Activity Scheduling with Deep Generative Models

F Shone[a*], T Hillel[a]

[a] UCL Behaviour and Infrastructure Group, London, UK
ucfnfjs@ucl.ac.uk, tim.hillel@ucl.ac.uk
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 26, 2024

Keywords: Activity Modelling, Sequence Modelling, Deep Generative Modelling, Modelling

1 INTRODUCTION

The generation of realistic samples of synthetic activity schedules is a critical component of both Activity-Based Models (ABMs) and simulation-based transport approaches such as MATSim. Conventional models decompose the scheduling process into series of discrete choices, applied sequentially. Realistic interactions between these choices are therefore limited. The combination of multiple interacting sub-models is also challenging to develop and use.

We demonstrate an alternative integrated approach to activity sequence modeling using deep generative learning. Our model is able to rapidly synthesise activity sequences for a population of socio-demographic attributes. This allows applications for (i) data anonymisation, (ii) diverse up-sampling, (iii) bias correction, and (iv) forecasting based on demographic shift. We evaluate the quality of output activity schedules both individually and in aggregate.

We identify three primary benefits of our integrated approach; (i) simplicity and speed, (ii) diversity in output sequences, and (iii) the potential for more realistic interaction of choice components. We have created the open-source Python software CAVEAT (Shone, 2023) for the development and evaluation of generative activity models.

1.1 Activity Modelling

The planning of activity schedules for an individual can be thought of as a set of interacting choices. These choices interact with each other, the choices of others, and with the environment. These choices are also influenced by agent attributes, such as age, employment and car ownership. The dominant approach to model activity scheduling is with systems of discrete choice models. Choices are then combined into coherent schedules using bespoke rules or scheduling algorithms, such as by Manser *et al.* (2022).

Discrete choice models require measurement or estimation of costs for every possible alternative choice. They also commonly rely on significant simplifications of the choice sets, for example limiting possible tours to a set of most common activity sequences. This simplification restricts their realism, fails to capture true diversity, and limits realistic interaction between choices.

1.2 Deep Generative Models

Deep generative models have made headlines for the generation of realistic images and sequences of text. They have already been applied in the transport domain for generating synthetic population attributes (Borysov *et al.*, 2019, Kim & Bansal, 2023). There has also been application of deep generative models to more complex data structures; for example by Choi *et al.* (2021), for vehicle trajectories.

Koushik *et al.* (2023) predict activity sequences conditional on agent attributes using a discriminative approach. They find aggregate realism challenging, particularly the correct representation of infrequently observed activities. We build on this work by developing a a generative model with new data encoding, model structure and evaluation framework.

2 Methodology

We consider an observed population of individuals each with a recorded 24-hour activity sequence, e.g. from a travel diary survey. Each individual is described by a set of socio-demographic attributes, such as age and gender. We train a model to learn both the joint and conditional distributions of attributes and sequences in this observed population. We are then able to apply the model to generate synthetic sequences for a given set of individual attributes.

2.1 Data

We extract daily activity sequences from 2021 UK National travel Survey (NTS) using PAM (Shone *et al.*, 2024). We filter the sample of activity sequences to only include those that start and end at home. This enables evaluation of the structural quality of synthesised sequences.

The finished dataset contains 39850 individuals with corresponding activity sequences and attributes describing their age, gender, education, employment, and car ownership. We refer to this data as the *observed* sample. Output sequences are then called *synthetic* samples. Models are trained on 90% of the observed data. We use the remaining 10% for validation during model training. We encode each activity schedule into a sentence-like sequence, listing each activity in order alongside its associated duration (with travel time omitted). For example, the sequence *start; home: 8.5; work: 8; shop: 1.5; home: 6; end* would describe someone leaving to work at 8:30, working until 16:30, participating in leisure until 18:00, and then staying at home until the end of the day.

2.2 Model Design Summary

Our model is based on the Conditional Variational Auto-encoder (CVAE) (Sohn *et al.*, 2015). During training the model learns to map from a normally distributed latent space to the observed sequence distribution, conditional on attributes for each individual. After training, by sampling from the latent space and providing conditional attributes, the model can then generate new, potentially novel sequences.

Our model is composed of an encoder and decoder. Both use a recurrent-structure of LSTM units (Hochreiter & Schmidhuber, 1997) as per Figure 1. The full architecture is available in CAVEAT (Shone, 2023).

2.3 Evaluation Summary

We evaluate the model by comparing each output synthetic sequence to target observed data. We measure both the **correctness** and **creativity** of the synthetic population sample. For **correctness** we measure the distance between distributions in the target and synthetic populations.

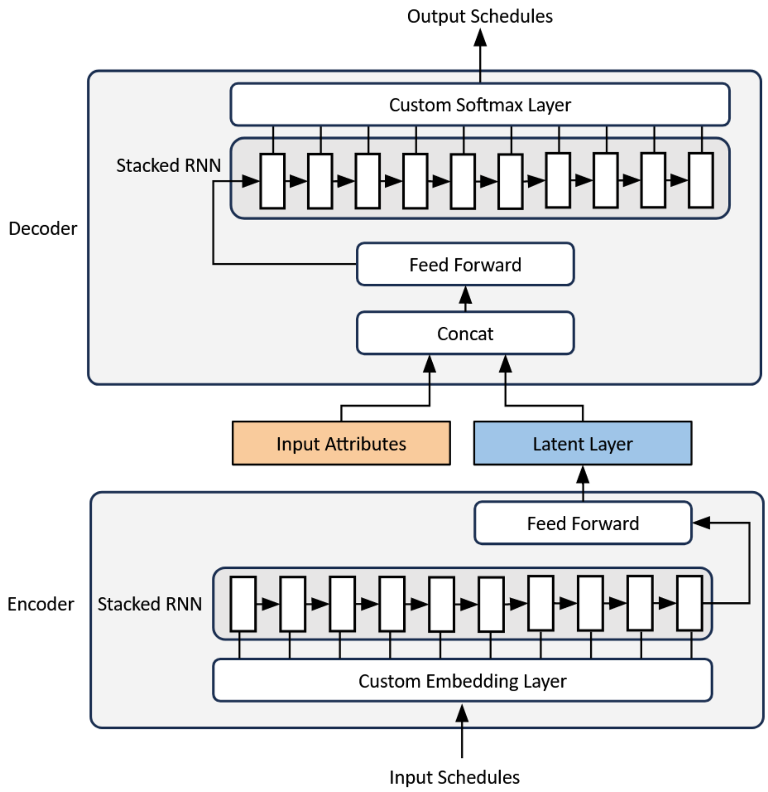

Figure 1 – *Model design*

For example, the number of times each individual goes shopping (participation rate) or the start time of each work activity. For **creativity** we measure the diversity of output sequences and the probability they are *novel* (unobserved in the training data).

This results in a large number of metrics which we combine using weighted averages into the themes; **creativity**, **structure**, **frequency**, **participation**, **transitions** and **timing**. Full details of which can be found in the CAVEAT evaluation module (Shone, 2023). We evaluate model performance on two different tasks; (i) interpolation, where all target attributes (such as combined age, gender and income) occurred in the training set, and (ii) extrapolation, where some target attribute combinations are previously unseen.

3 Preliminary Results

We find that the model is able to rapidly generate realistic sequences, both individually and in aggregate for a synthetic population. We are also able to produce diverse and novel sequences through the latent sampling process. We present preliminary mean and variance of evaluation metrics for the correctness themes for five model runs in Table 1. Figure 2 shows a sample of synthetic sequences.

These initial results promise a practical option for more realistic activity scheduling within an activity or simulation-based modelling framework. However our approach is limited in it's ability to forecast as it can only extrapolate or interpolate patterns in observed data. Additionally the model is likely to require safeguarding in application as its performance cannot be guaranteed for all possible samples of the latent space. Further work will look to extend the approach to additional choice components such as mode choice, and extend the model to multi-day scheduling.

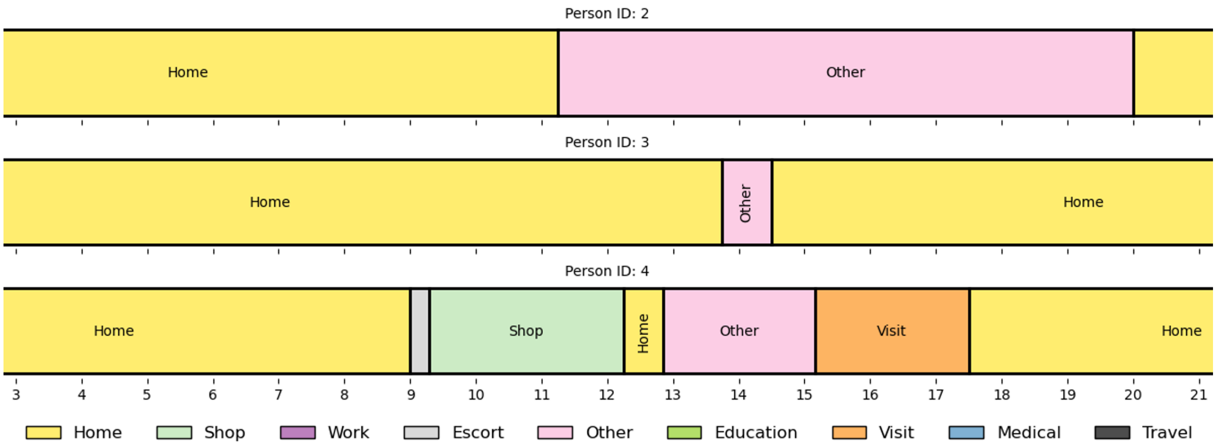

Figure 2 – *Example synthetic sequences*

Table 1 – *Correctness evaluation summary*

	distance to target*		metric	unit
Evaluation theme	mean	variance		
structural feasibility of schedule	0.075	0.001	EMD	-
frequency of activity throughout day	0.039	0.000	MAPE	%
participation rate of activities	0.007	0.000	EMD	rate
transitions between activities	0.006	0.000	EMD	rate
timing and durations of activities	0.037	0.000	EMD	hours

∗mean and variance of evaluation metrics from 5 runs of model.

EMD = Wasserstein or *Earth Movers* Distance.

MAPE = Mean Absolute Percentage Error.

References

Borysov, Stanislav S., Rich, Jeppe, & Pereira, Francisco C. 2019. Scalable Population Synthesis with Deep Generative Modeling. *Transportation Research Part C: Emerging Technologies*, **106**(Sept.), 73–97.

Choi, Seongjin, Kim, Jiwon, & Yeo, Hwasoo. 2021. TrajGAIL: Generating Urban Vehicle Trajectories Using Generative Adversarial Imitation Learning. *Transportation Research Part C: Emerging Technologies*, **128**(July), 103091.

Hochreiter, Sepp, & Schmidhuber, Jürgen. 1997. Long Short-Term Memory. *Neural Computation*, **9**(8), 1735–1780.

Kim, Eui-Jin, & Bansal, Prateek, AP. 2023. A Deep Generative Model for Feasible and Diverse Population Synthesis. *Transportation Research Part C: Emerging Technologies*, **148**(Mar.), 104053.

Koushik, Anil, Manoj, M, Nezamuddin, N, & Prathosh, AP. 2023. Activity Schedule Modeling Using Machine Learning. *Transportation Research Record*, **2677**(8), 1–23.

Manser, Patrick, Haering, Tom, Hillel, Tim, Pougala, Janody, Krueger, Rico, & Bierlaire, Michel. 2022. Estimating Flexibility Preferences to Resolve Temporal Scheduling Conflicts in Activity-Based Modelling. *Transportation*, Oct.

Shone, Fred. 2023. *CAVEAT*. https://github.com/fredshone/caveat.

Shone, Fred, Chatziioannou, Theodore, Pickering, Bryn, Kozlowska, Kasia, & Fitzmaurice, Michael. 2024. PAM: Population Activity Modeller. *Journal of Open Source Software*, **9**(96), 6097.

Sohn, Kihyuk, Lee, Honglak, & Yan, Xinchen. 2015. Learning Structured Output Representation using Deep Conditional Generative Models. *In:* Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., & Garnett, R. (eds), *Advances in Neural Information Processing Systems*, vol. 28. Curran Associates, Inc.