Generation of Activity Sequences using Dynamic Bayesian Networks with Latent Variables

Ioanna Arkoudi*[a], Isabel Viegas de Lima[a], Carlos Lima Azevedo*, Francisco C. Pereira*

ioaar@dtu.dk, isabel@viegasdelima.com, climaz@dtu.dk, camara@dtu.dk

DTU Technical University of Denmark*, [a] *equal contribution*

1 | Introduction

Activity-based models consider travelers' behaviors as a sequence of trips resulting from their choice of activity-making [1], which can be used to extract aggregate demand metrics and to simulate travel behavior. While travel behavior in itself has been studied and modeled for decades, modeling the activity sequence generation process is less straightforward. There is a number of endogenous and external factors, such as congestion, unplanned events and incidents, that constantly change the system's conditions and ultimately affect travelers' decisions introducing stochasticity in their observed behavior. In this study, we propose a dynamic latent approach, which leverages on Model-Based Machine Learning (MBML), for generating daily activity sequences. This approach offers a coherent framework, based on probability theory and Bayesian Inference (BI), which formally accounts for the uncertainty and dynamic issues that are inherent to travel decision making, and is less prone to overfitting compared to point estimate approaches, such as Maximum Likelihood Estimation (MLE) [2].

2 | Approach

The model we present (M), is a multivariate dynamic Bayesian network with timestep-specific latent states corresponding to continuous-valued vectors $\mathbf{z}_t$ of fixed dimensionality D, and K number of timesteps. Let TS the number of performed timesteps at time t, C the number of target classes/activities $\mathbf{y}$, and N the number of observations in the data. The graphical structure of M is presented below in Figure 1, followed by the generative process of the model.

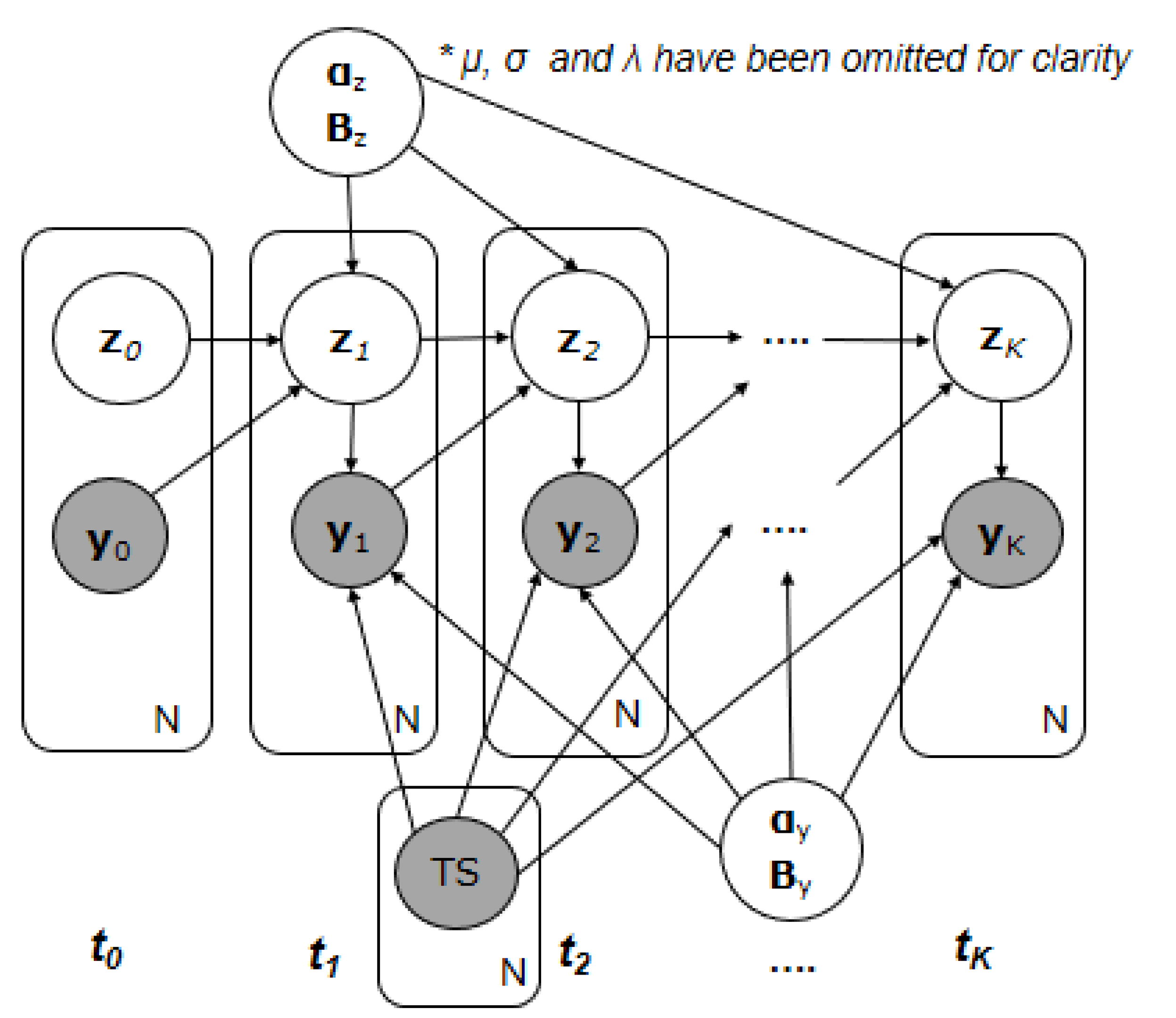

Figure 1:Probabilistic graphical structure of the proposed model M.

1. Draw latent state coefficients $\mathbf{B}_z^{Dx(C+D)} \sim N(\mu, \sigma)$ intercept $\boldsymbol{\alpha}_z^D \sim N(\mu, \sigma)$
2. Draw observed activity coefficients $\mathbf{B}_y^{Cx(D+1)} \sim N(\mu, \sigma)$ and intercept $\boldsymbol{\alpha}_y^C \sim N(\mu, \sigma)$
3. For $t = 0$:
   For each $n = 1, ..., N$:
   Draw first latent state $\mathbf{z}_0^D \sim N(\mu, \mathbf{I}\lambda)$
4. For each timestep $t = 1, ..., K$:
   For each $n = 1, ..., N$:
   Draw latent state at time t:

$$\mathbf{z}_t^D \sim N\left(\boldsymbol{\alpha}_z^D + \mathbf{B}_z^{Dx(C+D)}[\mathbf{y}_{t-1}^C, \mathrm{Softmax}(\mathbf{z}_{t-1}^D)]\right) \tag{1}$$

Draw observed activities at time t:

$$\mathbf{y}_t \sim Multinomial\left(\mathrm{Softmax}(\boldsymbol{\alpha}_y^C + \mathbf{B}_y^{Cx(D+1)}[exp(\mathbf{z}_t^D), TS]\right) \tag{2}$$

Mathematical Notation

Letters which are not bold (e.g. μ, D or σ) denote scalars. A lower case bold letter (e.g. α or $\mathbf{z}$) denotes a vector, while an upper case bold letter (e.g. $\mathbf{B}$) denotes a matrix. $\mathbf{I}$ denotes an identity matrix, with dimensions determined by context.

3 | Model Estimation

We used the Danish National Travel Survey (TU) dataset [3] and extracted 138,102 trip sequences comprising of 6 major activities: home, leisure, workplace, business, errand, educational. We used $N = 3,817$ sequences of maximum length $K = 8$, to train M, and assumed priors for its hyperparameters $\mu = 0.0$, $\sigma = 5.0$ and $\lambda = 0.01$. We experimented with the number of the latent dimensions (D) within a range from 1 to 7. We used the resulting models to run approximate BI in Pyro [5] following an Elbo-based Stochastic Variational Inference (SVI) approach. Then we estimated the values of $\boldsymbol{\alpha}_z, \mathbf{B}_z, \boldsymbol{\alpha}_y, \mathbf{B}_y$ by drawing samples from the posterior distribution.

4 | Results

As comparison baselines we considered 2 models (B_1 and B_2) with similar dynamic structures to M but without containing a latent layer. For estimating their parameters we used BI for B_1 and MLE for B_2. For each model configuration we generated 138,102 activity sequences, and compared their distribution to the true sequence distribution observed in the data. We used Cross Entropy as a loss metric to evaluate the generative performance of the models. Given two discrete probability distributions P and Q on the same probability space $\mathcal{X}$, the CE from Q to P is defined as:

$$CE(P\|Q) = -\sum_{x\in\mathcal{X}} P(x)\log\left(Q(x)\right). \tag{3}$$

The results are presented below in Figure 2.

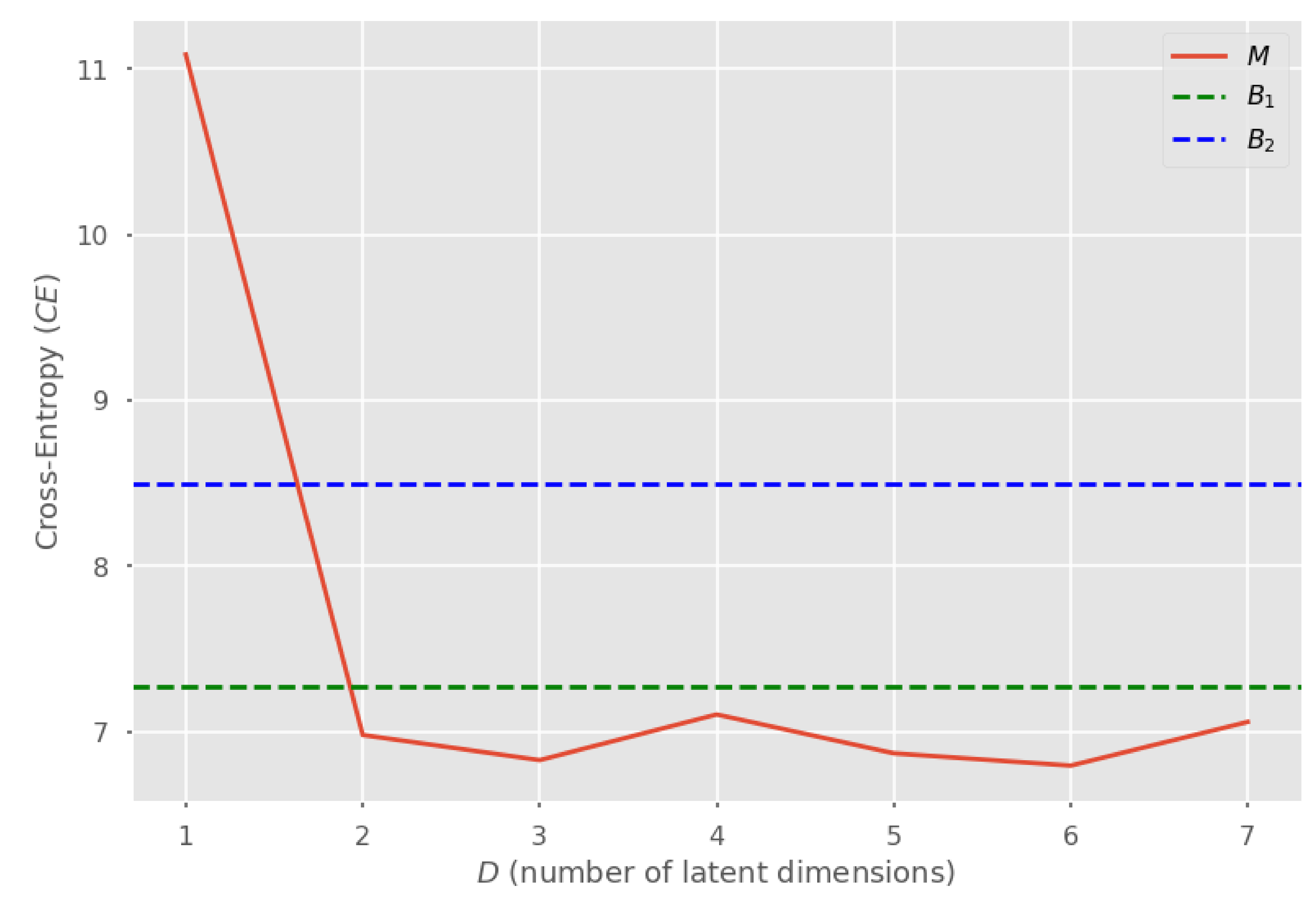

Figure 2:Generative Performance of M and comparison to baselines B_1, and B_2.

5 | Conclusions & Future work

The proposed model exhibits higher generative capability compared to both baseline models for $D = 1$. The results suggest that including a latent layer within a dynamic MBML framework allowed to capture the sequential dependencies between the latent states that give rise to the observed activity-trip sequences of heterogeneous travelers. In the future we intend to extend the current model by including socioeconomic explanatory variables and time-related features (time of day an activity is performed, duration of the previous activity etc.). Further work will focus on interpreting the latent dimensions by introducing different constraints, e.g. as the ones suggested in [4], in order to provide a model that is more intuitive to understand and closer resembling travelers' decision-making processes.

References

[1] John L Bowman and Moshe E Ben-Akiva. Activity-based disaggregate travel demand model system with activity schedules. *Transportation research part a: policy and practice,* 35(1):1–28, 2001.

[2] Inon Peled, Filipe Rodrigues, and Francisco Câmara Pereira. Model-based machine learning for transportation. *In Mobility patterns, big data and transport analytics,* pages 145–171. Elsevier, 2019.

[3] Hjalmar Christiansen. Documentation of the danish national travel survey. 2012.

[4] Isabel Viegas de Lima. *The dynamics of activity-making and travel.* Massachusetts Institute of Technology, Jun 2018.

[5] Bingham, E., J. P. Chen, M. Jankowiak, F. Obermeyer, N. Pradhan, T. Karaletsos, R. Singh, 28 P. Szerlip, P. Horsfall, and N. D. Goodman, *Pyro: Deep universal probabilistic programming,* pages 973–978. The Journal of Machine Learning Research, Vol. 20, No. 1, 2019.