MaxCutPool: Differentiable Feature-Aware MAXCUT for Pooling in Graph Neural Networks

Carlo Abate[1,2], Filippo Maria Bianchi[3]

[1]Alma Mater Studiorum, University of Bologna, Italy [2]Fondazione Istituto Italiano di Tecnologia, Genoa, Italy [3]UiT The Arctic University of Norway, Tromsø, Norway

CODE

MAIN IDEA

- Adjacent nodes in a graph contain redundant information due to smoothing effects of message passing (MP)
- MAXCUT finds complementary groups of nodes by maximizing dissimilarity between connected nodes
- This enables improved information preservation while reducing graph size

KEY CONTRIBUTIONS

- MAXCUT computation for attributed graphs
- New hierarchical pooling layer especially effective for heterophilic graphs
- General scheme for node-to-supernode assignment
- First heterophilic dataset for graph classification

HETEROPHILIC MESSAGE PASSING

Consider the MP operator $X' = \sigma(PX\Theta)$.

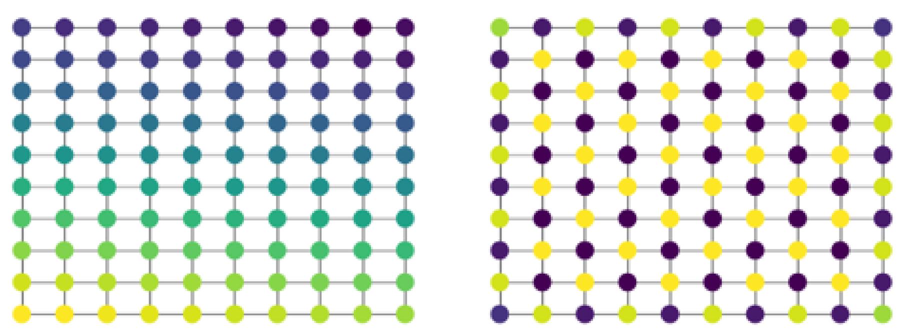
- Standard MP:

$$P = \hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}$$

- Heterophilic MP (HetMP):

$$P = I - \delta L^{\text{sym}}, \delta > 1$$

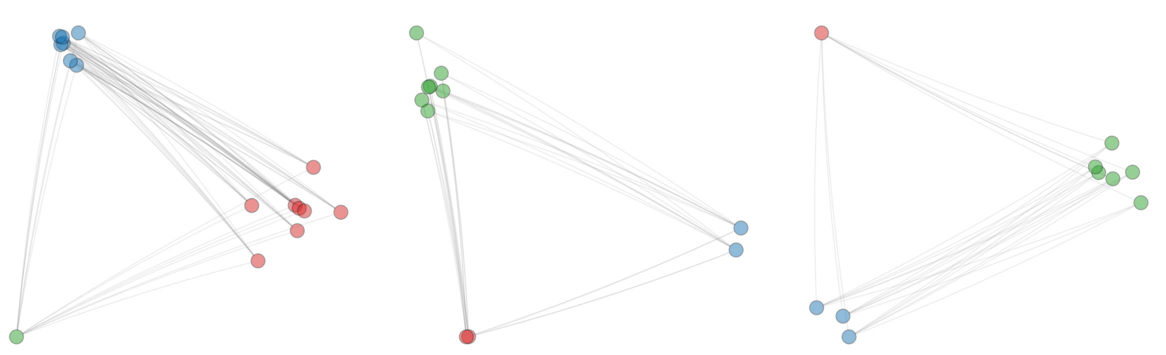
- HetMP can learn non-smooth graph signals



Left: standard MP; right: HetMP

MULTIPARTITE DATASET

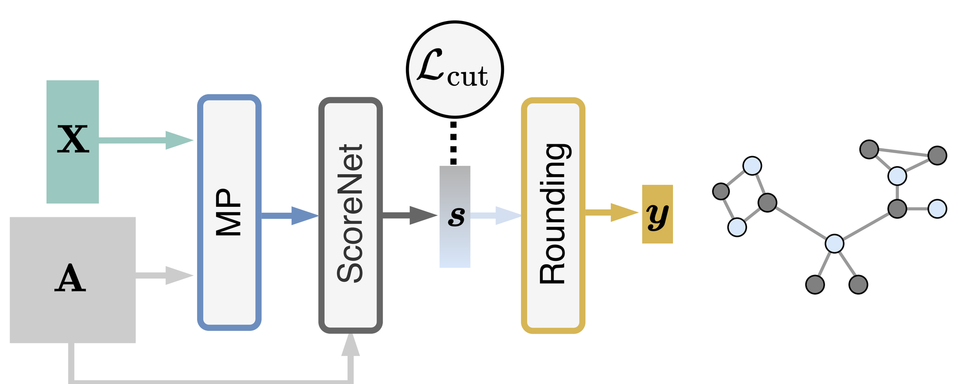
- First heterophilic benchmark for graph classification
- Complete C-partite graphs
- Nodes only connect to different-colored clusters
- Class determined by rightmost cluster color
- Tests GNNs to distinguish between relevant (node features) and irrelevant (connectivity) information



3-partite graphs

MAXCUT EVALUATION

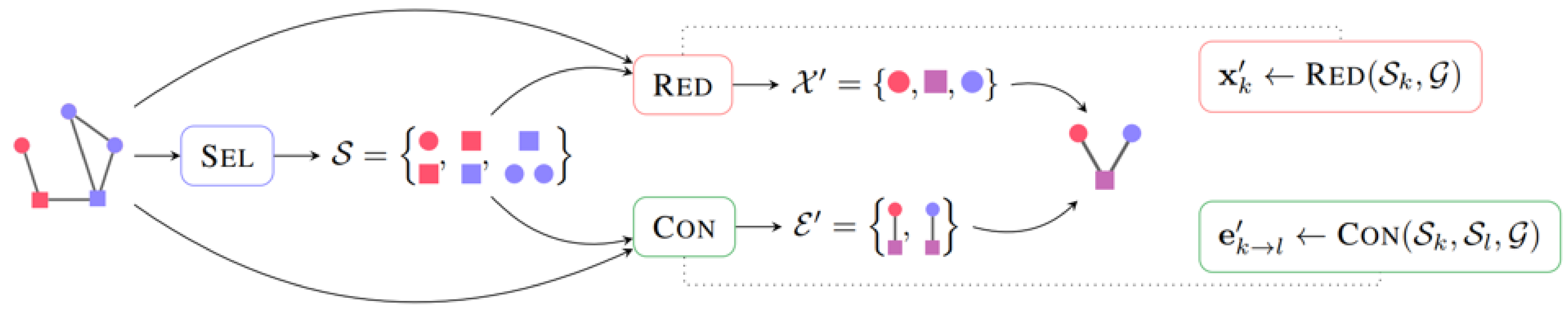
Dataset	GW	NDP	GCN	MaxCutPool
BarabasiAlbert	0.6875	0.6589	0.7240	0.7292
Community	0.6767	0.6429	0.6805	0.6814
ErdősRenyi	0.6920	0.6858	0.6797	0.7105
Grid (10×10)	1.0000	1.0000	0.9222	1.0000
Grid (60×40)	-	0.9787	0.1862	0.9815
Minnesota	-	0.9104	0.8904	0.9130
RandRegular	0.4827	0.8760	0.8733	0.9040
Ring	1.0000	1.0000	0.4200	1.0000
Sensor	0.6000	0.5719	0.6281	0.6406



REFERENCES

1. D. Grattarola et al., "Understanding Pooling in Graph Neural Networks," *IEEE TNNLS*, 2024

SELECT-REDUCE-CONNECT (SRC) [1]

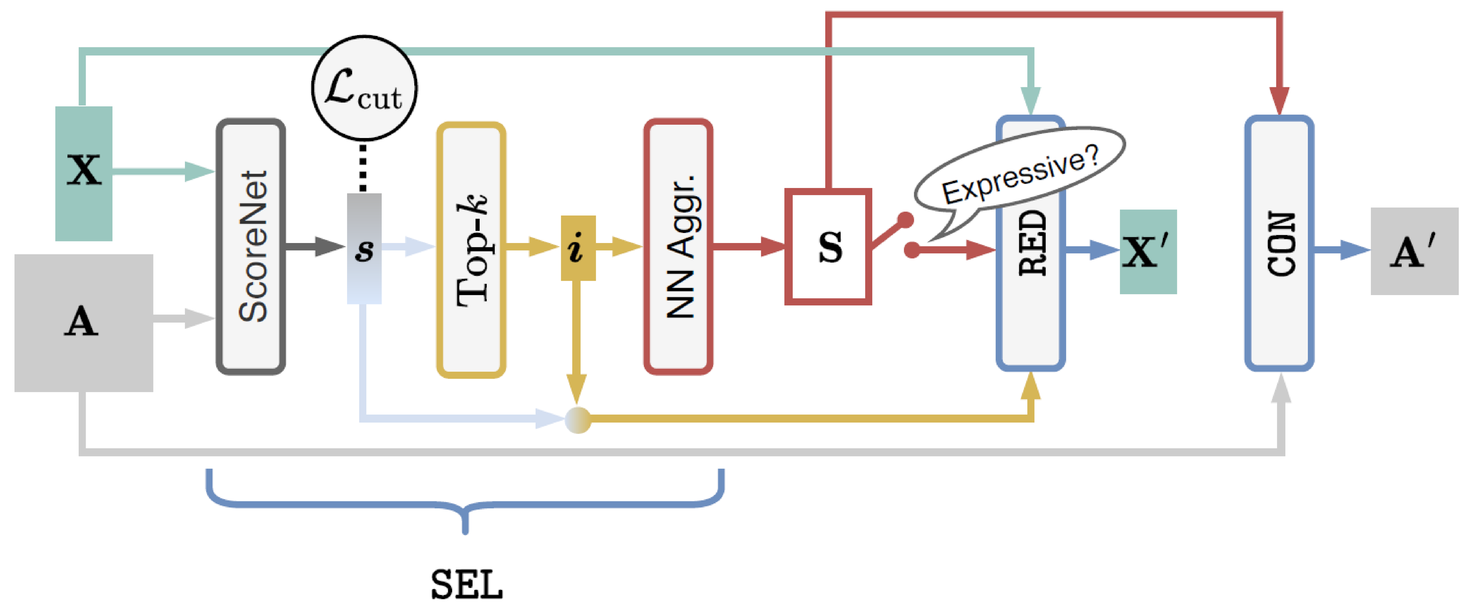


MAXCUTPOOL STRUCTURE

We introduce an auxiliary loss function defined as

$$\mathcal{L}_{\text{cut}} = \frac{s^\top A s}{|\mathcal{E}|}$$

where $s \in [-1, 1]^N$ is the score vector, A is the adjacency matrix and $|\mathcal{E}|$ is the total edge weight.



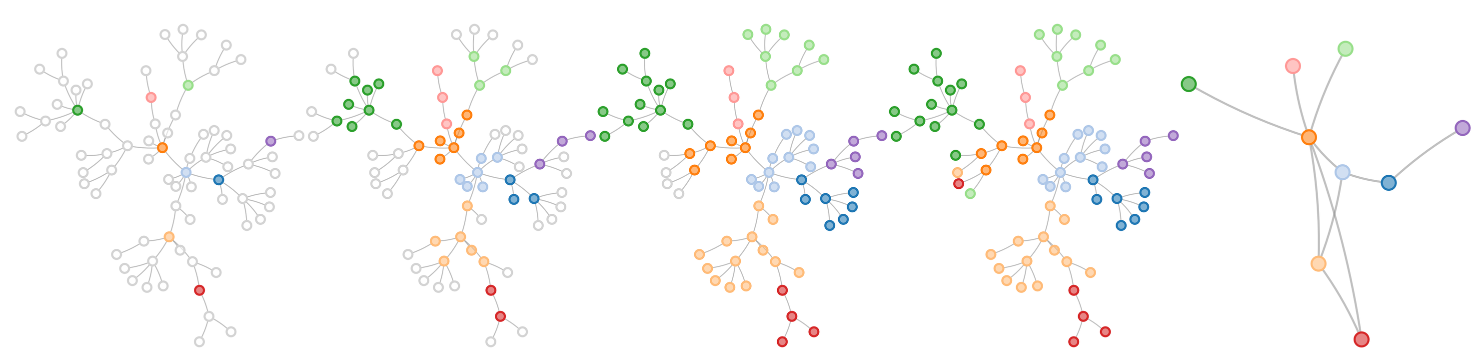
- Optimizes partition to maximize cut edges
- Connected nodes have opposite scores
- Enables end-to-end differentiable training
- Integrates with task-specific objectives

SELECT

- A ScoreNet with HetMP layers generates a score vector s
- Top-K scores identify supernodes: $i = \text{top}_K(s)$
- Builds assignment matrix S via breadth-first propagation
- Each node is assigned to the closest supernode

REDUCE

- MaxCutPool: $[X']_{i:} = s_i \odot [X]_{i:}$
- MaxCutPool-E: $X' = s \odot S^\top X$

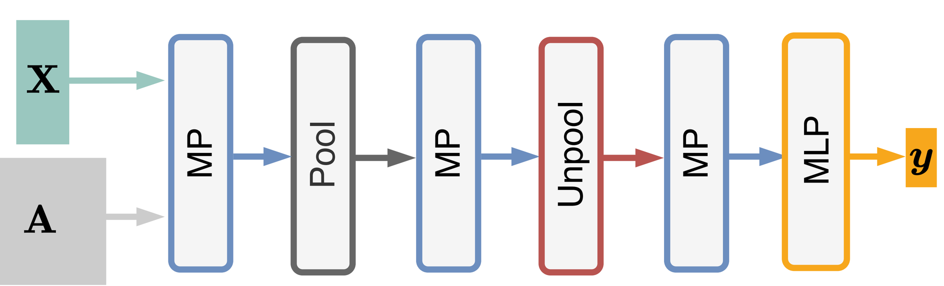


CONNECT

- Pooled adjacency: $A' = S^\top A S$

NODE CLASSIFICATION

Pooler	Roman-e.	Amazon-r.	Minesw.	Tolokers	Questions	Score
Top-k	26±7	46±4	94±1	89±5	64±3	1
k-MIS	23±3	48±2	75±2	84±2	83±1	1
NDP	22±5	53±2	98±0	88±6	68±4	3
MaxCutPool	56±3	53±1	96±1	87±3	82±4	4
MaxCutPool-E	60±4	53±2	97±1	91±2	85±5	**5**



GRAPH CLASSIFICATION

Pooler	GCB-H	COLLAB	EXPWL1	Mult.	Mutag.	NCI1	REDDIT-B	Score
DiffPool	51±8	70±2	69±3	9±1	78±2	75±2	90±2	1
DMoN	74±3	68±2	73±3	52±2	80±2	77±2	88±2	3
EdgePool	75±4	72±3	90±2	55±3	80±2	77±3	91±2	4
Graclus	75±3	72±3	90±2	25±18	80±2	77±2	90±3	4
k-MIS	75±4	71±2	99±1	58±2	79±2	75±3	90±2	4
MinCutPool	75±5	70±2	71±3	56±3	78±3	73±3	87±2	1
Top-k	56±5	72±2	73±2	43±3	75±3	73±2	77±2	0
MaxCutPool	73±3	77±2	100±0	90±2	77±2	75±2	89±3	5
MaxCutPool-E	74±3	77±2	100±0	87±5	79±1	76±2	89±2	**7**
MaxCutPool-NL	61±6	77±3	100±0	91±1	76±3	74±2	86±3	3