

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:

Fatema T. Johora, Hao Cheng, Jörg P. Müller and Monika Sester. 2020. An Agent-Based Model for Trajectory Modelling in Shared Spaces: A Combination of Expert-Based and Deep Learning Approaches. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020*, IFAAMAS, 3 pages.

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{Y}_i^t of agent $i \in N$ for N agents at prediction time $t \in \{k+1, \dots, m\}$ based on the locations X_i^t at observation time $t \in \{1, \dots, k\}$ for both expert-based and DL models. The objective is to minimise $L(Y, \hat{Y})$ 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

* equal contribution by F.T. Johora and H. Cheng.

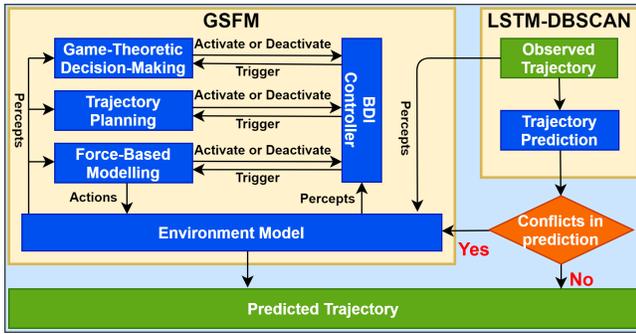


Figure 1: The structure of the combined model GSFM-w-LSTM of GSFM and LSTM-DBSCAN. A conflict checking box selects the final prediction from either of the two models.

handles simple interactions by using and extending the classical SFM and the game module 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 in any time step t is: $\hat{Y}_i^{t+\Delta t} = f(Z_i, (\frac{d\vec{v}_i}{dt} + X_i^t))$. Here, Z_i , $\frac{d\vec{v}_i}{dt}$, X_i^t , and $Y^{t+\Delta t}$ 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_i 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 neighbourhood 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 $\hat{Y}_{i \in N} = f(X_{i \in N}, \phi(\psi(X_{i \in N}, X_{j \in N, j \neq i})))$, where $f(\cdot, \cdot)$ stands for LSTM, $\phi(\cdot)$ for PDM, and $\psi(\cdot, \cdot)$ 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.

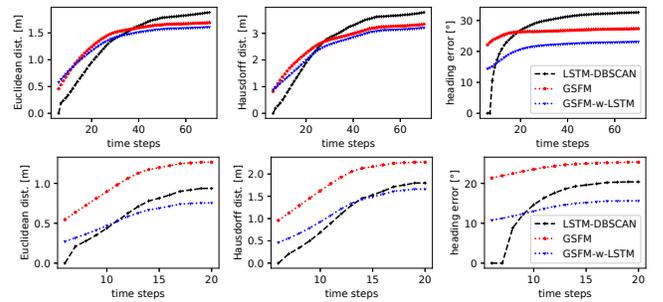


Figure 2: The performance of each model, validated on HBS [21] (row 1) and DUT [28] (row 2) data sets.

3 EXPERIMENTS AND RESULTS

Data sets: We use the train station data set (HBS) from Germany [21] and the DUT data set from a university campus in China [28]. HBS was recorded in a street with pedestrian crossing and DUT was recorded in a roundabout and an intersection. 89 scenarios that involve interactions between pedestrians and vehicles were manually extracted from the data sets for evaluating and the rest of the data sets are used for calibrating/training the proposed models.

Evaluation Metrics: The average Euclidean distance error measures the aligned error for each time step and we report the value averaged over the path [1, 10]. For the accumulated error, we use Hausdorff distance to measure the largest distance from the set of the predicted positions of a trajectory to the set of true positions [19]. Heading (from the previous position to the next position) error measures the pairwise absolute heading difference over all positions between the predicted and ground truth trajectories.

Results: In general, as the time step increases, the performance of all models decreases on both data sets, shown by Fig. 2. While, the errors of GSFM-w-LSTM increase with a much slower speed compared with GSFM and, especially, LSTM-DBSCAN. In comparison with GSFM, GSFM-w-LSTM makes smaller errors by all evaluation metrics and shows a similar pattern on both data sets. In comparison with LSTM-DBSCAN, GSFM-w-LSTM falls behind for short-term trajectory prediction. However, with the increment of steps i.e. after 25 time steps on HBS and 13 on DUT, the gain of the combined model becomes more profound.

Unlike GSFM, LSTM-DBSCAN learns collision avoidance from the training data with PDM automatically, which does not guarantee collision-free predictions due to incomplete data. Thanks to the collision checking mechanism of the combined model in the post-processing, any collisions in the predictions are prevented by switching to the expert-based model. On the other hand, rather than having a limited number of behaviour patterns like GSFM (e.g. a Gaussian distribution of speed), the DL model generates heterogeneous trajectories using the motion planning module with the encoded information from the observation of the respective agent.

To conclude, the combined model hoards the collective advantages of both models and outperforms the expert-based and DL models in terms of more realistic and collision-free trajectories.

4 ACKNOWLEDGEMENTS

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

REFERENCES

- [1] Alexandre Alahi, Kratharth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, and Silvio Savarese. 2016. Social lstm: Human trajectory prediction in crowded spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 961–971.
- [2] Malte Aschermann, Philipp Kraus, and Jörg P. Müller. 2017. LightJason: A BDI Framework inspired by Jason. In *Multi-Agent Systems and Agreement Technologies: 14th Europ. Conf., EUMAS 2016 (LNCS)*, Vol. 10207. Springer, 58–66. <https://lightjason.github.io>
- [3] Stefania Bandini, Luca Crociani, and Giuseppe Vizzari. 2017. An Approach for Managing Heterogeneous Speed Profiles in Cellular Automata Pedestrian Models. *Journal of Cellular Automata* 12, 5 (2017).
- [4] Hao Cheng, Yao Li, and Monika Sester. 2019. Pedestrian Group Detection in Shared Space. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 1707–1714.
- [5] Hao Cheng and Monika Sester. 2018. Mixed Traffic Trajectory Prediction Using LSTM-Based Models in Shared Space. In *The Annual International Conference on Geographic Information Science*. Springer, 309–325.
- [6] Hao Cheng and Monika Sester. 2018. Modeling mixed traffic in shared space using lstm with probability density mapping. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 3898–3904.
- [7] Arthur Choi, Ruo Cheng Wang, and Adnan Darwiche. 2019. On the Relative Expressiveness of Bayesian and neural networks. *International Journal of Approximate Reasoning* (2019).
- [8] Emma Clarke. 2006. Shared Space-: the alternative approach to calming traffic. *Traffic Engineering & Control* 47, 8 (2006), 290–292.
- [9] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. AAAI, 226–231.
- [10] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. 2018. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2255–2264.
- [11] Ben Hamilton-Baillie. 2008. Shared space: Reconciling people, places and traffic. *Built environment* 34, 2 (2008), 161–181.
- [12] John C Hayward. 1972. Near miss determination through use of a scale of danger. In *51st Annual Meeting of the Highway Research Board*. Highway Research Board, 24–34.
- [13] Dirk Helbing and Peter Molnar. 1995. Social force model for pedestrian dynamics. *Physical review E* 51, 5 (1995), 4282.
- [14] Zhiting Hu, Xuezhe Ma, Zhengzhong Liu, Eduard Hovy, and Eric Xing. 2016. Harnessing deep neural networks with logic rules. *arXiv preprint arXiv:1603.06318* (2016).
- [15] Boris Ivanovic, Edward Schmerling, Karen Leung, and Marco Pavone. 2018. Generative modeling of multimodal multi-human behavior. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 3088–3095.
- [16] Fatema T Johora and Jörg P Müller. 2018. Modeling Interactions of Multimodal Road Users in Shared Spaces. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 3568–3574.
- [17] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436.
- [18] Namhoon Lee, Wngun Choi, Paul Vernaza, Christopher B Choy, Philip HS Torr, and Manmohan Chandraker. 2017. Desire: Distant future prediction in dynamic scenes with interacting agents. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 336–345.
- [19] James R Munkres. 2000. *Topology*. Prentice Hall.
- [20] Kai Nagel and Michael Schreckenberg. 1992. A cellular automaton model for freeway traffic. *Journal de physique I* 2, 12 (1992), 2221–2229.
- [21] F Pascucci, N Rinke, C Schiermeyer, V Berkahn, and B Friedrich. 2017. A discrete choice model for solving conflict situations between pedestrians and vehicles in shared space. *arXiv preprint arXiv:1709.09412* (2017).
- [22] Dino Pedreschi, Fosca Giannotti, Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, and Franco Turini. 2019. Meaningful explanations of black box AI decision systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. AAAI, 9780–9784.
- [23] N Rinke, C Schiermeyer, F Pascucci, V Berkahn, and B Friedrich. 2017. A multi-layer social force approach to model interactions in shared spaces using collision prediction. *Transportation Research Procedia* 25 (2017), 1249–1267.
- [24] Amir Sadeghian, Vineet Kosaraju, Ali Sadeghian, Noriaki Hirose, Hamid Reza Tofighi, and Silvio Savarese. 2019. Sophie: An attentive gan for predicting paths compliant to social and physical constraints. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 1349–1358.
- [25] Jürgen Schmidhuber. 2015. Deep learning in neural networks: An overview. *Neural networks* 61 (2015), 85–117.
- [26] Robert Schönauer. 2017. *A Microscopic Traffic Flow Model for Shared Space*. Ph.D. Dissertation. Graz University of Technology.
- [27] Kota Yamaguchi, Alexander C Berg, Luis E Ortiz, and Tamara L Berg. 2011. Who are you with and where are you going?. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 1345–1352.
- [28] Dongfang Yang, Linhui Li, Keith Redmill, and Ümit Özgüner. 2019. Top-view trajectories: A pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus. In *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 899–904.
- [29] Dongfang Yang, Ümit Özgüner, and Keith Redmill. 2018. Social force based microscopic modeling of vehicle-crowd interaction. In *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 1537–1542.