MSGNN: A Spectral Graph Neural Network Based on a Novel Magnetic Signed Laplacian

Yixuan He [1] Michael Perlmutter [2]
Gesine Reinert [1,3] Mihai Cucuringu [1,3]

[1]Department of Statistics, University of Oxford, United Kingdom
[2]University of California, Los Angeles, United States
[3]The Alan Turing Institute, United Kingdom

Many important and interesting phenomena are naturally modeled as signed and/or directed graphs, i.e., graphs in which objects may have either positive or negative relationships, and/or in which such relationships are not necessarily symmetric. Examples:

- In the analysis of social networks, positive and negative edges could model friendship or enmity, and directional information could model the influence of one person on another [Kumar et al., 2016, Huang et al., 2021].
- Signed/directed networks also arise when analyzing time-series data with lead-lag relationships [Bennett et al., 2021], detecting influential groups in social networks [He et al., 2021], and computing rankings from pairwise comparisons [He et al., 2022a].

Motivation for Signed Directed GNNs

- Many signed networks are directed.
- Graph neural networks (GNNs) can be broadly classified into spatial or spectral.
- Most existing works on signed directed networks are spatial, such as SDGNN [Huang et al., 2021], SiGAT [Huang et al., 2019], SNEA [Li et al., 2020], and SSSNET [He et al., 2022b].
- A principal challenge for spectral GNNs: need to design a notion for signed directed Laplacian matrix.

- $\mathcal{V}$: a set of n nodes
- $\mathcal{E}$: the set of (directed) edges or links, that could be divided in positive and negative parts $\mathcal{E}^+$ and $\mathcal{E}^-$ for a signed network so that $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^-$
- adjacency matrix $\mathbf{A} = (A_{ij})_{i,j\in\mathcal{V}}$
- $\mathbf{X}_\mathcal{V} \in \mathbb{R}^{n\times d_{\text{in}}}$: node feature matrix
- For a *directed* network, $\mathbf{A}$ is usually *asymmetric*
- A **clustering** into K clusters: a partition of the node set into disjoint sets $\mathcal{V} = \mathcal{C}_0 \cup \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_{K-1}$
- Semi-supervised: seed nodes for each cluster
- Self-supervised: no label supervision

Problem Definition (cont.)

Link prediction tasks considered:

1. **Link sign prediction (SP)**: predict the edge sign of pairs of vertices u, v for which either $(u, v) \in \mathcal{E}^+$ or $(u, v) \in \mathcal{E}^-$.
2. **Direction prediction (DP)**: predict the edge direction of pairs of vertices u, v for which either $(u, v) \in \mathcal{E}$ or $(v, u) \in \mathcal{E}$.
3. **Three-class classification (3C)**: classify an edge $(u, v) \in \mathcal{E}$, $(v, u) \in \mathcal{E}$, or $(u, v), (v, u) \notin \mathcal{E}$.
4. **Four-class classification (4C)**: classify an edge $(u, v) \in \mathcal{E}^+$, $(v, u) \in \mathcal{E}^+$, $(u, v) \in \mathcal{E}^-$, or $(v, u) \in \mathcal{E}^-$;
5. **Five-class classification (5C)**: in addition to the classes in (4C), add a class $(u, v), (v, u) \notin \mathcal{E}$.

We evaluate the performance by **accuracy**.

- A novel magnetic signed Laplacian:
  - Hermitian
  - positive semidefinite
  - eigenvalues of its normalized counterpart lie in $[0, 2]$
  - reduce to existing Laplacians when the network is unsigned and/or undirected
- An efficient spectral GNN architecture, MSGNN, based on this magnetic signed Laplacian, which attains leading performance on extensive tasks:
  - node clustering
  - link prediction (sign + directionality)
- A novel synthetic model: Signed Directed Stochastic Block Model (SDSBM)
- New real-world data sets constructed from lead-lag relationships of financial time series data.

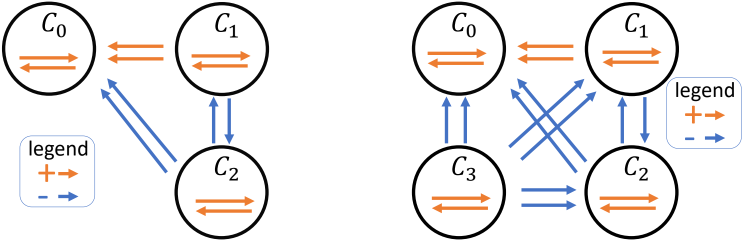


Figure: This toy example models groups of athletes and sports fans on social media. Here, signed, directed edges represent positive or negative mentions. $\mathcal{C}_0$: players of a sports team; $\mathcal{C}_1$: a group of their fans who typically say positive things about the players; $\mathcal{C}_2$: a group of fans of a rival team; $\mathcal{C}_3$: a group of fans of a third, less important team.

Hermitian Adjacency Matrix

- Phase matrix for direction distinguishment

$$\mathbf{\Theta}_{i,j}^{(q)} := 2\pi q(\mathbf{A}_{i,j} - \mathbf{A}_{j,i}), q \geq 0$$
$$\exp(\mathbb{i}\mathbf{\Theta}^{(q)})_{i,j} := \exp(\mathbb{i}\mathbf{\Theta}_{i,j}^{(q)})$$

- Hermitian adjacency matrix and the absolute degree matrix

$$\tilde{\mathbf{A}}_{i,j} := \frac{1}{2}(\mathbf{A}_{i,j} + \mathbf{A}_{j,i}), \;\; 1 \leq i, j \leq n,$$
$$\tilde{\mathbf{D}}_{i,i} := \frac{1}{2}\sum_{j=1}^{n}(|\mathbf{A}_{i,j}| + |\mathbf{A}_{j,i}|), \;\; 1 \leq i \leq n,$$
$$\mathbf{H}^{(q)} := \tilde{\mathbf{A}} \odot \exp(\mathbb{i}\mathbf{\Theta}^{(q)})$$

- Magnetic Signed Laplacian

$$\mathbf{L}_U^{(q)} := \tilde{\mathbf{D}} - \mathbf{H}^{(q)} = \tilde{\mathbf{D}} - \tilde{\mathbf{A}} \odot \exp(\mathbb{i}\boldsymbol{\Theta}^{(q)}),$$

- Normalized Magnetic Signed Laplacian

$$\mathbf{L}_N^{(q)} := \mathbf{I} - \left(\tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-1/2}\right) \odot \exp(\mathbb{i}\boldsymbol{\Theta}^{(q)}) \, .$$

Theorem 1
For all $q \geq 0$, both $\mathbf{L}_U^{(q)}$ and $\mathbf{L}_N^{(q)}$ are positive semidefinite.

Theorem 2
For all $q \geq 0$, the eigenvalues of $\mathbf{L}_N^{(q)}$ are contained in the interval $[0, 2]$.

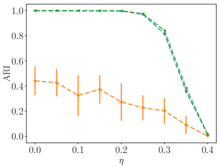
- MSGNN in ChebNet form [Defferrard et al., 2016]

$$Z = \sum_{k=0}^{K} T_k(\widetilde{\mathbf{L}})\mathbf{X}\Theta_k$$

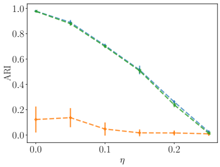
- The computation complexity is comparable with GCN [Kipf and Welling, 2017].

- Established synthetic data sets: Signed Stochastic Block Models (SSBMs) and polarized SSBMs (POL-SSBMs) introduced in [He et al., 2022b] (Signed Undirected Networks)
- A novel synthetic model: Signed Directed Stochastic Block Model (SDSBM)
- Standard real-world data sets: *BitCoin-Alpha*, *BitCoin-OTC* [Kumar et al., 2016], *Slashdot* [Ordozgoiti et al., 2020], and *Epinions* [Massa and Avesani, 2005]. These networks range in size from 3783 to 131580 nodes. Only *Slashdot* and *Epinions* are unweighted ($|w_{i,j}| = 1, \forall (v_i, v_j) \in \mathcal{E}$).
- Novel financial data sets from stock returns: *FiLL*

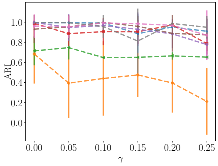
UNIVERSITY OF OXFORD



(a) SSBM$(n = 10000, C = 5, p = 0.01, \rho = 1.5)$



(b) Pol-SSBM$(n = 5000, r = 5, p = 0.1, \rho = 1.5)$



(c) SDSBM$(\mathbf{F}_1(\gamma), n = 1000, p = 0.1, \rho = 1.5, \eta = 0)$

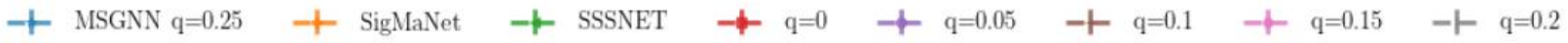
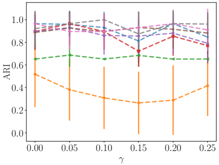
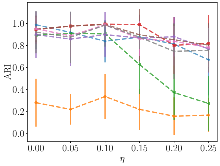


Figure: Node clustering test ARI comparison on synthetic data. Error bars indicate one standard error. Results are averaged over ten runs — five different networks, each with two distinct data splits.

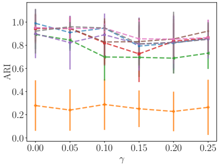
Node Clustering Results



(a) SDSBM($\mathbf{F}_1(\gamma), n = 1000, p = 1, \rho = 1.5, \eta = 0$)



(b) SDSBM($\mathbf{F}_2(\gamma = 0), n = 1000, p = 0.1, \rho = 1.5$)



(c) SDSBM($\mathbf{F}_2(\gamma), n = 1000, p = 0.1, \rho = 1.5, \eta = 0$)

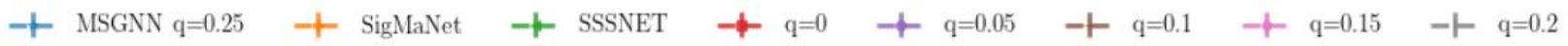


Figure: Node clustering test ARI comparison on synthetic data. Error bars indicate one standard error. Results are averaged over ten runs — five different networks, each with two distinct data splits.

Table: Test accuracy (%) comparison the signed and directed link prediction tasks. The best is marked in **bold red** and the second best is marked in underline blue.

Data Set	Link Task	SGCN	SDGNN	SiGAT	SNEA	SSSNET	SigMaNet	MSGNN
BitCoin-Alpha	SP	64.7±0.9	64.5±1.1	62.9±0.9	64.1±1.3	67.4±1.1	47.8±3.9	**71.3±1.2**
	DP	60.4±1.7	61.5±1.0	61.9±1.9	60.9±1.7	68.1±2.3	49.4±3.1	**72.5±1.5**
	3C	81.4±0.5	79.2±0.9	77.1±0.7	83.2±0.5	78.3±4.7	37.4±16.7	**84.4±0.6**
	4C	51.1±0.8	52.5±1.1	49.3±0.7	52.4±1.8	54.3±2.9	20.6±6.3	**58.5±0.7**
	5C	79.5±0.3	78.2±0.5	76.5±0.3	81.1±0.3	77.9±0.3	34.2±6.5	**81.9±0.9**
BitCoin-OTC	SP	65.6±0.9	65.3±1.2	62.8±1.3	67.7±0.5	70.1±1.2	50.0±2.3	**73.0±1.4**
	DP	63.8±1.2	63.2±1.5	64.0±2.0	65.3±1.2	69.6±1.0	48.4±4.9	**71.8±1.1**
	3C	79.0±0.7	77.3±0.7	73.6±0.7	82.2±0.4	76.9±1.1	26.8±10.9	**83.3±0.7**
	4C	51.5±0.4	55.3±0.8	51.2±1.8	56.9±0.7	57.0±2.0	23.3±7.4	**59.8±0.7**
	5C	77.4±0.7	77.3±0.8	74.1±0.5	80.5±0.5	74.0±1.6	25.9±6.2	**80.9±0.9**
Slashdot	SP	74.7±0.5	74.1±0.7	64.0±1.3	70.6±1.0	86.6±2.2	57.9±5.3	**92.4±0.2**
	DP	74.8±0.9	74.2±1.4	62.8±0.9	71.1±1.1	87.8±1.0	53.0±4.0	**93.1±0.1**
	3C	69.7±0.3	66.3±1.8	49.1±1.2	72.5±0.7	79.3±1.2	42.0±7.9	**86.1±0.3**
	4C	63.2±0.3	64.0±0.7	53.4±0.2	60.5±0.6	72.7±0.6	25.7±8.9	**78.2±0.3**
	5C	64.4±0.3	62.6±2.0	44.4±1.4	66.4±0.5	70.4±0.7	19.3±8.6	**76.8±0.6**
Epinions	SP	62.9±0.5	67.7±0.8	63.6±0.5	66.5±1.0	78.5±2.1	53.3±10.6	**85.4±0.5**
	DP	61.7±0.5	67.9±0.6	63.6±0.8	66.4±1.2	73.9±6.2	49.0±3.2	**86.3±0.3**
	3C	70.3±0.8	73.2±0.8	52.3±1.3	72.8±0.2	72.7±2.0	30.5±8.3	**83.1±0.5**
	4C	66.7±1.2	71.0±0.6	62.3±0.5	69.5±0.7	70.2±5.2	29.9±6.4	**78.7±0.9**
	5C	73.5±0.8	76.6±0.7	52.9±0.7	74.2±0.1	70.3±4.6	22.1±6.1	**80.5±0.5**
FiLL (avg.)	SP	88.4±0.0	82.0±0.3	76.9±0.1	90.0±0.0	88.7±0.3	50.4±1.8	**90.8±0.0**
	DP	88.5±0.1	82.0±0.2	76.9±0.1	90.0±0.0	88.8±0.3	48.0±2.7	**90.9±0.0**
	3C	63.0±0.1	59.3±0.0	55.3±0.1	64.3±0.1	62.2±0.3	33.7±1.3	**66.1±0.1**
	4C	81.7±0.0	78.8±0.1	70.5±0.1	83.2±0.1	80.0±0.3	24.9±0.9	**83.3±0.0**
	5C	63.8±0.0	61.1±0.1	55.5±0.1	**64.8±0.1**	60.4±0.4	19.8±1.1	**64.8±0.1**

- Nonzero q values usually boost performance.
- In general, we see that there are no significant differences in performance between normalizing and not normalizing for the Laplacian; and in the vast majority of cases, these differences are less than one standard deviation.
- In most cases, adding slightly more layers (from 2 to 4) yields a modest increase in performance. However, we again note that these increases in performance are typically quite small and often less than one standard deviation.
- When we further increase the number of layers (up to 10), performance of MSGNN begins to drop slightly. Overall, we do not see much evidence of severe oversmoothing. However, there do not seem to be significant advantages to very deep networks.

Conclusion and Outlook

Our work for signed directed networks:

- a novel magnetic signed Laplacian matrix
- a spectral GNN based on the proposed Laplacian
- novel synthetic network model and new real-world data sets
- leading performance on node clustering & link prediction tasks with faster speed

Future work:

- more properties of the proposed Laplacian
- an extension to temporal/dynamic graphs
- encode nontrivial edge features
- develop objectives which explicitly handle heterogeneous edge densities throughout the graph
- hypergraphs and other complex network structures

Full paper (LoG 2022): `https://proceedings.mlr.press/v198/he22c.html`

Code: `https://github.com/SherylHYX/MSGNN`