Delft University of Technology

Delft Center for Systems and Control

Technical report 12-026

Dynamic optimal routing based on a reverse Stackelberg game approach*

N. Groot, B. De Schutter, and H. Hellendoorn

If you want to cite this report, please use the following reference instead:
N. Groot, B. De Schutter, and H. Hellendoorn, "Dynamic optimal routing based on a reverse Stackelberg game approach," *Proceedings of the 15th International IEEE Conference on Intelligent Transportation Systems (ITSC 2012)*, Anchorage, Alaska, pp. 782–787, Sept. 2012.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

Dynamic Optimal Routing Based on a Reverse Stackelberg Game Approach

Noortje Groot, Bart De Schutter, and Hans Hellendoorn

***Abstract*— A game-theoretic approach to dynamic routing is proposed in order to maximize the traffic throughput on a freeway network. While existing methods of informing drivers of the approximate travel times for the alternative routes do not in general yield the system optimum, we can achieve a better performance by introducing a leader-follower game with monetary incentives. In particular, a control strategy is proposed in which the traffic authority (the leader) proposes a function that maps the possible travel times for a certain destination to positive or negative monetary incentives. Based on this function that is communicated via on-board computers, the drivers (followers) will rationally choose those travel times associated with an optimal distribution over the available routes. Finally, in return for the associated monetary value, the drivers are presented with a route that they should follow to the desired destination.**

I. Introduction

A significant number of traffic networks around the world suffer from congestion caused by a disbalance between traffic demand and road capacity. Dynamic traffic routing aims at addressing this problem by guiding traffic to make use of alternative routes in the network [1]. Here, one strives after a system-optimal state of the network, e.g., in which the total travel time is minimized or throughput is maximized, respectively. However, when the individual drivers choose the route with the lowest cost, which can measured in travel time or distance, a different, user-equilibrium assignment is obtained in which the costs of the alternative routes are equally high [2].

In order to spread the traffic flow over the alternative routes to reach the system optimum, a possible approach is to adopt dynamic route guidance information panels (DRIP) that indicate the travel times when taking alternative routes [3], which can be integrated with ramp metering [4] and variable speed limits [5]. The assumption of the adoption of these panels is that traffic indeed spreads accordingly. However, there is inefficiency caused by the fact that the travel time between the alternative routes should be similar for drivers to deviate from the – originally shorter – route. The majority of traffic will thus choose the popular, shortest route until congestion applies and – due to the now equal or even shorter travel time of the alternative route on the DRIP – traffic will start to use the alternative route. This process leads to the user equilibrium under dynamic routing by providing information on the DRIPs.

In order to prevent this suboptimal behavior from appearing and to achieve an optimal distribution of traffic over the network, in this paper a hierarchical game-theoretic approach is proposed in which traffic is routed by the road authority based on monetary incentives. In this so-called reverse Stackelberg game [6], also known as ‘incentives’ [7] or as the ‘inverse Stackelberg game’ [8], the leader (road authority) is able to propose a function that maps the follower decision space into the leader decision space, i.e., that proposes a monetary incentive (the leader’s decision variable) for a certain choice of desired travel time by the drivers (the followers).

This hierarchical game has been applied in traffic control before, in assigning tolls to roads in order to accomplish a certain traffic assignment [9], [10]. However, there either the tolls are taken to be time-variant but fixed, i.e., constant according to the regular Stackelberg game [9], or tolls are a function of the traffic flow on the corresponding road [10]. Since the individual driver cannot know nor significantly influence this flow by his individual route choice, also in the latter case, the drivers basically make their route decision based on a fixed toll. Moreover, by tolling specific road stretches, the splitting rate of traffic at a specific intersection is influenced, but this is essentially a greedy approach and it does not necessarily lead to an optimal distribution of traffic flow over the alternative *routes*, as was also the case with the DRIP panels.

In order to do enforce a user equilibrium that coincides with the system optimum with respect to the overall throughput of traffic, we suggest a novel reverse Stackelberg approach to dynamic traffic routing. In this approach (i) more general monetary incentives are adopted that can also assume negative values, i.e., rewards instead of penalties; (ii) drivers decide upon their desired travel time rather than upon the route to travel, and (iii) complete routes rather than individual roads are associated with a monetary value.

The paper is organized as follows. After an introduction to the reverse Stackelberg game in Section II, the routing problem at hand is described in Section III. Subsequently, the reverse Stackelberg approach to routing is explained in Section IV, by first defining the basic elements of the game, followed by the overall dynamic framework. An illustrative case study is presented in Section V and the paper is concluded in Section VI.

Research supported by the BSIK project “Next Generation Infrastructures (NGI)”, the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 257462 HYCON2 Network of Excellence, and by the European COST Action TU1102 and TU0702.

N. Groot, B. De Schutter, and H. Hellendoorn are with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. E-mail: n.b.groot@tudelft.nl

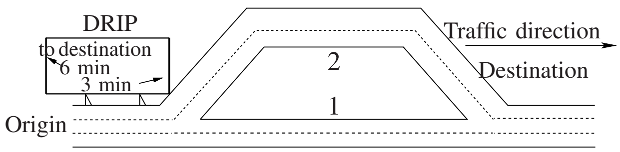


Fig. 1: Traffic network of Example 1.

II. The Reverse Stackelberg Game

The basic reverse Stackelberg game can be described as follows. A leader player proposes a leader function $u_{\mathrm{L}} = \gamma_{\mathrm{L}}(u_{\mathrm{F}})$, i.e., $\gamma_{\mathrm{F}} : \Omega_{\mathrm{F}} \to \Omega_{\mathrm{L}}$, after which the follower determines his optimal response $u_{\mathrm{F}} \in \Omega_{\mathrm{F}} \subseteq \mathbb{R}^{n_{\mathrm{F}}}$ with the associated leader decision variable $u_{\mathrm{L}} \in \Omega_{\mathrm{L}} \subseteq \mathbb{R}^{n_{\mathrm{L}}}$.

Here, the leader aims to achieve a desired equilibrium

$$(u_{\mathrm{L}}^{\mathrm{d}}, u_{\mathrm{F}}^{\mathrm{d}}) \in \arg \min_{u_{\mathrm{L}} \in \Omega_{\mathrm{L}}, u_{\mathrm{F}} \in \Omega_{\mathrm{F}}} \mathcal{J}_{\mathrm{L}}(u_{\mathrm{L}}, u_{\mathrm{F}}),$$

where $\mathcal{J}_{\mathrm{L}}(u_{\mathrm{L}}, u_{\mathrm{F}}) : \Omega_{\mathrm{L}} \times \Omega_{\mathrm{F}} \to \mathbb{R}$. Similarly, given the leader function $\gamma_{\mathrm{L}}(u_{\mathrm{F}})$, the follower optimizes his objective function $\mathcal{J}_{\mathrm{F}}(\gamma_{\mathrm{L}}(u_{\mathrm{F}}), u_{\mathrm{F}}) : \Omega_{\mathrm{L}} \times \Omega_{\mathrm{F}} \to \mathbb{R}$.

A well-known special case of this hierarchical leader-follower game is the original Stackelberg game [11]. In this game, the follower player determines his optimal decision variable $u_{\mathrm{F}} \in \Omega_{\mathrm{F}}$ as a response to the leader's (constant) decision variable $u_{\mathrm{L}} \in \Omega_{\mathrm{L}}$, thus not to the more general leader function $\gamma_{\mathrm{L}}(u_{\mathrm{F}})$. The reverse Stackelberg game therefore has an important advantage to the regular Stackelberg game, as there, the leader cannot control the follower's response in case it is not unique [6].

III. The Routing Problem

We now focus on achieving a system-optimal distribution of vehicle flows over a traffic network with respect to the throughput. In order to optimize the use of the available routes in a traffic network by drivers, we will strive after congestion avoidance by keeping the flow on a route below the bottleneck capacity or critical density as long as possible.

The current approach of guiding traffic by means of predicted travel times to reach a certain destination, shown on panels along the nodes of a highway network, does not in general lead to an optimal spread of traffic. The following simple example shows this scenario:

Example 1 (Reverse Stackelberg approach versus DRIP): Consider an origin-destination (OD) pair for which two routes are available as depicted in Fig. 1, and a traffic demand d^{OD} [veh/h]. The two alternative routes contain homogeneous road segments of total length $l^1 = 6\,\mathrm{km}$, $l^2 = 10\,\mathrm{km}$ and speed limits $v_{\max}^1 = 120\,\mathrm{km/h}$, $v_{\min}^2 = 100\,\mathrm{km/h}$, leading to minimum travel times of $\tau^1 = l^1/v_{\max}^1 = 3\,\mathrm{min}$ and similarly $\tau^2 = 6\,\mathrm{min}$.

Let q_{crit}^r [veh/h] and ρ_{crit}^r [veh/lm/lane] be the critical flow (capacity) and density of a freeway segment r. If the demand is high enough such that the flow and density on route 1 exceed the critical flow or capacity q_{crit}^r: $q_{\mathrm{crit}} < q^1 \leq q_{\max}$, and the density $\rho^1 \leq \rho_{\mathrm{crit}}^1$ the unstable traffic state causes a reduction in speed, hence an increase in total travel time. In order for the traffic to consider the alternative route, the predicted travel time for route 1 should be at least as high as the – originally longer – travel time for route 2. If this is indeed the case, i.e., if the average speed on route 1 is reduced to 60 km/h, the flow on route 1 is highly suboptimal, yielding a throughput for the OD-pair of

$$\zeta = d^{OD} \cdot (6/60) \cdot t_{\mathrm{tot}}^{-1} \text{ [veh/h]}$$

with t_{tot} [h] the time window or horizon during which the traffic flow is considered. Note that for simplicity we assumed that traffic splits evenly over the two alternative routes of similar perceived travel time. For the likelihood of changing the driver's preferred route based on a difference in travel time, see e.g., [12]. By indicating predicted travel times on the DRIP instead of the current, instantaneous times, the drivers will deviate to the alternative road earlier, leading to a quicker convergence to the equilibrium assignment.

If instead drivers are assigned to one of the alternative routes before the traffic conditions on the main route lead to a suboptimal travel time, the system-optimal distribution with total throughput

$$\begin{aligned}\zeta &= (d^1 \cdot \zeta^1 + d^2 \cdot \zeta^2) \cdot t_{\mathrm{tot}}^{-1} \\ &= \big((3/60) q_{\mathrm{crit}}^1 + (6/60)(d^{OD} - q_{\mathrm{crit}}^1)\big) \cdot t_{\mathrm{tot}}^{-1} \text{ [veh/h]}\end{aligned}$$

could be reached. A monetary incentive can accomplish such a distribution (d^1, d^2) as will be shown in Section IV below.

IV. The Reverse Stackelberg Routing Approach

The problem of reaching a maximal throughput of traffic over a network is modeled by means of a reverse Stackelberg game approach. The scheme of Fig. 2 illustrates the process that leads to a dynamic route assignment, using a leader-follower approach where the road authority associates monetary incentives θ with the driver's choice of the time τ in which he desires to reach his destination. Based on the drivers', i.e., follower's choice of the pair (θ, τ) according to the relation $\theta = \gamma_{\mathrm{L}}(\tau)$ proposed by the leader, the drivers are assigned to a route. The leader's aim is therefore to compose a leader function γ_{L} such that a system-optimal distribution of traffic can be achieved. The characteristics of the drivers as well as the incentive functions and travel time choice of the drivers are communicated via an on-board computer.

Before the overall approach is elaborated upon, first the basic elements of the reverse Stackelberg game are translated to the traffic domain. Here it should be noted that a receding horizon approach is adopted, where k_{c} indicates the time instant $t = k_{\mathrm{c}} T_{\mathrm{c}}$, with T_{c} the sample or control time step of the dynamic routing approach. Similarly, k indicates the time instant $t = k T_{\mathrm{s}}$, with T_{s} the time step for the simulation of the traffic behavior based on a prediction model that will be described in Section IV-B.1. Finally, $T_{\mathrm{c}} = M T_{\mathrm{s}}, M \in \mathbb{N}$.

A. Basic Elements of the Reverse Stackelberg Approach

1) The Players and Their Decision Variables:

- The single **leader player** represents the road authority responsible for accomplishing an optimal use of a given traffic network.

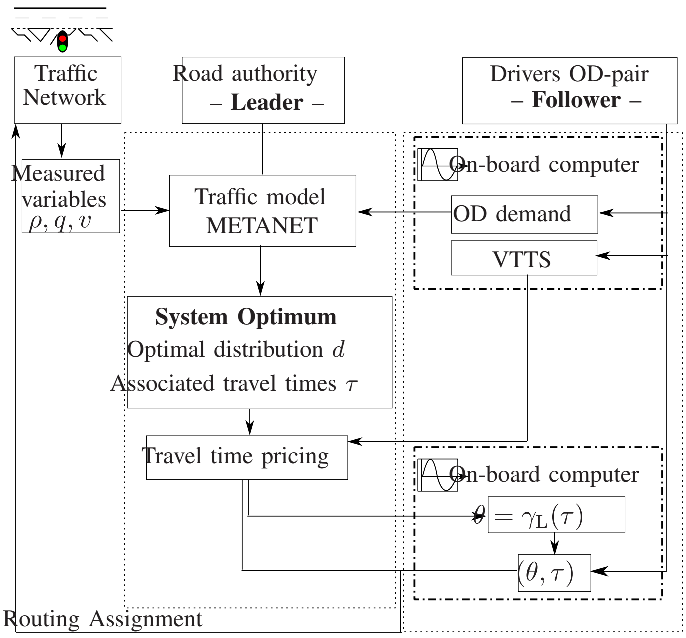


Fig. 2: Schematic framework for the reverse Stackelberg approach to dynamic route assignment.

- A **follower player** represents a homogeneous group of vehicles that desire to travel according to a certain origin-destination (OD) pair $(o,d)\in\mathcal{O}\times\mathcal{D}$, with $\mathcal{O},\mathcal{D}$ the set of origin and destination nodes, respectively. The total number of OD-pairs is denoted by $N_{\mathrm{OD}}=|\mathcal{O}|\cdot|\mathcal{D}|$, with $|X|$ the cardinality of X. The number of alternative routes, i.e., paths without cycles, for the origin destination-pair $i\in\{1,...,N_{\mathrm{OD}}\}$ is denoted by $n_i\in\mathbb{N}$. Alternatively, each individual driver could be treated as a separate follower player. However, this involves a more computationally demanding configuration, while it is intuitively more clear to characterize a follower player as a group of drivers with the same properties.
  Further, the group of drivers should be homogeneous in the sense that they have a similar monetary value of time, as will be elaborated upon in Section IV-A.2 below. We denote the value-of-time-class by $h\in\{1,\ldots,H\}$, with $\mathcal{H}_i$ the set of classes of drivers with a particular destination $i\in\{1,\ldots,N_{\mathrm{OD}}\}$.
  It should be noted that in the reverse Stackelberg approach, the road authority will want to distribute drivers over different available routes. Therefore we also add the property of a follower group to represent a specific route $j\in\{1,\ldots,n_i\}$ which the leader desires the drivers of the group to take. The leader can accomplish this by presenting a different leader function γ_{L} to n_i fractions of the drivers within value-of-time-classes $h\in\{1,\ldots,H\}$ and OD-pair $i\in\{1,\ldots,N_{\mathrm{OD}}\}$, as explained in Section IV-B.3 below.
  The total number of follower players is denoted by $N_{\mathrm{F}}=\sum_{i=1}^{N_{\mathrm{OD}}} n_i\cdot|\mathcal{H}_i|$ where we can represent the set of followers by

$$\mathcal{F}=\{(h,i,j)|i\in\{1,\ldots N_{\mathrm{OD}}\},j\in\{1,\ldots,n_i\},h\in\mathcal{H}_i\}.$$

Further, the **decision variables** represent respectively:

- A monetary incentive $\theta^{hij}(k_{\mathrm{c}})\in\Omega_{\mathrm{L}}^{hij}$ in \$ to be paid by or received by the follower player $(hij)\in\mathcal{F}$ where $\Omega_{\mathrm{L}}^{hij}:=[\theta_{\min}^{hij},\theta_{\max}^{hij}]$ denotes the range of monetary incentives that is accepted by the drivers.
- A choice of desired travel time $\tau^{hij}(k_{\mathrm{c}})\in\Omega_{\mathrm{F}}^{hij}$ in minutes for follower player $hij\in\{1,\ldots,N_{\mathrm{F}}\}$ to reach the associated destination d, where $\Omega_{\mathrm{F}}^{hij}(k_{\mathrm{c}}):=[\tau_{\min}^{hij}(k_{\mathrm{c}}),\tau_{\max}^{hij}(k_{\mathrm{c}})]$ denotes the range of possible, i.e., realizable travel times for a specific OD-pair at the current time step k_{c}, that is provided by the leader.

2) Leader and Follower Objective Functions:

- The leader player aims to maximize traffic throughput:

$$\mathcal{J}_{\mathrm{L}}(k_{\mathrm{c}})=\sum_{i=1}^{N_{\mathrm{OD}}}\sum_{h\in\mathcal{H}_i}\sum_{j=1}^{n_i}\left(d^{hij}(k_{\mathrm{c}})\cdot\tau^{hij}(k_{\mathrm{c}})\right)t_{\mathrm{tot}}^{-1},\quad(1)$$

  where $d^{hij}(k_{\mathrm{c}})\in\mathbb{R}_+$ denotes the part of the total demand $d^i(k_{\mathrm{c}})$ [veh/h] for the OD-pair $i\in\{1,\ldots,N_{\mathrm{OD}}\}$ that involves the vehicles from one of the $h\in\{1,..,H\}$ classes of drivers with a certain monetary value of time, that is distributed over route $j\in\{1,\ldots,n_i\}$. Recall that $\mathcal{H}_i$ denotes the set of value-of-time classes that apply to vehicles for the particular OD-pair $i\in\{1,\ldots,N_{\mathrm{OD}}\}$. In case the class $\mathcal{H}$ is not relevant in the context, superscript h will be disregarded, i.e., $d^{ij}:=\sum_{h\in\mathcal{H}_i}d^{hij}$.
  Instead of optimizing $\mathcal{J}_{\mathrm{L}}$ at each control time step, the leader can also consider a time horizon, as will be described in the dynamic framework in Section IV-B.2.
- The followers' aim is to minimize the travel cost as a function of monetary incentives and travel time:

$$\mathcal{J}_{\mathrm{F}}^{hij}(k_{\mathrm{c}})=\alpha_{\mathrm{F}}^{h}\cdot\tau^{hij}(k_{\mathrm{c}})+\theta^{hij}(k_{\mathrm{c}}),\qquad(2)$$

  with $\alpha_{\mathrm{F}}^{h}\in\mathbb{R}_+$ the possibly time-variant monetary value of travel time also known as value of travel time savings (VTTS).

Remark 1: Here, we assume a linear mapping of travel time to monetary value as is often adopted in the literature, e.g., in [9], [10]. However, (2) could be replaced by a more involved, nonlinear relation as considered in e.g., [13], [14]. The only consequence of a different follower objective function is in the type of leader function γ_{L} that is needed to reach the optimal distribution of tolls to arrive at the system optimum, as will be elaborated upon in Section IV-B.3.

Note that the monetary value of time of the driver classes could be differentiated between given a particular car type, or it could be determined by an iterative learning process in which the monetary value of time is adapted over time based on the choice (τ,θ) of the particular driver.

B. The Dynamic Game Framework

The dynamic reverse Stackelberg routing approach consists of the following main steps:

- Given the current traffic state and the demand for the OD-pairs as indicated by the drivers, a **system-optimal distribution** of the new vehicles over the available

routes is computed, together with the corresponding predicted mean travel times.
- Given the desired distribution of vehicles and the according travel times to the respective destinations, **leader functions** $\gamma_{\mathrm{L}}^{hij}(k_{\mathrm{c}}) : \Omega_{\mathrm{F}} \rightarrow \Omega_{\mathrm{L}}$ for each of the $(hij) \in \mathcal{F}$ followers are computed. Thus, in order for the leader to achieve an optimal distribution $d^{ij}(k_{\mathrm{c}})$ with associated travel times $\tau^{ij}(k_{\mathrm{c}})$, a specific fraction of the drivers for the i-th OD-pair and with a monetary value of time-class $h \in \mathcal{H}_i$ will be associated with one of the $j \in \{1, \ldots, n_i\}$ routes.
- As a response to the optimal leader functions, the follower will choose a combination of monetary incentive and travel time, which the leader associates with a certain route that the follower is obliged to follow.

1) The Prediction Model: In order to capture the behavior of traffic over time, a prediction model is used to track the traffic states, i.e., to analyze the impact of a traffic routing assignment. Based on the current and future state of the network and future demand, the desired traffic flow distribution can be determined and the associated decision variables that can lead to this state.

Here, the routes are further divided into homogeneous freeway stretches or road links $r_1, \ldots, r_{n_{ij}}, r_l \in \mathcal{R}_l, l \in \{1, \ldots, n_{ij}\}$, $i \in \{1, \ldots, N_{\mathrm{OD}}\}, j \in \{1, \ldots, n_i\}$ – potentially overlapping with alternative routes of the same or another OD-pair – of length L^l [km], capacity $q_{\max}^l$ [veh/km], and maximum speed $v_{\max}^l$ [km/h]. Here, $\mathcal{R}_l$ denotes the set of links present in route l, where $\mathcal{R}$ represents the set of all links in the network. By $\mathcal{N}$ we denote the set of internal nodes that connect the links.

For each link $r \in \mathcal{R}$ the flow of vehicles at time step k traveling towards destination d is denoted by $q_{r,d}(k)$. Then, the total inflow $Q_{n,d}(k)$ of node $n \in \mathcal{N}$ with a destination d can be computed by

$$Q_{n,d}(k) = \sum_{r \in I(n)} q_{r,d}(k),$$

with $I(n)$ the set of incoming links for node n. Similarly, the set of outgoing roads is denoted $O(n)$, with the traffic flow represented by

$$q_{r,d}(k) = \beta_{n,r,d}(k) Q_{n,d}(k)$$

where $\beta_{n,r,d}(k)$ denotes the splitting or turning rate for link r at node n with destination d. The total flow on link r can be summed

$$q_r(k) = \sum_{d \in \mathcal{D}} q_{r,d}(k).$$

In particular, we chose to adopt the macroscopic METANET traffic flow model [15], [16]; the reader is referred to the literature for details. In order to be able to accurately model the traffic behavior, the road links $r \in \mathcal{R}$ that are part of the possible routes of one or more OD-pairs are further divided into N_r segments of equal length L_r, typically between 300-1000 m. Further, the number of lanes in link r is denoted by λ_r. The state of the traffic network is now described by the following macroscopic variables, where the evolution is described by the respective update equations:

- average traffic density $\rho_{r,i}(k)$ [veh/km/lane] in segment i of link r at time $t = kT_{\mathrm{s}}$;
- mean speed $v_{r,i}(k)$ [km/h] of vehicles in segment i of link r at time $t = kT_{\mathrm{s}}$;
- traffic flow $q_{r,i}(k)$ [veh/h] leaving segment i of link r in time interval $[kT_{\mathrm{s}}, (k+1)T_{\mathrm{s}}]$.

Note that for the static game, it suffices to make use of a static prediction model that is based on the fundamental diagram of traffic flow, which represents the equilibrium relation between speed and density [17].

2) Optimal Distribution of Traffic Flow: Using the prediction model, given a driver demand pattern $d_{\mathrm{in}}^{hi}(k) \,\forall\, i \in \{1, \ldots, N_{\mathrm{OD}}\}, h \in \mathcal{H}_i$ in veh/h, a system-optimal – with respect to the throughput – distribution of traffic flow $d^{ij}(k)$ over the road network can be computed, as well as the associated mean travel times $\tau^{ij}(k)$ for the routes $j \in \{1, \ldots, n_i\}$.

This can be done by solving a dynamic version of the minimum cost flow problem, where the following constraints are needed to correctly represent the flows of the traffic network:

$$\sum_{h \in \mathcal{H}_{(o,d)}} d_{\mathrm{in}}^{h,(o,d)}(k) = q_{r_{\mathrm{out}}(o),d}(k) \;\forall (o,d) \in \mathcal{O} \times \mathcal{D} \tag{3}$$

$$\sum_{d \in \mathcal{D}} \sum_{r \in I(n)} q_{r,d}(k) = \sum_{d \in \mathcal{D}} \sum_{r \in O(n)} q_{r,d}(k) \;\forall n \in \mathcal{N} \tag{4}$$

$$\sum_{d \in \mathcal{D}} q_{r,d}(k) \leq q_{\mathrm{cap},r} \;\forall r \in \mathcal{R} \tag{5}$$

$$q_{r,d}(k) \geq 0 \;\forall r \in \mathcal{R}, \forall d \in \mathcal{D}, \tag{6}$$

where $r_{\mathrm{out}}(o), d$ denotes the single outgoing road segment to node o; in case of multiple outgoing roads, a virtual road can be created with zero length and travel time.

Finally, in order to optimize the throughput the following cost function is adopted at time step k:

$$\min_{q_{r,d}} \sum_{j=1}^{N_{\mathrm{p}}} \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} q_{r,d}(k+j) t_r(k+j) t_{\mathrm{tot}}^{-1}, \tag{7}$$

for $j = 1, \ldots, N_{\mathrm{p}}$ with N_{p} the prediction horizon that is incorporated to take into account not only the present but also the future traffic conditions. The desired distribution of traffic flow $d^{ij}(k)$ over the road network now follows straightforwardly from the optimal flows $q_{r,d}(k)$.

Note that the travel time for a particular road depends on the current density and flow, i.e., we assume the average travel time $t_r(k)$ of road $r \in \mathcal{R}$ to be $L_r / \tilde{v}_r(k)$, with $\tilde{v}_r(k)$ the desired speed according to the METANET model or according to the fundamental diagram expression $\tilde{v}_r(k) = v_{\mathrm{free},r} \exp[-\frac{1}{a_r}(\frac{\rho_r(k)}{\rho_{\mathrm{cr},r}})^{a_r}]$ with $v_{\mathrm{free},r}$ the free-flow speed, $\rho_{\mathrm{cr},r}$ the critical density, and a_r a model parameter. Alternatively, one could assume the average speed to be (I) the maximum or free-flow speed that applies during uncongested traffic conditions, or to be (II) the mean speed that is currently measured, which leads to a simplified linear programming problem (3)-(7) in both cases.

3) Optimal Distribution of Monetary Incentives: Given the desired traffic flow distribution d^{hij} at the control time step k_c, the leader can associate with each feasible travel time a fee or incentive $\theta^{hij}(k_\mathrm{c})$ such that the optimal response of the follower $(hij) \in \mathcal{F}$ coincides with the desired travel time of the respective i-th OD-pair and j-th route of the follower, $\tau_\mathrm{d}^{hij}(k_\mathrm{c})$. The optimal leader functions are those that cause the followers to choose the associated desired travel time $\tau_\mathrm{d}^{hij}(k_\mathrm{c})$ that has the same value for the followers of each class $\mathcal{H}_i$.

In Fig. 3, several level curves are plotted of the travel cost function (2) for $\alpha_\mathrm{F} = 0.5$, i.e., each point on the straight lines results in a same objective function value for the follower. The optimal response of the follower to one of the two leader functions $\gamma_\mathrm{L}^1(\tau^1), \gamma_\mathrm{L}^2(\tau^2)$ is respectively $\tau_\mathrm{d}^1, \tau_\mathrm{d}^2$ with the associated monetary incentives $\theta_\mathrm{d}^1, \theta_\mathrm{d}^2$.

In order to achieve the desired distribution of travel times $\tau^{ij}(k_\mathrm{c})$ associated with the distribution of flow $d^{ij}(k_\mathrm{c})$ over the n_i alternative routes for OD-pair $i \in \{1, \dots, N_\mathrm{OD}\}$, multiple, i.e., N_F leader functions are necessary that reach the associated monetary incentives for each of the $\mathcal{H}_i$ classes of drivers, i.e., $\theta^{hij}(k_\mathrm{c}) \forall (hij) \in \mathcal{F}$. Hence, it may occur that due to the different leader functions $\gamma_\mathrm{L}^{hij}(k_\mathrm{c})$ posed to drivers with a different value of time, a different monetary value is set on the same travel time.

Remark 2: As it could be deemed unfair to associate different monetary incentives with the same travel time value, one can prevent this by adopting a single leader function that leads to the same optimal response for all followers with the same OD-pair and associated route. Due to the similar – affine – shapes of the level curves of (2) for all classes $h \in \mathcal{H}$, it is indeed possible to adopt a function $\gamma_\mathrm{L}^{ij}(k_\mathrm{c})$ that leads to the same follower behavior. Moreover, if the classes of drivers with a homogeneous value of traffic time are categorized by the type of vehicle (personal car, bus, motorcycle), such a distinction in monetary incentive may be accepted.

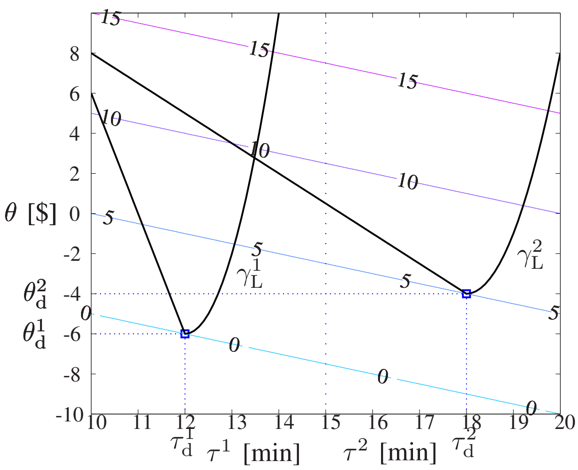


Fig. 3: Optimal leader functions: An example.

Finally, in this context it should be noted that the leader may accomplish any desired amount $\theta(k_\mathrm{c}) \in \Omega_\mathrm{L}$ by constructing the leader function $\gamma_\mathrm{L}(k_\mathrm{c})$ in such a way that the optimal follower response $\tau_\mathrm{d}(k_\mathrm{c})$ is associated with the desired toll $\theta_\mathrm{d}(k_\mathrm{c})$. Thus, the vertical axis in Fig. 3 can be shifted for both leader functions separately such that for both τ_d^1 and τ_d^2, e.g., $\theta_\mathrm{d}^1 = \theta_\mathrm{d}^2 = 0$ is the associated toll.

4) Taking Care of Deviations: The proposed approach to dynamic routing relies on several models, i.e., the traffic model, an approximation of the drivers' monetary value of travel time, and the assumption that the drivers respond in a fully rational manner to the proposed incentive function. As a result, several inaccuracies and possible unpredicted behavior should be taken into account, e.g.:

- *Inaccurate estimate of the monetary value of travel time.* In order to make drivers conscious of their perceived value of time, they may be asked to indicate this value through the on-board computers.
- *Deviation of drivers from the optimal response* (θ, τ). In case of a leader function that is symmetric around the desired pair (θ, τ), we can assume the deviations from the rational follower response to γ_L to be random and independent and identically distributed, spreading symmetrically around the optimum. Further, an adapted speed measure can be taken to speed up or slow down the follower group. Next to a suboptimal response of the drivers to the leader function, one can use such speed measure or another traffic control measure to take care of a misestimation in the required travel time for a certain route.
- *Deviation of drivers from the imposed route.* In order to induce the drivers to follow the particular route as is instructed by the leader, a posterior monetary penalty term can be implemented:

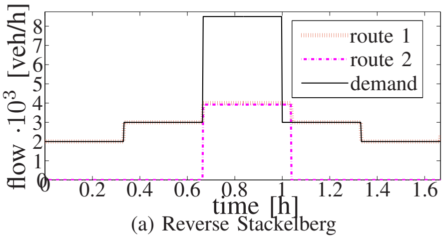
$$\theta_\mathrm{pen}^{hij} := \alpha_\mathrm{pen} |\tau^{hij}(k_\mathrm{c}) - \tau_\mathrm{dev}^{hij}| + c_\mathrm{pen},$$

  with τ_dev^{hij}, $(hij) \in \mathcal{F}$ the travel time resulting from the deviated route. Here, $\alpha_\mathrm{pen} \in \mathbb{R}_+$ is a weighting factor and $c_\mathrm{pen} \in \mathbb{R}_+$ represents a constant monetary penalty that applies in case of route deviation. Note that this involves the implementation of the game, not the game set-up itself.

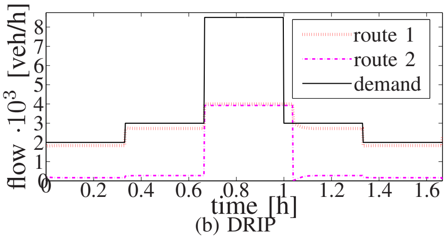
V. Case Study

A. Set-up

In order to illustrate the performance of the reverse Stackelberg approach as compared to route guidance methods based on travel time information or road tolls, we adopt the simple network of two non-overlapping routes and a single OD-pair as depicted in Fig. 1. Recall that a three-lane 3 km freeway splits into a two-lane freeway of 10 km and a one-lane freeway of 6 km. These routes join again in a three-way freeway of 2 km length. For the sake of simplicity, we only consider the two alternative freeways, both with a capacity $q_\mathrm{crit} = 4000\,\mathrm{veh/h}$.



(a) Reverse Stackelberg



(b) DRIP

Fig. 4: A comparison of traffic assignment.

We simulated the traffic behavior for 100 min for a demand scenario with three levels of demand as depicted in Fig. 4, i.e., where the highest demand results in congestion on the shortest route. Here we adopted the METANET model with the parameter values taken from [16]. Finally, the distribution of demand when using the DRIP approach is determined by the logit model [18], in which the difference in the predicted total travel time between the routes influences the split rate.

B. Results

The distribution of demand over the two routes in receding horizon is shown in Fig. 4. While in the reverse Stackelberg case the traffic is allocated to the shortest route as long as the flow remains below the capacity, in the routing based on DRIP, a part of the flow splits to the longer route, which influences the average travel time. The methods lead to a total throughput of $1.2466 \cdot 10^5$ versus $6.4077 \cdot 10^4$ veh/h for the reverse Stackelberg and the DRIP approach respectively.

Basically, this simple case-study shows that route guidance based on a difference in travel time does not yield the system-optimal throughput, which can be achieved by using the suggested reverse Stackelberg approach. This benefit should become more clear when a larger network is considered with several overlapping routes, where the optimal throughput may not be realized by influencing the split-rates at the individual nodes.

VI. Discussion and Future Work

From the traffic guidance point of view, methods like dynamic route guidance information panels are not able to make optimal use of the available routes in traffic networks. While monetary incentives are recognized as a tool to accomplish traffic states that are closer to the system optimum with respect to total travel time [9], [10], the current methods of time-variant and flow-dependent road tolls do not in general achieve the optimal distribution either. However, with the advancement of intelligent vehicles equipped with on-board computers, a reverse Stackelberg approach can be adopted to make optimal use of the traffic network while satisfying the users' individual preferences. Here, the road authority can communicate monetary incentives associated with certain travel times to the driver, who – upon the choice of a travel time and monetary value – is presented with a certain route to the desired destination. The proposed approach can be extended e.g., by considering also the separate lanes. Finally, a more elaborate case study will be considered, including an analysis of computational efficiency and a comparison of the method not only to the information-based DRIP approach but also to methods in which time-varying road tolls are adopted.

References

[1] S. Peeta and A. K. Ziliaskopoulos, "Foundations of dynamic traffic assignment: The past, the present and the future," *Networks and Spatial Economics*, vol. 1, no. 3–4, pp. 233–265, 2001.

[2] J. G. Wardrop, "Some theoretical aspects of road traffic research," in *Proceedings of the Institute of Civil Engineers, PART II*, vol. 1, no. 2, 1952, pp. 325–378.

[3] F. P. Deflorio, "Evaluation of a reactive dynamic route guidance strategy," *Transportation Research Part C*, vol. 11, no. 5, pp. 375–388, 2003.

[4] A. Karimi, A. Hegyi, B. De Schutter, J. Hellendoorn, and F. Middelham, "Integrated model predictive control of dynamic route guidance information systems and ramp metering," in *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*, Washington, DC, USA, 2004, pp. 491–496.

[5] M. van den Berg, T. Bellemans, B. De Schutter, B. De Moor, and J. Hellendoorn, "Anticipative ramp metering control using dynamic traffic assignment," in *Proceedings of the 7th International IEEE Conference on Intelligent Transportaion Systems*, Washtingon, DC, USA, 2004, pp. 503–508.

[6] Y.-C. Ho, P. B. Luh, and R. Muralidharan, "Information structure, Stackelberg games, and incentive controllability," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 454–460, 1981.

[7] Y.-C. Ho, P. B. Luh, and G. J. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, no. 2, pp. 167–179, 1982.

[8] G. J. Olsder, "Phenomena in inverse Stackelberg games, part 1: Static problems," *Journal of Optimization Theory and Applications*, vol. 143, no. 3, pp. 589–600, 2009.

[9] D. Joksimovic, M. C. J. Bliemer, P. H. L. Bovy, and Z. Verwater-Lukszo, "Dynamic road pricing for optimizing network performance with heterogeneous users," in *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, Tucson, AZ, USA, 2005, pp. 407–412.

[10] K. Staňková, G. J. Olsder, and M. C. J. Bliemer, "Comparison of different toll policies in the dynamic second-best optimal toll design problem: Case study on a three-link network," *European Journal of Transport and Infrastructure Research*, vol. 9, no. 4, pp. 331–346, 2009.

[11] H. von Stackelberg, *The Theory of Market Economy*. Oxford, UK: Oxford University Press, 1952.

[12] R. B. Dial, "A probabilistic multipath traffic assignment model which obviates path enumeration," *Transportation Research*, vol. 5, no. 2, pp. 83–111, 1971.

[13] A. J. DeSerpa, "Microeconomic theory and the valuation of traveltime: some clarifications," *Regional Urban Economics*, vol. 2, pp. 401–410, 1973.

[14] T. Blayac and A. Causse, "Value of travel time: a theoretical legitimization of some nonlinear representative utility in discrete choice models," *Transportation Research Part B: Methodological*, vol. 35, no. 4, pp. 391–400, 2001.

[15] A. Messmer and M. Papageorgiou, "METANET: A macroscopic simulation program for motorway networks," *Traffic Engineering and Control*, vol. 31, no. 9, pp. 466–470, 1990.

[16] A. Kotsialos, M. Papageorgiou, and A. Messmer, "Optimal coordinated and integrated motorway network traffic control," in *Proceedings of the 14th International Symposium of Transportation and Traffic Theory (ISTTT)*, Jerusalem, Israel, Jul. 1999, pp. 621–644.

[17] A. May, *Traffic Flow Fundamentals*. Englewood Cliffs, New Jersey: Prentice-Hall, 1990.

[18] J. S. Cramer, *The Logit Model: An Introduction for Economists*. London, UK: Edward Arnold, 1991.