Cellular Automata Simulating Experimental Properties of Traffic Flow

Dirk Helbing 1 and Michael Schreckenberg 2

1II. Institute of Theoretical Physics, University of Stuttgart, Pfaffenwaldring 57/III, 70550 Stuttgart, Germany
2Theoretische Physik, Gerhard-Mercator-Universität Duisburg, Lotharstraße 1, 47048 Duisburg, Germany

A model for 1D traffic flow is developed, which is discrete in space and time. Like the cellular automaton model by Nagel and Schreckenberg [J. Phys. I France 2, 2221 (1992)], it is simple, fast, and can describe stop-and-go traffic. Due to its relation to the optimal velocity model by Bando et al. [Phys. Rev. E 51, 1035 (1995)], its instability mechanism is of deterministic nature. The model can be easily calibrated to empirical data and displays the experimental features of traffic flow recently reported by Kerner and Rehborn [Phys. Rev. E 53, R1297 (1996)].

05.50.+q,05.40.+j,47.55.-t,89.40.+k

Cellular automata (CA) are interesting for their speed and their complex dynamical behavior, including such fascinating phenomena as self-organized criticality, formation of spiral patterns, or oscillatory and chaotic sequences of states. Their enormous computation speed and efficiency is a consequence of the following properties, which are ideal preconditions for parallel computing: (i) discretization of space into identical sites, (ii) a finite number of possible states \( f(x) \), (iii) the (parallel) update at times \( T = t \Delta T \) with an elementary time step \( \Delta T \), (iv) globally applied update rules, based on (v) short-range interactions with a finite (small) number of neighbouring sites. Despite these simplifications, cellular automata have a broad range of applications, reaching from realistic simulations of granular media or fluids up to the modeling of avalanches. Their application to traffic dynamics has stimulated an empirical traffic data recently reported by Kerner and Rehborn [Phys. Rev. E 53, R1297 (1996)].

Recently, Kerner and Rehborn have reported on empirical high-speed traffic flow, which a realistic traffic model should display: (i) At small densities, traffic flow is stable, i.e., arbitrarily large disturbances of homogeneous traffic will disappear in the course of time. (ii) Above a certain critical density, any small perturbation will give rise to the formation of a traffic jam. (iii) Between the stable and the unstable regions, there exists a density interval beginning at about 20 vehicles per kilometer, where traffic flow is metastable. That is, sufficiently small disturbances (so-called ‘sub-critical perturbations’) will fade away, whereas ‘super-critical’ perturbations exceeding a certain minimal amplitude will cause a traffic jam. (iv) The outflow from traffic jams has a typical value which is independent of the initial conditions and, to a large extent, independent of the average surrounding traffic density. It varies only with road or weather conditions, and the average vehicle characteristic properties (regarding their lengths and acceleration capabilities). (v) The typical outflow is considerably smaller than the maximum flow and lies at about 1600 vehicles/km for slow lanes, and on fast lanes between 1800 vehicles/km (measured in Germany) and 2100 vehicles/km (on Dutch motorways). It is associated with a density of \( 20 \pm 5 \) vehicles per kilometer. (vi) Downstream jam fronts move with a velocity of \(-15 \pm 5\) km/h. Notice that properties (iv) and (vi) originate from the uniform acceleration behavior of queued vehicles, resulting from the similar distances and velocities that they share inside traffic jams. While the propagation velocity \( C \) of traffic jams is given by the dissol-
fraction $\Delta X = \ell/n$ of $\ell$ ($n \in \{1, 2, 3 \ldots \}$). Velocity steps of the order 5–10 km/h require $\Delta X \leq \ell/2$. In the following, the spatial coordinate $X = x_1 \Delta X$ and distances $D = d_\Delta X$ are measured in units of $\Delta X$, time $T = t \Delta T$ in units of $\Delta T$, and any velocity in units of $\Delta V$.

Assume we want to distinguish $A$ different vehicle types $a \in \{1, 2, \ldots, A\}$, e.g., cars and trucks. Then, at each time $t \in \{0, 1, 2, \ldots \}$, every site $x \in \{1, 2, \ldots, l\}$ can be in one of the states $f_t(x) = (a, v)$, where $f_t(x) = (0, 0)$ corresponds to an empty site, and $f_t(x) = (a, v)$ with $a > 0$ represents a vehicle of type $a$ with velocity $v \in \{0, 1, 2, \ldots, v_{\text{max}}^a\}$. For safety reasons, the velocity $v_t$ must be smaller than the distance $d_t$ to the vehicle ahead.

The states of the cellular automaton are updated in parallel according to the following successive steps: First, each vehicle is moved by its actual velocity $v_t$ to position $x + v_t$, which means $f_{t+1}(x + v_t) = f_t(x)$, if $a > 0$ and $v_t > 0$. Then, the states of the previously occupied positions are reset to $f_{t+1}(x) = (0, 0)$. Any other site keeps its previous state. Finally, all vehicle velocities are modified along with the proposed acceleration rule

\begin{align*}
\nu_{t+1}^a &= v_t^a + \left[ \lambda_a [v_a(d_{t+1}) - v_t^a] \right], \\
\nu_{t+1} &= \nu_{t+1}^a - \left\{ \begin{array}{ll}
1 & \text{with probability } p, \text{ if } \nu_{t+1}^a > 0 \\
0 & \text{otherwise}
\end{array} \right.
\end{align*}

where $\lfloor y \rfloor$ is defined by the largest integer $i \leq y$. Therefore, the above equation implies $v_{t+1}^a \leq \lambda_a v_a(d_{t+1}) + (1 - \lambda_a) v_t^a$, meaning that the new velocity is a weighted average of the previous velocity $v_t$ and the optimal velocity $v_a$ of vehicle type $a$, or somewhat less. A small value of the model parameter $\lambda_a \geq 0$ relates to a great inertia of vehicle motion, whereas a large value $\lambda_a \leq 1$ implies a fast adaptation to the distance-dependent optimal velocity $v_a(d)$. The corresponding adaptation time is $\tau_a = \Delta T/\lambda_a$. If the (back-bumper-to-back-bumper) distance $d_{t+1}$ to the next vehicle exceeds a certain finite value $d_{\text{min}}$, the vehicles do not interact, and $v_{t+1}$ is given by the maximum velocity $v_{\text{max}}^a$ of vehicle type $a$. For small distances, $v_{t+1}$ should be determined by the velocity-dependent safe distance $d(v_a) \approx \ell/\Delta X + v_a$, required to avoid accidents. The model parameter (‘slowdown probability’) $p$ describes individual velocity fluctuations due to delayed acceleration (imperfect driving). Here, we are interested in the limit $p \to 0$.

In order to compare this discrete model with the observed properties of traffic flows, it is necessary to investigate aggregate quantities $(h)$. These are defined by

\[ \langle h \rangle_{x,t} = \frac{1}{\Delta t} \sum_{t'} f_{t'}(x) \sum_{x'=x-\Delta x}^{x+\Delta x} h(f_{t'}(x')). \]

The vehicle density is given by $\rho(x, t \Delta T) = \Theta(a)/\Delta X$, the traffic flow by $Q(x, t \Delta T) = \langle v \Theta(a) \rangle/\Delta T$, and the average velocity by $V(x, t \Delta T) = Q(x, t \Delta T)/\rho(x, t \Delta T)$, where $\Theta(a) = 1$ for $a \geq 0$, otherwise $= 0$. During the simulation runs, the density minima and maxima, their propagation speed, and the associated traffic flows were evaluated automatically, as well as the spatially averaged vehicle velocity and traffic flow on the circular road of length $L = 20\text{km}$. The corresponding values were averaged over several hours after a sufficiently long transient period. Averages over the whole street are indicated by an overbar.

First, let us discuss the case of one type $a = 1$ of vehicles. Our simulation results can be summarized as follows: At small average densities $\pi$, homogeneous traffic flow is stable, and the spatially averaged velocity $\overline{v}$ is given by the density-independent value $\langle v \rangle_{\pi} = \langle v \rangle_{\pi}^0 - \left\lfloor \frac{1}{\lambda_1 - 1} - p \right\rfloor \Delta V$. It is zero for high densities [Fig. 3(a)], since $\lambda_1 v_1(d) > 1$ is necessary for the acceleration of a vehicle from standstill.

At medium densities, the resulting velocity-density relation is largely dependent on $\lambda_1$: Regime I: For $0 \leq \lambda_1 \leq \lambda_{\text{stab}}(v_{\text{max}}^1, p)$ (corresponding to a quasi-instantaneous adaptation to the optimal velocity), $\overline{v}(\pi)$ is close to the piecewise constant relation $v_1(1/\pi)$. Accordingly, the associated average traffic flow $\overline{Q} = \langle v \rangle_{\pi}$ is a piecewise linear function of $\pi$. Regime II: In a certain interval $\lambda_{\min}(v_{\text{max}}^1, p) \leq \lambda_1 \leq \lambda_{\text{max}}(v_{\text{max}}^1, p)$, traffic flow is unstable for a certain range $\rho_{\text{out}} \leq \pi \leq \rho_{\text{max}}$ of medium densities, and $\overline{Q}(\pi)$ becomes the self-organized linear relation

\[ \overline{Q}(\pi) = \frac{1}{\pi} \left( 1 - \frac{\pi}{\rho_{\text{jam}}} \right) \]

demanded by Kerner [16, Figs. 1(a) and 2(b)]. $T$ denotes the average time gap between the acceleration of successive vehicles. The linear relation $\overline{Q}(\pi)$ reflects a mixture of free and jammed traffic with characteristic densities $\rho_{\text{out}}$ and $\rho_{\text{jam}}$, respectively, where the jammed regions grow with increasing density. $\rho_{\text{out}} = \overline{Q}(\rho_{\text{out}})$ is the typical outflow from traffic jams [16]. The slope

\[ C = \frac{\partial \overline{Q}}{\partial \rho} = -\frac{1}{T \rho_{\text{jam}}} \]

corresponds to their dissolution velocity [16]. The dependence of the spatially averaged velocity $\overline{v}$ on $\pi$ is given by $\overline{v}(\pi) = \overline{Q}(\pi)/\pi$. Regime III: In a parameter range $\lambda_{\min}(v_{\text{max}}^1, p) \leq \lambda_1 \leq \lambda_{\text{stab}}(v_{\text{max}}^1, p)$, there is still an unstable range of traffic, but $\overline{Q}(\pi)$ is piecewise linear with different slopes $C$ and different values of $\rho_{\text{jam}}$ or $\rho_{\text{out}}$ [like in Fig. 3(b)], where the above relations are separately fulfilled for each linear piece. This may be understood as crossover behavior between the cases I and II. Notice that the discretization of vehicle dynamics implies $1/\rho_{\text{max}} = k_1 \Delta X$ and $T = k_2 \Delta T/k_3$ with small integers $k_3$ (see below). Thus, $C$ is restricted to a few discrete values $h_n/h_{\Delta V}$. Regime IV: For $\lambda_1 v_{\text{max}}^1 < 1$, $\lambda_1$ is so small that $v_{t+1} \leq v_t$, which, for $p \neq 0$, implies that traffic eventually comes to rest.
The characteristic quantities can explicitly be calculated. Let us show this for the simple but nontrivial case \( p \to 0 \) and \( v_1(d) = \min(d - 1, 3) \) modeling city traffic. With \( \Delta X = 6.25 \text{ m} \) we have \( \rho_{\text{jam}} = 1/\Delta X = 160 \text{ vehicles/km} \). Now, let us assume that a disturbance has produced a queue of vehicles with velocities \( v = 0 \) and distances \( d = 1 \) to the respective vehicles ahead, and a free road in front of the first vehicle. We can characterize the acceleration behavior of a vehicle by the sequence

\[
(v_t', d_t') \rightarrow (v_{t+1}', d_{t+1}) \rightarrow (v_{t+2}', d_{t+2}) \rightarrow \ldots \,
\]

where, for \( \lambda = 0.77 \), \( v_{t+1} = \max(d_t - 2, 0) \) if \( (v_t + 1) \leq d_t \leq 4 \), and \( d_{t+1} = (d_t + v_{t+1}' - v_{t+1}) \) (\( v_{t+1}' \) being the velocity of the vehicle ahead). Denoting with \( t' \) the time when the state \( (0,1) \) of the respective vehicle ahead has changed to another state, we find two alternating sequences: \( (0,1) \rightarrow (0,1) \rightarrow (0,3) \rightarrow (1,4) \rightarrow (2,4) \rightarrow \ldots \) and \( (0,1) \rightarrow (0,2) \rightarrow (0,4) \rightarrow (2,4) \rightarrow \ldots \). That is, cars start to accelerate alternatingly every one or two time steps \( \Delta T \). With \( \Delta T = 1 \text{ s} \), we find an average of \( T = 1.5 \text{ s} \), which implies the dissolution velocity \( C = -15 \text{ km/h} \). Moreover, the resulting density in front of jams is given by the evolving maximal distance \( 4\Delta X \), which gives \( \rho_{\text{out}} = 40 \text{ vehicles/km} \) and, according to Eq. [1], an outflow of \( Q_{\text{out}} = 1800 \text{ vehicles/h} \). All this is in total agreement with simulation results [Fig. 2(a)].

Now, let us investigate the behavior for fixed \( \lambda_1 = 0.77 \) and density \( \overline{\rho} = 80 \text{ vehicles/km} \), but various \( p \) [Fig. 2(a)]. If \( p \) is continuously reduced, the backward dissolution velocity \( C \) is decreasing, whereas the outflow \( Q_{\text{out}} \) from traffic jams, and the difference between the jam density \( \rho_{\text{jam}} \) and the self-organized density \( \rho_{\text{out}} \) downstream of traffic jams are rapidly growing towards an almost constant value. In particular, the amplitude \( \rho_{\text{jam}} - \rho_{\text{out}} \) of traffic jams is constant over more than five decades.

In the deterministic case \( p = 0 \), the formation of traffic jams requires some initial inhomogeneity. Let us investigate the response to localized perturbations of the form \( \rho(x,0) = \overline{\rho} + \Delta \rho \cos^2[(x - L/2)/w_+ - (w_+/w_-) \cos^{-2}[(x - L/2 - w_+ - w_-)/w_-]] \), as suggested in Ref. [23]. Typically, one observes a piecewise linear response like in regime III [Fig. 2(b)]. However, its dependence on the perturbation amplitude \( \Delta \rho \) indicates multistability, i.e. the coexistence of a variety of solutions. These correspond to periodic patterns of vehicle updates (which naturally result in deterministic systems with a finite number of states). In the presence of noise \( (p > 0) \), only one of them survives [Fig. 2(a)], i.e. most of them are unstable with respect to fluctuations. This explains the role of randomization for the behavior resulting in regime II.

Finally, let us focus on the density-dependent behavior for fixed \( \lambda_1 = 0.77 \) and \( p > 0 \). We find that initial localized perturbations of the above form, regardless of their amplitude \( \Delta \rho \), are damped out for average densities \( \overline{\rho} \) below some value \( \rho_{c1}(\nu^1_{\text{max}} ; p) \) and above some value \( \rho_{c4}(\nu^1_{\text{max}} ; p) \). In a certain density range \( \rho_{c2}(\nu^1_{\text{max}} ; p) < \overline{\rho} < \rho_{c3}(\nu^1_{\text{max}} ; p) \), the perturbation grows for any finite amplitude. In the density regimes \( \rho_{c1}(\nu^1_{\text{max}} ; p) \leq \overline{\rho} < \rho_{c2}(\nu^1_{\text{max}} ; p) \) and \( \rho_{c3}(\nu^1_{\text{max}} ; p) < \overline{\rho} \leq \rho_{c4}(\nu^1_{\text{max}} ; p) \), we observe metastability [Fig. 3], i.e. perturbations with an amplitude \( \Delta \rho \geq \Delta \rho_{c1}(\overline{\rho}) \) will grow, otherwise they will fade away in the course of time (‘local cluster effect’ [20,23]). Notice that \( \rho_{c2} \approx \rho_{\text{out}} \) and \( \rho_{c3} \approx \rho_{\text{max}} \).

These findings can be understood by analogy with the continuous optimal velocity model by Bando, Sugiyama et al. [21] that results in the limit \( \Delta T \to 0 \) and \( \Delta V \to 0 \). It displays stable traffic at low and high vehicle densities, unstable traffic on the condition \( \partial v_1(d)/\partial d > \lambda_1/2 \), and metastable regimes between the unstable and stable ones [21]. However, in the continuous optimal velocity model, the jam density \( \rho_{\text{jam}} \) is not independent on how a traffic jam is formed, since fast cars are more crowded than slow ones, after they had to stop. This undesired property can be avoided by a refined model [24] or just by a suitable discretization, like in the proposed cellular automaton, when operated in regime II.

In summary, we have developed a cellular automaton for one-lane traffic which reproduces many of the empirically observed features of traffic flow in a “unified” way. In particular, the model showed the characteristic quantities of traffic flow, which we managed to calculate analytically. By appropriate specification of the tabular functions \( v_{\gamma}(d) \) and the parameters \( \lambda_1, p, \Delta T, \) and \( \Delta V \), the model can be calibrated to empirical data. The most interesting case is to operate the model in regime II, since this guarantees the desired properties (iv) and (vi). The characteristic quantities like \( C \) and \( Q_{\text{out}} \) are determined by \( \rho_{\text{out}}, \rho_{\text{jam}} = k_1 \Delta T, \) and \( T = k_2 \Delta T/k_3 \). The latter can be enforced by a suitable choice of \( \Delta X \) and \( \Delta T \). \( v_{\gamma}(d) \) and \( \lambda_1 \) determine the maximum velocity, the approximate velocity-density relation, the instability region, and the amplitude \( \rho_{\text{jam}} - \rho_{\text{out}} \) of traffic jams. The metastable regimes and the difference between the maximal possible traffic flow and the self-organized outflow \( Q_{\text{out}} \) from traffic jams grow with increasing \( \nu^0_{\text{max}} \). Finally, \( p \) allows to influence the characteristic ‘wave length’ between successive traffic jams [Fig. 2(a)]. Suitable choices are \( \Delta T \in [1 \text{ s}, 1.3 \text{ s}] \), \( \lambda_1 \approx 0.77 \) and \( p \leq 0.01 \). The optimal velocity functions \( v_{\gamma}(d) \) were chosen proportional to relations that were determined from traffic data of the Dutch freeway A9. The results are in good agreement with macroscopic traffic data [Fig. 2(b)–(d)].

The authors want to thank for financial support by the BMBF (research project SANDY) and the DFG (Heisenberg scholarship He 2789/1-1). They are grateful to B. Kerner for valuable comments, and to H. Taale and the Dutch Ministry of Transport, Public Works and Water Management for supplying the freeway data.
[1] S. Wolfram, Nature 311, 419 (1984).
[2] P. Bak et al., Phys. Rev. Lett. 59, 381 (1987).
[3] P. Bantay and I. M. Janosi, Phys. Rev. Lett. 68, 2058 (1992).
[4] Z. Olami et al., Phys. Rev. Lett. 68, 1244 (1992).
[5] M. Markus and B. Hess, Chania, Greece, 1997, Vol. II.
[6] M. A. Nowak and R. M. May, Nature 359, 826 (1992).
[7] G. Engel and H. J. Herrmann, Phys. Rev. E 49, 1796 (1994).
[8] U. Frisch et al., Phys. Rev. Lett. 56, 1505 (1986); S. Chen et al., Phys. Rev. Lett. 67, 3776 (1991).
[9] D. Dab et al., Phys. Rev. Lett. 66, 2535 (1991).
[10] M. Bando et al., Phys. Rev. E 51, 2939 (1995).
[11] T. Nagatani, Phys. Rev. E 51, 922 (1995).
[12] M. Schreckenberg and D. E. Wolf (eds.) Traffic and Granular Flow ’97 (Springer, Singapore, 1998).
[13] D. Helbing, Verkehrsdynamik (Springer, Berlin, 1997).
[14] D. Helbing and M. Treiber, Phys. Rev. Lett. 81, 3042 (1998).
[15] B. S. Kern and H. Rehborn, Phys. Rev. E 53, R1297 and R4275 (1996).
[16] B. S. Kern, in Transportation Systems, ed. by M. Papageorgiou and A. Poulios (International Federation of Automatic Control, Chania, Greece, 1997), Vol. II.
[17] K. Nagel and H. J. Herrmann, Physica A 199, 254 (1993); M. Takayasu and H. Takayasu, Fractals 1, 860 (1993).
[18] S. Krauß et al., Phys. Rev. E 54, 3707 (1996) and 55, 5597 (1997); S. Krauß, Microscopic Modelling of Traffic Flow (DLR, Cologne, 1998), FB 98-08.
[19] B. Barlovic et al., European Journal of Physics B, in print (1998).
[20] B. S. Kern and P. Konhäuser, Phys. Rev. E 48, R2335 (1993) and 50, 54 (1994).
[21] M. Bando et al., Phys. Rev. E 51, 1035 (1995) and J. Phys. I France 5, 1389 (1995).
[22] D. Helbing and B. A. Huberman, Nature, in print (1998).
[23] M. Herrmann and B. S. Kerner, Physica A 255, 163 (1998).
[24] D. Helbing and B. Tilch, Phys. Rev. E 58, 133 (1998).

FIG. 1. Simulation results for $v_1(d) = \min(d-1.3, \lambda_1 = 0.77, \Delta T = 1 \text{s}, \Delta X = 6.25 \text{m}$, and (a) $p = 0.001$, (b) $p = 0, w_+ = w_- = 200 \text{ m}$, and $\Delta \rho = \rho/2$. Illustration (a) shows the “optimal flow” $v_1(1/\rho)$ (−) and the resulting average flow $Q(\rho)$ (□) as a function of the average density $\rho$, (b) the densities inside (−−) and in front of (−−) traffic jams.

FIG. 2. (a) Characteristic parameters of traffic flow as a function of the slowdown probability $p$ for the model parameters specified in Fig. 1. (b) Comparison of the model corresponding to $\lambda_1 = 0.77, p = 0.001, \Delta T = 1 \text{s}, \Delta X = 2.5 \text{ m}$, and the optimal flow relation specified by ‘+’ with 1-minute-averages of single-vehicle data from the left lane of an undisturbed cross section of the Dutch highway A9 (−). Boxes illustrate the simulated average flow resulting in the limit of long times. (c) Simulation of ‘stop-and-go waves’ at an average density of 28 vehicles/km for a mixture of 90% cars and 10% trucks (the optimal velocities of which are only 70% of the cars). The occurring minimal and maximal velocities, the minimal densities, and the largely varying time intervals of successive breakdowns of velocity are in good agreement with the Dutch freeway data (1-minute-averages) displayed in (d). The mixture of vehicle types also explains the observed fluctuations in the density and average velocity of vehicles.

FIG. 3. (a) Critical amplitudes of localized perturbations with $w_+ = 200 \text{ m}$ and $w_- = 800 \text{ m}$ as a function of average density for the model specified in Fig. 2(b). Whereas larger perturbations cause the formation of traffic jams, smaller ones will fade away ($\rho_1 = 21 \text{ vehicles/km}$, $\rho_2 = 23 \text{ vehicles/km}$, $\rho_3 = 150 \text{ vehicles/km}$, $\rho_4 \geq 164 \text{ vehicles/km}$). (b) The average wave length of emerging stop-and-go waves diverges at $\rho_2$ (checked for large system sizes). This is why they are not triggered by fluctuations below this density.