Formulation and validation of a car-following model based on reinforcment learning

Fabian Hart[a], Ostap Okhrin[a,b], Martin Treiber[a,b,*]

[a] *TU Dresden*
[b] *Possible second address*

Abstract

To be written at the end

Keywords: reinforcement learning, car-following model, stochastic processes, string stability, validation, trajectory data

1. Introduction

[problem statement]

[references for state-of-the art] references RL: Farazi et al. (2020); Qu et al. (2020) references classical, ACC, stochastic CF model: Treiber et al. (2000); Treiber and Kesting (2013, 2018) references AR(1), e.g. Honerkamp (1993)

[central statement] To our knowledge, no string stable neuronal-network car-following model has been proposed that can self-learn based on generated trajectories which has the advantage of unlimited supply of training data.

In this contribution, we propose a novel reinforcement learning (RL) car-following model that is trained on leading-vehicle trajectories generated by an AR-1 process with parameters reflecting the kinematics of real leaders. We validate the trained model on experimental and naturalistic trajectory data, and on artifical speed profiles bringing the model to its limits. In all cases, the model proved to be accident free and string stable. Unlike other variants of AI models such as LSTM models, the proposed model is not completely blackbox

*Corresponding author
Email address: `Martin.treiber@tu-dresden.de` (Martin Treiber)
URL: `www.mtreiber.de` (Martin Treiber)

since the reinforcement learning reward function reflects driving style attributes such as desired time gap and speed, maximum acceleration, and comfortable deceleration.

[short textual enumeration of the sections to come]

2. Model specification

2.1. RL artitecure

2.2. Reward function

learning input (leader speed time series)

[also relate parameters to driving style attributes such as desired speed, accelerations, decelerations, desired time gap, minimum gap]

3. Model training

3.1. Synthetic trajectories

(truncated) AR(1) process of the leading speed

parameters and statistical properties (expectation, variance, auto-correlation function, typical accelerations

figure of realisation

3.2. Evaluation of the reward function

[implementation of the AR(1) process, numerical integration of the model output (acceleration?): numerical scheme, update time etc]

3.3. The reinforcement learning process

[things to look out for]

[typical figure of increasing reward over #steps, then saturation]

[number of steps, computing time]

[figure of following trajectory instance at the beginning and after saturation of the learning process]

4. Validation

The goal is not to minimize some error measure as in usual calibration/validation but to check if the driving style is safe, effecetive, and comfortable. Reference for this is the reward function

4.1. string stability

many trained RL vehicles behind the AR(1) realisation

4.2. Response to an external leading vehicle speed profile

[decribe profile with episodes of free driving, dynamic approaching, car-following, stopping, accelerating, and traffic waves]

[figure with leader and several trained RL followers]

[discussion: free: desired speed; following: desired time gap; dynamic situations: accelerations, desired and maximum decelerations, jerk; comfort: maximum accelerations, decelerations, jerk; safety: no crashs, minimum TTC; stability: string stable]

4.3. Response to experimental leaders

[describing the Napoli data]

[figure with several followers]

[cross comparison with IDM calibrated to this set] 2x2 table; rows: RL model and IDM; columns: reward function and calibration GoF (goodness-of-fit) function

4.4. Simulation of collective phenomena

[open system with speed-limit or on-ramp bottleneck (simplest vehicle dropping), increase inflow until breakdown to determine capacity, stability: no traffic waves, just congested traffic, maximum deceleration at the upstream jam front, propagation velocities]

5. Conclusion/Discussion

References

N. P. Farazi, T. Ahamed, L. Barua, B. Zou, Deep reinforcement learning and transportation research: A comprehensive review, arXiv preprint arXiv:2010.06187 (2020).

X. Qu, Y. Yu, M. Zhou, C.-T. Lin, X. Wang, Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach, Applied Energy 257 (2020) 114030.

M. Treiber, A. Hennecke, D. Helbing, Congested traffic states in empirical observations and microscopic simulations, Physical Review E 62 (2000) 1805–1824.

M. Treiber, A. Kesting, Traffic Flow Dynamics: Data, Models and Simulation, Springer, Berlin, 2013.

M. Treiber, A. Kesting, The intelligent driver model with stochasticity – new insights into traffic flow oscillations, Transportation Research Part B: Methodological 117 (2018) 613 – 623. TRB:ISTTT-22.

J. Honerkamp, Stochastic dynamical systems: concepts, numerical methods, data analysis, John Wiley & Sons, 1993.