EUROGRAPHICS 2025 / A. Bousseau and A. Dai
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 44 (2025), Number 2

CEDRL: Simulating Diverse Crowds with Example-Driven Deep Reinforcement Learning
Supplementary Material

A. Panayiotou[1,2], A. Aristidou[1,2] and P. Charalambous[2]
[1]Department of Computer Science, University of Cyprus, Cyprus [2]CYENS - Centre of Excellence, Cyprus

Content. We include additional findings in this supplementary document. Initially, we conduct an in-depth study to identify the optimal weight combination for our model's reward function (Section 4.4.1)). Then we provide two additional evaluation studies, one involving dynamic environments with obstacles (Section 5.6), and another focusing on w_{comp} distribution reproduction performance (Section 5.7).

4.4. Reward Function Design

4.4.1 Evaluating Imitation Performance

Weight Combination Analysis. We perform a study to find the combination of $\{w_u, w_p, w_o\}$ that yields the best results. We train four models, each with different combinations of weights, presenting the training curves in Figure 1. To ensure consistency, we run

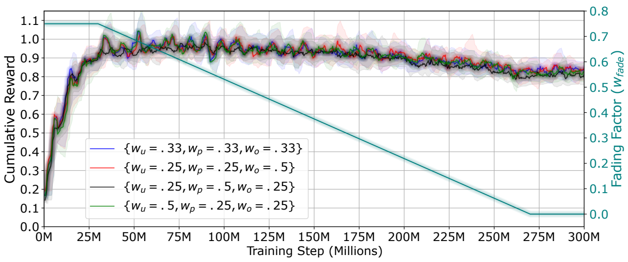

Figure 1: *Training curves for different weight combinations.*

five simulations per dataset (Ped-1 and UC-3), using each trained model, and calculate the distributions of w_{comp}. Then we compute the cross-entropy [Sha48] between the ground truth and simulated distributions to assess which model most closely resemble the characteristics of the real-world data; lower values indicate better alignment between the two distributions. The results presented in Figure 2 indicate that the model trained with the combination $\{w_u = .25, w_p = .5, w_o = .25\}$ achieves the best performance on both datasets. This combination places a greater emphasis on proximity, compelling agents to more precisely align their positions with those of their assigned real agents. Thus, we select this weight combination for our model.

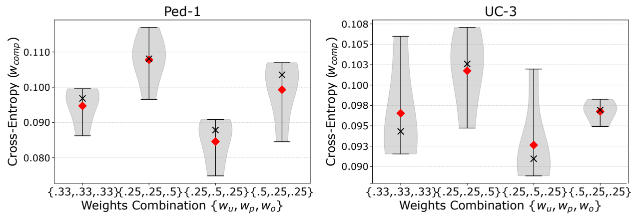

Figure 2: *Reward Parameters Analysis. We show median and mean with X and diamond symbols respectively.*

5.6. Adaptability in Dynamically Changing Environment

We test our model performance on synthetic dynamic scenarios utilizing the *Infinite* environment, enabling constant agent flow, and dynamically spawn obstacles using pseudo-random placement. Specifically, every 10 seconds we use the current agents' positions and apply Delaunay triangulation [Del34] to discretize the environment. We then identify triangles with an area above a set threshold, randomly select 15-25, and place one randomly sized and oriented obstacle in each, ensuring obstacles do not disrupt agents' state.

The setup forces agents to face scenarios not encountered during training, and is applied to assess their ability to perform the in-

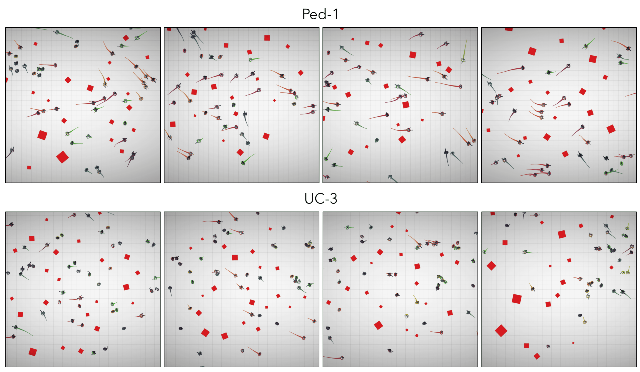

Figure 3: *Synthetic Dynamically Changing Environment. We utilize the Infinite environment, dynamically placing random obstacles at regular intervals. Rendered results are presented by sampling w_{comp} values from Ped-1* ***(top)*** *and UC-3* ***(bottom)*** *datasets.*

© 2025 Eurographics - The European Association
for Computer Graphics and John Wiley & Sons Ltd.

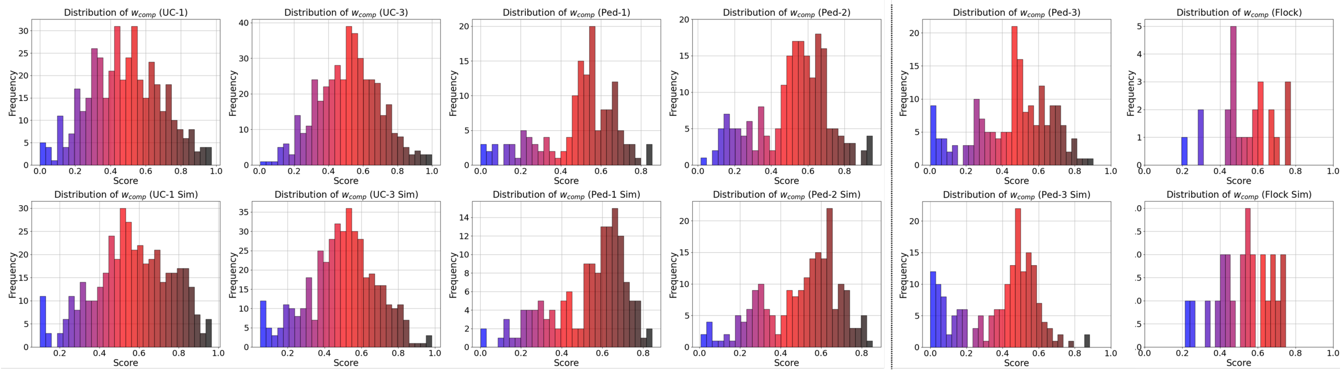

Figure 4: *We present the distribution of calculated complexity scores for all real-world training and unseen datasets* ***(top)****, accompanied by the distributions calculated over their generated simulations using our model* ***(bottom)****.*

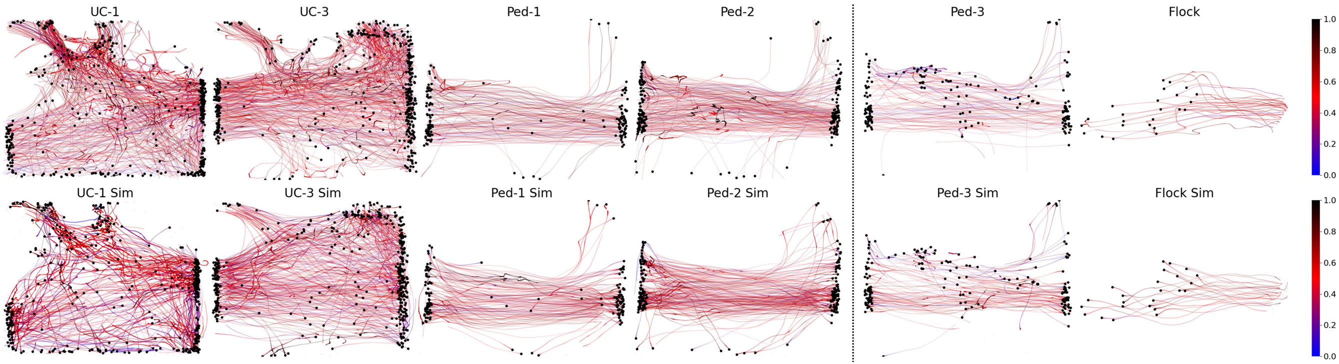

Figure 5: *We present trajectories, color-coded based on their w_{comp} score, for all real-world training and unseen dataset* ***(top)****, accompanied by sample color-coded trajectories from their generated simulations using our model* ***(bottom)****.*

dicated behaviors in novel dynamic environments. We select two datasets, Ped-1 and UC-3, sample w_{comp} values from their distributions, and run individual simulation. In Figure 3 we present rendered results, showing the environment's state at four different timesteps over time. In the upper section (Ped-1), we show that most agents are moving in pairs or groups, mimicking pedestrians on a shopping street, with a few agents remaining stationary. On the other hand, for the UC-3, a greater proportion of agents are engaged in stationary group interactions, similar to students gathering on a university campus, while others are seen moving towards specific destinations, like their classes. Finally, we argue that despite the changing structure of the environment over time, agents are still able to perform their designated behaviors; animated results are presented in the supplementary video material.

5.7. Behavior Complexity Reproduction

We apply an additional study where we evaluate if our model can reproduce the input data distribution w_{comp}. We run individual simulations using the characteristics of each dataset, both training and unseen. Then, we characterize the simulation trajectories and compute their w_{comp} distributions. We present the results in Figure 4, while also Figure 5 shows the generated trajectories, colored by their corresponding complexity value. The results demonstrate that our model effectively replicates the complexity distributions, suggesting that the agents accurately exhibit the behaviors presented in each dataset. In addition, the behaviors in each simulation are significantly influenced by the varying input data, as evidenced by the clear structural distribution differences observed between various crowd datasets. Specifically, in UC, we show a better coverage of the whole complexity spectrum, as these real-world datasets contain a variety of behaviors, both simple and complex, due to their nature. Individuals on a university campus may walk towards their classes alone or with their classmates, engage in stationary conversations, wander around, and more. However, the complexity of behaviors in the Pedestrians datasets falls mainly within the range of $(.4, .7)$. According to our behavior analysis, there is a high frequency of goal-seeking behavior, with people walking in pairs or groups. This behavior is typical and expected on a small shopping street near a sidewalk. Finally, in the Flock dataset, since the number of tracked trajectories is low (24), the distribution may not be as informative as the other datasets. Nonetheless, our model still reproduces w_{comp} values within the ground truth range.

References

[Del34] DELAUNAY B.: Sur la sphère vide. *Bulletin de l'Academie des Sciences de l'URSS.*, 6 (1934), 793–800. 1

[Sha48] SHANNON C. E.: A mathematical theory of communication. *The Bell System Technical Journal 27*, 3 (1948), 379–423. 1

© 2025 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd.