ENABLING CONGESTION AVOIDANCE AND REDUCTION IN THE MICHIGAN-OHIO TRANSPORTATION NETWORK TO IMPROVE SUPPLY CHAIN EFFICIENCY: Freight ATIS

PROGRESS REPORT

October 1, 2008

Principal Investigators:

Wayne State University
Industrial & Manufacturing Engineering
4815 Fourth Street, Detroit, MI 48202, USA

- Ratna Babu Chinnam (PI), Associate Professor
  Tel: 313-577-4846; Fax: 313-578-5902
  E-mail: r_chinnam@wayne.edu
- Alper E. Murat (Co-PI), Assistant Professor
  Tel: 313-577-4846; Fax: 313-578-5902
  E-mail: r_chinnam@wayne.edu

University of Detroit Mercy
School of Business Administration

- Gregory Ulferts (Co-PI), Professor
  Tel: 313-993-1219; Fax: 313-993-1673
  E-mail: ulfertgw@udmercy.edu

Submitted to:

Dr. Leo Hanifin

Director, MIOH University Transportation Center

University of Detroit Mercy

Table of Contents

SUMMARY

Our project has six major mile-stones for the first two years:

i. Mile-stone #1 [Year 1]: Data collection for MI-OH road network structure and historical incident data form MDOT, ODOT, and other agencies.
ii. Mile-stone #2 [Year 1]: Developing road network models representative of major freight transportation routes.
iii. Mile-stone #3 [Year 1]: Developing static re-routing optimization model and implementation for a limited set of scenarios.
iv. Mile-stone #4 [Year 2]: Develop dynamic re-routing optimization models and efficient heuristic solution methods.
v. Mile-stone #5 [Year 2]: Develop computationally efficient and effective parametric incident delay models.
vi. Mile-stone #6 [Year 2]: Develop extensive scenarios based on vehicle, incident, distribution strategy, and road network characteristics, etc.

The Research Team has made outstanding progress with respect to all these milestones during the last two years of funding.

Mile-stone #1: We have approached the data collection goal from multiple directions. On the ***network structure*** (network topology, design parameters, link characteristics) side, we have developed collaborative relationships with managers of two MDOT projects.[3] Currently, we are in the process of acquiring these models and a license for Paramics software. Through these calibrated micro-simulation models of the SouthEast-Michigan corridor, we will be able to test our dynamic routing decisions as well as incident (i.e. accidents, breakdowns) delay models in the next two years.

For Southeast-Michigan corridor ***link velocity data***, we have collaborated with the MITS Center and signed a data-sharing agreement with Traffic.com. To date we have had multiple meetings with MITS Center to develop a better understanding for their traffic monitoring system (for Southeast-Michigan highways) and have also received data representing several months of traffic flow (such as velocity, occupancy) for the southeast Michigan highways. We have analyzed this data to improve the quality of the models being developed for dynamic vehicle routing decision support when operating with access to Advanced Traveler Information Systems (ATIS) information. For instance, through our analyses, we identified the need for representing each link's congestion with a different number of states (and not force all links to be modeled with two states – i.e., congested and uncongested). Accordingly, we have refined the recurrent congestion state modeling by employing the Gaussian Mixture Model clustering method for automated detection of number of states and state velocity thresholds. In addition to MITS Center data, we now have access to Traffic.com's sensor database covering majority of high-ways in the Southeast-Michigan corridor. These datasets (and the networks resulting from them) are playing a critical role for evaluating and refining our dynamic routing algorithms. For ***incident data*** collection, we are collaborating with the MITS Center. We received several months of incident data from Monroe Pendelton and Mark Burrows of MITS Center. This dataset allowed us to initiate modeling of non-recurring congestion (incidents and special events) in routing applications. In addition to MITS

[3] METSIM and Gateway projects aim to develop tools for strategic and operational planning of highway projects through micro-simulation models. Both projects are utilizing the Paramics Suite software package for traffic simulation.

Center data, Traffic.com also has an extensive archive of incident data which we are currently using to develop parametric incident delay models, models of particular interest to SEMCOG.

Mile-stone #2: We have initially constructed a simple hypothetical road network simulator to build, test, and validate our algorithms in Matlab. This simulator allowed us to experiment with various network, velocity and incident scenarios. In the second year, we have developed a Southeast-Michigan road network model that covers the sensors from both MITS Center and Traffic.com. We constructed these networks using archived historical traffic ITS data provided by our research partners, MITS-Center and Traffic.com. One instance from this network encompasses main freeways and arterials extending from the intersection of I-94 and I-275 to the intersection of I-696 and I-75. In addition to network construction, we developed a data extraction and network configuration tool that allowed us to automate the loop sensor velocity and incident data extraction from the ITS databases. This tool takes in the origin-destination coordinates as inputs to identify and locate loop sensors. Subsequently, this tool first extracts sensor velocity data and incident data and then configures our routing models by determining such model inputs as link travel time distributions by departure time of day. This tool encompasses data extraction, filtering, and cleaning procedures and is based on MS Access database with Matlab interface for efficient network configuration and algorithmic implementation.

Mile-stones #3 and #4: The vast majority of our efforts in the first year went toward developing static and dynamic routing algorithms that enable congestion avoidance and travel time reduction in commercial cargo transportation networks. We have gone beyond Mile-stone #3 in that our emphasis was not just static but both static and dynamic algorithms. In the second year (Mile-stone #4), we have concentrated our efforts on improving the efficiency of the exact dynamic routing algorithm and developing heuristic algorithms.

We have followed a two-phase approach in developing routing algorithms for the base case (routing for one-to-one shipment). In the first phase we focused on developing ***Stochastic Dynamic Programming (SDP)*** based algorithms for optimal routing under ATIS. While SDP algorithms yield optimal routing policies, they are not computationally efficient. However, we need these solutions for testing and benchmarking the effectiveness of fast heuristic algorithms to be developed over the course of this multi-year project. In the second phase, we have adapted the more computationally efficient ***AO* algorithms*** for developing optimal policies. Previously, AO* algorithms have been applied to stochastic routing problems with single congestion states. In our implementation, we have extended this heuristic algorithm to include congestion states as in SDP. The efficiency of the AO* algorithm significantly depends on the heuristic estimate of a lower bound representing travel cost to the destination node. In the absence of a quality lower bound, the SDP's performance is comparable with the AO*. Hence, our algorithmic framework currently uses both of these methodologies. To improve the efficiency of the Stochastic Dynamic Programming algorithm we developed a variety of smart routines for different tasks.[4]

Mile-stone #5: Towards the end of the first year, we developed preliminary incident delay models. In the second year, we have extended and refined these models by calibrating according to the incident data obtained from MITS Center and Traffic.com.

[4] For instance, we developed an efficient pre-processing model (based on Dijkstra's algortihm) for extracting a sub-network from the full network given an O-D pair. We have further transformed the state space representation for more efficient data access. In terms of presentation and ease of use, we are planning to implement our algorithms using Google Maps API interface beginning with the third year.

Our effort has thus gone into developing compact yet effective parametric models representing real-world incident delay signatures. In addition, we have also extended the algorithmic framework by incorporating more realistic "non-recurring congestion" modeling and exploitation logic into the algorithms. Given that nearly 50% of all traffic congestion and about 50-60% of non-recurring delays in urban areas are attributable to incidents[5], and that vast majority of dynamic routing algorithms reported in the literature do not exploit this information, this extension has greatly enhanced the fidelity of our dynamic routing algorithms in reducing trip completion times. Extensive evaluations of our SDP algorithms on hypothetical networks revealed significant reductions in trip completion times in comparison with deterministic algorithms and static stochastic algorithms that do not account for non-recurring congestion information. The current version of our incident model is a parametric multiplicative model for the incident delay and accounts for the real-time traffic congestion, incident duration, incident severity, incident response. In the second year, we have further extended our incident model by coupling the parametric delay model with an incident clearance Markov model. The incident clearance model is a non-stationary Markov chain model in which the incident clearance probability increases with the duration of the incident. Our incident model is currently integrated within the recurring congestion modeling and algorithmic framework and further improvements (shockwave propagation, traffic behavior) are in our project plan.[6]

Mile-stone #6: We have developed the road-network model for the Southeast-Michigan region and identified some set of origin-destination pairs for major freight routes. On these routes, we have extracted the road-network recurring and non-recurring congestion data sets and calibrated these links accordingly. Beginning with the second half of the second year, we have been implementing our static and dynamic models and algorithms in these major freight routes and comparing the performance differences between typical base-line routing algorithms and our stochastic dynamic routing algorithms. In the last quarter, we will be meeting with our collaborators Ford MP&L, UPS and C.H. Robinson to obtain their current distribution strategies (frequency of shipments, origin-destination pairs, vehicle, driver and freight characteristics) on this network. Via these efforts, we will develop our extensive scenarios for testing and comparison of our models and algorithms.

Thus far, our emphasis has been on the routing of a vehicle delivering from an origin to a destination. However, milk-run deliveries are critical to some of the logistics companies and our partners. In the remainder quarter and during next year, we will extend our models and routing algorithms to begin to support milk-run based deliveries with and without time-windows. We realize that the computational complexity and other challenges associated with this problem and thus seek efficient yet effective heuristic approaches to cope with these challenges.

The rest of the report describes our efforts and results to date in more detail. It is organized as follows: Section I describes our efforts, results, and next steps in developing dynamic routing algorithms in great detail. Section II outlines our dissemination efforts. Section III provides references.

---

[5] According to American Association of State Highway and Transportation Officials (www.aashto.org) and National Traffic Incident Management Coalition (www.timcoalition.org)

[6] Further improvements to our incident model include accounting for the shockwave-based congestion state transitions and propagations for accurate congestion dependency between links affected by incidents. Another improvement under consideration for our incident model is to develop models for anticipating changes in regular traffic patterns in response to an incident for more efficient avoidance and/or navigation.

I: DEVELOPING STATIC AND DYNAMIC ROUTING ALGORITHMS UNDER ATIS AND REAL-TIME INFORMATION

The overall goal here is to develop effective static and dynamic routing algorithms for congestion avoidance and reduction for commercial cargo carriers given real-time information regarding recurring and non-recurring congestion by Advanced Traveler Information Systems (ATIS). Vast majority of our R&D efforts over the past two years targeted this goal. We have extensively reviewed the literature on state-of-the-art static and dynamic routing algorithms, tested promising algorithms, recognized their strengths/weaknesses, and identified means to improve their performance.

I.1. Introduction

Just-in-time supply chains require reliable deliveries. However, travel times on road networks are unfortunately stochastic in nature. This randomness might stem from multiple sources. One of the most significant sources is the high volume of traffic due to commuting. This kind of traffic congestion is called *recurrent congestion* for it usually occurs at similar hours and days on a given network. The most used approach to deal with recurrent congestion is building 'buffer time' into the trip, i.e. starting the trip earlier to end the trip on time. However, these buffers significantly increase driver and equipment idle time (i.e., reduce utilization). For example, the 2007 mobility report notes that congestion causes the average peak period traveler to spend an extra 38 hours of travel time and consume an additional 26 gallons of fuel, amounting to a cost of $710 per traveler per year (Urban Mobility Report 2007). Another disturbance to traffic networks stems from 'irregular events'—crashes, stalled vehicles, work zones, weather problems and special events—that cause unreliable travel times and also contribute significantly to the overall congestion problem. The resulting congestion is labeled *non-recurrent congestion* for the time and frequency of this kind of congestion is unpredictable. The combined effect of recurrent and non-recurrent congestion is very significant, making trip travel times increasingly unreliable and increasing travel times by as much as 50% in some highly congested urban areas (Urban Mobility Report 2007). The authoritative report also concludes that congestion continues to worsen in American cities of all sizes, creating a $78 billion annual drain on the U.S. economy.

Intelligent Transportation Systems (ITS) that collect and provide real-time traffic data are now available in most urban areas and traffic monitoring systems are beginning to provide real-time information regarding incidents. In-vehicle communication technologies, both GPS and non-GPS based, are also enabling drivers access to this information, facilitating vehicle routing and re-routing for congestion avoidance. We are proposing dynamic vehicle routing models that use ITS traffic information to avoid both recurrent and non-recurrent congestion in stochastic transportation networks.

Our basic model is a non-stationary stochastic shortest path problem and we present savings results for several network scenarios. We developed a dynamic vehicle routing model based on Markov decision process (MDP) formulation. The state set of the MDP is based on the position of the vehicle, the time of the day, and the traffic congestion states of the roads. Recurrent congestion states of the roads and their transition patterns are determined using historic and real-time traffic data from M-DOT's ITS (MITS) Center. In particular, states are determined using Gaussian mixture model (GMM) based clustering. To address issues of 'curse of dimensionality' common to MDPs and the recognition that information from distant links are unreliable and less likely to influence 'optimal' path selection, we formulated the MDP state space such that only the roads/links that are in proximity to the vehicle affect local decisions.

Our dynamic routing models also account for non-recurring congestion stemming from incidents. Our incident models attempt to address two questions: 1) Estimate the affect of incident on travel time (incident-induced travel time delay) and 2) Estimate the incident clearance time (incident clearance time). We estimate incident-induced link travel time delay using a decay function based on incident severity and duration parameters. Time required to clear the incident and restore the traffic is usually defined as incident clearance time and most of the delay due to incident is experienced during this period. We model the incident-clearance process using a Markov chain with an eventual absorbing state of incident clearance.

Given that a road network may encounter both types of congestion concurrently, our dynamic routing models integrally account for both types of congestion and their interactions. In summary, our contributions are fourfold: 1) More accurate representation of recurring congestion (i.e., identification of congestion states and their transition patterns) through GMM based clustering, 2) Efficient modeling of recurring congestion (through limited-look ahead modeling), 3) Integrated modeling of recurring and non-recurring congestion for dynamic routing, and 4) Nonrecurring congestion modeling representative of realistic incident delays.

The rest of this section is organized as follows. Relevant literature is reviewed in section I.2. Section I.3 establishes a dynamic vehicle routing model for the problem. In section I.4, experimental settings and results are presented. Section I.5 provides some concluding remarks and next steps.

I.2. Literature Review

I.2.1. Shortest Path Problem

In the classic deterministic shortest path (SP) problem, the cost of traversing a link is deterministic and independent of the arrival time to the link. The stochastic SP problem (S-SP) is a direct extension of this deterministic counterpart where the link costs follow a known probability distribution. In S-SP, there are multiple potential objectives, the two most common ones being minimization of the total expected cost and maximization of the probability of being lowest cost (Sigal et al 1980). To find the path with minimum total expected cost, Frank (1969) suggested replacing link costs with their expected values and subsequently solving as a deterministic SP. Loui (1983) showed that this approach could lead to sub-optimal paths and proposed using utility functions instead of the expected link costs. Eiger et al. (1985) showed that Dijkstra's algorithm (Dijkstra, 1959) can be used when the utility functions are linear or exponential.

Stochastic SP problems are referred as stochastic time-dependent shortest path problems (STD-SP) when link costs are time-dependent. Hall (1986) first studied the STD-SP problems and showed that the optimal solution has to be an adaptive decision policy (ADP) rather than a single path. In an adaptive decision policy (ADP), the node to visit next depends on both the current node and the time of arrival at that node, and therefore the standard SP algorithms cannot be used. Hall

(1986) employed the dynamic programming (DP) approach to derive the optimal policy. Bertsekas and Tsitsiklis (1991) proved the existence of optimal policies for (STD-SP) when arc costs are positive and/or negative. Later, Fu and Rilett (1998) modified the method of Hall (1986) for problems with link costs as continuous random variables. They showed the computational intractability of the problem based on the mean-variance relationship between the travel time of a given path and the dynamic and stochastic travel times of the individual links. They also proposed a heuristic in recognition of this intractability. Bander and White (2002) modeled a heuristic search algorithm AO* for the problem and demonstrated significant computational advantages over dynamic programming, when there exists known strong lower bounds on the total expected travel cost between any node and the destination node. Fu (2001) discussed real-time vehicle routing based on the estimation of immediate link travel times and proposed a label-correcting algorithm as a treatment to the recurrent relations in DP. Waller and Ziliaskopoulos (2002) suggested polynomial algorithms to find optimal policies for stochastic shortest path problems with one-step link and limited temporal dependencies. Gao and Chabini (2006) designed an ADP algorithm and proposed efficient approximations to their algorithm by using the value of information in a both time and link dependent stochastic network. An alternative routing solution to the adaptive decision policy is a single path satisfying an optimality criterion. For identifying paths with the least expected travel time, Miller-Hooks and Mahmassani (2000, 2003) proposed a modified label-correcting algorithm.

I.2.2. Recurrent Congestion Modeling and Real-Time Information

All of the studies on STD-SP assume deterministic temporal dependence of link costs, with the exception of Waller and Ziliaskopoulos (2002) and Gao and Chabini (2006). In most urban transportation networks, however, the change in the cost of traversing a link over-time is stochastic and there are very few studies addressing this issue. Most of these studies model this stochastic temporal dependence through Markov chain modeling and propose using the real-time information available through ITS systems for observing Markov states. In addition, all of these studies assume that recourse actions are possible such that the vehicle's path can be re-adjusted based on newly acquired congestion information. Accordingly, they identify optimal adaptive decision policies. Psaraftis and Tsitsiklis (1993) is the first study to consider stochastic temporal dependence of link costs and to suggest using online information en route. They considered an acyclic network where the cost of outgoing links of a node is a function of the environment state of that node and the state changes according to a Markovian process. They assumed that the link's state is learned only when the vehicle arrives at the source node and that the states of nodes are independent. They also proposed a DP procedure to solve the problem. Polychronopoulos and Tsitsiklis (1996) consider a problem when recourse is possible in a network with dependent undirected links and the link costs are time independent. They proposed a DP algorithm to solve the problem and discussed some non-optimal but easily computable heuristics. Azaron and Kianfar (2003) extended Psaraftis and Tsitsiklis (1993) by evolving the states of current node as well as its forward nodes with independent continuous-time semi-Markov processes for ship routing problem in a stochastic but time invariant network. Kim et al. (2005a) studied a similar problem as in Psaraftis and Tsitsiklis (1993) except that the information of all links are available real-time. They assume that a link can be in two states, either congested or uncongested, based on the link velocities. They proposed a DP formulation where the state space includes states of all links, time, and the current node. They reported substantial cost savings from a computational study based on the Southeast-Michigan's road network. They however stated that the state space of the proposed formulation becomes quite large making the problem intractable. To address the intractable state-space issue, Kim et al. (2005b) proposed state space reduction methods.

I.2.3. Non-recurring Incidents and Incident Clearance

All of the shortest-path studies reviewed consider the stochastic link costs which are mostly attributable to recurring congestion. However, about 55% of all traffic congestion is attributable to non-recurring incidents such as accidents, bad weather, work zones and special events (FHWA report, 2004). Incident-induced delay time estimation models are widely studied in the transportation literature. These models can be categorized into three based on the approaches followed: shockwave theory (Wirasinghe, 1978; Al-Deek et.al, 1995; Mongeot and Lesort, 2000), queuing theory (Cohen and Southworth, 1999; Olmstead, 1999; Henderson et al., 2004, Baykal-Gursoy et.al. 2008), and statistical (regression) models (Gaver, 1969; Lindley, 1987; Giuliano, 1989; Garib et.al., 1997). All of these modeling approaches have certain requirements such as loop-sensor data or assumptions regarding traffic/vehicle behavior. For instance, the shockwave theory based models require extensive loop sensor data for accurate positioning and progression of the shockwave. Both the queuing and shockwave theory based models require assumptions about the vehicle arrival process. Regression models often cannot handle missing data without compromising accuracy.

In all these three types of models, the delay due to an incident is often a function of incident duration. Thus, the estimation of incident duration is fundamental and there are various distributions suggested. Gaver (1969) derived probability distributions of delay under flow stopping. Gamma and exponential distributions are also suggested as good representations of incident duration distribution (Noland and Polak, 2002).

Integrated Modeling of Recurrent and Non-recurring Congestion

Modeling incident delay in conjunction with vehicle routing is in its nascence. Ferris and Ruszcznski (2000) present a problem in which links with incidents fail and become permanently unavailable rather than incidents being cleared after some time. They model the problem as an infinite-horizon Markov decision process (MDP). Thomas and White (2007) consider the incident clearance process and adopt the models in Kim et al. (2005a) for routing under non-recurring congestion. However, they do not account for recurring congestion and assume link costs are time-invariant and deterministic. They model the incident clearance time as a non-stationary Markov chain with transition probabilities following a Weibull distribution with an increasing instantaneous clearance rate. To model incident-induced delay, they multiply the link's cost by a constant and time-invariant scalar.

I.3. Dynamic Routing Model

I.3.1 Recurrent Congestion Modeling

Let the graph $G=(N,A)$ denote the road network where N is the set of nodes (intersections) and $A \subseteq N \times N$ is the set of directed arcs between nodes. For every node pair $n',n \in N$, there exists an arc $a \equiv (n,n') \in A$ if and only if there is a direct road that permits traffic flow from node n to n'. Given an origin-destination (OD) node pair, we treat the trip planner's decision problem to be that of selection of an arc at each decision node such that the expected total trip travel time is minimized. We denote the origin and destination nodes with n_0 and n_d, respectively. We formulate this problem as a finite horizon Markov decision process (MDP) where the travel time on each arc follows a non-stationary stochastic process.

An arc $a \equiv (n, n') \in A$ is labeled as observed if its real-time traffic data (e.g., velocity) is available through the real-time traffic information system. An observed arc's traffic congestion can be in $r \in Z^+$ different states at time t. These states represent arc's congestion level and are associated with the real-time traffic velocity on the arc. We begin with discussing how to determine an arc's congestion state given the real-time velocity information and defer the discussion on estimation of the congestion state parameters to Section 1.4.2. Let $c_a^{i-1}(t)$ and $c_a^i(t)$ for $i = 1, 2, \ldots, r$ denote the cut-off or threshold velocities used to determine the state of arc a given $v_a(t)$, e.g., the velocity at time t on arc a. We further define $s_a^i(t)$ as the i^{th} traffic congestion state of arc a at time t, i.e. $s_a^1(t) = \{\text{Congested}\} = \{1\}$ and $s_a^r(t) = \{\text{Uncongested}\} = \{r\}$. Congestion state, $s_a^i(t)$ of arc a at time t can then be determined as:

$$s_a(t) = \{i, \text{if } c_a^{i-1}(t) \le v_a(t) < c_a^i(t)\}$$

Consistent with much of the literature, we assume that the congestion state of an arc evolves according to a non-stationary Markov chain. We also assume that the link traverse time given the congestion state follows a statistical distribution, in particular Gaussian, the parameters can be dependent on time of travel.

Let us suppose that all arcs in the network are observed and let $S(t)$ denote the traffic congestion state vector for the entire network, i.e., $S(t) = \{s_1(t), s_2(t), \ldots, s_{|A|}(t)\}$ at time t. For presentation clarity, we will suppress the (t) in the notation whenever time reference is obvious from the expression and when we suppress the arc subscript then it represents the whole network. We denote a realization of $S(t)$ by $s(t)$.

It is assumed that arc traffic congestion states are independent from each other and exhibit first-order Markovian property. In order to estimate the pattern of state transitions for any arc at time t, we model the distribution of arc velocities from two consecutive periods, t and $t+1$, jointly. Accordingly, time-dependent single-period state transition probability from state $s_a^i(t)$ to state $s_a^j(t+1)$, denoted by $\alpha_a^{ij}(t)$, is estimated from this joint velocity distribution as follows:

$$\begin{aligned}\alpha_a^{ij}(t) &= P\left(s_a(t+1) = j \mid s_a(t) = i\right) \\ &= \frac{P\left(\left(c_a^{i-1}(t) \le V_a(t) < c_a^{i-1}(t)\right) \cap \left(c_a^{j-1}(t+1) < V_a(t+1) < c_a^{j-1}(t+1)\right)\right)}{P\left(c_a^{i-1}(t) \le V_a(t) < c_a^{i-1}(t)\right)}\end{aligned}$$

Let $T_a(t,t+1)$ denote the matrix of state transition probabilities for arc a from time t to time $t+1$ ($T_a(t,t+1)=\left[\alpha_a^{ij}(t)\right]_{ij}$). Given the assumption that the congestion state transitions for any given arc follow a first-order Markov chain, $P\{s_a(t)\mid s_a(t-1),s_a(t-2),...,s_a(t_0)\}=P\{s_a(t)\mid s_a(t-1)\}$. Given the assumption that arc congestion states are independent of other arcs' states, $P\{s_{a_1}(t)\mid s_{a_2}(t),s_{a_1}(t-1)\}=P\{s_{a_1}(t)\mid s_{a_1}(t-1)\}$. Let us suppose that the network is in state $S(t)$ at time t and we want to find the network state $S(t+\delta)$ where δ is a positive integer number. Given our arc independence assumption, the joint state transition probability is simply the product for the state transition probabilities for individual arcs:

$$P\left(S(t+\delta)\mid S(t)\right)=\prod_{a=1}^{|A|}P\left(s_a(t+\delta)\mid s_a(t)\right)$$

The δ period transition probability for each arc can be calculated from the standard Kolmogorov equation:

$$T_a(t,t+\delta)=\left[\alpha_a^{ij}(t)\right]_{ij}\times\left[\alpha_a^{ij}(t+1)\right]_{ij}\times...\times\left[\alpha_a^{ij}(t+\delta)\right]_{ij}$$

With the assumption that arc velocities follow normal distribution, the time to traverse an arc can be modeled as non-stationary normal distribution. We further assume that the arc's travel time depends on the congestion state of the arc at time of departure (equivalent to arrival time whenever there is no waiting). It can be determined according to the corresponding normal distribution:

$$\delta(t,a,s_a)\sim N\left(\mu_a(t,s_a),\sigma_a^2(t,s_a)\right)$$

where, $\delta(t,a,s_a)$ is travel time on arc a at time t with congestion state $s_a(t)$, and $\mu_a(t,s_a)$ and $\sigma_a(t,s_a)$ are the mean and standard deviation of travel time on arc a at time t, respectively, under congestion state $s_a(t)$.

We assume that the objective of dynamic routing is to minimize the expected travel time based on real-time information. The nodes (intersections) of the network represent decision points where a routing decision can be made. Since our algorithm is also applicable for a network with incident, in the next section, we will discuss incident modeling and then integration of recurring congestion and incident driven non-recurring congestion.

I.3.2 Incident Modeling

The incident modeling section addresses two questions: 1) With what probability the incident will be cleared and 2) How much delay the driver will experience on an incident arc. First we will explain the logic behind incident clearance modeling, then, our incident delay function, and finally, we will integrate incident modeling with the dynamic routing problem.

I.3.2.1 Incident Clearance

We assume that incidents undergo a series of stages (e.g., fresh incident, response team arrives to incident scene, partial incent clearance …) and that these stage transitions exhibit the Markov property. For ease of presentation, we limit the number of states to two $\{Incident, Clear\}$. We also assume that information regarding these states is available to the driver instantaneously without any delay. Let, $x_a(t)$ be a random variable representing the incident state of arc a at time t, i.e. $x_a(t) = \{Incident, Clear\} = \{1,2\}$. Then, the status of the observed arcs is given by the vector $X(t) = \{x_1(t), x_2(t), ..., x_{|A|}(t)\}$. Although the transitions occur in continuous time, we discretize the transition times for consistency with our whole model. Since the incidents that cause congestion appear infrequently and are cleared, such as accidents, we assume that incident carry an absorbing state 'clear'.[7] Furthermore, if an incident occurs en route, we may simply re-optimize the path treating the approaching node of the driver is the starting node. With these assumptions, a one-step incident transition matrix can be setup as follows:

$$IT(t, t+1) = \begin{bmatrix} 1-\gamma(t) & \gamma(t) \\ 0 & 1 \end{bmatrix}.$$

where $\gamma(t)$ is the one-step transition of a link state from incident state to cleared state at time t. Note that the transition probability element is independent from arc.[8] Intuitively, the longer the presence of a particular incident the higher the likelihood it will clear in the next interval. This suggests that $\gamma(t) > \gamma(t+1)$. The δ period state transition probability of incident clearance can be calculated once again using the Kolmogorov equation:

$$IT(t, t+\delta) = \begin{bmatrix} 1-\gamma(t) & \gamma(t) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1-\gamma(t+1) & \gamma(t+1) \\ 0 & 1 \end{bmatrix} \times ... \times \begin{bmatrix} 1-\gamma(t+\delta) & \gamma(t+\delta) \\ 0 & 1 \end{bmatrix}.$$

I.3.2.2 Incident-Induced Delay

We assume incident delay function, $\Theta(\kappa, \rho, \varphi, x)$, is based on incident severity (κ), duration/stage (ρ), incident response (φ), and incident realization (x) of the arc. Our incident delay model is a multiplicative model in that $\Theta(\cdot)$ represents the factor by which the arc travel time under same conditions (congestion state and the time) will be increased. Specifically, given the arc travel time without incident, $\delta(t, a, s_a, x_a = 2)$ and incident parameters (κ, ρ, α), the arc travel time under incident is expressed as:

$$\delta(t, a, s_a, x_a = 1) = (1 + \Theta(\kappa, \rho, \varphi))\, \delta(t, a, s_a, x_a = 2)$$

In the expected sense the travel time of arc a at time t is:

---

[7] This is not appropriate for settings such as lane closures in work zones and can be handled using an incident Markov chain with a single persistent 'incident' state.

[8] Given data limitations (i.e., infrequent occurrence of identical incidents at same location), it is difficult to create arc dependent incident parameters statistically.

$$\delta(t,a,s_a,x_a)=\sum_{x=1}^{2}P(x_a(t)=x)\big(1+\Theta(\kappa,\rho,\varphi,x)\big)\delta(t,a,s_a,x_a=2)$$

where $\Theta(\kappa,\rho,\varphi,x=2)=0$. Note that incident duration/stage, ρ, depends on the duration between incident occurrence time (t_{inc}^{0}) and arrival time to the arc (t), i.e. $\rho=\{t-t_{inc}^{0}:t\geq t_{inc}^{0}\}$. We make the following assumptions for the incident delay function:

1. Incident delay is only experienced on the incident arc (no propagating delay effect in the remainder of the network)
2. Incident delay function, $\Theta(\cdot)$, is non-increasing function of ρ
3. Incident delay function, $\Theta(\cdot)$, is such that total delay associated by deciding to wait at a node, waiting time plus the incident delay, is not less than the case without waiting.
4. Incident delay function is a multiplicative factor which amplifies the incumbent arc travel time. This factor is independent of the arc.

In practice, the incident effect propagates in the network in the form of a shockwave after a certain duration following the incident. In this incident model, our goal is investigate the impact of incidents in the travel time and therefore we chose to focus on the most important ingredient, namely delay on in the incident arc. Hence, we believe *Assumption 1* is acceptable under certain scenarios. One scenario is where the incident duration is not long enough that vehicles divert to alternative arcs or the capacity of alternative arcs is sufficiently large to accommodate the diversion without any change in their congestion state. *Assumption 2* is based on the fact that there is an incident response and clearance mechanism which mediates the incident delay over time. *Assumption 3* is consistent with our network and travel time assumptions where we assume that waiting at a node (or on an arc) is not permitted and/or does not provide travel time savings.

Multiplicative model assumption (*Assumption 4*) is reasonable since the travel time delay of a particular incident depends on the both the incident characteristics and incumbent travel time on the arc. We assume that arcs are comparable (i.e., same number of lanes) hence the multiplicative incident delay factor is independent of the arc.

Herein, we assume the following exponential function form for the incident delay.

$$\Theta(\kappa,\rho,\alpha,x)=\kappa e^{-\left(\frac{1}{\varphi}\right)\rho}=\kappa e^{-\left(\frac{1}{\varphi}\right)\left(t-t_{inc}^{0}\right)}$$

where $\rho,\kappa,\varphi\geq 0$, $x=1$, and $t\geq t_{inc}^{0}$.

I.3.3 Integrating Recurrent and Non-Recurrent Congestion and Cost Calculation

We assume that the objective of dynamic routing is to minimize the expected travel time based on real-time information where the travel starting point is node n_0 and destination point is node n_d. Assume there is a feasible path between (n_0,n_d) where a path $p=(n_0,..,n_k,..,n_{K-1})$ is defined as sequence of nodes such that $a_k\equiv(n_k,n_{k+1})\in A$, $k=0,..,K-1$ and K is the number of nodes on the path. We define set $a_k\equiv(n_k,n_{k+1})\in A$ as the current arcs set of node n_k, and denoted with

$CrAS(n_k)$. That is, $CrAS(n_k) \equiv \{a_k : a_k \equiv (n_k, n_{k+1}) \in A\}$ is set of arcs those emanating from node n_k.

Assume each node (intersections) on a path is a decision stage (or epoch) where a routing decision (which node to select next) can be made. Let $n_k \in N$ be the location of the k^{th} decision stage, t_k is the time at k^{th} decision stage where $t_k \in \{1,...,T\}$, $T > t_{K-1}$. Note that we discretized the time. We also define the following sets for arcs to reduce the state space. $ScAS(a_k)$, successor arcs set of arc a_k, $ScAS(a_k) \equiv \{a_{k+1} : a_{k+1} \equiv (n_{k+1}, n_{k+2}) \in A\}$ i.e., set of outgoing arcs from the destination node (n_{k+1}) of arc a_k. $PScAS(a_k)$, post-successor arcs set of arc a_k, $PScAS(a_k) \equiv \{a_{k+2} : a_{k+2} \equiv (n_{k+2}, n_{k+3}) \in A\}$ i.e., set of outgoing arcs from the destination node (n_{k+2}) of arc a_{k+1}.

Since total trip travel time is an additive function of individual arc travel times on the path plus a penalty function measuring earliness/tardiness of arrival time to the final destination, dynamic route selection problem can be modeled as a dynamic programming model. The state of the system at k^{th} decision stage is denoted by $\Omega(n_k, t_k, s_{a_{k+1} \cup a_{k+2}, k}, X_k)$. It is composed of the state of the vehicle and network thus characterized by the current node (n_k), arrival time (t_k), and congestion ($s_{a_{k+1} \cup a_{k+2}, k}$) state of arcs $a_{k+1} \cup a_{k+2}$ where $\{a_{k+1} : a_{k+1} \in ScAS(a_k)\}$ and $\{a_{k+2} : a_{k+2} \in PScAS(a_k)\}$, and incident states ($X_k$) of the network, i.e. $X_k \equiv X(t_k)$. Action space for $\Omega\left(n_k, t_k, s_{a_{k+1} \cup a_{k+2}, k}, X_k\right)$ is the current arcs set of node n_k, denoted with $CrAS(n_k)$.

At every decision stage, the trip planner evaluates the alternative arcs from $CrAS(n_k)$ based on the remaining expected travel time. The expected travel time at a given node is composed of minimum expected travel time on the next outgoing arc chosen and expected travel time of the next node. Let's $\pi = \{\pi_0, \pi_1, ..., \pi_{K-1}\}$ be the set of policies for the trip. For a given state $\Omega\left(n_k, t_k, s_{a_{k+1} \cup a_{k+2}, k}, X_k\right)$, policy $\pi_k(\Omega_k)$ is a deterministic Markov policy which chooses the outgoing arc from node n_k, i.e., $\pi_k(\Omega_k) = a \in CrAS(n_k)$. Therefore, the expected travel cost given the policy vector $\pi = \{\pi_0, \pi_1, ..., \pi_{K-1}\}$ is as follows:

$$F_0(n_0, t_0, S_0, X_0) = \underset{\delta_k}{E}\left\{ g_{K-1}(\Omega_{K-1}) + \sum_{k=0}^{K-2} g_k(\Omega_k, \pi_k(\Omega_k), \delta_k) \right\}$$

where (n_0, t_0, S_0, X_0) is the starting state of the system. δ_k is random travel time at decision stage k, i.e., $\delta_k \equiv \delta(t_k, \pi_k(\Omega_k), s_a(t_k), x_a(t_k))(1 + \Theta(\kappa, \rho, \varphi, x))$ and $\Theta(\kappa, \rho, \varphi, x) = 0$ if $x = 2$ (an arc without incident). $g(\Omega_k, a, \delta_k)$ is cost of travel on arc $a_k = \pi_k(\Omega_k) \in CrAS(n_k)$ at stage k, i.e., if travel cost is a function (ϕ) of the travel time, then $g(\Omega_k, \pi_k(\Omega_k), \delta_k) \equiv \phi(\delta_k)$.

Minimum expected travel time can be found by minimizing $F\left(n_0, t_0, S_0, X_0\right)$ over the policy vector $\pi = \{\pi_0, \pi_1, ..., \pi_{K-1}\}$ as follows:

$$F^*\left(n_0, t_0, S_0, X_0\right) = \min_{\pi=\{\pi_0,\pi_1,...,\pi_{K-1}\}} F\left(n_0, t_0, S_0, X_0\right)$$

where $\pi^* = \underset{\pi=\{\pi_0,\pi_1,...,\pi_{K-1}\}}{\arg\min} F\left(n_0, t_0, S_0, X_0\right)$. Hence, the Bellman (cost-to-go) equation for the dynamic programming model can be expressed as follows:

$$F^*\left(\Omega_k\right) = \min_{\pi_k} \underset{\delta_k}{E}\left\{g(\Omega_k, \pi_k\left(\Omega_k\right), \delta_k) + F^*\left(\Omega_{k+1}\right)\right\}$$

For a given policy decision $\pi_k\left(\Omega_k\right) = a_k \in CrAS\left(n_k\right)$, we can re-express the cost-to-go function by writing the expectation in explicit form such that:

$$\begin{aligned} F\left(n_k, t_k, s_{a_{k+1}\cup a_{k+2},k}, X_k \mid a_k\right) = \sum_{\delta_k} P\left(\delta_k \mid n_k, t_k, S_k, X_k, a_k\right)\Big[g\left(\Omega_k, a_k, \delta_k\right) + \\ \sum_{s_{a_{k+1},k+1}} P\left(s_{a_{k+1},k+1}\left(t_{k+1}\right) \mid s_{a_{k+1},k}\left(t_k\right)\right) \sum_{s_{a_{k+2},k+1}} P\left(s_{a_{k+2},k+1}\left(t_{k+1}\right)\right) \\ \sum_{X_{k+1}} P\left(X_{k+1}\left(t_{k+1}\right) \mid X_k\left(t_k\right)\right) F\left(n_{k+1}, t_{k+1}, S_{k+1}, X_{k+1}\right)\Big] \end{aligned}$$

where $P\left(\delta_k \mid n_k, t_k, S_k, X_k\right)$ is the probability of travelling arc a_k in δ_k periods and calculation of δ_k is explained above for different states. $P\left(s_{a_{k+2},k+1}(t_{k+1})\right)$ is the state probability of arcs $a_{k+2} : a_{k+2} \in PScAS\left(a_k\right)$ in stage $k+1$. This probability is calculated from the frequency of a state at a given arc and time.

Using the backward induction we could solve $F_k^*\left(\Omega_k\right)$ for $k = K-1, K-2, .., 0$, where $\Omega_{K-1} = \Omega\left(n_{K-1}, t_{K-1}\right)$, n_{K-1} is destination node, and $F_{K-1}\left(\Omega_{K-1}\right) = 0$ if $t_{K-1} \le T$. A cost of penalty is accrued whenever $t_{K-1} > T$.

I.4. Experimental Studies

In this section we will demonstrate our proposed algorithm and methods solution quality on a network from South-east Michigan with real-time traffic data from M-DOT's MITS Center and Traffic.com. All algorithm and methods were coded in Matlab 7 and executed on a Pentium IV machine with 1.6 GHz speed processor and 1024 MB RAM under the Microsoft Windows XP operating system environment.

Our experimental study is outlined as follows: Section I.4.1 describes two road networks and traffic data information. Section I.4.2 explains the experimental settings and modeling recurrent congestion. Savings of dynamic policy under recurrent-congestion for a sub-network and 5 OD pairs are given in section I.4.3. The experimental setup for the sub-network with incident and its savings are given in the Section I.4.4. The last section summarizes the results and gives some insights about them.

I.4.1 Networks and Traffic Data

This section describes two road networks considered for our experiments along with traffic data information extraction and processing steps. Our methodology is mainly applied to the South-East Michigan freeway and highway road network, an urban area, that includes Detroit metro area (Figure 1). The network has 30 nodes and a total of 98 arcs with many observed arcs and few unobserved arcs (all the Michigan freeways, including M-24 and M-39, and city/local roads). The real-time traffic data of observed arcs is collected by MDOT ITS Center for 23 weekdays from January 21 to February 20, 2008 with a one minute resolution (involving literally hundreds of loop sensors). This raw speed data is cleaned with a series of procedures from Cambridge Systematics [1] to assure quality standards of the data.

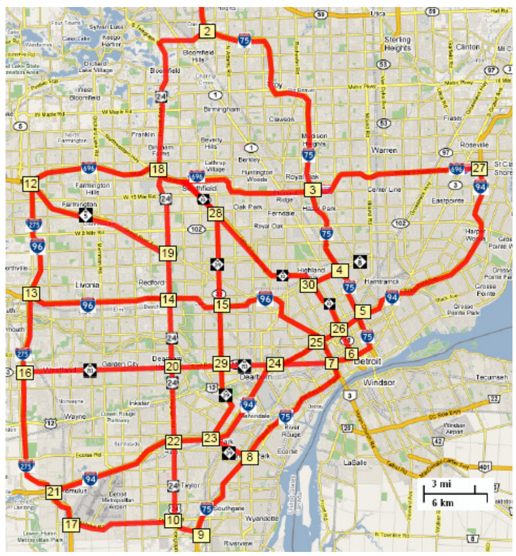


Figure 1 : South-East Michigan road network

We took a small part of our full network and labeled as *stylized network* (Figure 2) to illustrate the incident modeling methods and results better. The stylized network (Figure 2) has 5 nodes and 6 observed arcs. Arcs lengths and information are as shown in Table 1.

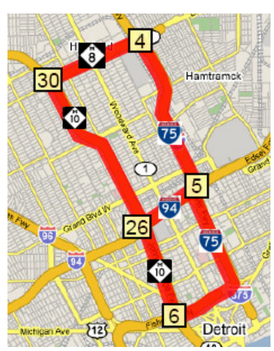


Figure 2 : The South-East Wayne County road *stylized network* used in Experiment Set #1

Table 1 : Stylized Network Information for Experiment Set #1

Arc ID	Freeway	Length (miles)	FROM		TO	
			Node #	Description (Exit #)	Node #	Description (Exit #)
1	I-94	1.32	5	216	26	215
2	M-8	1.75	4	56A (I-75)	30	7C (M-10)
3	I-75	3.13	4	56A	5	53B
4	I-75	2.81	5	53B	6	50
5	M-10	3.26	30	7C	26	4B
6	M-10	1.42	26	4B	6	2A

We consider node 4 as the origin node and node 6 as the destination node of the trip. Speed data of the arcs are plotted in Figure 3. As can be seen from the figure, the traffic speeds are highly stochastic and non-stationary, varying by time of the day. The mean and standard deviations of arcs speeds during the day are plotted in Figure 4.

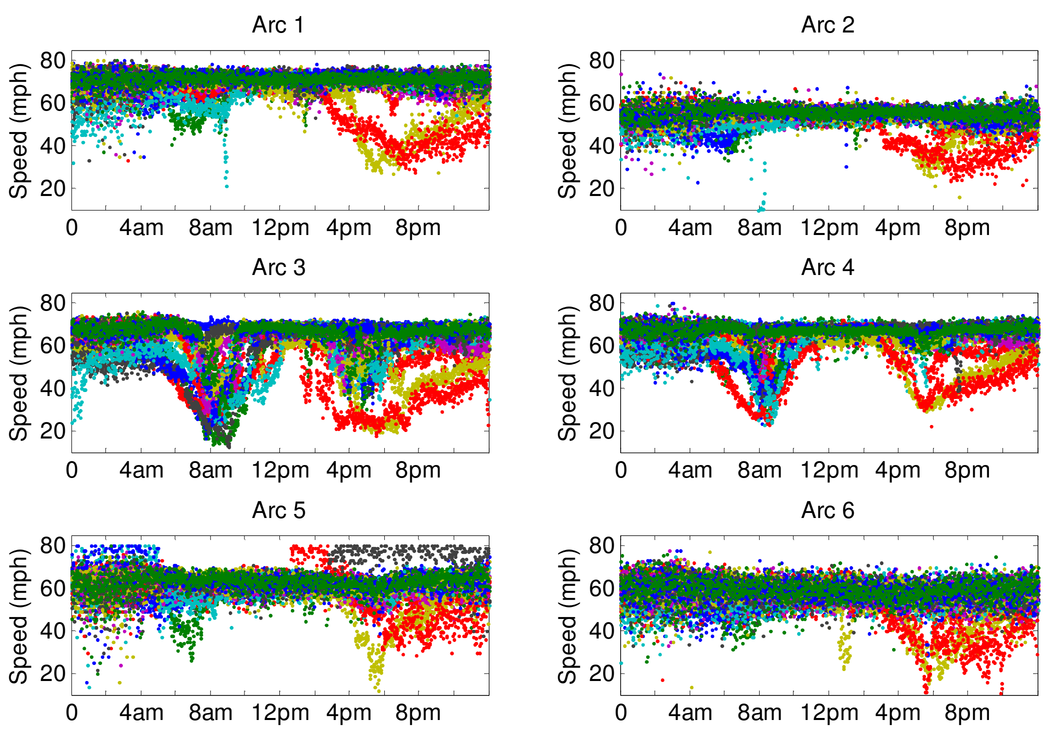


Figure 3 : Raw traffic speeds for arcs on stylized network (mph) at different times of the day.
Data: Weekday traffic from January 21 to February 20.
Each color represents a distinct day of 23 days.

I.4.2 Recurrent Congestion Modeling

The methodology of implementing real data to the dynamic policy algorithm needs estimating transition probabilities between states for each arc and parameters of arcs travel time for each state. Number of congestion states on each arc is determined using the Gaussian Mixture Models

(GMMs). In particular, it is according to the greedy learning GMMs of Verbeek (2003). In order to estimate the classes, two consecutive periods' velocities are modeled as bi-variate joint Gaussian distribution, i.e: velocity data at time t as x axes and time $t+1$ as y axes. GMMs identify the classes as ellipses where center of ellipse is class mean $\left(\mu_t^i, \mu_{t+1}^i\right)$. Ellipse has the eigenvectors of the covariance matrix as axes and radii of twice the square root of the corresponding eigen value. The ellipses are labeled such that $(\mu_t^0, .., \mu_t^i, .., \mu_t^r)$ is an increasing series. The mid-point of the line from i th ellipse center to $(i+1)$ th ellipse center is denoted with $c_a^i(t)$ (i th cut-off speed for arc a at time t), that is $c_a^i(t) = (\mu_t^i + \mu_t^{i+1})/2$. We assume $c_a^0(t) = 0$ mph and $c_a^r(t) = 80$ mph.

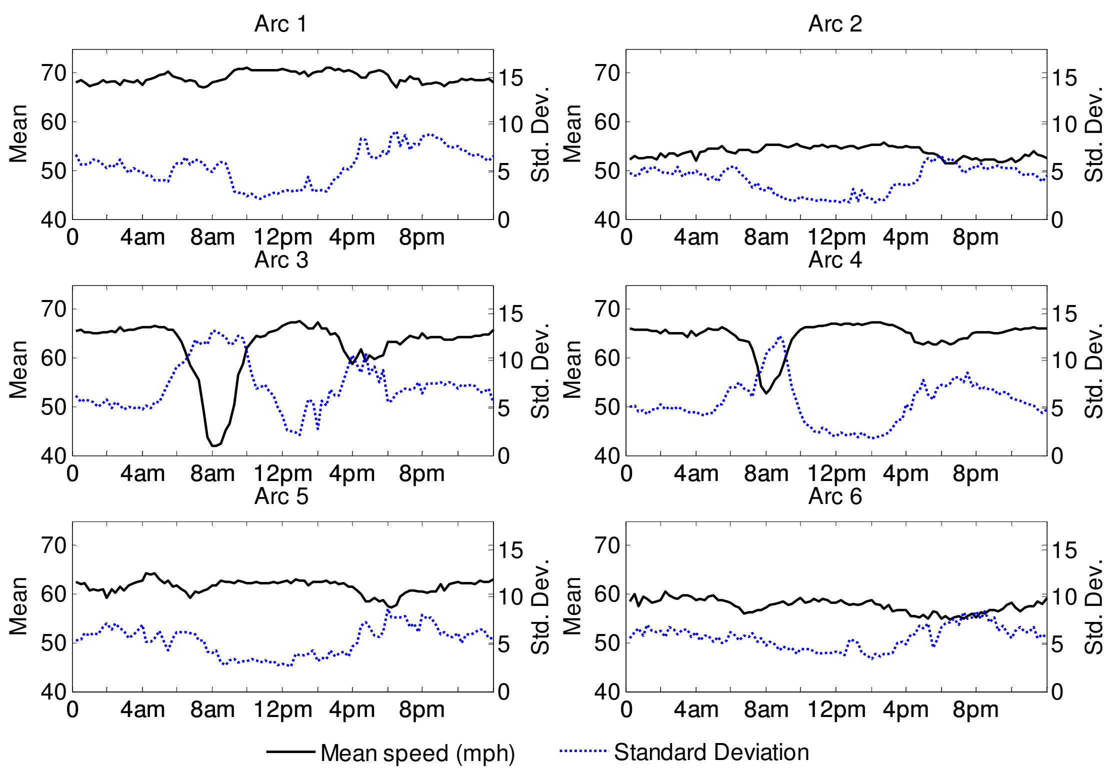


Figure 4 : Traffic mean speeds (mph) and standard deviations by time of the day for arcs on stylized network. (15 minute time interval resolution)

Although our method can handle any number of recurring congestion states, for the sake of simplicity, we limit here the number of states to two states: congested and uncongested. Two different examples are given which depicts the evolution of classifications at two different times of the day. In Figure 5 (a), joint plots of traffic speeds in consecutive periods at 8:30 am, at arc 1 is plotted and in Figure 5 (b) the classes are shown after partitioning with GMMs. Each class is shown as an ellipse which has the eigenvectors of the covariance matrix as axes and radii of twice the square root of the corresponding eigenvalue.

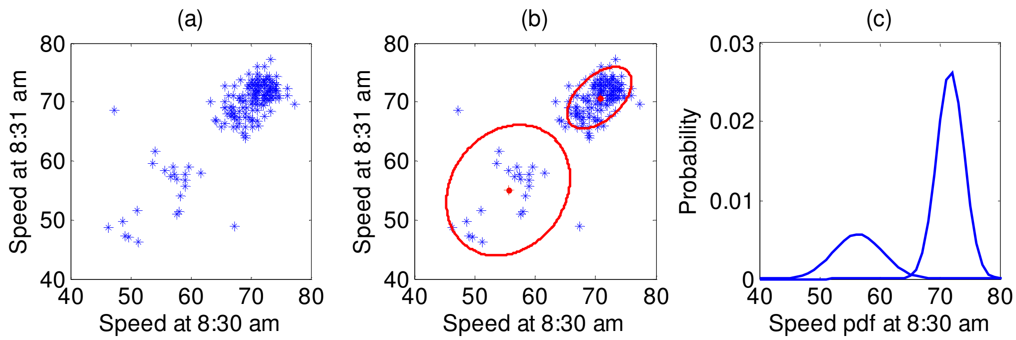


Figure 5 : (a) Joint plots of traffic speeds in consecutive periods for modeling state-transitions at 8:30 am, at arc 1 (b) Probability distribution of speed at 8:30 am, at arc 1 generated by GMM (c) Classes partitioned with GMM

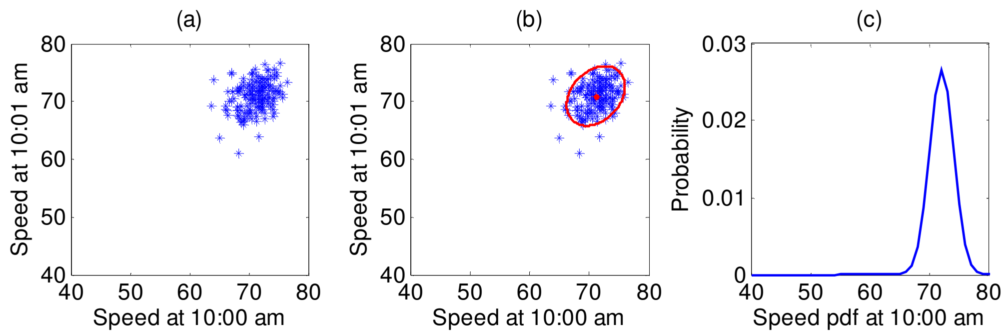


Figure 6 : (a) Joint plots of traffic speeds in consecutive periods for modeling state-transitions at 10:00 am, at arc 1 (b) Probability distribution of speed at 10:00 am, at arc 1 generated by GMM (c) Class identified with GMM

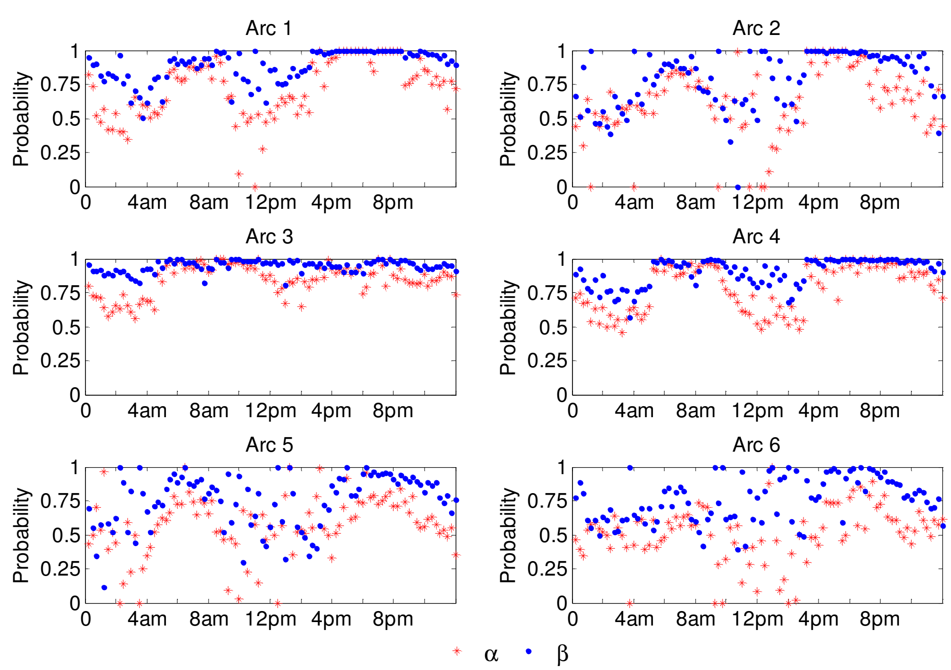


Figure 7: Recurrent congestion state-transition probabilities for arcs on stylized network. α: congested to congested transition; β: uncongested to uncongested transition probability (15 minute time interval resolution)

We assume if GMMs suggested two classes than the mid-point of the line from one ellipse center to the other is the cut-off speed. For instance, in Figure 5 (c) the congested class' center x-axes value (time: 8:30) is 53.3mph and the uncongested class' center x-axes value is 70.5mph. This implies that the cut-off speed at 8:30 am for arc 1 is 61.9mph.

In Figure 6 (a) joint plots of traffic speeds in consecutive periods at 10:00 am, at arc 1 is plotted. In Figure 6 (c) there is only one class suggested by GMMs. So, we assume there is no cut-off speed and this class belongs to the uncongested state. Using these cut-off speeds and speeds at each class, we calculated travel time distribution parameters and the transition matrix elements as explained earlier. The state transition elements alpha (congested to congested) and beta (uncongested to uncongested) are illustrated in Figure 7. Note that the state transitions to the same states are more likely in high traveler demand time periods, which is the case in practice.

The parameters mean and standard deviations, of each state for all arcs during the day are illustrated in Figure 8 and Figure 9 respectively.

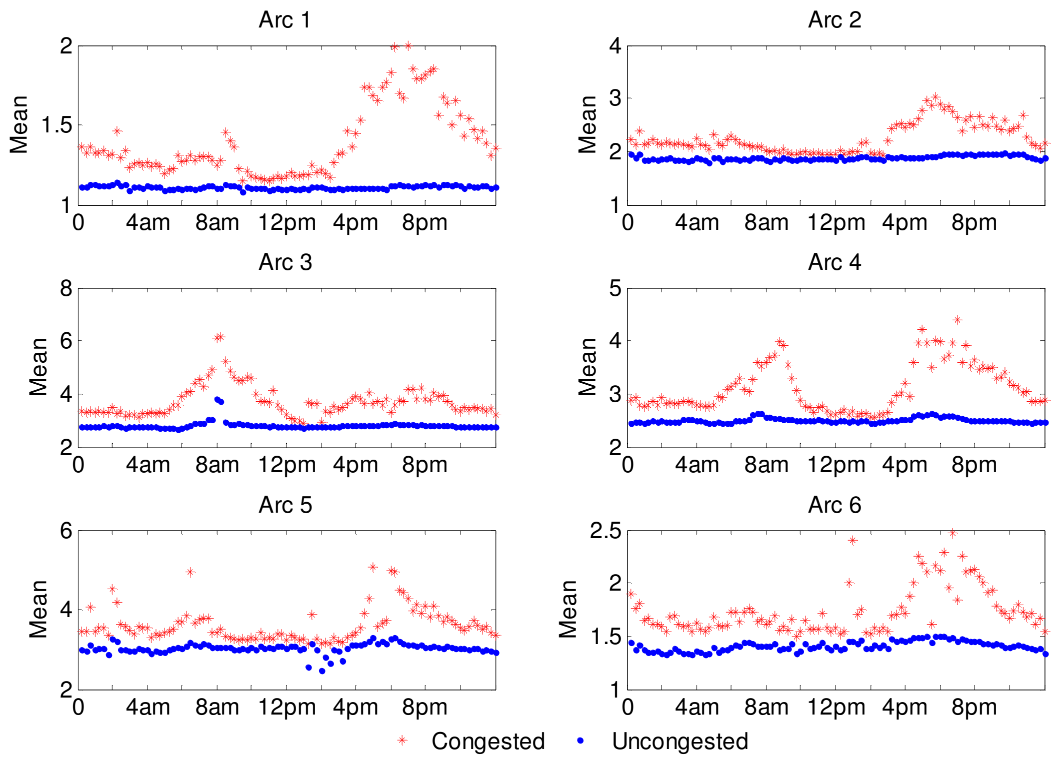


Figure 8 : Mean arc travel times of arcs on stylized network in minutes.
(15 minute time interval resolution)

I.4.3 Results of Modeling Recurrent Congestion

In this section the savings of our methodology is shown based on two different network under recurrent congestion. First we will discuss the results of stylized network. As stated earlier, we consider node 4 as the origin node and node 6 as the destination node of the trip for this network. From origin to destination there are 3 different path options could be taken (*path 1: 4-5-6; path 2: 4-5-26-6;* and *path 3: 4-30-26-6*). Note that, our aim is not to identify an optimal path; our aim is to identify the best policy based on the time of the day, location of the vehicle and the traffic state of the network. However, any commercial logistics software tries to define a best (static) path from an origin to the destination. In this context, the commercial logistics software selects the best path as *path1: 4-5-6*, since it is dominant to others most of the time during the day with assuming the historical speed data is available to the software. We will identify the *path 1* as *baseline path* and show the savings with regard to baseline path. We simulated the trip 10.000 times for each starting state combinations (since there are 6 arcs, there are 2^6=64 network traffic state combinations) throughout the day with 15 minutes intervals. Figure 10 (a) shows baseline path and Figure 10 (b) shows dynamic policy simulated travel times for every state combinations of the stylized network with sample size 10.000. We define savings as following:

$$\text{Savings}_t(\%) = \left(\frac{T_t^{DP} - T_t^B}{T_t^B} \right) \times 100 \ .$$

where T_t^{DP} is time to complete the travel under dynamic vehicle routing policy and T_t^B is time to complete the travel on the baseline path at time t. Figure 11 (a) shows the savings of each state combinations of dynamic policy over baseline path during the day and Figure 11 (b) shows the average savings for all states during the day.

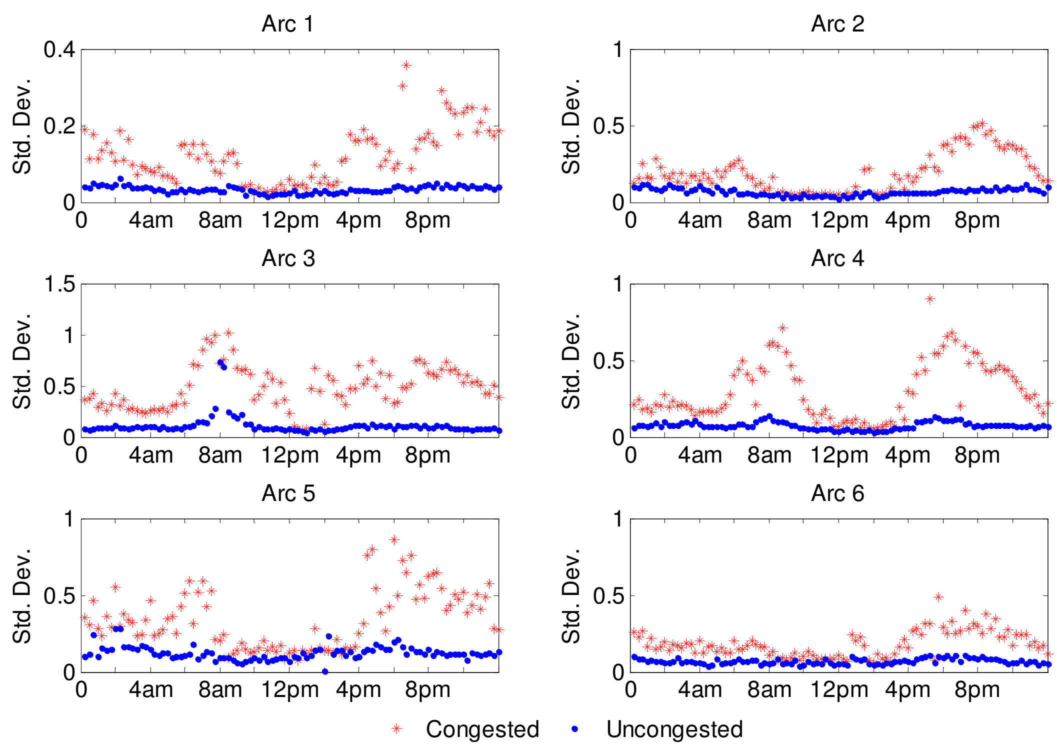


Figure 9 : Arc travel times standard deviations of arcs on stylized network in minutes. (15 minute time interval resolution)

Beside stylized network we identified also 5 other origin and destination (OD) pairs (Table 2) in Southeast Michigan road network (Figure 1) to investigate the savings by using real-time traffic information in dynamic policy. Unlike the *stylized network* these OD pairs has both observed and unobserved paths and each has several alternative paths from origin node to destination node.

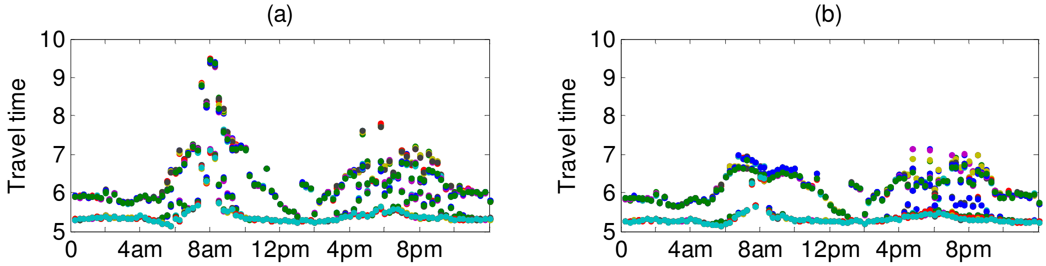


Figure 10: (a) Baseline path and (b) dynamic vehicle routing policy simulated travel times for all state combinations of the stylized network (each color represents a different state combination).

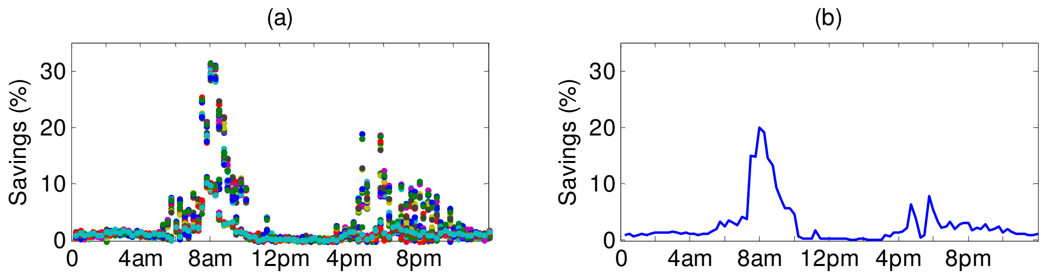


Figure 11: (a) For all state combinations, and (b) the average savings of all state combinations in percentages of dynamic policy over baseline policy during the day. (15 minute time interval resolution)

Table 2 : Origin-Destination pairs selected from South-East Michigan road network

OD Pair	ORIGIN		DESTINATION	
	Node #	**Description (Intersection)**	**Node #**	**Description (Intersection)**
1	2	I-75 & US-24	21	I-275 & I-94
2	12	I-96 & I-696	25	I-96 & I-94
3	19	M-5 & US-24	27	I-696 & I-94
4	23	I-94 & M-39	13	I-96 & I-275
5	3	I-75 & I-696	15	I-96 & M-39

We identify the *baseline path* of each OD pair as explained for stylized network and show the savings with regard to their baseline paths. We simulated the trip 10.000 times for each starting state combinations of corresponding OD pair throughout the day with 15 minutes intervals.

Figure **12** shows the savings of each state combination of dynamic policies over baseline paths during the day for each OD pairs.

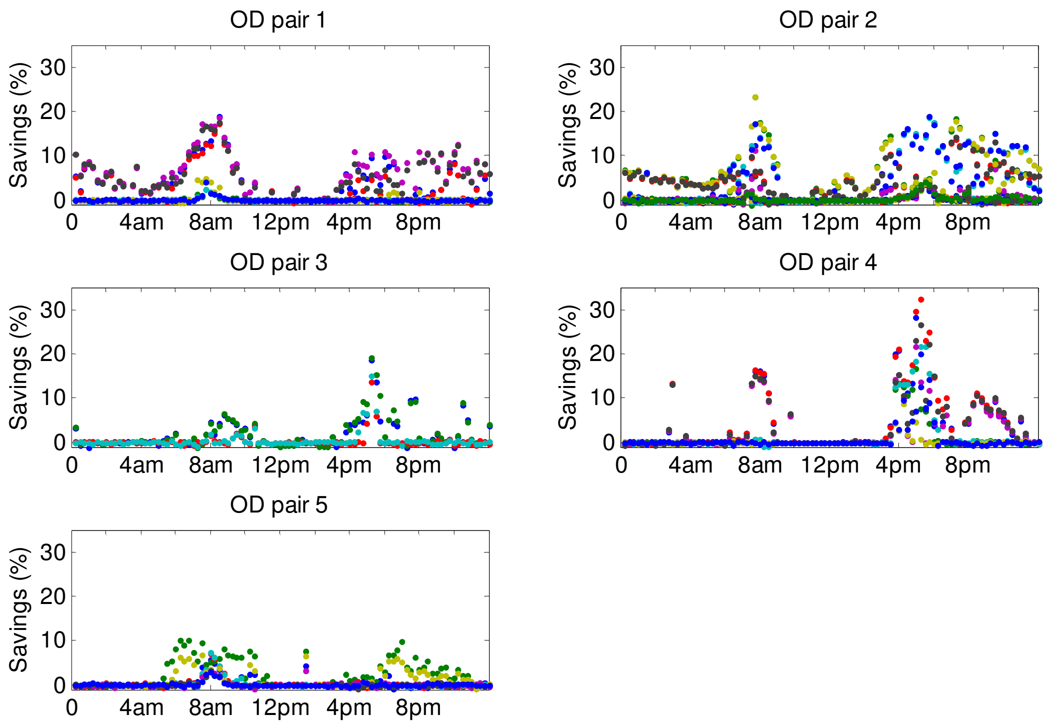


Figure 12 : For different OD pairs, savings percentages of dynamic policy over baseline path during the day for all states

Figure 13 shows the average savings for all states during the day.

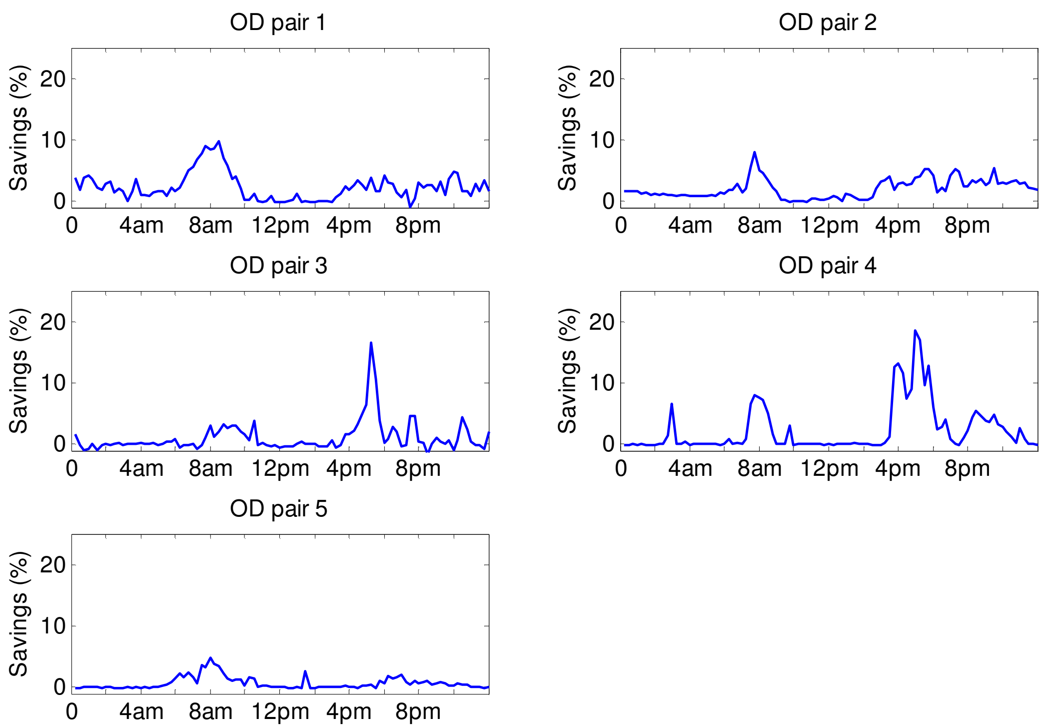


Figure 13 : For different OD pairs, average savings percentages of dynamic policy over baseline path during the day for all states

I.4.4 Impact of Modeling Incidents

In this section, the stylized network (Figure 2) is experiencing nonrecurring congestion beside recurrent congestion. We derive the dynamic routing policies in two ways. Initially, the dynamic policy does not account for non-recurring congestion while there is one incident on the network. Later, we allow the dynamic policy to explicitly account for non-recurring congestion information to dynamically update the route. The incident settings are assumed to be the same for all incidents. We show the results for 6 starting times during the day. The incident clearance decay function severity parameter is set to, κ=2 and response or clearance rate parameter is set to α=10 (Fast). The Weibull distribution, models how much time the incident arc spends in the incident state, shape parameter is set to 2, and scale parameter is set to 2.

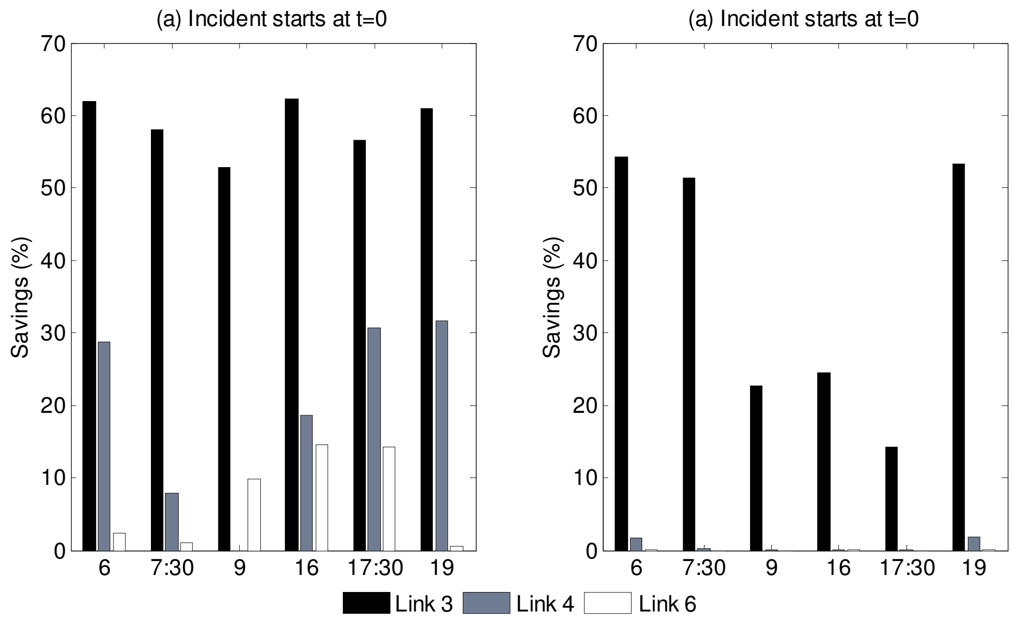


Figure 14: Savings realized from modeling non-recurring incidents (besides recurring congestion) under incidents on arc 3, 4 or 6. (a) Trip starts at the time of the incident; (b) trip starts 3 minutes after incident has occurred

The results presented here (Figure 14) pertain to incidents on arc 3, 4 or 5. We studied two different scenarios: first, the incidents are created on one of these arcs at the same instant the vehicle departs the starting node; second is the incidents are created on one of these arcs 3 minutes before the vehicle departs the starting node. For example, if the vehicle departs the origin node at 10AM, we create an incident on arc 3 at 10AM as well. Here we presented the savings realized from modeling non-recurring incidents only (besides recurring congestion) and calculated as follows:

$$\text{Savings}_t(\%) = \left(\frac{T_t^{R\&NR} - T_t^R}{T_t^R} \right) \times 100 .$$

where $T_t^{R\&NR}$ is time to complete the travel with dynamic policy that models both recurring and non-recurring congestion and T_t^R is time to complete the travel with dynamic policy that models only recurring congestion at time t. The savings for the first scenario is presented in Figure 14(a). Since the arc 3 is very close to the origin node and on the baseline path the effect of incident is high and savings are higher. Arc 4 is also on the baseline path and given that it is a downstream arc (i.e., it is not connected to the origin node), by the time the vehicle reaches there, the incident is partially cleared, reducing the impact of the incident on arc travel time and the savings is less. Arc 6 is also a downstream arc but not on the baseline path. However, the dynamic policy that not take into account the non-recurrent congestion takes this arc more (Figure 15) than the dynamic policy that take into account the non-recurrent congestion and this leads to savings. Due to space constraints, we are not presenting results from incidents on other arcs. The results for other arcs vary based on similar reasons. The results for the second scenario are presented in Figure 14(b). The savings for this scenario is less than the other since the incident partially or fully cleared until we reach the incident arcs. Similarly, since the arc 3 is very close to the origin node and on the baseline path the effect of incident is high and savings are higher. Arc 4 and 6 are downstream arcs (i.e., it is not connected to the origin node) and by the time the vehicle reaches those arcs the incident is partially/fully cleared. Thus, there is little savings or not.

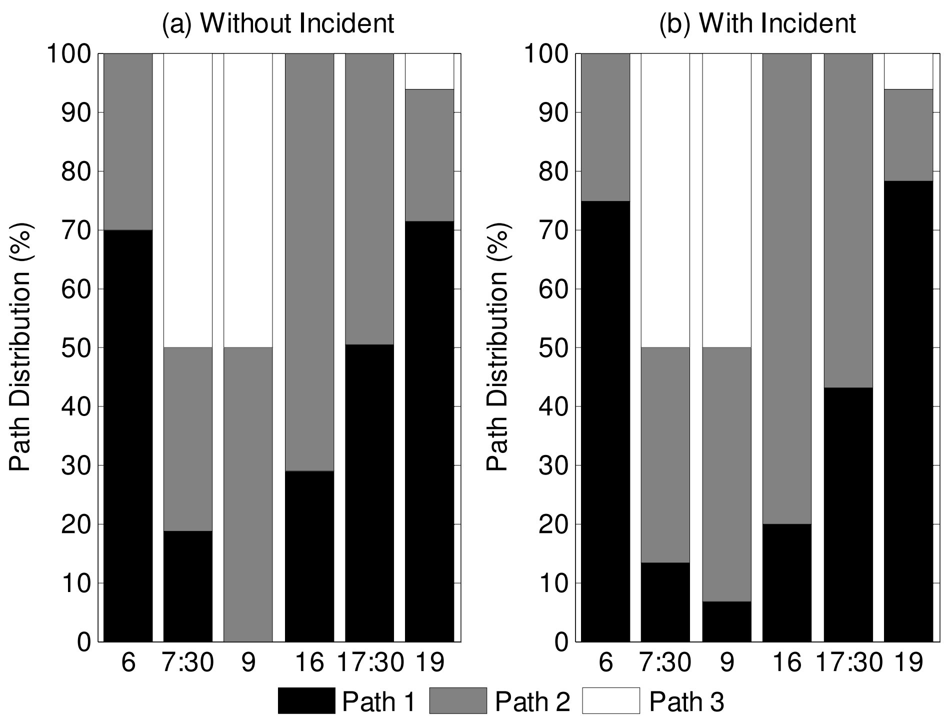


Figure 15 : Path distribution during the day (a) without incident, (b) with an incident on arc 3 and trip starts at the time of incident.

I.4.5 Experiment Conclusions

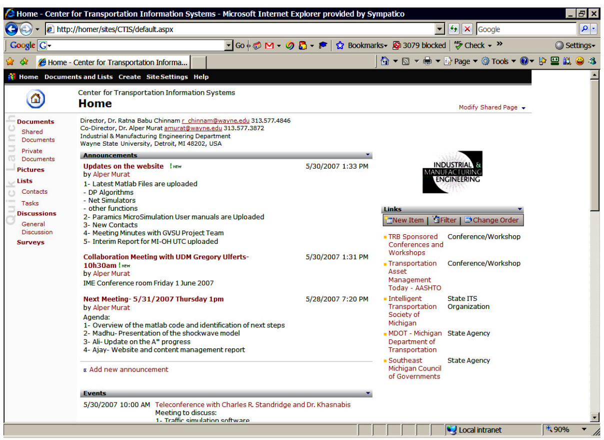
Experiments clearly illustrate the superior performance of the SDP derived dynamic routing policies when they accurately account for recurrent congestion (i.e., they differentiate between congested and uncongested states) and non-recurrent congestion attributed to incidents (e.g., accidents). To show the effects of taking into account recurrent and non-recurrent congestion first, we present results for the case where the network experiences no incidents but experiences recurrent congestion. Later, we present results for the case where the network is also experiencing non-recurring incidents besides recurrent congestion. Future work will work to extend the quality and efficacy of our dynamic routing models and work to relax some of the assumptions (e.g., traffic conditions on adjacent arcs are independent).

II: Results Dissemination

II.1 Project Website

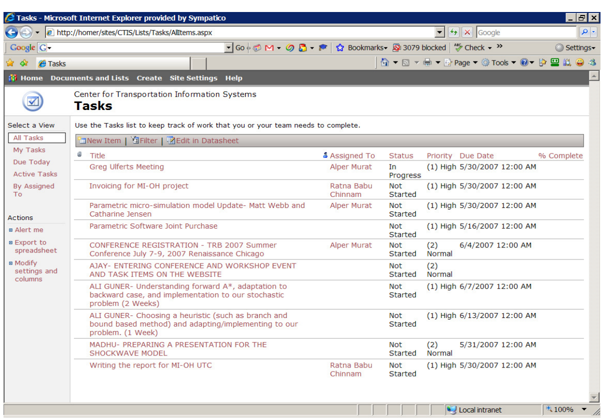
We have established a Microsoft SharePoint Website for the project that helps us track/store all project related documents/information in one place. Currently, it carries all our literature, data sets, code, weekly research group meeting minutes, long-term mile-stones, short-term tasks, calendar, and contacts. While we currently control access to this website through password protection, we are in the process of opening parts of the website for anonymous access. The screen shots below highlight different parts of our website.

Homepage

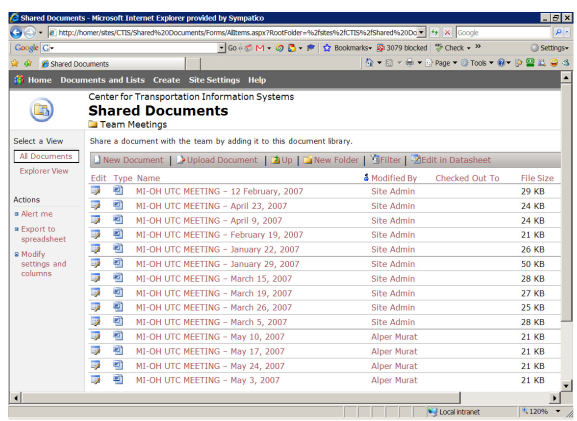


Literature

Tasks



Meeting Minutes



II.2 Conference Activity

Conference Presentations:

1. Guner, A., Chinnam, R.B., and Murat, A., "Dynamic Vehicle Routing under Real-time Congestion and Incident Information for JIT Logistics," *INFORMS 2008 Annual Meeting,* Washington D.C. (Oct 12-15, 2008).

2. Guner, A., Chinnam, R.B., and Murat, A "Modeling Traffic Incidents for Dynamic Vehicle Routing Applications," *INFORMS 2008 Annual Meeting,* Washington D.C. (Oct 12-15, 2008).
3. Guner, A., Chinnam, R.B., Murat, A., and Saripalle, M., "Enabling Congestion Avoidance in Stochastic Transportation Networks Under ATIS," *INFORMS 2007 Annual Meeting,* Seattle (Nov 3-7, 2007).
4. Saripalle, M., Chinnam, R.B., Murat, A., and Guner, A., "Modeling Incidents for Dynamic Vehicle Routing Applications," *INFORMS 2007 Annual Meeting,* Seattle (Nov 3-7, 2007).
5. Murat, A., Chinnam, R.B., Guner, A., Saripalle, M., and Azadian, F., "Dynamic Routing under ATIS for Congestion Avoidance," *Research Issues in Freight Transportation -- Congestion and System Performance Conference,* Seattle (Oct 22-23, 2007).

Conference Sessions Organized:

1. We are organizing a special session titled "Dynamic Routing and Logistics under Real-Time ITS Information" at the *INFORMS 2008 Annual Meeting* in Washington D.C. (Oct 12-15, 2008) under the Cluster: Real-time Systems. The session is Chaired by our Co-PI – Dr. Alper Murat.
2. "Urban Transportation Planning Models: Dynamic Routing with Real-time ITS Information" at the *INFORMS 2007 Annual Meeting* in Seattle (Nov 3-7, 2007) under the Cluster: Transportation Science & Logistics. The session is Chaired by our Co-PI – Dr. Alper Murat.

Conferences Attended:

1. Drs. Chinnam and Murat attended the *Michigan Intelligent Transportation Systems Conference*, May 16-17, 2007. Dr. Khasnabis, Associate Dean for Research College of Engineering has presented on our project along with other WSU efforts related to transportation.
2. Dr. Murat attended the *Meeting Freight Data Challenges Conference*, July 9–10, 2007 Renaissance Chicago Hotel, Chicago.
3. Dr. Murat attended the 2nd *Annual National Urban Freight Conference* - December 5-7, 2007 Long Beach, CA.

Conferences Planning to Attend:

1. INFORMS Annual Meeting- October 12-15, 2008 Washington, D.C.

II.3 Journal Publications

A manuscript has been dispatched to *Computers & Operations Research* journal that reports our SDP algorithms and their performance. A second manuscript based on the AO* algorithms is currently under preparation.

III: References

1. A.P. Eiger, P.B. Mirchandani, H. Soroush, *Path preferences and optimal paths in probabilistic networks*, Transportation Science 19 (1985) 75–84
2. D. P. Bertsekas, J. N. Tsitsiklis, *An analysis of stochastic shortest path problems,* Mathematics of Operations Research 16 (1991) 580-595
3. R. Hall, *The fastest path through a network with random time-dependent travel time*. Transportation Science 20(3) (1986) 182-188
4. L. Fu, L.R. Rilett, *Expected shortest paths in dynamic and stochastic traffic networks*, Transportation Research Part B 32 (1998) 499–516
5. E.D. Miller-Hooks, H.S. Mahmassani, *Least expected time paths in stochastic, time-varying transportation networks,* Transportation Science 34 (2000) 198–215
6. E.D. Miller-Hooks, H. S. Mahmassani, *Path comparisons for a priori and time-adaptive decisions in stochastic, time-varying networks,* European Journal of Operational Research 146 (2003) 67–82
7. J.L. Bander, C.C. White III, *A heuristic search approach for a nonstationary shortest path problem with terminal costs,* Transportation Science 36 (2002) 218–230
8. H.N. Psaraftis, J.N. Tsitsiklis, *Dynamic shortest paths in acyclic networks with Markovian arc costs,* Operations Research 41 (1993) 91–101
9. G. H. Polychronopoulos, J. N. Tsitsiklis, *Stochastic shortest path problems with recourse,* Networks 27(2) (1996)133–143
10. A. Azaron, F. Kianfar, *Dynamic shortest path in stochastic dynamic networks: Ship routing problem,* European Journal of Operational Research 144 (2003) 138–156
11. L. Fu, *An adaptive routing algorithm for in vehicle route guidance systems with real-time information,* Transportation Research Part B 35(8) (2001) 749–765
12. S.T. Waller, A.K. Ziliaskopoulos, *On the online shortest path problem with limited arc cost dependencies,* Networks 40 (2002) 216–227
13. S. Gao and I. Chabini, Optimal routing policy problems in stochastic time-dependent networks, Transportation Research Part B 40 (2006) 93-122
14. S. Kim, M.E. Lewis, C.C. White III, *Optimal vehicle routing with real-time traffic information,* IEEE Transactions on Intelligent Transportation Systems 6 (2) (2005) 178–188
15. S. Kim, M.E. Lewis, C.C. White III, *State space reduction for non-stationary stochastic shortest path problems with real-time traffic congestion information,* IEEE Transactions on Intelligent Transportation Systems 6(3) (2005) 273–284
16. B.W. Thomas, C.C. White III, *The dynamic shortest path problem with anticipation,* European Journal of Operational Research 176 (2007) 836–854
17. FHWA Report, *Traffic Congestion and Reliability: Linking Solutions to Problems,* prepared by Cambridge Systematics Inc. for the Federal Highway Administration Office of Operations, Washington, D.C., July 2004.
18. D.P. Gaver, *Highway Delays Resulting From Flow Stopping Conditions*, Journal of Applied probability 6 (1969) 137-153
19. R.B. Noland, J.W. Polak, *Travel Time Variability: A Review of Theoretical and Empirical Issues* Transport Reviews 22(1) (2002) 39-54,

20. J.A. Lindley, *Urban freeway congestion: quantification of the problem and effectiveness of potential solutions,* Institute of Transportation Engineers Journal 57(1) (1987) 27-32
21. H. Cohen, F. Southworth, *On the Measurement and Valuation of Travel Time Variability Due to Incidents on Freeways*, Journal Of Transportation And Statistics 2(2) (1999) 123-131
22. T. Olmstead, *Pitfall to avoid when estimating incident-induced delay by using deterministic queuing models,* Transportation Research Record 1683 (1999) 38-46
23. J. Henderson, L. Fu, S. Li , *Optimal CMS: A Decision Support System for Locating Changeable Message Signs*, *Transportation Engineering* 2004, Beijing, China, (2004)
24. H. Al-Deek, A. Garib, A. E. Radwan, *New method for estimating freeway incident congestion*, Transportation Research Record 1494, (1995) 30-39.
25. G. Giuliano, *Incident Characteristics, Frequency, and Duration on A High Volume Urban Freeway*, Transportation Research 23 (1989) 387-36.
26. H .Mongeot, J.B. Lesort , *Analytical expressions of incident-induced flow dynamics perturbations,* Transportation Research Record 1710 (2000) 58-68
27. M. Baykal-Gürsoy, W. Xiao, K. Ozbay, *Modeling traffic flow interrupted by incidents, European Journal of Operational Research*, In Press, (2008)
28. A. Garib, A.E. Radwan, H. Al-Deek, *Estimating magnitude and duration of incident delays*, Journal of Transportation Engineering 123 (6) (1997) 459–466
29. C.E. Sigal, A.A.B. Pritsker, J.J. Solberg, *The stochastic shortest route problem*, Operations Research, 28, (1980) 1122-1129
30. H. Frank, *Shortest paths in probabilistic graphs*, Journal of Operational Research 17 (1969) 583-599
31. R.P. Loui, *Optimal paths in graphs with stochastic or multidimensional weights*, Communications of the ACM, 26(9) (1983) 670-676
32. Mobility Monitoring Program, Texas Transportation Institute. http://mobility.tamu.edu/mmp/
33. JJ. Verbeek, Vlassis N., Kröse B. *Efficient Greedy Learning of Gaussian Mixture Models,* Neural Computation 2003; 5(2): 469-485.
34. Urban Mobility Report, Texas Transportation Institute, 2007. http://mobility.tamu.edu/ums/