PERGAMON

Transportation Research Part B 34 (2000) 147–156

TRANSPORTATION RESEARCH PART B

www.elsevier.com/locate/trb

Fictitious play for finding system optimal routings in dynamic traffic networks ☆

Alfredo Garcia, Daniel Reaume, Robert L. Smith *

Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109-2117, USA

Received 8 May 1998; received in revised form 9 February 1999; accepted 9 April 1999

Abstract

We introduce a novel procedure to compute system optimal routings in a dynamic traffic network. Fictitious play is utilized within a game of identical interests wherein vehicles are treated as players with the common payoff of average trip time experienced in the network. This decentralized approach via repeated play of the fictitious game is proven to converge to a local system optimal routing. Results from a large-scale computational test on a real network are presented. © 2000 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

The impact of modern telecommunication technologies, real time route guidance and driver information systems on many of the traditional transportation policy and analysis questions has been substantial. For instance, to account for time varying congestion, models based on classical procedures for dynamic transportation systems analysis have increasingly used complex analytical setups (see for example, Friesz et al., 1993; Ran et al., 1996). It is usually the case that to render tractable the analysis, regularity properties like linearity or convexity are assumed.

For the design and operation of these networks, from a classical engineering perspective, one would attempt to find *optimal* policies with respect to an *overall* or *system* efficiency criterion. However, the decentralized structure of ownership and/or decision making, has also called for analysis centered around the notion of *equilibrium* or *user* optimality. In this paper, we explore a

☆ This work was supported in part by the Intelligent Transportation Systems Research Center of Excellence at University of Michigan.

* Corresponding author. Tel.: +313-763-2060; fax: +313-764-3451; e-mail: rlsmith@umich.edu

PII: S0191-2615(99)00018-1

new computational procedure to compute system optimal routings in the sense that, average trip time experienced by vehicles in the network is minimized. The a priori advantages of this procedure are twofold. First, no traditional regularity properties are invoked in the theoretical analysis, which is carried out in a very simple analytical setup. Secondly, the procedure's implementation fully exploits the real time parallel processing potential present in modern dynamic transportation networks. However, the procedure does have a flavour of *equilibrium* models in that it mimics fictitious play within an artificial dynamic traffic game.

Following Rosenthal's (Rosenthal, 1973) setting, in a traffic network game, players are identified with vehicles and payoffs are computed through an assignment mapping that, given routing decisions for all players, calculates the resulting travel times. A routing mapping assigns to each vehicle a time-dependent shortest path from its origin to destination given the route choices of all other vehicles. In game theoretic terminology, the routing mapping is a *best reply* for each vehicle to the routing decisions of the other vehicles.

The theory of learning in games (see Fudenberg and Levine, 1998) attempts to explain equilibrium as the result of a dynamic process of adjustment in which players grope for optimality over time. One example of such a dynamic process is fictitious play.

Fictitious play is an iterative procedure in which at each step, players compute their best replies based on the assumption that opponents' decisions follow a probability distribution in agreement with the historical frequency of their past decisions (Brown, 1951). In the dynamic traffic network context, this procedure can be interpreted as an iterative routing-assignment algorithm, in which at each step, for each player, the routing mapping computes time-dependent expected shortest paths given that other players' decisions are distributed according to the historical frequency of routing decisions. Unfortunately, convergence of such a process is not ensured in general. However, Monderer and Shapley (1996) have demonstrated that when players share a common objective function, fictitious play does converge in a certain sense. From a descriptive viewpoint, this is not a very interesting result, since the kind of strategic interactions game theory attempts to model seldom involve the case of identical interests. Nonetheless, from a normative viewpoint, one could envision an artificial game in which players are forced to assume the common objective (such as in our case, average trip time of vehicles in the network) as theirs. One would expect then, that the set of routings to which this artificial process converges must have some optimality properties.

The major shortcoming of other iterative techniques for solving system or user optimal routings has been the failure to guarantee convergence (see e.g. Kaufman et al., 1998). The fictitious play algorithm we explore in this paper is unique in this regard since convergence is established. The resulting routing, while not necessarily system optimal, is locally optimal in the sense that no player can, by changing its assigned route, diminish any further the average trip time experienced in the network.

2. Preliminaries

2.1. Notation

We introduce the dynamic traffic network game where:
$N = \{1, 2, \ldots, n\}$ is the index set of vehicles.

Every vehicle has a finite number of routing choices, that is to say: for every $i \in N$ there is a finite set R_i of possible routes to take. Let us denote by $R = \prod_{i \in N} R_i$, the set of all vehicles feasible route choices.

$A: R \to \mathscr{R}^n$ is the *assignment* mapping. For any $r \in R$, $A_i(r)$ is the sum over the path determined by r_i of the resulting dynamic link travel times given that the other vehicles take routes r_j, $j \neq i$. In other words, $A_i(r)$ is the total travel time for vehicle i if the vehicles follow routes specified by r. We will denote by Δ^i the set of mixed routing decisions, i.e.:

$$\Delta^i = \left\{ f^i: R_i \mapsto [0,1] \text{ such that } \sum_{r_i \in R_i} f^i(r_i) = 1 \right\},$$

where $f^i(r_i)$ is the probability that vehicle i selects route r_i, $i \in N$. Notice that the extreme points of Δ^i are exactly the elements of R_i.

Let $\Delta = \prod_{i \in N} \Delta^i$ be the set of all mixed routing strategies over all vehicles. We extend the domain of the assignment mapping so that for $f \in \Delta$, we have:

$$A_i(f) = \sum_{r \in R} A_i(r) \cdot f^1(r_1) \cdot f^2(r_2) \cdots f^n(r_n).$$

In words, $A_i(f)$ is the expected value of total travel time for vehicle i when all vehicles adopt mixed routing strategy f.

Let $\Delta^{-i} = \prod_{j \neq i} \Delta^j$ be the Cartesian product of the sets of mixed routing strategies for all vehicles other than i.

2.2. Equilibrium definition

We say that a mixed routing vector f^* is a Nash equilibrium iff for every vehicle $i \in N$, the probabilities assigned to routes for vehicle i by f^*, yield its minimum expected total travel time, provided that f^*_{-i}, the mixed routing choice of all other vehicles, is held fixed, i.e.:

$$f^*_i \in \arg\min_{f^i \in \Delta^i} A_i(f_i, f^*_{-i}).$$

2.3. Example: mixed strategies

To illustrate the need to incorporate *mixed* routing decisions let us consider the simple example where $n = 2$, $R_1 = R_2 = \{a, b\}$ and $A_i(a,a) = A_i(b,b) > A_i(a,b) = A_i(b,a)$, $i = 1, 2$. This could be the case of a situation faced by two vehicles departing at the same time, with same origin and destination, with two route choices available, namely a and b. The outcomes (a,b) or (b,a) (they take different routes) are clearly, Nash equilibrium outcomes.

Consider now the case where vehicles are *uncertain* of other vehicles routing decisions. For instance, let us suppose that vehicle 1 believes vehicle 2 is equally likely to choose routes a or b, or equivalently, that it would follow a mixed strategy assigning equal probabilities to each route. In this case, the best decision for vehicle 1 (i.e. minimizes expected value of his/her travel time) is to also randomize routing choice in the same way vehicle 2 does.

This simple example illustrates the importance of considering mixed strategies. In short, the equilibrium set is enlarged. In many applications, there may only exist equilibria in *mixed* deci-

sions (for a more complete treatment, the reader is referred to Fudenberg and Levine, 1991). Moreover, the quintessential existence theorem in game theory that we invoke below, only ensures existence of mixed equilibria.

3. The fictitious dynamic traffic game

In the dynamic traffic network game described above, fictitious play can be interpreted as an iterative routing-assignment algorithm, in which at each step, for each player, time-dependent shortest paths are computed given that other players' decisions are distributed according to the historical frequency of their earlier routing decisions.

Let us illustrate the workings of the fictitious play process in the example introduced in Section 2.3.

If the first set of routes chosen is $r^0 = (a, b)$ then the first iteration would yield exactly $r^1 = (a, b)$. This instantaneous convergence is due to the fact that (a, b) is in fact, a Nash equilibrium outcome. If $r^0 = (a, a)$ then $r^1 = (b, b)$. That is, b is the best routing decision corresponding to the belief that the other vehicle will repeat a as route choice. At this point, the historical play says that vehicles are equally likely to choose a or b. Hence, the best reply is to randomize route choice with equal probabilities. If the realization is $r^2 = (a, b)$ then, historical frequencies of play suggest that vehicle 1 will pick with probability $2/3$ route a. Hence, vehicle 2 choice will be biased towards route b. Now, let us assume that historical frequencies of play converge. The limit can then be seen as a Nash equilibrium in *mixed* strategies.

Unfortunately, fictitious play processes do not converge in general. Monderer and Shapley (1996) have demonstrated that when players share a common objective function, fictitious play does converge in the sense introduced in the example, i.e., the historical frequencies of routing decisions by players stabilize. This is a weak form of convergence. In words, the law of large numbers ensures that the historical frequencies generated by the process approach the ones generated by a large number of samplings from the limiting mixed routing distribution.

In order to use Monderer and Shapley's result we now define a *different* game; the *Fictitious Dynamic Traffic Game* (which we shall refer to as FDTG), by imposing the average trip time experienced in the network as the common payoff function for all vehicles, i.e. $U: \Delta \to \mathscr{R}$ where:

$$U(f) = \sum_{i \in N} \frac{A_i(f)}{n}.$$

An equilibrium outcome for this game is defined as follows:

Definition 1. Mixed routing strategy f^* is a Nash Equilibrium for the *FDTG* game if for every vehicle $i \in N$, we have:

$$f_i^* \in \arg \min_{f^i \in \Delta^i} U(f_i, f_{-i}^*).$$

Let us now examine the meaning of a Nash Equilibrium for this game. Intuitively, given that all other vehicles $j \neq i$ follow f_j^*, vehicle i cannot reduce any further the average trip time experienced by the vehicles in the network by deviating from the prescribed routing f_i^*. In other words, for the optimization problem:

$$\text{(P)} \quad \min_{f \in \Delta} U(f)$$

the mixed routing f^* is a type of a *local* optimal solution. Every optimal solution to (P) must be an equilibrium routing for the FDTG game, whereas there may be equilibrium routings for the FDTG game that are not optimal solutions to problem (P). In other words;

Optimal Solution Set(P) $\subset$ Equilibrium Set(FDTG).

3.1. Example

Let us consider the case when vehicles have only two choices. In this case, $\Delta = [0,1]^n$. If f^* is a *mixed* equilibrium routing for the FDTG game and $U(\cdot)$ is assumed differentiable then a first order necessary condition is:

$$\frac{\partial U}{\partial f_i}(f^*) = 0. \tag{3.1}$$

If in addition, $U(\cdot)$ is assumed convex then:

$$U(f_1^* + \lambda_1, \ldots, f_i^* + \lambda_i, \ldots, f_n^* + \lambda_n) \geqslant U(f^*) + \sum_i \frac{\partial U}{\partial f_i}(f^*)\lambda_i \tag{3.2}$$

for $\lambda_i \in [0,1)$. Thus, (1) and (2) imply that:

$$U(f_1^* + \lambda_1, \ldots, f_i^* + \lambda_i, \ldots, f_n^* + \lambda_n) \geqslant U(f^*).$$

That is, *mixed* equilibrium solutions of the FDTG game are also optimal solutions of problem (P) when $U(\cdot)$ is convex and differentiable.

3.2. Existence of Nash equilibrium in mixed strategies for the FDTG game

The Nash equilibrium existence theorem (1950) states that there exists an equilibrium solution in mixed strategies for finite games, that is games with finite decision sets. Such a result when applied to the FDTG game translates into:

Theorem 1. *The FDTG game has a Nash equilibrium in mixed routing decisions.*

Proof. See Nash (1950). □

4. Fictitious play convergence for FDTG game

We now briefly review Monderer and Sharpley's (Monderer and Shapley, 1996) results which constitute the theoretical basis for the computational procedure we will introduce.

4.1. Notation

Let us denote by $K \subset \Delta$ the equilibrium set for the FDTG traffic game above presented and $\|\cdot\|$ any fixed euclidean norm on Δ.

For $\delta > 0$ let $\mathbf{B}_\delta(K)$ the open ball with radii δ, i.e.:

$$\mathbf{B}_\delta(K) = \left\{ g \in \Delta : \min_{f \in K} \|g - f\| < \delta \right\}.$$

A pure routing sequence is a sequence of the form $\mathbf{r} = \{r_t\}_{t=1}^{\infty}$ of elements of R. A *historical routing sequence* is a sequence of mixed routing strategies of the form $\mathbf{f} = \{f_t\}_{t=1}^{\infty}$ in Δ.

Definition 2. We say that a historical routing sequence converges to equilibrium if each limit point is an equilibrium point. Formally, for every $\delta > 0$ there exists T such that $f_t \in \mathbf{B}_\delta(K)$ for all $t \geqslant T$.

This notion of convergence is rather weak in that it is only implied by Definition 2 that limits of converging subsequences are equilibrium points (the historical routing sequence itself may not converge).

To every pure routing sequence $\mathbf{r}$ we can associate a historical routing sequence $f_{\mathbf{r}}$ by simply computing the historical frequency of the various pure routing decisions in $\mathbf{r}$, i.e. for given $r \in R_i$:

$$f^i_{\mathbf{r},t}(r) = \frac{\#\{1 \leqslant s \leqslant t : r_i(s) = r\}}{t}. \tag{4.1}$$

In Eq. (4.1), $f^i_{\mathbf{r},t}(r)$ is simply the ratio of the number of times vehicle i has chosen route r, over total number of iterations t. Note that if we define $I^i_{\mathbf{r},t}(r)$ to be the indicator function of the route r in the pure routing sequence $\mathbf{r}$, a simple analytical expression for updating historical frequencies is:

$$f^i_{\mathbf{r},t+1}(r) = f^i_{\mathbf{r},t}(r) + \frac{(I^i_{\mathbf{r},t}(r) - f^i_{\mathbf{r},t}(r))}{t+1}, \tag{4.2}$$

where

$$I^i_{\mathbf{r},t}(r) = \begin{cases} 1 & \text{if } r_i(t) = r, \\ 0 & \text{otherwise.} \end{cases}$$

A pure routing sequence $\mathbf{r}$ is the result of a fictitious play process if for every $i \in N$ and every t:

$$r_i(t+1) \in \arg \min_{r_i \in R_i} [U(r_i, f^{-i}_{\mathbf{r},t})].$$

In words, at each t the route prescribed for player i, is the pure best response to the mixed strategy for all other vehicles consisting of the historical frequency of routes they have chosen up to time period t.

We now state the important Monderer and Shapley's result applied to the FDTG game.

Theorem 2. *Every pure routing sequence* $\mathbf{r}$ *which is the result of a fictitious play process converges to equilibrium.*

Proof. See Monderer and Shapley (1996). □

5. Algorithm

We formally present the algorithm motivated by Monderer and Shapley's result when applied in the context of the FDTG game: It presents, however, a major difficulty; Theorem 2 only asserts

that for a *converging sequence of mixed strategies* generated by fictitious play the limit is a Nash equilibrium of the original game. It is in a sense, a "limsup" set convergence result, and for computational purposes we need a "liminf" type of result. However, it is worth pointing out that whenever the equilibrium set of our artificial game is a singleton, the algorithm is guaranteed to converge in the stronger sense. In any other case, the algorithm will compute routings that will be arbitrarily close to the equilibrium set. A continuity assumption of the *assignment* mapping will ensure a good approximation of the optimal value of average trip time in the network.

Algorithm

1. Set $t = 0$. Pick an initial pure routing strategy f_0.
2. Compute a best reply for each $i \in N$:

$$r_i(t+1) \in \arg \min_{r_i \in R_i} [U(r_i, f_t^{-i})].$$

3. Update historical frequencies of route choices, f_t according to Eqs. (4.1) and (4.2).
4. If $\|f_{t+1} - f_t\| \leqslant \varepsilon$ then Stop, otherwise, set $t = t + 1$ and go to 2.

5.1. Implementation

We have implemented the above algorithm in a software package called *Alliance*. To implement the assignment mapping, we use a version of the *INTEGRATION* traffic simulator for link-time prediction (Van Aerde et al., 1989), called *INTEGRATION-UM*, which was developed by researchers at the University of Michigan. It basically follows a deterministic mesoscopic approach employing macroscopic travel time and flow relationships and macroscopic individual vehicle control and link queueing.

To ensure tractability, we made several simplifications to the implementation of the dynamic router or best reply subroutine. First, it is extremely difficult to, for each vehicle i analytically compute a best response to the historical frequencies of the routings of the other vehicles. Instead we simulate the passage of these other vehicles through the network where each vehicle chooses its route with probability in accordance with the historical frequencies of its routings. Using the time-dependent link travel time profile produced by this simulation, vehicle i may then be assigned to a path minimizing the increase to total system travel time. We use a variant of Dijkstra's algorithm introduced in Kaufman and Smith (1993). It is worth emphasizing here that the *Alliance* basic iteration (simulation + best replies), although not done in its present form, can be processed in parallel.

To further simplify matters, since a single vehicle will have little effect on congestion, we also simulate the travel vehicle i, in this simulation, thus avoiding the need to run a separate simulation for each vehicle. It is hoped that for a heavily congested network with many vehicles leaving at the same time with the same origin and destination this approach will adequately approximate true best responses.

The second simplification we perform is to discretize time into a sequence of slices. Within each slice, vehicles are routed by the simulation according to the routing tables.These tables assign, for each slice s and node n, the probability distribution with which a vehicle arriving at node n with destination d at time s, will choose its next link. Note that this simplification allows vehicles to

follow routes they may have historically never taken, contradicting the algorithm's scheme. Here again, we appeal to the large number of vehicles flowing through the network to justify this simplification since the effective congestion should remain relatively unchanged while greatly reducing storage and computation requirements.

5.2. Computational tests

To validate *Alliance* we applied the algorithm to the Troy, Michigan traffic network (Wunderlich et al., 1997). Approximately 16,500 vehicles were allowed to flow into the network in 24 min according to travel patterns approximating those actually observed in Troy. After 24 min the flow into the network was halted and the vehicles were allowed to travel for a further 36 min, thus allowing the network to clear.

We will compare the performance of *Alliance* performance to *SAVaNT* (see Kaufman et al., 1998), which is an iterative routing-assignment procedure intended to compute *user* optimal routings that has been perceived to provide (whenever it converges) reasonably good routings in terms of average trip time in the network.

To account for the impact of different market penetration levels of ITS technologies we define three classes of vehicles. *Class 1*, consisted of those vehicles following the free flow shortest path. *Class 2*, consisted of those vehicles that perform a periodic update of the free flow shortest paths, and finally, *Class 3* vehicles were guided by the *Alliance* algorithm. The initial routing given to all classes corresponded to shortest paths under free flow conditions.

In the first test with high market penetration (i.e., Class 3 vehicles account for 25% of the total number of vehicles) we observe that *Alliance* computes routings roughly as good as those computed with *SAVaNT*, in terms of system average trip time, using considerably fewer iterations and less c.p.u time.

Test 1: Average trip time per class (min)				
	C1 (50%)	C2 (25%)	C3 (25%)	# Iterations
Alliance	8.85988	8.85677	8.71779	14
SAVaNT	8.87245	8.84866	8.68266	34

In the second test, we assume a low market penetration (i.e. Class 3 vehicles account for 5% of the total number of vehicles) we observe reductions in travel time for intelligent vehicles, here again at a substantially lower computational effort when compared to *SAVaNT* (see also, Fig. 1).

Test 2: Average trip time per class (min)				
	C1 (95%)	C3 (5%)	Average	# Iterations
Alliance	17.30430	15.59930	17.21905	20
SAVaNT	17.49250	15.49160	17.39240	68

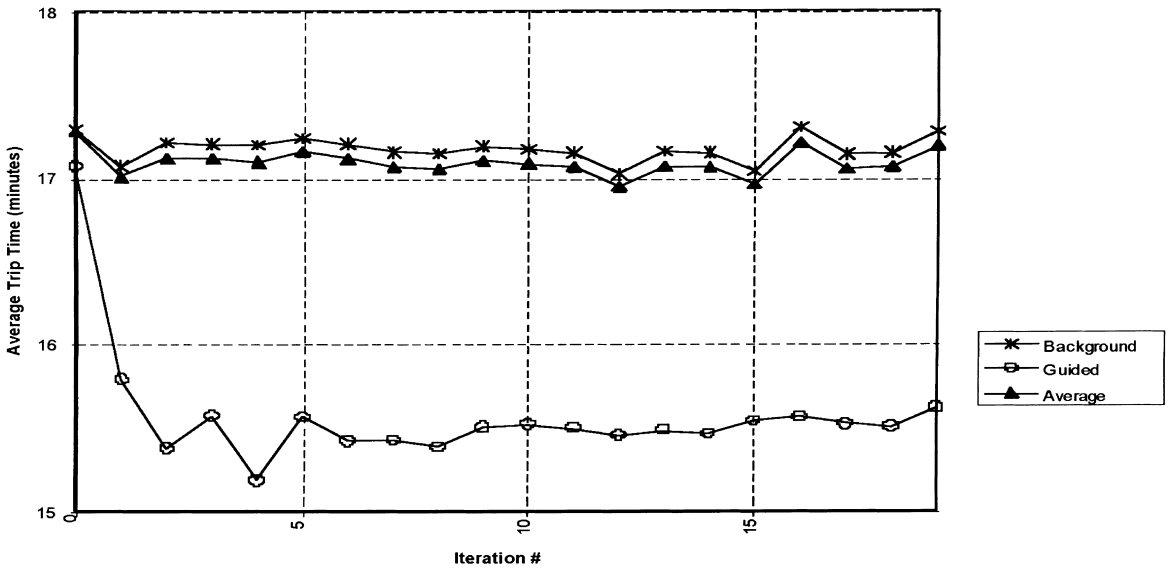

Fig. 1. Illustration of speed of convergence.

6. Conclusions

Through the application of recent results in the theory of learning in games and the extension of Rosenthal (1973) framework to the formulation of a dynamic traffic game, we have implemented a decentralized iterative procedure to compute system optimal routings.

By focusing on discrete routing decisions and with the help of a dynamic travel time simulator we have avoided the technicalities of a more thorough analytical development.

First large-scale empirical results are encouraging, yielding substantial reductions in computational requirements.

References

Brown, G.W., 1951. Iterative solution of games by fictitious play. Activity Analysis of Production and Allocation. Wiley, New York.

Friesz, T., Bernstein, D., Smith, M., Tobin, R., Wie, B., 1993. A variational inequality formulation of the dynamic network user equilibrium problem. Operations Research 41, 179–191.

Fudenberg, D., Levine, D., 1991. Game Theory. MIT Press.

Fudenberg, D., Levine, D., 1998. The Theory of Learning in Games. MIT Press, Cambridge, MA.

Kaufman, D., Smith, R.L., 1993. Fastest paths in time-dependent networks for intelligent vehicle-highway systems application. IVHS Journal 1 (1), 1–11.

Kaufman, D., Smith, R.L., Wunderlich, K., 1998. Dynamic user-equilibrium properties of fixed-points in iterative routing-assignment methods. Transportation Research, Part C, vol. 6 (1).

Monderer, D., Shapley, L., 1996. Fictitious play property for games with identical interests. Journal of Economic Theory 68, 258–265.

Nash, J., 1950. Equilibrium points in *n*-person games. Proceedings of the National Academy of Sciences 36, 48–49.

Ran, B., Hall, R., Boyce, D., 1996. A link-based variational inequality model for dynamic departure time/route choice. Transportation Research, Part B, vol. 30.

Rosenthal, R.W., 1973. The network equilibrium problem in integers. Networks 3, 53–59.

Van Aerde, M., Voss, J., McKinnon, G., Integration Simulation Model User's Guide, Version 1.1. Department of Civil Engineering, Queen's University, Kingston, Ontario, Canada.

Wunderlich, K., Kautman, D., Smith, R.L., 1997. Link travel time prediction techniques for convergent iterative anticipatory Route guidance methods. Technical report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, 1997.