in microsimulation models. Another challenge in applying microsimulation models to a large network is because of the difficulty of obtaining reliable travel demand data (usually as time-dependent origin-destination matrices).

Although planning models are traditionally used to address these demand-side problems, they are criticized for (1) assigning traffic flow over capacity, and (2) being unable to capture operational improvements such as better signal timing. As planning models move from the aggregate four-step models into more realistic individual-based models, more details on travel experience (e.g., time-dependent travel time) are required to make these models operational in practice. Surveys based on hypothetical scenarios, which are heavily relied on during the model development and calibration process, can only support analysis for a limited number of origin-destination (OD) pairs because of budget and labor constraints. However, these inputs can be easily extracted from microscopic traffic simulation models. Therefore, it becomes very attractive for both researchers and practitioners to develop integrated models to benefit from strengths of both sides.

The integration of microscopic traffic simulation with demand models has been proposed for years. For example, Esser and Nagel (2001) studied the home-based work trips in Portland, Oregon by implementing a systematic feedback between demand generation and traffic simulation. Route and destination decisions were considered, and the objective was to match the overall commute time distribution. The demand model lacked behavioral foundation, and explicit departure time choice was not considered.
In recent years, several studies have been conducted to combine the dynamic traffic assignment (DTA) model with advanced demand models, such as activity-based models. Examples include Lin et al. (2008), Hao et al. (2010), and Miller and Roorda (2003). These studies generated activity patterns and thus travel demand from synthesized travelers, and then conducted a simulated-based DTA to decide traffic pattern that would in turn be fed back into the activity generator. This iterative process was repeated until a convergence in behavior was reached. Departure time was selected from predetermined sets for each individual. Most of these models were calibrated against aggregate travel demand instead of individual travel activities. In contrast, Flötteröd et al. (2009) explicitly modeled learning behavior in travel plan generation. In a more recent study, Flötteröd et al. (2011) further discussed the calibration of behavioral model parameters (such as the coefficients of a utility function) from traffic counts. Two previous studies by de Palma and Marchal (2002) and Ettema et al. (2005) were specifically centered around the departure time choice problem in an integrated model. The former modeled the aggregate probability of departing during each time slot for each OD pair by applying Vickery’s model, whereas the latter used a mental model to describe the effect of learning and adaptation processes on departure time decisions.

Most previous studies on integrating travel demand models with microscopic simulation are either limited in scale [e.g., only one corridor is modeled in Ettema et al. (2005)] or not detailed enough to capture the impact of traffic operation improvements (e.g., DTA-based analysis cannot model traffic signal with sufficient detail). Several pioneering studies are trying to bridge the gap by integrating activity-based travel demand models with traffic simulation on large networks. For example, a recent study (Resource Systems Group et al. 2009) funded by the Strategic Highway Research Program (SHRP2) proposed to integrate the DaySim, an activity-based travel demand simulation model first applied in Sacramento (Bradley et al. 2009) and Seattle, with a microscopic traffic simulator called the transportation analysis and simulation system. The integrated model will be applied in both Jacksonville, Florida, and Burlington, Vermont. Similarly, Bekhor et al. (2011) also developed an integrated model of activity-based demand models and dynamic traffic assignment models based on the multi-agent transport simulation framework. The integrated model was then applied in Tel Aviv, Israel. These pioneering studies are still ongoing, and more experiments with various models of both dynamic network supply (DTA or microscopic traffic simulation models) and individual network demand simulation [e.g., a learning based transportation oriented simulation system, Urbansim-activity-mobility simulator, the comprehensive econometric micro-simulator for daily activity-travel patterns, and travel activity scheduler for household agents] are required to advance understanding in this field (for more discussions in different research directions, also see Hao et al. (2010) and Horni and Axhausen (2012)).

Significant efforts have been dedicated to the normative approach, and previous studies exclusively follow the utility maximization paradigm in describing individual responses to traffic condition changes. However, some researchers, e.g., Banister (1978), criticize random utility maximization approach, suggesting it describes how travelers should behave instead of how they actually do. It is crucial to realize that people are limited to knowledge about ever-evolving network conditions and willingness and/or capacity to optimize their travel decisions. Many behavioral studies (Golledge and Stimson 1997) suggest that decisions are guided by simple rules rather than a complex evaluation process.

Therefore, this study aims at bridging the gap in both research directions by integrating a dynamic network supply model with positive/descriptive travel behavior models, both of which are calibrated with locally collected high-quality data. Although a full-fledged positive/descriptive travel supply/demand model system that covers multiple choice dimensions is preferable, this study starts with the departure time choice problem to keep the modeling process simple and leave other choice dimensions for future studies. This step-by-step modeling process also allows the flexibility to enhance the conventional four-step models by replacing some steps with agent-based travel demand submodels, which may make this transition smoother and more appealing for planning and traffic management agencies.

In this study, a framework is developed to integrate agent-based travel behavior models with large-scale microscopic traffic simulation models. Both the simulation model and behavioral models are calibrated using field data. The model is then applied to the I-270/I-495/I-95 corridor in the north Washington, DC metropolitan area for a case study. The new intercounty connector (ICC), the most significant and high-profile highway project in Maryland since the completion of the existing interstate freeway system several decades ago, also traverses through the study area. Framework set up in this study can support future analysis of transportation policy and operation strategies in this area.

The next section will describe the overall framework of the integrated model, followed by a detailed description about how the microscopic traffic simulation model is developed and calibrated. The agent-based departure time choice model is then presented and integrated with the microsimulation model. Issues during the calibration and integration phases of this large scale simulation model are discussed. The paper will be concluded after a thorough discussion of current findings and future research possibilities.

**Modeling Framework**

The framework of the proposed integrated model is presented in Fig. 1. A microscopic traffic simulator will be built with TransModeler, one of the major commercial softwares for microsimulation. Models will be constructed to build the OD matrices for the

---

**Fig. 1.** Framework of the integrated model

---

simulator based on demand data from the regional planning model. The dynamic OD matrices and parameters for the microsimulation model will then be calibrated using field traffic counts. An agent-based departure time choice model will be developed separately and then integrated to capture the behavioral reactions to network changes. The integrated model will operate iteratively until nobody is willing or able to adjust their travel decisions and a stable network condition is reached. Details for each component of the model will be discussed in the following sections.

Traffic Simulation Model

Many microscopic traffic simulators, e.g., CORridor SIMulation (Prevedouros and Wang 1999), TransModeler (Wojtowicz et al. 2011), Vissim (Gomes et al. 2004), Aimsun (Barceló and Casas 2005), or noncommercial traffic simulators (Chen et al. 2002), have been used in previous studies. These models differ in the underlying car-following models and the way in which different traveler/driver modules are implemented. No consensus is reached about the superiority of any simulators in literature. TransModeler is selected in this study because it has a well-developed interface with the geographic information system (GIS), which is important when working with different data sources.

A simulation model that includes all freeways, major arterials, most minor arterials, and some local streets along the I-270/I-495/I-95 corridor in the north Washington, DC metropolitan area was developed. It covers the central and eastern Montgomery County and the northwestern Prince George’s County of the state of Maryland, where several new developments, such as the Great Seneca Science Corridor (GSSC) in West Gaithersburg and military bases in Fort Meade, have been proposed. The intercounty connector, a new freeway currently under construction, also traverses this area. The simulated network (see Fig. 2), which includes 7,121 links and 3,521 nodes, is developed on top of satellite images provided by Google Earth, and conforms to the true geometry with high accuracy.

Overview of Model Calibration

It has always been a great challenge to calibrate large-scale microsimulation models. The reasons are three-fold: (1) it is extremely hard to obtain the dynamic OD tables on a large network; (2) complete information of traffic control strategies (such as signal timing plans) are usually not available; and (3) the number of parameters are so large that it is hard to identify which one (or ones) should be adjusted to match the field counts.

Many previous research efforts have been dedicated to the first problem, in which researchers seek to solve time-dependent OD tables based on two major information sources, static OD tables from the regional planning model and traffic counts from field observation. For example, Van Zuylen and Willumsen (1980) estimated the static OD trip table by using the entropy model. Fisk (1988) extended the entropy model under the user equilibrium assumption, whereas Liu and Fricker (1996) investigated the same problem under the stochastic user equilibrium assumption. More recent examples include Cascetta et al. (1993), Ashok and Ben-Akiva (1993), Tavana and Mahmassani (2001), and Lin and Chang (2006). Although the literature in this field is abundant, many problems have not yet been addressed. Jintanakul et al. (2011) pointed out four common difficulties for existing studies:

![Fig. 2. Simulated network in the study area](image-url)
(1) high sensitivity to seed matrix, (2) nonuniqueness in the link-path flow pattern, (3) lack of monotonicity between traffic counts and OD flows, and (4) lack of ground true OD matrices for model validation. Given these difficulties, this study does not seek a comprehensive solution with high accuracy to the dynamic OD estimation problem. Instead, a systematic approach based on the gradient projection (GP) algorithm will be proposed to derive the seed matrix for the study area, and an algorithm extended from a previous study by Nielson (1997) will be used to obtain the dynamic OD matrices. Differing from most previous studies that focus on single mode, multimodel traffic, single occupancy vehicles (SOV), high occupancy vehicles (HOV), and trucks are considered in this study. This difference is crucial because a HOV facility (I-270 HOV lane) exists in the study area, and more are under construction/consideration in the United States.

Surprisingly, the second problem related to traffic control information is not widely discussed in previous studies on large-scale traffic simulation. Ideally, the signal timing plan used in the field should also be used in simulation. However, the complete information is usually not available for a large network such as the one in this study. Because of the complexity of various intersection geometry designs in the field, it is almost infeasible to apply the state-of-the-art optimization algorithms in literature, most of which require detailed inputs about demand patterns and turning movement designs. In this study, the field signal timing plan is applied wherever it is available. There are in total 466 signalized intersections, 80 of which use the true signal plan. For the rest of them, a stylized ring-and-barrier plan (with a cycle length of 150 s typically seen from field plans) is applied, and the length of green phases is assigned proportionate to the turning demand. Future research will explore more effective approaches to implement optimized signal plans on large networks.

The third difficulty, the large number of parameters controlling both driving and traveling behavior, has been extensively studied in previous research focusing on the development of microscopic traffic simulation models. This study adopts a hierarchical calibration strategy to avoid confounding facts at different levels. Parameters for driving behavior, including free flow speed, car-following models, distribution of critical gaps, and parameters controlling lane-changing behavior (critical distance to start strategic lane changing and critical headway for the gap acceptance model) are considered in the calibration. The latter is critical for the traffic flow pattern at weaving segments. There are in total 6 free-way bifurcations in the study areas (where I-270 and I-95 interchange with I-495 Capital Beltway). Unfortunately, no fixed loop detector systems were available in the region, and no speed contour is available. Instead, the major calibration objective is to replicate the flow rate at these bottlenecks and the corridor travel time. After calibration, the largest hourly flow rate difference between the observation and model prediction is approximately 7% at the I-95/I-495 bifurcation area, and the number is smaller at other locations. These parameters for driving behavior will be fixed once the study is moved from corridor to network level calibration, in which the major concern is the network OD tables.

### Multimodal Static OD Estimation

The static OD estimation process uses the GP algorithm (Chen et al., 2002), an efficient path-based traffic assignment algorithm, to derive the OD for the study area from the regional planning model. The metropolitan Washington council of government (MWCOG) planning model is used as the basis. It includes 27,743 links, 10,505 nodes, and 2,019 traffic analysis zones (TAZ). The study area contains 162 TAZ centroids and 39 external stations through which the subnetwork is connected with the rest of the MWCOG model. For simplicity, only the morning peak period is considered in this study. The OD tables for the study area can be obtained by following these steps:

1. Assign HOV demand on the network using the GP algorithm;
2. Keep the HOV path flow, and assign truck demand by excluding HOV lanes from available links;
3. Keep the HOV and truck path flow, and assign SOV demand by excluding HOV lanes from available links;
4. Compare the longest path travel time with the shortest one between each OD pair; if any of them is beyond the predetermined threshold, go back to step 1; and
5. Compare each path with the positive flow with the external stations and centroids. If any part of the path locates in the study area, assign path flow to the corresponding OD pair and vehicle class for the subarea. Repeat this for each vehicle class and for all OD pairs.

Because the number of HOV and truck trips is smaller than that of SOV, the algorithm starts by the assignment of HOV trips. The algorithm converges very fast and reaches a relative error less than 2% within 10 iterations. The results are summarized in Table 1, showing the subarea demand is approximately one tenth of that for the entire MWCOG model.

### Calibration of Dynamic OD

Given the static OD matrices from previous studies, this study seeks to match the spatial and temporal traffic pattern with field observations by adjusting the time-dependent OD tables. Consistent with the behavioral models to be integrated in the next section, 20 min will be used as the standard time interval for each OD matrix. A demand profile based on aggregated travel demand during each 20-min time period is developed and applied to divide the initial 3-h static demand into a series of OD matrices, each of which represents OD demand by vehicle class for the corresponding 20 min. These matrices will serve as seed matrices for further adjustment.

The dynamic OD estimation algorithm is a variation of the multiple path matrix estimation (MPME) method proposed by Nielsen (1997), which is designed for static OD estimation. The proposed algorithm first evaluates the OD demand adjustment factor $\alpha_{ij,r,t}$ associated with each path $r$ between an OD pair $i, j$ and for a given time slot $t$ by using the following:

$$
\alpha_{ij,r,t} = \frac{\sum_{a \in S(i,j,r)} \zeta_{ij,r,a,t} \left( F_{a,t} + \Delta t_{ij,r,a,t} \right) / \left( F_{a,t} + \Delta t_{ij,r,a,t} \right) + \Delta t_{ij,r,a,t}}{\sum_{a \in S(i,j,r)} \zeta_{ij,r,a,t}}
$$

(1)

where $ij = \text{OD pair from origin } i \text{ to destination } j; r \in R(i,j,t)$ where $R(i,j,t)$ is set of all used paths of OD pair $i, j$ at time $t$; $S(i,j,r,t)$ is link set of path $r$ at time $t$; $F_{a,t}$ is actual link flow on link $a$ at time interval $t$; $\Delta t_{ij,r,a,t}$ = travel time from origin $i$ to link $a$ starting at time $t$. $\zeta_{ij,r,a,t}$ is given by:

$$
\zeta_{ij,r,a,t} = \begin{cases} 
1, & \text{when } a \in S(i,j,r,t) \\
0, & \text{otherwise}
\end{cases}
$$

(2)

With all factors $\alpha_{ij,r,t}$ the superscript representing iteration $n$ is omitted in Eqs. (1) and (2) to avoid unnecessary complexity in
notation], the OD demand between each $i$ to $j$ during iteration $n$ based on demand of the previous iteration is updated as follows:

$$d_{ij}^n = \frac{\sum_{r \in R(i,j)} a_{ij,r} d_{ij}^{n-1}}{\sum_{r \in R(i,j)} a_{ij,r}} \quad (3)$$

The algorithm seeks to match the observed link flow by adjusting the OD demand, which uses this specific link along its path. If more than one observation is available, an average is taken. It then updates the OD demands by considering the impacts of all paths. It differs from its static counterpart by considering the time required to reach the specific link along the path and mapping its impact to the OD demand with the corresponding departure time. Because the OD demand tables are discretized into slices of 20 min in this study, only the departure time slot of the majority of trips is considered. Therefore, only one mapping per iteration will be conducted, and stochasticity in the travel time is not considered because of the constraints of computing the efficiency of large-scale traffic simulation. Starting from the seed matrices, the traffic pattern is first simulated with the model built in TransModeler, and then the complete trip table for all paths is obtained. The time-dependent OD matrices are updated by applying Eqs. (1) and (2) for all OD pairs and for all time slots.

Field traffic counts provided by the Maryland State Highway Administration are applied to calibration the model. Freeway traffic counts can be accessed through the online interactive annual average daily traffic (AADT) locator, whereas arterial traffic counts were collected in previous turning movement studies. The total counts from 50 stations located on both freeways (15 stations) and arterial roads were used (see Fig. 3 for their locations). Consistent with previous studies, the relative root mean square error (RRMSE) was used as the convergence measure. It is defined as follows:

$$\text{rmse} = \sqrt{\frac{\sum_{i,j} (f_{ij} - \bar{f}_{ij})^2 / N}{\bar{f}}} \quad (4)$$

where $f_{ij} = \text{actual traffic flow count at station } j$, from time interval $I$; $\bar{f}_{ij} = \text{simulated traffic flow count}$; and $\bar{f} = \text{average actual traffic flow count over time and locations}$.

After 10 iterations, the overall root mean square error (RMSE) measure falls to 16.8%, whereas the RRMSE based on 15 freeway stations is 12.1%. If two stations in which traffic data are not consistent are excluded (significant jump in hourly counts between adjacent stations), the RRMSE on freeway is only 8.7%. This quality of calibration is comparable to many corridor-level simulation studies in which more accurate inputs of travel demand based on either turning movement tables or OD patterns from a limited number of entrances to the network are used. The problem investigated in this study, which considers a large-scale network and has only limited prior OD information based on the regional planning models, is much more complicated. Ideally, a better match to the field data could be reached by continuing this iterative process. However, this process is very time-consuming because of the size of the network. Therefore, the current results are satisfactory, and further calibration efforts are reserved for future research.

### Agent-Based Travel Behavior Model

#### Theoretical Framework

As previously mentioned, there is an imperative need for incorporating departure time choices into microscopic traffic operations analysis. The novel positive model employed in this study theorizes...
departure time choice as a continual search process and tracks the departure time changes of each individual user in the transportation system, and therefore is especially suitable for integration with microscopic traffic simulators, simulation-based dynamic traffic assignment models, and activity/agent-based travel demand models. The theoretical framework is illustrated in Fig. 4.

As depicted in Fig. 4, from the prior travel experience, any individual accumulates information about travel conditions corresponding to different departure times. The individual forms a certain degree of spatial knowledge, which produces subjective beliefs. As a result, at any given time, an individual has an aspiration level for potential gain, which will influence travel decisions such as when to depart.

If an individual decides not to search, repetitively learned behavior or habitual behavior is executed. If otherwise the individual decides to search, a search method (or heuristics) is employed to identify alternatives, which constitute a mapping from knowledge to one feasible alternative departure time. Then, the decision step employs decision rules to pick up an alternative option. The decision rules constitute a mapping from perceived attributes of alternatives to a choice. Decided by these rules, the individual may prefer the currently used alternative because of habit, or prefer the new alternative because of the desire for shorter travel time, or less delay, for instance. The outcome of the decision step is a provisional try behavior. The execution of the provisional try behavior provides first-hand experience about the actual travel attributes of the temporarily-chosen alternative at the time of trial.

**Knowledge and Learning**

Assume that an individual’s perception about departure time is based on utility, which is separated into I categories based on prior perception, and that the utility $u_i$ has been experienced $n_i$ times. Therefore, the individual’s knowledge about departure times can be quantified as a vector $K(n_1, \ldots, n_i, \ldots, n_I)$. According to Bayesian learning rules, the perceived weights of past observations are the same. Let vector $P(p_1, \ldots, p_i, \ldots, p_I)$ represent an individual’s subjective beliefs, where $p_i$ is the subjective probability that an additional search would lead to an alternative departure time with utility $u_i$. It is assumed that an individual’s prior beliefs follow a Dirichlet distribution. Thus, the posterior beliefs will also be a Dirichlet distribution (Rothschild 1974). This assumption is equivalent to Eq. (5), where $N$ denotes the total number of observations $[N = \sum(n_i)]$.

$$p_i = n_i/N$$  \hspace{1cm} (5)

**Search Gain versus Search Cost**

The decision to search a new alternative is based on the subjective search gain (which is hereby assumed to be based on the predicted utility improvement). It is assumed that an individual’s utility associated with the current departure time is $u$. The expected gain ($g$) in terms of utility improvement per trip from an additional search is

$$g = \sum_{i(u_i > u)} p_i \cdot (u_i - u)$$  \hspace{1cm} (6)

Eq. (6) shows how the subjective search gain is calculated as the search process proceeds in the model. To initiate the search process, the perceived search cost is needed to be compared with the search gain. The perceived search cost is assumed to be constant for the same traveler. If an individual stops searching after $n$ rounds of search, the perceived search cost for this individual must be lower than the expected search gain after $(n - 1)$ searches, such that search $n$ is meaningful, and must be higher than the expected search gain after $n$ searches such that search $(n + 1)$ does not occur. These lower and upper bounds of search cost can be calculated using Eq. (7). Let the average be the estimate of the perceived search cost ($c$)

$$c_{\text{LOW}} = g_n = \frac{u^* - u_{\text{max},n}}{n + 1}$$  \hspace{1cm} (7a)

$$c_{\text{HIGH}} = g_{n-1} = \frac{u^* - u_{\text{max},n-1}}{n}$$  \hspace{1cm} (7b)

$$c = \frac{1}{2}(c_{\text{LOW}} + c_{\text{HIGH}})$$  \hspace{1cm} (7c)

The utility function adopted here has been empirically estimated using survey data. More details about this survey can be found in Zhang and Xiong (2012). The function consists of three explanatory variables including travel time, schedule delay early, and schedule delay late. It is estimated following Small’s (1982) multinomial logit model specification. In the same survey, a number of individuals have been asked about the order by which alternative departure times are sequentially searched. This information is used to empirically derive the distribution of perceived search costs.

**Search Rules**

The search for alternatives is obviously not random, because travelers usually tend to first adapt their choice to their schedule and then avoid extremely high congestion. If then rules are selected to represent these departure time search heuristics for their capability of replicating human decision-making processes and their minimum computational resources requirement. The latter is especially important for this microscopic traffic simulation model, which is large enough to involve half a million independent decision agents.

The data set used to derive search rules was collected from commuters on the Capital Beltway. They were then asked to recall the order of the alternative departure times they have considered and actually used for their commute. The variable used in the classifier includes arrival schedule delay early (ASDE), arrival schedule delay late (ASDL), travel time (TT), and free flow travel time (TT').
The following equations define the delay variables (i.e., ASDE, ASDL, and Delay). In the equations, PAT denotes the preferred arrival time, AT denotes the actual arrival time, and Delay measures the difference between the actual travel time (TT) and the free flow travel time (TT*), which indicates the congestion level.

\[
ASDE = \max(0, PAT - AT) \quad (8a)
\]

\[
ASDL = \max(0, AT - PAT) \quad (8b)
\]

\[
\text{Delay} = \frac{(TT - TT^*)}{TT^*} \quad (8c)
\]

Various machine learning algorithms (Witten and Frank 2000) are able to derive if-then rules using the collected survey data. From four popular algorithms for deriving if-then classification rules, including C4.5 (Quinlan 1986), Programming in Statistical Modeling (Cendrowska 1987), Partial Decision Tree (PART) (Frank and Witten 1998), and RIPPER (Cohen 1995), PART is chosen for its superior cross validation accuracy on the data set. The output of PART derived from data locally collected is listed subsequently. The numbers in the bracket provide a measure of accuracy similar to that provided by the confusion matrix. The first number represents the total number of cases that fall in the situation applicable to the current rule, and the second number represents the cases that have not been accurately predicted by the current rule. Therefore, a lower ratio between the second number and the first number implies a higher level of prediction accuracy following the current set of rules. If all cases are correctly predicted by the rule, the second number is omitted (or can be noted as zero). The numbers in the bracket after decision rules carry similar meanings.

Rule 1: Search 60+ min earlier, if
[ASDL > 70] (11.0)

Rule 2: Search 30–60 min earlier, if
[45 < ASDL <= 70] (12.0/4.0)

Rule 3: Search 0–30 min earlier, if
[ASDL > 0 AND Delay > 0] (11.0/1.0)

Rule 4: Search 0–30 min later, if
[0 < ASDL <= 30 AND Delay > 40%] (4.0)

Rule 5: OR [ASDL <= 10 AND ASDE <= 40 AND Delay <= 50% AND TT <= 65] (18.0/3.0)

Rule 6: Search 30–60 min later, if
[ASDL = 0] (13.0/2.0)

Rule 7: Search 60+ min later, if
[ASDE > 75] (12.0/1.0)

Rule 8: OR[ASDE > 45 AND Delay > 10%] (6.0/1.0)

Rule 9: Otherwise, search 0–30 min earlier. (3.0)

### Decision Rules

As discussed in the previous subsection, a new departure time alternative is identified after each round of search. The alternative will be either accepted or rejected. This decision is determined by a set of decision rules used to describe departure time switching behavior. Unlike the utility maximization theory, this assumption about the decision step does not presume complete information processing and allows for historical dependencies. The decision rules are again derived from a survey experiment that was conducted in spring and summer 2011. Subjects’ actual departure time changing behaviors were observed from the survey. Decision rules were extracted using a machine learning algorithm. A more detailed explanation about the experiment can be found in Zhang and Xiong (2012).

The final decision rule set consists of six rules, as presented subsequently. The RIPPER model was chosen for its better predictive performance on the data set. The variables used in the decision rules include preferred arrival time (PAT), departure time (DT), preferred departure time (PDT), travel time (TIME), household income (INCOME), trip purpose (PURPOSE), fuel cost (FC), and toll (TC). The variable none-peak is a dummy variable that equals one if the trip occurs in off-peak hours. The symbol “denotes changes or percentage changes (i.e., alternative departure time attributes—original departure time attributes).

Switch to the alternative departure time, if
Rule 1: [\Delta\text{TIME} <= -35\% \text{ and } \Delta\text{FC} <= -8\%] (126.0/60.0)
Rule 2: [\Delta\text{TC} <= 2.5 \text{ and } \Delta\text{ASDL} <= -48\%] (43.0/13.0)
Rule 3: [\Delta\text{TC} <= 2.4 \text{ and } \text{INCOME} >= $150K \text{ and } \Delta\text{ASDL} <= -31\%] (21.0/3.0)
Rule 4: [\text{none-peak} = 1 \text{ and } \text{PURPOSE} = \text{Other and } \Delta\text{TIME} <= -8\% \text{ and } \Delta\text{ASDL} <= -53\%] (17.0/0.0)
Rule 5: [\Delta\text{ASDE} <= -20\% \text{ and } \Delta\text{TC} <= $0.7] (12.0/1.0)
Rule 6: Otherwise, continue to use the current departure time. (1203.0/181.0)

### Scenario Analysis and Model Demonstration

This scenario test demonstrates how the aforementioned departure time choice model is integrated with the traffic microsimulator and how the peak spreading is achieved by the integrated model. Overall, the model is capable of predicting individuals’ departure time changes responding to various road conditions (e.g., different congestion level and varied toll cost) for the entire network when considering developments that are planned for the next 10–30 years. This scenario considers the additional trips generated in the study area because of planned growths in Gaithersburg.

West Gaithersburg is part of the Great Seneca Science Corridor (GSSC) master plan, which is proposing the development of 4,360 acres in the center of the I-270 corridor, as shown in Fig. 5. This development will be under construction for the next 25–30 years and is to be named the Life Science Center (LSC). The LSC will add 5,750 additional residences and is projected to attract 52,500 jobs (based on existing, approved, and proposed development) (The Maryland-National Capital Park and Planning Commission 2009). This development will add a substantial amount of automotive volume to the surrounding network. Therefore, for demonstration purposes, in this study it is assumed that the West Gaithersburg developments will generate 30% more travel demand. The additional demand will change the congestion pattern in the region, and thus motivate some travelers to adjust their departure time accordingly. The updated OD tables will be simulated in TransModeler to generate a new traffic pattern, which will in turn be fed back into the agent-based peak spreading model until convergence is achieved. Other scenario features are listed as follows:

- The base scenario adopts an original dynamic travel demand (20-min interval);
- The tested scenario assumes 30% higher travel demand generated from the LSC study area (i.e., centroid 77, 78, 168, and 169 in Fig. 5) based on the base case; and
- Peak spreading effect for traffic going through this area is also considered and tested. As an explorative process, this study only considers the closest external station (centroid 170) while keeping a departure time pattern for demand from other more remote TAZ centroids unchanged. This assumption can be easily relaxed later.

As previously mentioned, for comparison purposes, the scenario has tested the response of both the demand generated from the LSC study area and the travelers going through this area. Assuming that individuals’ preferred arrival time is their individual arrival time in...
the base case, travelers who are now experiencing higher schedule delay and longer travel time because of the demand growth will consider alternative departure schedules. After seven iterations, the departure pattern becomes stable, and the results are summarized in Fig. 6.

Depicted in Fig. 6(a), as the congestion deteriorates, a number of travelers originally traveling between 8 a.m. and 9 a.m. leave earlier to avoid being late. This forms a new peak in the demand pattern. Evidently, travelers from the LSC zones begin to consider shoulder hours (time periods that are slightly earlier or later than the a.m. peak period) as their alternative departure time. Additionally, as the demand grows, the congestion in the LSC area is expected to be exaggerated. Thus, the external travelers traversing this region during peak hours are also likely to consider switching departure times to avoid possible delay. As illustrated in Fig. 6(b), the model is capable of capturing this behavioral change as well. Although the travel demand remains unchanged, the travelers going through the study area also adjust their departure schedule to adapt to the new travel condition in the LSC region.

This scenario test demonstrates the capability of the integrated model in representing individuals’ behavioral response to various travel conditions. The results overall illustrate a reasonable departure time shift in response to the assumed more congested situation. As displayed in Fig. 6, approximately 6.2% of the peak-hour trips have switched to depart either earlier or later to avoid higher congestion and delay. Moreover, the original 3-h peak period (6 a.m. to 9 a.m.), as predicted by the model, has spread out to cover a wider time period [approximately from 5:30 a.m. to 9:30 a.m.; Fig. 6(a)].

Fig. 7 illustrates the evolution of the congestion level. The x-axis represents the number of iterations. In each iteration, travelers would reevaluate the network conditions based on their previous travel experience and make decisions about whether to try an alternative departure time based on the theory proposed in the “Agent-Based Travel Behavior Model” Section. As more alternative departure time options are explored, drivers get more familiar with the

![Fig. 5. LSC study area](image-url)
time-dependent travel condition, and better decisions are made to avoid extreme delay during the peak hours. The overall delay is gradually mitigated.

Conclusions

This study integrates agent-based travel demand models with large-scale traffic simulation to capture the behavioral responses to traffic condition changes. Although many challenges still exist for the calibration of a large-scale microscopic traffic simulation model, the intuitive iterative calibration process proposed in this study works reasonably well. However, many factors could affect the simulation results, including parameters controlling driving behavior, routing, signal control plans, and parameters related to travel choices. This study calibrated a subset of these modeling components separately with traffic data and high-quality behavioral data. However, the model also needs to replicate reality and be calibrated and validated based on aggregate simulation results. This process requires a deep understanding of the sensitivity of overall simulation results to each component, modeling technique to separate different effects, and high-quality data to support the calibration. Future research will address these challenges.

The proposed agent-based departure time choice model differs from the utility-based modeling paradigm widely used in previous research. The proposed rule-based model is intuitive to interpret, and their implementation only involves a series of logic evaluation. By integrating more choice dimensions such as mode and destination, this agent-based modeling approach provides an easier way to maintain consistency in simultaneous decisions of different choices [the preferred arrival time of an individual agent could affect both departure time and mode choice (such as HOV for higher reliability) in demand modeling and route choice in dynamic traffic assignment process]. With future research efforts, all behavior rules concerning demand and routing will be empirically estimated from longitudinal behavior data with considerations for learning, expectation, and behavior adjustments. Examples in this study demonstrate its potential in capturing various behavioral reactions to network improvement and policy initiatives. Comparison between the utility-based models and the agent-based travel supply/demand model system will be considered in future research.

The integration of agent-based travel behavior models with microsimulation models also provides a low-cost resource for capturing individual experience and network conditions, and thus makes such models operational in policy analysis. This framework also provides a valuable tool for the evaluation of new transportation infrastructure, such as the ICC corridor currently under construction, and its operation strategies. More research efforts are needed to explore the full potentials of the integrated model.

Acknowledgments

This research was funded partially by the Maryland State Highway Administration, Federal Highway Administration Exploratory Advanced Research Program, and the Center for Integrated Transportation Systems Management at the Univ. of Maryland. The views in this paper do not necessarily reflect the views of the funding agencies. The authors are solely responsible for all statements in the paper.

References


