Macroscopic traffic model for large scale urban traffic network design

Elvira Thonhofer a,*, Toni Palau b, Andreas Kuhn b, Stefan Jakubek a, Martin Kozek a

a TU Wien, Institute of Mechanics and Mechatronics, E325/Getreidemarkt 9, 1060 Wien, Austria
b Andata Development Technology GmbH, Straubengasse 16/6, 1040 Wien, Austria

A R T I C L E   I N F O

Article history:
Received 5 June 2017
Revised 28 September 2017
Accepted 30 September 2017
Available online 10 October 2017

Keywords:
Traffic network
Traffic control
Traffic simulation
Parameter estimation

A B S T R A C T

This paper presents a flexible macroscopic traffic model that is applicable for large scale urban traffic network simulations. The proposed approach combines several advantages. First, the discretized partial differential equation of the macroscopic model is solved numerically with low computational effort. The spatial and temporal resolution of the discretization are tuned parameters to balance computational effort with model accuracy. Second, the model allows arbitrary functional forms of the fundamental diagram defined by a small number of parameters. Thereby, moving density gradients (jam fronts) are represented accurately. The model parameters are physically meaningful and can readily be estimated from measurement data. Third, two general types of intersection handling are proposed and can be combined with different merge models. The first intersection approach is a binary traffic light, similar to real traffic lights. Detailed insights concerning queue length, flow across intersections and routing decisions can be investigated. The second approach is a continuous valve-like approach that allows investigation of averaged effects and large scale interaction and feedback effects. Fourth, the proposed model scales linearly with spatial and temporal resolution as well as network size and can be partially solved in parallel to increase computational efficiency. To demonstrate the above mentioned qualities of the presented model, two realistic example situations and a comprehensive study on the scaling properties are provided.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

In the last decades urban traffic is steadily increasing, causing large jammed regions in urban traffic networks. This not only limits life quality but causes significant economic loss due to lost time in traffic and increases air pollution. As a consequence, ways to optimize traffic flow in both urban regions and along highways are an active topic of research. Methods generally aim at accurate traffic prediction and smart routing. However, methods to optimally utilize existing infrastructure, i.e. deriving control algorithms for traffic networks, are scarce. In this work, a macroscopic traffic model that can be utilized to develop and test control algorithms for intelligent network control in large scale traffic applications is proposed.

Inherent in every modeling task, a trade-off between modeling effort, model accuracy and computational effort must be balanced. The proposed model has the following characteristics:

* Corresponding author.
E-mail addresses: e.thonhofer@tuwien.ac.at, elvira.thonhofer@tuwien.ac.at (E. Thonhofer), t.palau@andata.at (T. Palau), a.kuhn@andata.at (A. Kuhn), stefan.jakubek@tuwien.ac.at (S. Jakubek), martin.kozek@tuwien.ac.at (M. Kozek).
https://doi.org/10.1016/j.simpat.2017.09.007
1569-190X/© 2017 Elsevier B.V. All rights reserved.
Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>boundary condition</td>
</tr>
<tr>
<td>FD</td>
<td>fundamental diagram</td>
</tr>
<tr>
<td>α</td>
<td>foresight coefficient</td>
</tr>
<tr>
<td>D</td>
<td>spatial computational domain</td>
</tr>
<tr>
<td>f</td>
<td>reduction coefficient</td>
</tr>
<tr>
<td>i</td>
<td>cell index</td>
</tr>
<tr>
<td>n</td>
<td>index of time step</td>
</tr>
<tr>
<td>φ</td>
<td>traffic flow in vehicles \cdot s^{-1}</td>
</tr>
<tr>
<td>ρ</td>
<td>traffic density in vehicles \cdot m^{-1}</td>
</tr>
<tr>
<td>s</td>
<td>traffic signal flow multiplicator</td>
</tr>
<tr>
<td>t, Δt</td>
<td>time, temporal increment</td>
</tr>
<tr>
<td>t_{green}</td>
<td>green time</td>
</tr>
<tr>
<td>t_{cycle}</td>
<td>cycle time</td>
</tr>
<tr>
<td>θ</td>
<td>vector of model parameters</td>
</tr>
<tr>
<td>θ</td>
<td>model parameter</td>
</tr>
<tr>
<td>τ, Υ</td>
<td>traffic regulation matrix, turn rate matrix</td>
</tr>
<tr>
<td>v</td>
<td>driving speed in m\cdot s^{-1}</td>
</tr>
<tr>
<td>x, Δx</td>
<td>length of road, spatial increment</td>
</tr>
</tbody>
</table>

- It is based on a non-linear partial differential equation (PDE) that can be solved numerically with low computational effort.
- It is capable of representing (moving) density shock waves where the accuracy is defined by the spatial and temporal resolution of the discretization.
- It is tuned by a small set of physically meaningful model parameters that can be estimated from measurement data.
- The computational effort scales linearly with temporal and spatial discretization and the network size.
- Its modularity allows an efficient work flow in real life applications, e.g. developing and testing control algorithms.

In this work a macroscopic traffic simulation approach is proposed. Traffic density is treated like a compressible fluid, hence, can be described by the transport equation, a non-linear PDE. Existing macroscopic traffic models, such as the Cell Transmission Model (CTM proposed in [1,2] are well established in highway traffic prediction and analysis. The model parameters of the CTM represent a triangular flux function specifying the forward and backward wave propagation speed and the critical and maximum traffic density. The model has been extended to allow traffic density estimation [3]. Recently, the Stochastic CTM (SCTM) has been developed to account for the stochastic nature of inflow and outflow as well as fluctuations in the model parameters [4,5]. The (extended) CTM is well suited for highway traffic state estimation and prediction. To distribute the computational load of large network simulations cloud computing techniques [6] and auto-scaling methods [7] have recently been developed. A flux function describes the relationship between traffic mass flow and density. In a different parameterization but otherwise equivalent the Fundamental Diagram (FD) [8] describes the relationship between driving speed and traffic density. It is well known from both measurement data and literature [9,10] that the FD is a general, nonlinear function. The CTM’s limitation to triangular flux functions is overcome by the traffic model presented in this work, where no limit on the shape of the flux function is imposed. Thereby, shock waves at jam fronts can be represented accurately. This is important for the application in urban traffic network simulations, where the flux functions are often better approximated by non-convex functions of arbitrary shape.

For the proposed model the set of parameters that define the functional form of the FD are the model parameters. It is important to note that the exact parameterization of the functional form is not critical. However, in this work a particular choice of parameters is proposed, such that model parameters are physically meaningful. A method to estimate the model parameters from measurement data is proposed in [11].

With regard to the computational effort the proposed macroscopic traffic model fares well compared to other concepts. For each road in the network the transport equation is discretized and solved numerically with a first-order Godunov finite volume approach. The spatial and temporal resolution are coupled via the Courant–Friedrichs–Lewy (CFL) condition and provide a tuning parameter for the trade-off between model accuracy and computational effort. The computational effort scales linearly with the number spatial cells and temporal increments. Note that the complexity of the traffic situation with regard to changing boundary conditions, evolution of jam fronts, etc. has no impact on computation time. Furthermore the proposed model can be partially solved in parallel in each time step: the discretized transport equation can be solved individually for each road within the traffic network.

Microscopic traffic models represent the other end of the spectrum with respect to modeling effort. Each vehicle is modeled and tracked individually. A driver model (car-following-model) for each vehicle is defined [12], providing the advantage that mixed-mode traffic can be specified readily. A major drawback is the large number of model parameters that are difficult to estimate and require a considerable calibration effort. Due to the high computational load, microscopic models are
generally used in off-line applications such as intersection planning [13], optimization of phase plans [14,15] and evaluation of emergency evacuations [16]. Commercial software packages exist (e. g. VISSIM [17], CORSIM [18], AIMSUN [19]) that employ a variety of more or less complex car-following-models. In this work, VISSIM traffic simulations are used as a reference solution for model validation.

Recently, mesoscopic models [20] have been successfully utilized in traffic network control [21], travel time estimation [22] and emission modeling [23]. Mesoscopic models track individual (packets of) vehicles with aggregated level of service variables, e. g. space mean speed. All vehicles belonging to a packet are assumed to occupy the head of the packet. The speed of all packets on a link is the same during a time interval \( t + \Delta t \). Furthermore, acceleration of all vehicles is assumed to be constant and uniform. These simplifications result in a relatively low computational effort for mesoscopic models. However, in the proposed approach these simplifications are dropped for the sake of improved representation of density gradients along links.

In order to model traffic in urban networks, efficient and accurate intersection handling is a key requirement. Merge models have been developed for unregulated intersections (on- and off-ramps) along highways [22,24,25] and the literature therein. These merge models are not applicable to regulated intersections per se but require a method to respect traffic signals. Therefore, two approaches to efficiently model traffic signals in the scope of a macroscopic traffic model are presented here. First, a straightforward approach is presented, where the traffic light is binary as in reality, switching from red to green and vice versa. Second, the signal is treated like a valve, allowing a fraction of the outgoing traffic flow in each time step to pass across the intersection. The ratio of green-time to cycle-time is converted to a decimal number defining the valve opening.

Lastly, to describe traffic within a traffic network completely, both topology and admissible routes must be known. Traffic network topology as well as all valid routing options are supplied by Geographic Information System (GIS) data, often readily available for European countries and North America. In this work a simulation framework is presented. An interface that allows direct import of GIS data of a selected area for modeling purposes is implemented. The traffic network is a directed graph, where edges represent all lines of one direction of roads and nodes represent intersections. Topology and routing options are fully defined by the adjacency matrix. Knowledge of the global origin-destination (OD) traffic is necessary in order to divert traffic at intersections properly. OD data is traditionally gathered by filed counts. However, automated vehicle counting and tracking systems (toll collection) are in place along highways and in many urban regions. Mobile sensing via mobile phone data and applications [26] as well as via vehicle data [27] has successfully been utilized to collect OD data. A method for automated evaluation of video data is presented in [28]. In this work a turn-rate-matrix as proposed by Coclite et al. [29] is utilized to describe the diverging of traffic flow at intersections.

Summarizing, this work is based on a traffic model presented in [11,30] to bridge the gap between the large scale applicable CTM and the microscopic traffic models. The model parameters comprise of the parameters to define the FD and the foresight coefficient, which describes the amount of responsible foresight, that the average driver exhibits. A method to estimate all model parameters from both Eulerian and Lagrangian measurement data is presented in [11]. The work presented in [11] is extended to a simulation framework that is readily applicable for on-line application in large traffic networks. An interface to directly import GIS data of the network area is provided. Traffic data provided by Eulerian and Lagrangian measurement systems can be utilized as model input. The concept of a turn-rate-matrix provides efficient OD traffic description. Additionally, two methods to model traffic signals are presented. The spatial and temporal discretization can be interpreted as tuning parameters that balance model accuracy and computational effort. Finally, the simulation can partly be solved in parallel, allowing large network simulations with high spatial and temporal resolution while maintaining on-line applicability.

This paper is organized as follows. The traffic model is introduced in Section 2. The intersection management is discussed in detail in Section 3 and the network definition is introduced in Section 4. Demonstration examples and a detailed analysis of results are presented in Section 5. Finally, conclusions are presented in Section 6.

2. Traffic model

The traffic model as well as the parameter estimation and sensitivity analysis is described in detail in [11]. A brief introduction is presented here, so that the reader can follow the extension of the approach to a modular traffic network application.

2.1. Macroscopic traffic description

The macroscopic traffic model is based on the idea that vehicular traffic along a road can be treated like a compressible fluid. The road network resembles a pipe system with valves, merges and junctions. It is based on the non-linear transport equation, a hyperbolic partial differential equation

\[
\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial \phi(\rho(x, t))}{\partial x} = 0, \quad x \in \mathbb{R}, t \geq 0
\]
where \( \rho \) represents the vehicle density in [vehicles/m], \( x \) and \( t \) represent the spatial and temporal coordinates, respectively, and \( \phi(\rho(x, t)) \) represents the (arbitrary) flux function

\[
\phi(\rho, x, t) = f(x, t, \rho) = \rho(x, t) \cdot v(\rho) \quad \text{[vehicles/s].}
\]  

(2)

The flux function \( \phi \) describes the relationship between traffic density \( \rho(x, t) \) and driving speed \( v(\rho) \) and is modeled via the so-called fundamental diagram (FD). In the approach presented in this work and in \([11,30]\), the functional form of the FD is arbitrary but it must be a single-valued function.

The parameters of the FD are the traffic model parameters and are estimated using measurement data. With the presented approach both Eulerian (stationary sensors, e.g. loop sensor data) and Lagrangian sensing (floating sensors, e.g. probe vehicle or mobile phone data) can be applied and combined. In the introduced, conventional traffic model, a FD with a simple piecewise-linear functional form, cf. Fig. 1, is advantageously chosen.

Additionally, a foresight-coefficient \( \alpha \) is introduced:

\[
\alpha = \alpha(\rho, x, t, \ldots), \quad \alpha \in [0, 1].
\]  

(3)

The foresight coefficient describes the extent to which a driver is influenced by the traffic situation ahead. A driver will drive at speed \( v \) based on the local density and to some extent on the density ahead. The driving speed in the discretized spatial domain is defined by

\[
v_i = (1 - \alpha) v(\rho_i) + \alpha v(\rho_{i+1})
\]  

(4)

where \( v(\rho_i) \) is the driving speed \( v \) based on the density \( \rho_i \) in cell \( i \in \{1, 2, \ldots, I\} \) and \( I \) denotes the last cell within the computational domain.

The foresight coefficient is estimated together with the remaining model parameters describing the FD.

The set of parameters that is used in the traffic model is collected in a vector \( \theta \)

\[
\theta = [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5 \ \theta_6 \ \alpha]
\]  

(5)

where \( \theta_i \) are the parameters defining the functional form of the FD, cf. Fig. 1, and \( \alpha \) is the foresight coefficient. Note that the individual parameters \( \theta_i \) represent the arbitrary but complete parameterization of the FD and, for the FD chosen in this work, represent slope, axis intercept and points of discontinuity.

2.2. Nomenclature

Eq. (1) is solved in a discretized spatial domain with cells along the length of a given road. A traffic network contains several discretized roads which are connected by nodes. All lanes along such a discretized road are lumped into a cell. In this work, uniform cell lengths and constant time steps are proposed, however, the model is not limited to these restrictions. The following nomenclature is introduced:

- **cells** represent the basic elements of the model, denoted by subscript \( i \) and represent discretized sections of roads
- **nodes** represent intersections, denoted by index \( w \)
- **links** are roads connecting nodes (intersections), denoted by left-superscript \( k \)
- **road networks** are roads connected by regulated or unregulated nodes (intersections)

Consequently, \( k^i \rho^n \) denotes the vehicle density in a specific cell \( i \) and a parameter vector \( k^i \theta \) on road \( k \) within the road network at a discrete time \( n \). To avoid cluttering of the formulas the left-superscript is dropped when only a single road is considered.
2.3. Link transport

Eq. (1) is approximated by the vehicle (mass) balance for each cell $i$ in discrete time

$$\rho_i^{n+1} = \rho_i^n + (\phi_i^{n+1} - \phi_i^n) \frac{\Delta t}{\Delta x}, \quad i \in \mathcal{D}, n \geq 0$$  \hspace{1cm} (6)

where $\phi_i^n$ denotes the flux exiting cell $i - 1$ (and entering cell $i$) at the discrete time $n$. The flux exiting cell $i$ is denoted $\phi_i^n$ and $\mathcal{D}$ denotes the discrete spatial domain. The temporal and spatial increments $\Delta t$ and $\Delta x$, respectively, are restricted by the Courant-Friedrichs-Lewy (CFL) condition for stability [31]

$$\frac{\Delta x}{\Delta t} \leq v_{\text{max}}.$$  \hspace{1cm} (7)

The maximum driving speed $v_{\text{max}}$, is defined by the FD, cf. Fig. 1, and valid for all cells for which a specific FD is valid. Generally, $v_{\text{max}}$ is close to the (legal) speed limit. Note that different roads within a traffic network may be modeled with different fundamental diagrams. Hence, the cell length $\Delta x$ may vary between roads within a network while the simulation time step $\Delta t$ is a global setting. General constraints apply to all cells at all time steps

$$0 \leq \rho_i^n \leq \rho_{\text{max}}$$  \hspace{1cm} (8)

which are fulfilled if both the initial conditions as well as the FD are defined properly. The maximum allowable density in a cell $\rho_{\text{max}}$ represents the capacity of the physical road that is represented by the cell and is linked to the number of (open) lanes.

The fluxes $\phi_i^n$ crossing cell-edges are defined by the model parameters, i.e., the FD. The desired driving speed at cell $i$ is defined by (4), where the foresight coefficient $\alpha$ is a weight for each velocity-term. The introduction of $\alpha$ effectively smooths the kinks of the FD in cases where two neighboring cells contain densities that correspond to different segments of the FD.

The discretized mass balance given by (6) and subject to boundary conditions (BC) and constraints given by (8) is solved sequentially, starting at the downwind end of $\mathcal{D}$. To ensure that (8) is always true, an optimization problem is formulated, that prevents cell overload (overshoot). The largest possible inflow $\phi_{i-1}^n$ into a given cell is computed while observing constraints. For clarity, the equations of the optimization problem described above are summarized

$$\max_{0 \leq f_i^n \leq 1} \phi_{i-1}^n = f_i^n \cdot (\rho_{i-1}^n v_{i-1}^n)$$  \hspace{1cm} (9)

s.t. $$\rho_i^{n+1} = \rho_i^n + (\phi_i^n - \phi_i^n) \frac{\Delta t}{\Delta x}, \quad i \in \mathcal{D}, n \geq 0$$

$$0 \leq \rho_i^n \leq \rho_{\text{max}}$$

where $f_i^n \in [0, 1]$ is the flux limiting coefficient. Any optimization algorithm may be utilized to solve the optimization problem. However, in this particular case the solution can be found analytically. The inflow $\phi_{i-1}^n$ is defined by

$$\phi_{i-1}^n = f_i^n \cdot \rho_{i-1}^n v_{i-1}^n = f_i^n \cdot \rho_{i-1}^n \left(1 - \alpha\right) v(\rho_{i-1}) + \alpha v(\rho_i)\right).$$  \hspace{1cm} (10)

Inserting the definition in (10) into (6) and applying the maximum constraint given in (8) yields

$$\rho_i^n + (f_i^n \cdot \rho_{i-1}^n v_{i-1}^n - \phi_i^n) \frac{\Delta t}{\Delta x} \leq \rho_{\text{max}}.$$  \hspace{1cm} (11)

After rearranging the above equation one arrives at

$$f_i^n \leq \frac{(\rho_{\text{max}} - \rho_i^n) \frac{\Delta x}{\Delta t} + \phi_i^n}{\rho_{i-1}^n \cdot v_{i-1}^n}.$$  \hspace{1cm} (12)

Since (9) is a maximization problem the flux limiting coefficient $f_i^n$ is

$$f_i^n = \frac{(\rho_{\text{max}} - \rho_i^n) \frac{\Delta x}{\Delta t} + \phi_i^n}{\rho_{i-1}^n \cdot v_{i-1}^n}$$  \hspace{1cm} (13)

where all terms on the right hand side are known if and only if (13) is solved sequentially for each cell, starting at the downwind end of $\mathcal{D}$.

2.4. Boundary conditions

In this work, the transport equation (1) is solved numerically. For this purpose the computational domain is discretized both in space and time. The numeric scheme defined by (6) is based on a 3-point stencil that requires attention at the edges of the computational domain, see Fig. 2.

In order to compute the updated traffic densities $q_1^+$ and $q_1^-$ at the first and last cell, respectively, of the physical domain one additional cell is required outside of the physical domain - the so-called ghost cells [32]. The density values at ghost
cells are set depending on the desired boundary behavior. Similar to the application of discretized Neumann and Dirichlet boundary conditions, absorbing boundary conditions, see [33] for an overview, can be applied. The numeric value (traffic density) for ghost cells are set at the beginning of each time step. For traffic simulation the following BC are relevant and can readily be realized

- absorbing BC: representation of an open domain outlet without reflection of information at the domain boundary
- inlet BC: given (measured) inflow or feed from preceding intersection
- outlet BC: regulated outflow at intersections

In this work, the absorbing BC is a special case of Dirichlet BC, due to the 3-pint stencil of the numeric updating scheme. Absorbing BC serve to emulate an open end of the physical domain. Traffic density will leave the computational domain without spurious reflections at the domain edge.

\[ \rho^n_{i+1} = \begin{cases} \rho^n_i - \rho^n_{i+1} & \text{restricting outflow} \\ \rho^n_{i+1} & \text{absorbing BC} \\ 0 & \text{accelerating outflow} \end{cases} \]

These effects are due to (2) and (4): The density of the outlet ghost cell affects the driving speed in the last cell of the physical domain (4) which in turn affects the outflow (2). Consequently, setting the outlet BC to \( \rho^n_i < \rho^n_{i+1} \leq \rho_{\text{max}} \) restricts the outflow, emulating a situation where a high traffic density affects the computational domain, so-called spill-back.

Specific inflow \( k^I \rho^n_{\text{data}} \) is realized at the inlet ghost cell. Traffic signals are discussed in detail in Section 3 below.

3. Intersection handling

In order to simulate traffic in networks an intersection model is necessary. Two general intersections are distinguished:

- unregulated intersection: e. g. on- and off-ramps, merge model and turn rate matrix, see Section 4.3
- regulated intersection: merge model, traffic light and turn rate matrix

Their respective components are discussed in detail below.

3.1. Merge models

The purpose of a merge model is to define how traffic density from different sources (links) move across an intersection and finally yields the input BC for the destination link.

A merge model based on the supply-demand concept as proposed by Daganzo [2] is utilized. Looking at a single node, each incoming link represents a demand, i.e., traffic density that is supposed to pass the intersection. Each outgoing link represents a supply, i.e., the capacity to accept traffic density. However, the user of the proposed traffic network simulation framework may implement different merge models, see [24] for instance. An in-depth analysis of merge models is provided e. g. in [25] and in the literature cited therein, where the use cases for different approaches are outlined.

The merge model is formulated as an optimization problem for each intersection, where a flux limiting coefficient \( k^I \) for each upwind demand must be found, cf. (9). The inflow into the first cell \( i \) of a downwind road is composed of the outflow of all connected upwind roads weighted by a flux reduction coefficient for each incoming road

\[ s^I \phi^n_0 = \sum_k k^I f^I_k s^I \phi^n_I \]

where the summation index \( k \) collects all connected incoming roads and the superscript \( s \) indicates the supply capacity of downwind links. The flux capacity of a cell is denoted \( s^I \phi^n_{0,\text{max}} \) and depends on the current traffic density in the corresponding link. Based on the unconstrained inflow \( s^I \phi^n_{0,\text{unrestricte}} \)

\[ s^I \phi^n_{0,\text{unrestricted}} = \sum_k k^I \phi^n_I \]
two cases are distinguished:

\[ k f^n_l = \begin{cases} 
1 & \text{merge model rule} \\
\text{else} & \text{for } 2 \phi^n_{0,\text{unrestricted}} \leq \phi^n_{0,\text{max}} 
\end{cases} \]

(16)

For the fairness-based approach outflows from each road are allowed to pass fractionally, based on their relative volume. Assuming that the unconstrained inflow exceeds the target cell’s capacity, the outflow of all incoming links is reduced by one flux reduction coefficient. The flux reduction coefficient is defined by

\[ k f^n_l = 2 f^n_l = \cdots = k f^n_l = \frac{2 \phi^n_{0,\text{max}}}{\phi^n_{0,\text{unrestricted}}} \]

(17)

where the superscript \( k \) denotes all incoming roads. Thereby, the relative volume of density from each link is kept constant, while the flux reduction coefficient ensures, that the maximum capacity of the downstream cell is not exceeded. Since the merge model for each intersection is formulated as an optimization problem, all incoming links are considered simultaneously. The order in which intersections within a network are processed is irrelevant, given, that the optimization problem for each node yields a feasible solution.

Note here, that other merge models can readily be utilized by replacing (17) with the appropriate rule. Different merge models can be in effect for individual intersections throughout the traffic network.

### 3.2. Turn rate matrix

At intersections, both regulated and unregulated, vehicle traffic diverges. A turn rate matrix for each intersection as proposed in [29] is utilized to relate incoming traffic to outgoing traffic and is of the general form

\[ \Upsilon^n_w = \begin{bmatrix} 
a_{1,1} & \cdots & a_{1,j} \\
\vdots & \ddots & \vdots \\
a_{j,1} & \cdots & a_{j,j} 
\end{bmatrix} \quad j \geq 1, \quad n \in [0, t] \]

(18)

where the index \( j \) represent the total number of links connecting to the node \( w \). The subscript \( w \) indicates the intersection for which the turn rate matrix is valid. A coefficient \( a_{k,l} \) of \( \Upsilon^n_w \) describes the fraction of traffic density that travels from (incoming) link \( k \) to (outgoing) link \( l \). The superscript \( n \) indicates, that the turn rate matrix is not necessarily constant over the course of the simulation time. It can readily be observed that at a particular intersection traffic with respect to origin-destination (OD) behavior during morning rush hours differs significantly from the OD behavior during evening rush hours.

The turn rate matrix \( \Upsilon^n_w \) for each intersection \( w \) can be acquired by several means. Conventional traffic counts provide accurate information, however, care must be taken as to how well the particular situation during the counting matches the representative mean for a given period of time. Both the actual time of the day as well as the weekday (weekend, work week or holiday) or even season (spring break, holiday season) may affect the result. Alternatively, video data can be evaluated (semi) automatically [28] and provide accurate turn rate information. Again, to collect a representative data set, time of day, day of the week and season must be taken into account. Ideally, several data sets are collected for characteristic large scale traffic patterns. In the future, Car2InfraStructure communication may be successfully utilized to provide accurate on-line information. In Section 5 this work the numeric values of the turn rate matrices are arbitrary, due to the generic nature of the test cases.

### 3.3. Traffic light models

Traffic lights serve to interrupt traffic at intersections to ensure that otherwise conflicting traffic flows pass sequentially. A conventional traffic light is fully defined by its phase plan, where cycle time, green time, offset with respect to a global time frame and a number of transition times are specified. However, in the framework of macroscopic traffic models for urban networks simplified traffic signal definitions are used. A traffic regulation matrix \( \mathcal{T}^n_w \) is conveniently defined similar to the turn rate matrix \( \Upsilon^n_w \)

\[ \mathcal{T}^n_w = \begin{bmatrix} 
\frac{1}{s} & \cdots & \frac{1}{s} \\
\frac{1}{s} & \ddots & \vdots \\
\frac{1}{s} & \cdots & \frac{1}{s} 
\end{bmatrix} \quad j \geq 1, \quad n \in [0, t] \]

(19)

The coefficients \( \frac{1}{s} \) of the matrix indicate the status of the traffic signal, relating a particular incoming direction \( k \) to a particular outgoing direction \( l \). Note that \( \mathcal{T}^n_w \) must be updated for each new signal phase and is therefore highly variable.

Two methods to implement traffic signals are presented below. Both concepts are illustrated and analyzed in two examples presented in Section 5.
3.3.1. Binary traffic light

The binary traffic light is essentially a conventional traffic light, albeit with minimal complexity to ensure compliance with the macroscopic traffic model. The phase plan for an intersection is defined by these variables:

- cycle time \( t_{cycle} \)
- green time for primary direction \( t_{green} \)
- offset \( t_{shift} \)

The primary direction is assigned explicit green time. The secondary (and all lower priority) directions are assigned fractions of the remaining cycle time. To keep complexity at a minimum transition phases (orange, orange flashing) are omitted.

A binary traffic signal for an incoming link is represented by an integer variable \( k_{s^n} \in \{0, 1\} \), where

\[ k_{s^n} = \begin{cases} 0 & \text{red} \\ 1 & \text{green} \end{cases} \]  

(20)

is proposed. In the above formula, \( k_{s^n} = 1 \) represents a completely open valve that allows traffic to pass unhindered. In contrast, \( k_{s^n} = 0 \) represents a completely closed valve that interrupts traffic flow.

The binary traffic light allows a detailed analysis of the interaction of phase plans of multiple intersections.

3.3.2. Continuous approximation of a traffic light

In order to investigate the overall traffic situation in a large scale network, a continuous traffic light is proposed. The general idea is to convert the relation of green time \( t_{green} \) to cycle time \( t_{cycle} \) into a decimal number:

\[ k_{s^n} = \frac{t_{green}}{t_{cycle}} \quad \forall n \in t_{cycle} \]  

(21)

The traffic light represents a valve that is partially open, i. e. it allows traffic to pass, albeit, only a fraction of the total traffic density. Thereby the average effect of a certain phase plan is represented.

The continuous traffic light approximation provides a good overview of the general traffic situation throughout large traffic networks, see Section 5 for an in-depth analysis of the continuous traffic light. It provides accurate throughput information for time scales exceeding several cycle times and provides accurate information at the domain boundaries. Thereby, coupling such a system with other simulation tools, e. g. microscopic models, is possible.

3.4. Intersection model

The general idea of the intersection model is to distribute incoming traffic towards outgoing links, taking into account the traffic signal itself, observing the turn rates and finally, achieve a compromise with respect to total amount of traffic transported with the traffic conditions at the receiving link as well as establishing compliance with the merge model. For a single intersection, the traffic flow is described by

\[ \phi^n_{w} = k^n_{w} \phi^n_{w} T^n_{w} \gamma^n_{w} \]  

(22)

and schematically depicted in Fig. 3. The vector

\[ k^n_{w} = \begin{bmatrix} k^n_{w} \phi^n_{1} \cdots k^n_{w} \phi^n_{j} \end{bmatrix}, \quad k^n_{w} \in \mathbb{R}^{1 \times j} \]  

(23)

collects all inflows from incoming links \( k \) at the intersection \( w \), cf. (10) and

\[ \phi^n_{w} = \begin{bmatrix} k^n_{w} \phi^n_{1} \cdots k^n_{w} \phi^n_{j} \end{bmatrix}, \quad k^n_{w} \in \mathbb{R}^{1 \times j} \]  

(24)
collects the outflow towards each outgoing link \( l \) at the intersection \( w \). Note that both vectors are of dimension \([1 \times j]\), where only coefficients that correspond to the index \( k \) of an incoming link or \( l \) of an outgoing link, respectively, are non-zero elements.

A general intersection is depicted in Fig. 3. A generic node, two incoming roads and two outgoing roads are shown. Ghost cells are not depicted, since they are only used to apply the respective boundary conditions. For the purpose of illustrating the concept, the traffic signal is set to be binary. That means, outflow from an upwind link is either allowed to pass fully during a green phase, or blocked completely during a red phase, cf. (20). The binary traffic signal is therefore symbolically depicted as a hard switch.

For the given example, road 1 is in a green phase and contributes the content of its last cells \( l \) to the preliminary pool of upwind road-specific outflow. The upwind road-specific outflows \( k \cdot w \phi^k_l \) are cumulated into downwind road specific inflows \( l \cdot w \phi^l \) as indicated by the green arrows. The fractions by which a single upwind outflow is diverged are defined by the turn rate matrix of that particular node.

The merge model processes the actual transport of density \( l \cdot w \phi^l \). Note that this includes both the forward transport indicated by black, dashed arrows, provided that the downwind capacity is sufficient as well as incomplete or canceled transport in case the downwind capacity is insufficient, cf. (16).

4. Network definition

The usability of the proposed model strongly depends on the modular framework. The straightforward extension of simple node and link transport to a connected network description is a key requirement. In this section, the necessary components are described in detail.

4.1. Network topology

Network topology can readily be imported from Geographic Information System (GIS) data. GIS data generally provides sufficient information on link length, number of lanes, speed limits and inter-connectivity. It is important to note that the network graph is indeed a directed graph. Lanes going in the same direction are lumped to a single link. Each node is associated with a number of incoming and outgoing links. The network graph does not necessarily contain admissible routing information. The corresponding information must be provided by the user and is conveniently encoded via turn rate matrices, see Section 3.2. In Fig. 3 the left panel shows a minimalist example of a directed graph, including routing information indicated by the dashed arrows.

4.2. Global inlets and outlets

Global inlets and outlets serve as sources and sinks for traffic density. They are commonly located at the border of the modeled traffic network. However, internal nodes may also be declared as global inlets and outlets, e.g. parking lots, garages.

Global inlets are defined for the entire simulation time. Traffic sensors provide on-line information about the current status at the domain borders of traffic networks. For off-line models, historical data or traffic census data may be used alternatively.

Prescribing Dirichlet BC at global outlets allows the application of absorbing BC, the emulation of spill-back from down-stream road sections or accelerated outflow emulating empty downstream roads. For a detailed discussion on the application of BC see Section 2.4.

4.3. Routing

The simulation framework presented in this work supports direct import of GIS data for the generation of a traffic network system including nodes, links and admissible routes, which is state-of-the-art for commercial traffic simulation frameworks. The admissible routes across a node are partly defined by the geometry of an intersection (number of incoming and outgoing links) specified by GIS data and partly defined by legal rules (one-way roads, turn restrictions) which must be gathered from inspection or local authorities. Each node is associated with a number of admissible routes that connect incoming links with outgoing links. Each route is signaled individually, similar to real live traffic signals.

For both regulated and unregulated diverges the use of a turn-rate matrix as suggested by Coclite et al. [29] is proposed, see also Section 3.2. Thereby, origin-destination flows for each incoming and outgoing direction are described. Turn rate matrices are collected from measurement data. Conventional traffic census provides relevant data. Alternatively, crowd sensing can be utilized [26,27].

5. Simulation results - efficiency and flexibility

Traffic model features are demonstrated via examples. First, an intersection with a conventional binary traffic light is presented. Then, the same intersection is shown with a continuous traffic light. Finally, both traffic light options are utilized to model a generic traffic network.
All shown examples are validated against a state-of-the-art commercial microscopic traffic simulation (VISSIM). Since VISSIM uses stochastic variation when generating traffic inflow and driver characteristics, ten simulation runs are performed for each experiment. The result data set is split in two parts, one for parameter estimation and one for validation. For parameter estimation the average of five simulation runs is used as reference data. The model parameters in VISSIM are calibrated with live data from a local test site collected for a current project of the industry partner. However, due to non-disclosure agreements, the authors are not permitted to publish details on the process of data acquisition and calibration. For the methodological validation of the proposed model and parameter estimation the exact set of driver-model parameters set in VISSIM is not critical. The proposed methods are capable of representing simulation results generated by VISSIM. The interested reader is referred to [11] for details on parameter estimation in this context. The remaining five simulation results are averaged and used in validation. First, the BC for the macroscopic simulation are extracted. Then, the macroscopic simulation results are compared against the VISSIM validation data.

Computations are performed on an i7-6700HQ CPU at 2.6GHz with Matlab 2016a on a Windows 7 64bit platform. The code is not optimized for performance, hence, a significant performance boost is to be expected from full parallelization.

5.1. Single intersection

A regular x-junction consisting of 4 one-way links serves to demonstrate the effects of binary and continuous traffic light application. Two incoming and two outgoing links are defined, each with a length of \( x = 1000 \) m and a spatial increment of \( \Delta x = 50 \) m. The configuration is depicted on the left panel in Fig. 3. At the intersection both straight-through and turning traffic is allowed, yielding four possible routes. A turn-rate matrix \( Y \) defines the corresponding turn rates, specifying which fraction of incoming traffic exits straight and which fraction turns

\[
Y^n = \begin{bmatrix}
0 & 0 & 0.7 & 0.3 \\
0 & 0 & 0.3 & 0.7 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\forall n \in [0, t_{\text{sim}}].
\] (25)

The rows of \( Y^n \) represent all incoming roads, the columns of \( Y^n \) represent all outgoing roads. Here, traffic enters only via road 1 and road 2 and exits only via road 3 and road 4. The total simulation time \( t_{\text{sim}} = 1200 \) s, with a time step of \( \Delta t = 3 \) s.

The estimated model parameters, cf. (5), are

\[
\theta = \begin{bmatrix}
-1.8 & 13.8 & 0.0598 & 0.095 & 2.3 & 0.13 & 0.295
\end{bmatrix}.
\] (26)

These parameters correspond to a speed limit of \( v_{\text{max}} = 50 \) km/h in VISSIM.

5.1.1. Intersection with binary traffic light

The signal is defined by a cycle time \( t_{\text{cycle}} \) and a green-time fraction \( 1^s \) defining how much of the cycle time is allocated to a green phase for the primary road (road 1), indicated by the superscript 1.

\[
t_{\text{cycle}} = 80[\text{s}] \quad 1^s = 0.5
\] (27)

The corresponding green-time fraction for the secondary road is \( 2^s = 1 - 1^s \). Then, green-time is

\[
1^t_{\text{green}} = 2^t_{\text{green}} = 1^s \cdot t_{\text{cycle}} = 40[\text{s}]
\] (28)

for fairness between the incoming roads. Note that the secondary road is automatically allotted red time while the primary road is in its green phase and vice versa. For clarity, the signal definition is depicted in Fig. 4.

Simulation results are shown for the primary road (road 1), the results for the secondary road are similar and differ only by the exact input due to the stochastic variation of traffic inflow in VISSIM. The left panel in Fig. 3 shows the intersection geometry. Note that the length of each road is \( x = 1000 \) m. The traffic density evolution over time for road 1 is depicted in Fig. 5. The x-axis points along the direction of the road, the y-axis represents simulation time. Darker shades of gray correspond to higher traffic densities. Additionally, a detailed comparison between the macroscopic approach presented here and the reference solution is provided. Two static sensors (blue solid lines) and three snapshots in time (magenta solid lines) indicate the data to be compared. The snapshots at given instances in time are depicted in Fig. 6 on the left panels. The snapshots at given locations over the entire simulation are shown in Fig. 6 on the right panels.

The presented model with the binary traffic light formulation captures the traffic density evolution over time well, in comparison to the microscopic traffic model utilized in a commercial state-of-the-art software package (VISSIM). A rigorous comparison for complex situations is presented in the second example, see Section 5.2.
5.1.2. Intersection with continuous traffic light approximation

The continuous traffic light formulation is based on the green-time fraction $\lambda_{\text{gt}}$ for the primary road and the corresponding green-time fraction of the secondary road. To ensure comparability, the setup of the experiment (geometry, simulation time, model parameters and input boundary conditions) is identical to the setup described above.

The traffic density evolution with a continuous traffic light over time for road 1 and 3 (primary road and arterial connection to road 3) is depicted in Fig. 7. This plot is directly comparable to Fig. 5. The obvious difference is the evolution of the jams at the traffic light. With the continuous traffic light formulation, there is no jam due to cyclic red-phases. Hence, lower traffic densities pass the intersection unhindered. As a consequence, for lower traffic densities, the jams at the junction disappear completely. Only after a certain density is exceeded in one of the incoming links a high-density region builds up and persists until the incoming traffic density falls below a certain level for a longer period of time. This effect is shown in Section 5.2.

However, on average the total throughput of traffic density is similar to the throughput of both the binary traffic light formulation and the reference solution generated by VISSIM. Fig. 8 shows cumulative sums of the input BC for road 1 and road 3 (green) generated via VISSIM, the input BC as generated with the binary traffic light formulation (gray, dotted) and the input BC as generated with the continuous traffic light formulation (blue). Note that the input density of roads 3 is shifted by 30 samples, which is exactly the number of discrete cells along the length of the road. The discrete traffic light formulation causes a hardly noticeable alternating wiggle-pattern in the cumulative sum, whereas the cumulative sum for the continuous traffic light formulation is smooth. The average slopes of all curves are similar, suggesting that the input-output behavior of the continuous traffic light formulation is equivalent to the discrete formulation.
For a better comparison, the relative error of both traffic light formulations with respect to the VISSIM reference solution is shown in Fig. 9. The reference solution is computed with a binary traffic light formulation. Hence, in Fig. 6 the cumulated sum of throughput is a line with alternating higher and lower slope, corresponding to green and red traffic light settings. As a consequence, the relative error between the simulation results and the reference solution is large during the initial phase of the simulation. A small absolute value of the cumulative sums in combination with relatively large differences due to slightly different signal timing (a 3-s simulation time step causes rounding differences compared to VISSIM, where signals change every 100 s) for the binary traffic light and a systematic difference for the continuous traffic light formulation render the initial phase irrelevant. The measure of relative error provides information about the long-term performance. Therefore, the graphs y-axis is capped to provide high resolution in areas, where a comparison of the relative error is meaningful. The graph also shows the non-stationary part of the simulation, where starting from initially empty roads, traffic density enters the computational domain and travels across it. It is important to note, that a true steady state will never be reached, due to the stochastic inflow boundary conditions. This is also illustrated in Fig. 8 where the input BC for link 1 is approximately a straight line with constant slope.

In Fig. 9 the solid lines represent the input BC at road 3. The relative error of the binary traffic light formulation (solid, gray) is below 10% after 200 s and the cumulative sum of throughput approaches the reference solution. The relative error for the continuous traffic light approximation is below 5% after 300 s.

Generally, when starting with BC conditions that are far from the approximately steady state of the traffic situation, the binary traffic light formulation requires longer simulation times before it accurately replicates throughput with respect to a reference solution. The continuous traffic light approximation represents throughput with less than 5% relative error after about 300s. The quality of the initial condition significantly improves the convergence rate for both formulations.
5.2. Multiple intersections

An arterial with four x-junctions is analyzed and represents a more realistic application scenario. The geometry of the test configuration is depicted in Fig. 10.

The route-options, traffic light settings as well as the turn-rates at each intersection are identical to the options used in Section 5.1, cf. (27), (28). The turn rate matrices are defined in analogy to (25), as indicated by the arrows in Fig. 10. The estimated parameter vector $\theta$ is identical to (26) in the previous example.

5.2.1. Multiple intersections with binary traffic lights

Simulation results are shown in Fig. 11. The traffic density over time for all roads is depicted. Darker shades of gray represent areas of high traffic density. At each junction the binary traffic light interrupts the arterial traffic, corresponding jams form and dissolve again, once the traffic light turns green.

The reference solution (average of five simulation runs) for the arterial roads generated by VISSIM is shown in Fig. 12. The proposed model yields a similar traffic density evolution over time and space.

5.2.2. Multiple intersections with continuous traffic lights

The corresponding results for the continuous traffic light are shown in Fig. 13. The temporary high-density areas at red traffic lights does not form, however, traffic is slowed down sufficiently, such that the average throughput of both configurations is equivalent.

A detailed analysis of the cumulative sums of traffic density throughputs along the main artery (roads 1, 2, 3, 4 and 5) is shown in Fig. 14.

5.2.3. Discussion of results

As seen in the previous example, the cumulated sum of density throughput with binary traffic lights is initially below the cumulated sum of the throughput with continuous traffic lights. However, as the simulation progresses, the difference...
between the binary traffic light formulation and the reference solution decreases. Results suggest, that the binary traffic light formulation takes longer to represent the ground truth quantitatively, however, it provides a high resolution of jam formation. The initial phase where the model is inaccurate can be decreased by improving the initial conditions on all links.

The relative errors of the cumulative sums for road 3 (dashed) and road 5 (solid) are shown in Fig. 15. As stated previously at the interpretation of Fig. 9 the large relative error in the beginning of the simulation is due to the fact that the absolute value of the cumulative sum is small compared to the methodical difference between the reference solution and the binary traffic light formulation and the difference due to rounding for the binary traffic light formulation. Therefore, the graphs y-axis is capped to provide high resolution in areas, where a comparison of the relative error is meaningful. The first part of the diagram represents the initial phase where, starting from empty initial conditions, traffic density travels across the computational domain. It takes 100 time steps (5 roads of 20 discrete cells each) for the first portion of traffic density to travel the entire computational domain. At this point, a steady state is generally not yet reached. Hence, the first 300 s of simulated time can safely be ignored.

The cumulative sum of throughput with the continuous traffic light approximation drops below 10% after 135 time steps, indicating that the simulation is converging to an approximately steady state. The cumulative sum of throughput with the binary traffic light drops below 10% after about 280 time steps, indicating that this formulation takes significantly longer to quantitatively resemble ground truth, however, it provides a high resolution insight on traffic jam formation almost immediately. Both traffic light formulations converge to the ground truth.
5.3. Network application

A comprehensive study on the scaling of CPU time with respect to network size is conducted and summarized in Table 1. Network grids of dimension \(1 \times 1\) (a single intersection as described in Section 5.1) up to a grid size of \(16 \times 16\) intersections, comprising of 320 nodes, 544 links with a link length of \(l = 1000\) m each and 1024 possible routes are investigated.

An exemplary 16x16 grid is shown on a road map of Manhattan,\(^1\) c.f. Fig. 16, just south of Central Park. Note, that the links in the real-live urban network are significantly shorter (30 m–100 m) than the link length used here.

Traffic enters the computational domain at all left and top nodes and leaves the domain at all right and bottom nodes. For a fair comparison, all incoming links are supplied with constant inflow, while all outgoing links allow traffic to exit unhindered.

The CPU time increases linearly with respect to the number of entities.

---

\(^1\) Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.
Fig. 15. The relative error of the cumulated sums of the continuous traffic light formulation for road 5 (solid, blue) and road 3 (dashed, blue) and of the binary traffic light formulation for road 5 (solid, gray) and road 3 (dashed gray) is shown with respect to the reference simulation results generated by VISSIM. The 5%- and 10%-error margins are shown in red (dotted, dash-dotted). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
CPU time for generic intersection models and 250 time steps, $t_{\text{end}} = 250$ s.

<table>
<thead>
<tr>
<th>Entity</th>
<th>$1 \times 1$</th>
<th>$2 \times 2$</th>
<th>$4 \times 4$</th>
<th>$8 \times 8$</th>
<th>$16 \times 16$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_{\text{nodes}}$</td>
<td>7</td>
<td>12</td>
<td>32</td>
<td>96</td>
<td>320</td>
</tr>
<tr>
<td>$n_{\text{links}}$</td>
<td>6</td>
<td>12</td>
<td>40</td>
<td>144</td>
<td>544</td>
</tr>
<tr>
<td>$n_{\text{routes}}$</td>
<td>6</td>
<td>16</td>
<td>64</td>
<td>256</td>
<td>1024</td>
</tr>
<tr>
<td>sum</td>
<td>19</td>
<td>40</td>
<td>136</td>
<td>496</td>
<td>1888</td>
</tr>
<tr>
<td>cpu time [s]</td>
<td>3.50</td>
<td>7.00</td>
<td>22.05</td>
<td>79.04</td>
<td>298.28</td>
</tr>
</tbody>
</table>

Fig. 16. Map of Manhattan [34] with an overlay of a $16 \times 16$ grid of regulated intersections (black circles with white fill). Light gray nodes indicate the full extent of the $16 \times 16$ grid.
6. Conclusion

In this work a modular macroscopic traffic model with two traffic light formulations is presented. The binary traffic light formulation provides direct insight into jam formation at red traffic lights and interactions of such jams across several links of a traffic network. A binary traffic light approximation is introduced, which provides accurate simulation results concerning the throughput across links. It is well suited to provide a general overview of the traffic situation across the network, however, does not allow a detailed analysis at intersection level. Both formulations are compared extensively via two examples.

Network geometry can readily be imported from GIS data, origin-destination traffic is efficiently handled by turn rate matrices. The model is capable of representing traffic in networks of significant size while maintaining real-time applicability. A study on generic grids shows that the model scales linearly with the number of entities (sum of nodes, links and allowable routes) in the network.

As part of a future research the proposed model must be validated with real traffic data. Then, the simulation framework can be used in the development of intelligent control algorithms for urban traffic networks.

Acknowledgment

This research was partially funded by Wirtschaftsagentur Wien (ID 1464714).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.simpat.2017.09.007.

References