A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks

2003 J. Phys. A: Math. Gen. 36 L593

(http://iopscience.iop.org/0305-4470/36/46/L03)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) L593–L598 PII: S0305-4470(03)66060-9

LETTER TO THE EDITOR

A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks

Dirk Helbing

Institute of Economics and Traffic, Dresden University of Technology, Andreas-Schubert-Str 23, 01062 Dresden, Germany

E-mail: helbing@trafficforum.org

Received 1 July 2003, in final form 2 September 2003
Published 5 November 2003
Online at stacks.iop.org/JPhysA/36/L593

Abstract

While many classical traffic models treat the spatial extension of streets continuously or by discretization into cells of a certain length Δx, we will subdivide roads into comparatively long homogeneous road sections of constant capacity with an inhomogeneity at the end. The related model is simple and numerically efficient. It is inspired by models of dynamic queueing networks and takes into account essential features of traffic flows. Instead of treating single vehicles or velocity profiles, it focusses on flows at specific cross sections and average travel times of vehicles.

PACS numbers: 89.40.+k, 47.40.Nm, 47.55.−t

In the last decade, physicists have made many significant contributions to traffic modelling [1]. The modelling approaches reach from cellular automata [2] over microscopic car-following models [3] up to macroscopic or fluid-dynamic traffic models [4–6]. They have, for example, been the basis for understanding the various breakdown phenomena of traffic flow in terms of non-equilibrium phase transitions [7].

Classical traffic models are mostly based on the treatment of interacting vehicles, their statistical distribution, or their density and average velocity as a function of space and time. For numerical reasons, space and time are often discretized into small intervals Δx and Δt. Depending on the model, typical values of Δx are between 5 and 500 m. Although drivers are normally interested in travel times, most models focus on the velocities (from which the travel times are sometimes derived). An alternative approach is queueing models (see [8–11] and references therein): for example, Cremer and Landenfeld have developed a microscopic model for saturated urban road networks which is discrete in time. It distinguishes only vehicles moving with the maximum allowed speed $V_i^0 = 50\ \mathrm{km\ h^{-1}}$ and standing vehicles with a space requirement of 6.7 m, assuming discharge rates of $Q_{\mathrm{out}} = 1/2$ s and parallel forward motion of all queued vehicles when a site becomes empty [9]. The model by Eissfeldt *et al* eliminates vehicle motion and translates macroscopic flows into microscopic time headways. It is based on a section travel time and additional traffic-state-dependent waiting times of cars, taking into account the finite velocity of the upstream propagation of emptied road space [10]. It can reproduce empirical flows quite well [10]. In contrast to these microscopic queueing models, Kerner proposes a macroscopic approach to city traffic relating to the continuity equation [11]. It assumes a constant density of saturated (queued) traffic and a time-dependent, but space-independent density of unsaturated traffic. Therefore, it neglects the spatio-temporal propagation of density variations.

In the following, we will propose a model which exactly integrates the Lighthill–Whitham model over finite road sections. Although it formally eliminates the traffic dynamics on the road sections by expressing it through the dynamics at a few cross sections of the road network, it takes into account continuous (and discontinuous) changes in the density of free or congested traffic and their characteristic finite propagation velocities. The model was inspired by a dynamic queueing model of supply and production networks [12]. When we now specify road traffic as a queueing system, we will take into account essential traffic characteristics such as the flow-density relation or the properties of extended congestion patterns at bottlenecks. In fact, traffic congestion is usually triggered by spatial inhomogeneities of the road network [7], and queueing effects are normally not observed along sections of low capacity, but upstream of the beginning of a bottleneck. Therefore, we will subdivide roads into sections i of homogeneous capacity and length L_i, which start at place x_i and end with some kind of inhomogeneity (i.e. an increase or decrease of capacity) at place $x_{i+1} = x_i + L_i$. In other words, the end of a road section i is, for example, determined by the location of an on- or off-ramp, a change in the number of lanes, or the beginning or end of a gradient.

We will derive our model from the continuity equation

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial Q_i(x,t)}{\partial x} = \text{source terms} \tag{1}$$

describing the conservation of the number of vehicles. Here, $\rho(x,t)$ denotes the vehicle density per lane at place x and time t, and $Q_i(x,t)$ the traffic flow per lane. The source terms originate from ramp flows $Q_i^{\text{rmp}}(t)$, which enter the road at place x_{i+1}. Let us define the arrival rate at the upstream end of road section i by $Q_i^{\text{arr}}(t) = I_i Q_i(x_i + \mathrm{d}x, t)$, where $\mathrm{d}x$ is a differential space interval and I_i the number of lanes of road section i. Analogously, the departure rate from the downstream end of this section is defined by $Q_i^{\text{dep}}(t) = I_i Q_i(x_{i+1} - \mathrm{d}x, t)$. The conservation of the number of vehicles implies that the departure rate plus the ramp flow determine the arrival rate in the next downstream section $i+1$:

$$Q_{i+1}^{\text{arr}}(t) = Q_i^{\text{dep}}(t) + Q_i^{\text{ramp}}(t). \tag{2}$$

In order to guarantee non-negative flows, we will demand for the ramp flows that the consistency condition $-Q_i^{\text{dep}}(t) \leqslant Q_i^{\text{ramp}}(t) \leqslant Q_{i+1}^{\text{arr}}(t)$ is always met.

Integrating the continuity equation over x with $x_i < x < x_{i+1}$ provides a conservation equation for the number $N_i(t) = \int_{x_i}^{x_{i+1}} \mathrm{d}x\, I_i \rho(x,t)$ of vehicles in road section i. It changes according to

$$\frac{\mathrm{d}N_i(t)}{\mathrm{d}t} = Q_i^{\text{arr}}(t) - Q_i^{\text{dep}}(t) = Q_{i-1}^{\text{dep}}(t) + Q_{i-1}^{\text{ramp}}(t) - Q_i^{\text{dep}}(t). \tag{3}$$

In the following, we will try to express the traffic dynamics and the travel times only through the flows at the cross sections x_i. For this, we will take into account the following simplified features of traffic flow: (i) for free flow, i.e. below some critical vehicle density ρ_{cr} per lane, the relation between the traffic flow Q_i per lane and the density ρ per lane can be approximated by an increasing linear relationship, while above it, a falling linear relationship is consistent with congested flow-density data (in particular, if the average time gap T is treated as a time-dependent, fluctuating variable) [13]. This implies $Q_i(x,t) \approx Q_i(\rho(x,t))$ with

$$Q_i(\rho) = \begin{cases} Q_i^{\text{free}}(\rho) = \rho V_i^0 & \text{if} \quad \rho < \rho_{\text{cr}} \\ Q_i^{\text{cong}}(\rho) = (1-\rho/\rho_{\text{jam}})/T & \text{otherwise.} \end{cases} \tag{4}$$

Here V_i^0 denotes the average free velocity, T the average time gap and ρ_{jam} the density per lane inside traffic jams. Moreover, we define the free and congested densities by

$$\rho_i^{\text{free}}(Q_i) = Q_i \big/ V_i^0 \qquad \text{and} \qquad \rho_i^{\text{cong}}(Q_i) = (1 - TQ_i)\rho_{\text{jam}}. \tag{5}$$

The quantity $Q_{\text{out}} = (1-\rho_{\text{cr}}/\rho_{\text{jam}})/T$ corresponds to the outflow per lane from congested traffic [14]. Depending on the parameter specification, the model describes a continuous flow-density relation (for $\rho_{\text{cr}} V_i^0 = Q_{\text{out}}$) or a capacity drop at the critical density ρ_{cr} and high-flow states immediately before (if $\rho_{\text{cr}} V_i^0 > Q_{\text{out}}$). (ii) According to shock wave theory [4], density variations at place x propagate with velocity $C(t) = [Q_i(x+\mathrm{d}x,t) - Q_i(x-\mathrm{d}x,t)]/[\rho(x+\mathrm{d}x,t) - \rho(x-\mathrm{d}x,t)]$. Accordingly, the propagation velocity is $C = V_i^0$ in free traffic, and $C = -c = -1/(T\rho_{\text{max}})$ in congested traffic. Therefore, it takes the time period $T_i^{\text{free}} = L_i \big/ V_i^0$ for a perturbation to travel through free traffic, while it takes the time period $T_i^{\text{cong}} = L_i/c$, when the entire road section i is congested. (iii) Now, remember that congestion in section i starts to form upstream of a bottleneck, i.e. at place x_{i+1}. Let $l_i(t)$ denote the length of the congested area and $x(t) = x_{i+1} - l_i(t) = x_i + L_i - l_i(t)$ the location of its upstream front. Then, we have free traffic between x_i and $x_i + L_i - l_i(t)$, i.e. $Q_i(x-\mathrm{d}x,t) = Q_i^{\text{arr}}\big(t - (x-x_i)\big/V_i^0\big)$ (considering $\mathrm{d}x \to 0$), and congested traffic downstream of $x(t)$, i.e. $Q_i(x+\mathrm{d}x,t) = Q_i^{\text{dep}}(t - (x_{i+1}-x)/c)$. With $\mathrm{d}x/\mathrm{d}t = -\mathrm{d}l_i/\mathrm{d}t = C(t)$ and equation (5) we find

$$\frac{\mathrm{d}l_i}{\mathrm{d}t} = -\frac{Q_i^{\text{dep}}(t - l_i(t)/c)/I_i - Q_i^{\text{arr}}\big(t - [L_i - l_i(t)]\big/V_i^0\big)\big/I_i}{\rho_i^{\text{cong}}\big(Q_i^{\text{dep}}(t-l_i(t)/c)/I_i\big) - \rho_i^{\text{free}}\big(Q_i^{\text{arr}}\big(t - [L_i - l_i(t)]\big/V_i^0\big)\big/I_i\big)}. \tag{6}$$

(iv) The capacity of a *congested* road section i is approximated as the outflow $Q_{\text{out}} = (1-\rho_{\text{cr}}/\rho_{\text{jam}})/T$ from congested traffic per lane times the number I_i of lanes, minus the maximum bottleneck strength at the end of this section. This may be given by an on-ramp flow $Q_i^{\text{ramp}}(t) > 0$ or analogously by $(I_i - I_{i+1})Q_{\text{out}}$ in the case of a reduction $I_{i+1} - I_i < 0$ in the number of lanes, or in general by some potentially time-dependent value $\Delta Q_i(t)$ in the case of another bottleneck such as a gradient:

$$Q_i^{\text{cap}}(t) = I_i Q_{\text{out}} - \max\big[Q_i^{\text{ramp}}(t), (I_i - I_{i+1})Q_{\text{out}}, \Delta Q_i(t), 0\big]. \tag{7}$$

Analogously, the maximum capacity $Q_i^{\text{max}}(t)$ of the road section i under free flow conditions is given by the maximum flow $I_i \rho_{\text{cr}} V_i^0$ minus the reduction by bottleneck effects:

$$Q_i^{\text{max}}(t) = I_i \rho_{\text{cr}} V_i^0 - \max\big[Q_i^{\text{ramp}}(t), (I_i - I_{i+1})\rho_{\text{cr}} V_i^0, \Delta Q_i(t), 0\big]. \tag{8}$$

Note that an off-ramp flow $Q_i^{\text{ramp}}(t) < 0$ does not contribute to these formulae, but one may reflect the bottleneck effect of weaving flows (i.e. frequent lane changes at exits) by some $\Delta Q_i(t) \geqslant 0$. (v) We can distinguish three states $S_i(t)$ of road section i: $S_i(t) = 0$ corresponds to free traffic, which assumes that the length $l_i(t)$ of the congested area of road section i is zero and that the maximum capacity $Q_i^{\text{max}}(t - \mathrm{d}t)$ in the last time step $t - \mathrm{d}t$, where $\mathrm{d}t$ denotes a small time interval, was not reached by the arrival rate at time $t - \mathrm{d}t - T_i^{\text{free}}$, i.e. $Q_i^{\text{arr}}\big(t - \mathrm{d}t - T_i^{\text{free}}\big) < Q_i^{\text{max}}(t - \mathrm{d}t)$. $S_i(t) = 1$ corresponds to a completely congested road section i, which assumes with $l_i(t) = L_i$ that the congested area expands over the full section length L_i and that the arrival rate $Q_i^{\text{arr}}(t - \mathrm{d}t)$ in the last time step $t - \mathrm{d}t$ was not below the departure rate at time $t - \mathrm{d}t - T_i^{\mathrm{cong}}$, i.e. $Q_i^{\mathrm{dep}}\big(t - \mathrm{d}t - T_i^{\mathrm{cong}}\big) \leqslant Q_{i-1}^{\mathrm{arr}}(t - \mathrm{d}t)$. Otherwise, we have partially congested traffic in road section i and set $S_i(t) = 2$. Altogether this implies:

$$S_i(t) = \begin{cases} 0 & \text{if} \quad l_i(t) = 0 \text{ and } Q_i^{\mathrm{arr}}\big(t - \mathrm{d}t - T_i^{\mathrm{free}}\big) < Q_i^{\mathrm{max}}(t - \mathrm{d}t) \\ 1 & \text{if} \quad l_i(t) = L_i \text{ and } Q_i^{\mathrm{dep}}\big(t - \mathrm{d}t - T_i^{\mathrm{cong}}\big) \leqslant Q_{i-1}^{\mathrm{arr}}(t - \mathrm{d}t) \\ 2 & \text{otherwise.} \end{cases} \tag{9}$$

(vi) Finally, we have to specify the departure rate $Q_i^{\mathrm{dep}}(t)$ as a function of the respective traffic situation. Focussing on the cross section at location x_{i+1} and considering the directions of information flow (i.e. the propagation direction of density variations), we can distinguish three different cases:

1. If we have free traffic in the upstream section i and free or partially congested traffic in the downstream section $i + 1$, density variations propagate downstream and the departure rate $Q_i^{\mathrm{dep}}(t)$ at time t is given as the arrival rate $Q_i^{\mathrm{arr}}\big(t - T_i^{\mathrm{free}}\big) = Q_{i-1}^{\mathrm{dep}}\big(t - T_i^{\mathrm{free}}\big) + Q_{i-1}^{\mathrm{ramp}}\big(t - T_i^{\mathrm{free}}\big)$, since the vehicles entering section i at time $t - T_i^{\mathrm{free}}$ leave the section after an average travel time T_i of T_i^{free}.
2. In the case of partially or completely congested traffic upstream and free or partially congested traffic downstream, the departure rate $Q_i^{\mathrm{dep}}(t)$ is given by the capacity $Q_i^{\mathrm{cap}}(t)$ of the congested road section i.
3. In the case of congested traffic on the entire downstream road section $i + 1$, the departure rate $Q_i^{\mathrm{dep}}(t)$ is given by the departure rate $Q_{i+1}^{\mathrm{dep}}\big(t - T_{i+1}^{\mathrm{cong}}\big)$ from the downstream section at time $t - T_i^{\mathrm{cong}}$ minus the ramp flow $Q_i^{\mathrm{ramp}}(t)$ entering at location x_{i+1}.

Summarizing this, we have

$$Q_i^{\mathrm{dep}}(t) = \begin{cases} Q_i^{\mathrm{arr}}\big(t - T_i^{\mathrm{free}}\big) & \text{if} \quad S_{i+1}(t) \neq 1 \quad \text{and} \quad S_i(t) = 0 \\ Q_i^{\mathrm{cap}}(t) & \text{if} \quad S_{i+1}(t) \neq 1 \quad \text{and} \quad S_i(t) \neq 0 \\ Q_{i+1}^{\mathrm{dep}}\big(t - T_{i+1}^{\mathrm{cong}}\big) - Q_i^{\mathrm{ramp}}(t) & \text{if} \quad S_{i+1}(t) = 1. \end{cases} \tag{10}$$

A numerical solution of the above-defined section-based queueing-theoretical model is carried out as follows: first, determine the states of the road sections according to equation (9). Second, calculate the new arrival and departure rates by means of equations (2) and (10), taking into account the boundary conditions for the flows at the open ends of the road network. Third, if $S_i(t) = 0$, then set $l_i(t+\mathrm{d}t) = l_i(t) = 0$. If $S_i(t) = 1$, then set $l_i(t+\mathrm{d}t) = l_i(t) = L_i$. Otherwise, if $S_i(t) = 2$, determine the length $l_i(t + \mathrm{d}t) = l_i(t) + \mathrm{d}t\, \mathrm{d}l_i(t)/\mathrm{d}t$ of the congested area in road section i with formula (6). Next, continue with the first step for the new time point $t + \mathrm{d}t$. It is obvious that this numerical solution is significantly more simple and robust than the numerical solution of the Lighthill–Whitham model, as shock waves (i.e. the interfaces between free and congested traffic) are treated analytically.

Let us now derive a general relationship for the average travel time $T_i(t)$ of a vehicle that enters road section i at time t. The average travel time $T_i(t)$ is given by the fact that this vehicle will leave section i, when (on average) the $N_i(t)$ vehicles that are in section i at time t have passed the downstream end of that section. Since the number of vehicles passing the downstream end of section i is given as the time-integral over the departure rate $Q_i^{\mathrm{dep}}(t)$, we have the implicit relationship

$$N_i(t) = \int_t^{t+T_i(t)} \mathrm{d}t'\, Q_i^{\mathrm{dep}}(t') = \int_{-\infty}^{t+T_i(t)} \mathrm{d}t'\, Q_i^{\mathrm{dep}}(t') - \int_{-\infty}^{t} \mathrm{d}t'\, Q_i^{\mathrm{dep}}(t'). \tag{11}$$

Identifying equation (3) with the time derivative of this equation finally leads to the delay-differential equation

$$\frac{\mathrm{d}T_i(t)}{\mathrm{d}t} = \frac{Q_i^{\mathrm{arr}}(t)}{Q_i^{\mathrm{dep}}(t + T_i(t))} - 1 = \frac{Q_{i-1}^{\mathrm{dep}}(t) + Q_{i-1}^{\mathrm{ramp}}(t)}{Q_i^{\mathrm{dep}}(t + T_i(t))} - 1. \tag{12}$$

According to this, the travel time $T_i(t)$ increases with time, when the arrival rate Q_i^{arr} at the time t of entry exceeds the departure rate Q_i^{dep} at the leaving time $t + T_i(t)$, while it decreases when it is lower. It is remarkable that this formula does not explicitly depend on the velocities on the road section, but only on the arrival and departure rates. The calculation of the travel time based on the velocity $v(t)$ is considerably more complicated as it depends on the traffic density $\rho(x(t), t)$ at its respective location $x(t) = x_i + \int_{t_0}^{t} \mathrm{d}t'\, v(t')$.

In summary, we have proposed a novel queueing theoretical model, which facilitates to simulate the departure rates $Q_i^{\mathrm{dep}}(t)$ of road sections i and the average travel times $T_i(t)$ of vehicles in an efficient way. Assuming a fundamental diagram with linear free and congested branches (i.e. constant propagation velocities V_i^0 and $-c$ in free and congested traffic, respectively) allowed us to eliminate the traffic dynamics within homogeneous road sections apart from the location of shock fronts, i.e. the moving interfaces between free and congested traffic. By determining the travel times from the arrival and departure rates of vehicles, the model considerably differs from classical traffic models, in which they are determined from the spatio-temporal vehicle speeds. However, note that there is some relationship with cell-based approaches [15], if we subdivide our large homogeneous road sections into smaller subsections (cells).

Apart from numerical efficiency, the proposed section-based model is promising for analytical investigations of traffic in road networks and of dynamic assignment problems. Moreover, it can describe the hysteretic breakdown of traffic flow and reproduce typical congestion patterns. Nevertheless, it is not as accurate as other macroscopic traffic models [6], which do not assume constant propagation velocities and can describe emergent stop-and-go waves. Despite this, the queueing-theoretical traffic model is expected to provide reasonable estimates of the average travel times.

The above model can be extended in several ways: (i) while the model can describe boundary-induced stop-and-go waves (characterized by the temporary disappearance of a bottleneck $I_i Q_{\mathrm{out}}(t) - Q_i^{\mathrm{cap}}(t)$), emergent stop-and-go waves require the formulation of a generalized model containing a dynamic velocity equation, which will reduce analytical tractability. (ii) Accidents may be treated by a splitting of sections at the location and for the duration of an accident. (iii) Stochastic effects in traffic flows could be easily incorporated. (iv) Forecasts of travel times would be possible, if the above model would be combined with a model for the prediction of origin-destination flows. In such kinds of simulations, it makes sense to extend the model to multi-destination flows [15]. (v) Finally, the above model may be generalized to the treatment of urban road networks with traffic lights, if we specify the capacities $Q_i^{\mathrm{cap}}(t)$ in a time-dependent way which reflects the effects of amber and red lights. Note that the actual section capacity does not change immediately with the traffic light, but with some delay due to reaction and acceleration or deceleration times.

Acknowledgments

The author is grateful to Martin Treiber, Martin Schönhof, Thomas Seidel and an anonymous referee for useful comments.

References

[1] Helbing D 2001 *Rev. Mod. Phys.* **73** 1067
Chowdhury D, Santen L and Schadschneider A 2000 *Phys. Rep.* **329** 199
Nagatani T 2002 *Rep. Prog. Phys.* **65** 1331
[2] Biham O, Middleton A A and Levine D 1992 *Phys. Rev.* A **46** R6124
Nagel K and Schreckenberg M 1992 *J. Physique* **2** 2221
[3] Bando M, Hasebe K, Nakayama A, Shibata A and Sugiyama Y 1995 *Phys. Rev.* E **51** 1035
Krauß S, Wagner P and Gawron C 1997 *Phys. Rev.* E **55** 5597
Treiber M, Hennecke A and Helbing D 2000 *Phys. Rev.* E **62** 1805
[4] Lighthill M J and Whitham G B 1955 *Proc. R. Soc.* A **229** 317
[5] Kerner B S and Konhäuser P 1994 *Phys. Rev.* E **50** 54
Herman R and Prigogine I 1979 *Science* **204** 148
[6] Helbing D, Hennecke A and Treiber M 1999 *Phys. Rev. Lett.* **82** 4360
[7] Kerner B S and Rehborn H 1997 *Phys. Rev. Lett.* **79** 4030
Lee H Y, Lee H-W and Kim D 1998 *Phys. Rev. Lett.* **81** 1130
Helbing D and Treiber M 1998 *Phys. Rev. Lett.* **81** 3042
[8] Jain R and Smith J 1997 *Transp. Sci.* **31** 324
Vandaele N, Woensel T V and Verbruggen A 2000 *Transp. Res.* D **5** 121
Lan C-J and Davies G A 1997 *Transp. Res. Rec.* **1591** 31
[9] Cremer M and Landenfeld M 1998 *Traffic and Granular Flow '97* ed M Schreckenberg and D E Wolf (Singapore: Springer) p 169
[10] Eissfeldt N, Gräfe J and Wagner P 2003 *Transportation Research Board Proc. of the Annual Meeting 2004 Preprint* zaik2003-456
[11] Kerner B 2001 *German Patent* DE 199 40 957 C2
[12] Helbing D 2003 *New J. Phys.* **5** 90.1
[13] Nishinari K, Treiber M and Helbing D 2003 *Phys. Rev.* E submitted (*Preprint* cond-mat/0212295)
[14] Kerner B S and Rehborn H 1996 *Phys. Rev.* E **53** R1297
[15] Daganzo C F 1995 *Transp. Res.* B **29** 79
Hilliges M and Weidlich W 1995 *Transp. Res.* B **29** 407
Newell G F 1993 *Transp. Res.* B **27** 281