Edge-Based Graph Component Pooling

Thijs Snelleman, Bram M. Renting, Holger H. Hoos, Jan N. van Rijn

Motivation

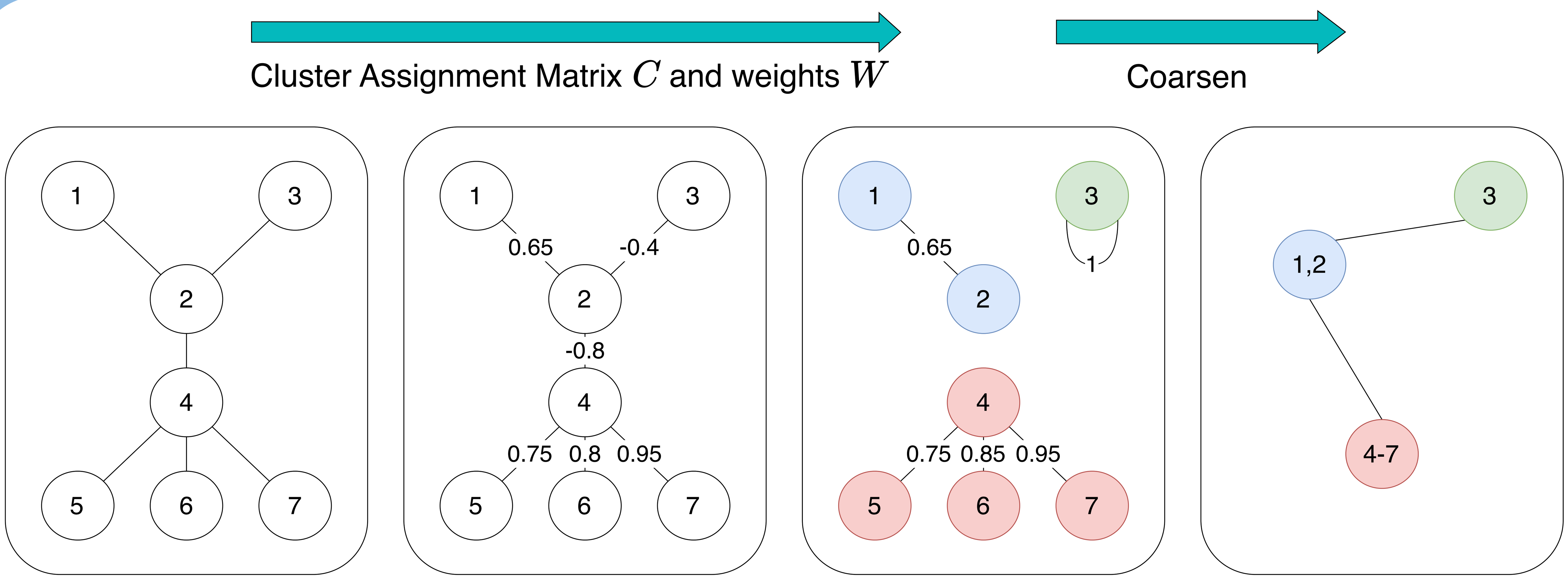
Message-passing propagates features to local neighbourhoods
- But **costly** for large and sparse graphs

Pooling improves efficiency by changing the graph structure
- Improves connectivity by reducing the longest path
- Crucial when handling <u>large, sparsely connected graphs</u>
- Two categories: Node drop pooling and cluster pooling

Contributions

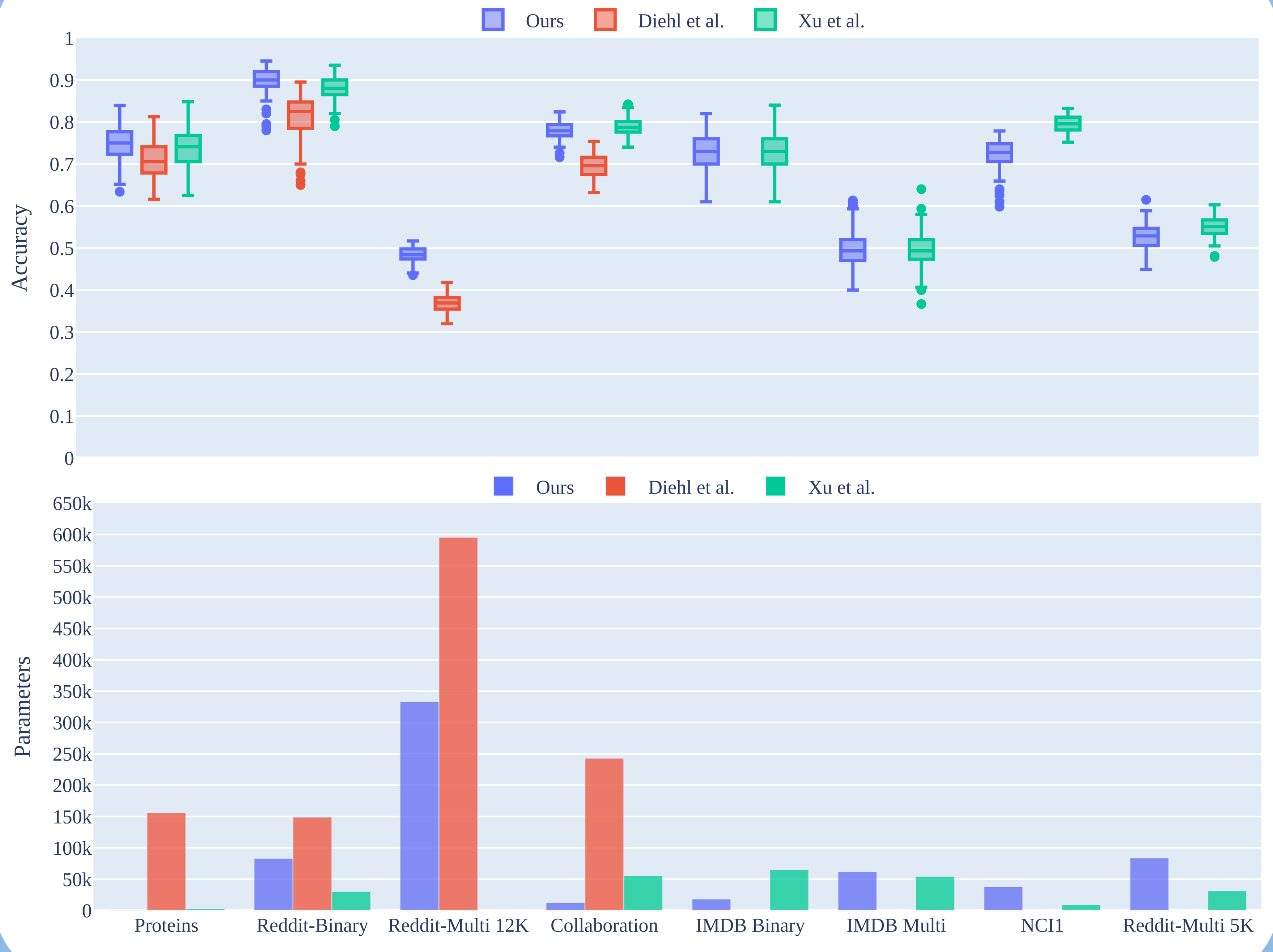
We propose an improved pooling method by Diehl et al. (2019):

Hard constraints removal: # nodes merged, # nodes per cluster
Reduced **complexity**: from $\mathcal{O}(n^2 \cdot \log(n))$ to $\mathcal{O}(n^2)$
Improved accuracy in seven benchmarks
Efficient implementation
Bridges the gap between drop pooling and cluster pooling
Comparable performance to message-passing GNN



$$W_{ij} = S_{ij} \quad \text{if } S_{ij} > t$$
$$W_{ij} = 1 \quad \text{if } i = j \text{ and not merging } i$$
$$W_{ij} = 0 \quad \text{otherwise}$$

$$S_{ij} = \sigma\left(\psi(x_i, x_j) + b\right)$$



Results

Compared to Diehl et al. (2019):
- Significantly outperforms Diehl et al. (2019) in terms of accuracy and learnable parameters

Compared to Xu et al. (2019):
- Higher accuracy on **two**, outperformed on **three** benchmarks
- Smaller model in three cases, but larger in four of which three very substantially larger

Concluding

- Efficient, maximally expressive, information retaining method
- Siginificant increase of accuracy and reduction of learnable parameters compared to the original method
- Competes with MP GNN, but is outperformed in several benchmarks in terms of accuracy and model size

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG