GNNs Getting ComFy: Community and Feature Similarity Guided Rewiring

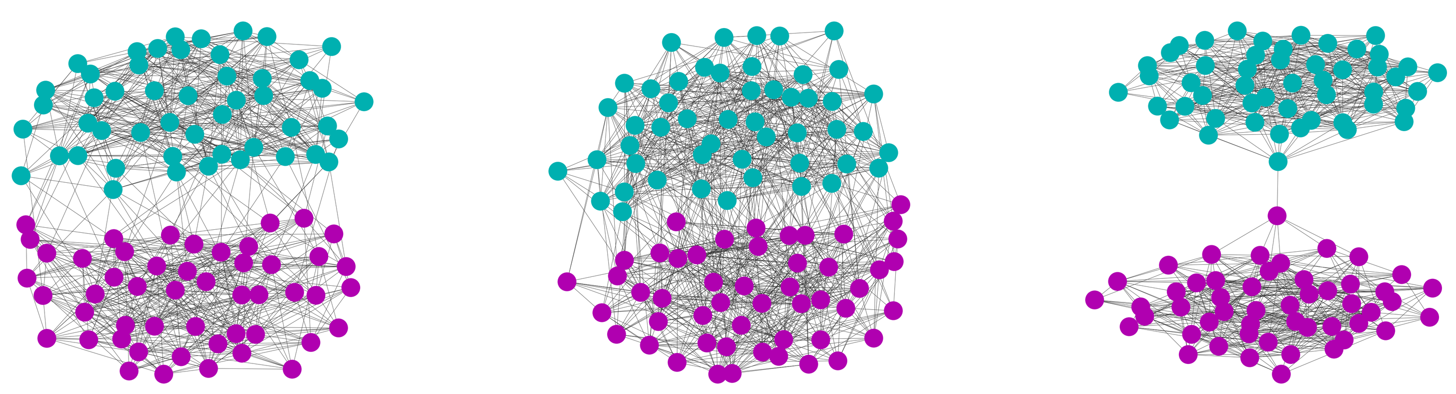
Celia Rubio-Madrigal*[1] Adarsh Jamadandi*[1,2] Rebekka Burkholz[1]

*Equal contribution [1]CISPA Helmholtz Center for Information Security [2]Universität des Saarlandes

Background

GNNs fail to capture long-range interactions due to topological bottlenecks resulting in over-squashing. Graph rewiring (e.g. spectral gap max.) can help alleviate such bottlenecks.

Insight: It can also destroy community structure, which can be harmful when node labels align well with communities.



(a) Original graph. (b) Maximization. (c) Minimization.

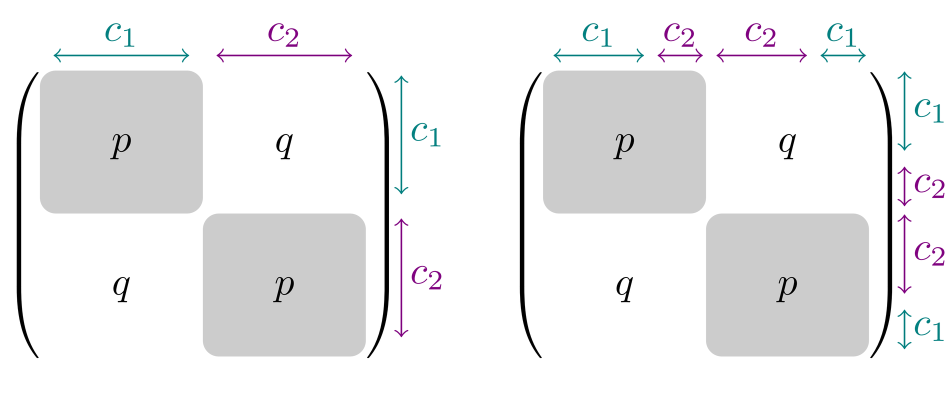
Research questions

- When is spectral gap maximization/minimization beneficial?
- How does the graph-task alignment ψ influence performance?
- Can the community structure of the graph and the node features form better graph rewiring criteria?

(p, q)-SBM adjacency matrices for different alignments ψ.

In Fig. (d), the blocks match classes c_1 and c_2, so $\psi = 1$.

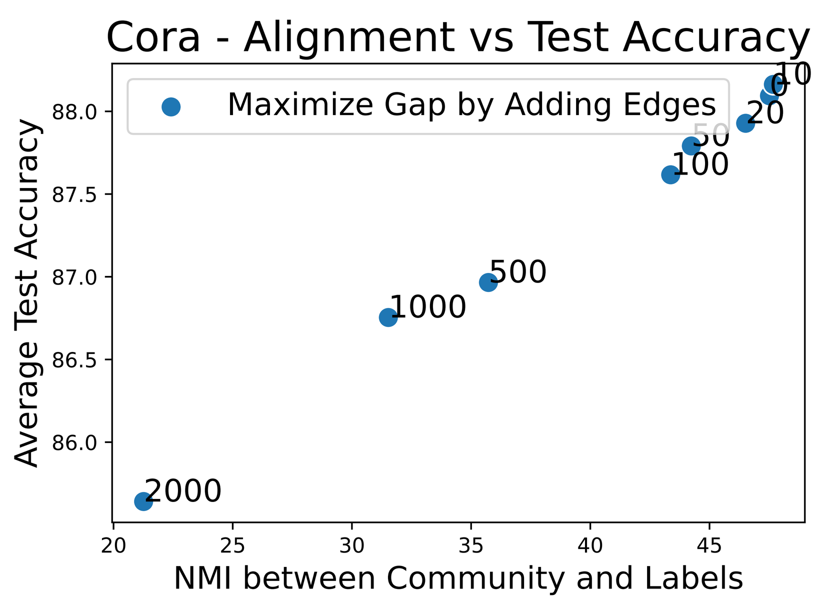
In Fig. (e), a third of each block is of opposite class, so $\psi = 2/3$.

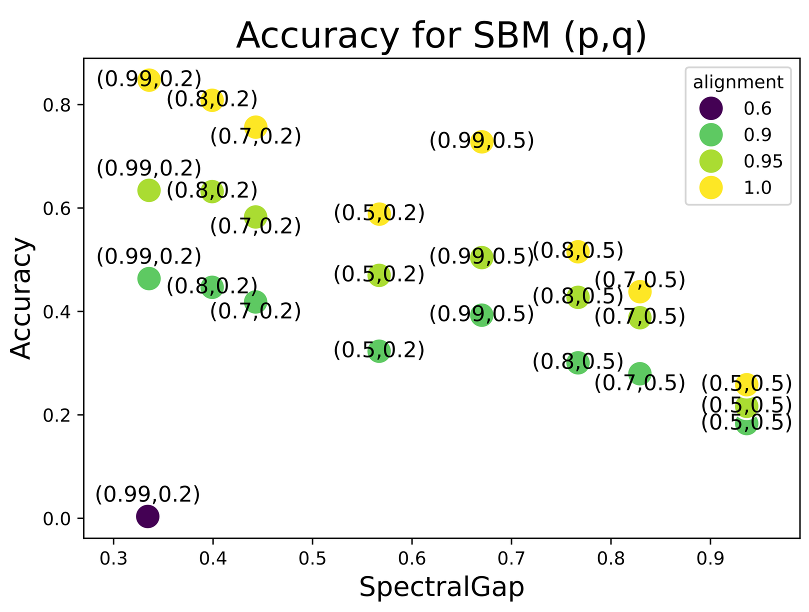


(d) $\psi = 1$. (e) $\psi = \frac{2}{3}$.

Theory on (p, q)-SBMs

1. Spectral gap $\lambda_1 \sim -(p-q)/(p+q)$. Maximizing it means $\downarrow p$ and $\uparrow q$, which destroys communities.
2. If task labels = community membership labels (high alignment), destroying them is harmful.
3. If the alignment changes, this is not necessarily the case. But spectral gap rewiring cannot tackle this.





Rewiring methods

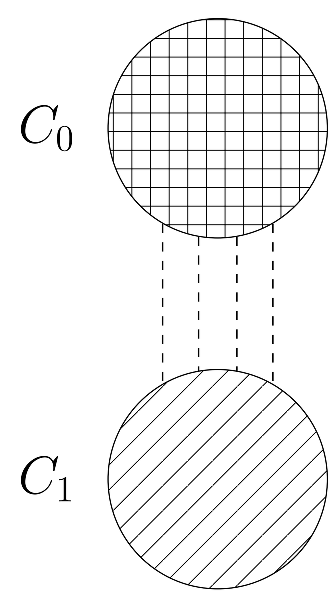
ComMa: Draws edges to $\uparrow$ or $\downarrow$ community strength.

FeaSt: Prioritizes edges that $\uparrow$ feature similarity.

ComFy: $\uparrow$ similarity proportionally to each community.

Experiments

Methods	Cora	Citeseer	Chameleon	Roman-Empire
GCN	86.12±0.36	77.83±0.35	39.33±0.59	70.30±0.73
GCN+SPAddMax	85.92±0.43	79.25±0.35	38.20±0.70	77.54±0.74
GCN+SPAddMin	84.10±0.39	78.77±0.40	39.33±0.55	79.18±0.06
GCN+SPDelMax	86.32±0.38	**81.84±0.38**	39.33±0.70	77.45±0.68
GCN+SPDelMin	85.92±0.37	79.01±0.34	39.89±0.59	79.09±0.05
GCN+FeaStAdd	87.73±0.39	78.54±0.34	43.26±0.62	**79.67±0.07**
GCN+FeaStDel	**90.74±0.39**	81.60±0.39	42.70±0.69	78.99±0.05
GCN+ComFyAdd	87.73±0.26	77.36±0.38	41.57±0.83	79.53±0.07
GCN+ComFyDel	88.13±0.27	78.07±0.35	**45.51±0.76**	79.17±0.07



	ComMa		FeaSt		ComFy	
	LowerComMa	HigherComMa	Add	Del	Add	Del
$C_0 \times C_0$	If Del, draw and delete $\lfloor \frac{A}{A+C} \cdot k \rfloor$ edges	If Add, draw and add $\lfloor \frac{A}{A+C} \cdot k \rfloor$ edges	Add top k of sim(u,v) for $(u,v) \in \bar{\mathcal{E}}$	Delete bottom k of sim(u,v) for $(u,v) \in \mathcal{E}$	FeaSt Add $\left(\lfloor \frac{A}{A+B+C} \cdot k \rfloor\right)$	FeaSt Del $\left(\lfloor \frac{A}{A+B+C} \cdot k \rfloor\right)$
$C_0 \times C_1$	If Add, draw and add k edges	If Del, draw and delete k edges	u v	u v	FeaSt Add $\left(\lfloor \frac{B}{A+B+C} \cdot k \rfloor\right)$	FeaSt Del $\left(\lfloor \frac{B}{A+B+C} \cdot k \rfloor\right)$
$C_1 \times C_1$	If Del, draw and delete $\lfloor \frac{C}{A+C} \cdot k \rfloor$ edges	If Add, draw and add $\lfloor \frac{C}{A+C} \cdot k \rfloor$ edges			FeaSt Add $\left(\lfloor \frac{C}{A+B+C} \cdot k \rfloor\right)$	FeaSt Del $\left(\lfloor \frac{C}{A+B+C} \cdot k \rfloor\right)$

Conclusions

We show that leveraging node features to rewire the graph can significantly boost GNN performance. Moreover, spectral gap based rewiring and other topology-based methods are insufficient because they fail to account for the alignment between the graph and the task.

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY

ICLR