Technical report on the SEM4 traffic-flow model

Stephen Taylor

14th May 2021

0.1 Notation used

q	traffic-flow, in vehicles per hour
ρ	traffic-density across all lanes of a road-link, in vehicles per kilometre
v	traffic-speed, in kilometres per hour
C_{ab}	flow-capacity of road-link (a, b)
D_{ab}	number of lanes in road-link (a, b)
ρ_C	traffic-density when flow is at its capacity
v_0	free speed of a given road-link
T	time-gap between consecutive vehicles, in hours
ρmax	maximum traffic-density per lane (a constant); occurs in stopped traffic
$q_{w,\text{bh}}$	flow in the zone behind wavefront w
$q_{w,\text{ah}}$	flow in the zone ahead of wavefront w
$\rho_{w,\text{bh}}$	density in the zone behind wavefront w
$\rho_{w,\text{ah}}$	density in the zone ahead of wavefront w
$q^-_{ab}(t_i)$	flow leaving node a on link (a, b), at ith time-step t_i
$q^+_{ab}(t_i)$	flow entering node b on link (a, b), at ith time-step t_i
$\rho^-_{ab}(t_i)$	density leaving node a on link (a, b), at ith time-step t_i
$\rho^+_{ab}(t_i)$	density entering node b on link (a, b), at ith time-step t_i

0.2 The problem: to detect points of traffic-congestion during an evacuation

As an example, consider the four-node tree network, shown in figure 1; flow-capacities on the links are $C_{02} = 100$, $C_{12} = 1000$, and $C_{23} = 500$ vehicles per hour.

The flows q shown on the links are the steady-state values, but we are typically more interested in the flows exiting the links as averaged over the duration of the period of interest. If this duration is one hour, the time-averaged flows are $\overline{q}_{02} = 99.881$, $\overline{q}_{12} = 298.072$, and $\overline{q}_{23} = 494.589$ vehicles per hour.

0.2.1 Hydrodynamic relationship between flow, density, and speed

Let $q_l(x, t)$ be the rate of traffic-flow at time t hours at a position x kilometres along a link l from its start-node; flow is measured in vehicles per hour.

Note that for fluid-flow, we have the hydrodynamic relationship

$$q = \rho v \tag{1}$$

with fluid-flow q measured in vehicles per hour, density ρ in vehicles per kilometre, and the speed v in kilometres per hour.

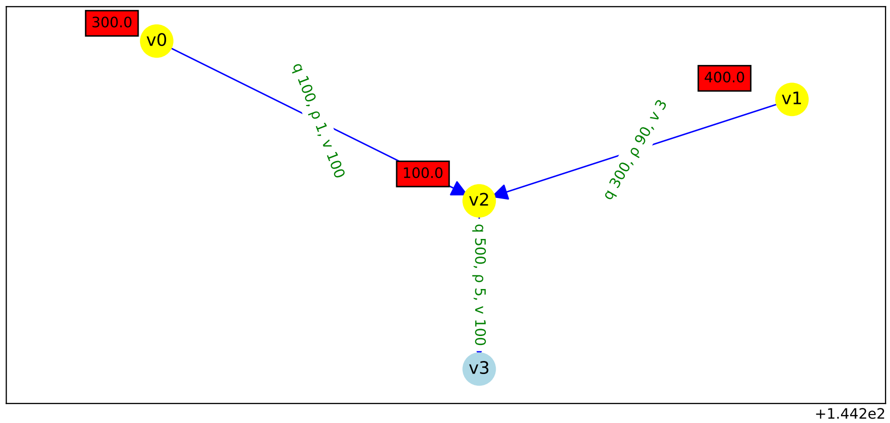

Figure 1: Four-node tree network (injection-flows e_0, e_1, and e_2 are shown in red, and exit-nodes in blue) with triangular fundamental diagram. The resulting flows q_{rs}, densities ρ_{rs}, and speeds v_{rs} are shown on the links (r, s).

0.2.2 Inputs for a scenario

An evacuation scenario is defined by a road-network, the properties of the network's road-links, a partition of the network's nodes into injection- and exit-nodes, the inflow of external traffic into the network at each injection-node, and a splitting of inflow at each injection-node into subflows that are directed to specific exit-nodes.

The SEM4 model also has an input the duration (in hours) of the time-period of interest, while the simpler SEM3 model seeks the steady-state flows and therefore assumes that all density-variations (see below) have propagated out of the network.

Solutions to the SEM3 model have proved to be of little practical interest, and they can be obtained in any case using SEM4 by setting the duration of the period of interest to be sufficiently long that all wavefronts representing traffic-flow variations shall have exited the network; we shall not consider SEM3 further.

The road-network is given as a graph $G = (V, A)$ with a set V of nodes and set A of directed arcs. The network is embedded in a geographical space containing features occupied by humans, such as dwellings, parks, and offices, and there is a corresponding inflow-rate e_r into each injection-node r from proximate such features. In the case of evacuation of population to escape a bushfire, the inflow is proportional to the fleeing population in the vicinity of the node and therefore e_r is a non-negative quantity. Note that typically, many injection-nodes r will have an inflow-rate of $e_r = 0$ vehicles per hour.

Each link has a given flow-capacity ρ_C, a *free speed* v_0 corresponding to the speed at zero density (the free speed reflects road-surface quality, slopes, etc.), the number of lanes D, and the direction of permitted travel (both directions might be permitted).

Let S be the subset of injection-nodes and T be the subset of exit-nodes; every node is either an injection-node or an exit-node, but not both, which can be expressed by $V = S \cup T$ and $S \cap T = \emptyset$.

The total traffic-inflow that enters the network at an injection-node is subdivided into subflows, each of which is directed to a given exit-node.

0.2.3 The outputs

The desired output of running a model on a scenario gives, at a minimum, those injection-nodes where the specified inflow is unable to enter the network as a result of congestion.

Another desired output might be the flow in each link averaged over the duration of the time-period of interest; this is calculated as the cumulative number of vehicles to have exited the link by divided by the duration.

0.3 The SEM4 model in overview

The SEM4 traffic-flow model is macroscopic, in that the traffic's flow q, density ρ, and speed v vary across space and time as dynamic fields, rather than being properties of individual vehicles as in a microscopic model; thus different types of vehicle and driver, and phenomena such as lane-changes, are abstracted away.

Before the model proper is run, the routes that traffic will take from injection-nodes to exit-nodes are determined; traffic might always take its shortest-path route to its assigned exit-node, for example. Once the routes have been determined, the graph G is restricted to the subgraph induced by the links that appear in the routes, and all other links are dropped; we assume the resulting subgraph is a directed acyclic graph.

The simplifying assumption is made (which implies a Lighthill-Whitham-Richards model [3]) that flow is everywhere and always a static function of the density, i.e.

$$q(x, t) = q_f(\rho(x, t)). \tag{2}$$

This implies that traffic instantaneously assumes (i.e. accelerates to) the value of flow that corresponds to its density, in other words that flow is always in local equilibrium with respect to the the density. The function $q_f(\rho)$ is known as the *fundamental diagram*; an example of a piecewise-linear "triangular" fundamental diagram is shown in figure 2.

Every link in the network is partitioned into a finite number of spatio-temporal regions called "zones," within each of which it's assumed that all traffic has the same flow and speed (and hence density). The spatio-temporal extents of the zones are defined by a set of moving boundaries between zones, called

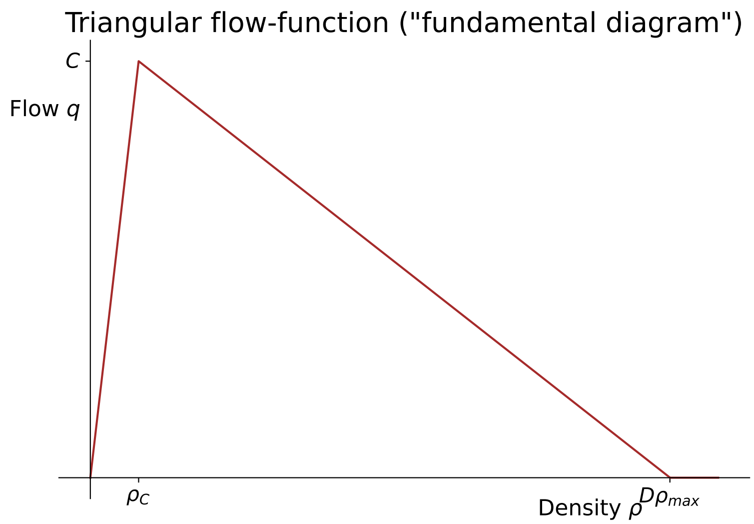

Figure 2: Flow q_{rs} as a piecewise-linear "triangular" function of ρ. The lefthand branch with positive gradient corresponds to free traffic, and the righthand branch with negative gradient to congested traffic.

"wavefronts." Each wavefront can propagate along a link, transit in modified form to one or more neighbouring links, disappear between two other converging wavefronts, or exit the network at an exit-node.

The model is initialised at time $t = 0$ with wavefronts at the injection-nodes, corresponding to the inflows of traffic at those nodes. The duration of the next time-step is calculated as the amount of time that will elapse before the soonest intersection of a wavefront either with another wavefront or with the node at the end of its link. The positions of all wavefronts are updated to those at the next time-step; thus the set of (continuous) state-variables $q(t_{i+1})$ and $\rho(t_{i+1})$ at the $(i+1)-th$ time-step for each wavefront is a function of the state-variables at the i-th time-step for both the same wavefront and neighbouring wavefronts.

Each traffic-subflow entering the network at an injection-node is directed to a specified exit-node; this traffic might be directed to take its shortest-path route to that exit-node, or to take another route, for example that given by a minimum spanning-tree on the network.

For every link (a, b) and every time-step t_i, we keep track of the values of both the flow q^-_{ab} leaving node a on (a, b), and the flow q^+_{ab} entering node b on (a, b). The densities $\{\rho^-_{pa}\}$ and $\{\rho^+_{pa}\}$ are similarly updated.

0.4 Properties of the triangular fundamental diagram

This section follows explanations given in [1] and [4].

0.4.1 The continuity equation

We present one of the ways to derive the continuity equation for traffic-flow within a road-link. For a road or link being traversed in a given direction, let x be the distance along the link from the origin-node, and let t be time; the quantities q, ρ, and v are functions of x and t. Let $N(x,t)$ be the number of cars that have passed point x by time t. It is seen that

$$\frac{\partial N(x,t)}{\partial t} = q(x,t)$$

and

$$\frac{\partial N(x,t)}{\partial x} = -\rho(x,t)$$

which implies that if N is twice continuously differentiable, then

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0. \tag{3}$$

0.4.2 Propagation of density-variations

Substituting (2) into (3) and applying the chain rule gives

$$\frac{\partial \rho}{\partial t} + \frac{dq_f}{d\rho}\frac{\partial \rho}{\partial x} = 0. \tag{4}$$

The equation

$$\rho(x,t) = \rho_0(x - ct)$$

describes a density-wave travelling uniformly with velocity c, where $\rho_0(x) = \rho(x,0)$ defines the initial density-distribution. Let $\rho_0'(x)$ be the derivative of $\rho_0(x)$; application of the chain rule yields

$$\frac{\partial \rho}{\partial t} = -c\rho_0'(x-ct) \text{ and } \frac{\partial \rho}{\partial x} = \rho_0'(x-ct).$$

Substitution of these into (4) gives

$$-c\rho_0'(x-ct) + \frac{dq_f}{d\rho}\rho_0'(x-ct) = 0,$$

which holds for all x and t; therefore we obtain

$$c(\rho) = \frac{dq_f}{d\rho} \tag{5}$$

which states that the propagation-speed of density-variations is equal to the gradient of the fundamental diagram.

0.4.3 Triangular fundamental diagram: traffic-flow as a non-monotonic function of density

The triangular fundamental diagram consists of two piecewise-linear branches; the leftmost branch (low densities) corresponds to *free traffic* and the rightmost branch (high densities) to *congested traffic*. Here the densities ρ and ρ_C are across all lanes of the given road-link (differing from [4] where ρ_C is density *per lane* at capacity).

$$q_f(\rho) \;=\; \begin{cases} v_0\rho, & 0 \leq \rho \leq \rho_C; \\ \frac{D}{T}\left(1 - \frac{\rho}{D\rho_{\text{max}}}\right), & \rho_C < \rho \leq D\rho_{\text{max}} \end{cases} \tag{6}$$

where ρ_C is the density that gives capacity (i.e. maximum) flow, v_0 is the road-link's free speed, T is the time-gap between consecutive vehicles, D is the number of lanes in the link, and ρ_{max} is the maximum density per lane (which occurs in stopped traffic and can be measured, and a value for which [4] gives as 120 vehicles per kilometre on a highway or city-road).

The inputs include the link-capacity C and free speed v_0, from which is calculated $\rho_C = C/v_0$.

At capacity we have $C = \frac{D}{T}\left(1 - \frac{\rho_C}{D\rho_{\text{max}}}\right)$, from which is obtained

$$T = \frac{D}{C} - \frac{1}{v_0\rho_{\text{max}}}.$$

0.4.4 Special properties of the triangular fundamental diagram

From (5) the propagation-speed of density-variations is equal to the slope of the fundamental diagram, which in free traffic is given by (6) as

$$c_{\text{free}} = \frac{dq_f}{d\rho}(\rho < \rho_C) = v_0.$$

In congested traffic, by contrast, we have

$$c_{\text{congested}} = \frac{dq_f}{d\rho}(\rho > \rho_C) = -\frac{1}{TD\rho_{\text{max}}},$$

meaning that congested conditions propagate backwards (upstream).

The triangular fundamental diagram is particularly advantageous when the density-profile along the length x of a link is a continuous function of x. In the present case of a discontinuous density-profile, nevertheless, the triangular fundamental diagram has the useful property that all freely-flowing traffic within the same link has identical speed, which implies (see below) that no pair of distinct downstream-propagating wavefronts will ever intersect.

0.5 Definition of the next time-step

The next time-step is defined to be the time until either the next intersection of two wavefronts on the same link, or the next arrival by a wavefront at the node ahead of it on its current link, whichever occurs soonest.

0.6 Propagating wavefronts

A wavefront is a moving "forward" boundary of a zone. A zone is contained entirely within one road-link and within it all traffic has identical flow, density, and speed; these values of flow, density, and speed apply "behind" the wavefront (i.e. in the direction opposite to the wavefront's direction of travel), as far back as the nearest link-node or other wavefront.

The SEM4 model keeps track of the positions through time of all propagating wavefronts; a wavefront's position is given by the link it's on currently, its direction of travel, and its current distance down the link.

Also stored with each wavefront w are the values of flow $q_{w,\mathrm{bh}}$ and density $\rho_{w,\mathrm{bh}}$ that characterise the zone behind w, i.e. the zone that expands at the wavefront as it propagates, and (for calculation of the propagation-speed of a wavefront, and of the adjusted flow as traffic flows through a node) the values of flow $q_{w,\mathrm{ah}}$ and density $\rho_{w,\mathrm{ah}}$ that characterise the zone ahead of the wavefront w, i.e. the zone in whose direction w is propagating and which contracts at the wavefront as it propagates.

0.6.1 Types of wavefronts

A wavefront has one of two types, based on whether it propagates in the same direction as traffic, i.e. downstream, or in the opposite direction, i.e. upstream.

A downstream-propagating wavefront is the downstream boundary of a zone of freely-flowing traffic, so we call such a wavefront a "free-front," a "downfront," or sometimes simply a "wavefront."

An upstream-propagating wavefront is the upstream boundary of a zone of congested traffic-flow, and corresponds to the upstream flow of "backed-up" traffic-conditions. Congestion results when a free-flow wavefront reaches a downstream link with insufficient capacity to carry its entire flow, forcing a reduction in flow upstream of the "shock," after which the boundary of the reduced congested flow propagates upstream. Thus an upstream-propagating wavefront is called a "shockfront," or sometimes an "up-front." An example of upstream-propagating congestion is shown in figure 3 (taken from [4]).

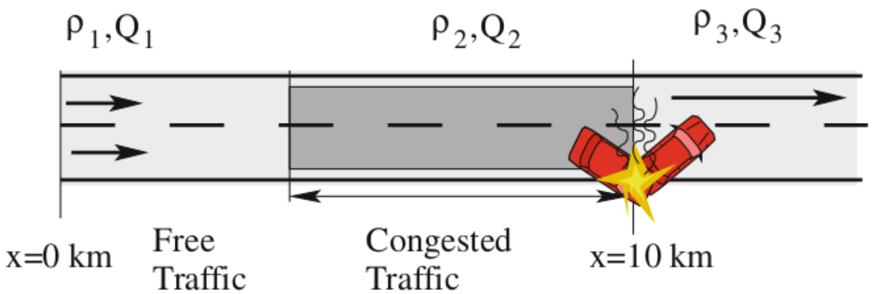

Figure 3: On a two-lane highway there is free flow $q1$ of traffic, until at position $x = 10$ km one lane is closed due to a collision; the resulting reduced (congested) flow q_2 upstream of the closure propagates upstream. Downstream of the closure, traffic flows freely.

0.6.2 Propagation-speeds of wavefronts

Consider, as shown in figure 4 (taken from [4]), a discontinuous transition between state 1 (constant flow q_1 and density ρ_1) in a short road-section 1, and state 2 (constant flow q_2 and density ρ_2) in a short road-section 2, where the two road-sections lie between $x = 0$ and $x = L$, and $x_{12}(t)$ is the location of the transition-point, or wavefront. Consider an interval of time sufficiently short that $x_{12}(t)$ does not propagate beyond the two road-sections.

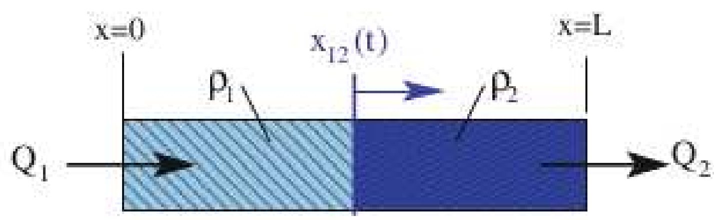

Figure 4: A wavefront at location $x_{12}(t)$, with the road-section to its left in state 1 (constant flow q_1 and density ρ_1) and the road-section to its right in state 2 (constant flow q_2 and density ρ_2).

Let N be the number of vehicles in both road-sections. From the conservation of vehicles, we have

$$\frac{dN}{dt} = q_1 - q_2.$$

But from the definition of density it is seen that

$$N = \rho_1 x_{12} + \rho_2 (L - x_{12}),$$

which implies

$$\frac{dN}{dt} = (\rho_1 - \rho_2) \frac{dx_{12}}{dt}.$$

Thus the propagation-speed of the wavefront at $x_{12}(t)$ is deduced to be

$$c_{12} \equiv \frac{dx_{12}}{dt} = \frac{q_2 - q_1}{\rho_2 - \rho_1}. \tag{7}$$

Note that a downstream-propagating wavefront or "down-front" w always has ahead of it a zone of free flow, because the down-front would have been removed if it had intersected with a shockfront. Thus the propagation-speed c_w of a down-front w is equal to

$$\frac{q_{w,\mathrm{ah}} - q_{w,\mathrm{bh}}}{\rho_{w,\mathrm{ah}} - \rho_{w,\mathrm{bh}}},$$

and noting that $\rho_{w,\mathrm{ah}} = q_{w,\mathrm{ah}}/v_0$ and $\rho_{w,\mathrm{bh}} = q_{w,\mathrm{bh}}/v_0$, we obtain

$$c_w = v_0, \tag{8}$$

with v_0 the free speed on that link.

0.7 Transition of a wavefront through a node

0.7.1 Downstream transition of a wavefront through a node

When a down-front w reaches node b at the (downstream) end of its current link (a, b), then if b is an exit-node the down-front exits the network at b.

But if b is not an exit-node, then there is a next node c on the shortest path to an exit (more generally, if traffic at any injection-node is directed to more than one exit-node then b might have more than one successor-node in G; this sub-section is specific to the case when there is only one successor c, but the method can be generalised without much difficulty). As w passes through b the flow q^-_{bc} leaving b on link (b, c) will increase by $q_{w,\mathrm{bh}} - q_{w,\mathrm{ah}}$, unless this would cause q^-_{bc} to exceed the flow-capacity C_{bc}. If it would, then a shockfront s is created at b and propagates upstream towards a, trailing behind (downstream) of it a congested flow $q_{s,\mathrm{bh}}$ equal to $C_{bc} - q^-_{bc}$, which might be zero.

The shockfront s will expand upstream into the zone behind (upstream of) w, so we set $q_{s,\mathrm{ah}} = q_{w,\mathrm{bh}}$ and $\rho_{s,\mathrm{ah}} = \rho_{w,\mathrm{bh}}$. The density $\rho_{s,\mathrm{bh}}$ behind s is obtained from the inverse function of the congestion-branch of the fundamental diagram.

From (7), the shockwave propagates with speed

$$\frac{q_{s,\mathrm{ah}} - q_{s,\mathrm{bh}}}{\rho_{s,\mathrm{ah}} - \rho_{s,\mathrm{bh}}}. \tag{9}$$

0.7.2 Upstream transition of a shockfront through a node

When an up-front (shockfront) s reaches the node a at the (upstream) end of its current link (a, b), then if a is a root-node of the directed acyclic graph G induced by the traffic-routes (these might be shortest paths), the shockfront exits the network at a.

But if a is not a root-node of G, then there is a set $P(a)$ of parent-nodes of a in G. Let the flow-reduction factor γ_a at a be

$$\gamma_a \equiv \frac{q_{s,\text{bh}}}{e_a + \sum_{p \in P(a)} q^+_{pa}},$$

where the denominator is the initial (i.e. before s reaches node a) total flow entering a, equal to the sum of the injection-flow at a and the flows entering a from its parents. It can be seen that the congested flow $q_{s,\text{bh}}$ satisfies

$$q_{s,\text{bh}} \;\; \leq \;\; q^-_{ab} \;\; \leq \;\; e_a + \sum_{p \in P(a)} q^+_{pa},$$

and therefore that $\gamma_a \leq 1$. Each of the initial flows q entering a, where $q \in \{e_a\} \cup \{q^+_{pa}\}$, is multiplied by γ_a to reflect the effect of the upstream-propagating congested flow, and the densities $\{\rho^+_{pa}\}$ are also re-calculated, from the congestion-branch of the fundamental diagram, to reflect that all flows entering a are now congested rather than free. If the injection-flow e_a has been reduced, i.e. if $e_a > 0$ and $\gamma_a < 1$, then an alert-message is generated.

The shockfront s is removed, but a modified copy of it propagates up each parent-link of node a on which the initial flow q^+_{pa} was greater than zero.

0.8 Intersections between wavefronts

Two wavefronts on the same link cannot intersect if they are both down-fronts, because by (8) they have identical propagation-speeds equal to the free speed v_0 on that link. Hence any intersection must involve at least one up-front (shockfront).

0.8.1 A shockfront intersects a free-front

If a shockfront s intersects a free-front w, the shrinking zone that was between the two intersecting wavefronts disappears, and w is likewise removed from the list of propagating wavefronts. We set $q_{s,\text{ah}} = q_{w,\text{bh}}$ and $\rho_{s,\text{ah}} = \rho_{w,\text{bh}}$, and re-calculate the shockfront's propagation-speed in accordance with (9).

0.8.2 A shockfront catches up with a slower-propagating shockfront

If a faster-propagating shockfront f catches up from behind with a slower-propagating shockfront s, once again the shrinking zone that was between the two intersecting wavefronts disappears, and s is likewise removed from the list of propagating wavefronts. We set $q_{f,\mathrm{ah}} = q_{s,\mathrm{ah}}$ and $\rho_{f,\mathrm{ah}} = \rho_{s,\mathrm{ah}}$, and recalculate the propagation-speed of shockfront f in accordance with (9).

Bibliography

[1] Stephen Boyles. *Lighthill-Whitham-Richards traffic-flow model*, lecture slides, 2018. https://sboyles.github.io/teaching/ce392d/5-lwrmodel.pdf.

[2] Wenlong Jin. *Traffic-flow Models and their Numerical Solutions*. University of California, Davis, 2000.

[3] Michael James Lighthill and Gerald Beresford Whitham. *On kinematic waves II: A theory of traffic flow on long crowded roads*. Proc. R. Soc. Lond. A, 229(1178):317–345, 1955.

[4] Martin Treiber, Arne Kesting, and Christian Thiemann. *Traffic Flow Dynamics: Data, Models, and Simulation*. Springer, Berlin, 2012.