

Visualization of Pedestrian Interaction through Attention-based Pedestrian Trajectory Prediction

Wen-Xin Qiu^{*1}, Takashi Fuse¹

¹ Department of Civil Engineering, The University of Tokyo

Analysis of pedestrian trajectory from observational data is an important approach to understanding microscopic pedestrian behaviors at the operation level. Based on the understanding, pedestrian simulation and trajectory prediction could facilitate pedestrian space development and pedestrian safety study. The studies can be categorized as conventional approaches and deep learning approaches. The conventional approaches model pedestrian behaviors based on known features, such as avoiding collision, and further improve the knowledge of those features. The deep learning-based approaches learn various features from data and model the pedestrian interactions through designed mechanisms rather than treat them as independent time series data. Although deep learning-based approaches achieved higher accuracies in prediction, the lack of interpretability due to its black-box nature is an obstacle to improving generalizable knowledge of pedestrian behaviors. This study aims to improve the deep learning-based pedestrian trajectory prediction method with the consideration of accuracy, computational cost, and interpretability. A spatial-temporal graph is constructed to model the coordinates and interactions of observed pedestrians. The graph attention network (GAT) is introduced into the proposed approach to obtain attention scores. GAT is effective in the number of learnable parameters, a measure of computational costs, and can handle bidirectional edges. The learned attention scores represent the degree how much a pedestrian is aware of one another, so they can be considered as an explicit quantitative representation of the interactions. With the visualization of the scores, the users, such as space planners or traffic engineers, can perceive how the deep learning model learned the interactions. Our proposed approach is validated on a benchmark dataset, ETH/UCY. Compared to the baseline models, the low computational cost is achieved owing to the efficiency of the GAT; the high accuracy is shown by evaluating average displacement error (ADE) and final displacement error (FDE). Finally, the predictions and the attention scores are visualized to provide an interpretation of pedestrian interaction learned by the deep learning model.

Keywords: Pedestrian trajectory, Deep learning, Spatial-temporal graph, Attention mechanism