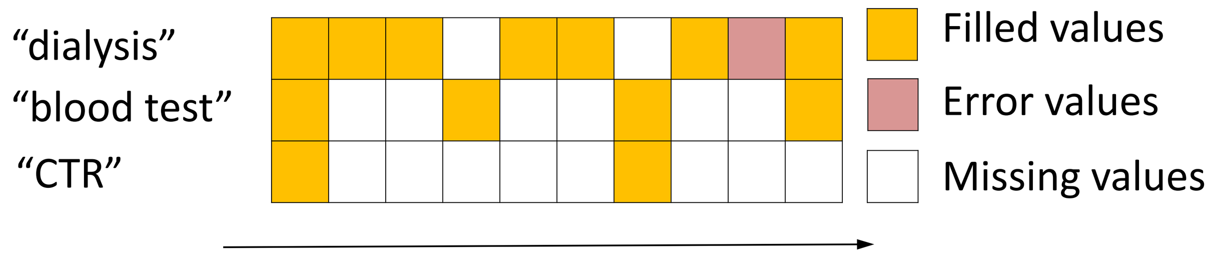
Dynamic Gaussian Mixture based Deep Generative Model For Robust Forecasting on Sparse Multivariate Time Series

Yinjun Wu[1], Jingchao Ni[2], Wei Cheng[2], Bo Zong[2], Dongjin Song[3], Zhengzhang Chen[2], Yanchi Liu[2], Xuchao Zhang[2], Haifeng Chen[2], Susan Davidson[1]
[1]University of Pennsylvania, [2]NEC Laboratories America, [3]University of Connecticut

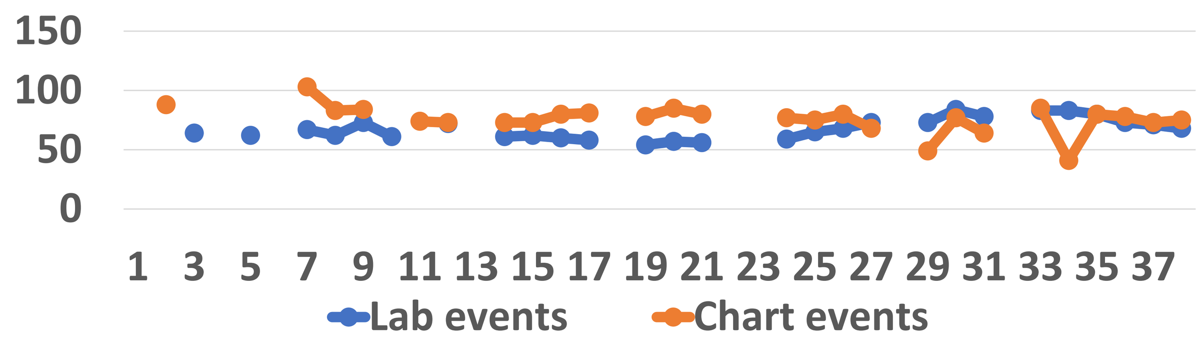
Background

High sparsity of the multivariate time series (MTS) in Medical Domain:

- Different sampling rates, e.g., dialysis is three times per week, blood test is biweekly
- Irregular measurements: patients may miss some visits for tests for a variety of reasons

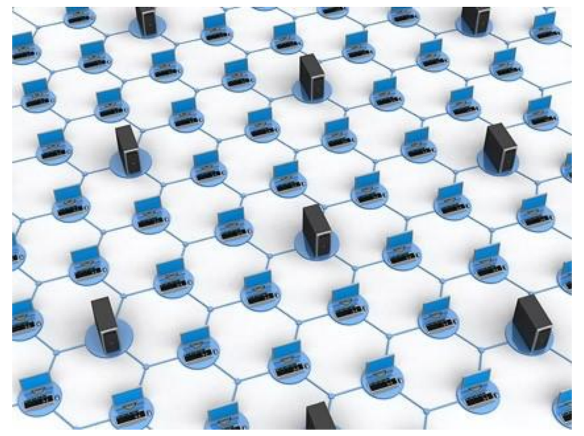


- Dialysis: 3 times / week
- Blood test: 2 times / month
- CTR: 1 time / month



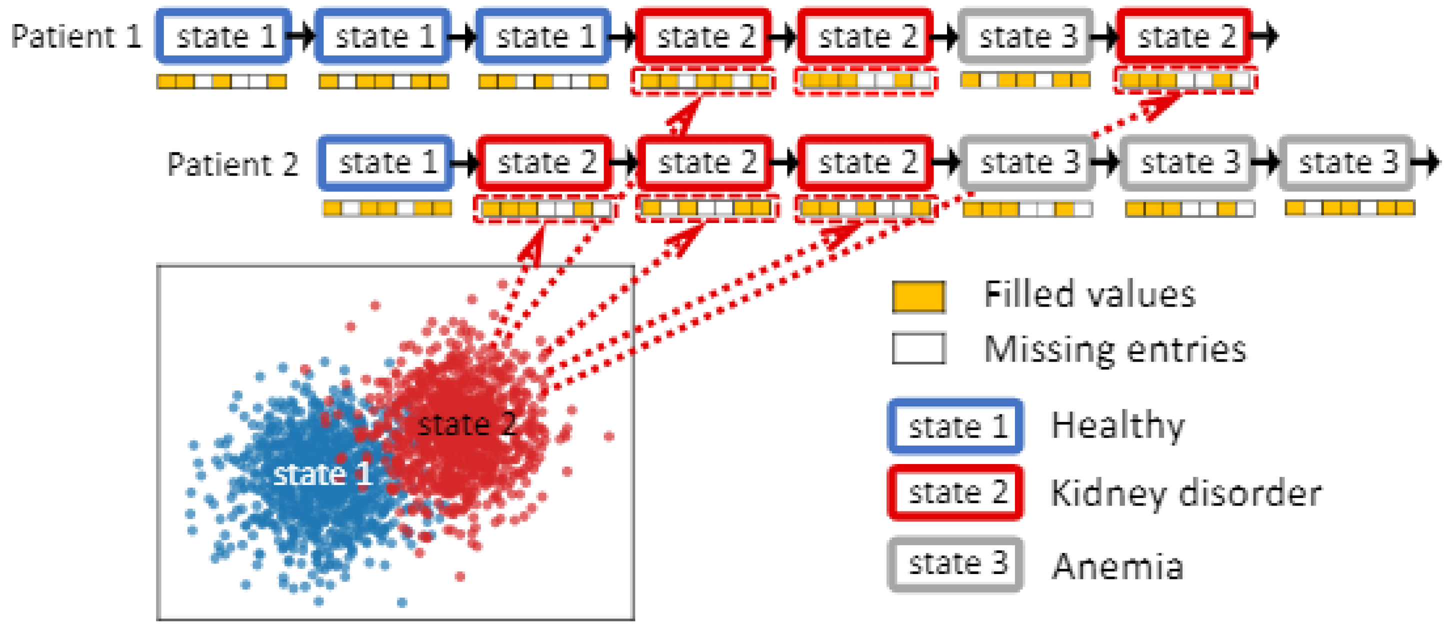
ICU signals (MIMIC-III): lab events and chart events are different sources.

Finance: different sources such as financial news, analysis and market-makers in investment banks



Complex monitoring systems: sensors of different components have different rates

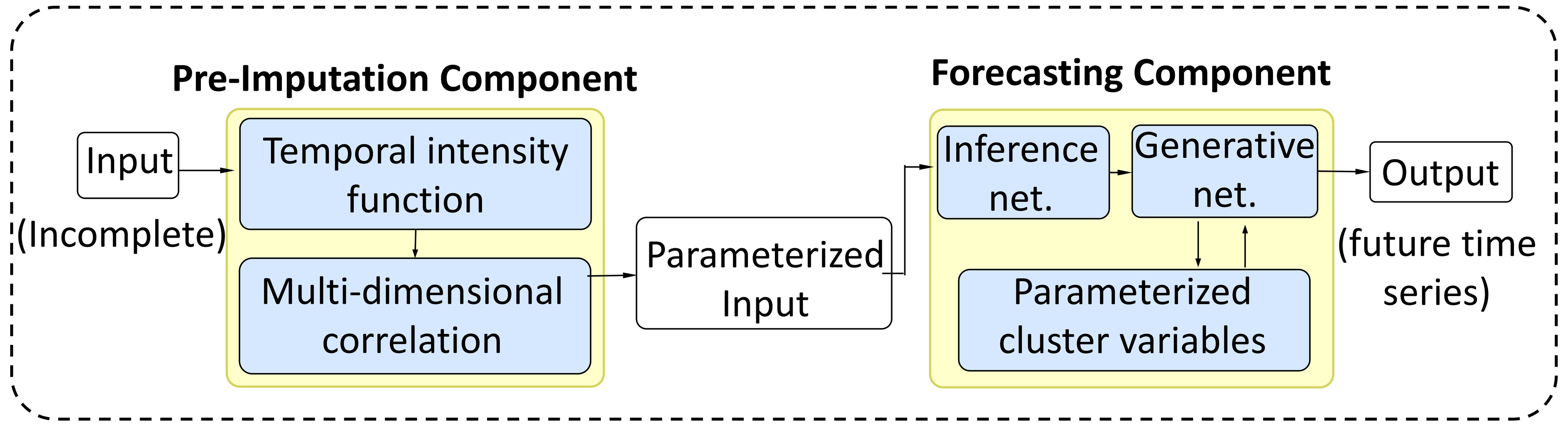
Our perspective



Similar patients may have similar latent states
-> cluster latent states from different MTS's

Our framework

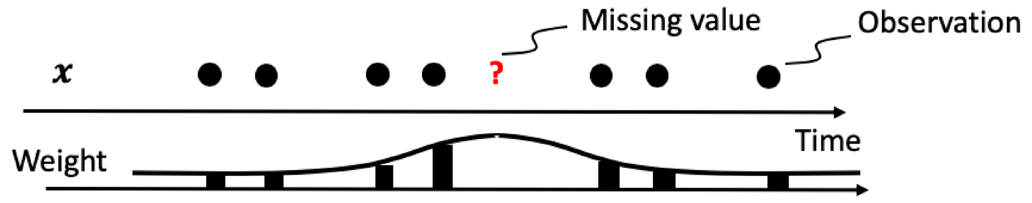
Joint imputation and forecasting framework



Imputation component

Parameterizing missing entries - interpolation

- Temporal intensity: $\bar{x}_{t^*}^i = \dfrac{1}{\lambda(t^*, \boldsymbol{m}^i; \alpha_i)} \sum_{t=1}^{w} \kappa(t^*, t; \alpha_i) m_t^i x_t^i$



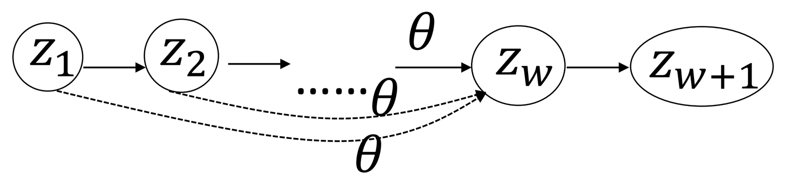
- Multi-dimensional correlation:

$$\hat{x}_{t^*}^i = \left[\sum_{j'=1}^{d} \rho_{ij} \lambda(t^*, \boldsymbol{m}^i; \alpha_j) \bar{x}_{t^*}^j\right] \Big/ \sum_{j'=1}^{d} \lambda(t^*, \boldsymbol{m}^i; \alpha_{j'})$$

Forecasting component

Transition Process

Sample $x_{1:w}$



$$p(z_{t+1}|z_{1:t}) = f_{\theta}(z_{1:t})$$

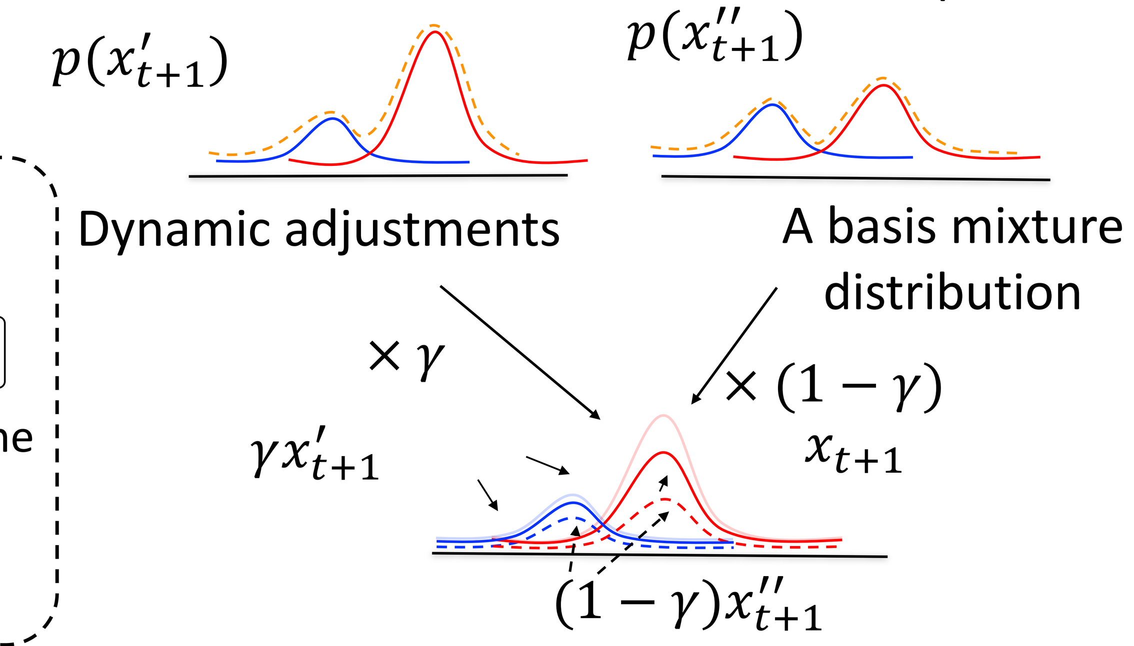
Emission Process

Emission of a new observation from a Dynamic Gaussian mixture

- Draw a latent cluster variable z_{t+1} from $Categorical(p(\psi_{t+1}))$
- Draw x_{t+1} from $N(\mu_{z_{t+1}}, \sigma^{-1} I)$

$$\psi_{t+1} = \underbrace{(1-\gamma)p(z_{t+1}|z_{1:t})}_{\text{dynamic adjustment}} + \underbrace{\gamma p(\boldsymbol{\mu})}_{\text{basis mixture}}$$

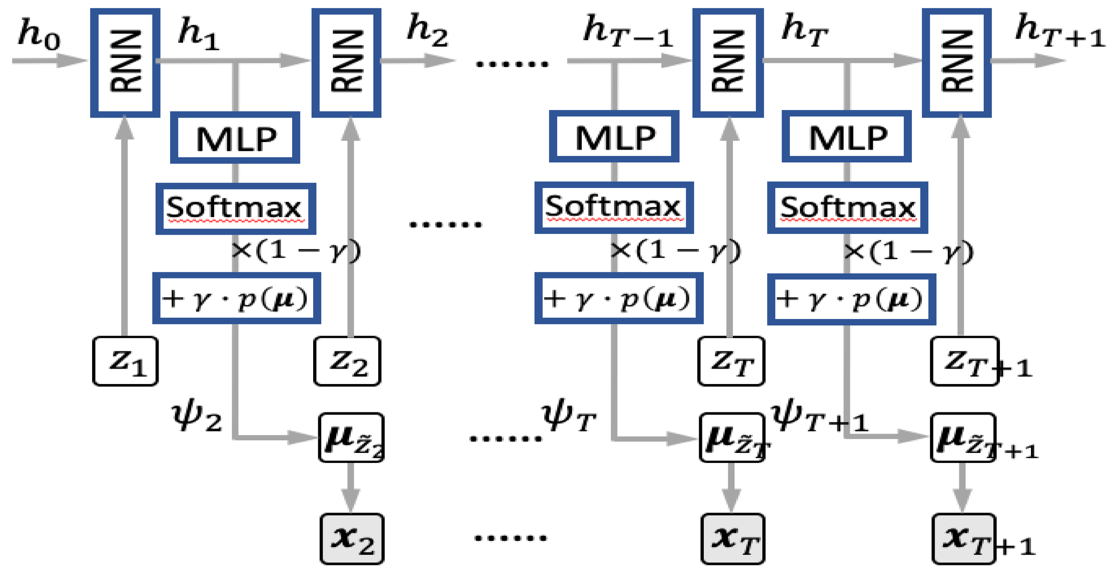
Gaussian Mixture with two mixture components:



Pamameterization

Generative network

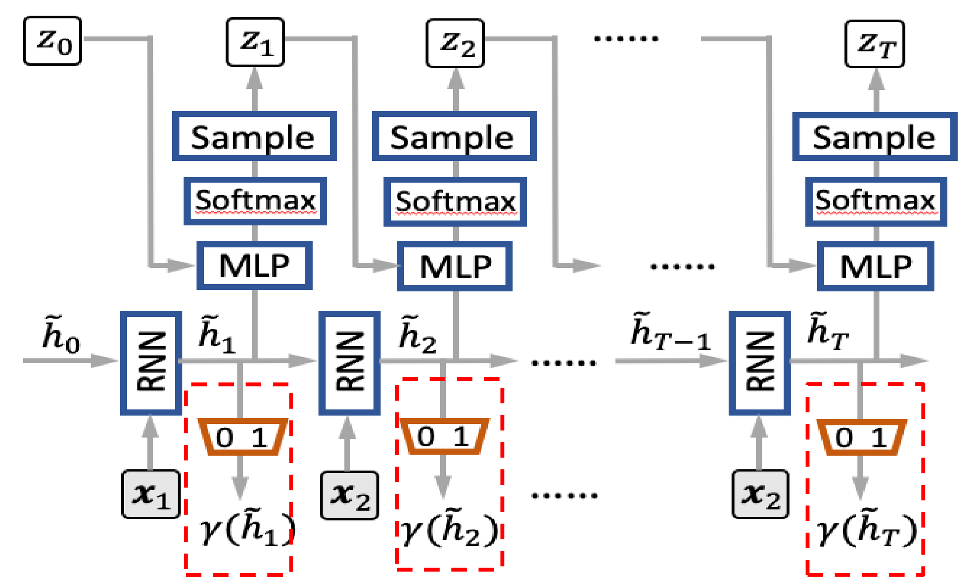
$$p(z_{t+1}|z_{1:t}) = f_{\theta}(z_{1:t}) \Rightarrow p(z_{t+1}|z_{1:t}) = softmax(MLP(\boldsymbol{h}_t)),$$
where $\boldsymbol{h}_t = RNN(z_t, \boldsymbol{h}_{t-1})$



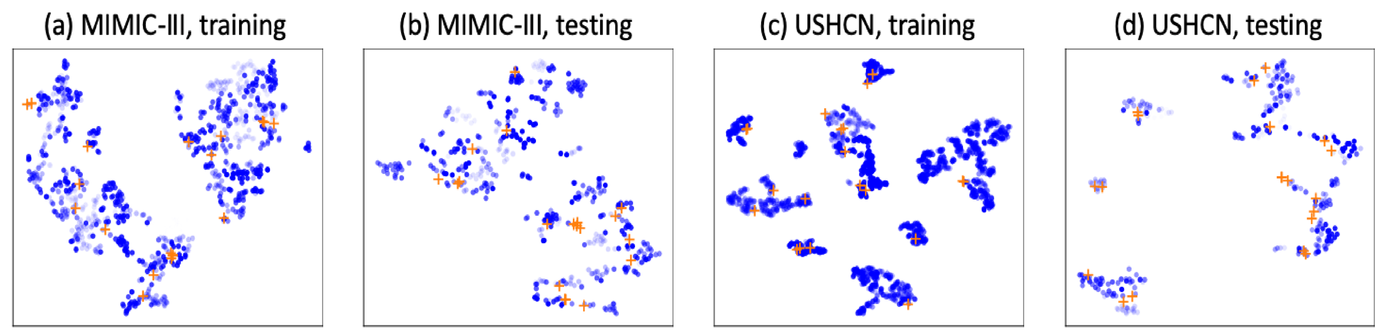
Inference network

Variational inference:

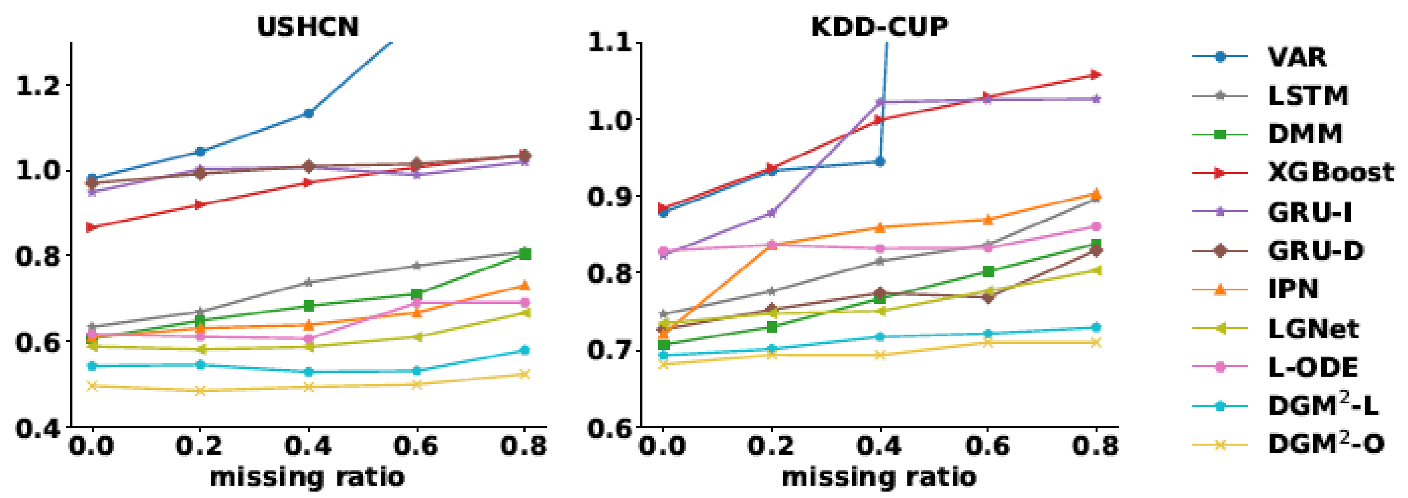
- introduce approximate posterior $q_{\phi}(\boldsymbol{z}_{1:T}|\boldsymbol{x}_{1:T}) = q_{\phi}(\boldsymbol{z}_1|\boldsymbol{x}_1)\sum_{t=1}^{T-1} q_{\phi}(\boldsymbol{z}_{t+1}|\boldsymbol{x}_{1:t+1}, \boldsymbol{z}_t)$



Experiments



DGM² is able to fit the underlying clustering structures



DGM² can generate time series that are close to the true values