DTU Compute
Department of Applied Mathematics and Computer Science
Technical University of Denmark

Piecewise-Velocity Model for Learning Continuous-time Dynamic Node Representations

Abdulkadir Çelikkanat, Nikolaos Nakis, Morten Mørup

LOG CONFERENCE

Introduction

Graph representation learning (GRL) aims to encode the structure of a given network into low-dimensional vectors.

Applications: classification, community detection and link prediction.

In this work:

- We propose a novel scalable GRL method to flexibly learn **continuous-time dynamic node representations**.
- It balances the trade-off between the smoothness of node trajectories and model capacity accounting for the temporal evolution.
- It can embed nodes accurately in very low dimensional spaces (D=2).
- We show that it outperforms well-known baseline methods.

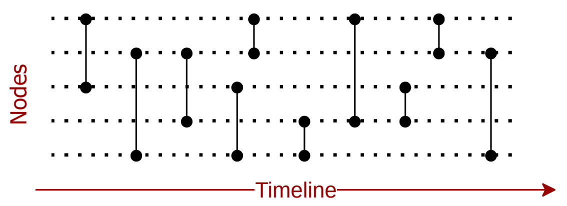


Illustration of the links of a continuous-time dynamic network

Proposed Approach

Nonhomogeneous Poisson Point Process

- We assume that the number of links follows a **Nonhomogeneous Poisson Point Process** with intensity function $\lambda_{ij}(t)$ on the time interval $[t_l, t_u)$.
- Then, the log-likelihood function can be written by

$$\mathcal{L}(\Omega) := \log p(\mathcal{G}|\Omega) = \sum_{\substack{i<j \\ i,j \in \mathcal{V}}} \left(\sum_{e_{ij} \in \mathcal{E}_{ij}} \log \lambda_{ij}(e_{ij}) - \int_0^T \lambda_{ij}(t)dt \right)$$

where $\mathcal{E}_{i,j} \subseteq \mathcal{E}[0,T]$ is the set of links of (i,j)-pair on the timeline $[0,T]$, and $\Omega = \{\lambda_{ij}\}_{1 \leq i < j \leq N}$ is the set of intensity functions.

Problem Formulation

Objective. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a continuous-time dynamic network and $\lambda^* : \mathcal{V}^2 \times \mathcal{I}_T \longrightarrow \mathbb{R}$ be an unknown intensity function of a nonhomogeneous Poisson point process. For a given metric space $(\mathsf{X}, d_{\mathsf{X}})$, our purpose is to learn a function or representation $\mathbf{r} : \mathcal{V} \times \mathcal{I}_T \to \mathsf{X}$ satisfying

$$\frac{1}{(t_u - t_l)} \int_{t_l}^{t_u} \psi^+\Big(d_{\mathsf{X}}(\mathbf{r}(i,t), \mathbf{r}(j,t))\Big)\, dt \approx \frac{1}{(t_u - t_l)} \int_{t_l}^{t_u} \lambda^*(i,j,t)dt$$

for a continuous function $\psi^+ : \mathbb{R} \to \mathbb{R}^+$ for all $(i,j) \in \mathcal{V}^2$ pairs, and for every interval $[t_l, t_u] \subseteq \mathcal{I}_T$.

PiVeM: Piecewise-Velocity Model

We learn continuous-time node representations by defining the intensity function by

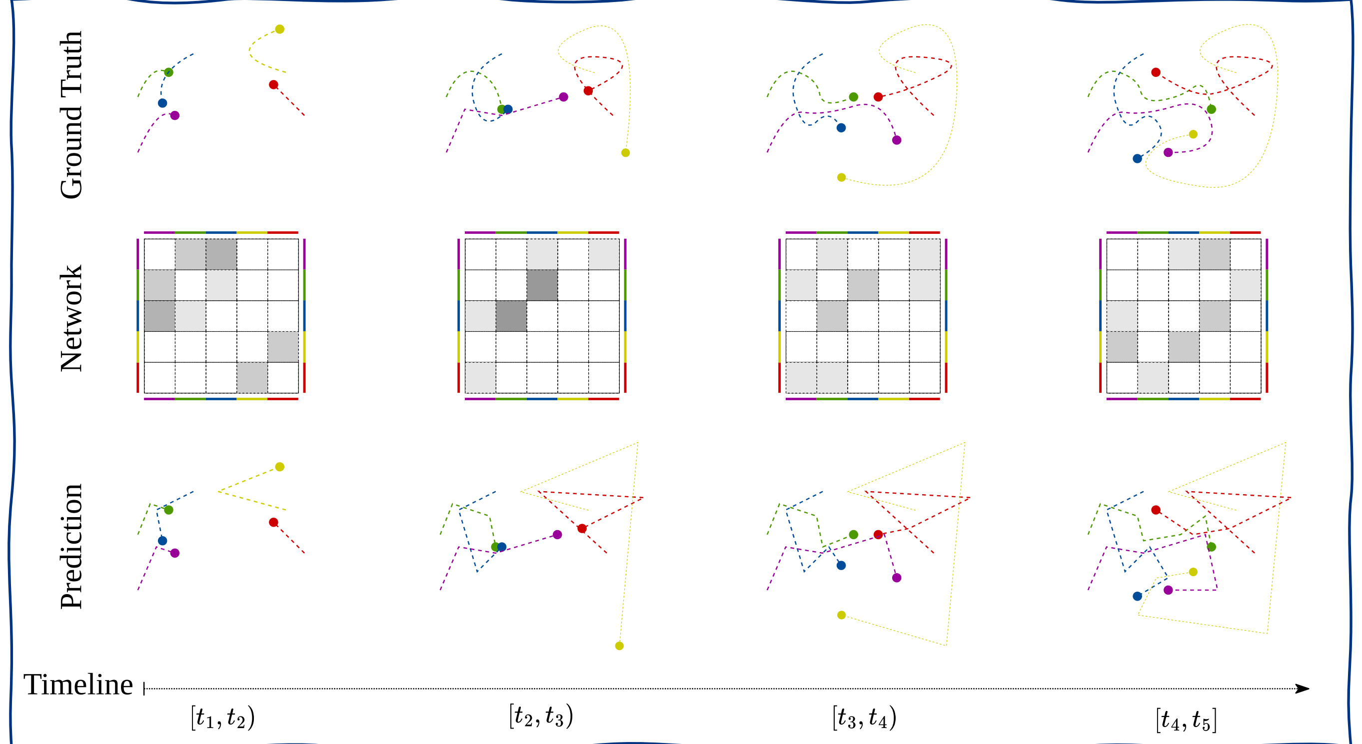
$$\lambda_{ij}(t) := \exp\left(\beta_i + \beta_j - ||\mathbf{r}_i(t) - \mathbf{r}_j(t)||^2\right)$$

where $\mathbf{r}_i(t) \in \mathbb{R}^D$ and $\beta_i \in \mathbb{R}$ denote the **embedding vector** at time t and the **bias term** of node $i \in \mathcal{V}$, respectively.

We define the representation of node $i \in \mathcal{V}$ at time $t \in [0,T]$ as follows:

$$\mathbf{r}_i(t) := \mathbf{x}_i^{(0)} + \Delta_B \mathbf{v}_i^{(1)} + \Delta_B \mathbf{v}_i^{(2)} + \cdots + (t \bmod(\Delta_B))\mathbf{v}_i^{(\lfloor t/\Delta_B \rfloor + 1)}$$

where $\mathbf{x}_i^{(0)}$ is the initial position, $\mathbf{v}_i^{(b)}$ the velocity for bin $b \in \{1, \ldots, B\}$ and Δ_B is the bin width.



Illustrative comparison of the ground-truth embeddings, the adjacency matrices constructed based on aggregating the links within the intervals and learned node representations.

Prior probability

We suppose that $\text{vect}(\mathbf{x}^{(0)}) \oplus \text{vect}(\mathbf{v}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma} := \lambda^2(\sigma_\Sigma^2 \mathbf{I} + \mathbf{K})$ is the covariance matrix with a scaling factor λ.

$\mathbf{K} := \mathbf{B} \otimes \mathbf{C} \otimes \mathbf{D}$ accounts for temporal-, node-, and dimension specific covariance structures.

We can express our objective relying on the piecewise velocities with the prior by:

$$\hat{\Omega} = \arg\max_{\Omega} \sum_{\substack{i<j \\ i,j \in \mathcal{V}}} \left(\sum_{e_{ij} \in \mathcal{E}_{ij}} \log \lambda_{ij}(e_{ij}) - \int_0^T \lambda_{ij}(t)dt \right) + \log \mathcal{N}\left(\begin{bmatrix} \mathbf{x}^{(0)} \\ \mathbf{v} \end{bmatrix}; \mathbf{0}, \boldsymbol{\Sigma} \right)$$

Experiments

Network reconstruction

	Synthetic(π)		*Synthetic(μ)*		*College*		*Contacts*		*Email*		*Forum*		*Hypertext*	
	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR
LDM	.563	.539	.669	.642	**.951**	.944	.860	.835	.954	.948	**.909**	.897	.818	.797
Node2Vec	.519	.507	.503	.509	.711	.655	.812	.756	.853	.828	.677	.619	.696	.648
CTDNE	.613	.580	.539	.544	.661	.622	.787	.760	.854	.840	.657	.622	.725	.725
HTNE	.614	.591	.599	.571	.721	.683	.846	.823	.871	.867	.723	.691	.775	.787
MMDNE	.582	.565	.600	.576	.725	.692	.844	.825	.867	.863	.737	.712	.778	.787
PiVeM	**.762**	**.713**	**.905**	**.869**	.948	**.948**	**.938**	**.938**	**.978**	**.977**	.907	**.902**	**.830**	**.823**

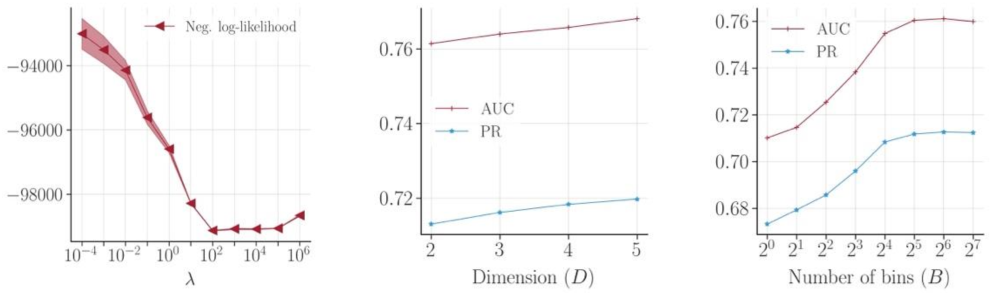
Network completion

	Synthetic(π)		*Synthetic(μ)*		*College*		*Contacts*		*Email*		*Forum*		*Hypertext*	
	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR
LDM	.535	.529	.646	.631	.931	.926	.836	.799	.948	.942	.863	.858	.761	**.738**
Node2Vec	.519	.511	.747	.677	.685	.637	.787	.744	.818	.777	.635	.592	.596	.588
CTDNE	.608	.573	.531	.539	.601	.556	.752	.703	.831	.812	.568	.539	.554	.537
HTNE	.605	.583	.573	.557	.673	.651	.792	.759	.853	.834	.596	.581	.602	.633
MMDNE	.587	.570	.592	.571	.677	.662	.819	.811	.844	.829	.596	.570	.587	.614
PiVeM	**.750**	**.696**	**.874**	**.851**	**.935**	**.934**	**.873**	**.864**	**.951**	**.953**	**.879**	**.875**	**.770**	.712

Network prediction

	Synthetic(π)		*Synthetic(μ)*		*College*		*Contacts*		*Email*		*Forum*		*Hypertext*	
	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR	ROC	PR
LDM	.562	.539	.498	**.642**	**.951**	**.944**	.860	.835	.954	.948	**.909**	**.897**	**.819**	.797
Node2Vec	.518	.506	.498	.502	.705	.676	.783	.716	.825	.807	.635	.605	.748	.739
CTDNE	.680	.629	.481	.487	.691	.711	.842	.815	.824	.815	.664	.642	.699	.734
HTNE	.573	.569	.491	.493	.715	.684	.864	.824	.838	.837	.764	.747	.785	**.820**
MMDNE	.591	.575	**.506**	.515	.717	.703	.874	.847	.827	.832	.762	.746	.795	.813
PiVeM	**.716**	**.689**	.474	.485	.891	.887	**.876**	**.884**	**.964**	**.964**	.894	.895	.756	.767

Influence of the model hyperparameters.



Datasets

	$\lvert\mathcal{V}\rvert$	M	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{E}_{ij}\rvert_{max}$
Synthetic(μ)	100	4,889	180,658	124
Synthetic(π)	100	3,009	22,477	32
College	1,899	13,838	59,835	184
Contacts	217	4,274	78,249	1,302
Hypertext	113	2,196	20,818	1,281
Email	986	16,064	332,334	4,992
Forum	899	7,036	33,686	171

For the implementation, datasets, animations and other details, please visit the address:
https://abdcelikkanat.github.io/projects/pivem/

SCAN ME