An Agent-Based Model for Trajectory Modelling in Shared Spaces: A Combination of Expert-Based and Deep Learning Approaches

Extended Abstract

Fatema T. Johora, Hao Cheng, Jörg P. Müller and Monika Sester
fatema.tuj.johora,joerg.mueller@tu-clausthal.de; Hao.Cheng,Monika.Sester@ikg.uni-hannover.de

ABSTRACT

Realistically modelling behaviour and interaction of mixed road users (pedestrians and vehicles) in shared spaces are challenging due to the heterogeneity of transport modes and the absence of classical traffic rules. Existing models have mostly used the expert-based approach, combining symbolic modelling and reasoning paradigm with the hand-crafted encoding of the decision logic. Recently, deep learning (DL) models have been largely used to predict trajectories based on e.g. video data. Studies comparing expert-based and DL-based micro-simulation of shared spaces concerning their accuracy are missing, and so are proven methodologies for combining these approaches into a single agent-based system. In this paper, we propose and compare an expert-based and a DL model and then combine them for trajectory prediction in shared spaces. Simulation results show the combined model to outperform both pure approaches in predicting realistic and collision-free trajectories.

KEYWORDS

Mixed-traffic; Intent detection; Deep learning; Game theory

ACM Reference Format:

1 INTRODUCTION

Shared space, introduced by Monderman [8] as an alternative to classical traffic design, largely removes road signs, signals, and markings to prompt direct interaction among mixed traffic participants, guided by social protocols and negotiation. The absence of explicit traffic rules and thereby caused vagueness makes it critical to investigate safeness and traffic efficiency of shared spaces [11].

Understanding how road users behave and how their actions can be predicted is far from trivial. There is a considerable body of research to tackle these challenges. In particular, we can distinguish two classes of methodologies: expert-based approaches [3, 13, 16, 20, 23, 26, 29] and data-driven approaches [1, 5, 6, 10, 15, 18, 24]. Expert approaches involve human design crafting explicit decision rules to tackle the modelling problem [13, 26], which makes it difficult to scale up for large or new problems. Whereas, data-driven approaches can be trained by processing the data extracted from real-world situations and deriving a complex neural network structure with associated parameters or weights optimised via training [17]. These models are often black boxes, difficult to understand and explain for humans; adding human modeller’s intention to guide the models to capture specific desired patterns is difficult [14] and computational cost can also be a bottleneck [25].

To our knowledge, there are no studies that compare and analyse the strengths and weaknesses of these two types of approaches for microscopically modelling shared spaces. To address this gap, in this work, we firstly propose an expert-based model called GSFM that combines Social Force Model (SFM) and Game (G) theory and a DL model called LSTM-DBSCAN that manipulates Long Short-Term Memories (LSTM) with Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [9] for multi-agent trajectory prediction. Their accuracy (in terms of realistic behaviour) is tested on real-world shared-space scenarios, using the same evaluation metrics. Secondly, based on our empirical results and motivated by some initial studies [7, 14, 22], we proposed a combined model to hoard the collective advantages of both kinds of approaches.

2 METHODOLOGY

The prediction task is to generate realistic and collision-free future trajectories of multi-agents, mathematically, to predict the locations \( \hat{X}_i^k \) of agent \( i \in N \) for \( N \) agents at prediction time \( t \in \{ k+1, \cdots, m \} \) based on the locations \( X_i^t \) at observation time \( t \in \{ 1, \cdots, k \} \) for both expert-based and DL models. The objective is to minimise \( L(Y^k, \hat{Y}^k) \) for all agents, where \( \hat{Y} = f(X) \) and \( Y \) is the ground truth, \( f(\cdot) \) stands for the prediction models, and \( L(\cdot, \cdot) \) the loss function.

The expert-based GSFM model consists of three modules with different roles: trajectory planning, force-based modelling, and game-theoretic decision-making. GSFM is built on a BDI (Belief, Desire, Intention) platform, LightJason [2], to design and explain the control flow among the modules. The BDI controller acts as the brain of the agent to perceive the environment and activate one of these modules based on the situation. Each module triggers the controller on the completion of their task(s). The GSFM component in Fig. 1 visualises the overall structure of GSFM. The trajectory planning module computes free-flow trajectories. The force-based and game modules model interactions among agents. In GSFM, these interactions are classified into two categories: simple interaction (percept \( \rightarrow \) act) and complex interaction (percept \( \rightarrow \) choose an action among many alternatives \( \rightarrow \) act). The force-based module

---

handles simple interactions by using and extending the classical SFM and the game mode conduct complex interactions using a sequential leader-follower game, a.k.a. Stackelberg game, to guarantee collision-free trajectories explicitly. The overall process of GSFM for predicting the movement of any target agent $i$ at any time step $t$ is: $Y_{i}^{t+\Delta t} = f(Z_{t}, \frac{d\psi_{t}}{dt} + X_{i}^{t})$. Here, $Z_{t}$, $\frac{d\psi_{t}}{dt}$ and $X_{i}^{t}$ depict the input to the model, change in velocity of $i$ (measured by force/game modules), the position of $i$ in current and next time step, respectively. $Z_{t}$ contains start, predicted goal, speed profile, and minimum distance acceptance of $i$ with others, derived from the observation of $X_{i}$. More details of GSFM in [16].

The DL model LSTM-DBSCAN takes $X_{i}$ as input and outputs $\hat{Y}_{i}$. It has a mapping module for interaction pooling and an LSTM module for motion planning. The mapping module pools the interactions between the target and other neighbourhood agents at each time step. It maps the collision probability based on safety distance maintained by each other, denoted as probability density mapping (PDM) [6]. Similar to the repulsive force in SFM [13], if two agents approach each other, PDM increases exponentially. To differentiate the impact from non-group and group members (if any) on the target agent [23, 27], a density-based cluster DBSCAN [9] is incorporated to detect pedestrian groups so as to cancel out erroneous collision and relax on close interactions for group members [4]. A neighbour-hood agent is defined as a group member for the target agent if they co-exist in the same cluster over a certain duration. PDM is then reset to zero for group members. The LSTM module is used for motion planning, which takes the target agent’s coordinates and PDM as input at each observed time step to predict the distribution of the next positions [1]. The prediction process for the target agent $i$ is denoted as $\hat{Y}_{i\in N} = f(X_{i\in N}, \phi(\psi(X_{e\in N}, X_{i\in N}, j_{i\in N})))$, where $f(\ldots)$ stands for LSTM, $\phi(\ldots)$ for PDM, and $\psi(\ldots)$ for DBSCAN.

The combined model GSFM-w-LSTM of the expert and DL models is visualised in Fig.1. Its workflow is as follows: (1) GSFM and LSTM-DBSCAN predict the trajectories of respective road users by sharing the same observation. (2) The predicted trajectories of LSTM-DBSCAN are then cross-checked for collision avoidance. If the time to collision (TTC [12]) of the predicted trajectories of two users is less than one second, the prediction is considered as collided. (3) If the predicted trajectories of LSTM-DBSCAN are collision-free then these trajectories will be executed, otherwise the predicted trajectories of GSFM will be selected to execute.

4 ACKNOWLEDGEMENTS

This research is funded by the German Research Foundation (DFG) through the Research Training Group SocialCars (GRK 1931).
REFERENCES


