Propagating agents with macroscopic dynamic network loading: challenges and possible solutions

Jeroen P.T. van der Gun*, Adam J. Pel, Bart van Arem

Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

Abstract

While agent-based modelling of traffic demand is gaining attention, a macroscopic dynamic network loading model may be beneficial, particularly in large-scale applications. We investigate the implications of coupling such models, with inclusion of en-route choices, for the modelling of links and the determination of turning fractions, yielding useful recommendations to help select an appropriate solution scheme of the macroscopic traffic flow theory and overcome other practical challenges specifically associated with the coupling of agent-based traffic demand and macroscopic traffic propagation.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of the Conference Program Chairs

Keywords: agent-based modelling, dynamic network loading, macroscopic traffic model, en-route choice

1. Introduction

Dynamic network loading models represent the roads and intersections of a transportation network as a graph consisting of links and nodes, and use this graph to simulate the propagation of traffic over time. The traffic demand is specified as a dynamic route vector, defining the amount of traffic departing on each route for each moment of time. With this routing information, the dynamic network loading model is able to determine the dynamic turning fractions at each node endogenously. The dynamic route vector itself may also be determined endogenously from a dynamic origin-destination matrix and a route choice model, in which case the overall model becomes a dynamic traffic assignment model.

E-mail address: j.p.t.vandergun@tudelft.nl

1877-0509 © 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.04.185
Within dynamic network loading, the propagation of traffic can be simulated by means of agents, each representing individual vehicles, such that the overall traffic pattern emerges from the interactions between vehicles. Helbing and Balietti\(^1\) show that complex traffic phenomena that can be reproduced this way. The dynamic network loading model is said to be microscopic in this case\(^2\). On the other hand, macroscopic models describe traffic at an aggregated level. As discussed by Van Wageningen-Kessels\(^3\), macroscopic models tend to have less parameters that are more easily observable, while they are also getting better at reproducing complex traffic phenomena. Furthermore, compared to microscopic approaches, the computation time of a macroscopic model can be much shorter because one does not need to move all vehicles in the network individually, every small time step. Therefore, a macroscopic model can be advantageous, especially for large-scale applications.

Notwithstanding, for modelling purposes other than driving behavior, agent-based modelling in transportation is becoming increasingly attractive. The primary reason for this is the rise of activity-based modelling, which calculates traffic demand by generating a synthetic population of agents with activity-travel patterns\(^4\). The coupling of agent-based demand modelling with microscopic dynamic network loading is relatively straightforward: an agent of the population simply maps to a vehicle agent on the road, as illustrated by e.g. Illenberger et al.\(^5\). Although one could use macroscopic dynamic network loading here by aggregating the activity-travel patterns to a dynamic origin-destination matrix or route vector, the agent is then unable to exercise any choice behavior while en-route.

This en-route choice behavior of agents is important in various applications dealing with disruptions and resilience of the transportation system, which can be divided in three categories. Firstly, rescheduling of activity-travel patterns in case of unexpected delays is a good example, where it is particularly important for the traffic dynamics near the location of the disruption\(^6\). A related example is the simulation of transportation systems during emergency conditions, including regional evacuations, as these also benefit from an activity-based approach\(^7\).

Secondly, outside of activity-based modelling, agents with en-route choices are also needed in other cases where parameters of choice behavior are sampled. In route choice under either uncertainty or risk, agent-based choice modelling allows to ensure consistent choice behavior of drivers over time. This applies regardless of whether route choice is modelled as en-route route choice or, as proposed by Gao et al.\(^7\), as en-route implementation of a pre-trip route policy choice. Thirdly, in multimodal transportation systems, public transport vehicles, e.g. buses, can be modelled as agents in the traffic network, operating the public transport service en-route as efficiently as possible. Public transport passengers can be represented as agents as well and simulated including access, egress and transfers\(^8\). This allows studying the interactions between modes, and the consequent impacts on reliability and resilience, in detail.

In this paper, we therefore explore the fundamental challenges and opportunities that need to be addressed when using agents to represent traffic demand in a macroscopic dynamic network loading model while retaining the option for agents to effectuate individual choice behavior while en-route. By doing so, we seek to combine the advantages of agent-based modelling of choice behavior and non-agent-based modelling of driving behavior.

2. Overview of the problem

Let us first start with an overview of the problem at hand, and the general principles behind a solution. In macroscopic dynamic network loading models, flows and densities are continuous variables, which can be disaggregated into so-called commodities, each of which employing specific routing behavior\(^9,10\). Hence, with our agent-based traffic demand, we will need to use commodities to link parts of the aggregate traffic flow to individual agents. However, since all flows and densities are continuous variables, an agent does not have an unambiguous location in the network. Its vehicle, represented by disaggregate continuous densities, may consist of multiple vehicle parts that are spread out over multiple links, possibly even with a gap in between. For the purposes of en-route choices, we will use the front of the vehicle, i.e. the first vehicle part, as the location of the agent, and we require that any remainder of the vehicle always follows the same route as the front. This ensures en-route choices are possible and yield a unique route for the agent.

We now proceed to define the interface between the choice modelling and the dynamic network loading. For en-route choices, the current location of the agent, as defined above, will always be a link. Here, the agent needs to choose which turning movement to take at the node downstream of this link, or choose to leave the network there. In other words, the agent needs to have an intended next turn, where leaving the network can be included in the list of
possible turns. The intended next turn may be derived from a previously chosen route, where the route choice is periodically updated, or chosen ad hoc whenever the dynamic network loading model requests to use it. Note that an intended next turn must be known before an agent actually leaves the link and the intention is made final, since, depending on the downstream congestion, the intention itself may affect the ability of an agent to leave the link.

Finally, let us look at the possible structures of the macroscopic dynamic network loading model itself. We will here focus on the Lighthill-Whitham-Richards theory of traffic flow\(^1\),\(^2\), which, if needed, can be extended to include more complex higher-order phenomena\(^3\). Then, any traffic pattern can be defined as the relation between three continuous variables: time \(t\), location \(x\) and cumulative number of vehicles \(n\)\(^4\). Given such a relation, the flow \(\frac{dn}{dt}\) and density \(-\frac{dn}{dx}\) are obtained by differentiation. We can now categorize the dynamic network loading models based on which of these three continuous variables are discretized. Furthermore, we can distinguish models based on whether they serve as solution methods to the Lighthill-Whitham-Richards equations or whether they were formulated separately. Fig. 1 lists the resulting six discretization categories with example models, indicating for each of them whether or not it involves the network loading according to the Lighthill-Whitham-Richards theory. We have included several models that advertise themselves as mesoscopic as these are based on macroscopic traffic propagation rules. In the next section, we will see that each discretization has important implications for our use of agents.

\[
\begin{align*}
\text{Shockwave analysis} & \quad \frac{dn,dt,dx}{\Delta n,dt,dx} \\
\checkmark \text{Raadsen et al.}^{14} & \quad \checkmark \text{Mahut}^{13} & \quad \checkmark \text{Mezzo}^{16,17} & \quad \checkmark \text{MATSim-DEQSim}^{18} & \quad \checkmark \text{CONTRAM}^{19} \\
\text{Link Transmission Model}^{10,20} & \quad \checkmark \text{Fastlane}^{4} & \quad \checkmark \text{Anisotropic Mesoscopic Simulation}^{22} & \quad \checkmark \text{DYNASMART}^{23} & \quad \checkmark \text{MATSim-QSim}^{24} \\
\text{Dynamic Queuing Model}^{21} & \quad \checkmark \text{Cell Transmission Model}^{9,25} & \quad \checkmark \text{Nagel and Schreckenberg}^{26} & \quad \checkmark \text{TRANSIMS}^{27} \\
\end{align*}
\]

Fig. 1. Categorization of macroscopic dynamic network loading models by discretizations. Models indicated with a checkmark serve as solution methods for Lighthill-Whitham-Richards theory.

3. Specific challenges and possible solutions

We will now list specific challenges encountered when coupling the agent-based traffic demand and the macroscopic dynamic network loading, and identify possible solutions to these challenges. The challenges below reflect the most fundamental issues we experienced ourselves while constructing such a coupled model. They relate to modelling links and determining turning fractions at nodes. Both categories are discussed below.
### 3.1. Link modelling

The first challenge for link modelling is to avoid systematic errors causing deviations from the exact solution according to traffic flow theory. A large number of models in Fig. 1 do not provide a solution method to Lighthill-Whitham-Richards theory, but are merely inspired by it. Hence, they will have systematic errors no matter how fine the discretizations are. Among the listed models with continuous \( n \), this applies only the Dynamic Queuing Model, which splits the link into a running part and a congested part, the latter having a constant density that does not depend on flow\(^{21}\). The same simplification is commonly encountered among the models with discrete \( n \), e.g. Mezzo\(^{16}\), CONTRAM\(^{19}\) and MATSim-QSim\(^{24}\). Models like Mezzo, CONTRAM, Anisotropic Mesoscopic Simulation and DYNA SMART furthermore compute the vehicle speed as a function of some average link density rather than the current local density\(^{16,19,22,23}\), with CONTRAM even lacking any explicit vehicle interaction during queuing\(^{19}\). The accuracy of the traffic propagation in a cellular automaton is limited due to its stringent discretization. Thus, while one may intuitively prefer models with discrete \( n \) for agent-based applications, these are often limited in their accuracy of traffic dynamics. In some cases, the simplifications causing the systematic errors do reduce the computational complexity of the model.

Even if one selects a model without such systematic errors, a follow-up challenge is to limit numerical diffusion that causes the numerical solution to deviate from the exact one. Although this is an issue with all applications of macroscopic dynamic network loading, it is of particular interest in agent-based applications as the vehicles of agents should diffuse as little as possible. For Eulerian-discretized models, i.e. the Cell Transmission Model, this is minimal if the ratio \( \Delta x / \Delta t \) matches the free speed, while for Lagrangian-discretized models, i.e. Fastlane, it is minimal if the ratio \( \Delta n / \Delta t \) matches the congested wave speed, outperforming the Eulerian discretization\(^1\). The variational methods have even less numerical diffusion because the traffic dynamics within the interior of links are not discretized. The continuous event-based model of Raadsen et al.\(^{14}\) is an extreme example that can eliminate numerical diffusion completely, albeit potentially at a high computational cost. Therefore using variational methods for link modelling that are free of systematic errors can be recommended to tackle these first two challenges.

For completeness, we remark that if one seeks to find an equilibrium in the choices of agents, numerical diffusion may also be somewhat helpful in achieving this as it softens the changes in travel times, which can be particularly sharp when using first-order Lighthill-Whitham-Richards traffic flow theory – hence one may not want to go at great lengths to completely eliminate numerical diffusion in such applications. On the other hand, it is theoretically better to calculate the traffic propagation without numerical diffusion and apply some kind of moving average filter when resulting travel times are used in agent choice models.

A next challenge is to prevent agents from traversing links faster than the free speed \( v \), which is one consequence of numerical diffusion that is particularly troublesome, because we do not want agents to experience impossibly low travel times. For this assessment, we use the location of an agent as defined previously, hence we investigate to what extent numerical diffusion can speed up the front of the vehicle. Naturally, this problem does not apply to models with continuous \( t \). An Eulerian-discretized model makes a speed error of up to \( \Delta t \sqrt{\left( \frac{\Delta x}{\Delta t} \right)^2 + \left( \frac{\Delta v}{\Delta t} \right)^2} \), which vanishes at \( \Delta x / \Delta t = v \) when the numerical diffusion in general is minimal. In general, models with continuous \( x \) and discrete \( t \) have an identical formula for the error, with the link length substituted for \( \Delta x \). However, for the Link Transmission Model, the overall numerical diffusion is smaller and it is computationally less costly to increase the accuracy by reducing \( \Delta t \), since one does not need to simultaneously reduce \( \Delta x \) too. Appendix A further mitigates the issue by proposing to modify the interpolation of the cumulative inflow curve such that the speed error becomes unbiased. Some of the models with discrete \( n \) keep track of the exact time within a time step a vehicle packet passes a node, making the error unbiased as well or even eliminating it, e.g. Fastlane\(^3\), Anisotropic Mesoscopic Simulation\(^{22}\) and DYNA SMART\(^{23}\).

A final, but not unimportant link modelling challenge is to accept agents to flow into the link. This may seem trivial, but it is actually problematic for models with discrete \( n \), as they need to collect \( \Delta n \) vehicles before these can enter the link as a vehicle packet. Unless \( \Delta n \leq 1 \), an individual agent is therefore unable enter a link on its own. Using such a small \( \Delta n \) is problematic for those models that discretize \( t \) in addition to \( n \), as their computational complexity then becomes similar to that of a microscopic model. An exception is MATSim-QSim which avoids this by calculating the travel speed of a vehicle packet only when it enters a link\(^{24}\), loosening adherence to Lighthill-Whitham-Richards theory. While CONTRAM reduces its computational burden by varying \( \Delta n \) over origin-
destination pairs depending on the size of their demand\textsuperscript{19}, this approach is not suitable if en-route choices need to be possible\textsuperscript{23}.

3.2. Turning fractions

At nodes, turning fractions of an incoming link are determined based on the order in which vehicles entered the incoming link, using the either the intended next turn or a past finalized turn of each corresponding agent, depending on whether or not the agent location is still on the incoming link. An important challenge is to specify this ordering of vehicles on a link. One can choose between an order of agents and an order of vehicle parts that correspond to agents. Moreover, one can choose between a strict weak order and a strict total order, that is, it may or may not be possible for multiple elements of the order to have an equivalent priority for leaving the link. For example, single-lane microscopic models use a strict total order of agents, whereas macroscopic models typically use a strict weak order of vehicle parts of which the size and commodity are known.

In our application, it seems attractive to use an order of agents, since it prevents the sequence of vehicles from becoming fuzzy due to numerical diffusion. It is also relatively easy to implement, especially in case of a strict total order of agents for models with discrete \( n \). However, this results in highly fluctuating discrete turning fractions, and in case of a strict total order even in binary turning fractions. This can cause congestion even if there is sufficient road capacity. For example, consider the situation of a motorway with an off-ramp. Highly fluctuating turning fractions may cause the demand for the off-ramp to temporarily exceed its capacity, even though there would be sufficient capacity on average. The fluctuations will decrease the average throughput of the node, and will have even larger consequences if the dynamic network loading includes a capacity drop. In principle, such fluctuations should hence be avoided, although in models with discrete \( n \) the issue may be obscured by a large \( \Delta t \). The issue may also not appear if the capacity constraint is not applied to the total link outflow, as in CONTRAM where vehicle packets never block each other at all\textsuperscript{19}, but also as in e.g. Mezzo where different outgoing turns from a link process queued vehicles independently within a specified “look-back limit”\textsuperscript{16}. Either way, this makes the queuing process difficult to interpret physically.

Therefore, we should prefer a strict weak order of vehicle parts, as common in macroscopic network loading. This means that multiple vehicle parts can enter a link simultaneously, and that they also leave the link simultaneously. For models with discrete \( n \), this implies that the composition of vehicle packets must be modified at nodes, rather than simply forwarding vehicle packets one-by-one. Unfortunately, this is unusual\textsuperscript{3}. The effects of a strict total order may be mitigated by explicitly considering individual lanes like Mahut does, at the cost of having a more complex model\textsuperscript{15}.

Another challenge is to choose a node model of an appropriate form, that is, the dynamic network loading component that computes the flows through nodes. One can chooses between an incremental node model, that starts its computation with zero flow and gradually increases the flows, or a squeezing node model, that starts with the maximum outflows permitted by the incoming links and gradually decreases these flows. A good example is the comparison of Flötteröd and Rohde\textsuperscript{28} and Tampère et al.\textsuperscript{29}, which, from the same assumptions on traffic behavior, develop an incremental algorithm and a squeezing algorithm respectively. While these produce the same results if turning fractions are constant, it is possible to adapt the turning fractions in the middle of the computation of an incremental node model. Consequently, an incremental model can ensure first-in-first-out behavior when the outflow of a link is congested, whereas a squeezing model cannot. This reduces numerical diffusion at nodes and hence an incremental model is to be preferred. Note that for efficiency, the order of vehicle parts on a link should be stored as an explicit queue in computer memory, so that the vehicle parts can be incrementally extracted from the front of the queue, unlike e.g. the formulation of Yperman\textsuperscript{10} that needs to iterate over all commodities in a link to disaggregate its outflow – which can be many in an agent-based application.

A final challenge related to turning fractions is to prevent “agent-based gridlock”. As stated before, parts of the vehicle of an agent may lag behind the location of the agent, and are required to follow the agent. Hence, this part of the flow cannot exercise en-route choices, potentially leading to a gridlock-like situation where agents are stuck on a circle of network links. Despite that the agents may choose to leave such a circle in the en-route choice process, none of the agents are able to effectuate such a choice as they are not located exactly at the downstream end of their link. One possible solution, similar to what is proposed by e.g. Charypar et al.\textsuperscript{18}, is to always accept some small
percentage of the capacity as inflow into a link, even if that means exceeding the jam density. Although this method, which partially disables spillback, can also speed up convergence in cases where one seeks an equilibrium between choices of agents, it is of course not ideal for the accuracy of traffic propagation. Hence, we believe a better solution is to relax the first-in-first-out condition such that in cases where the outflow of a link would otherwise be jammed, vehicle parts belonging to agents that have already left the link can be overtaken. This allows agents to drive out of a circle of congested links while maintaining the first-in-first-out property on the level of agents. This seems to be a balanced solution to the problem, although it does increase the computation time because it precludes grouping multiple agents into a single commodity at departure as proposed by Van der Gun et al. 8.

4. Conclusions

In this paper, we investigated whether and how macroscopic dynamic network loading models may be coupled with an agent-based representation of traffic demand, whose agents are explicitly propagated throughout the network such that they can exercise en-route choices. We conclude that this is indeed possible, also with less computational complexity than microscopic dynamic network loading models, but there are several of pitfalls that a modeler must consider. Many existing traffic flow models that are commonly used in agent-based applications turn out to have systematic inconsistencies in the modelling of links and nodes if we compare them to macroscopic traffic flow theory. Perhaps surprisingly, models which discretize the number of vehicles do not seem ideal for the task at hand, due to limitations of current node models and the need for complete vehicle packets. In particular, we find that models that discretize both time and the number of vehicles are unable to outperform microscopic models in this context. Based on the above, we recommend using a dynamic network loading model with a continuous number of vehicles. Here, the modeler can reduce numerical diffusion by avoiding discretizations of space and time. Currently, discretization of time will be necessary as Raadsen et al. 14 have not yet formulated a multi-commodity version of their continuous-time model, while their model is also restricted to triangular fundamental diagrams and can thus not handle complex traffic phenomena. The Link Transmission Model with small time steps hence seems a suitable recommendation for now. We further recommend to use a strict weak order of vehicle parts and an incremental node model to determine the turning fractions, with a relaxed first-in-first-out constraint to prevent unnatural gridlock-like situations.

Acknowledgement

This research effort is funded by the NWO-NSFC project “Optimal Multimodal Network Management for Urban Emergencies”, part of the China-Netherlands joint research program “The Application of Operations Research in Urban Transport”.

Appendix A. Enhanced interpolation for the Link Transmission Model

In this appendix we formulate an alternative interpolation procedure for the cumulative inflow curves of the Link Transmission Model that prevents the maximum speed violations reported in Section 3.1 from accumulating over links. Instead of linear interpolation within the time step, it uses linear extrapolation of both surrounding time steps if these extrapolated lines intersect within the surrounded time step. This reduces diffusion of the shape of the curve while translated to downstream nodes. The interpolation formula becomes:

\[
I'(t) = \frac{[I(t) - 2\Delta t_s]N(0, \hat{t} - 2\Delta t_o) + [I(t) + \Delta t_s]N(0, \hat{t} - \Delta t_o) - [I(t) - \Delta t_s]N(0, \hat{t}) - [I(t) + 2\Delta t_s]N(0, \hat{t})}{N(0, \hat{t}) - 2\Delta t_s} - [I(t) + \Delta t_s]N(0, \hat{t} - \Delta t_o) - [I(t) + 2\Delta t_s]N(0, \hat{t})}
\]

\[
N(0, t) = \begin{cases} 
\frac{[I(t) - \Delta t_s]N(0, \hat{t} - 2\Delta t_o) + \frac{I(t) + 2\Delta t_s}{\Delta t_o}N(0, \hat{t}) - \Delta t_o}{N(0, \hat{t}) - 2\Delta t_s} - [I(t) + \Delta t_s]N(0, \hat{t} - \Delta t_o)} & \text{if } [\hat{t} - \Delta t_o] < t \leq I'(t) \leq [\hat{t}] \\
\frac{[I(t) - \Delta t_s]N(0, \hat{t} - \Delta t_o) + \frac{I(t) + 2\Delta t_s}{\Delta t_o}N(0, \hat{t}) - \Delta t_o}{N(0, \hat{t}) - 2\Delta t_s} - [I(t) + \Delta t_s]N(0, \hat{t} - \Delta t_o)} & \text{if } [\hat{t}] - \Delta t_o < t < I'(t) \leq [\hat{t}]
\end{cases}
\]

(1)
Here \( N(0, \cdot) \) denotes the cumulative inflow curve of a link, \( \Delta t_0 \) denotes the time step of the node upstream of the link and \( \lceil \cdot \rceil \) denotes ceiling to a multiple of \( \Delta t_0 \). Now, let \( L \) be the link length, \( v \) the maximum speed, \( \Delta t \) the time step of the downstream node and \( N(L, \cdot) \) the cumulative outflow curve. Since the determination of \( N(L,t+\Delta t) \) requires the value of \( N(0,t+\Delta t_0-L/v) \) as input, \( N(0,t+\Delta t_0-L/v) + \Delta t_0 \) must already be known. In general, this is the case if \( \Delta t_0 + \Delta t \leq L/v \). The enhanced interpolation thus poses more stringent requirements on the time steps. An easy method to determine satisfactory maximum time steps for all nodes in a network is requiring \( \Delta t_0 \leq L/2v \) and \( \Delta t \leq L/2v \).

References