Predicting the unseen: Improving robustness in Koopman surrogate models for crowd dynamics at a bottleneck

Sabrina Kern*[1] and Gerta Köster[1]

[1]Hochschule München - University of Applied Sciences

Abstract Crowds moving through bottlenecks form a dynamical system, with its density fluctuating in time and space. The system dynamics can be learned and predicted using the Koopman operator framework. But how reliable are predictions for previously unseen situations? How significant is the impact of stochastic observations? In this work, we show that using Taken's embedding and diffusion maps as part of our learning pipeline facilitates robustness of the surrogate model.

Keywords pedestrian bottleneck, dynamical system, data-driven modeling, Koopman operator, diffusion maps, manifold learning

Koopman surrogate models for crowds in a bottleneck

Modeling and predicting the behavior of pedestrian flows during evacuation, especially in constrained environments such as bottlenecks, is crucial for improving public safety. However, to effectively support practitioners in their decision making, models must provide reliable predictions even for unseen situations.

This work investigates the robustness of the Koopman operator, a promising AI technique suitable for predicting the behavior of various dynamical systems [1]. In the context of pedestrian dynamics at a bottleneck, a Koopman surrogate model may allow us to forecast potentially critical levels of density upfront. Training such a surrogate model can use various learning pipelines, including the Dynamic Mode Decomposition (DMD) [2] or the Extended Dynamic Mode Decomposition (EDMD) [3]. While DMD is designed to predict only linear dynamics, EDMD allows to include non-linear transformations such as Taken's embedding [4] or diffusion maps [5].

Given the great flexibility of the learning pipeline, this work focuses on the following question: How does using techniques such as Taken's embedding and diffusion maps enhance the robustness of Koopman surrogate models? Here, robustness refers to the model's ability to accurately predict unseen data variants, in particular: (1) Unseen crowd sizes in pedestrian inflow, being larger, smaller or between crowd sizes present in the training data, and (2) unseen stochastic observations, containing trajectory-induced noise which is not present in the training data.

Taken's temporal embedding enriches the state space of a dynamical system by adding past snapshots to each observation. We expect that this additional information supports inter- and extrapolation. Diffusion maps on the other hand are a dimensionality reduction technique. They learn a low-dimensional embedding via a diffusion process which aims to preserves essential geometric structures of the data. Overall, reducing the dimensionality increases computational efficiency; in extreme cases, this makes working with large temporal embeddings tractable at all.

Experiments on surrogate model robustness

We simulate the crowd dynamics at a bottleneck using the microscopic simulator Vadere [6]. All virtual pedestrians want to exit through a single bottleneck simultaneously. We generate data for seven crowd sizes, from 10 to 70 people, creating varying density patterns. For each size, we generate two time series: one average of 50 simulations, and one stochastic from a single run. The data is discretized into a spatial grid of 20cm $\times$ 20cm cells, leading to $55 \times 40 = 2200$ cells per snapshot, with snapshots taken every 0.2 seconds.

To analyze surrogate model robustness, we use three learning pipelines. The first employs the classical Dynamic Mode Decomposition (DMD). The second integrates Taken's embedding, adding four frames to each state. In the third and most advanced pipeline, the Taken's embedding is followed by diffusion maps which embed the enriched observations into a lower dimensional state space. This learning approach was previously suggested in [7].

*Email of the corresponding author: sabrina.kern@hm.edu

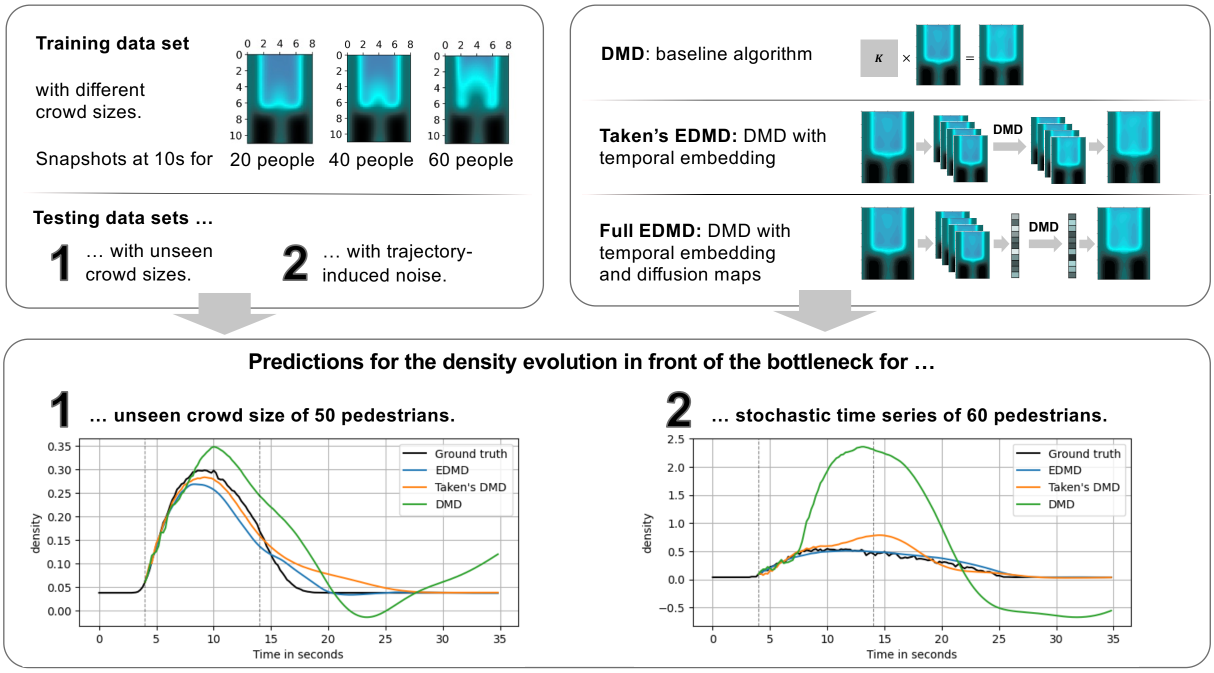

Figure 1: Graphical abstract, showcasing the improvements in prediction accuracy for two exemplary time series from the training data set.

Starting from pure DMD, which leads to infeasible predictions especially for stochastic time series, our experiments show significantly better predictions for the two EDMD learning pipelines. Here, we can largely attribute the main improvements in robustness to using Taken's embedding. At the same time, diffusion maps are able to reduce the dimensionality by a factor of 100, while maintaining the prediction accuracy achieved with Taken's embedding alone. Such a reduction becomes inevitable once we need to increase the amount of training data, increase the spatial resolution, or add more snapshots to the temporal embedding to capture more complex dynamics.

In summary, our comparative analysis shows that using Taken's embedding and diffusion maps significancy improves the accuracy in predictions for unseen crows sizes and stochastic observations while keeping the problem computationally tractable. Further investigations will also incorporate data from laboratory experiments as the stochastic time series.

References

[1] Marko Budišić, Ryan Mohr, and Igor Mezić. Applied koopmanism. *Chaos*, 22:047510, 2016.

[2] Jose Nathan Kutz, Brunton, Steven L. Brunton, Bingni W. Brunton, and Joshua L. Proctor. *Dynamic Mode Decomposition: data-driven modeling of complex systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2016.

[3] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data-driven approximation of the koopman operator: Extending dynamic mode decomposition. *J. Nonlinear Sci.*, 25(6):1307–1346, 6 2015.

[4] Floris Takens. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, pages 366–381, 1981.

[5] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 07 2006.

[6] Benedikt Kleinmeier, Benedikt Zönnchen, Marion Gödel, and Gerta Köster. Vadere: An open-source simulation framework to promote interdisciplinary understanding. *Collective Dynamics*, 4, 2019.

[7] Daniel Lehmberg, Felix Dietrich, and Gerta Köster. Modeling Melburnians—Using the Koopman operator to gain insight into crowd dynamics. *Transportation Research Part C: Emerging Technologies*, 133:103437, December 2021.