Comparison of assignment methods for simulation-based dynamic-equilibrium traffic assignment

Michael Mahut\(^1,2\), Michael Florian\(^2,1\), Nicolas Tremblay\(^1\)

\(^1\) INRO Consultants Inc, Montreal, Canada
\(^2\) Centre for Research on Transportation (CRT), University of Montreal, Canada

Dynamic traffic assignment (DTA), which concerns the assignment of a time-varying origin-destination (O-D) demand matrix to a transportation network, is a growing field of transportation science that was originally motivated by the desire to extend static network assignment algorithms to a time-varying context. As with static assignment, dynamic assignment has stimulated a strong interest from academics and engineers due to its potential to solve, or partially solve, complex real-world transportation problems that have not been addressed satisfactorily by other methodologies.

A number of distinct DTA model formulations have been proposed, which can be differentiated by the properties of the assignment (e.g., equilibrium vs. non-equilibrium), assumptions about drivers' awareness of traffic conditions (e.g., average daily conditions vs. unexpected conditions due to incidents), as well as contextual aspects of the problem being solved (e.g., off-line planning vs. real-time analysis).

DTA models can be used to evaluate impacts of changes to the network infrastructure that may be long term (new roads) or short term (temporary loss of infrastructure; road closures and detours due to construction or maintenance), or changes to network control or management (traffic signals and ramp meters; reserved lanes). Demand scenarios also vary widely, from long-term changes in land-use, to special events (e.g. sporting events), to evacuation planning. Other applications include the study of network performance under unexpected traffic incidents, particularly in critical incident locations (bridges and tunnels), and the off-line development of response plans for such incidents. These different applications require various levels of detail in the DTA models or model components that should be used to solve them.

Generally speaking, a DTA model can be broken down into two components: a routing model, which determines how path flows vary with travel times (or potentially with a generalized cost that includes travel time), and a traffic model, which determines how these path demands interact in the network, giving rise to congestion and thus travel time. Early work in this field was focused on developing time-varying extensions of the static network assignment paradigms, and thus used the notion of link-cost functions to represent the impact of traffic flow on link travel time. Due to the complexity of many of the problems mentioned above, the last decade has seen a distinct trend towards DTA models with a
more detailed and realistic representation of the transportation system and traffic phenomena. The level of detail is mainly determined by the choice of traffic model used, which may be based on hydrodynamic theory (e.g. cell-transmission method), approximations thereof (link speed-density functions), or microscopic models, which consider the interactions between individual vehicles on roadways and at intersections.

These dynamic traffic models share a basic property, which is their ability to reproduce a fundamental relationship between traffic flow, speed and density. Of particular significance, in the context of traffic assignment, is that this relationship exhibits a bi-modal relationship between traffic flow and speed. Low flows may be accompanied by high speeds (free-flow conditions), or by low speeds (congested conditions). Another significant feature of realistic traffic models is that of congestion spill-back (or blocking back), as a result of which the link travel times are no longer separable.

As a result of these properties, the assignment map is discontinuous and difficult to characterize analytically, and the existence and uniqueness of an equilibrium solution are no longer guaranteed. Nevertheless, algorithms inspired from static network assignment have been found to work in practice for finding approximate dynamic equilibrium conditions on real-world networks of significant size. DTA (routing) algorithms reported in the literature for use with dynamic traffic models generally fall into two categories: path-based and splitting-rate formulations. Splitting rates are time-varying turning-movement proportions that are defined by destination.

This paper reports on the evaluation of alternative routing models used for pre-trip assignment (expected traffic conditions) with a detailed traffic simulation model. The tests are executed on several real-world networks of significant size (between 1000 and 6000 network links) from North America and Europe.

One dynamic network equilibrium model that we consider is formulated in the space of path flows. The time varying path flows, denoted \( h_k(t) \), are defined over a demand period \((0, T)\) and the time varying path costs satisfy the dynamic user equilibrium conditions as follows:

\[
\begin{align*}
  s_k(t) = u_i(t) & \text{ if } h_k(t) > 0 \\
  s_k(t) \geq u_i(t) & \text{ otherwise } \\
  \forall k \in K_i, i \in I, t \in (0, T)
\end{align*}
\]

where \( s_k(t) \) is the path travel time, \( I \) is the set of all O-D pairs, \( K_i \) is the set of paths connecting O-D pair \( i \), and \( u_i(t) \) is the minimum path travel time for O-D pair \( i \), defined as

\[
u_i(t) = \min_{k \in K_i} \left\{ s_k(t) \right\} \forall k \in K_i, t \in (0, T)\]
The path flows, comprising the set $\Omega$, satisfy conservation of flow and non-negativity constraints:

$$\Omega = h(t) : \sum_{k \in K} h_k(t) = g_i(t), i \in I; h_k(t) \geq 0, t \in (0,T)$$

The above conditions have been shown to be equivalent to a variational inequality problem (by Friesz et al, 1993), which is to find $h^*$ in $\Omega$, such that:

$$\left(S(h^*), h - h^* \right) \geq 0, \forall h \in \Omega$$

Averaging methods, which are inspired from optimization methods used in static equilibrium assignment algorithms, are used to solve a discretized version of this problem. These methods include variants in which the step size varies for each O-D pair, as a function of the path travel times.

An alternative formulation is in the space of time-varying splitting rates, denoted $f_j^m(t)$, which is the percentage of vehicles headed for destination $j$ on the upstream (incoming) link of turning movement $m$, that uses this turning movement $m$, at time $t$. (N.B. Models of this kind in the literature are formulated in the space of link flows leaving a node, rather than movement flows leaving a link: this makes no difference to the formulation, it only changes the graph to which it is applied). The user equilibrium conditions can be stated as:

$$c_j^m(t) = d_j^l(t) \text{ if } f_j^m(t) > 0$$
$$c_j^m(t) \geq d_j^l(t) \text{ otherwise}$$
$$\forall m \in M_j, l \in L, j \in J, t \in (0,T)$$

where $c_j^m(t)$ is the travel time from movement $m$ to destination $j$, $M_j$ is the set of turning movements leaving link $l$, $L$ is the set of network links, $J$ is the set of destinations, and $d_j^l(t)$ is the minimum travel time from link $l$ to destination $j$ at time $t$, defined as:

$$d_j^l(t) = \min_{m \in M_j} \left[d_j^m(t) \right] \forall t \in (0,T)$$

The splitting rates satisfy conservation of flow and non-negativity constraints:

$$\sum_{m \in M_j} f_j^m(t) = 1$$
$$0 \leq f_j^m(t) \leq 1$$
$$l \in L, j \in J, t \in (0,T)$$
The model definition is completed with the specification of the time-varying O-D demands. A discretized version of this model formulation is solved using an original approach that is based on notions from traffic flow theory. In particular, the following properties of real traffic are considered:

- The relationship between link flow and travel time is bi-modal: travel times at low flows may either be low (no congestion) or high (congestion);
- As a function of link out-flow, there is a corresponding maximum density and maximum travel time, which will be observed if the outflow remains constant for a period of time;

The routing algorithms are evaluated primarily on the proximity of the converged assignment to the dynamic equilibrium conditions. Other measures of algorithm performance include the speed of convergence (number of iterations), computational burden (per iteration), and computer memory requirements. Applicability of these models to solving *en-route* assignment problems, that involve unexpected traffic conditions due to incidents, is also discussed.