Variational Graph Recurrent Neural Networks

Introduction

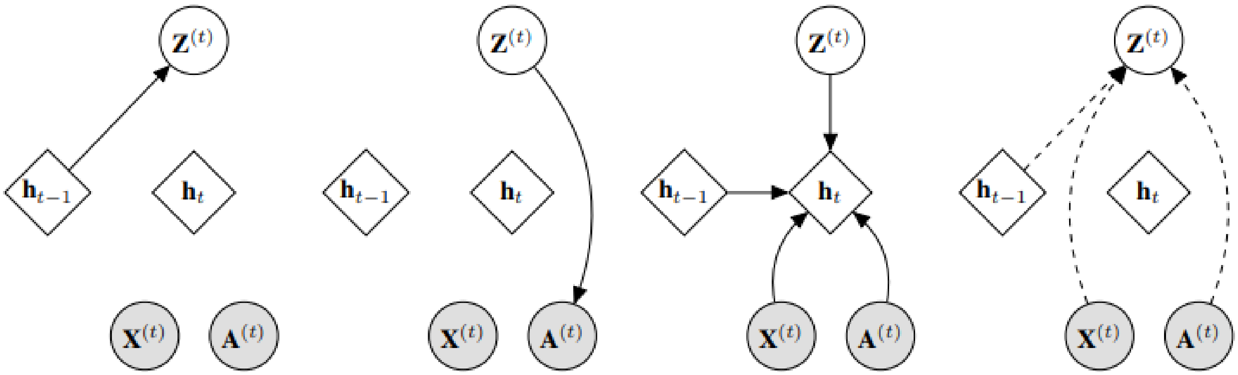
- Existing dynamic graph embedding approaches represent each node by a **deterministic** vector in
  a low-dimensional space
- Such deterministic representations lack the capability of modeling **uncertainty** of node embedding
- This paper first introduces dynamic graph autoencoder **(GRNN)**, then to increase the expressive power of GRNN in addition to modeling the uncertainty of node latent representations, variational graph recurrent neural network **(VGRNN)** is introduced

Background

- GCRN aims to model dynamic node attributes defined over a **static graph** whereas GRNN applies on **dynamic graph**
- GRNN reconstruct the graph at time t by **decoding** hidden state h_t
- By introducing VGRNN, we can capture time dependencies between graphs, but also each node is represented with a distribution in the latent space (similar to variational autoencoder).

VGRNN model

- VGAE in our VGRNN learns the prior distribution parameters based on the hidden states in previous time steps.
- Hence, our VGRNN allows more flexible latent representations with greater expressive power that captures dependencies between and within topological and node attribute evolution processes



(a) Prior (b) Generation (c) Recurrence (d) Inference

Figure 1: Graphical illustrations of each operation of VGRNN; (a) computing the conditional prior by (2); (b) decoder function (3); (c) updating the GRNN hidden states using (4); and (d) inference of the posterior distribution for latent variables by (3.2).

- Objective functions:
  The first term is the reconstruction lost, while second term is approximating prior to posterior

$$\mathcal{L} = \sum_{t=1}^{T} \Big\{ \mathbb{E}_{\mathbf{Z}^{(t)} \sim q\left(\mathbf{Z}^{(t)} \mid \mathbf{A}^{(\leq t)}, \mathbf{X}^{(\leq t)}, \mathbf{Z}^{(<t)}\right)} \log p\left(\mathbf{A}^{(t)} \mid \mathbf{Z}^{(t)}\right) - \mathbf{KL}\left(q\left(\mathbf{Z}^{(t)} \mid \mathbf{A}^{(\leq t)}, \mathbf{X}^{(\leq t)}, \mathbf{Z}^{(<t)}\right) \| \, p\left(\mathbf{Z}^{(t)} \mid \mathbf{A}^{(<t)}, \mathbf{X}^{(<t)}, \mathbf{Z}^{(<t)}\right)\right) \Big\}. \tag{7}$$

Reference

[Original paper]