The (Near-Term) Future of Dynamic Traffic Assignment

Terry L. Friesz

7^{th} International DTA Symposium

6-8 June 2018

What is DTA?

- DTA modeling concerns both descriptive (positive, behavioral) and prescriptive (normative) modeling.
- Idealy travel demand theory and traffic flow theory are "integrated".
- DTA models emphsize passenger transport.
- The intention is to control and/or predict traffic patterns on vehicular networks.

From a collegial point of view, DTA is “nicely making way”.

1. ORSA/TIMS 1980 Washington, DC (The first DTA paper session?)
2. 12th ISTTT 1993 Berkeley: first time there were lots of DTA papers
3. DTA2006 Leeds University, England
4. DTA2008 Katholieke Universiteit of Leuven, Belgium
5. DTA2010 Takayama, Japan
6. DTA2012 Martha’s Vineyard, US
7. DTA2014 Salerno, Italy
8. DTA2016 Sydney, Australia
9. DTA2018 Hong Kong, China
10. DTA2020 Seattle, US

Differential Game Theory: A Long Time in the Making

1. Differential game theory (DGT) provides an excellent vocabulary for discussing and understanding DTA.
2. The lineage of DGT is impressive, ancient, and still highly relevant:
    1. connects the 17th through 21st centuries, and
    2. involves Legendre, Newton, the Bernoulli brothers and their sister, l'Hôpital, Euler, Cournot, Stackelberg, Nash, Isaacs, Bellman, Pontryagin, Leitmann, Basar ...
    3. When studying dynamic optimization and dynamic games, do we have the right to be ignorant of the work of these luminaries?
3. The foundations of DGT are quite easy to master if one takes an introductory course oriented toward engineers and computation. Such a course may be offered online in the not too distant future ...

Differential Games: A Language for DTA

1. Differential game theory (DGT) involves
   1. noncooperative games: Cournot-Nash and Stackelberg
   2. state variables and control variables
   3. explicit dynamics
   4. side constraints
   5. feedback
   6. learning/competition, and
   7. multiple timescales.
2. As such DGT allows DSO, DUE and Mixed-DSO-DUE flows to be articulated in a natural rather than ad hoc way.
3. DGT is intertwined with the calculus of variations and modern optimal control theory, and has established computational methods.
4. Differential variational inequalities (DVIs) allow the powerful results of DGT to be exploited.

Comment 1

1. To better portray needed improvements in DTA modeling and computation, we need to look at a specific DTA model.
2. Dynamic user equilibrium (DUE) with simultaneous route and departure (SRD) choices remains one of the most widely studied DTA models.
3. SRD DUE requires selection of a departure time and route that results in equal effective delay for utilized paths between each given origin-destination pair.
4. We will consider the weaknesses of the DUE paradigm as vehicular networks and their supporting information technology evolve.
5. We will ask the question: Can the DUE paradigm be modernized? If not, what will replace it?

What is Dynamic User Equilibrium?

1. Wardrop's First Principle (WFP) is the dominant notion for static traffic assignment of passenger vehicles.
2. Dynamic user equilibrium as an extension of WFP. Does it make sense?
3. In particular, how can equilibrium be dynamic? (Samuelson's "moving equilibrium")
4. Is dynamic user equilibrium (DUE) dead?
5. There are various formulations of DUE, but all involve some dynamic extension of Wardop's First Principle (WFP).

Some Obviously Needed (but not easy) Extensions of Dynamic User Equilibrium (DUE)

1. Delay "operators" to replace the DNL problem while maintaining the simultaneous, consistent solution of the DNL and DUE problems
2. Greater realism in the DNL problem, especially mixed classes of flow that reflect autonomous and nonautonomous vehicles, including class-to-class transitions
3. Bounded rationality
4. Integration of Within-Day and Day-to-Day Timescales ♣
5. Collaborative games ("between" full noncooperation and full cooperation) not requiring equilibrium to be achieved.

Relaxation of SRD DUE Foundation Assumptions

1. Differential Nash game among users
2. Open loop
3. Atomic
4. Deterministic
5. Perfect information
6. Today the above define the problem class principally studied by "DTA/DUE scholars."
7. Main topic of debate has been what link-based traffic model to employ and how should that traffic model be integrated with the DUE formalism.
8. Shift to a more general problem class – relaxing some of items 1 thru 5 above – is imminent.
9. One would expect that if the DUE problem may be solved, then the dynamic system optimal (DSO) is solvable. There are complications, however. Foremost is the practicality of differentiating the effective delay operator.

Some Terminology: Differential Algebraic System, Slide 1

- An illustrative differential algebraic equation (DAE) is

$$F(x, \frac{dx}{dt}, u, t) = 0 \text{ plus boundary conditions}$$

- An illustrative partial differential algebraic equation (PDAE) is

$$F(x, \frac{\partial\rho(x,t)}{dx}, \frac{\partial\rho(x,t)}{\partial t}, u, t) = 0 \text{ plus boundary conditions}$$

- A constrained set of ordinary (or partial) differential equations is also a DAE (or PDAE) system.

Some Terminology: Differential Algebraic System, Slide 2

- A key issue in DAE theory is index reduction, where "index" is the number of differentiations required to obtain an implicit ODE system from a DAE system.
- DAE systems may be solved by:
  - conversion to a system of ODEs
  - finite element methods
- We express dynamic network loading (DNL) models as (P)DAE systems.
- There is a significant literature on (P)DAEs.

Comment 2

- To avoid mathematical detail we are going to asume a DNL model is formulated, and employ an effective delay operator that is derived from it.
- The DNL model: based on the so-called Lighthill-Whitham-Richards (LWR)/hydrodynamic theory of traffic, extended to a network. That has been done in different ways by several scholars.

Mathematical Road Map, Slide 1

- We will be employing symbols, equations, and inequalities along with notions like delay, optimization, games and equilibrium. In every case I will give a prose summary of the meaning of the mathematics presented.
- Mathematical games are contests over resources conducted according to some set of rules and offering some sort of payoffs.
- Why continuous time?
  - It is easy to formulate problems and notation is much simpler.
  - Allows immediate representation of two timescales through the use of time shifts, $t + \delta$
  - Computation in continuous time is possible, and holds much promise.

Mathematical Road Map, Slide 2

- The flow of the subsequent material is really quite simple; it will be something like this:

  Drivers selfishly optimize own delay
  $\Longrightarrow$ drivers are agents in a <u>noncooperative game</u>
  $\Longrightarrow$ models as <u>equations and inequalities</u>
  $\Longrightarrow$ <u>equations and inequalities are manipulated</u>
  $\Longrightarrow$ <u>recognizable problem categories</u> (to be named)
  $\Longrightarrow$ <u>numerical and qualitative analyses</u>
  $\Longrightarrow$ <u>understanding, vetting and solution</u> of the models.

SRD DUE with Exogenous Delay: Basic Notation

- Path set, departure rates, and time:

$$\begin{aligned} &\mathcal{P} \text{ is the set of all paths employed by travelers} \\ &t \text{ denotes departure time} \\ &h \text{ is a vector of departure rates (“path flows”)} \end{aligned}$$

- The interval of continuous time:

$$t_0 \leq t \leq t_f$$

- The path unit delay operator:

$$D_p(t, h)$$

SRD DUE with Exogenous Delay: Effective Delay Operators I

- Suppose

$$
\begin{aligned}
T_A &= \text{is the desired arrival time} \\
T_A &< t_f
\end{aligned}
$$

SRD DUE with Exogenous Delay: Effective Delay Operators II

- So-called schedule delay:

$$F\left[t + D_p(t,h) - T_A\right]$$

- The effective unit path delay operator:

$$\Psi_p(t,h) = D_p(t,h) + F\left[t + D_p(t,h) - T_A\right] \qquad \forall p \in P$$

SRD DUE with Exogenous Delay: Flow Conservation

- Fixed trip matrix

$$Q = (Q_{ij} : (i,j) \in \mathcal{W})$$

where

$$\begin{aligned} Q_{ij} &\in \Re^1_{++} \text{ fixed travel demand for OD pair } (i,j) \in \mathcal{W} \\ \mathcal{W} &= \text{the set of all origin-destination pairs} \end{aligned}$$

- Recall the flow conservation constraints from the previous slide:

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \quad \forall (i,j) \in \mathcal{W}$$

where

$\mathcal{P}_{ij}$ = subset of paths that connect origin-destination pair $(i,j) \in \mathcal{W}$.

SRD DUE with Exogenous Delay: DUE Defined

- We define the set of feasible flows by

$$\Lambda_F = \left\{ h \geq 0 : \sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \; \forall (i,j) \in \mathcal{W} \right\}$$

- Dynamic user equilibrium:

Definition

Dynamic user equilibrium $DUE(\Psi, \Lambda_F, t_0, t_f)$. A vector of departure rates (path flows) $h^* \in \Lambda_F$ is a dynamic user equilibrium if

$$h_p^*(t) > 0, p \in P_{ij} \Longrightarrow \Psi_p[t, h^*(t)] = \min_h \Psi_p[t, h(t)] = v_{ij}$$

SRD DUE with Exogenous Delay: Variational Inequality Formulation

- $DUE\left(\Psi, \Lambda_F, t_0, t_f\right)$ is equivalent to the following variational inequality (VI) under mild regularity conditions:

$$\left.\begin{array}{c} \text{find } h^* \in \Lambda_F \text{ such that} \\ \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p(t, h^*)(h_p - h_p^*)dt \geq 0 \\ \forall h \in \Lambda_F \end{array}\right\}$$

- However, as we shall next see, the above is equivalent to a differential variational inequality (DVI), without introducing ODEs for arc dynamics.

SRD DUE with Exogenous Delay: Isoperimetric Constraints

Constraints constitute a two-point boundary-value problem:

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \;\forall (i,j) \in \mathcal{W}$$

$$\Updownarrow$$

$$\begin{cases} \dfrac{dy_{ij}}{dt} = \displaystyle\sum_{p \in \mathcal{P}_{ij}} h_p(t) \quad \forall (i,j) \in \mathcal{W} \\ y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \\ y_{ij}(t_f) = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \end{cases}$$

SRD DUE with Exogenous Delay: DVI Formulation

Thus, the DUE problem is expressible as the following DVI:

$$\left.\begin{array}{c} \text{find } h^* \in \Lambda_F \text{ such that} \\ \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p(t,h^*)(h_p - h_p^*)dt \geq 0 \\ \forall h \in \Lambda_F \end{array}\right\} DVI(\Psi, \Lambda_F, t_0, t_f)$$

where

$$\begin{aligned} \Psi &= \text{the effective delay operator} \\ \Lambda_F &= \left\{ \begin{array}{c} h \geq 0 : \dfrac{dy_{ij}}{dt} = \sum_{p\in P_{ij}} h_p(t), \; y_{ij}(t_0) = 0, \; y_{ij}(t_f) = Q_{ij} \\ \forall (i,j) \in \mathcal{W} \end{array} \right\} \\ Q &= (Q_{ij}) = \text{the trip matrix} \\ t_0 &= \text{ start time} \\ t_f &= \text{ end time} \end{aligned}$$

-

Note that $\forall h \in \Lambda_F$:

$$DVI \Longleftrightarrow \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p dt \geq \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p^* dt$$

$$\Uparrow$$

$$\begin{aligned}
\min J_0 &= \sum_{(i,j)\in\mathcal{W}} v_{ij}\left[Q_{ij} - y_{ij}\left(t_f\right)\right] + \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p dt \\
\frac{dy_{ij}}{dt} &= \sum_{p\in\mathcal{P}_{ij}} h_p\left(t\right) \quad \forall\left(i,j\right)\in\mathcal{W} \\
y_{ij}\left(t_0\right) &= 0 \quad \forall\left(i,j\right)\in\mathcal{W} \\
h &\geq 0
\end{aligned}$$

DUE DVI Analysis with Exogenous Delay

Theorem

Dynamic user equilibrium equivalent to a differential variational inequality. Assume $\Psi_p(\cdot, h)$ is measurable and strictly positive for all paths and all feasible departure. A vector of departure rates (path flows) $h^* \in \Lambda_F$ is a dynamic user equilibrium if and only if h^* solves $DVI(\Psi, \Lambda_F, t_0, t_f)$.

Proof.

Requires some background in the theory of optimal control . . . $\square$

The Notion of a Singular Control, Slide 1

- Consider an optimal control problem with state x, adjoint λ, and control u obeying the constraint

$$L \leq u \leq U \quad \text{where} \quad L, U \geq 0 \quad \text{and} \quad L < U$$

- The Hamiltonian (of Pontryagin's minimum principle) expresses the instananeous primal value of the objective plus the shadow value of as yet unrealized motion.
- Assume the Hamiltonian is linear in its control u:

$$H = \Phi(x, \lambda, t) + S(x, \lambda, t)u$$

- For such a problem the minimum principle requires u^* be "bang-bang-singular":

$$u^* = \begin{cases} L & \text{if} \quad S > 0 \\ U & \text{if} \quad S < 0 \\ u_s & \text{if} \quad S = 0 \end{cases}$$

The Notion of a Singular Control, Slide 2

- Recall we are considering the Hamiltonian to be linear in its control u:

$$H = \Phi(x, \lambda, t) + S(x, \lambda, t)u$$

- If the switching function S vanishes on some $[t_1, t_2] \subseteq [t_0, t_f]$ for $t_1 < t_2$, then u_s is a singular control that must be determined.
- Generally speaking the solution of such a problem requires:
  1. A "bang" control to get on the singular trajectory
  2. Navigation along the singular trajectory
  3. Waiting until the last moment at which another "bang" control may be used to get off the singular trajectory (in order to satisfy stipulated terminal time conditions)
  4. Complete control law is a synthesis of bang-bang and singular controls.

Analysis of DUE Solutions 1

- The Hamiltonian is linear in the controls h_p:

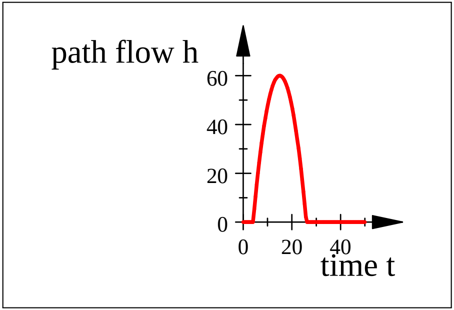
$$\begin{aligned} H &= \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \left[\Psi_p(t,h^*) + \lambda_{ij}\right] h_p \\ &= \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \left[\Psi_p(t,h^*) - v_{ij}\right] h_p \end{aligned}$$

- Thus, <u>all bounded controls are either zero or singular</u>!!
- That is

$$h_p > 0 \Longrightarrow \begin{cases} \Psi_p(t,h^*) - v_{ij} = 0 \\ \text{and} \\ \dfrac{d^n \Psi_p(t,h^*)}{dt^n} = 0 \qquad n = 1, 2, \ldots \end{cases}$$

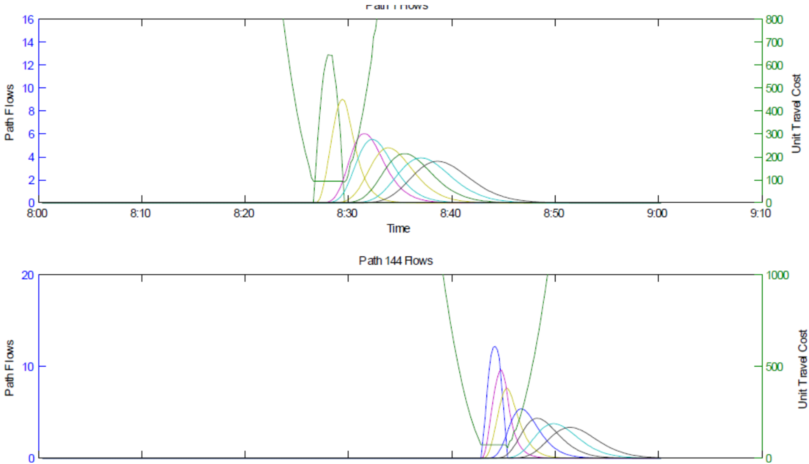
Analysis of DUE Solutions 2

- The departure rates (path flows) are either singular or zero.
- The singular solution is smooth.
- The departure rates (path flows) are piecewise smooth.
- The departure rates and arc exit flows typically have an inverted parabolic shape:

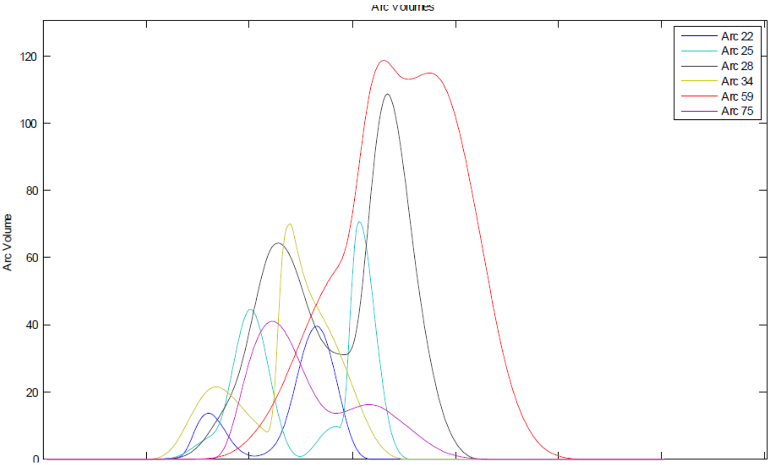


- Implications?

Typical DUE Solution, Arc Exit Flows



Typical DUE Solution, Superposition of Arc Volumes



Relationship of Network Loading and Pure DUE

- Note that the pure DUE problem and the DNL problem are not meant to be solved sequentiall nor have we solved them sequentially.
- Rather, for each instant of time considered during the solution process, the chosen DUE algorithm will need to know the delay for a specific path. That information is obtained by solving the DNL model one selects. That is, delay is treated as an operator that has no closed form.
- Consequently, DNL calculations are the most expensive aspect of dolving DUE models.
- This fact focuses attention on an alternative method of calculating delays $\Longrightarrow$ Kriging, to be discussed shortly.

Algorithms in Either Discrete or Continuous Time

- Fixed-point
- Complementarity woth successive linearization
- projection
- gap function
- proximal point
- duality/subgradient

Fixed Point Formulation and Algorithm

- A (D)VI may be re-stated as a fixed point problem:

$$h = P_{\Lambda_F}\left[h - \alpha\Psi\left(t, h\right)\right]$$

- The associated continuous-time algorithm for arbitrary $\alpha \in \Re^1_{++}$:

$$h^{k+1} = \arg\min_h \left\{ \frac{1}{2} \left\| h^k - \alpha\Psi\left(t, h^k\right) \right\|^2 : h \in \Lambda_F \right\}$$

The Fixed Point Projection Subproblem

- At each iteration k, we must solve a linear-quadratic optimal control problem:

$$\min_{h} J^{k}(h)$$

$$= \sum_{(i,j)\in W} v_{ij}^{k}\left[Q_{ij} - y_{ij}(t_f)\right] + \int_{t_0}^{t_f} \frac{1}{2}\left[h^{k} - \alpha\Psi\left(t, h^{k}\right) - h\right]^{2} dt$$

$$\frac{dy_{ij}}{dt} = \sum_{p\in P} h_p(t) \quad \forall (i,j) \in \mathcal{W}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W}$$

$$h \geq 0 \quad \forall (i,j) \in \mathcal{W}$$

- When $h \geq 0$, our subproblem is solved by descent in Hilbert space.
- Note: derivatives of $\Psi\left(t, h^{k}\right)$ are not needed.
- But how do we find the current dual variables v_{ij}^{k}?

Computing the Dual Variables

- A critical challenge in optimal control is finding dual variables.
- Fortunately this is easy for our present circumstance:

$$\begin{aligned} h_p^{k+1} &= \arg\left\{\frac{\partial H^k}{\partial h_p} = 0\right\} \qquad\qquad \forall\,(i,j)\in W,\;\; p\in P_{ij} \\ &= \arg\left\{\left[h_p^k - \alpha\Psi_p\left(t,h^k\right) - h_p\right](-1) - v_{ij}^k = 0\right\}_+ \\ &\qquad\qquad\qquad \Downarrow \\ h_p^{k+1} &= \left[h_p^k - \alpha\Psi_p\left(t,h^k\right) + v_{ij}^k\right]_+ \end{aligned}$$

- Flow conservation requires:

$$\int_{t_0}^{t_f}\sum_{p\in P_{ij}} h_p^{k+1}(t)\,dt = \int_{t_0}^{t_f}\sum_{p\in P_{ij}}\left[h_p^k - \alpha\Psi_p\left(t,h^k\right) + v_{ij}^k\right]dt = Q_{ij}$$

- The above is easily solved by line search for each $(i,j)\in\mathcal{W}$; the searches may be performed in parallel.

Convergence and the Effective Delay Operator

- The effective delay operator has been studied by many scholars and its "shape" remains a topic of basic research.
- To apply the classical results on convergence, the effective delay operator must be strongly monotone increasing. The same sufficiency condition is associated with most other algorithms
- Examples of non-monotone delay exist.
- Also monotonic delay operators that are consistent with LWR PDE (hydrodynamic traffic flow theory) have been reported by Perakis and Roels (2007), for certain traffic environments.
- What to do?

Traditional Regularity Condition Assuring Fixed-Point Convergence

- Using the notation $\Psi^k \equiv \Psi\left(\cdot, h^k\right)$, the unit path delay operator $\Psi\left(\cdot, h\right)$ is called strongly monotone on Λ_F with constant $K_\Psi > 0$ when
$$\langle \Psi^{k+1} - \Psi^k, h^{k+1} - h^k \rangle \geq K_\Psi \left\| h^{k+1} - h^k \right\|^2$$
- Using notions of pseudo- and quasimonotonicity, this condition has only been slightly relaxed in proofs of convergence for all known DUE algorithms put forward to date.
- Thus DUE algorithms to date have generally been considered heuristics.

Admonishment

- Unfortunately some reviewers do not understand the limitations of presently available mathematics for proving convergence of NLPs , VIs, and DVIs. I have started sending reviewer reports displaying such ignorance back to reviewers for rewriting.
- The proof of convergence for DUE algorithms whose operators have no essential properties is doomed before it is begun. It is "a fool's errand".
- What has been missing is articulation of the behaviors that might assure convergence.
- Let me give an example.

Elastic Demand as a DVI

When Θ_{ij} is inverse travel demand, it is now known that elastic demand DUE may be formulated as follows: find $(h^*, Q^*) \in \Lambda_E$ such that

$$\int_{t_0}^{t_f} \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \Psi_p(t, h^*) \left(h_p - h_p^*\right) dt - \sum_{(i,j)\in\mathcal{W}} \Theta_{ij}(Q^*) \left(Q_{ij}^1 - Q_{ij}^2\right) \geq 0 \quad \forall (h, Q) \in \Lambda_E ,$$

which is recognized as a differential variational inequality (DVI), where

$$\Lambda_E = \left\{ (h, Q) \geq 0 : \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \;\forall (i,j) \in \mathcal{W}; \quad \frac{dQ_{ij}}{dt} = \sum_{p\in P_{ij}} h_p(t),\;\; Q_{ij}(t_0) = 0 \;\forall (i,j) \in \mathcal{W} \right\},$$

See Friesz and Meimand (2014) and Han et al (2015).

Regularity Conditions for the Inverse Demand

- It is quite realistic to assume the inverse travel demand is strongly monotone decreasing:

$$\int_{t_0}^{t_f} \sum_{(i,j)\in\mathcal{W}} \left[\Theta_{ij}\left(Q^1\right) - \Theta_{ij}\left(Q^2\right)\right]\left(Q_{ij}^1 - Q_{ij}^2\right) dt \leq -K_\Theta \left\|Q^1 - Q^2\right\|^2$$

where K_Θ is a positive scalar. An alternative form is

$$(-1)\int_{t_0}^{t_f} \sum_{(i,j)\in\mathcal{W}} \left[\Theta_{ij}\left(Q^1\right) - \Theta_{ij}\left(Q^2\right)\right]\left(Q_{ij}^1 - Q_{ij}^2\right) dt \geq K_\Theta \left\|Q^1 - Q^2\right\|^2$$

Regularity Conditions for Effective Delay Operators

- We also want to introduce the notion of a weakly monotone increasing effective delay operator for some scalar $K_\Psi > 0$:

$$\int_{t_0}^{t_f} \sum_{(i,j)\in\mathcal{W}} \left[\Psi_p(t,h^1) - \Psi_p(t,h^2) \right] \left(h_p^1 - h_p^2 \right) \Big] dt \geq -K_\Psi \cdot \left\| h_p^1 - h_p^2 \right\|^2$$

- Note that weakly monotone increasing operators are quite general functions, and they may be explicitly monotone decreasing for the appropriate constant K_Ψ.

Strong Monotonicity

- From the previous slides, we have

$$\begin{aligned}\langle \Psi^1 - \Psi^2, h^1 - h^2 \rangle_E &\geq -1 \cdot K_\Psi \left\| h^1 - h^2 \right\|^2 \\ \langle (-1) \cdot \Theta^1 - (-1) \cdot \Theta^2, Q^1 - Q^2 \rangle_E &\geq K_\Theta \left\| Q^1 - Q^2 \right\|^2\end{aligned}$$

$$\Longrightarrow (-1) \cdot \langle \Theta^1 - \Theta^2, Q^1 - Q^2 \rangle_E \geq K_\Theta \left\| Q^1 - Q^2 \right\|^2$$

- The effective delay and inverse demand operators obey

$$\begin{aligned}\langle \Psi^1 - \Psi^2, h^1 - h^2 \rangle_E &- \langle \Theta^1 - \Theta^2, Q^1 - Q^2 \rangle_E \\ &\geq -1 \cdot K_\Psi \left\| h^1 - h^2 \right\|^2 + K_\Theta \left\| Q^1 - Q^2 \right\|^2 \qquad (1)\end{aligned}$$

- We now note that the righthand side of (1) may be positive or negative.

Strong Monotonicity, Slide 2

- In fact, it <u>may be that there exists $\bar{K}_\Psi > 0$ and $\bar{K}_\Theta > 0$</u> such that

$$-1 \cdot K_\Psi \left\| h^1 - h^2 \right\|^2 + K_\Theta \left\| Q^1 - Q^2 \right\|^2 \geq \bar{K}_\Psi \left\| h^1 - h^2 \right\|^2 + \bar{K}_\Theta \left\| Q^1 - Q^2 \right\|^2$$

with the immediate consequence that the joint operator

$$\left(\Psi^T, -\Theta^T \right)$$

is strongly monotone increasing, and convergence is provable. Note, however, that there is no a priori reason $\bar{K}_\Psi > 0$ and $\bar{K}_\Theta > 0$ will exist.

DNL by Statistical Learing/Kriging, Slide 1

- We are working on a nonlinear-response-surface/statistical-learning approach to providing closed form expressions for effective path delay operators.
- The statistical approach to Metamodeling we are using is known as Kriging.
- Named after the South African mining engineer Krige (Krige, 1951); revived in the 1960's by geostatisticians (Matheron, 1963) for modeling spatial data.
- Extensively studied in the 1980s.
- More recently, has become an important class of statistical learning and metamodeling methods.
- Kriging tries to approximate a deterministic function by a realization of a Gaussian random process. It has many variants (Fang et al, 2005).

DNL by Statistical Learning/Kriging, Slide 2

- We assume some familiarity with nonlinear response surfaces, statistical learning, kriging, and/or metamodeling.
- School of thought A: Since DNL is the most difficult aspect of DUE modeling, it is natural to look for an alternative way of determining effective delay operators.
- School of thought B: replace the entire DTA modeling exercise with a metamodel.
- Following "B" will not allow decision structure evolution or paradigm shift without rebuilding the metamodel from scratch.
- Following "A" will allow complex tolling and incentivizing, the emergence of new technology, and a spectrum of gaming behaviors.

Toy Numerical Example of SRD Computation: Sioux Falls amd Other Networks

Sioux Falls Network: 76 arcs, 20 nodes, 10 to 300 origin-destination pairs
Fixed travel demand for each OD pairs:

$$Q_{ij} = 100 \quad \forall\,(i,j) \in \mathcal{W}$$

There are 200 paths associated with the 10 origin-destination pair problem.

Numerical Example: Sioux Falls and Other Networks, Comparisons by Major Iterations

Network	LDM	KP-LWR	CTM
3 arcs, 4 nodes	13	31	14
6 arcs, 5 nodes	11	-	12
19 arcs, 13 nodes	14	-	10
Sioux Falls	14	-	13

Table: DUE Fixed-Point Major Iterations by Problem Type

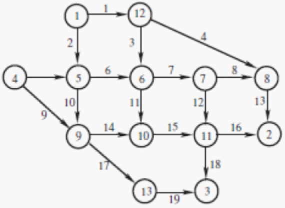
Numerical Example: Sioux Falls Network, Comparisons by CPU Time

Network	**LDM**	**KP-LWR**	**CTM**
3 arcs, 4 nodes	16.1	37.8	21.0
6 arcs, 5 nodes	26.4	-	31.7
19 arcs, 13 nodes	152.2	-	159.1
Sioux Falls	1975	-	3136

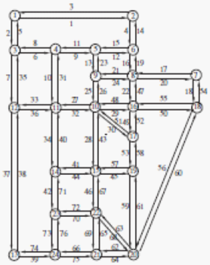
Table: Computation Time (seconds) by Problem Type

$$1975 \text{ sec} \approx 33 \text{ min}$$
$$3136 \text{ sec} \approx 52 \text{ min}$$

Larger Networks, Slide 1

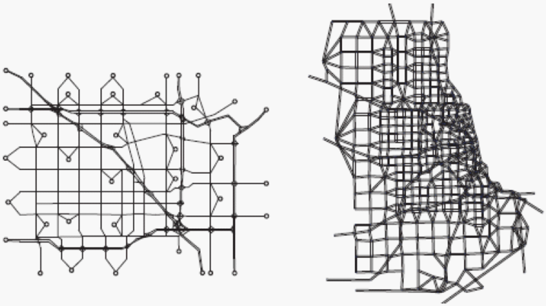


Nguyen network
(13 nodes, 19 links, 4 zones)



Sioux Falls network
(24 nodes, 76 links, 24 zones)

Larger Networks, Slide 2



Anaheim network
(416 nodes, 914 links, 38 zones)

Chicago sketch network
(933 nodes, 2950 links, 387 zones)

Numerical Results, Larger Networks

	Nguyen network	Sioux Falls	Anaheim	Chicago Sketch
No. of iterations	75	43	28	42
Computational time	2.2s	246s	761s	8543s
Avg. time per DNL	0.04s	4.0s	23.3s	111.2s
Avg. time per FP update	0.006s	1.4s	3.1s	59.1s

$$8543 \text{ seconds } = 2.373 \text{ hours}$$

Achieved with a single-processor desktop, without sophisticated programming and based on LWR DNL

Report Card for Dynamic Traffic Assignment, Slide 1

- A fully satisfying behavioral foundation for DTA has not been attained. In particular, DTA should describe departure time, route, and route-updating choices as trajectories, but presently a theory of route updating is missing.
- DTA is moving toward the description of traditional driver-controlled, autonomous, and partially autonomous vehicular flows and volumes. This work on mixed flows is embryonic.
- DTA is a multi-time-scale, multi-spatial-scale and multi-physics problem, but "multiness" has not been adequately investigated.
- Feedback solutions of DTA models have not been adequately investigated.

Report Card for Dynamic Traffic Assignment, Slide 2

- DTA models are sometimes not consistent with basic microeconomics (travel demand theory and utility maximization) and traffic physics.
- DTA modernizes the static 4-step transportation planning paradigm, but has not been applied for that purpose.
- Prescriptive (normative) control/tolling must be constrained by appropriate descriptive (positive, behavioral) DTA models. The result is Stackelberg games (SGs) and mathematical programs with equilibrium constraints (MPECs).
- Differential SGs and MPECs for prescriptive control/tolling have only been partially investigated.
- DTA remains an attractive research topic, but it must adapt to and better reflect information technology innovations in order to survive.

Report Card for Dynamic Traffic Assignment, Slide 3

- Test problems and test data need to be made public.
- DTA is not attached to or recognized by a permanent governing body.
- The language of differential game theory is directly relevant to DTA, but has not been widely adopted.

Report Card for Dynamic Traffic Assignment, Slide 4

- All the relevant physical analogs of traffic have yet to be considered in the DTA context. These are: kinematics, kinetics/entropy, diffusions, and mechanics.
- LWR (kinematic wave) theory said to be the only way to address spillbacks and shock waves. An overstatement?
- However, there are no classical solutions of the LWR equation that are physically real. So numerical solution is filled with subtleties.
- Also, kinematic wave theory does not lend itself to modeling multiple lanes, passing, left turns, and other notions requiring side constraints when analytical DTA modeling is attempted.
- Simulation is said to be the only way to incorporate such considerations. Give up on analytical modeling?
- Statistical Learning has a significant role to play. Can metamodeling overcome such obstacles?

DTA Innovation: Targets and Opportunities

1. Improved Mathematical Representation of DTA – DAEs and DVIs
2. Use the language of Differential Game Theory for discourse
3. New Notions of Differential Traffic Games and Their Solution (going beyond Wardrop)
4. Determine whether Metamodels and Statistical Learning will work for large problems
5. Simulation: How do we look inside the black box and apply the scientific standared of reproducibility?
6. Improved Agorithms, Convergence and Heuristics
7. Weaken the Notion of Monotonicity of Travel Delay
8. Transportation Planning Applications
9. Traffic Network Control and Tolling
10. New Applications – urban supply chains, freight systems, and the congestion impacts of e-commerce
11. Determine a Role for Data Science.

Issues and Questions, Slide 1

1. Existence without a priori bounds. (recently solved) ✓
2. Suitability of different traffic/arc delay models.

1. Value of expressing dynamic network loading (DNL) through the notion of an operator. ✓
2. Simultaneous solution of DNL and pure DUE. ✓
3. Theoretical and computational properties of network loading by (P)DAE systems. ✓
4. What are the properties of path delay operators?
5. Can closed form path delay operators be articulated?
6. Algorithm convergence with nonmonotonic operators. (recent progress) ✓

Issues and Questions from the Literature, Slide 2

8. Establish and share a test problem database for all DTA scholars.
9. Benchmarking of algorithms.
10. Spillbacks (seems to be solved for the DVI approach) ✓
11. Elastic demand. (solved) ✓
12. Uncertain demand.
13. Uncertain delay.
14. Enroute updating for greater behavioral realism.
15. Feedback control/closed loop games for greater behavioral realism.
16. Integrating day-to-day and within-day time scales.

Findings, Slide 1

1. SRD DUE with exogenous delayis easily stated as a VI, NCP, fixed-point problem.
2. Isoperimetric constraints allow restatement of SRD DUE as a DVI.
3. The DVI is especially easy to analyze using optimal control theory/theory of DVIs in continuous time.
4. The network loading subproblem is a (P)DAE system.
5. Network loading is in general expressible as a DAE system.
6. For PQM, the DAE system may be restated as an ODE system.

Findings, Slide 2

7. Fixed-point algorithm works and has distinct advantages:
    1. differentiation of the effective delay operator is avoided
    2. subproblems are LQ optimal control problems
    3. dual variables are easily found.
8. Continuous time computation for SRD DUE is a promising approach, which we did not have time to discuss. Detail to be provided in another conference.
9. Numerical DUE examples demonstrate computability.
10. We are putting together a special issue dedicated soley to large-scale DTA calculations..

Special Topic: Collaborative Games and Enhanced Behavioral Realism

1. Collaborative (cooperative) games consider a set of joint actions that any group of agents can take.
2. Coalitions of agents are formed and their payoffs computed.
3. The solution concept is known as the *core*, for which no smaller coalition of players has an incentive to deviate. Such vectors are said to make up the core.
4. The precise mathematical definition of the core depends on whether we are considering a game with transferable or nontransferable payoffs.
5. In DTA and extended DUE, there are various notions of spatial and temporal coalitions:
    1. coalitions of traffic controls
    2. coalitions of drivers
    3. example: coordinating the traffic control schemes of adjacent MPOs
    4. example: using social networking to form/dissolve car pools

Unanswered Questions/Research Opportunities

There are many opportunities, but from the various comments made previously, I would say the most pressing theoretical issues on which I will be working involve the following:

1.
    1. Models with elastic demand (solved)
    2. Models with uncertain demand
    3. Models with uncertain delay
    4. more general DAE systems for LWR
    5. Enroute updating and feedback games ★
    6. models with dual day-to-day and within-day time scales
    7. Existence without a priori bounds
    8. Algorithm convergence with more behaviorally plausibe operators
    9. Computation with dual time scales

Successful Transfer of this Formalism to Other Applications

1. revenue management, especially computation ✓
2. service pricing ✓
3. electric power pricing ✓
4. urban freight and city logistics ✓
5. dynamic congestion tolls ✓
6. models of the Internet and electronic commerce - ongoing

Thank you!!!!

Questions?