# Learning Expanding Graphs for Signal Interpolation

Bishwadeep Das,[@*] Elvin Isufi[@*]

[@]b.das@tudelft.nl, [@]e.isufi@tudelft.nl, [*]Multimedia Computing Group, TU Delft

## Introduction

- Graph data processing, i.e., interpolation, classification usually done over fixed-size static and dynamic graphs.
- New nodes often emerge, increasing graph size, often without information (e.g. cold start recommendation).
- Existing works either assume known connectivity of new nodes, or independent attachment and data-processing.
- We propose a data-driven, task-aware stochastic attachment of new nodes, focusing on signal interpolation.

## Problem Formulation

- Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, adjacency matrix $\mathbf{A}$, signal $\mathbf{x}$. Directed edge of weight $w_i$ from $v_i \in \mathcal{V}$ to $v_+$ with prob $p_i$.
- Attachment vector $\mathbf{a}_+$ with $[\mathbf{a}_+]_k = w_k$ if $v_+$ connects to $v_k$. $\mathbf{a}_+$ has mean $\mathbf{w} \circ \mathbf{p}$, covariance $\mathbf{\Sigma}_+$. $\mathbf{p}$, $\mathbf{w}$ attachment parameters.
- Estimate $(\mathbf{p}, \mathbf{w})$ over training set $\mathcal{T} = \{(v_{t+}, x_{t+}, \mathbf{a}_{t+})\}_t$ with $x_{t+}$, $\mathbf{a}_{t+}$ observed signals and attachments.
- New adjacency matrix $\mathbf{A}_+ = \begin{bmatrix} \mathbf{A} & \mathbf{0}_+ \\ \mathbf{a}_+^\top & 0 \end{bmatrix}$

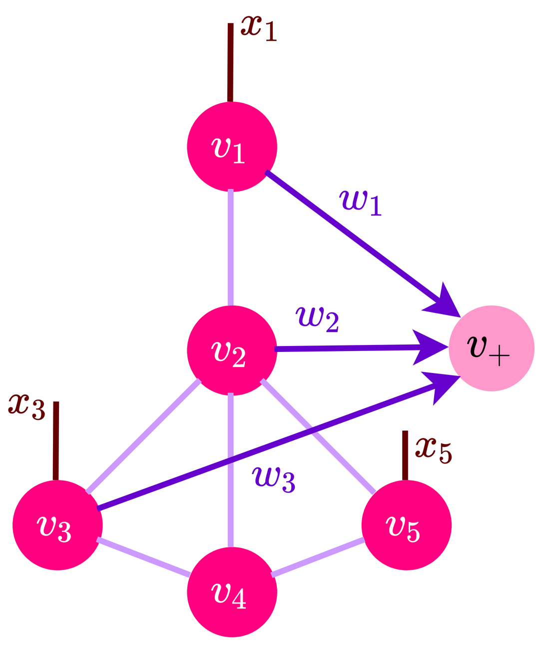


*Figure 1: Incoming node attaching to an existing graph of five nodes.*

## Signal Interpolation on Incoming Nodes

- Order $L$ graph filter with coefficients $\mathbf{h}$ and shift $\mathbf{A}_+$ outputs at $v_+$

$$y_+ = \mathbf{a}_+^\top \sum_{l=1}^{L} h_l \mathbf{A}_+^l \mathbf{x} = \mathbf{a}_+^\top \mathbf{A}_x \mathbf{h} \quad (1)$$

We now solve $(\mathbf{p}, \mathbf{w})$ via

$$\min_{\mathbf{p},\mathbf{w}} \ \text{MSE}_{\mathcal{T}}(\mathbf{p}, \mathbf{w}) + \sum_{t=1}^{|\mathcal{T}|} \left( \mu_p ||\mathbf{p} - \mathbf{b}_{t+}||_q^q + \mu_w ||\mathbf{w} - \mathbf{a}_{t+}||_q^q \right)$$
$$\mathbf{p} \in [0,1]^N, \mathbf{w} \in \mathcal{W}$$

$$\text{MSE}_{\mathcal{T}}(\mathbf{p}, \mathbf{w}) = \|(\mathbf{w} \circ \mathbf{p})^\top \mathbf{A}_x \mathbf{h} - x_+^\star\|_2^2 + \mathbf{h}^\top \mathbf{A}_x^\top \mathbf{\Sigma}_+ \mathbf{A}_x \mathbf{h} \quad (2)$$

- First term bias, second variance of prediction at $v_+$.
- Not always convex, solved using projected gradient descent

## Results

- *Synthetic*: ER and BA graphs with $100$ nodes, band-limited signals for order three filter.

| | Erdős-Rényi | | | Barabasi-Albert | | |
|---|---|---|---|---|---|---|
| | Prop. | Pref. | Rand. | Prop. | Pref. | Rand. |
| **MSE** | **0.03** | 0.06 | 0.06 | **0.05** | 0.1 | 0.08 |
| **Std.** | **0.003** | 0.003 | 0.003 | **0.006** | 0.006 | 0.006 |

*Table 1: Caption*

- Proposed outperforms topology-aware random attachments
- *Real*: Item cold start item collaborative filtering on Movielens 100K
- Trained on three sets of items-with low, medium and high number of interactions

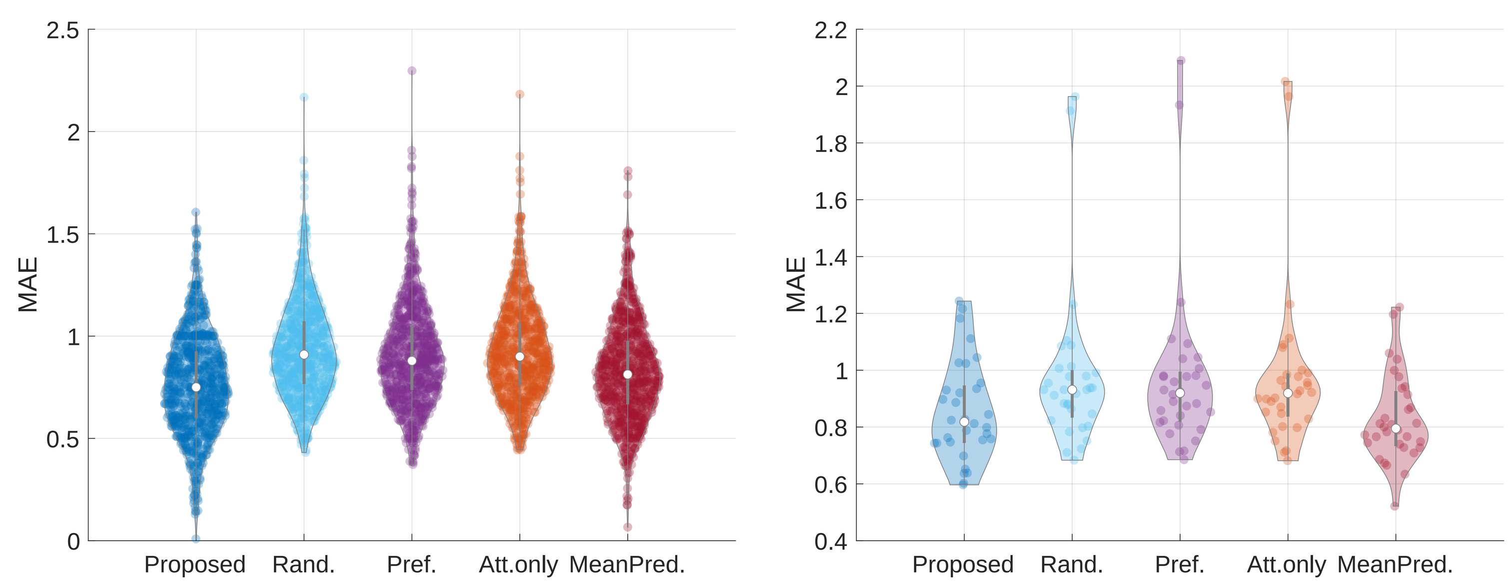


*Figure 2: [Left] Low items, [Right] high items*

- Lowest error with low items (challenging scenario), robust for items with many interactions. Shows benefit of proposed model and problem formulation.

## Conclusion

- We proposed learning a topology, data and task-aware attachment model for growing graphs.
- Solve for parameters via projected gradient descent.
- Outperform other attachment rules on real and synthetic data.
- Future work to address sequence of incoming nodes.

TU Delft — Delft University of Technology

Multimedia Computing Group