Information Flow in Graph Neural Networks

Pradeep Kr. Banerjee
MPI MiS
pradeep@mis.mpg.de

Institute for Data Science Foundations
Hamburg University of Technology (TUHH)
24th January, 2023

Information Flow in Graph Neural Networks

- **B.**, K. Karhadkar, Y.G. Wang, U. Alon and G. Montúfar, "Oversquashing in GNNs through the lens of information contraction and graph expansion."
  58th Annual Allerton Conference on Communication, Control, and Computing, 2022
- K. Karhadkar, **B.** and G. Montúfar, "FoSR: First-order spectral rewiring for addressing oversquashing in GNNs."
  International Conference on Learning Representations (ICLR), 2023

GNNs as message passing networks

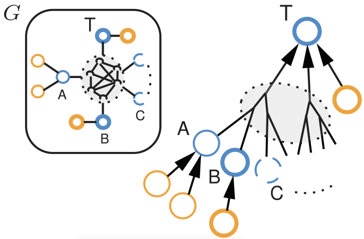
- Input to the GNN is a graph $G = (\mathcal{V}, \mathcal{E})$ endowed with node embeddings or features.
- G is used both as a part of the *data* and the *computational structure*.
- Each message-passing step is parameterized by a neural network layer.
- At each layer:
  - Every node computes a message and sends it to its neigbors.
  - Every node *aggregates* messages from its neighbors and *combines* them with its own representation.
- After L layers, a node's representation captures structural information within its L-hop neighborhood.

$$a_v^{(l)} = \text{AGGREGATE}^{(l)} \left(\left\{ h_u^{(l-1)} : u \in N_G(v) \right\} \right)$$
$$h_v^{(l)} = \text{COMBINE}^{(l)} \left(h_v^{(l-1)}, a_v^{(l)} \right)$$

Courtesy: Stefanie Jegelka, Representational Power of GNNs

- Unrolling the recursion at node v gives v's *computation graph*: A tree of depth L rooted at v that represents the L-hop neighborhood of v, where the children of any node u in the tree are the nodes adjacent to u.

Long-range tasks and information oversquashing



The NeighborsMatch problem (Alon and Yahav, 2021):

- Predict the label for the target node T. Correct label is the label of the blue node that has the same number of orange neighbors as T. In the example, the correct label is B.
- For all examples in the training dataset, there is a unique blue node with a matching number of neighbors as the target.
- With increasing L, recursive nature of the neighborhood aggregation process leads to *information oversquashing*, when an exponential amount of information is “compressed” into fixed-size node vectors.

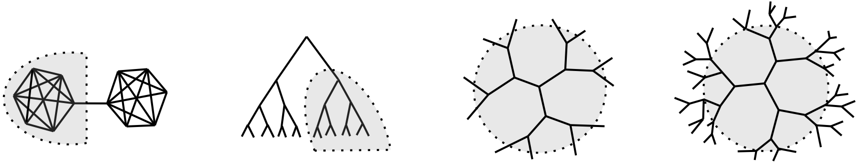
Structural bottlenecks and expander graphs

Definition

The isoperimetric ratio or the Cheeger constant of $G = (\mathcal{V}, \mathcal{E})$ *is*

$$h(G) = \min_{\mathcal{S} \subset \mathcal{V} \colon |\mathcal{S}| \leq n/2} \frac{|\partial \mathcal{S}|}{|\mathcal{S}|},$$

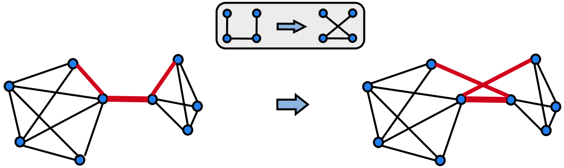
where $\partial \mathcal{S} = \{(u, v) \colon u \in \mathcal{S}, v \in \mathcal{V} \setminus \mathcal{S}, (u, v) \in \mathcal{E}\}$ *is the edge boundary of* $\mathcal{S} \subset \mathcal{V}$.
For fixed d *and* $\beta > 0$, *a* d*-regular graph* G *on* n *nodes is a* (n, d, β)*-expander if* $h(G) \geq \beta$.
An infinite family $(G_i)_{i \geq 1}$ *of* (n_i, d, β)*-expanders forms an expander family if* $h(G_i) \geq \beta \ \forall i$.



- Cheeger's inequality relates the spectral gap of G to it's Cheeger constant:

$$\frac{\lambda_2(L)}{2} \leq h(G) \leq \sqrt{2d\lambda_2(L)}, \quad L = dI - A.$$

Expansion using Greedy Random Local Edge Flips (G-RLEF)



- For d-regular input G on n nodes, repeatedly applying RLEF produces an expander in $O(d^2 n^2 \sqrt{\log n})$ steps with high probability (Allen-Zhu et al., 2016).
- Greedy version of RLEF: Sample the hub edge in proportion to their effective resistance.
- Intuition: Effective resistance captures the "electrical importance" of an edge. High resistance paths span structural bottlenecks and should be sampled with higher probability.
- For still faster convergence, optimize directly for the rate of graph expansion (FoSR, Karhadkar, **B.**, Montúfar, ICLR 2023).
- Navigate the Oversquashing vs. Oversmoothing trade-off using relational GNNs.

References I

Noga Alon and Vitali D Milman.
Eigenvalues, expanders and superconcentrators.
In *25th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 320–322, 1984.

Zeyuan Allen-Zhu, Aditya Bhaskara, Silvio Lattanzi, Vahab Mirrokni, and Lorenzo Orecchia.
Expanders via local edge flips.
In *Proc. 27th Annual ACM-SIAM Symposium on Discrete algorithms*, pages 259–269, 2016.

Pradeep Kr Banerjee, Kedar Karhadkar, Yu Guang Wang, Uri Alon, and Guido Montúfar.
Oversquashing in GNNs through the lens of information contraction and graph expansion.
In *58th Annual Allerton Conference on Communication, Control, and Computing*, pages 1–8, 2022.

Jeff Cheeger.
A lower bound for the smallest eigenvalue of the Laplacian.
Problems in Analysis, 625(195-199):110, 1970.

Shlomo Hoory, Nathan Linial, and Avi Wigderson.
Expander graphs and their applications.
Bulletin of the American Mathematical Society, 43(4):439–561, 2006.

Kedar Karhadkar, Pradeep Kr Banerjee, and Guido Montúfar.
FoSR: First-order spectral rewiring for addressing oversquashing in GNNs.
In *International Conference on Learning Representations (ICLR)*, 2023.