2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Game-Theoretic Routing of GPS-Assisted Vehicles for Energy Efficiency

Anil Aswani and Claire Tomlin

Abstract— **Congestion on roads and highways is an issue that leads to reductions in the energy-efficiency of travel. Current GPS navigation devices include features which provide turn-by-turn directions to vehicles based on real-time traffic conditions, and these features provide an opportunity to improve average fuel consumption. Routing strategies in these devices optimize individual travel times, but theoretical (e.g., Braess's paradox) and empirical results show that this can actually increase congestion and average travel times. We model traffic routing in the game-theoretic framework of Stackelberg games, which is a simplification of the true information patterns, and then use this model to provide an algorithm for turn-by-turn directions. One advantage of our algorithm is that it can be easily incorporated into existing GPS devices by modifying the traffic information sent to them. Our framework is used to qualitatively study the effectiveness of traffic routing on a specific road network topology. If roughly 60% of users follow GPS directions implementing our strategy, then the average delay will be close to the optimal average delay for the road network. This poses social and technological challenges for reduction in congestion through routing. The situation is not hopeless though, because our qualitative results indicate that having a small percentage of compliant users may still lead to large reductions in congestion.**

I. Introduction

Vehicle traffic congestion is a growing problem with many associated costs. In addition to economic costs from lengthy travel times, there are external, environmental costs from increased fuel consumption and vehicle emissions. Reducing congestion can lead to improvements, and the associated design of intelligent transportation systems (ITS's) is an important area of research. ITS's combine computation, communication, and infrastructure to improve safety and congestion [1], [2]. An example of design in an ITS is the scheduling of traffic lights for highway on-ramps [3], [4].

ITS's typically make significant use of fixed infrastructure such as electronic road signs, but the growing computational power of cellular phones and GPS navigation devices will enable more efficient systems [5], [6], [7]. Magnetic detectors in roads have been used to measure traffic levels on highways [8], and a recent demonstration has shown that cellular phones equipped with GPS can also be used for this purpose [6]. An advantage of systems based on cellular phones is a reduction in fixed costs due to construction of infrastructure.

GPS navigation devices have begun to provide traffic congestion reports [6], in addition to planning routes for vehicles. These devices indicate the location of accidents and the level of congestion on highways, and they assist with planning routes which take traffic levels into account in order to minimize expected travel time. The impact of such devices has not been explicitly considered in the literature, and one aim of this paper is to do so. The other aim of this paper is to provide a basis for improving their operation.

This material is based upon work supported by the National Science Foundation under Grant No. 0931843.

A. Aswani and C. Tomlin are with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, CA, USA{aaswani,tomlin}@eecs.berkeley.edu

A. Traffic Assignment

Traffic assignment is the selection of routes for different pairs of origin-destination, and it has historically been used to study the potential impact of adding new roads. The techniques can be broadly classified into being either static [9], [10], [11], which consider the assignment of routes under equilibrium conditions, or dynamic [12], [13], [14], [15], [16], [17], which consider the evolving nature of traffic patterns. Fundamentally, GPS navigation devices are an implementation of traffic assignment techniques.

The seminal work in [18] defined two notions of route assignment that are now prevalent. A traffic flow in which each vehicle minimizes individual travel time leads to a user equilibrium (UE) flow, which is also called a Wardrop or Nash equilibrium [19] and satisfies Wardrop's first principle of route choice. Flows in which vehicles cooperate to minimize average travel times lead to system optimal (SO) flows, and they satisfy Wardrop's second principle of route choice.

GPS navigation devices with traffic information encourage UE flows, but this is problematic because it can lead to perverse behavior. Traffic assignment which minimizes individual vehicle travel times can lead to increased congestion. This is most dramatically realized in Braess's paradox [19], [20] where the addition of roads can increase congestion, and it has been observed in real world situations [19]. On the other hand, ensuring SO flows corresponds to reducing the societal costs of traffic congestion by reducing the travel times for an average vehicle.

B. Overview

The purpose of this paper is to study how to best provide turn-by-turn directions on GPS navigation devices in order to promote reductions in congestion and energy consumption. We make proposals on the nature of what delay information to provide and what types of routes should be presented to GPS users. A benefit of our proposal is that existing GPS devices would not need to be updated, just the information sent to them would change. Additionally, we discuss what fraction of GPS users are needed in order to get appreciable improvements in congestion.

Studying these issues involves understanding the interplay between cooperative and non-cooperative behavior of users,

978-1-4577-0079-8/11/$26.00 ©2011 AACC

and this analysis is most easily framed in the language of game theory. We begin by describing a quasi-static model for vehicle traffic flow on uninterrupted highways, and we provide models for user behavior. A quasi-static model assumes that traffic levels are slowly varying [21], [19]. Results from game theory provide the tools to efficiently compute turn-by-turn directions which minimize congestion, and further results are used to qualitatively study the effectiveness of traffic routing. To make the analysis tractable, we make simplifications on the models of traffic flow and the information pattern provided to different drivers, and we consider a simple road network topology.

II. Quasi-Static Traffic Assignment Model

A road network is represented by a directed graph $G = (N, L)$ where $N = \{1, \ldots, n\}$ is a set of nodes representing junctions between roads and

$$L = \{(i, j, m) : \text{the } m\text{-th road going between node } i \text{ to node } j\} \quad (1)$$

is a set of links representing roads between junctions. Roads going in opposite directions are distinct links, meaning that $(i, j, \cdot) \neq (j, i, \cdot)$, and having a road $(i, j, \cdot)$ does not guarantee having a road in the opposite direction $(j, i, \cdot)$. We do not consider roads that loop back to the original node $(i, i, \cdot)$.

In our model, there are w origin-destination pairs $\{o_1, d_1\}, \ldots, \{o_w, d_w\}$. The k-th simple route for the i-th pair $\{o_i, d_i\}$ is a path which starts at node o_i and ends at node d_i such that there are no loops, and it is given by

$$\mathcal{P}_{i,k} = (o_i, j_1, l_0), (j_1, j_2, l_1), \ldots, (j_m, d_i, l_m), \quad (2)$$

where m is the number of links in the route and $l_0, \ldots, l_m$ specifies which road. The set of all routes for the pair $\{o_i, d_i\}$ is denoted by $\mathcal{P}_i = \bigcup_k \mathcal{P}_{i,k}$, and the set of paths for all pairs is given by $\mathcal{P} = \bigcup \mathcal{P}_i$.

We associate a finite and positive rate r_i for each pair $\{o_i, d_i\}$, and this rate denotes the number of vehicles entering and exiting $\{o_i, d_i\}$, when the network is at equilibrium. A flow $f_{\mathcal{P}_{i,k}}$ gives the number of vehicles following path $\mathcal{P}_{i,k}$, and a flow is feasible if $\sum_k f_{\mathcal{P}_{i,k}} = r_i$. The flow

$$f_{(i,j,m)} = \sum_{P_{i,k} : (i,j,m) \in P_{i,k}} f_{\mathcal{P}_{i,k}} \quad (3)$$

gives the number of vehicles on the m-th road going between nodes i and j. For each link, we define a delay function $D_{(i,j,m)} : f_{(i,j,m)} \to \mathbb{R}_+$ which gives the time required to travel over the road as a function of the number of vehicles on the road.

A. Delay Functions

In the game-theoretic framework we use, standard results require that the delay function $D_{(i,j,m)}$ be non-negative, differentiable, non-decreasing, and independent of the number of vehicles in other links [21], [22]. For reasons of tractability and analyzability, several simplifying assumptions are made about the behavior of traffic networks [4]. The delay in link (i, j, m) is interpreted as the time required for one vehicle to travel from the start to the end of the corresponding road, and the delay in the link is assumed to be only a function of the number of vehicles in the link (i, j, m).

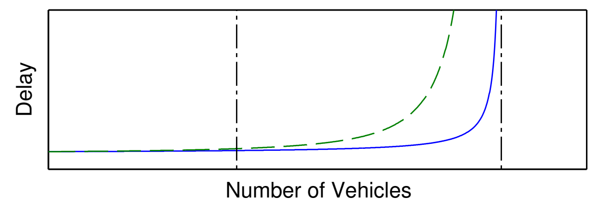


Fig. 1: The delay (solid line) $D_{(i,j,m)}(f_{(i,j,m)})$ is constant until the number of cars hits the critical density of the road, and then it increases up to an asymptote at the capacity of the road. The marginal delay time (dashed line) $\tilde{D}_{(i,j,m)}(f_{(i,j,m)})$ is monotonically increasing, and it has similar behavior.

One model for vehicle delay is shown in Fig. 1, and this model can be mathematically represented by

$$D_{(i,j,m)}(f) = K_1 + K_2/(C - f), \quad (4)$$

where K_1, K_2 are non-negative constants and $C > 0$. These constants can vary for different links (i, j, m). This model is representative of the fundamental flow diagram in macroscopic-scale, fluid-flow models of traffic networks, and the delay function implicitly enforces constraints on the maximum number of vehicles on a link. The number of vehicles on the road cannot exceed the capacity, otherwise the delay would be infinite.

B. User Equilibrium Flow

In a user equilibrium (UE) flow, each vehicle tries to minimize its individual delay [18], [19]. Mathematically, it corresponds to a feasible flow (it is a constraint) which minimizes the following objective:

$$D_{UE} = \min_{f_{(i,j,m)}} \sum_{(i,j,m) \in L} \int_0^{f_{(i,j,m)}} D_{(i,j,m)}(f) df \quad (5)$$

subject to the constraints:

$$\begin{aligned}
&f_{(i,j,m)} \geq 0 \\
&r_i + \sum_{j,m} f_{(j,o_i,m)} = \sum_{j,m} f_{(o_i,j,m)}, \qquad \forall \{o_i, d_i\} \\
&\sum_{j,m} f_{(j,d_i,m)} = r_i + \sum_{j,m} f_{(d_i,j,m)}, \qquad \forall \{o_i, d_i\} \\
&\sum_{j,m} f_{(j,k,m)} = \sum_{j,m} f_{(k,j,m)}, \qquad \forall k \neq o_i, d_i, \quad \forall i.
\end{aligned} \quad (6)$$

The constraints (6) are convex, and they ensure non-negativity of flows and conservation of flows (i.e., flow into a node equals the flow out of the node). This optimization problem can be reformulated as a variational inequality [22], and this inequality provides a more intuitive description of the UE flow. The UE flow is the flow such that each vehicle going between $\{o_i, d_i\}$ travels along a path $\mathcal{P}_{i,k}$ with minimum delay, otherwise the vehicle would reroute to follow a path with smaller delay. Consequently, all vehicles going between $\{o_i, d_i\}$ experience equal delay.

The delay function (Fig. 1 and (4)) and constraints (6) are convex. Furthermore, the number of constraints is polynomial in terms of the number of edges [22]. As a result, this problem can be efficiently solved in reasonable time using either centralized or decentralized algorithms. If desired, we can compute $f_{\mathcal{P}_{i,k}}$ by solving the linear system of equations given in (3).

C. System Optimal Flow

In a system optimal (SO) flow, each vehicle tries to minimize the average vehicle delay [18], [19], [22]. The SO flow corresponds to minimizing the following objective:

$$D_{SO} = \min_{f_{(i,j,m)}} \sum_{(i,j,m)\in L} f_{(i,j,m)} D_{(i,j,m)}(f_{(i,j,m)}) \quad (7)$$

subject to the constraints (6). Because we are interested in minimizing average vehicle delay, the expression in (7) qualitatively looks like the expectation of some random variables. As discussed in the introduction, the SO flow corresponds to reductions in congestion (and indirectly to reductions in fuel consumption and vehicle emissions).

There is an important connection between SO and UE flows. A SO flow corresponds to a UE flow with the modified delay function

$$\tilde{D}_{(i,j,m)}(f_{(i,j,m)}) = D_{(i,j,m)}(f_{(i,j,m)}) + f_{(i,j,m)} D'_{(i,j,m)}(f_{(i,j,m)}), \quad (8)$$

where the prime indicates differentiation, and an example of this delay is shown in Fig. 1. This relationship can be shown by direct integration of delay (8) in the cost of (5). Note that $\tilde{D}_{(i,j,m)}(f_{(i,j,m)})$ is typically called the marginal delay function [22], and it corresponds to having a toll of $f_{(i,j,m)} D'_{(i,j,m)}(f_{(i,j,m)})$ in addition to the delay $D_{(i,j,m)}(f_{(i,j,m)})$. This connection is important because in an SO flow, the vehicles selfishly choose paths to minimize the marginal delay function. If all vehicles were to pick the path with minimal marginal delay, then the network would converge to an SO flow [22].

III. Stackelberg Flow

The goal of the traffic engineer is to design an ITS which solves (7), and this is challenging because it is at odds with the preferences of an individual vehicle which solves (5). A GPS navigation device can aid with this by providing vehicles with directions which correspond to the SO flow. Unfortunately, not all vehicles are GPS users, and not all GPS users will follow the directions.

This brings up the important question of what nature of delay information should be provided to GPS users. The number of vehicles on a link can be measured in a number of ways [8], [6], and these measurements can be used to display to users either delay $D_{(i,j,m)}(f_{(i,j,m)})$ or marginal delay $\tilde{D}_{(i,j,m)}(f_{(i,j,m)})$. Since vehicles tend to selfishly minimize their own delay, it is best to present GPS users with marginal delay information, because users individually minimizing marginal delay encourages the formation of an SO flow that minimizes average delay. The other benefit of displaying marginal delay to users is that it does not require giving false information in order to make vehicles behave favorably.

If GPS users are given marginal delay, then we can model vehicle behavior in a simple manner that is a considerable simplification of the true information pattern. Future approaches should consider more accurate notions such as Bayesian games and learning in policies. Vehicles can be classified into two groups: a) those that follow the GPS directions, and b) those that try to minimize individual delay $D_{(i,j,m)}(f_{(i,j,m)})$. The second group consists of both users without GPS devices and users with GPS devices who ignore the provided directions. Note that some users might try to use knowledge that the GPS device provides marginal delay to compute individual delay. To match the terminology to the game theory literature, users in group (a) will be called Stackelberg leaders, and users in group (b) will be called Stackelberg followers.

This can be modeled as a Stackelberg game in which the leaders determine their routes and the followers pick their routes based on what the leaders have picked [23], [11], [24]. A defining characteristic of a Stackelberg game is that the leader tries to pick a strategy in anticipation of the follower, such that after both players make moves, the objective of the leader is minimized. The traffic routing problem fits this framework well. The leader picks traffic routes, such that after the followers selfishly pick their routes to minimize individual delay, the average delay is minimized. The optimal way of doing this would be a leader which can anticipate how the followers will respond, and pick traffic routes in anticipation.

Unfortunately, it has been shown that picking the optimal Stackelberg strategy for an arbitrary network is NP-hard [25], and so several heuristic algorithms have been developed. Two common strategies are SCALE (scale the optimal flow) and LLF (largest latency first) [25], [23], [24]. These strategies can be improved by playing a finite number of rounds where the leader and follower successively select routes based on those of the previous player [26]. For simplicity, we focus on the SCALE strategy because it is computationally simpler and easier to generate qualitative results.

Suppose that $\alpha \in [0,1]$ is the fraction of vehicles which are Stackelberg leaders. Then, the SCALE strategy is as follows. Define $\tilde{f}_{(i,j,m)}$ to be SO flow, which minimizes (7) by definition. The Stackelberg leaders have flow $\alpha\tilde{f}_{(i,j,m)}$, meaning that their flow is the SO flow that has been scaled by the fraction of vehicles that are leaders. The Stackelberg followers have flow $f^*_{(i,j,m)}$ which is the UE flow for the delay functions

$$D^*_{(i,j,m)}(f^*_{(i,j,m)}) = D_{(i,j,m)}(f^*_{(i,j,m)} + \alpha\tilde{f}_{(i,j,m)}). \quad (9)$$

Algorithm 1: Turn-by-Turn Directions for Vehicle with GPS Navigation Device

input : Road Network – (N, L);
Marginal Delay – $\tilde{D}_{(i,j,m)}$;
Current Traffic – $\hat{f}_{(i,j,m)}$;
Origin-Destination – $\{o_i, d_i\}$;
output: Route – $\mathcal{P}_{i,k}$

foreach $(i, j, m) \in L$ **do**
  set $W(i, j, m) := \tilde{D}_{(i,j,m)}(\hat{f}_{(i,j,m)})$;
end
`Dijkstra's algorithm returns path of minimum length for an origin-destination pair on directed graph with weighted edges;`
set $\mathcal{P}_{i,k} := Dijkstra(N, L, W)$;
return $\mathcal{P}_{i,k}$

A. Algorithm for Turn-by-Turn Directions

Generating turn-by-turn directions for GPS users is simple when using the SCALE strategy, and it can be done in a decentralized manner as long as current traffic information is centrally generated and transmitted to everyone. Essentially, every GPS navigation device picks a route which minimizes the marginal delay between the desired origin-destination pair $\{o_i, d_i\}$ for the vehicle. This requires that the GPS device receive traffic information updates which give the number of vehicles on each link. It is formally described in Algorithm 1, which is simply Dijkstra's algorithm that returns a path of minimum length for an origin-destination pair on a directed graph with weighted edges.

This algorithm is decentralized in the sense that it routes individual users, though traffic information is aggregated centrally, and it leads to a Stackelberg flow in which the leaders approximately follow the SCALE strategy [27]. In fact, if α of these vehicles follow the directions generated by this algorithm, then the flow in network will equilibrate such that (7) will be no worse than it would have been with the SCALE strategy for the corresponding value of α [26]. This algorithm is simply an application of Dijkstra's algorithm which is a polynomial time algorithm. It can be implemented on a GPS navigation device. Furthermore, this approach allows easy incorporation of existing GPS devices by simply modifying the traffic information sent to the device: It does not require updating the GPS devices.

IV. Importance of User Compliance

Understanding the impact of turn-by-turn directions generated by traffic routing algorithms is important for designing an ITS. There are issues of cost-effectiveness and societal impact that need to be qualitatively studied, and the game-theoretic framework introduced above is very useful in being able to do this. The key question is how much reduction in average delays is possible as the fraction of Stackelberg leaders α varies from zero to one. By definition, if $\alpha = 0$, then there is no reduction in congestion from GPS users, because there are no GPS users. Similarly, if $\alpha = 1$, then congestion is at the minimum possible level, because all vehicles are GPS users.

There are two measures of improvement in congestion that can be used. The first measure is known as the *price of anarchy* [22], [23], [24], and it gives the worst possible ratio between average delay with Stackelberg routing and average delay with an SO flow:

$$\rho(\alpha) \triangleq \sup_{(G,N),D_{(i,j,k)}} \frac{\sum_{(i,j,m)\in L} S\Big(\alpha \tilde{f}_{(i,j,m)} + f^*_{(i,j,m)}\Big)}{\sum_{(i,j,m)\in L} S\Big(\tilde{f}_{(i,j,m)}\Big)}, \tag{10}$$

where $S_{(i,j,m)}(f_{(i,j,m)}) = f_{(i,j,m)} D_{(i,j,m)}(f_{(i,j,m)})$. The second measure is known as the *value of altruism*; though this is a new quantification that we have defined, it is motivated by ideas (and the same phrase) from [28]. The value of altruism gives the best possible ratio of improvement in average delay between Stackelberg routing and a UE flow, and it is defined as

$$\sigma(\alpha) \triangleq \inf_{(G,N),D_{(i,j,k)}} \frac{\sum_{(i,j,m)\in L} S\Big(\alpha \tilde{f}_{(i,j,m)} + f^*_{(i,j,m)}\Big)}{\sum_{(i,j,m)\in L} S\Big(\overline{f}_{(i,j,m)}\Big)}, \tag{11}$$

where $S(\cdot)$ is as defined above and $\overline{f}_{(i,j,m)}$ is the UE flow that minimizes (5).

There is an important point to note concerning the computation of these values for the delay function (4). The delay function is unbounded because $D_{(i,j,m)}(C)$ is infinite. To make these computations sensible, we have to impose an additional constraint on the maximum flow through any link [22]. Specifically, we assume that $f_{(i,j,m)} \leq f_{\max}$.

A. Price of Anarchy

The price of anarchy can be easily computed for $\alpha = 0$ and $\alpha = 1$. When $\alpha = 0$, we use a general method for computing the price of anarchy [22] which can be used with the delay function in (4). Some thought shows that this computation is equivalent to the price of anarchy for an M/M/1 queue, and so a result in [22] gives that $\rho(0) = (1+\sqrt{(\frac{C_{\min}}{C_{\min}-f_{\max}})})/2$, where $C_{\min}$ is the minimum value of C across the delay functions for each link (4). We trivially have $\rho(1) = 1$, because here the SCALE flow is equivalent to the optimal flow.

Computing the price of anarchy for $\alpha \in (0, 1)$ for the delay function in (4) is much more difficult. Existing results for the SCALE strategy only give tight bounds when the delay function has certain scaling properties [23], [24], and unfortunately the delay functions in (4) do not have these properties. Consequently, existing results provide bounds which are too weak to provide meaningful, qualitative insights. As a result, we focus on a particular network to derive bounds which are more useful for gaining a qualitative understanding.

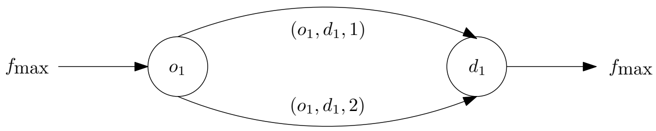


Fig. 2: Pigou's example [22] has certain worst-case properties when $\alpha = 0$.

The two node network with two parallel links shown in Fig. 2 is known as Pigou's example [22], and it has certain worst-case properties. Meaning, this simple network achieves the price of anarchy for when $\alpha = 0$. Consequently, we use a specific instance of this network to derive lower bounds on the price of anarchy for $\alpha \in (0, 1)$. We specifically pick delay function values of:

$$\begin{aligned} D_{(o_1,d_1,1)}(f_{(o_1,d_1,1)}) &= K \\ D_{(o_1,d_1,2)}(f_{(o_1,d_1,2)}) &= \frac{K(C_{\text{min}} - f_{\text{max}})}{C - f_{(o_1,d_1,2)}}. \end{aligned} \tag{12}$$

By construction, the second link has less delay

$$D_{(o_1,d_1,2)}(f_{(o_1,d_1,2)}) < D_{(o_1,d_1,1)}(f_{(o_1,d_1,1)}) \tag{13}$$

for $f_{(o_1,d_1,2)} < f_{max}$, and they are equal at f_{max}.

The average delay of different routing strategies can be computed by using the variational inequality formulation of UE flows. The summary of this formulation is that the flow follows paths of minimum delay, and all flows for a given origin-destination pair have the same delay [22]. Also, recall that the SO flow is given by the UE flow for the marginal delay. The UE flow is given by the path of minimum delay: $\overline{f}_{(o_1,d_1,1)} = 0$ and $\overline{f}_{(o_1,d_1,2)} = f_{\text{max}}$, and the average delay is $f_{\text{max}}K$. On the other hand, the SO flow occurs when the marginal delay of the two paths are equal. If amount of flow on the second link is $\tilde{f}_{(o_1,d_1,2)} = \lambda f_{\text{max}}$, then we can calculate λ by solving the following equation:

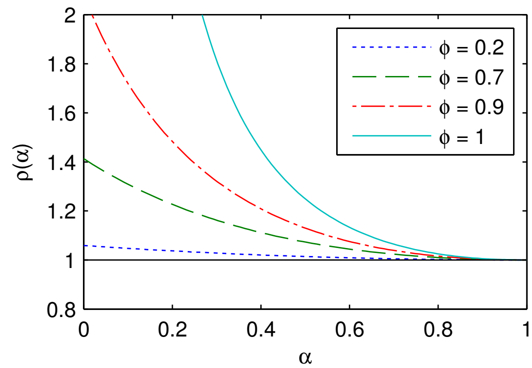
$$\begin{aligned} D^*_{(o_1,d_1,1)}(\lambda f_{\text{max}}) &= D^*_{(o_1,d_1,2)}((1-\lambda) f_{\text{max}}) \\ \frac{K(C_{\text{min}} - f_{\text{max}})}{C - \lambda f_{\text{max}}} &+ \lambda f_{\text{max}} \frac{K(C_{\text{min}} - f_{\text{max}})}{(C - \lambda f_{\text{max}})^2} = K. \end{aligned} \tag{14}$$

Simplifying the expression and using the quadratic equation gives that $\lambda = C_{\text{min}}\left(1 - \sqrt{1 - f_{\text{max}}/C_{\text{min}}}\right)/f_{\text{max}}$. Thus, the average delay is

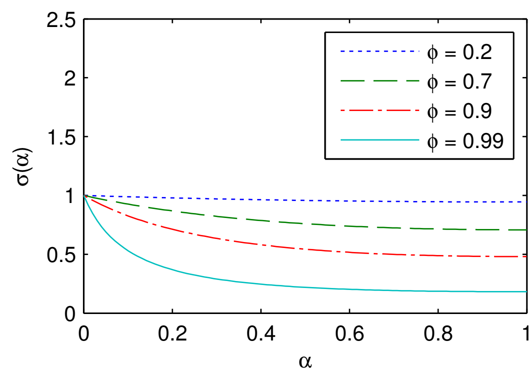
$$\lambda f_{\text{max}} \frac{K(C_{\text{min}} - f_{\text{max}})}{C_{\text{min}} - \lambda f_{\text{max}}} + (1-\lambda) f_{\text{max}} K. \tag{15}$$

We can use these results to compute the average delay for a SCALE flow where α fraction of the users are Stackelberg leaders. The Stackelberg leaders have flow given by $\alpha f_{(i,j,m)}$. The Stackelberg followers have flow given by $f^*_{(o_1,d_1,1)} = 0$ and $f^*_{(o_1,d_1,2)} = (1-\alpha) f_{\text{max}}$, because all of the Stackelberg followers use the second link which has lower delay. Thus, the average delay for a SCALE flow is

$$\begin{aligned} (\alpha\lambda + (1-\alpha)) f_{\text{max}} &\frac{K(C_{\text{min}} - f_{\text{max}})}{C_{\text{min}} - (\alpha\lambda + (1-\alpha)) f_{\text{max}}} \\ &+ \alpha(1-\lambda) f_{\text{max}} K. \end{aligned} \tag{16}$$



(a) Price of Anarchy



(b) Value of Altruism

Fig. 3: Lower bounds for the price of anarchy $\rho(\alpha)$ and upper bounds for the value of altruism $\sigma(\alpha)$ are plotted as functions of α, for four different values of ϕ.

This provides a lower bound on the price of anarchy: $\rho \geq (16)/(15)$. We can further simplify by setting $f_{\text{max}} = \phi C_{\text{min}}$, and the interpretation is that ϕ gives the largest fraction of utilization in the road links. After algebraic simplifications, we get that

$$\rho \geq \frac{\alpha(1-\lambda) + (\alpha\lambda + 1 - \alpha)\frac{1-\phi}{1-(\alpha\lambda+1-\alpha)\phi}}{1 - \lambda + \lambda\frac{1-\phi}{1-\lambda\phi}} \tag{17}$$

$$\lambda = \left(1 - \sqrt{1-\phi}\right)/\phi. \tag{18}$$

This expression is a function of the fraction of users α which follow the GPS directions and the maximum fraction of road utilization ϕ.

This lower bound on the price of anarchy ρ is plotted in Fig. 3a, and it gives several important qualitative insights. As road utilization ϕ increases, a larger fraction of compliant GPS users α are needed in order to make the average delay of the SCALE flow close to the average delay of the SO flow (i.e., make ρ close to one). Improvements in the price of anarchy ρ saturate at about $\alpha = 0.6$. This means that if roughly 60% of vehicles are compliant GPS users, then the average delay of the SCALE flow will be close to the optimal average delay of the SO flow.

B. Value of Altruism

Using the results from Sect. IV-A, we can provide a bound on the value of altruism. An upper bound on the value of altruism is given by: $\sigma \leq (16)/K$. Setting $f_{\max} = \phi C_{\min}$ and performing algebraic simplications, we get that

$$\sigma \leq \alpha(1-\lambda) + (\alpha\lambda + 1 - \alpha)\frac{1-\phi}{1-(\alpha\lambda+1-\alpha)\phi}, \quad (19)$$

where λ is given in (18). This upper bound on the price of anarchy ρ is plotted in Fig. 3b, and it gives several important qualitative insights on the reduction of average delay with the SCALE flow as compared to the UE flow (i.e., value of altruism). As road utilization ϕ increases, larger reductions in average delay with the SCALE flow are possible (i.e., $\rho(1)$ is smaller). Once again, improvements in the value of altruism σ saturate at about $\alpha = 0.6$. This means that if roughly 60% of vehicles are compliant GPS users, then no further reductions in average delay with the SCALE flow are possible. However, when ϕ is large, small values of α still lead to large reductions in congestion; thus, having even a small percentage of compliant users can still be advantageous when there is heavy congestion.

V. Conclusions

We modeled turn-by-turn directions of GPS navigation devices as an implementation of a Stackelberg game. Though we simplify existing traffic models and the information pattern, it allows for a tractable analysis of traffic routing and leads to a simple algorithm (which existing GPS devices can use) for providing turn-by-turn directions. The main qualitative insight from the model is that a large fraction of complaint GPS users are needed in order to effect real improvements in congestion and energy consumption. If roughly 60% of users follow GPS directions implementing a SCALE strategy, then the average delay will be close to the optimal average delay for the network. This poses social and technological challenges for the implementation of an ITS system which reduces congestion through routing. Fortunately, our qualitative results indicate that having a small percentage of compliant users can still lead to large reductions in congestion. We used a simple network topology to study these issues, and an open question is how these results translate to real road topologies.

VI. Acknowledgments

We thank Alexandre Bayen and Alex Kurzhanskiy for useful discussions and comments.

References

[1] L. Figueiredo, I. Jesus, J. A. T. Machado, J. Ferreira, and J. M. de Carvalho, "Towards the development of intelligent transportation systems," in *IEEE Intelligent Transportation Systems Conference Proceedings*, 2001, pp. 1206–1211.

[2] N.-E. El Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges," *Information Fusion*, 2010, to appear.

[3] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 271–281, 2002.

[4] G. Gomes, R. Horowitz, A. Kurzhanskiy, J. Kwon, and P. Varaiya, "Behavior of the cell transmission model and effectiveness of ramp metering," *Transportation Research Part C*, vol. 16, no. 4, pp. 485–513, 2007.

[5] Y. Zhao, "Mobile phone location determination and its impact on intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 1, pp. 55–64, 2000.

[6] J. Herrera, D. Work, R. Herring, X. Ban, Q. Jacobson, and A. Bayen, "Evaluation of traffic data obtained via gps-enabled mobile phones: the *Mobile Century* field experiment," *Transportation Research C*, 2009, to appear.

[7] A. Thiagarajan, L. R. Sivalingam, K. LaCurts, S. Toledo, J. Eriksson, S. Madden, and H. Balakrishnan, "VTrack: Accurate, Energy-Aware Traffic Delay Estimation Using Mobile Phones," in *ACM Conference on Embedded Networked Sensor Systems*, 2009.

[8] E. Schmitt and H. Jula, "Vehicle route guidance systems: Classification and comparison," in *IEEE Intelligent Transportation Systems Conference*, 2006, pp. 242–247.

[9] S. Dafermos and F. Sparrow, "The traffic assignment problem for a general network," *Journal of Research of the National Bureau of Standards: Part B*, vol. 73, pp. 91–118, 1969.

[10] F. B. Zhan and C. E. Noon, "Shortest path algorithms: an evaluation using real road networks," *Transportation Science*, vol. 32, pp. 65–73, 1998.

[11] H. Yang, X. Zhang, and Q. Meng, "Stackelberg games and multiple equilibrium behaviors on networks," *Transportation Research Part B: Methodological*, vol. 41, no. 8, pp. 841–861, 2007.

[12] R. Hall, "The fastest path through a network with random time-dependent travel time," *Transportation Science*, vol. 20, no. 3, pp. 182–188, 1986.

[13] T. Friesz, J. Luque, R. Tobin, and W. Byung-Wook, "Dynamic network traffic assignment considered as a continuous time optimal control problem," *Operations Research*, vol. 37, no. 6, pp. 893–901, 1989.

[14] A. Karimi, A. Hegyi, B. D. Schutter, J. Hellendoorn, and F. Middelham, "Integrated model predictive control of dynamic route guidance information systems and ramp metering," in *IEEE Conference on Intelligent Transportation Systems*, 2004, pp. 491–496.

[15] C. Papadimitriou and M. Yannakakis, "Shortest paths without a map," *Theoretical Computer Science*, vol. 84, no. 1, pp. 127–150, 1991.

[16] P. Sarachik and U. Ozguner, "On decentralized dynamic routing for congested traffic networks," *IEEE Transactions on Automatic Control*, vol. 27, no. 6, pp. 1233–1238, 1982.

[17] D. Bertsekas and J. Tsitsiklis, "An analysis of stochastic shortest path problems," *Mathematics of Operations Research*, vol. 16, no. 3, pp. 580–595, 1991.

[18] J. Wardrop, "Some theoretical aspects of road traffic research," *Proceedings of Institute of Civil Engineers: Part II*, vol. 1, pp. 325–378, 1952.

[19] F. P. Kelly, "Network routing," *Philosophical Transactions: Physical Sciences and Engineering*, vol. 337, no. 1647, pp. pp. 343–367, 1991.

[20] T. Roughgarden, "On the severity of Braess's paradox: designing networks for selfish users is hard," *Journal of Computer and System Sciences*, vol. 72, no. 5, pp. 922–953, 2006.

[21] R. Gallager, "A minium delay routing algorithm using distributed computation," *IEEE Transactions on Communications*, vol. 25, no. 1, pp. 73–85, 1977.

[22] T. Roughgarden, "The price of anarchy is independent of the network topology," *Journal of Computer and System Sciences*, vol. 67, no. 2, pp. 341–364, 2003.

[23] C. Swamy, "The effectiveness of Stackelberg strategies and tolls for network congestion games," in *ACM-SIAM Symposium on Discrete Algorithms*, 2007, pp. 1133–1142.

[24] V. Bonifaci, T. Harks, and G. Schäfer, "Stackelberg routing in arbitrary networks," *Mathematics of Operations Research*, vol. 35, no. 2, pp. 330–346, 2010.

[25] T. Roughgarden, "Stackelberg scheduling strategies," in *ACM Symposium on Theory of Computing*, 2001, pp. 104–113.

[26] V. Kumar and M. Marathe, "Improved results for Stackelberg scheduling strategies," in *Automata, Languages and Programming*, 2002, pp. 775–775.

[27] T. Roughgarden, "Selfish routing with atomic players," in *ACM-SIAM Symposium on Discrete Algorithms*, 2005, pp. 1184–1185.

[28] Y. Sharma and D. P. Williamson, "Stackelberg thresholds in network routing games or the value of altruism," *Games and Economic Behavior*, vol. 67, no. 1, pp. 174–190, 2009.