PATH PRIORITIZATION AND OPERATIONAL CONSISTENCY FOR DYNAMIC TRAFFIC ASSIGNMENT: AN ANALYTIC HIERARCHICAL APPROACH

Srinivas Bulusu1, Srinivas Peeta2, Hong Zheng3, Sushant Sharma4

1 E2open, Inc., Dallas, TX 75240, United States of America
2 Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, United States of America
3 NEXTRANS Center, Purdue University, West Lafayette, IN 47906, United States of America
4 Texas A&M Transportation Institute, San Antonio, TX 78213, United States of America

Received 18 August 2015; accepted 2 March 2016

Abstract: An operational consistency model for real-time dynamic traffic assignment (DTA) applications seeks to correct the time-dependent path assignment within a rolling horizon scheme. This study extends an existing consistency framework to develop a hierarchy for the time-dependent path set based upon their relative importance to ensuring consistency. Using the analytic hierarchy process, the eigenvalue associated with a path is identified as the parameter which enables the rank ordering of paths. The ability to identify a subset of dominant paths relative to enhancing consistency enhances the computational viability of the consistency framework for real-time implementation and has significant practical implications. Additionally, it provides insights on the complex dynamics that are inherent to the operational consistency problem.

Keywords: dynamic traffic assignment, analytic hierarchy process, path prioritization, eigenvalue.

1. Introduction

The dynamic traffic assignment (DTA) problem entails the consideration of several aspects including the accuracy, effectiveness, stochasticity, computational tractability, traffic realism, user heterogeneity, and system characteristics (information availability, type, and dissemination sources). The deployment of DTA entails using sensor data to adjust the DTA model in real-time. The real-time issues can be classified into two groups in general: those pertain to the system-wide operations (system-related issues), and those relevant to the solution procedure and quality (solution-related issues). The system-related issues include the functional integration, operational execution and its time requirements, measurements and reliabilities of system components, tolerances, and control architecture. The DTA solution-related issues address various nuances of consistency, robustness, and stability of solutions arising from a number of stochastic factors. While the two types of issues are dependent upon each other, this study focuses on the solution-related issues under a set of mild system-related assumptions.

Consistency refers to the degree of closeness between the predicted and observed traffic

1Corresponding author: peeta@purdue.edu
network states over time. It is a solution-related issue that can arise from unreliable origin-destination (O-D) demand estimation (Zhou and Mahmassani, 2005), simplified traffic flow dynamics (Zhang and Nie, 2005), inaccurate traffic flow model (Doan et al., 1999), detector measurement errors (Peeta and Anastassopoulos, 2002), behavior-related factors in route choice (Lotan and Koutsopoulos, 1993; Paz and Peeta, 2009), unpredictable incidents (Peeta and Zhou, 2002) and incorrect assumption of system-related parameters (Peeta and Yu, 2006). Without the correction of the above errors, the inconsistency would accumulate and propagate in the system making the network state estimation and prediction potentially unreliable. The inconsistency can also be viewed as the aggregate effect of random or unknown factors that introduce errors in the path assignment procedure (Peeta and Zhou, 2002).

Inconsistency issues for DTA have been mostly investigated in the context of specific contributing factors. For instance, studies (Antoniou et al., 2007; Balakrishna et al., 2007; Ben-Akiva et al., 2001; Zhou and Mahmassani, 2005) have sought to calibrate DTA models through optimization or statistical methods (He and Ran, 2007) to reduce the inconsistency by adjusting demand, supply, or demand and supply jointly (Huang et al., 2010; Omrani and Kattan, 2012; Vaze et al., 2009). These studies view inconsistency arising from incorrect O-D demand and/or behavior model on the demand side, as well as incorrect traffic flow model on the supply side. Incorrect prediction of travelers’ route choice behavior is another factor that can contribute to the inconsistency. In this context, travelers’ en route behavioral response to real-time information can be difficult to predict. The en route decisions describe their behavior in response to a change in traffic conditions and the need to deviate from their pre-trip routes. Hybrid route choice models, which combine both pre-trip route choice and en route route choice, have been used to address this inconsistency (Pel et al., 2009; Qian and Zhang, 2013). These studies highlight that behavior is a key source of inconsistency due to the spatio-temporal interactions resulting from travelers’ route choice decisions under information provision. Hence, behavioral aspects need to be carefully considered in consistency-checking procedures to factor travelers’ learning processes and en route adjustment in predicting the time-dependent network states (Paz and Peeta, 2009; Peeta and Yu, 2005).

Computational efficiency is another key challenge associated with consistency for real-time DTA. In a simulation-based DTA context, a brute force strategy to maintain system consistency is to re-simulate the network traffic conditions using the posteriori variables consistent with the state variable. Such a method is computationally intensive and not suitable for real-time DTA implementation (Zhou and Mahmassani, 2005). Kang (1999) developed a real-time long-term consistency updating module, which heuristically adjusts the demand level according to the discrepancy between simulated and observed link densities. Zhou and Mahmassani (2002) proposed a dynamic programming approach for real-time freeway flow propagation adjustment. Later, they further proposed an O-D demand consistency-checking and updating model for real-time DTA operations using feedback control to regulate the demand input (Hawas, 2002; Zhou and Mahmassani, 2005). Peeta and Bulusu (1999) proposed a mathematical programming approach
for ensuring consistency, which seeks to minimize deviations between real-time traffic measurements and predicted network states. All of these studies focus on real-time operational implementation, which requires the development of tractable and computationally efficient solution algorithms.

To address the inconsistency issue for real-time DTA, Peeta and Zhou (2002) developed a hybrid framework combining offline and real-time strategies to deal with two solution-related issues: (i) factoring the randomness in the O-D demand and/or network supply conditions, and (ii) enhancing the computational performance. The offline component uses historical O-D demand data to determine a robust initial solution, which is then updated dynamically in real-time based on unfolding O-D demand and incidents. The computationally intensive components are addressed offline to generate a robust initial solution that allows for efficient real-time updates, thereby enhancing the computational performance. Peeta and Yu (Peeta and Yu, 2004, 2005; Yu and Peeta, 2011) proposed and numerically analyzed a model that combines quantitative and qualitative variables in a single framework to more robustly predict travelers’ en-route route choice behavior under information provision. Later, Peeta and Yu (2006) combined their route choice model for network loading and a consistency-seeking model that updates driver behavior class fractions in real-time, called a behavior based consistency-seeking model (BBCS). BBCS captures the heterogeneous driver class fractions in real-time using the observed link traffic counts (Peeta and Yu, 2006). Although BBCS improves the solution quality, the computational effort is still significant relative to real-time deployment needs.

Peeta and Bulusu (1999) proposed a theoretical consistency-seeking model (CONS) to improve the operational consistency for real-time DTA with advanced traffic management and information systems. In this model, the inconsistency is measured by comparing predicted and observed link traffic counts, and is assumed to arise from a number of sources including inaccurate O-D demand, traffic incidents, inaccurate prediction of path choices for the unequipped user class (vehicles in this class are not equipped to receive real-time information), and inaccurate prediction of compliance for the equipped user class. The model is applied in a rolling horizon-based DTA approach (RH-DTA) (Peeta and Mahmassani, 1995) to improve the consistency by adjusting the path assignments. More specifically, the model first solves a deterministic DTA problem based on rolling-horizon technique to predict the traffic network performance, and then improves the consistency measured by the difference between the predicted network performance and the observed conditions unfolding in real-time. The consistency problem was then formulated as a constrained least squares model and solved using the generalized singular value decomposition (GSVD) method.

In their CONS model, Peeta and Bulusu (1999) observed that a considerable amount of computational effort is needed to determine a set of paths that contribute little to the consistency. It provides the motivation for this study; to rank order the time-dependent paths and identify a subset of dominant paths that significantly contribute to the consistency. By manipulating only the subset of dominant paths in the CONS model, the problem size can be reduced without compromising the solution quality significantly. This can improve the
computational efficiency considerably, which is synergistic with the real-time deployment objective.

In this study, we use the Analytic Hierarchy Process (AHP) to rank order the time-dependent paths in terms of their contribution to enhancing the consistency for real-time DTA deployment, by generating eigenvalues. Paths ranked high are grouped as the set of dominant paths. The rank ordering within the set of dominant paths signifies the importance of a path over the others in the consistency context. The inconsistency is corrected by manipulating only the set of dominant paths rather than all time-dependent paths.

The remainder of the paper is organized as follows. In the next section, the concept of AHP is briefly introduced, and the methodology to obtain priorities using eigenvectors of a pairwise comparison matrix is discussed. Section 3 applies AHP to the real-time consistency problem to obtain the path hierarchies. It also discusses the AHP-based path prioritization procedure and solution algorithm. Section 4 presents simulation experiments to demonstrate path prioritization in relation to real-time consistency. The final section summarizes the findings of this study and provides some concluding comments.

2. Analytical Hierarchy Process (AHP)

2.1. The Analytic Hierarchy Process (AHP)

The analytic hierarchy process (AHP) (Saaty, 1977; Saaty, 1980) is a technique widely applied in decision theory. It is leveraged in this study to identify a subset of dominant time-dependent paths (TDPs) that contribute to the inconsistency significantly, to the extent that considering the dominant subset of TDPs suffices to provide a similar degree of accuracy in the consistency context as when all paths are considered. From an operational perspective, this procedure suggests that manipulating the subset of dominant paths in the network at the current time can enhance the overall system performance in a CONS model. The resulting CONS model has a substantially reduced size that enables computational efficiency for real-time operational needs.

Saaty (1980) discusses aggregate decision-making using AHP, in which a set of alternatives is prioritized (or weighted) according to their importance. The AHP approach is as follows. First, in each level of a hierarchy with respect to an element of the next higher level, a pairwise comparison of \( n \) alternatives is constructed in an \( n \times n \) square matrix. In the current context, the set of alternatives are the time-dependent paths at a specific time. Each matrix element (\( i^{th} \) row and \( j^{th} \) column) indicates the strength with which path \( i \) dominates another path \( j \) for a given criterion. Next, the weights of the elements in each level of a hierarchy are scaled with respect to an element of the next higher level. This paper adopts Saaty’s method (Saaty, 1977) to scale the weights of the TDPs at a specific time with respect to another TDP of the next higher level. This is discussed in detail in the next subsection.

2.2. Determine Priorities: The Eigenvector Approach

AHP can model decision-making characterized by multiple objectives (multiple factors/criteria that are important in decision-making) (Arrow, 1963; Mirkkin, 1979; Sen, 1970; Wang and Shen, 1989; Wendell, 1980). There are two primary
goals in an analytic hierarchy process: (i) determine the relative importance of each objective involved in the decision process, and (ii) ensure that the qualitative judgment regarding each alternative to fulfill a certain objective is properly quantified.

Saaty (1980) proposed the eigenvector method for AHP. Let $P_1, P_2, \ldots, P_n$ be the set of alternatives, which are the set of TDPs in our context. The relative importance (or weight) of alternative $P_i$ compared to alternative $P_j$ is denoted by $p_{ij}$ (usually specified using a 1-9 scale). The pairwise comparisons of alternatives are represented in a reciprocal matrix:

$$P = \begin{bmatrix}
    1 & p_{12} & \ldots & p_{1n} \\
    1/p_{21} & 1 & \ldots & p_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    1/p_{n1} & 1/p_{n2} & \ldots & 1
\end{bmatrix}$$

(1)

This matrix has positive entries and is called the reciprocal matrix as it satisfies the reciprocal property $p_{ji} = 1/p_{ij}$.

The information conveyed by $P$ can be synthesized by methods such as least squares and logarithmic least squares methods (Harker and Vargas, 1987). $P$ is considered to be inconsistent when the reciprocal law is violated, that is, it does not satisfy the relation $p_{ij} p_{jk} = p_{ik}$ for every $i, j, k$. In such a case, the eigenvector method is useful to handle the inconsistency.

Let $W$ be the column vector that denotes the relative weights ($w_1, \ldots, w_n$) of the considered objective. There are $n$ eigenvalues in the square matrix $P$. Since all diagonal elements in $P$ are 1, the sum of all eigenvalues is $n$. If $\lambda_i$ is the $i$th eigenvalue, then:

$$\sum \lambda_i = \text{tr}(P) = \text{sum of the diagonal elements} = n$$

(2)

where $\text{tr}(P)$ represents trace of matrix $P$.

In a consistent matrix where the reciprocal law is strictly maintained for every $i, j$, and $k$, the largest eigenvalue is “$n$” and the other eigenvalues are zero. The relative weights in such a case can be obtained by solving the following Eq. (3):

$$P W = n W$$

(3)

where $W$ is the eigenvector corresponding to the eigenvalue $n$. But in most cases, the reciprocal matrices are not perfectly consistent. The degree of inconsistency is measured by the difference between the maximum eigenvalue $\lambda_{\text{max}}$ and $n$. The closer $\lambda_{\text{max}}$ is to $n$, the higher degree of consistency of $P$. When $\lambda_{\text{max}}$ is not significantly different from $n$, the relative weights $W$ can be estimated as follows:

$$P W = \lambda_{\text{max}} W$$

(4)

where $W$ is now the eigenvector with respect to $\lambda_{\text{max}}$. Let $w_i$ be the $i$th weight in $W$ and $\sigma = \sum_{i=1}^{n} w_i$. Then, the normalized weight is $w_i/\sigma$. The normalized weight is unique, and also satisfies $\sum_{i=1}^{n} w_i / \sigma = 1$. The lack of consistency can be measured by a consistency index (CI) defined as $\text{CI} = \lambda_{\text{max}} - n$. $\text{CI}$ is compared with the average consistency index of randomly generated reciprocal matrices, called the random index (RI) (Saaty, 1980). The consistency ratio (CR) is obtained by dividing CI by RI. The lower the value of CR indicates the higher degree of consistency. Saaty (1980) suggests that the consistency
ratio should be lower than 0.10 for every comparison matrix to ensure an accurate estimation. The relative weights of the alternatives under each of the objectives can be determined following the aforementioned steps. Let \( u_{1k}, u_{2k}, \ldots, u_{mk} \) be the relative weights of the \( m \) alternatives under the objective \( k \). Let \( \omega_k = \sum_{j=1}^{m} u_{jk} \); the normalized weights under the objective \( k \) are obtained as \( u_{jk} / \omega_k \). The normalized weights are aggregated to determine the priorities of the alternatives. The aggregated weights \( V \) for the alternatives are determined as follows:

\[
\begin{bmatrix}
    v_1 \\
    \vdots \\
    v_m
\end{bmatrix} = \begin{bmatrix}
    \frac{1}{\omega_1} u_{11} & \cdots & \frac{1}{\omega_n} u_{1n} \\
    \vdots & \ddots & \vdots \\
    \frac{1}{\omega_1} u_{m1} & \cdots & \frac{1}{\omega_n} u_{mn}
\end{bmatrix} \begin{bmatrix}
    \frac{1}{\omega} w_1 \\
    \vdots \\
    \frac{1}{\omega} w_n
\end{bmatrix}
\]

The alternative with the highest weight is then chosen as the most desirable alternative. The weights \( v_i \) are ranked in the descending order and reflect the priorities of the alternatives.

### 3. Application of AHP to the Real-Time Consistency Problem

#### 3.1. Overall Framework

This study extends Peeta and Bulusu’s CONS model (Peeta and Bulusu, 1999) using AHP to improve real-time DTA consistency. The developed model is labeled as AHP-based CONS (A-CONS). The AHP is used to rank order the TDPs using the eigenvector approach in terms of their significance to inconsistency.

Fig. 1 illustrates the A-CONS for real-time DTA in a stage-based rolling horizon framework. The objective is to reduce the real-time inconsistency between the observed and predicted system states. As illustrated in Fig. 1, the AHP model is triggered if the time-dependent error, measured by the discrepancy between the observed and predicted link traffic counts at specific time points, exceed a pre-defined threshold. The AHP determines the relative weights of the TDPs in the current stage \( \sigma \), which are used to rank order the TDPs according to their significance to inconsistency. A subset of TDPs with the top \( r \) ranks (i.e., 1 through \( r \) out of the total \( f \) paths in the current stage) is chosen as the dominant path set, denoted by \( c \). All O-D pairs that contain the dominant paths are used to determine various inputs to the A-CONS model, including a predicted link-path incidence matrix \( t_c^\sigma \), a predicted transition matrix \( \Phi^\sigma \), a predicted vector \( t_c^\sigma \) denoting the time-dependent number of travelers on the paths \( c \), and the observed link traffic counts \( x^\sigma \). The A-CONS model then uses the above inputs to determine the vector of “corrected” path flows in the current stage that minimize the time-dependent prediction errors in the network. The TDPs with the lower rankings \( r+1 \) to \( f \) are not used in the inconsistency correction procedure.
Fig. 1.
*An Analytic Hierarchical Approach Framework for the Consistent Real-Time DTA Problem*

![Diagram showing the AHP-Based Framework for the Consistent Real-Time DTA Problem](image)

STAGE $\sigma$

- $\eta_{l}^{l+1}$
- $\eta_{l}^{l}$
- $\eta_{l-l+1}$

Roll period (l units)

Stage length (h units)

$\eta = \sigma-1$

STAGE $\sigma+1$

- $\eta_{l}^{l}$
- $\eta_{l}^{l+1}$
- $\eta_{l}^{l+h}$

$\varphi =$ Time needed to solve the CONS model for stage $\sigma$

$\kappa =$ Time needed to solve the RH-DTA model for stage $\sigma+1$

$\gamma =$ Time at which predictions are available for stage $\sigma+1$

Fig. 2.
*The Rolling Horizon Framework for the Consistent On-Line DTA Problem*
The rest of the procedure is essentially identical to the framework in Peeta and Bulusu (1999). The rolling horizon based DTA (RH-DTA), as illustrated in Fig. 2, is used to solve for the predicted traffic states for stage \( \sigma \). It is used to predict the link traffic counts towards the end of stage, and determine the path assignments for the roll period of this stage. A comprehensive description of the RH-DTA algorithm and its implementation is provided in Peeta and Mahmassani (1995).

In the proposed A-CONS model, a subset of dominant TDPs is identified and utilized to correct inconsistency. Thereby, the dimensions of the A-CONS model and the corresponding computational effort are reduced substantially. The key aspect of A-CONS model is how to ensure that the significant factors influencing the consistency are appropriately identified for the rank ordering of the paths.

### 3.2. The Problem Statement

A traffic network is represented by a directed graph \( G(N, A) \) where \( N \) is the set of nodes and \( A \) is the set of directed arcs. Denote by \( \hat{\mathbf{r}}^f \) a \( f \)-dimensional vector of predicted (time-dependent) number of travelers on \( f \) paths up to time \( t \), \( \mathbf{tt}^f \) a vector of predicted (time dependent) travel times on \( f \) paths, and \( \Psi^f_k \) a vector of number of links on each path \( k \).

Denote by \( \hat{\mathbf{d}}^o \) the predicted number of vehicles at the origin at time \( t = \gamma - \kappa - \varphi \) (see Fig. 2) in stage \( \sigma \). The AHP model seeks rank order the TDPs in terms of their contribution to the real-time consistency.

### 3.3. The AHP Model

This section discusses the AHP model for identifying a subset of dominant paths according to their significance in contributing to inconsistency for the real-time DTA problem. First, a pairwise comparison matrix is developed for each objective in AHP. An objective represents a factor contributing to inconsistency. In A-CONS two factors are considered, O-D demand and TDPs for each O-D pair. The two factors are modeled in the following hierarchy approach. At the first hierarchy level, a pairwise comparison matrix of all O-D pairs is established based upon the demands. This incorporates the effect of how O-D demand influences the path characteristics. Suppose the rest factors have equal effects, an O-D with larger demand has more contribution to congestion (link and path flows) and inconsistency (link and path flow errors between observed and predicted states). The second hierarchy level addresses the relative importance of a path in an O-D pair relative to consistency. The weight for each TDP is determined and its priority is computed based upon its contribution to the inconsistency in a specific O-D pair.

In the first hierarchy, the O-D comparison matrix has the following structure:

\[
\hat{\mathbf{d}}^o = \begin{bmatrix}
\hat{d}_{11} & \hat{d}_{12} & \cdots & \hat{d}_{1p} \\
\hat{d}_{21} & \hat{d}_{22} & \cdots & \hat{d}_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{d}_{k1} & \hat{d}_{k2} & \cdots & \hat{d}_{kp}
\end{bmatrix}
\]
where \( b \) is the number of O-D pairs. \( \hat{D}_p \) is an element of the matrix specifying the ratio of demand between O-D pairs \( p \) and \( q \) by Eq. (7):

\[
\hat{D}_p = \frac{\hat{D}_q}{D_q}
\]

(7)

where \( \hat{D}_j = \sum_{i=1}^{k_j} D_{i,j} \) denotes the number of paths for O-D pair \( j \), and \( \hat{D}_j \) is the total predicted time-dependent O-D demand for O-D pair \( j \) at time \( t \) in stage \( \sigma \).

\( \hat{\Theta}^{\sigma} \) is an \( b \times b \) square matrix with diagonal elements of value 1, and satisfies the reciprocal property (see Eq. (8)).

Due to the reciprocal feature of \( \hat{\Theta}^{\sigma} \), the vector of relative weights \( \hat{w}^{\sigma} \) can be computed by solving the equation \((\hat{\Theta}^{\sigma} - \lambda I)\hat{w}^{\sigma} = 0 \). As discussed in Section 2.2, the eigenvector associated with the maximum eigenvalue \( \lambda_{max} \) specifies the weight \( \hat{w}^{\sigma}_{max} \). The normalized vector is unique, and also \( \sum_{i=1}^{k} \hat{w}^{\sigma}_{max} = 1 \).

\[
\hat{w}^{\sigma}_{i,max} = \frac{\hat{w}^{\sigma}_{i,max}}{\sum_{i=1}^{k} \hat{w}^{\sigma}_{i,max}}, \quad \text{where} \quad \hat{w}^{\sigma}_{i,max} \quad \text{is the element in vector} \ \hat{w}^{\sigma}, \ \sum_{i=1}^{k} \hat{w}^{\sigma}_{i,max} \quad \text{is the summation of all the elements in the vector} \ \hat{w}^{\sigma}_{i,max}. \ \text{The normalized vector is unique, and also} \ \sum_{i=1}^{k} \frac{\hat{w}^{\sigma}_{i,max}}{\sum_{i=1}^{k} \hat{w}^{\sigma}_{i,max}} = 1.
\]

(8)

The second hierarchy level addresses the relative importance of a path in an O-D pair relative to consistency. For each O-D pair \( s \), a comparison matrix \( \hat{c}^{\sigma}_{i} \) is developed that comprises of \( k_s \) paths for \( s \). Hence, \( n \) comparison matrices are developed. An element of the path comparison matrix for \( s \) is given by \( \hat{c}^{\sigma}_{i,j} \), where \( i, j \in k_s \) and is computed as follows:

\[
\hat{c}^{\sigma}_{i,j} = \frac{\hat{r}^{\sigma}_{i,j}}{\hat{r}^{\sigma}_{j,i}}
\]

(9)

The factor \( \hat{r}^{\sigma}_{i,j} \) for path \( i \in k_s \) in the O-D pair \( s \) is computed as:

\[
\hat{r}^{\sigma}_{i,s} = \frac{\hat{R}^{\sigma}_{s,i}}{\hat{T}^{\sigma}_{s,i}/\Psi^{\sigma}_{s,i}}
\]

(10)

where

\( \hat{R}^{\sigma}_{s,i} = \) the predicted number of vehicles on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

\( \hat{T}^{\sigma}_{s,i} = \) the predicted travel time on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

\( \Psi^{\sigma}_{s,i} = \) number of links on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

The second hierarchy level addresses the relative importance of a path in an O-D pair relative to consistency. For each O-D pair \( s \), a comparison matrix \( \hat{c}^{\sigma}_{i} \) is developed that comprises of \( k_s \) paths for \( s \). Hence, \( n \) comparison matrices are developed. An element of the path comparison matrix for \( s \) is given by \( \hat{c}^{\sigma}_{i,j} \), where \( i, j \in k_s \) and is computed as follows:

\[
\hat{c}^{\sigma}_{i,j} = \frac{\hat{r}^{\sigma}_{i,j}}{\hat{r}^{\sigma}_{j,i}}
\]

(9)

The factor \( \hat{r}^{\sigma}_{i,j} \) for path \( i \in k_s \) in the O-D pair \( s \) is computed as:

\[
\hat{r}^{\sigma}_{i,s} = \frac{\hat{R}^{\sigma}_{s,i}}{\hat{T}^{\sigma}_{s,i}/\Psi^{\sigma}_{s,i}}
\]

(10)

where

\( \hat{R}^{\sigma}_{s,i} = \) the predicted number of vehicles on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

\( \hat{T}^{\sigma}_{s,i} = \) the predicted travel time on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

\( \Psi^{\sigma}_{s,i} = \) number of links on path \( i \) of O-D pair \( s \) at time \( t \) in stage \( \sigma \)

The structure of \( \hat{r}^{\sigma}_{i,s} \) in Eq. (10) is based on the factors that determine the relative importance of one path compared to the others in terms of contribution to consistency. The numerator, \( \hat{R}^{\sigma}_{s,i} \), is a straightforward indicator of the path’s contribution to the traffic flow. A large value of \( \hat{R}^{\sigma}_{s,i} \) implies a larger upper bound on the potential error between predicted and observed path flows. A better measure relative to consistency would be the errors in the link or path flow counts rather than just the predicted path flow \( \hat{R}^{\sigma}_{s,i} \). The denominator represents the average travel time per link of path \( i \). \( \Psi^{\sigma}_{s,i} \) is used to remove the potential bias introduced by the number of links in a path. Thereby, congestion, rather
than the distance between the origin and destination, is the focus of consistency. A large denominator value implies more severe congestion on path $i$ due to the congestion on one or more of its links.

Based on Eqs. (9) and (10), the path comparison matrix is also a reciprocal matrix and has the following structure:

$$
\hat{C}_{\sigma} = \begin{bmatrix}
\frac{1}{\hat{c}_{i1}} & \hat{c}_{i2} & \cdots & \hat{c}_{ik}
\frac{1}{\hat{c}_{i2}} & \frac{1}{\hat{c}_{i1}} & \cdots & \hat{c}_{ik}
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{\hat{c}_{ik}} & \frac{1}{\hat{c}_{ik}} & \cdots & \frac{1}{\hat{c}_{ik}}
\end{bmatrix}
$$

(11)

The weights of the path comparison matrix are obtained by solving the equation $(\hat{C}_{\sigma} - \lambda_{max,1})\hat{\Upsilon}_s = 0$. The eigenvector is normalized to obtain the relative weights of the paths for the O-D pair $s$, as shown in Eq. (12):

$$
\hat{\Upsilon}_s = \begin{bmatrix}
\frac{\hat{u}_{i1}}{\sum_{i=1}^{K} \hat{u}_{is}} \\
\frac{\hat{u}_{i2}}{\sum_{i=1}^{K} \hat{u}_{is}} \\
\vdots \\
\frac{\hat{u}_{ik}}{\sum_{i=1}^{K} \hat{u}_{is}}
\end{bmatrix}
$$

(12)

where $\sum_{i=1}^{K} \hat{u}_{is}$ is the sum of values in the eigenvector associated with the maximum eigenvalue for the O-D pair $s$. The maximum eigenvalues and the corresponding eigenvectors are obtained for all O-D pairs and normalized. The matrix $\hat{U}^{\sigma}$, defined as the relative weights for all $S$ O-D pairs, is of the dimension $(K \times S)$, where $K$ equals $\sum_{s} k_s$.

The path priority values in Eq. (13) are used to determine the rankings of the TDPs in the network according to their significance to enabling consistency. These paths are used to determine the predicted state of the system $(\hat{R}^{\sigma}_s, \hat{I}^{\sigma}_s, \hat{\phi}^{\sigma}_s, \hat{x}^{\sigma}_s)$ for the A-CONS model. The subscript “c” indicates that only dominant paths are used to predict the system state. $\hat{R}^{\sigma}_s$ is a vector of predicted (time-dependent) number of travelers on the set of dominant paths, $\hat{I}^{\sigma}_s$ is a predicted time-dependent enhanced link-path incidence matrix. It indicates the proportion of the number of travelers on a set of TDPs which passes a given link. The elements of $\hat{I}^{\sigma}_s$ are fractions between 0 and 1. $\hat{\phi}^{\sigma}_s$ is a predicted transition matrix which indicates the time-dependent existence of paths for every O-D pair. It consists of binary values 0 and 1, where 1 indicates the existence of a path. $\hat{x}^{\sigma}_s$ is a vector of predicted link count.

The simulation-based numerical experiments in Section 4 explore the impact of the set of dominant paths in terms of how much they can correct the discrepancy between the observed and predicted system states (in terms of travel times and the consistency indices).

3.4. Algorithm for Consistent Real-Time DTA Problem Including Path Prioritization

This section discusses the solution procedure of the A-CONS model leveraging the set of dominant TDPs identified by AHP, as illustrated in Fig. 3. The algorithm for consistent real-time DTA consists of two components: (i) a rolling horizon based
deterministic DTA solution algorithm (Peeta and Mahmassani, 1995), and (ii) an AHP-based consistency solution algorithm. The consistency solution component fits seamlessly with the notion of execution cycles within the context of real-time DTA system operations. It can be triggered at discrete (and unequal) time points based on the logic for consistency check internally embedded with the real-time DTA solution algorithm, as discussed hereafter.

Solution Algorithm

**Step 0:** Set stage \( \sigma = 0 \). It represents the time at the start of the planning horizon of interest.

**Step 1:** Obtain the O-D demand for stage \( \sigma + 1 \) and information on current network conditions at time \( t = \gamma - \kappa \) (see Fig. 2) prior to the start of stage \( \sigma + 1 \). Initial paths for users in stage \( \sigma + 1 \) are obtained either exogenously from historical data or from current network conditions.

**Step 2:** Execute the deterministic RH-DTA algorithm for the stage \( \sigma + 1 \) at time \( \kappa \) before in stage \( \sigma \) to determine the set of predicted path flow vector \( \hat{R}^\sigma \) for stage \( \sigma + 1 \).

**Step 3:** AHP Algorithm

Let \( \sigma = \sigma + 1 \). From Step 2, obtain the predicted state of the network \( \hat{R}^\sigma \), in terms of number of users on each TDP, \( \hat{T}^\sigma \), the predicted time dependent travel times on each path, the predicted time-dependent O-D demand for each O-D pair \( \hat{\Phi}^\sigma \), the number of links on each path \( \Psi^\sigma \), and link-path incidence matrix \( \hat{I}^\sigma \) for time \( t = \gamma - \kappa - \varphi \) of the current stage \( \sigma \). Also, obtain the corresponding predicted link traffic counts \( \hat{X}^\sigma \) from the deterministic DTA solution.

**Step 3.1:** Using \( \hat{R}^\sigma \), obtain the pairwise comparison matrix \( \hat{\Theta}^\sigma \) using Eq. (8).

**Step 3.2:** Obtain the eigenvalues and eigenvectors of the pairwise comparison matrix by solving the equation \( [\hat{\Theta}^\sigma - \lambda I] \hat{W}^\sigma = 0 \). The maximum eigenvalue and the corresponding eigenvector are used to obtain the relative weights of the O-D pairs. The eigenvector is normalized to obtain the normalized relative weights \( \hat{w}_{max}^\sigma \) of all O-D pairs.

**Step 3.3:** Using \( \hat{R}^\sigma \), \( \hat{T}^\sigma \), and \( \Psi^\sigma \) determine the path comparison matrix \( \hat{C}^\sigma \).

**Step 3.4:** The maximum eigenvalue and the corresponding eigenvector of the path comparison matrix \( \hat{C}^\sigma \) for each O-D pair are used to provide the relative weights of the paths for each O-D pair. The relative weights are normalized by Eq. (11). The matrix \( \hat{U}^\sigma \) denotes the normalized path relative weights for all O-D pairs.

**Step 3.5:** Compute the vector of path priorities \( \hat{V}^\sigma \) using Eq. (12). They are used to rank TDPs in the network at time \( t = \gamma - \kappa - \varphi \) in stage \( \sigma \).

**Step 3.6:** A subset of TDPs with the top rank orders determines the dominant path set in the network. These TDPs are used to obtain the predicted network state described by \( \hat{R}^\sigma \), \( \hat{I}^\sigma \), \( \hat{\Phi}^\sigma \), \( \hat{X}^\sigma \), which is used as input to the A-CONS in Step 6.1.

**Step 4:** Implement the path assignments for the roll period of stage \( \sigma \).

**Step 5:** Obtain the observed link traffic counts, \( X^\sigma \) at time \( t = \gamma - \kappa - \varphi \) in the current stage. If \( |X^\sigma - \hat{X}^\sigma| \leq \varepsilon \) for all links
in the network (where \( \epsilon \) is a positive scalar specifying a threshold), consistency is preserved; go to Step 7. Otherwise, go to Step 6.

**Step 6:** Use Consistency Solution Algorithm (Peeta and Bulusu, 1999) to obtain the corrected network state \( \tilde{R}^{\sigma} \) at time \( t = \gamma - \kappa \). The consistency solution algorithm applies a generalized singular value decomposition (GSVD) approach to determine \( \tilde{R}^{\sigma} \), \( \tilde{I}^{\sigma} \), and \( \tilde{\Phi}^{\sigma} \) simultaneously that are consistent with the observed link counts \( x^{\sigma} \).

**Step 6.1:** \( \tilde{I}_i = \tilde{I}^{\sigma} \), \( \tilde{R}_i = \tilde{R}^{\sigma} \), \( \tilde{X}_i = \tilde{X}^{\sigma} \), where \( \sim \) represents the corrected states. \( \tilde{I}_i \) is a vector of corrected time-dependent enhanced link-path incidence matrix at the initial iteration; \( \tilde{R}_i \) is a vector of corrected time-dependent transition matrix at the initial iteration; \( \tilde{X}_i \) is a vector of corrected number of travelers on TDPs at the initial iteration; and \( \tilde{X}_s \) is a vector of correct link counts at the initial iteration. Also, observed link counts \( x^{\sigma} \) and O-D demand \( d^{\sigma} \) at time \( t \) are obtained. Set iteration number \( i = 0 \).

**Step 6.2:** Append \( \delta \) \( I \) to the enhanced link-path incidence matrix \( \tilde{L}_i \), where \( \delta \) is a small scalar, and \( I \) is an identity matrix with dimension of \( \tilde{R}_i \)

\[
\tilde{L}_i = \begin{bmatrix} \tilde{L}_i \end{bmatrix} \quad \text{and} \quad \tilde{\Phi}_i = \tilde{\Phi}_i
\]

**Step 6.3:** Decompose the matrix \( \tilde{L}_i \) and transition matrix \( \tilde{\Phi}_i \) using an orthogonal transformation (Peeta and Bulusu, 1999) to obtain an upper triangular invertible matrix \( M \) and orthogonal matrices \( Q_1 \) and \( Q_2 \) such that \( \tilde{L}_i = Q_1 M \) and \( \tilde{\Phi}_i = Q_2 M \).

**Step 6.4:** Decompose \( Q_1 \) and \( Q_2 \) using the CS decomposition (Peeta and Bulusu, 1999), a simultaneous diagonalization algorithm, to obtain diagonal matrices \( C \) and \( S \).

**Step 6.5:** Obtain the corrected state of the network \( \tilde{R}_{i+1}^{\sigma} \).

**Step 6.6:** Using \( \tilde{R}_{i+1}^{\sigma} \) as the input to the traffic flow model, determine the corresponding link-path incidence matrix \( \tilde{L}_{i+1}^{\sigma} \) and transition matrix \( \tilde{\Phi}_{i+1}^{\sigma} \).

**Step 6.7:** The convergence of this algorithm is assumed if the following criteria are satisfied:

(i) If errors in link traffic counts are within the threshold \( a \) of observed counts, and
(ii) If the error in terms of the average vehicular travel time in the roll period of stage \( \sigma \) is within \( \% \) of the observed time \( T \) at time \( t \).

Criterion (ii) is used only if the travel times can also be observed. If convergence is achieved, go to Step 7. If not, set \( i = i + 1 \), and go to Step 6.2 of the consistency solution algorithm.

**Step 7:** If the end of the planning horizon is reached, stop the rolling horizon procedure. If not, store the positions and other characteristics of all vehicles in the network at \( t = \gamma - \kappa \). This information is obtained from Step 6 if the state is corrected for, or from Step 2 if consistency execution is not required in stage \( \sigma \). Go to Step 1.
4. Simulation Experiments

The performance of the proposed AHP-based CONS model for the real-time DTA is analyzed using simulation-based numerical experiments. The traffic and network characteristics are discussed in Sections 4.1 and 4.2, respectively. The performance measures used to analyze the consistency solution framework are as follows:

1. Percentage error in average travel time:
This index provides the percentage error in average travel time for all vehicles in the roll period of a stage. It provides the percentage deviation of the travel time of the predicted state from the observed state, and is computed as:

$$\frac{T_{\sigma t} - T_{e_{\sigma t}}}{T_{e_{\sigma t}}} \times 100\%$$

where:

- $T_{\sigma t}$ = Observed average travel time at time $t$ in stage $\sigma$
- $T_{e_{\sigma t}}$ = Predicted average travel time at time $t$ in stage $\sigma$

2. Consistency in link traffic counts: is defined as the two-norm of the difference
in the number of users on a link under the observed and predicted states. It is computed as follows:

\[
I_i = \left[ \frac{Z^{a\sigma} - \hat{Z}^{a\sigma}}{\Psi} \right] = \frac{\sum_{i=1}^{\Psi} (Z^{m\sigma}_i - \hat{Z}^{m\sigma}_i)^2}{\Psi} \tag{15}
\]

where:
- \( I_i \) = Link consistency measurement
- \( Z^{a\sigma} \) = Observed state of the network (link count) at time \( t \) in stage \( \sigma \)
- \( \hat{Z}^{a\sigma} \) = Predicted state of the network (link count) at time \( t \) in stage \( \sigma \)
- \( \Psi \) = Total number of links

The consistency measurement indicates the average error per link in terms of the number of users on that link. The normalization allows comparison of the consistency measurement across stages.

3. Path Consistency measurement (\( I_p \)): \( I_p \) is defined similar to the link consistency measurement (Eq. (15)), except that link counts are replaced by number of users on TDPs, and \( \Psi \) is replaced by number of paths. It is pertinent to note that while the first two indices are measurable, current technology precludes complete determination of \( I_p \) in the real world. It is used here primarily to derive insights on the consistency solution algorithm.

4.1. Network Configuration and Traffic Characteristics

As shown in Fig. 4, the test network includes a freeway and the surrounding arterials. It consists of 50 nodes and 168 links, 32 origins (all nodes except freeway nodes 38-43 and 45-50, and nodes 9, 10, 15, 16, 21 and 22) and 10 destinations (nodes 2, 5, 13, 18, 25, 30, 35, 36, 37, 44).

![Fig. 4. The Test Network](image-url)
Table 1 specifies the user class characteristics of the vehicles generated in the network. In the experiments, all four user classes are assumed to be equally represented. The parameters $\alpha$ and $\beta$ denote the threshold errors in terms of link traffic counts and average vehicular travel time, respectively. $\alpha$ and $\beta$ are determined through sensitivity analysis; they are 15% and 6%, respectively, in these experiments. Also, those users who have not reached their destination in the current stage and have an expected remaining travel time greater than 15% of the roll period are used in the consistency analysis.

Table 1
User Class Characteristics

<table>
<thead>
<tr>
<th>User Class</th>
<th>Percentage (%)</th>
<th>Type of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Optimal(^1) (SO)</td>
<td>25</td>
<td>Pre-trip</td>
</tr>
<tr>
<td>User Equilibrium(^2) (UE)</td>
<td>25</td>
<td>Pre-trip</td>
</tr>
<tr>
<td>Boundedly Rational(^3) (BR)</td>
<td>25</td>
<td>Pre-trip + En route</td>
</tr>
<tr>
<td>Pre-Specified(^4) (PS)</td>
<td>25</td>
<td>No Information</td>
</tr>
</tbody>
</table>

1. Equipped users following prescribed system optimal paths. Users who are either unfamiliar with the typical network traffic conditions or affected by a severe incident are likely to follow this information.
2. Equipped users following prescribed user equilibrium paths. Users who are very familiar with the typical network traffic conditions are likely to follow this information.
3. Equipped users following the boundedly rational switching rule in response to descriptive information on the prevailing network conditions. The switching rule states that users switch from their current path at decision points if travel time savings on an alternative route exceed a threshold value.
4. Unequipped users following pre-specified paths. These paths are either determined from historical data (representing past network experience) or solved for exogenously based on current conditions. It is assumed that in the absence of information, users do not switch routes en route.

The AHP-based framework is analyzed at the third stage of the planning horizon, which comprises of 1,560 paths and 4,031 users, as shown in Table 2. It is assumed that traffic conditions are normal and no incidents occur in the network over the horizon of interest.

Table 2
Stage Traffic Characteristics

<table>
<thead>
<tr>
<th>Stage Number</th>
<th>Number of Paths</th>
<th>Number of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1377</td>
<td>3008</td>
</tr>
<tr>
<td>3</td>
<td>1560</td>
<td>4031</td>
</tr>
<tr>
<td>4</td>
<td>1430</td>
<td>3031</td>
</tr>
<tr>
<td>5</td>
<td>1458</td>
<td>3072</td>
</tr>
<tr>
<td>6</td>
<td>746</td>
<td>1245</td>
</tr>
</tbody>
</table>

4.2. Scenarios Addressed

Two scenarios are considered to investigate the effectiveness of the proposed A-CONS model.

Scenario 1: The effectiveness of the solution algorithm with the dominant path set is investigated for three different sets of rank-ordered paths. The first set comprises of the most dominant paths (150 paths), ranked...
from 1 to 150. The second set consists of paths ranked 250 to 400, and the third set comprises of the least dominant paths that are rank ordered from 1,410 to 1,560. The objective is to analyze whether more dominant paths are more effective in correcting the gaps between the predicted and observed states.

**Scenario 2:** The objective of this experiment is to study the marginal improvement in terms of travel time error and the consistency index when more paths are included in the dominant path set. The paths are included in increments of 150 based on their rank-orders.

### 4.3. Analysis of Results

Figs. 5-7 show the results for Scenario 1. Fig. 5 illustrates the AHP-based performance relative to consistency in terms of percentage errors in travel times for the three different sets of rank ordered paths for the 3rd stage. The initial error between the observed and predicted states is 58%. When the A-CONS is used with the paths ranked from 1 to 150, the percentage error of average travel times is reduced to 10% of the observed travel time over 11 iterations. Similarly, the error is reduced to 21% of the observed travel time when paths ranked 250 through 400 are used, and to 47% when the least significant paths (ranked 1,410 to 1,560) are used in the consistency solution procedure. The results indicate a significant improvement in terms of average travel time errors when only the top 150 ranked paths are used.

Figs. 6(a) and 6(b) illustrate the link and path consistency measurements, respectively, for this scenario. The average error per link in terms of traffic counts is less than 4.1 for the first set, about 4.6 for the second, and about 5.0 for the third. Similarly, the average error in terms of number of users on TDPs is 4.2, 4.8 and 5.1 for the three sets, respectively. The results clearly demonstrate the hierarchical nature of the TDPs relative to enabling consistency. They suggest important paradigms for the real-time deployment of DTA for ATIS/ATMS equipped networks in terms of achieving computational feasibility and efficiency.

Scenario 2 investigates the performance in terms of enabling consistency when more paths are included, by analyzing the marginal improvement when each incremental set of paths is added to the cumulative set of paths considered. Table 3 illustrates the number of users in each path set. Each set comprises of 150 paths, except for the last one. In general, the more paths in the dominant set, the more the number of users.

### Table 3

**Path and Vehicle Characteristics: Stage Number =3, Number of Paths = 1560, Number of Vehicles = 4031**

<table>
<thead>
<tr>
<th>Path Ranking Relative to Enabling Consistency</th>
<th>Path Set Number</th>
<th>Number of Vehicles</th>
<th>Percentage of Vehicles in this Path Set</th>
<th>Path Ranking Relative to Enabling Consistency</th>
<th>Path Set Number</th>
<th>Number of Vehicles</th>
<th>Percentage of Vehicles in this Path Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 150</td>
<td>1</td>
<td>1014</td>
<td>25.16</td>
<td>751 – 900</td>
<td>6</td>
<td>277</td>
<td>6.87</td>
</tr>
<tr>
<td>151 - 300</td>
<td>2</td>
<td>629</td>
<td>15.60</td>
<td>901 – 1050</td>
<td>7</td>
<td>241</td>
<td>5.98</td>
</tr>
<tr>
<td>301 - 450</td>
<td>3</td>
<td>487</td>
<td>12.08</td>
<td>1051 – 1200</td>
<td>8</td>
<td>238</td>
<td>5.90</td>
</tr>
<tr>
<td>451 – 600</td>
<td>4</td>
<td>398</td>
<td>9.87</td>
<td>1201 – 1350</td>
<td>9</td>
<td>220</td>
<td>5.46</td>
</tr>
<tr>
<td>601 – 750</td>
<td>5</td>
<td>314</td>
<td>7.79</td>
<td>1351 - 1560</td>
<td>10</td>
<td>213</td>
<td>5.28</td>
</tr>
</tbody>
</table>
Fig. 7 illustrates the percentage difference in travel times between the observed and predicted network states with cumulative increments of ordered TDPs. A significant reduction in errors is obtained when the first set of rank-ordered paths is used as they involve about 25% of the demand. Also, by the 8th iteration, all cumulative paths sets other than the first 150 paths achieve convergence (within 6.25%). Including more paths in the consistency procedure significantly reduces the number of iterations to convergence, but increases the computational time. The total reduction in travel time errors is 11.4% when all paths in the network are used in the consistency procedure. Fig. 7 provides important insights for real-time paradigms and strategies that seek to trade-off computational efficiency and performance (in terms of consistency). As more paths are considered, the number of iterations required to achieve convergence also decreases. Thereby, when around 900 paths (58% of all paths) are considered, convergence is achieved in 4 iterations as compared to 3 iterations when all paths are considered. Also, because of the reduced number of paths, 4 iterations may take significantly less time than 3 iterations in which all paths are considered. Hence, the path set that considers paths 1 through 900 is a superior real-time consistency strategy than the one considering all paths. This improves the computational performance for on-line deployment while maintaining similar convergence characteristics. For the same reason, paths 1-600 and 1-750 are also attractive alternatives.

Figs. 8 and 9 illustrate the link and path consistency measurements, respectively. They reiterate the insights for real-time paradigms observed in Fig. 7. The average error per link is within 4 vehicles at convergence when 300 or more paths are used in the A-CONS model. Similarly, the average error in terms of number of users on each path is within 5 users at convergence for all paths. The AHP model uses about 1 to 5 seconds of CPU time depending on the number of paths. These results indicate that use of a subset of paths in the consistency solution procedure only marginally deteriorates the system performance in terms of travel time errors and consistency measurements while substantially improving the computational efficiency.

**Fig. 5.**
Comparison of the Three Sets of Rank Ordered Paths Relative to Enabling Consistency, in Stage 3

**Fig. 6a.**
Comparison of Link Consistency Indices for the Three Sets of Rank Ordered Paths, in Stage 3
5. Concluding Comments

The AHP is a widely-used approach for decision-making involving multiple criteria. In this paper, the AHP is used to prioritize the TDPs in a network for enabling on-line consistency. From an operational perspective, the AHP model forms a pre-processor for the CONS model and determines a subset of dominant paths that significantly contribute to enhancing the consistency. An eigenvector based approach is used to identify the relative weights of the demand characteristics of O-D pairs and the path flow characteristics in relation to the TDPs in the network. The relative weights are used to rank order the paths.

The experimental insights confirm a hierarchical structure for paths in terms of enabling consistency. An important paradigm for real-time deployment is the notion of considering a subset of dominant paths. The results emphasize the identification of such subsets to enhance consistency while significantly improving computational efficiency. The study insights also indicate that the eigenvalues associated
with paths represent parameters that can signify the relative importance of a path for enabling consistency.

The study experiments also establish the value of the hierarchical structure of TDPs in relation to the O-D demand. As illustrated in Table 3, 52% of the demand was observed on 29% of the paths. Therefore, minimizing the error between the observed and predicted states on 29% of the paths significantly improved consistency. The proposed A-CONS model can also be used in a decentralized network that lacks coordination among zones, and where the objective is to improve the local performance as opposed to the overall performance of the network.

References


