Optimizing Traffic Flow

Bernardo A. Huberman[+] and Dirk Helbing[*]
[+] *Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA*
[*] *II. Institute of Theoretical Physics, University of Stuttgart, Pfaffenwaldring 57/III, 70550 Stuttgart, Germany*

We present an economics-based method for deciding the optimal rates at which vehicles are allowed to enter a highway. The method exploits the naturally occuring fluctuations of traffic flow and is flexible enough to adapt in real time to the transient flow characteristics of road traffic. Simulations based on realistic parameter values show that this strategy is feasible for naturally occurring traffic, and that even far from optimality, injection policies can improve traffic flow. Our results also allow a better understanding of the high flows observed in "synchronized" congested traffic close to ramps.

47.62.+q,47.54.+r,02.70.Ns,89.40.+k

The advent of powerful traffic simulators [1–3] has led to a spate of new discoveries in the area of vehicular traffic that agree well with empirical observations [4]. From moving traffic jams [1] to transitions to states of coherent motion [3] these new results offer insights into a complicated and socially relevant many-body problem, while also suggesting ways of designing controls that can maximize the flow of vehicles in cities and highways [5,6].

The recent discovery of a new state of congested highway traffic, called "synchronized" traffic [7], has generated a strong interest in the rich spectrum of phenomena occuring close to on-ramps [8,9]. In this connection, a particularly relevant problem is that of choosing an optimal injection strategy of vehicles into the highway. Similar questions have recently been raised regarding the most efficient use of the Internet in light of its bursty congestion patterns [10]. While there exist a number of *heuristic* approaches to optimizing vehicle injection into freeways by on-ramp controls, the results are still not satisfactory. What is needed is a strategy that is flexible enough to adapt in real time to the transient flow characteristics of road traffic while leading to minimal travel times for all vehicles on the highway.

This paper presents a solution to this problem that explicitly exploits the naturally occuring fluctuations of traffic flow in order to enter the freeway at optimal times. This method, which leads to a more homogeneous traffic flow and a reduction of inefficient stop-and-go motions, is less sensitive to failures in the control mechanism than traditional approaches.

The basic performance criterion behind this optimization technique is the travel time distribution of vehicles [11] which is to be contrasted with the local velocity distribution applied in other approaches. The travel time distribution is a global measure of the overall dynamics on the whole freeway stretch. It allows the evaluation of both the expected arrival time of vehicles at a destination and its variance, the latter characterizing the likelihood that a particular vehicle will have a travel time different from the expected one. Both the average and the variance of travel times are influenced by the inflow of vehicles entering the freeway over an on-ramp. From these two quantities one can construct a relation between the average payoff (the negative mean value of travel times) and the risk (their variance), as is considered in many optimization problems in economics. The optimal strategy will then correpond to the point in the curve that yields the lowest risk at a high average payoff. In the following, we will show that the variance of travel times has a minimum for on-ramp flows that are different from zero, implying that traffic flow can be optimized by choosing the appropriate vehicle injection rate into the freeway.

In order to obtain the travel time distribution of vehicles on a highway, we simulated two-lane traffic flow via a discretized follow-the leader model, which describes the empirical known features of traffic flows quite well [2]. In our experiments, we extended the simulation to several lanes with lane-changing maneuvers and different vehicle types (cars and trucks). We determined the travel times of all vehicles by storing the times at which they pass two successive cross sections of the road. Empirical data of this kind could be collected at the gates of a toll road. The time difference between the passing times corresponds to the travel time needed to traverse the freeway stretch between the two cross sections. In our simulations, we chose a stretch of length $L = 10\,\mathrm{km}$.

The model distinguishes I neighboring lanes $i \in \{1,\dots,I\}$ of a unidirectional freeway. All lanes are subdivided into sites $z \in \{1,2,\dots,Z\}$ of equal length $\Delta x = 2.5\,\mathrm{m}$. Each site is either empty or occupied, the latter case representing the back of a vehicle of type $a \in \{1,\dots,A\}$ with velocity $v = u\Delta x/\Delta t$. Here, $u \in \{0,1,\dots,u_a^{\max}\}$ is the number of sites that the vehicle moves per update step Δt. We have distinguished cars ($a = 1$) and trucks ($a = 2$). These are characterized by different optimal velocities $U_a(d)$ with which the vehicles would like to drive at a distance d to the vehicle in front. Their lengths l_a correspond to the maximum distances satisfying $U_a(l_a) = 0$. The positions $z(T)$, velocities $u(T)$, and lanes $i(T)$ of all vehicles are updated [12] every time step $\Delta t = 1\,\mathrm{s}$ at times $T \in \{1,2,\dots\}$ according to the following successive steps [3]:

1. *Determine the potential velocities* $u_j(T+1)$ on the present and the neighboring lanes $i(T)+j$ with $j \in$

$\{-1, 0, +1\}$ according to the *acceleration law*

$$u_j(T+1) = \lfloor \lambda U_a\big(d_j(T)\big) + (1-\lambda)u(T) \rfloor . \quad (1)$$

Here, the floor function $\lfloor x \rfloor$ is defined by the largest integer $m \le x$, and $d_j(T) = [z_j^+(T) - z(T)]$ denotes the distance to the next vehicle ahead $(+)$ on lane $i(T)+j$ at position $z_j^+(T)$. The above equation describes the typical follow-the-leader behavior of driver-vehicle units. Delayed by the reaction time Δt, they tend to move with the distance-dependent optimal (safe) velocity U_a, but the adaptation takes a certain time $\tau = \lambda\,\Delta t$ because of the vehicle's inertia. Good results, in the sense of replicating highway data, are obtained for $\lambda = 0.77$.

2. *Change lane* to the left, i.e. set

$$i(T+1) = i(T) + k \quad (2)$$

with $k = +1$, if the vehicle considered can go fastest there, i.e. if

$$u_{+1}(T+1) > u_0(T+1) \quad \text{and} \quad > u_{-1}(T+1) . \quad (3)$$

This is usually the case if the headway on the left lane is greatest. Apart from the validity of this *incentive criterion,* we demand two extra *safety criteria* [13]: First, the current vehicle position should be ahead of the expected position $z_{+1}^-(T+1)$ of the following vehicle $(-)$ on the left lane, i.e.

$$z(T) > z_{+1}^-(T+1) . \quad (4)$$

Second, the potential velocity on the left lane should not be considerably less than the expected velocity $u_{+1}^-(T+1)$ of the following vehicle, i.e.

$$u_k(T+1) \ge q\, u_{+1}^-(T+1) . \quad (5)$$

Once again, realistic results are obtained for $q = 0.7$. A value $q < 1$ implies that drivers are ready to accept a braking maneuver of the follower on the destination lane at the next update step. Therefore, the values of q are a measure of how relentless drivers are in overtaking.

Assuming symmetrical ("American") lane changing rules for simplicity, a change to the right lane $(k = -1)$ is carried out, if the incentive criterion $u_{-1}(T+1) > u_0(T+1)$ and $\ge u_{+1}(T+1)$ as well as the safety criteria $z(T) > z_{-1}^-(T+1)$ and $u_k(T+1) \ge q\, u_{-1}^-(T+1)$ are fulfilled. Otherwise the vehicle stays on the same lane $(k = 0)$.

3. If the potential velocity $u_k(T+1)$ on the new lane $i(T+1)$ is positive, diminish it by 1 with probability $p = 0.001$, which accounts for *delayed adaptation* due to reduced attention of the driver and the variation of vehicle velocities:

$$u(T+1) = u_k(T+1) - \begin{cases} 1 & \text{with probability } p \\ & \text{if } u_k(T+1) > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The resulting value defines the updated velocity $u(T+1)$.

4. Update the vehicle position according to the *equation of motion*

$$z(T+1) = z(T) + u(T+1) . \quad (7)$$

Our simulations were carried out for a circular two-lane road. After the overall density was selected, vehicles were homogeneously distributed over the road at the beginning, with the same densities on both lanes. The experiments started with uniform distances among the vehicles and their associated desired velocities. The vehicle type was determined randomly after specifying the percentages of cars (90%) and trucks (10%).

At a given injection rate, vehicles enter the beginning of an on-ramp lane (i.e. a third lane) of 1 kilometer length with a uniform time headway and their optimal velocity related to the density in the destination lane. Our simulation stretch consists of a 1 km long 3-lane part and the $L = 10$ km long two-lane stretch on which the travel times are measured. No vehicle is allowed to change to the on-ramp from the main road. Injected vehicles try to change from the on-ramp to the main road as fast as possible, i.e. according to the above lane-changing rules, but they do not care about the incentive criterion. The end of the on-ramp is treated like a resting vehicle, i.e. any vehicle that approaches it has to stop, but it changes to the destination lane as soon as it finds a sufficiently large gap. If the on-ramp is completely occupied by vehicles waiting to enter the main road, the injected vehicles form a queue and enter the on-ramp as soon as possible. Injected vehicles that have completed their 10 km trip on the two-lane measurement stretch are automatically removed from the freeway (which would correspond to uncongested off-ramps adjacent to the lanes). Since our evaluations started after a transient period of two hours and continued until the 1000th injected vehicle finished its trip, the results are largely independent of the initial conditions.

If we plot the average of travel times as a function of their standard deviation (Fig. 1), we obtain curves parametrized by the injection rate of vehicles into the road. Several points are worth noting: 1. With growing injection rate $Q_{\text{rmp}} = 1/2^n$s $(n \in \{2, 3, \ldots, 10\})$, the travel time increases monotonically. This is because of the increased density caused by injection of vehicles into the freeway. 2. The average travel time of *injected* vehicles is higher, but their standard deviation lower than for the vehicles circling on the main road. This is due to the fact that vehicle injection produces a higher density on the lane adjacent to the on-ramp, which leads to smaller velocities. The difference between the travel time distributions of injected vehicles and those on the main road decreases with the length L of the simulated road, since lane-changes tend to equilibrate densities between lanes.

In addition, the standard deviation of travel times has a *minimum* for *finite* injection rates, as entering vehicles tend to fill existing gaps and thus homogenize traffic flow. This minimum is optimal in the sense that there is no other value of the injection rate that can produce travel with smaller variance. In particular, gap-filling behavior mitigates inefficient stop-and-go traffic at medium densities. Above a density of 45 vehicles per kilometer and lane on the main road (without injection), the minimum of the travel times' standard deviation occurs for $n = 6$. The reduction of the average travel time by smaller injection rates is less than the increase of their standard deviation. This result suggests that, in order to generate predictable and reliable arrival times, one should operate traffic at medium injection rates. Lastly, for the case of 40 vehicles per kilometer and lane, the minimum of the standard deviation of travel times is located at $n = 5$, while for 35 vehicles per kilometer and lane, it is at $n = 4$. Below 30 vehicles per kilometer and lane, vehicle injection does not reduce the standard deviation of travel times. This is because at these densities homogeneous traffic is stable anyway, so no stop-and-go traffic exists and therefore no large gaps that can be filled [3].

The curves displayed in Fig. 1 correspond to a given density $\rho_{\rm main}$ on the main road *without* injection of vehicles. The effective density $\rho_{\rm eff}$ on the freeway *resulting* from the injection of vehicles can be approximated by

$$\rho_{\rm eff} = \rho_{\rm main} + \frac{N_{\rm inj}}{IL}\,, \qquad (8)$$

where $I = 2$, $L = 10$ km. $N_{\rm inj}$ is the average number of injected vehicles present on the main road and can be written as

$$N_{\rm inj} = N_{\rm tot}\frac{\mathcal{T}_{\rm inj}}{\mathcal{T}_{\rm tot} - \mathcal{T}_{\rm inj}}\,, \qquad (9)$$

where $N_{\rm tot} = 1000$ is the total number of injected vehicles during the simulation runs, $\mathcal{T}_{\rm inj}$ is their average travel time, and $\mathcal{T}_{\rm tot}$ the time interval needed by all $N_{\rm tot} = 1000$ vehicles to complete their trip. We point out that, in addition to these measurements, we used two other methods of density measurement which yielded similar results.

We also investigated the dependence of the travel time characteristics on the resulting *effective* densities of vehicles. As Figure 2 shows, vehicle injection can actually reduce the average travel times of the vehicles on the main road, while the travel times of injected cars are about the same as those of vehicles on the main road without injection. This means that, for given $\rho_{\rm eff}$, one can actually increase the average velocity $V_{\rm main} = L/\mathcal{T}_{\rm main}$ of vehicles by injecting vehicles at a high rate without affecting their travel times. This result follows from the increased degree of homogeneity caused by entering vehicles that fill gaps on the main road, which mitigates the less efficient stop-and-go traffic. It might actually make sense to diverge a certain proportion of vehicles at intermittent times and to reinject them later following our methodology.

Finally, Figure 3 shows the average of the travel times for vehicles in the main road as a function of their standard deviation. In contrast to Fig. 1, the curves were computed for the resulting *effective* densities. This time, an increase of the injection rate (which corresponds to a smaller number of circling vehicles on the main road and a greater proportion of injected vehicles) *reduces* the average travel times. Once again, we observe a minimum of the standard deviation of travel times at high vehicle densities and medium injection rates.

In the limit of high injection rates, a traffic jam of maximum density $\rho_{\rm max}$ builds up at the end of the on-ramp, while downstream of it we find the typical density $\rho_{\rm out}$ related to the universal outflow $Q_{\rm out}$ from traffic jams [14,2]. We conjecture that the resulting structure consists of a block of density $\rho_{\rm max}$ and length L_1 containing $N_1 = \rho_{\rm max} L_1$ vehicles, and a block of density $\rho_{\rm out}$ of length $(L - L_1)$ containing $(N - N_1) = \rho_{\rm out}(L - L_1)$ vehicles at a mean density of $\rho_{\rm eff} = N/L$. The expected travel time $\mathcal{T}_{\rm main}$ would be

$$\begin{aligned}\mathcal{T}_{\rm main} &= \frac{L}{V_{\rm out}} + \frac{L_1}{C} + \mathcal{T}_{\rm acc} + \mathcal{T}_{\rm dec} \\ &= \frac{L}{V_{\rm out}} + \frac{\rho_{\rm eff} - \rho_{\rm out}}{\rho_{\rm max} - \rho_{\rm out}}\frac{L}{C} + \mathcal{T}_{\rm acc} + \mathcal{T}_{\rm dec}\,, \qquad (10)\end{aligned}$$

where $V_{\rm out}$ is the typical velocity emerging downstream of a traffic jam, and C is the universal dissolution velocity of traffic jams [14,2]. Notice that, for high injection rates, the average travel time should grow *linearly* with the mean density $\rho_{\rm eff}$, which is consistent with the results displayed in Fig. 2. For decreasing injection rates, travel times should increase, since the alternation of congested and free flow in the resulting stop-and-go traffic implies relevant acceleration times $\mathcal{T}_{\rm acc}$ and deceleration times $\mathcal{T}_{\rm dec}$ in total.

In conclusion, we have presented a strategy for optimizing traffic on highways in the sense of higher flows and more reliable predictions of individual travel times. The applied method is economics-based and resorts to the establishment of average payoff versus risk curves. Here, the average payoff corresponds to the negative mean value of travel times and the risk to their variance. The strategy exploits the naturally occuring fluctuations of traffic flow in order to allow the entry of new vehicles to the freeway at optimal times. Simulations based on realistic parameter values show that this strategy is feasible for naturally occurring traffic, and that even far from optimality, injection policies can improve traffic flow. The latter allows an interpretation of the high flows observed in "synchronized" congested traffic close to ramps [7], in contrast to the lower average flows in stop-and-go traffic.

Acknowledgments: D.H. wants to thank the DFG for financial support (Heisenberg scholarship He 2789/1-1).

[1] K. Nagel and M. Schreckenberg *J. Phys. I France* **2**, 2221 (1992).

[2] D. Helbing, Phys. Rev. Lett., submitted (1998).

[3] D. Helbing and B. A. Huberman, Nature, in print (1998).

[4] M. Schreckenberg and D. E. Wolf (eds.) *Traffic and Granular Flow '97* (Springer, Singapore, 1998).

[5] B. Faieta and B. A. Huberman, "Firefly: A synchronization strategy for urban traffic control" (Xerox PARC Internal Report, 1993).

[6] P. H. L. Bovy (ed.) *Motorway Traffic Flow Analysis.* (Delft University Press, Delft, 1998).

[7] B. S. Kerner and H. Rehborn, Phys. Rev. Lett. **79**, 4030 (1997) and Phys. Rev. E **53**, R4275 (1996).

[8] H. Y. Lee, H.-W. Lee, and D. Kim, Phys. Rev. Lett. **81**, 1130 (1998).

[9] D. Helbing and M. Treiber, Phys. Rev. Lett., in print (1998).

[10] B. A. Huberman, R. M. Lukose and T. Hogg, Science **275**, 51 (1997); R. M. Lukose and B. A. Huberman, Netnomics, in print (1998).

[11] K. Nagel and S. Rasmussen, in *Artificial Life IV*, edited by R. A. Brooks and P. Maes (MIT Press, Cambridge, MA, 1994).

[12] We applied sequential update in driving direction, which avoids conflicts of vehicles that like to change to the same lane and position from both neighboring lanes at the same time. For two-lane roads the sequential update leads to almost identical results as a parallel update, which is usually applied in cellular or lattice gas automata for flow simulations and gives realistic results.

[13] K. Nagel *et al.*, Phys. Rev. E **58**, 1425 (1998).

[14] B. S. Kerner and H. Rehborn, Phys. Rev. E **53**, R1297 (1996).

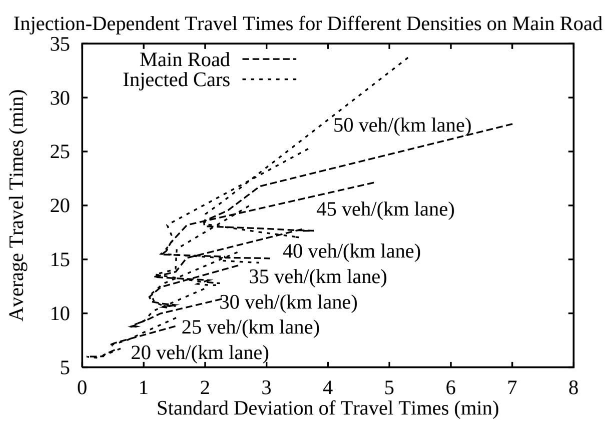

FIG. 1. Average and standard deviation of the travel times of vehicles on the main road and from the ramp as a function of the injection rate for various vehicle densities on the main road (measured without injection). With increasing injection rate, the average travel times are increasing due to the higher resulting vehicle density on the freeway.

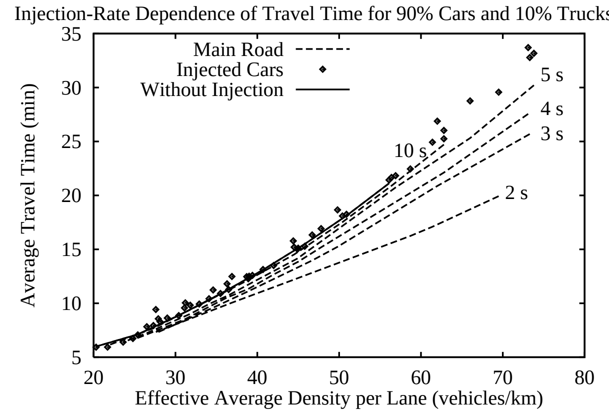

FIG. 2. Average travel times of vehicles on the main road as a function of the resulting effective vehicle density on the freeway, for various injection rates. In the limit of high injection rates, one observes the predicted linear dependence of average travel times on effective density, see Eq. (10). In contrast to the vehicles on the main road, the travel times of injected vehicles did not depend on the injection rate. However, when we checked what happens if the vehicles on the main road try to change to the left lane along the on-ramp in order to give way to entering vehicles, we found that both, injected vehicles and the vehicles on the main road, profited.

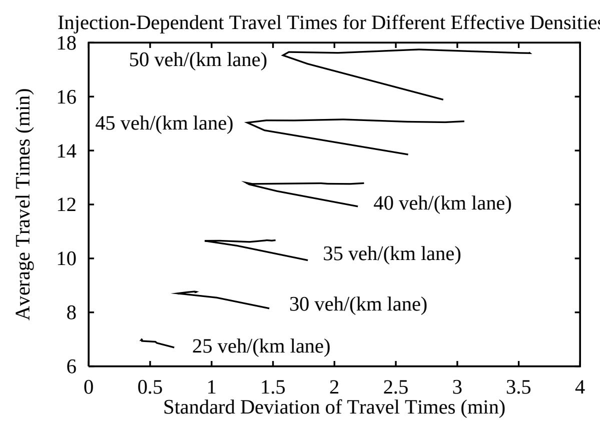

FIG. 3. As figure 1, but as a function of the resulting effective density on the freeway. We find shorter travel times at high injection rates because of the homogenization of traffic. The standard deviation of travel times is varying stronger than the average travel time, which indicates that medium injection rates are the optimal choice at high vehicle densities.