Transferable Hypergraph Neural Networks via Spectral Similarity

Mikhail Hayhoe, *Hans Riess*, Michael M. Zavlanos, Victor M. Preciado, Alejandro Ribeiro

Summary

Hypergraphs are higher-order graphs that describe relationships among arbitrarily sized groups. In this work, we introduce a novel Hypergraph Expansion Neural Networks (HENN) for machine learning on hypergraph-structured data by improving on recent stability results [1] on transferability of Graph Neural Networks (GNNs) to design a convolutional architecture for processing signals supported on hypergraphs. We establish bounds on the transferability error by exploiting multiple spectrally-similar graph representations of hypergraphs. Experimental results show the importance of considering multiple graph representations in HENNs and suggest superior performance when transferability is required.

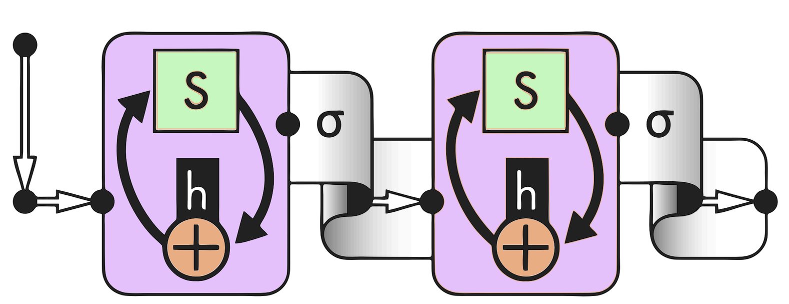
Graph Signal Processing & Graph Neural Networks

A **graph** is a tuple $\mathscr{G} = (\mathscr{V}, \mathscr{E}, W)$ where $\mathscr{V}$ is a set of n nodes, $\mathscr{E}$ is a set of m edges, and $W : \mathscr{E} \to \mathbb{R}$ is a weighting function

A **graph shift operator** (GSO) is a matrix $S \in \mathbb{R}^{n \times n}$ respecting the sparsity of $\mathscr{G}$, i.e. $[S]_{i,j} = 0$ if $(i, j) \notin \mathscr{E}$.

A **graph filter** with filter coefficients $\{h_k\}$is defined $H(S) = \sum_{k=0}^{\infty} h_k S^k$.

Graph neural networks (GNNs) are a cascade of L layers of graph filters, each followed by a pointwise nonlinearity. Let each layer have f_l graph signals (or features) $\mathbf{x}_l^1, \mathbf{x}_l^2, \ldots, \mathbf{x}_l^{f_l} \in \mathbb{R}^n$. At layer l, we apply $f_l f_{l-1}$ graph filters of the form $H_l^{ij}(S)$ followed by a pointwise (or elementwise) nonlinear function $\sigma : \mathbb{R} \to \mathbb{R}$ to process the f_{l-1} input features into the f_l output features via $\mathbf{x}_l^i = \sigma\left(\sum_{j=1}^{f_{l-1}} \sum_{k=0}^{\infty} h_{lk}^{ij} S^k \mathbf{x}_{l-1}^j\right), \quad i \in \{1,\ldots,f_l\}$.



Schematic of a (graph filter based-) Graph Neural Network (GNN)

Spectral Similarity

We say S and $\tilde{S}$ are ϵ-**spectrally similar** if

$$(1-\epsilon)S \preceq \tilde{S} \preceq (1+\epsilon)S$$

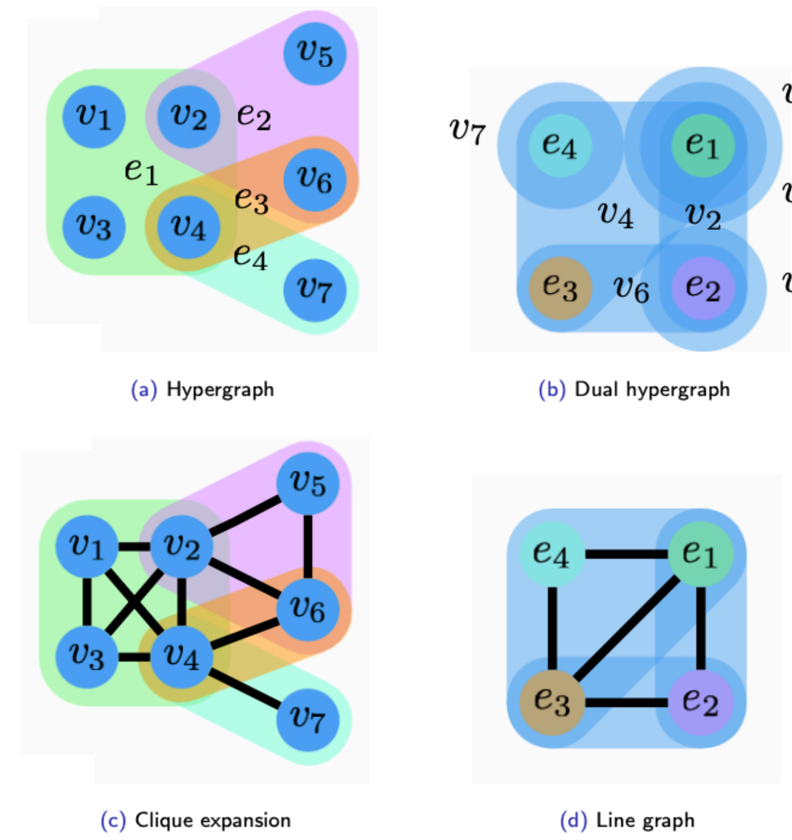
or equivalently,

$$(1-\epsilon)\mathbf{x}^\top S \mathbf{x} \leqslant \mathbf{x}^\top \tilde{S} \mathbf{x} \leqslant (1+\epsilon)\, \mathbf{x}^\top S \mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n .$$

Hypergraphs & Graph Representations

A **hypergraph** $\mathscr{H} = (\mathscr{V}, \mathscr{E})$ allows arbitrarily sized sets of nodes (not just two) in its hyperedges. Hypergraphs may have signals defined on the nodes $s_V = \{s_i\}_{i \in \mathscr{V}}$, and/or signals defined on the hyperedges $s_E = \{s_i\}_{i \in \mathscr{E}}$.

Graph representations, also called **expansions**, capture the essence of a hypergraph using a graph with some loss of information.



Hypergraph with several representations

The **energy (or total variation) of a node signal $\mathbf{x}$** on a hypergraph $\mathscr{H}$ with edge weight function $W(\cdot)$ is defined as

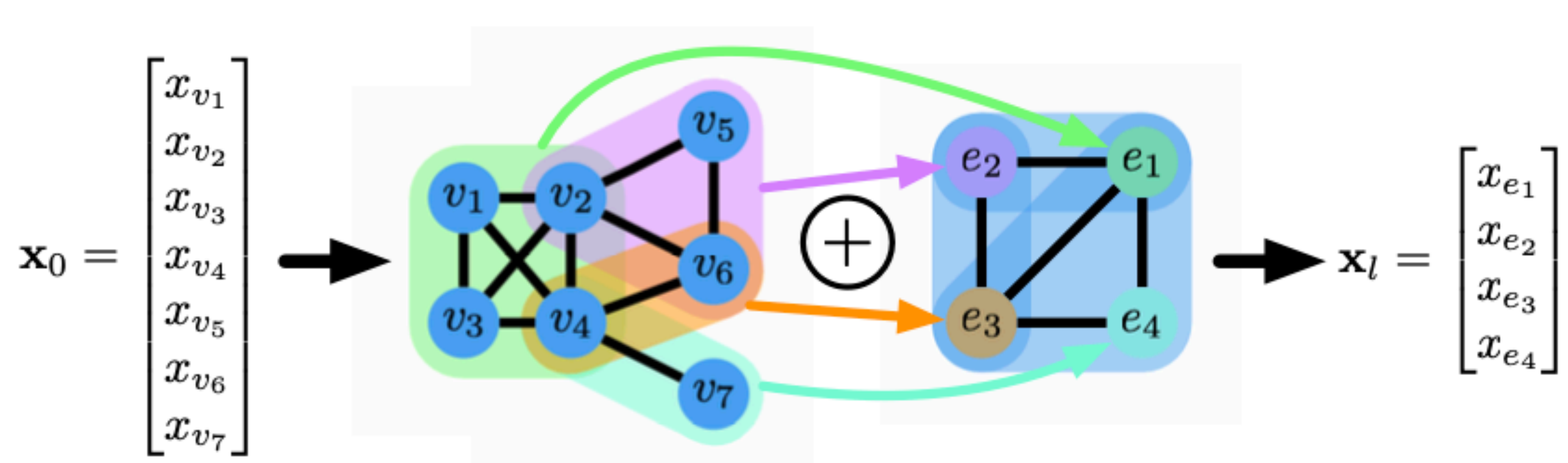
$$Q_{\mathscr{H}}(x) = \sum_{e \in \mathscr{E}} W(e) \max_{i,j \in e} (x_i - x_j)^2,$$

and the **hypergraph Laplacian** is then defined as the gradient of $Q_{\mathscr{H}}$, i.e.,

$$\mathscr{L}_{\mathscr{H}}(x) = \frac{1}{2} \nabla Q_{\mathscr{H}}(x)$$

Hypergraph Expansion Neural Networks (HENNs)

A **Hypergraph Expansion Neural Network (HENN)** is comprised of GNNs that use graph representations, pooled together. It may be comprised of an number of arbitrary graph representations. The example below uses the clique expansion and line graph with max-pooling between the GNNs to emulate diffusions via the hypergraph Laplacian:



Architecture of a Hypergraph Expansion Neural Network (HENN)

Results

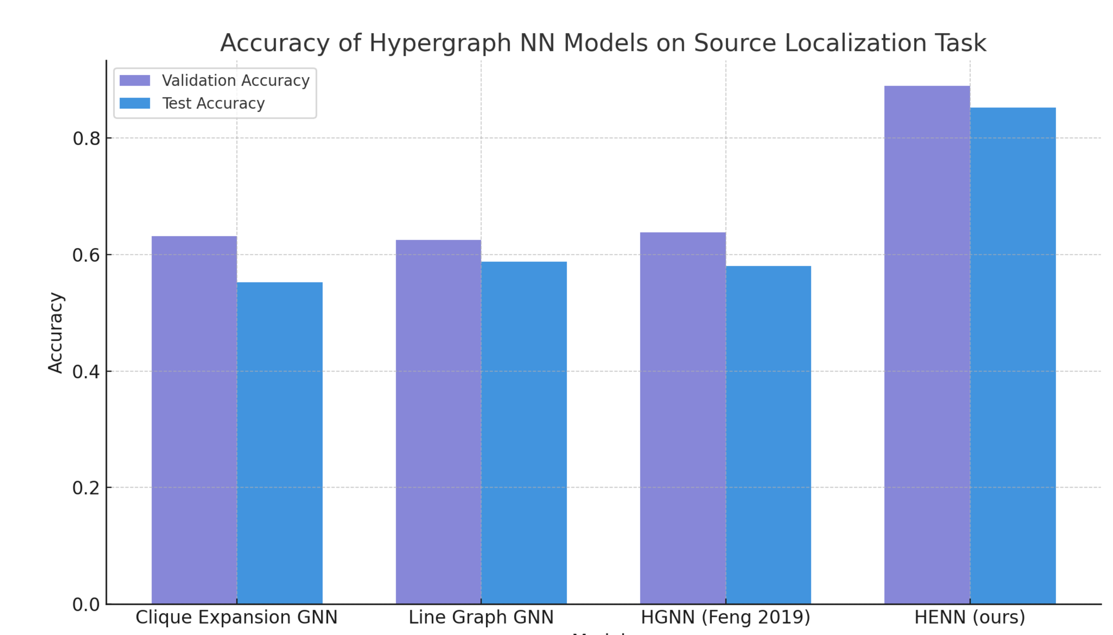
Our main result is a transferability guarantee between hypergraph expansion neural networks (HENNs) for hypergraphs that have arbitrary spectrally-similar graph representations.

Main Theorem. Consider two hypergraph $\mathscr{H}$ and $\tilde{\mathscr{H}}$ and r graph representations with GSOs $\{S_i\}_{i=1}^r$ and $\{\tilde{S}_i\}_{i=1}^r$ respectively, such that S_i and $\tilde{S}_i$ are ϵ_i-spectrally similar. If the hypergraph learning framework $\Phi(\cdot\,; \{S_i, \Theta_i\}_{i=1}^r)$ has a normalizing pooling function between graph representations with normalized Lipschitz nonlinearities, with L_i layers having f_i features for graph representation i, each with filters that have unit operator norm and are C-integral Lipschitz, then

$$\begin{aligned} &\|\Phi(\cdot\,; \{S_i, \Theta_i\}_{i=1}^r) - \Phi(\cdot\,; \{\tilde{S}_i, \Theta_i\}_{i=1}^r)\| \\ &\leq \sum_{i=1}^r C L_i \epsilon_i \prod_{j=1}^r f_j^{L_j} + O(\epsilon_1^2 + \cdots + \epsilon_r^2) . \end{aligned}$$

Experiments

We compared HENN with several other architectures on a hypergraph source localization problem, wherein the goal is to determine the source hyperedge of a diffusion process that has been progressing for an unknown period of time in the presence of both input and measurement noise. HENN shows over a 50% improvement in accuracy over individual graph representations alone, as well as improvement of 47% over HGNN [2] when filters are normalized and integral Lipschitz, suggesting substantially improved performance when transferability is desired.



References

[1] Gama, F., Bruna, J., & Ribeiro, A. (2020). Stability Properties of Graph Neural Networks. IEEE Transaction on Signal Processing 68, pp. 5680-5695.

[2] Feng, Y., You, H., Zhang, Z., Ji, R., & Gao, Y. (2019). Hypergraph Neural Networks. Proceedings of the AAAI Conference on Artificial Intelligence 33, pp. 3558-3565.