2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)
Windsor Oceanico Hotel, Rio de Janeiro, Brazil, November 1-4, 2016

Negative Externalities of GPS-Enabled Routing Applications: A Game Theoretical approach

Jérôme Thai[1] and Nicolas Laurent-Brouty[2] and Alexandre M. Bayen[1,2]

***Abstract*— This work studies the impact of the increasing penetration of routing apps on road usage. Its conclusions apply both to manned vehicles in which human drivers follow app directions, and unmanned vehicles following shortest path algorithms. To address the problem caused by the increased usage of routing apps, we model two distinct classes of users, one having limited knowledge of low-capacity road links. This approach is in sharp contrast with some previous studies assuming that each user has full knowledge of the network and optimizes his/her own travel time. We show that the increased usage of GPS routing provides a lot of benefits on the road network of Los Angeles, such as decrease in average travel times and total vehicle miles traveled. However, this global increased efficiency in urban mobility has negative impacts as well, which are not addressed by the scientific community: increase in traffic in cities bordering highway from users taking local routes to avoid congestion.**

I. INTRODUCTION

A. Motivation

Navigation applications such as Google Maps, Waze, INRIX, or Apple Maps, have deeply modified our approach of driving in the past. Pushed by the increasing penetration of smart phones and the rapid expansion of Mobility-as-a-Service systems such as Uber and Lyft, a significant percentage of drivers now use these tools daily, as they provide an easy way to optimize one's route choices and decrease one's travel time, specifically during peak hours. Since public agencies cannot indefinitely extend the capacity of urban road networks, these tools represent an opportunity to reallocate traffic in a way that might be more efficient (or not). Nonetheless, the impact of these applications on road traffic and urban congestion are not well-studied and understood. Cities bordering major highways in the United States have noticed an increase of traffic demand on their networks, presumably due to application users leaving highways to avoid congestion [1]. This alleged flow transfer is a challenge for public policy, as cities infrastructure, mostly financed by and for local taxpayers, receive a higher traffic demand.

[1]Electrical Engineering and Computer Science, University of California at Berkeley, CA
[2]Civil and Environmental Engineering, University of California at Berkeley, CA
Contact: jerome.thai@berkeley.edu

The aim of the present work is to propose and develop a framework to describe heterogeneous traffic in which a percentage of drivers use these applications. The main research question is the following: "how does the percentage of application users impact traffic redistribution and corresponding optimality of flows assignment?"

Historically, high-capacity roads, *e.g.* expressways and highways, have been developed to improve safety, comfort, and traveling speed. Today, a vast majority of drivers will consciously choose an expressway over a smaller road, because of all the previous benefits. The present work thus assumes that drivers, when traveling from an origin to a destination, will aim at minimizing the time spent on low-capacity (or low-speed) roads.

While the present work investigates the question of the impact of navigation applications on traffic, our framework encompasses heterogeneous traffic containing both classical manned vehicles and autonomous vehicles. Specifically, an autonomous vehicle can be modeled as a vehicle following real-time routing information, the same way a user follows instructions from routing services.

B. Approach and terminology

In order to address the research questions summarized above, we model the behavior of users on the road network with the established traffic assignment framework [2], in which each user traveling from their origin (*e.g.* their home) to their destination (*e.g.* their office) selfishly minimizes their own cost function. However, the transportation literature generally assumes that, for each user, the cost of traveling on a given route is the travel time of this route, see, *e.g.* [3]. Hence, state of the art work implicitly assumes that each user has access to the travel time of each link in the network and rationally chooses the shortest route to its destination. Contrasting from previous approaches, we model two types of users:

***Routed users*:** they have access to navigation information and thus follow the shortest route from their origin to their destination based on the network's current travel times. These vehicles can be drivers equipped

978-1-5090-1889-5/16/$31.00 ©2016 IEEE

with a GPS device (*e.g.* Garmin, TomTom, embedded navigation system), or a GPS-enabled mobile phone with a navigation app (*e.g.* Google maps, Waze, Apple maps), or they can be connected autonomous vehicle following routing directions from navigation services. Hence, for *routed users*, the cost of using a route is its travel time. In addition, users with expert knowledge of the network are also considered as *routed users* since they are able to find shortest routes without the use of navigation apps.

***Non-routed users*:** They do not have access to updated traffic information and thus have a limited knowledge of the travel times in the network. Since highways traditionally enable to travel with limited information and provide perceived benefits such as safety and higher travel speeds, *non-routed users* are assumed to choose high-capacity roads over low-capacity ones. The precise mathematical model of the behavior of these *non-routed users* will be introduced below.

The lack of information of users has been addressed previously in the field of transportation [4], [5], and in economics [6]. They collectively describe *bounded rational* users who make suboptimal choices due to the lack and/or price of information. Since local roads are arguably less known while major highways are in the information set of most of users, we choose an approach similar to studies modeling users with different objective functions than just minimizing travel times, *e.g.* seeking out less congested or scenic routes [7]. However, instead of using the nested logit model [8], we model the preference of *non-routed users* for larger roads segments and their limited knowledge of small streets. Hence, we define two types of road segments:

High-capacity road segments: highways and major arterial roads and avenues. High-capacity roads mainly serve users just passing through or nearby the city to go to their destination, hence they are maintained at a county or state level. We also assume that *non-routed users* favor this type of roads since, with limited knowledge on the local network, they represent a convenient way to move towards the destination by following signs.

Low-capacity road segments: They include small residential streets and small arterial streets. The low-capacity network is maintained by local taxpayers and is designed to provide mobility to local users, who either live or work in the area. It was originally not meant by planners to be used by through traffic, which should be confined to the high-capacity network.

Multiplicative cognitive cost to encode user choice: We add a multiplicative factor $C > 1$ to low-capacity links' cost functions to model the preference of *non-routed users* for high-capacity links. The multiplicative cognitive cost conserves the proportions between low-capacity links' travel times and models users that want to reduce the time spent on low-capacity links in favor to high-capacity ones. We also show that, in the Los Angeles network, in free flow, preference for highways is rational since it enables the users to choose routes that are close to being optimal without the use of GPS routing.

Heterogeneous game: To study the increasing penetration of GPS routing, we consider a heterogeneous routing game with two types of users: *routed users* for which the cost of using an edge is the travel time, and *non-routed users* for which the cost of using an edge is the travel time if it is high-capacity, or C times the travel time if it is low-capacity. Heterogeneous games have been studied before for the purpose of designing toll strategies [9], [10], and in a more general setting in [11]. To our knowledge, this is the first use of heterogeneous games to model the impact of routing via navigation apps, on flow allocation.

C. Outline and contributions

The main contribution of the article is twofold. In Section II, we introduce the concept of *multiplicative cognitive cost* to model *non-routed users*' preference for high-capacity roads and show that this choice is in general rational under low traffic demand. However, during peak hours, we show that this preference results in a poor allocation of the traffic with higher travel times, thus encouraging app based routing. In Section III, we expand on the established heterogeneous traffic assignment problem to quantify the road usage when there is a ratio α of *routed users* and $1 - \alpha$ of *non-routed users* in the urban network. We show that the use of app-based routing is rational since it decreases each user's travel time and allocates the flow efficiently throughout the network. However, this hidden cost is high as the low-capacity network sees a significant increase in traffic pressuring local governments to build additional infrastructure to reduce the nuisance related to it.

II. A MULTIPLICATIVE COGNITIVE COST MODEL

In this section, we present and motivate the multiplicative cognitive cost model using the traffic assignment framework.

A. Mathematical formulation and notations

We consider a given road network modeled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ with vertex set $\mathcal{V}$ and directed arc set $\mathcal{A}$. We note $\mathcal{W} \in \mathcal{V} \times \mathcal{V}$ the set of origin-destination vertex pairs. Each OD pair $w = \{s, t\} \in \mathcal{W}$

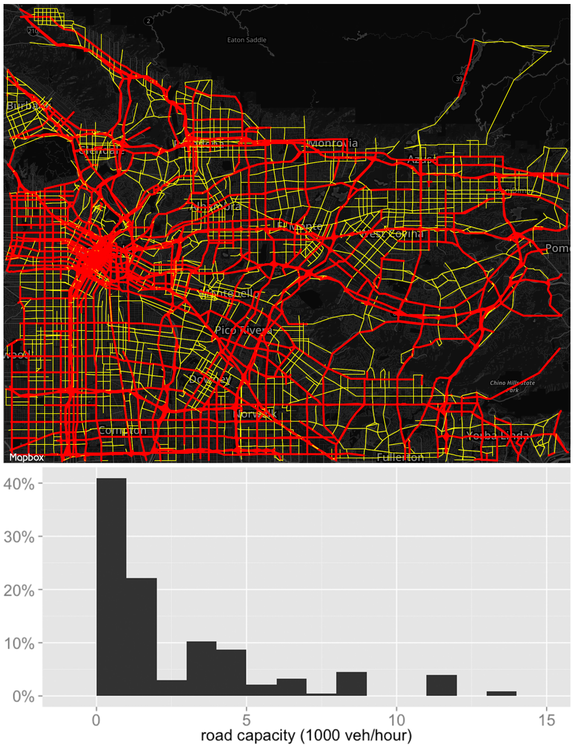

Fig. 1. **The map of Los Angeles, CA used for the present study composed of 28,376 arcs and 14,617 nodes extracted from OpenStreetMap. Information for each edge includes the free-flow travel time, length, capacity, and speed limit. Links with capacity less than 1000 vehicles per hour are considered low-capacity (in yellow) while links with 1000 vehicles per hour or more are considered high-capacity (red). The histogram of the different road capacities are shown in the bottom figure, with more than 40% of low-capacity links.**

models a population of drivers traveling from their origin $s \in \mathcal{V}$ to their destination $t \in \mathcal{V}$ at a rate d_w. They choose between routes $p \in \mathcal{P}_w$ such that their travel cost is minimized, where $\mathcal{P}_w$ is the set of all paths from s to t. Hence, the state of the network is described by the vector of route flows $f = [f_p]_{p \in \mathcal{P}} \in \mathbb{R}^{\mathcal{P}}$ where $\mathcal{P} = \cup_{w \in W} \mathcal{P}_w$ is the set of all paths in the network. A flow vector $f \in \mathbb{R}^P$ is then feasible if for all $w \in W$, $\sum_{p \in P_w} f_p = d_w, \quad f_p \geq 0, \forall p \in \mathcal{P}_w$. In matrix form, f said to be feasible if it belongs to the following set

$$\mathcal{X} := \{f \in \mathbb{R}^{\mathcal{P}} \,:\, f \succeq 0, \, \Lambda f = d\} \tag{1}$$

where Λ is the OD-path incidence matrix. *Non-routed users* have travel costs $\ell_p^{\text{nr}}(\cdot)$ along each path $p \in \mathcal{P}$ given by $\ell_p^{\text{nr}}(f) = \sum_{a \in p} c_a^{\text{nr}}(x_a)$ where $c_a^{\text{nr}}(x_a)$ is the *non-routed users*' cost of link a. We assume that the cost of a road segment a only depends on the flow x_a of vehicles on this segment, where x_a is expressed as $x_a = \sum_{p \in \mathcal{P}} I(a \in p) f_p$, the sum of the flows of every route passing through a, where $I(B)$ is the indicator function of the Boolean B, *i.e.* equal to 1 if B is true, and 0 otherwise. In matrix form, $x = \Delta f$ where the arc-path incidence matrix is given by $\Delta = [I(a \in p)]_{a \in \mathcal{A},\, p \in \mathcal{P}}$. Hence, we write that an arc flow vector $x = [x_a]_{a \in \mathcal{A}}$ is feasible if it is in the following set

$$\mathcal{K} := \{x \in \mathbb{R}^{\mathcal{A}} \,:\, \exists\, f \in \mathcal{X},\, x = \Delta f\} \tag{2}$$

We formalize the behavior of *non-routed users* by partitioning the arc set A into a set of low-capacity arcs $\mathcal{A}^{\text{lo}} := \{a \in \mathcal{A} \,:\, c_a < c_{\text{lo}}\}$ and a set of high-capacity arcs $\mathcal{A}^{\text{hi}} := \{a \in \mathcal{A} \,:\, c_a \geq c_{\text{lo}}\}$ where each arc has a capacity c_a and c^{lo} is an arbitrary threshold. Throughout our study, we consider road segments with capacities less than 1000 vehicles per hour as low-capacity, which amount for 40% of the road segments in the Los Angeles network, see Figure 1. The *non-routed users*' costs are then

$$c_a^{\text{nr}}(x_a) = \begin{cases} C \cdot t_a(x_a) & \text{if } a \in \mathcal{A}^{\text{lo}} \\ t_a(x_a) & \text{if } a \in \mathcal{A}^{\text{hi}} \end{cases} \tag{3}$$

This results in the following non-routed path costs

$$\ell_p^{\text{nr}}(f) = \sum_{a \in p^{\text{hi}}} t_a(x_a) + C \sum_{a \in p^{\text{lo}}} t_a(x_a) \tag{4}$$

where $t_a(x_a)$ is the travel time of road segment a under flow x_a, $C > 1$ is a constant that models how strongly non-routed users favor high-capacity roads over low-capacity roads, and p^{hi} (resp. p^{lo}) are the segments of roads in path p that are high (resp. low) capacity. Note that the multiplicative cognitive cost conserves the proportions between low-capacity links' travel times while increasing their costs.

B. Rationale behind preference for high-capacity links

Under low traffic demand, high-capacity roads generally enable to travel quickly between origins and destinations far apart. To validate this on the Los Angeles network, we collected the OD trip data from the American Community Survey (ACS), composed of a set W of 96,077 OD pairs and a demand vector $d \in \mathbb{R}^W$. In the Los Angeles network in free flow, we extracted a path p_w^{nr} with lowest non-routed cost $\min_{p \in \mathcal{P}_w} \ell_p^{\text{nr}}(0)$ for each OD pair $w \in W$ using python-igraph package, and found that that associated free-flow travel time $\sum_{a \in p_w^{\text{nr}}} t_a(0)$ is on average only 10% longer than the shortest route, as illustrated by Figure 2.a). In addition, travel times of non-routed users in the free-flow regime are not sensitive to the cognitive cost when it is above 1000. Hence, for the remainder of this work, we fix the non-routed costs c_a^{nr} with a cognitive cost $C = 3000$ and focus on the sensitivity of road usage to variations in the traffic demand and in the percentage of *routed users*. Moreover, Figure 2.b) shows a small shift of the travel time distribution in positive direction as the cognitive cost increases from 1 to 1000. Hence, without traffic, the Los Angeles high-capacity network provides a reliable and nearly optimal route for traversing cities with no information on local roads, thus justifying the rationale behind *non-routed users*' preference.

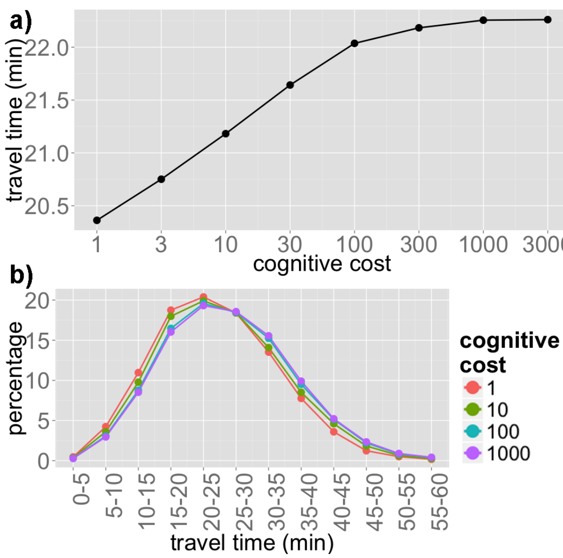

Fig. 2. **Travel times in Los Angeles when all edges are in free flow for *non-routed users* with perceived costs given by** (3)**, as a function of the cognitive cost** C**. Figure a) shows the average travel time, Figure b) shows the distribution of travel times.**

C. *Rationale behind routing on low-capacity links*

With increasing demand, high-capacity roads such as highways become congested since *non-routed users* choose them over low-capacity routes. We model flow of vehicles on roads using the traffic assignment framework [2] in which each *non-routed user*, represented as an infinitesimal amount of flow, selfishly chooses the path with the lowest cost $\ell_p^{\text{nr}}(f)$. This concept is known in the transportation literature as Wardrop's first principle [12]. The resulting flow is an equilibrium flow $f \in \mathbb{R}^{\mathcal{P}}$ for which the associated equilibrium edge flow $x = [x_a]_{a\in\mathcal{A}} \in \mathbb{R}^{\mathcal{A}}$ is unique when the travel time functions t_a are continuously differentiable, positive and strictly increasing [13]. Under these assumptions on the travel functions, Beckmann et al. [13] show that the equilibrium edge flow of the routing game can be expressed as the optimal solution of the following convex program

$$\min_x \ \phi(x) = \sum_{a\in\mathcal{A}} \int_0^{x_a} c_a^{\text{nr}}(u)du \quad \text{s.t.} \quad x \in \mathcal{K} \tag{5}$$

where ϕ is a potential function, c_a^{nr} is given by (3), and $\mathcal{K}$ is given by (2). We obtain different traffic demands by multiplying the demand vector $d \in \mathbb{R}^{\mathcal{W}}$ obtained from the ACS data by a scalar $\alpha \in [0.1, 1]$. We then solve (5) with a cognitive cost $C = 3000$ and different traffic demands to obtain various non-routed equilibrium flows x^{nr}. The network with 100% of *non-routed users* settles in a suboptimal state with imbalances in the flow allocation where high-capacity links are over-utilized and low-capacity links are under-utilized. We compare it to the *routed equilibrium* arc flow x^{r}, where every user follows the shortest path, with costs given by

$$\ell_p^{\text{r}}(f) = \sum_{a\in p} t_a(x_a), \quad \forall\, p \in \mathcal{P} \tag{6}$$

The equilibrium is obtained by solving (5) with arc costs $c_a^{\text{nr}}(\cdot)$ equal to the travel time functions $t_a(\cdot)$. The ratio of the respective total travel times $\sum_{a\in A} x_a^{\text{nr}} c_a^{\text{nr}}(x_a^{\text{nr}})$ and $\sum_{a\in A} x_a^{\text{r}} t_a(x_a^{\text{r}})$ are shown in turquoise in Figure 3. Figure 4.b) also shows that 20% of the users experience a 10-20% delay and 12% experience a 20-30% delay compared to the routed equilibrium.

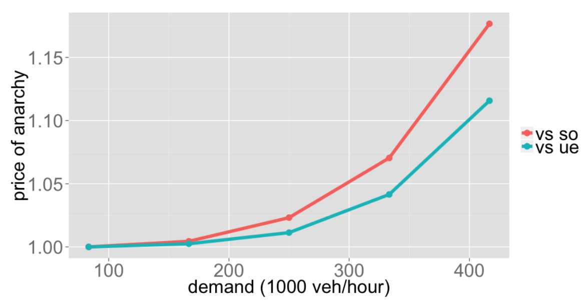

Fig. 3. **Ratio of the average travel time when the perceived non-routed costs are given by** (3) **with** $C = 3000$ **over the user equilibrium (blue) and the social optimum (red), as a function of the demand in the network.**

We also compare the non-routed equilibrium to the *social optimum* where the total cost incurred by all users in the network is minimized

$$\min \sum_{a\in A} x_a t_a(x_a) \quad \text{s.t.} \quad x \in \mathcal{K} \tag{7}$$

Figures 3 and 4 show that the preference for high-capacity links steers the equilibrium state further from the social optimum where 25% of users experience 10-20% delay and 18% of users experience a 20-30%

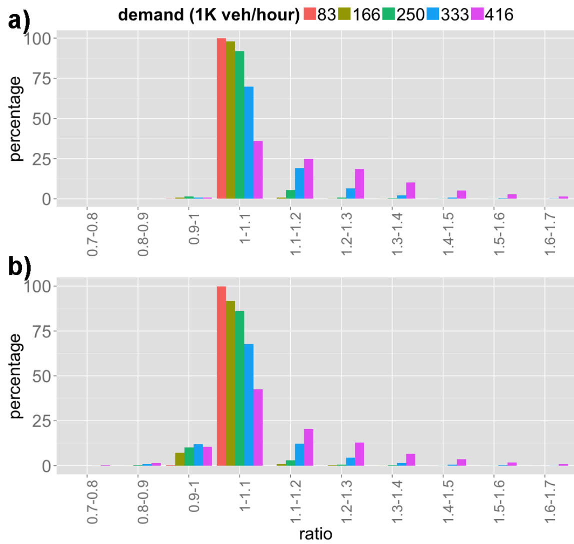

Fig. 4. **The distribution of the ratio of the travel times over the social optimum per OD pair (a), and the user equilibrium (b), when all users are non-routed, when the perceived costs are given by** (3) **with** $C = 3000$**.**

delay. Hence, rational users are pushed to choose low-capacity roads to avoid segments of high-capacity roads that are not along the shortest route due to congestion under heavy traffic demand.

III. Multiclass traffic assignment problem

The sharp increase of app-based routing spurred by the increasing penetration of navigation devices progressively increases the number of *routed users* on the road. It is likely that with the full advent of automated driving, this trend will accelerate in the future. This emerging behavior is in sharp contrast with non-routed users who favor high-capacity roads regardless to the level of congestion. To quantify the impact of *routed users* on traffic conditions, we introduce our heterogeneous traffic assignment problem with both *routed users* and *non-routed users*.

A. Multiclass traffic assignment problem

We consider a flow $d^{\mathrm{r}}_w \in \mathbb{R}^{\mathcal{W}}_+$ of *routed users* and a flow $d^{\mathrm{nr}}_w \in \mathbb{R}^{\mathcal{W}}_+$ of *non-routed users* between each OD pair w. The state of the network is described by the routed users' path flow vector $f^{\mathrm{r}} = [f^{\mathrm{r}}_p]_{p\in\mathcal{P}}$ and the non-routed users' path flow vector $f^{\mathrm{nr}} = [f^{\mathrm{nr}}_p]_{p\in\mathcal{P}}$. They are feasible if they are in $\mathcal{X}^{\mathrm{r}}$, $\mathcal{X}^{\mathrm{nr}}$ given by

$$\mathcal{X}^{\mathrm{r}} := \{f^{\mathrm{r}} \in \mathbb{R}^P \,:\, f^{\mathrm{r}} \succeq 0,\, \Lambda f^{\mathrm{r}} = d^{\mathrm{r}}\} \tag{8}$$

$$\mathcal{X}^{\mathrm{nr}} := \{f^{\mathrm{nr}} \in \mathbb{R}^P \,:\, f^{\mathrm{nr}} \succeq 0,\, \Lambda f^{\mathrm{nr}} = d^{\mathrm{nr}}\} \tag{9}$$

where Λ is the OD-path incidence matrix. With Δ the arc-path incidence matrix, we denote $x^{\mathrm{r}} = [x^{\mathrm{r}}_a]_{a\in\mathcal{A}} = \Delta f^{\mathrm{r}}$ and $x^{\mathrm{nr}} = [x^{\mathrm{nr}}_a]_{a\in\mathcal{A}} = \Delta f^{\mathrm{nr}}$ the routed and non-routed arc flow vectors respectively. Hence x^{r}, x^{nr} are feasible if they belong to the following sets respectively

$$\mathcal{K}^{\mathrm{r}} := \{x^{\mathrm{r}} \in \mathbb{R}^A \,:\, \exists\, f^{\mathrm{r}} \in \mathcal{X}^{\mathrm{r}},\, x^{\mathrm{r}} = \Delta f^{\mathrm{r}}\} \tag{10}$$

$$\mathcal{K}^{\mathrm{nr}} := \{x^{\mathrm{nr}} \in \mathbb{R}^A \,:\, \exists\, f^{\mathrm{nr}} \in \mathcal{X}^{\mathrm{nr}},\, x^{\mathrm{nr}} = \Delta f^{\mathrm{nr}}\} \tag{11}$$

The total path flow is $f = f^{\mathrm{r}} + f^{\mathrm{nr}} = [f^{\mathrm{r}}_p + f^{\mathrm{nr}}_p]_{p\in\mathcal{P}}$ and the total arc flow is $x = x^{\mathrm{r}} + x^{\mathrm{nr}} = [x^{\mathrm{r}}_a + x^{\mathrm{nr}}_a]_{a\in\mathcal{A}}$. As both routed and non-routed users make selfish choices by minimizing their associated costs, the resulting flow essentially describes the Nash equilibrium on road networks. Mathematically, the equilibrium flow are feasible flows $f^{\mathrm{r}} \in \mathcal{X}^{\mathrm{r}}$, $f^{\mathrm{nr}} \in \mathcal{X}^{\mathrm{nr}}$ such that $\forall\, w \in \mathcal{W}$

$$\forall\, p \in \mathcal{P}_w,\, f^{\mathrm{r}}_p > 0 \implies \ell^{\mathrm{r}}_p(f) = \min_{q\in\mathcal{P}_w} \ell^{\mathrm{r}}_q(f) \tag{12}$$

$$\forall\, p \in \mathcal{P}_w,\, f^{\mathrm{nr}}_p > 0 \implies \ell^{\mathrm{nr}}_p(f) = \min_{q\in\mathcal{P}_w} \ell^{\mathrm{nr}}_q(f) \tag{13}$$

where the routed and non-routed path costs ℓ^{r}_p and ℓ^{nr}_p are given by (6) and (4) respectively. Hence, only the least-cost paths are used between each origin and destination with respect to the associated type of users. The equilibrium f described in (12) and (13) can be expressed as a feasible solution $(f^{\mathrm{r}}, f^{\mathrm{nr}}) \in \mathcal{K}^{\mathrm{r}} \times \mathcal{K}^{\mathrm{nr}}$ of the following variational inequality problem

$$\ell^{\mathrm{r}}(f)^T g^{\mathrm{r}} + \ell^{\mathrm{nr}}(f)^T g^{\mathrm{nr}} \geq \tag{14}$$

$$\ell^{\mathrm{r}}(f)^T f^{\mathrm{r}} + \ell^{\mathrm{nr}}(f)^T f^{\mathrm{nr}}, \quad \forall\, (g^{\mathrm{r}}, g^{\mathrm{nr}}) \in K^{\mathrm{r}} \times K^{\mathrm{nr}} \tag{15}$$

Contrary to the homogeneous routing game, the general heterogeneous game cannot be formulated as a potential game of the form (5), see [14], [11]. However, by using the theory of variational inequality [15], it is possible to solve for the equilibrium described in (14), (15) with the Frank-Wolfe algorithm [16].

B. Positive impact

We apply the multi-class traffic assignment framework to the network of Los Angeles with a variable percentage α of *routed users*, and a cognitive cost $C = 3000$ for *non-routed users*, which means that their perceived cost on low-capacity links is 3000 times the real travel-time. We assume a uniform ratio of *routed users* for each OD pair, hence $d^{\mathrm{r}}_w = \alpha d$ and $d^{\mathrm{nr}}_w = (1-\alpha)d$, where $\alpha \in [0, 1]$ and the total traffic demand d is given by the ACS data. As the fraction α of *routed users* increases, Figure 5 shows a shift of the travel time distribution to the left as a result of users allocating themselves optimally (but selfishly) between the low-capacity and high-capacity networks. At an aggregate level, GPS routing can alleviate the road network with a possible decrease in Vehicle-miles Traveled (VMT) from 7.94 million miles per hour to

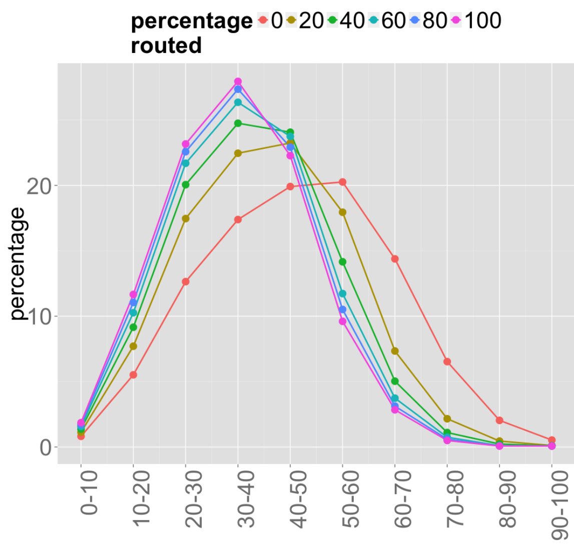

Fig. 5. **Distribution of travel times as a function of the percentage of routed users, with cognitive cost $C = 3000$ for *non-routed users*.**

7.15 million, hence a potential decrease of .79 million miles per hour, see Figure 6.b), thus corroborating the belief that GPS routing is able to alleviate gridlock in congested areas.

C. Negative externalities

Even though the increase in usage of app-based routing enables better navigation and time savings, they allegedly transfer large amounts of traffic in cities bordering highways, since navigation apps users have been reported to leave highways to avoid congestion [1]. For instance, in the Los Angeles network used for the present study and shown in Figure 1, we find that app-based routing can potentially increase the VMT on local roads by .34 million miles per hour, which represents a three-fold increase in traffic on low-capacity links, while there is only a 10% decrease in VMT on high-capacity roads, see Figure 6. Moreover, Figure 7 shows that an increase in *routed users*' ratio α is accompanied with a sharp increase in the percentage of users spending between 10 and 20 min on low-capacity links (we reiterate that we apply the framework to the Los Angeles network presented in Figure 1). Figure 8 shows that, despite a general decrease in VMT due to more efficient routing, the relative increase on low-capacity roads is very important for each 10% increase in routed users, due to the small traffic flow on the low-capacity network. This causes residential streets to be congested, encouraging cities to spend millions in infrastructure to steer the traffic away.

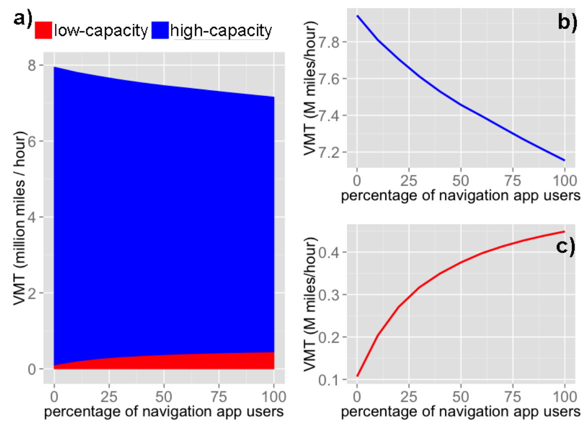

Fig. 6. **General VMT versus VMT on local roads as a function of the percentage of routed users.**

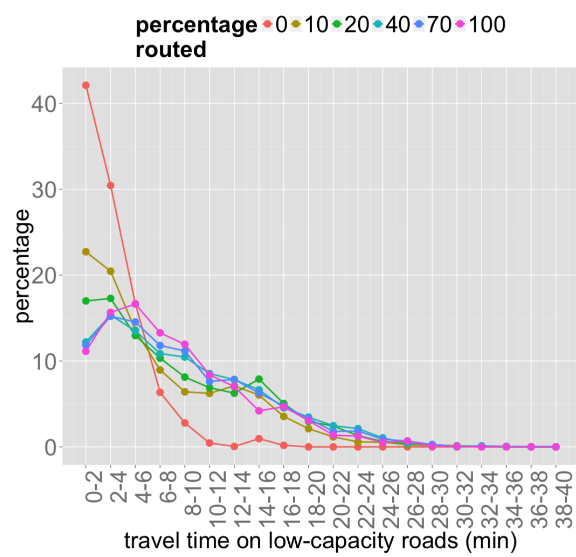

Fig. 7. **Distribution of travel times on local roads as a function of the percentage of routed users.**

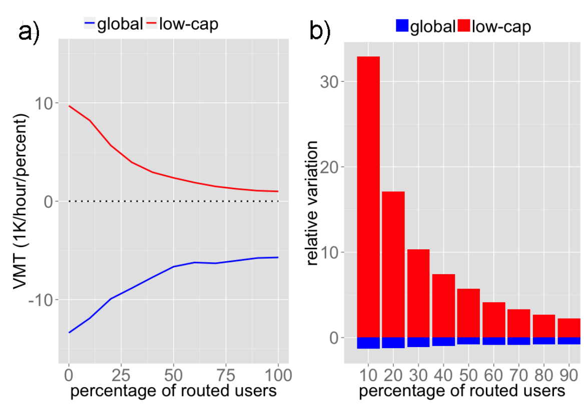

Fig. 8. **a) Variation in VMT for 1% increase in routed users. b) Relative variation in VMT for 10% increase in routed users.**

IV. Concluding remarks and future work

In practice it is of course difficult to accurately evaluate the current proportion of *routed users* in Los Angeles without accessing data from app-providers, like Google Maps or Waze. Hence we carry out a parametric study with a ratio of *routed users* varying from 0 to 1, with other parameters constant such as the traffic demand and the cognitive cost, to focus on the impact of app-based routing. In reality, traffic demand also increases in several urban areas such as in Los Angeles or in the Bay Area (along with the increasing penetration of app-based routing) hence we do not aim to predict the variation in VMT on low-capacity and high-capacity links. Instead, we put the emphasis on the fact that increasing penetration of app-based routing causes a sharp increase of traffic on local roads while there is little to no decrease of congestion on highways, as illustrated by Figure 8.

Finally, our work opens future avenues of research to mitigate the impact of app-based routing on cities. For a given city, a possible solution consists in reducing the capacity of its low-capacity network. However, this may push away traffic onto the low-capacity network of neighbouring cities, encouraging them to mimetically reduce their road capacity, thus exacerbating the congestion. Hence, this causes a 'prisoner's dilemma' effect that we would like to study. Another solution consists in implementing GPS-based (or 'Pay-as-you-go') road charges. However, such a plan may raise privacy issues that need to be addressed as well.

Appendix

Data collection: The road network of Los Angeles, CA is extracted from OpenStreetMap. The OD data is from the Census Transportation Planning Products database, and based on the 2006-2010 American Community Survey Data: ctpp.transportation.org/Pages/5-Year-Data.aspx.

Computing the OD costs: Once the equilibrium edge flow has been computed with the associated edge costs, the shortest paths between all pairs of nodes in the network can be computed very efficiently with the python-igraph package. The cost of a shortest path between each OD pair is then the OD cost.

Solving the homogeneous game: The equilibrium solver for the homogeneous game is based on Fukushima's modified Frank-Wolfe algorithm [17].

Solving the heterogeneous game: Similarly to the Frank-Wolfe algorithm applied to the homogeneous game, the descent direction for to the multi-class problem can be computed very efficiently by finding the shortest paths between all pairs of nodes in the network, which we implement with the python-igraph package.

ACKNOWLEDGMENT

We thank Alexei Pozdnoukhov and Madeleine Sheehan for insightful discussions and for providing us with data on Los Angeles.

References

[1] Joe Flint. In L.A., One Way to Beat Traffic Runs Into Backlash. *The Wall Street Journal*, 2016.
[2] M. Patriksson. *The Traffic Assignment Problem - Models and Methods*. VSP, Utrecht, 1994.
[3] Serdar Colak, Antonio Lima, and Marta C. Gonzalez. Understanding congested travel in urban areas. *Nature Communications*, 2016.
[4] Hani S. Mahmassani and Gang-Len Chang. On Boundedly Rational User Equilibrium in Transportation Systems. *Transportation Science*, 21, 1987.
[5] Song Gao, Emma Frejinger, and Moshe Ben-Akiva. Cognitive cost in route choice with real-time information: An exploratory analysis. *Transportation Research Part A*, 17:136–149, 2011.
[6] John Conlisk. Why Bounded Rationality? *Journal of Economic Literature*, 34(2):669–700, 1996.
[7] Moshe Ben-Akiva, M. J. Bergman, Andrew J. Daly, and Rohit Ramaswamy. Modelling Inter Urban Route Choice Behavior. *International Symposium on Transportation and Traffic Theory*, pages 299–330, 1984.
[8] Moshe E. Ben-Akiva and Steven R. Lerman. Discrete Choice Analysis: Theory and Application to Travel Demand. *MIT Press*, 1985.
[9] G. Karakostas and S. G. Kolliopoulos. Edge pricing of multicommodity networks for heterogeneous selfish users. *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, 4:268–276, 2004.
[10] P. Marcotte and D. L. Zhu. Existence and computation of optimal tolls in multiclass network equilibrium problems. *Operations Research Letters*, 37:211–214, 2009.
[11] Farhad Farokhi, Walid Krichene, Alexandre M. Bayen, and Karl H. Johansson. A Heterogeneous Routing Game. 2013.
[12] J. G. Wardrop and J. I. Whitehead. Correspondence. Some Theoretical Aspects of Road Traffic Research. *ICE Proceedings: Engineering Divisions 1*, 1952.
[13] M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation*. Cowles Commission Monograph, 1956.
[14] W. H. Sandholm. Potential Games with Continuous Player Sets. *Journal of Economic Theory*, 97:81–108, 2001.
[15] F. Facchinei and J. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York, 2003.
[16] Dao Li Zhu and Patrice Marcotte. Convergence Properties of Feasible Descent Methods for Solving Variational Inequalities in Banch SPaces. *Computational Optimization and Applications*, 10:35–49, 1997.
[17] Masao Fukushima. A Modified Frank-Wolfe Algorithm for Solving the Traffic Assignment Problem. *Transportation Research Part B*, 18:169–177, 1984.