Transportation Networks: Dynamics and Simulation

Stephen G. Eubank

*Basic and Applied Simulation Science Group,
D-2, MS-M997, LANL, Los Alamos, NM, 87545, USA*

Abstract. Analyzing the behavior of socio-technical systems – large urban populations interacting under constraints imposed by technical infrastructure – poses many interesting problems. We describe why simulation is an appropriate experimental method for studying these systems, drawing examples from the development and implementation of the Transportation Analysis and Simulation System (TRANSIMS). TRANSIMS is an urban planning tool that combines census data and household surveys with a description of the transportation network to produce estimates of human mobilty. It iteratively evaluates hundreds of thousands of coupled non-linear models to produce a solution to a million-person game.

INTRODUCTION

TRANSIMS[1] is a machine for turning readily available data into an estimate of a very complicated joint distribution – the location of every individual in an urban region during the course of a day. [2] The input data for TRANSIMS consists of:

- marginal distributions in many variables (demographics from census data)
- a small sample of some of the joint distributions among those variables (also from census data)
- a small sample of activity patterns (surveys, diaries)
- details of the transportation infrastructure (maps, transit schedules) in the form of a graph.

The output is the estimated distribution of a collection of demographic variables (in other words, people) at every location at every second.

The estimate respects travel time and cost constraints placed on people's behavior by the transportation infrastructure. It makes an assumption that in general people strive to minimize their cost, but neither assumes that they are perfectly "rational" nor that they all share the same cost function. In particular, they do not necessarily have complete information on which to base decisions.

The heart of the problem to be solved by TRANSIMS is a million-person game. Each player picks a strategy, in this case a route through the transportation network that allows him or her to execute a planned set of activities. A solution is a set of strategies in which nobody is too far from an optimal route. The game is solved iteratively using a traffic simulation to estimate the game's outcome given each player's strategy for that iteration. The solution is self-consistent in the sense that if everyone were allowed to make choices based on the results of their and everyone else's previous choices, the overall outcome would remain the same.

WHY TRAFFIC?

Traffic, or more generally transportation, has long been an important factor determining how urban environments grow. Traffic modeling provides an experimental testbed for urban planning. It makes possible evaluation of proposed infrastructure changes, including

construction projects
changes in the way infrastructure is used, such as HOV lanes
pricing changes for tolls, mass transit, or parking

Importantly, because of its self-consistency, TRANSIMS takes into acount changes in demand caused by hypothesized infrastructure changes rather than assuming the demand will remain fixed. Evaluation can take the form of total travel times, vehicle-miles, emissions, or many other variables. All these effects can be traced to the demographics of specific individuals, allowing estimates of the differential effect on different, possibly overlapping, subpopulations. For instance, urban planners can determine who will benefit from improvements in local streets (in the form of shorter travel time) and who will be exposed to degradations in air quality.

Important as it is for its own sake, estimating traffic is not the final goal of this simulation process. Human mobility is the basic data needed for understanding the characteristics of social networks. These in turn underlie estimates of mobile telecommunication demand, electric power demand, epidemiology, and many other poorly understood processes. Epidemiology, for instance, is essentially the study of diffusion on a time-dependent graph generated by human mobility. Until now, there has been no good characterization of these graphs. Only ad-hoc hypotheses such as "small-world" models have been studied.

INDIVIDUAL-BASED SIMULATION

Population, Activities, Routes, and Traffic

TRANSIMS synthesizes from census data a population with the correct joint distribution of demographics and locates this population on an actual transportation network. Each household is matched demographically to one of approximately 2000 households whose members filled out activity diaries (the "survey data"). The diaries define a set of activity patterns specifying activity type, start time, and duration – but not location – for each member of a household. Each synthetic household is assigned an activity pattern drawn from those of households with similar demographics. An initial assignment of important anchor locations such as work or school is made based on information such as known employment by location throughout the area.

Based on this set of activity location and time assignmnts, we use a novel shortest-constrained-path technique to find routes through the transportation network.[3] The novel aspects of the route-finder are that it makes use of time dependent travel times on each link of a mutli-modal network and that it easily handles path constraints expressed in a simple language. Examples of constraints that might be appropriate are:

find a path using the modes *walk* followed by *car* followed by *walk*;

find a path using at least one but no more than two bus rides;

find a path that avoids toll roads.

Finding shortest paths through a large network is computationally expensive, but feasible. Finding the shortest path while simultaneously accounting for travel delays induced by everyone else's routes is not feasible. Hence, each person is planned without regard to the rest of the population's plans except in so far as their plans are reflected in the estimated travel times. The first set of paths thus places most drivers on the small set of links with short travel times – everyone takes freeways as much as possible.

The "game" alluded to in the Introduction is to find a shortest path in the context of the choices made by the rest of the population. Here the context provided by other people's paths is summarized in the travel time on each link of the network, which in turn is mostly controlled by the presence or absence of jams.

The traffic simulator provides estimates of travel times for a given set of plans. If it were possible reliably to estimate these directly from the plan set, there would be no need for the traffic simulation. TRANSIMS requires a traffic model that can produce good travel time estimates while updating the entire region in much less than real time. (See Figure 1 for an idea of both the level of detail and the overall scale.) For Portland, Oregon, with population 1.6 million, this requires roughly 10^9 vehicle-seconds per 24 hours, or 10^4 vehicle-seconds per second, of simulated time.

Solving The Game

A perturbative approach to solving this problem seems reasonable, since any one person's travel hardly affects overall travel times, but finding an approximate solution for starting is difficult. Direct calculation is infeasible. TRANSIMS uses iterative approximation. In the first approximation, everyone makes plans as though they were alone on the network. In succeeding approximations, a gradually diminishing fraction of the population is allowed to replan based on the travel times computed from the previous set of plans.

It is tempting to think of each iterate as a new "day" in which people choose a new path based on the traffic they experienced the previous day. This analogy is appropriate in some circumstances, but should not be pushed too far. The iterative process is really a machine learning algorithm for solving a many-player game. For instance, a driver choosing a route in the real world typically does not have complete information about the state of the network on the previous day – only observations about links used on the previous day and only for the specific times at which the driver was present.

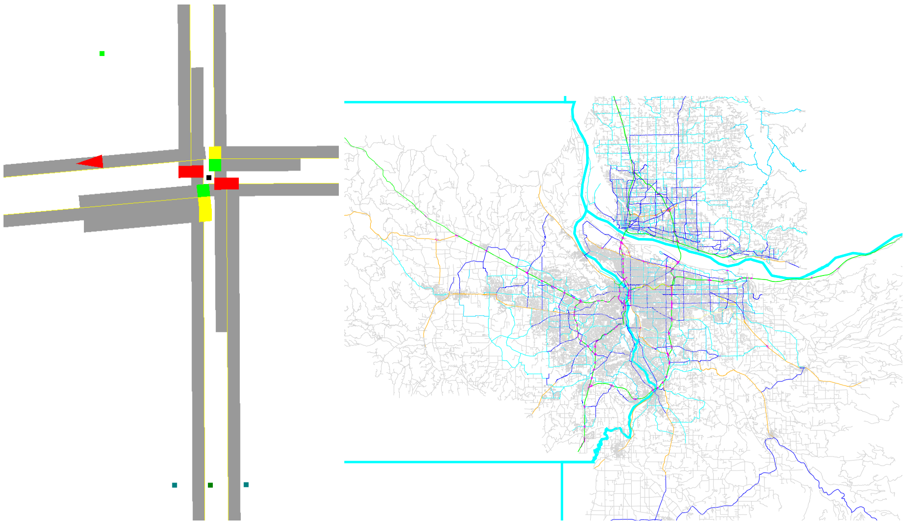

FIGURE 1. Detailed and overall views of the Portland, Oregon network. The simulation includes features such as pocket lanes, traffic signals, transit stops, activity locations, and parking lots. The full network contains about 120,000 edges and 100,000 vertices, all resolved to this level of detail.

The dynamics of this iterative process are more intricate than they at first appear. Many intuitively obvious shortcuts do not work well. For example, allowing the entire population to replan with the most recent travel time estimates can introduce periodic orbits in the state space of all plans. Consider the simplest case of two alternative routes, in which all the traffic oscillates between the two routes. Replanning a diminishing fraction of the population damps these oscillations. The optimal replanning fraction is not known, nor is it even clear what qualities of the network it depends on. Another example of the intricacy of the dynamics is that biasing the selection of those individuals allowed to replan towards those who travel on the most congested links does not seem to improve convergence.

Finally, once the iterative process has found an approximate solution for travel times given activity patterns, travelers can change unrealistic patterns, choose new times for specific activities, or change the locations at which activities are conducted. If the changes at this stage are large, it too can be done iteratively.

An Agent-Based Simulation

It has become almost required in some circles to declare simulations "agent-based". TRANSIMS is agent-based, but it involves an atypical notion of agency. An individual, identified by number, is variously represented as a set of demographics, a set of activities, or a route through the transportation network. Similarly, the route planner and the traffic simulator use radically different representations of the same section of roadway. The agent is not to be found in any of these representations, but rather in the whole collection of them. Moreover, since many of the attributes are the results of simulation, the agents are not predefined, but are endowed with important attributes through their interactions with others.

WHY SIMULATION?

Simulation provides the *only* way to determine the logical entailments of a given model on this scale. In many instances simulation of a particular model is provably more efficient than any other form of calculation. For example, consider the evolution of a cellular automaton. One way to determine the configuration at time t given an initial configuration at time t_0 is to simulate the rules. It can be shown that for a large class of update rules there is no way to determine the configuration at t in fewer steps. Even an analytical technique such as renormalization does not help, since the effective diameter of the rule grows larger and the rule table grows exponentially larger with every attempt to lengthen the time step. Approximation methods may give overall properties of the solution (an invariant distribution, for example), but these are not specific enough for our purposes. It is here that the lack of symmetries in real networks wreaks havoc with attempts to simplify solutions.

Some of these hardness results are most applicable to discrete models. But traffic is inherently discrete, and cannot be forced into a continuous description without losing the identity of the (distinguishable) travelers. Furthermore, determining the maximum flow through a network is the canonical example of a problem that has very different solutions over the integers (discrete flow) or the reals (continuous flow). Since even the lowest-level constituent model in TRANSIMS is best suited to simulation, it is likely that the entire system requires simulation.

MODELS OF TRAFFIC DYNAMICS

Historical Approaches

Because of the social and economic importance of transportation, models for traffic flow have a long history. It would not be surprising to find ideas about how to design efficient transportation networks in ancient civilizations. In the past century, several mathematical approaches have dominated the subject. A detailed review is beyond the scope of this paper. For an entry into the field, see the annotated bibliographies maintained in several places on the world-wide web. For example, `http://www.theo2.physik.uni-stuttgart.de/TrafficForum`.

Several authors have approached traffic from a macroscopic, statistical mechanical point of view. The most well-known among them are Prigogine, Herman, and Montroll. [4] [5][1] This approach develops a master equation for the distribution of speeds.

---

[1] It is interesting to note that one of these authors, R. Herman, was not content with theorizing but also conducted experiments with cars and buses on a racetrack.

Similarly, it has been tempting to use analogies to fluid dynamics, reasoning that smoothly flowing traffic is like laminar flow and jams are a form of turbulence. This approach was primarily codified by Lighthill and Whitham [6] who used an analogue of the Navier-Stokes equation to study the kinematics of density waves in traffic, including phenomena such as the upstream propagation of jams.

Moreover, though not strictly in the category of traffic models, there has been substantial effort invested in direct solution of the "traffic assignment" problem. Typically, these solutions use simple functions for travel time on a link as a function of density. A text illustrating the state of the art has recently been published by Sheffi. [7]

Recently, microscopic models of traffic have become popular.[8] These focus on vehicles as discrete particles rather than on a continuous velocity or density field. They may take the form of ordinary differential equations governing the speed of each vehicle. Nagel and Schreckenberg[9] introduced a model in which both space and time are discrete and the dynamics is represented by a cellular automaton (CA). Because of its simple update scheme, they were able to use the CA model to simulate traffic on the entire German autobahn network. Nagatani has also used a similar approach for studying phase transitions.

Both macroscopic and microscopic results have been important in the design of more efficient transportation systems. For example, an understanding of critical phenomena near the transition from free flow to jammed traffic suggests that designing transportation systems to operate near their critical point is very dangerous. Such designs are likely to produce very long-range (spatial) correlations in fluctuations, which translate into system-wide jams. Detailed models of driver behavior are also important for small-scale studies such as designing the entrance to a heavily used location, or the optimal signal timing for a network of a few intersections.

Nevertheless, these analyses are typically for long, isolated stretches of road – often, indeed, single lane roads. This is in the best reductionist tradition: break the world into its simplest components, find analytic solutions for those components, and let the engineers worry about how to put the pieces back together. A quick glance at the stated goals of the TRANSIMS project shows that we must proceed far beyond this idealized world. We must construct transportation networks consisting of hundreds of thousands of elements, each of which exhibits non-trivial behavior such as phase transitions.

There are three distinct hurdles to overcome in this effort:

1. The first is one of scale. A typical urban population of 10^6 people moves on a network with 10^5 intersections and a similar number of edges. This problem inhabits a kind of no-man's land, neither large enough to treat with statistical mechanics nor small enough to treat analytically. Roughly speaking, neither high-temperature nor low-temperature expansions are appropriate when $1 \ll log(N) \approx 6 \ll log(Avogadro's\ number)$.
2. We expect specific details of the input data to have important consequences. (Else we could not show the effects of changes in the infrastructure.) We cannot make any simplifying assumptions about homogeneity, isotropy, or any other kind of symmetry in the boundary conditions.
3. The particles in this case are not indistinguishable. Even with the drastic simplifications made by TRANSIMS – representing an individual as "economic man"

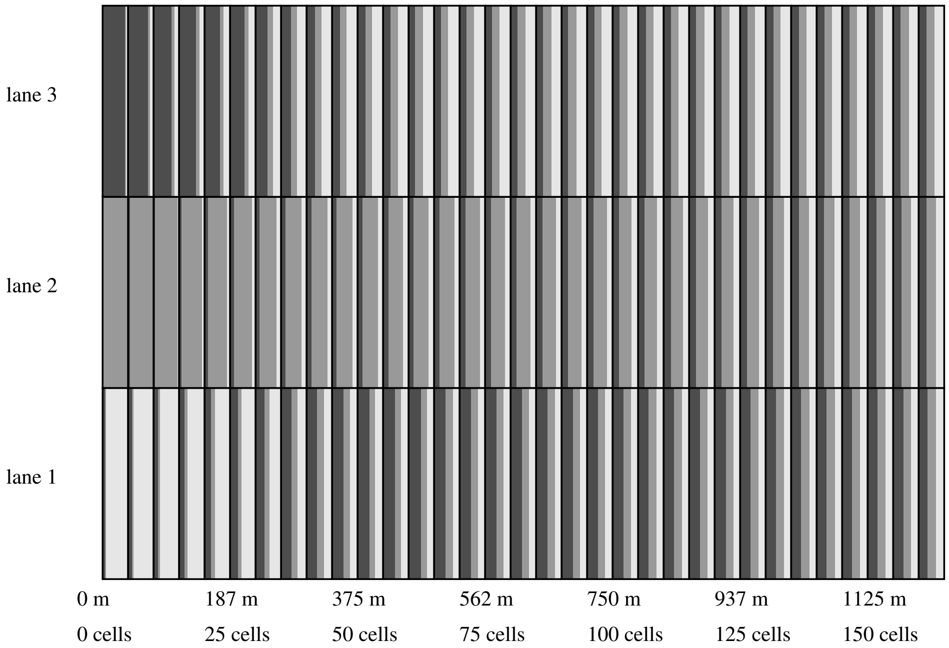

FIGURE 2. A simulated weaving region. Traffic enters from thre right and makes one of three choices at the intersection on the left. In each 37.5 meter segment of the roadway is shown, by lane, the fraction of vehicles in that lane intending to make each turn. The flow rate is such that not all vehicles can merge into the correct lane by the time they reach the intersection. Lane changing logic in the traffic model can be calibrated against observations of actual weaving regions.

with a few demographic variables and providing every driver with the same driving logic – different people can make different choices based on demographics and drivers behave differently in different traffic contexts. Worse yet, not only are these particles distinguishable, they have intentions and need not obey any conservation laws. Indeed, the introduction of drivers following planned routes into a simulation must be done carefully for the simulations to remain consistent with observations. Figure 2 shows the kinds of calibration required in this case.

TRANSIMS uses a CA traffic model modified to represent better multi-lane traffic, network use restrictions (e.g. turning movement or lane use restrictions) and plan-following behavior. Rather than turning bits on and off, a small set of rules emulating the CA rules moves pointers to individual vehicles around a grid.

The driver logic in TRANSIMS is as follows:

Choose a lane
- Pass slow vehicles if there is a gap available
- Move toward desired lane for upcoming turn

Choose speed
- Accelerate until the vehicle is at the speed limit or within range of an obstacle (other vehicles, the end of a lane, traffic signals, etc.).

– Decelerate as required to avoid obstacles.
– Decelerate to provide more opportunity for lane changing if required.
– Decelerate at random occasionally.

This model allows large decelerations in order to remain collision-free. Animation of the traffic may thus look unphysical, but the aspects important for jam dynamics are all correct.

Other elements of the simulation besides drivers must have rules. In particular, traffic signals must provide realistic flow rates. The data collection required to time correctly every signal in the network is too burdensome for most urban planning organizations. Instead, we have implemented adaptive signals that change their timing to accomodate demand. As with animations of the vehicles, close inspection of the traffic signals may reveal behavior that would be unlikely for real signals, but as long as they support the correct rates of flow through an intersection, these details are not crucial.

There are many epicycles introduced on top of these basic rules to handle parking, vehicles that cannot get in the correct lane for a turn because of congestion, priority at intersections, and so on. We have also introduced a few vehicle-specific parameters in circumstances where it seemed necessary to capture important behaviors. For example, trucks and other heavy vehicles have a grade-sensitive acceleration, rather than the uniform 1 $cell/timestep^2$ of other vehicles.

Contact with Experiment and Observation

There is ample observational data for comparison with these models. The quantities of interest are flow, density, and speed. Of these, only two are independent. Density is a natural parameter in most models, and the graph of flow vs. density is often called the "fundamental diagram" of traffic flow. Because of observational difficulties determining density, flow vs. speed is more readily available experimentally. The left side of Figure 3, taken from the Highway Capacity Manual, is representative of the data available. Analogous data from a calibration test of TRANSIMS is shown on the right side of Figure 3.

The modified CA model is up to the challenge of TRANSIMS. It reproduces all the necessary aspects of travel time as a function of demand, and does so quickly and efficiently. Other models are not as well suited to our needs. Analytic models often misplace traffic jams, leading to important differences in travel time estimates. For instance, jams may occur on links where demand is above capacity, but they typically occur at the entrance to those links and for the most part spill back onto the incoming links. Few queueing theory models capture this spillback correctly. Models that try to parameterize driver psychology, or incorporate extremely detailed physical models of car-roadway interaction, tend to run too slowly and in effect only add nuisance parameters.

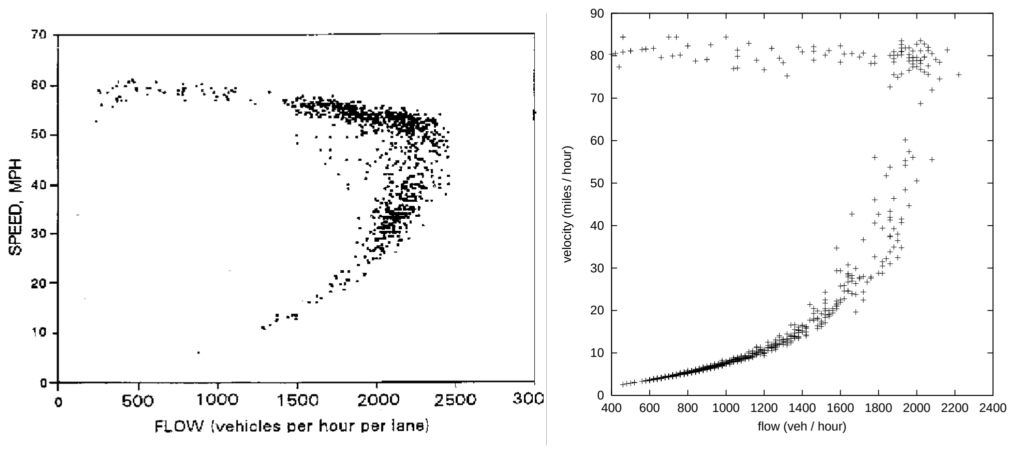

FIGURE 3. Observed[10] (left) and simulated (right) traffic flow characteristics. The y-axis scales differ because the speed limits in the two cases were different.

EXTRACTING MEANING

How are we to interpret the results of TRANSIMS? Ironically, although the defining characteristic of the simulation is that it is individual-based, we do not expect the results to be correct for individuals – I cannot find a specific synthetic "me" in the synthetic population. While the demographics of synthetic people living in specific areas are matched with the census, there is no explicit constraint forcing the synthetic people to work or play in the right locations at the right time.[2] Our assumption is that self-consistent travel time estimates together with cost-minimizing behavior supply the necessary constraints.

Of course, with every different set of initial conditions or pseudo-random number sequences, we obtain slightly different results on microscopic scales. This is no different from the world we are trying to model, in which each day brings a slightly different travel experience. Instead, we must consider macroscopic quantities or look at coarse-grained pictures of the outcomes. It is not clear, though, exactly how coarse our pictures must be. We can reliably say that traffic jams are likely to occur in particular locations and times. We can probably estimate the distribution of demographics for those caught in the jams. We probably cannot identify whether a specific individual becomes entangled in a jam. In the case of fluid dynamical simulations, the question does not arise. No one is particularly interested in which molecule of water is part of a shock wave, just where the shock is.

---

[2] This is an important feature of this approach – otherwise, problems of privacy immediately arise.

Not all traffic engineers understand this variability. Some would consider the entire TRANSIMS methodology invalid unless it reproduces traffic counts measured with a pressure hose on a rainy Tuesday, the 12th of May. But even a more generous interpretation of “valid” must provide some comparison of TRANSIMS results with daily variability. Is TRANSIMS invalid if it is outside a 3-sigma probability band on 10% of the available traffic counts? Is daily variability in traffic flow an appropriate determinant of variability to be expected in a simulation? What if the transportation system has multiple basins – sets of microscopically different initial conditions that produce macroscopically similar results? Is there any reason to expect that daily variability would explore the different basins?

It is useful in these circumstances to have an alternative method of determining the equivalence of two dynamical systems. All the results above attempt to compare properties of random samples of state space trajectories. It is possible to compare directly two dynamical systems represented as interaction graphs with local mappings at each vertex together with a schedule for updating the mappings. One can enumerate the possible topologies and multiplicities of state space trajectories from this representation without explicit evaluation of the trajectories themselves.[11] However, this program is not yet mature enough to apply to dynamical systems on this scale.

Sensitivity is intimately related to variability. At this stage of development for sociotechincal simulation, we are interested not so much in quantifying sensitivity to parameters and initial conditions or even classifying it as exponential or linear, but in simply *detecting* it. It is not hard to hypothesize that a certain set of parameters is important but determining its exact influence on a solution is time-consuming. What is really needed is a solution for the inverse problem: given several instances of behavior in a highly nonlinear system, determine the factors responsible.

CONCLUSIONS

Is TRANSIMS worth it? Comparing the total research and development expense for TRANSIMS, on the order of $30 million over 8 years, with the annual federal, state, and local expenditures on transportation, $129 billion in 1995[3], it clearly is. If even one construction project can be designed better using TRANSIMS, it will have paid for itself.

Does traffic depend sensitively on initial conditions? The separation between drivers increases linearly on average on long, isolated sections of roadway. Perhaps the introduction of interactions with the rest of the network introduces more sensitivity. Aggressive drivers’ willingness to risk life and limb for a few seconds’ gain on nearby traffic implies that they believe it does. Of course, those drivers who often catch up with the aggressive drivers at the next traffic signal probably believe otherwise. In any case, these observations are only relevant at a fine-grained, individual, level of resolution. Macroscopically, daily traffic jams appear in about the same locations and about the same times from day

[3] Source: US Department of Transportation, Bureau of Transportation Statistics, as quoted in “Transportation and the Economy”.

to day. The problem, once again, is that very few individuals care about the macroscopic view.

TRANSIMS is a good example, and perhaps the enabling technology, of the burgeoning field of socio-technical simulation. Is it Isaac Asimov's psychohistory, which predicts the course of societies through millennia? [12] Of course not, but it may turn out to be a significant first step.

REFERENCES

1. Barrett, C. L., and et al., R. J. B., TRANSIMS: Transportation Analysis Simulation System, Tech. Rep. LA-UR-00-1725, Los Alamos National Laboratory Unclassified Report (2001).
2. Eubank, S. G., "What Makes A Simulation Useful?", in *1999 IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, New York, 1999.
3. Barrett, C. L., Jacob, R., and Marathe, M. V., *SIAM Journal of Computing*, **30**, 809–837 (2001).
4. Prigogine, I., and Herman, R., *Kinetic Theory of Vehicular Traffic*, Elsevier, New York, 1971.
5. Chandler, R. E., Herman, R., and Montroll, E. W., *Operations Research*, **6**, 165–184 (1958).
6. Lighthill, M. J., and Whitham, G. R., *Proc. Roy. Soc A*, **229**, 317 (1955).
7. Sheffi, Y., *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, New Jersey 07632, USA, 1985, ISBN 0-13-939729-9.
8. Nagatani, T., *J. Phys. A*, **26**, L781–L787 (1993).
9. Nagel, K., and Schreckenberg, M., *Journal de Physique I France*, **2**, 2221 (1992).
10. Banks, J. H., *as reproduced in the Highway Capacity Manual, Transportation Research Record*, **1225** (1989).
11. Mortveit, H. S., and Reidys, C. M., *Discrete Mathematics*, **226**, 281–295 (2001).
12. Asimov, I., *The Foundation Trilogy*, Doubleday, Garden City, New York, 1951.