

# D-HYPR: Harnessing Neighborhood Modeling and Asymmetry Preservation for Digraph Representation Learning

**Honglu Zhou<sup>1</sup>, Advith Chegu<sup>1</sup>,  
Samuel S. Sohn<sup>1</sup>, Zuohui Fu<sup>1</sup>,  
Gerard de Melo<sup>2</sup>, and Mubbasir Kapadia<sup>1</sup>**



<sup>1</sup> Department of Computer Science, Rutgers University, Piscataway, NJ, USA

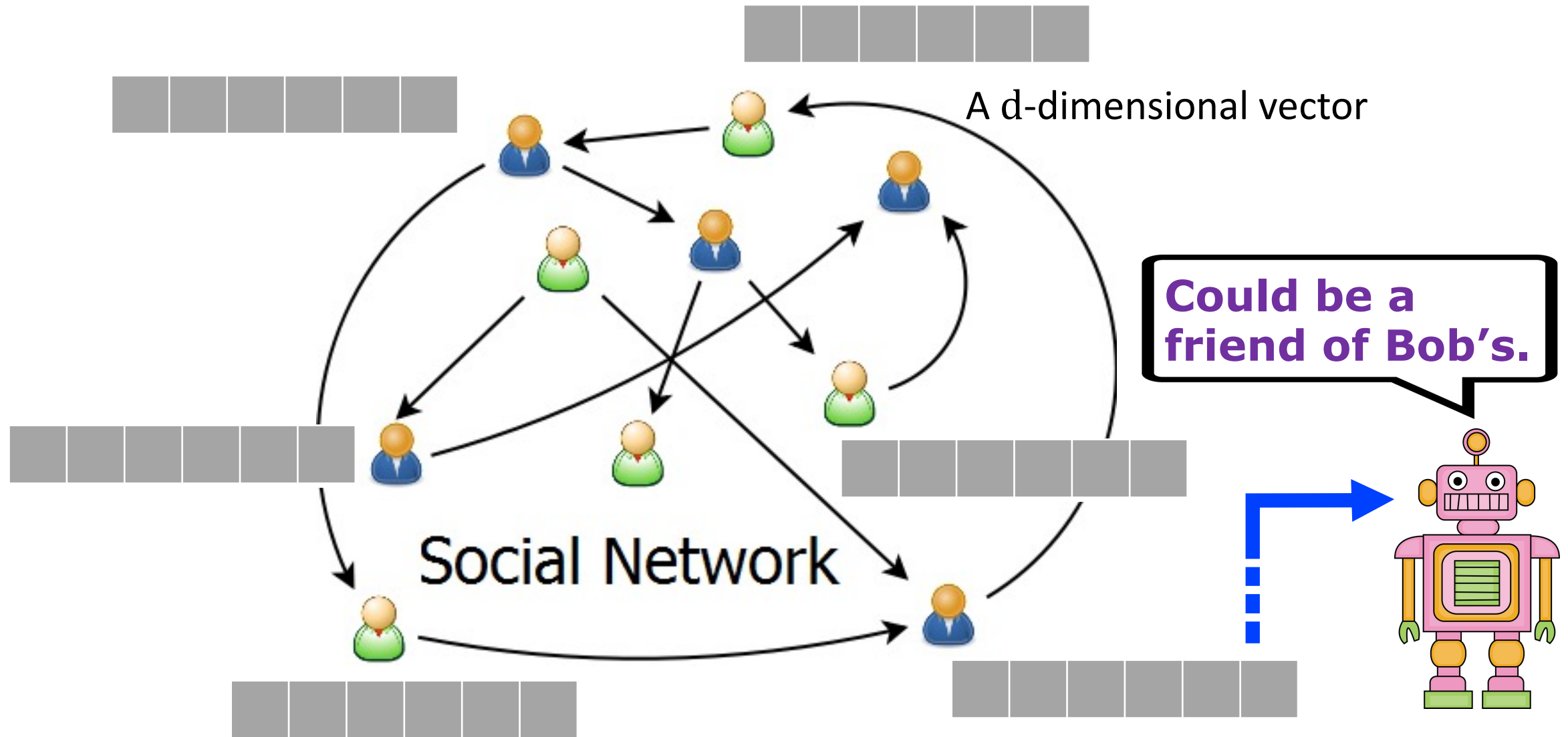
<sup>2</sup> HPI / University of Potsdam, Potsdam, Germany



**RUTGERS**  
THE STATE UNIVERSITY  
OF NEW JERSEY



# Directionality, fundamental characteristics!



**DRL aims to learn representations for directed homogeneous graphs (digraphs).**

# Challenges of DRL

## Challenge 1: Neighborhood Modeling

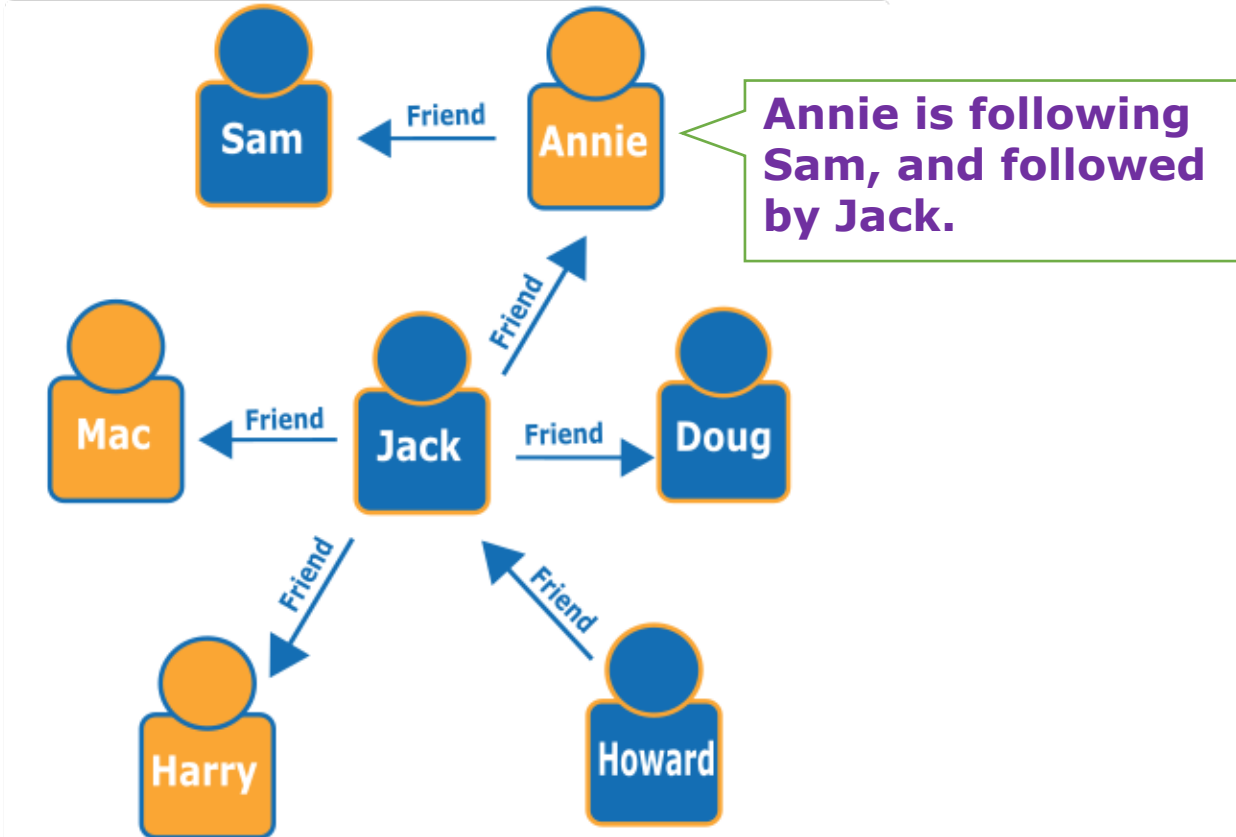
- The neighborhoods of a node may possess unique semantics.

# The neighborhoods of a node in digraphs may possess unique semantics.

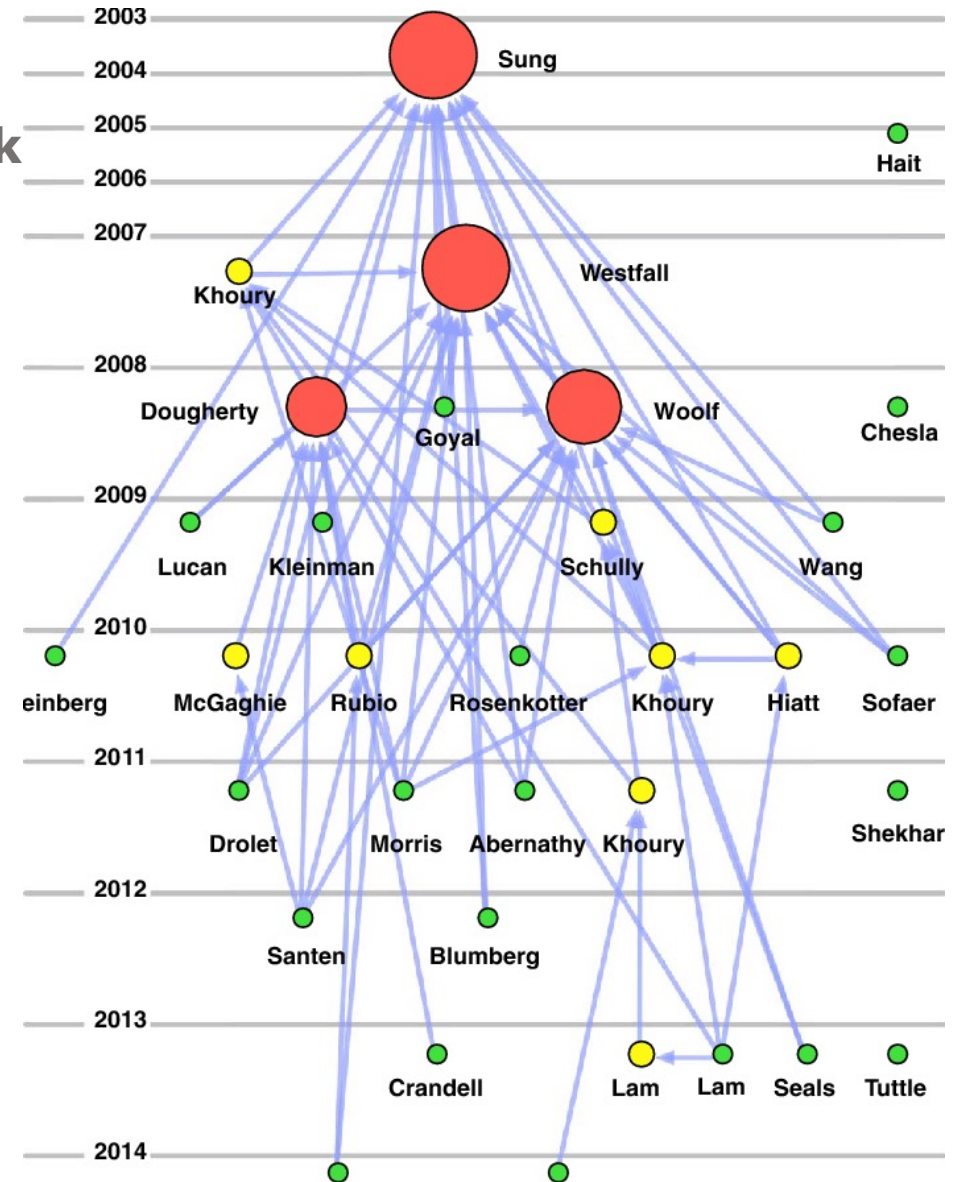


A horizontal timeline with two parallel lines. The top line has a tick mark labeled '2003'. The bottom line has a tick mark labeled '2004'. A red circle is positioned on the bottom line, between the '2003' and '2004' marks, and is labeled 'Sung' to its right.

# Social Network



## Citation Network



Sources:

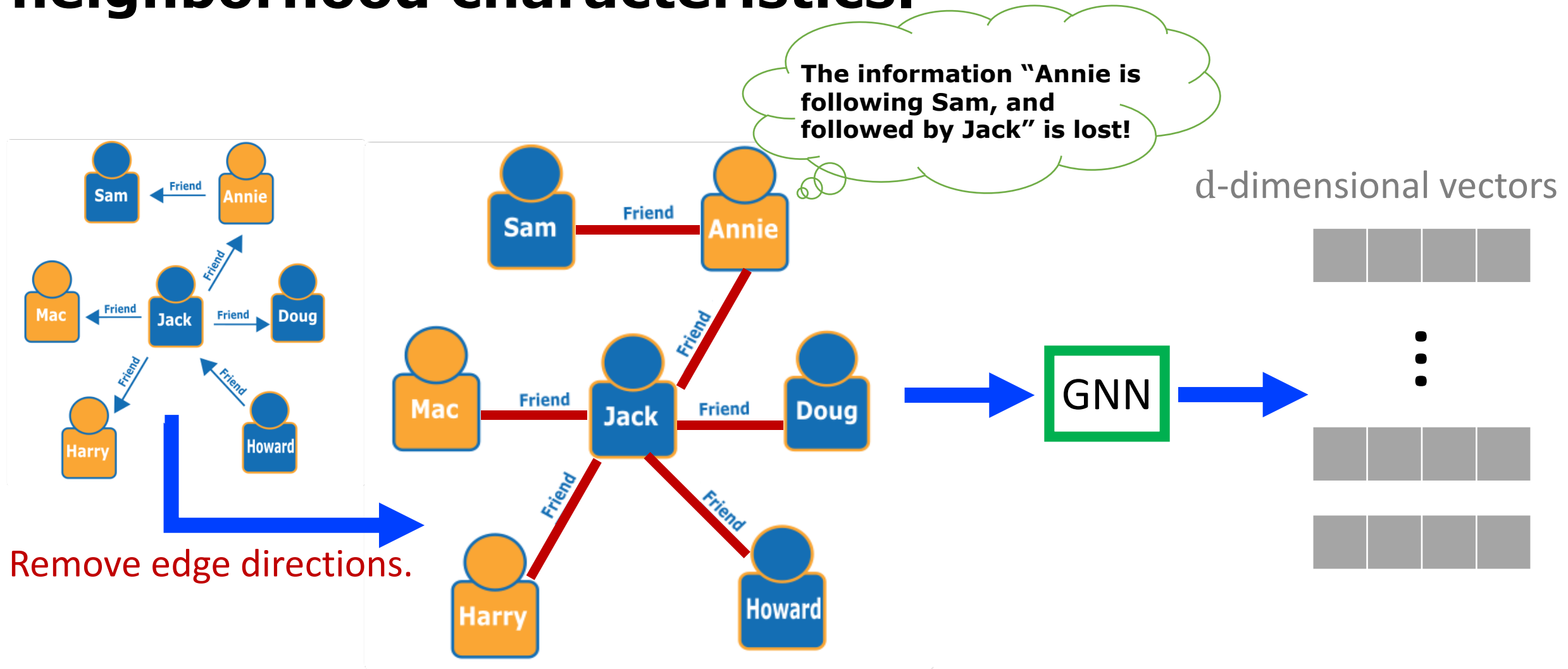
[1] An Example of A Social Network Graph. Friendship@seekpng.com.

[2] Daniel G. Fort, et al. "Mapping the Evolving Definitions of Translational Research".

Journal of Clinical and Translational Science 1, No. 1 (2017): 60-66.

## Challenge 1: Neighborhood Modeling

# Existing popular GNNs ignore the unique node neighborhood characteristics.



**Challenge 1: Neighborhood Modeling**

# Challenges of DRL

## Challenge 1: Neighborhood Modeling

- The neighborhoods of a node may possess unique semantics.
- Existing popular GNN techniques (e.g., GCN<sup>[1]</sup>, VGAE<sup>[2]</sup>, GAT<sup>[3]</sup>, HGCGN<sup>[4]</sup>, GIL<sup>[5]</sup>) transform digraphs to undirected graphs to enable running experiments, or only consider the direct out-neighbors in graph convolution.

## Challenge 2: Asymmetry Preservation

- Shall capture the asymmetric node connection probabilities for node pair  $(i, j)$  and  $(j, i)$ .

[1] Thomas N Kipf and Max Welling. "Semi-Supervised Classification with Graph Convolutional Networks." ICLR 2017.

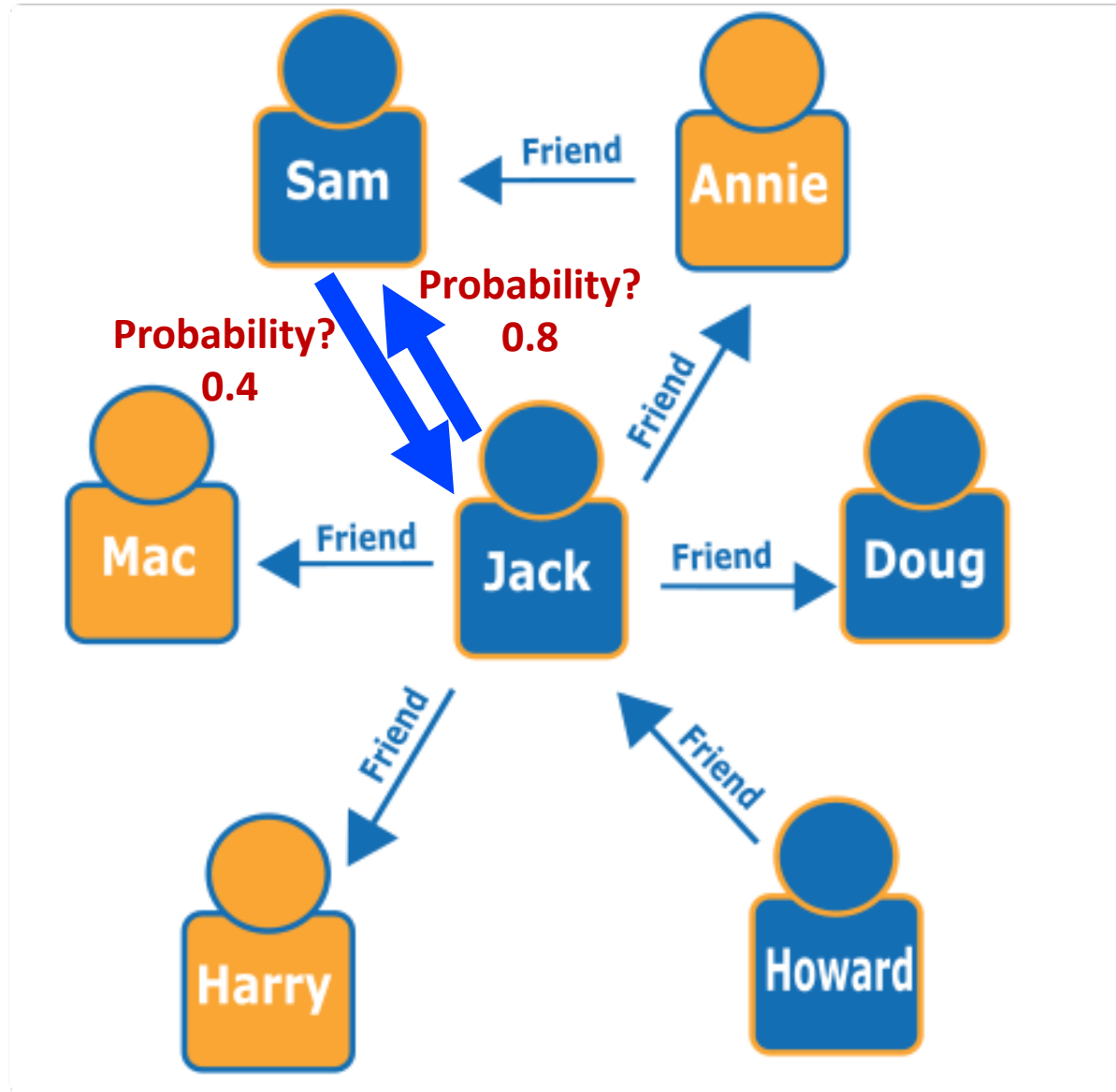
[2] Thomas N Kipf and Max Welling. "Variational Graph Auto-Encoders". arXiv preprint arXiv:1611.07308 (2016).

[3] Petar Veličković, et al. "Graph Attention Networks". ICLR 2018.

[4] Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec. "Hyperbolic Graph Convolutional Neural Networks". NeurIPS 2019.

[5] Shichao Zhu et al. "Graph Geometry Interaction Learning". NeurIPS 2020.

# Node connection probabilities are unequal in digraphs.



**Challenge 2: Asymmetry Preservation**

# What are the prior practices?

## Challenge 1: Neighborhood Modeling

- Spectral-based DRL GNNs<sup>[1-4]</sup> have been proposed.

## Challenge 2: Asymmetry Preservation

- View directions of edges as a kind of edge feature<sup>[5]</sup>.
- Parametrize the node pair likelihood function by a neural network<sup>[6-7]</sup>.

**Moreover, prior DRL techniques are often constrained to directed acyclic graphs (DAGs), are transductive, or have poor generalizability across tasks - some studies provide experimental evidence for a single task.**

[1] Yi Ma, et al. "Spectral-based Graph Convolutional Network for Directed Graphs". arXiv preprint arXiv:1907.08990 (2019).

[2] Zekun Tong, et al. "Digraph Inception Convolutional Networks". NeurIPS 2020.

[3] Zekun Tong, et al. "Directed Graph Convolutional Network". arXiv preprint arXiv:2004.13970 (2020).

[4] Xitong Zhang, et al. "MagNet: A Magnetic Neural Network for Directed Graphs". NeurIPS 2021.

[5] Liyu Gong and Qiang Cheng. "Exploiting Edge Features for Graph Neural Networks". CVPR 2019.

[6] Peter W Battaglia, et al. "Relational Inductive Biases, Deep Learning, and Graph Networks". arXiv preprint arXiv:1806.01261 (2018).

[7] Lei Shi, et al. "Skeleton-based Action Recognition with Directed Graph Neural Networks". CVPR 2019.

**Prior works fail to address both challenges.**

# Our solution: D-HYPR

## Challenge 1: Neighborhood Modeling

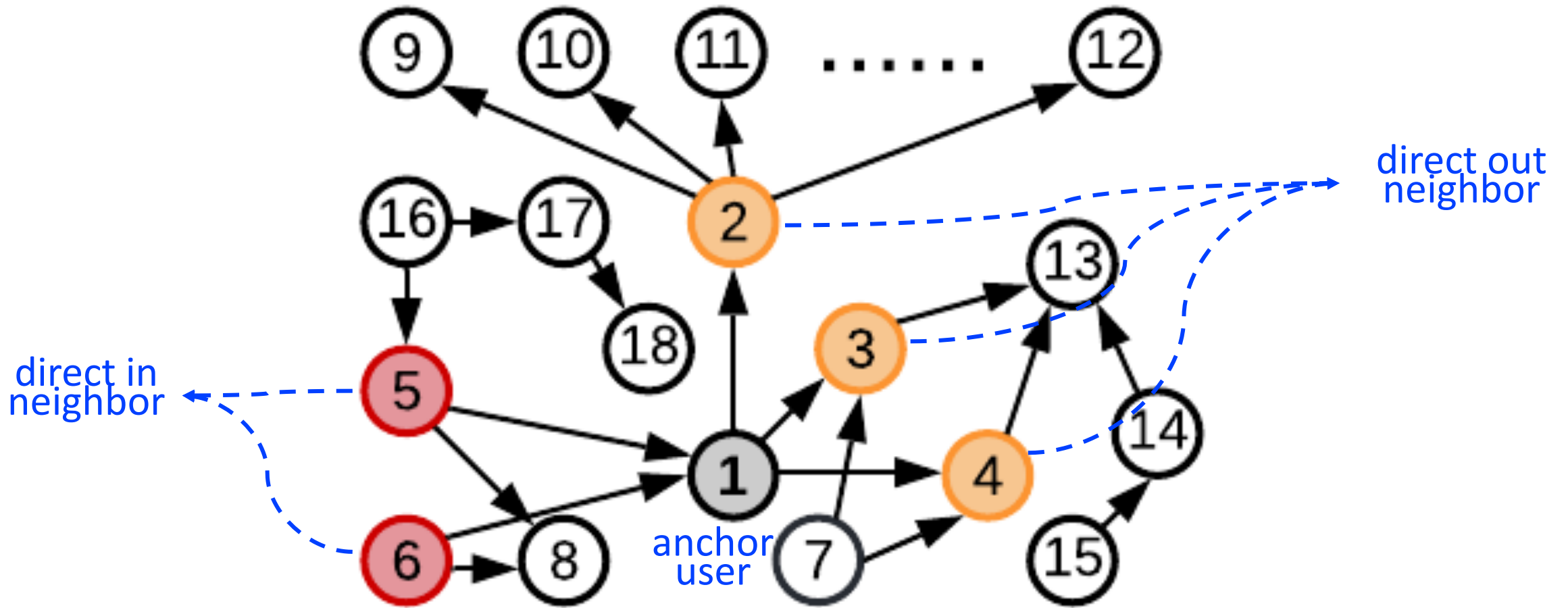
- D-HYPR utilizes **hyperbolic collaborative learning** from **multi-ordered and partitioned neighborhoods**.

## Challenge 2: Asymmetry Preservation

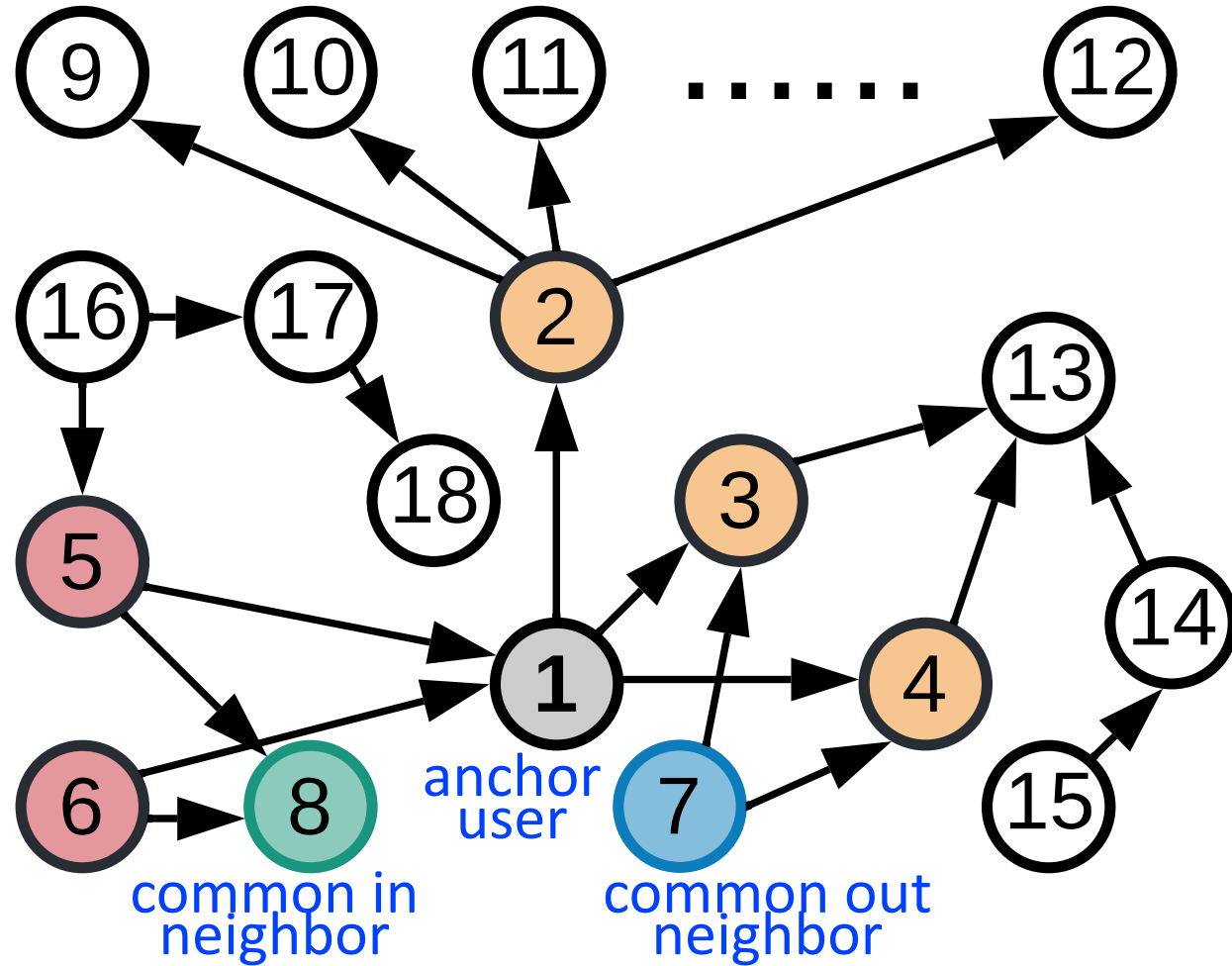
- D-HYPR takes advantage of **self-supervised learning**, using **asymmetry-preserving regularizers** supported by well-established socio-psychological theories.

We propose **Digraph HYPERbolic Networks (D-HYPR)** to address the problem.

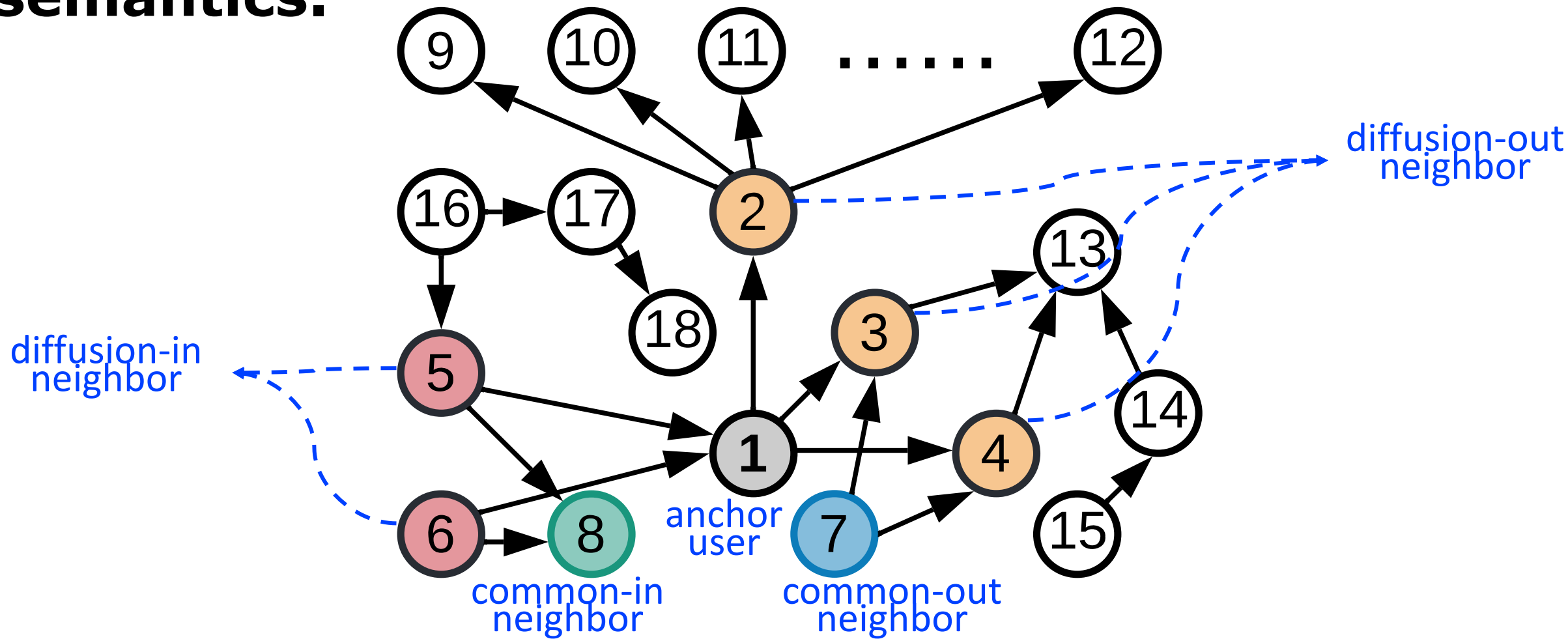
**The real-world inductive bias: neighbors of a node can be partitioned into groups based on the semantics.**



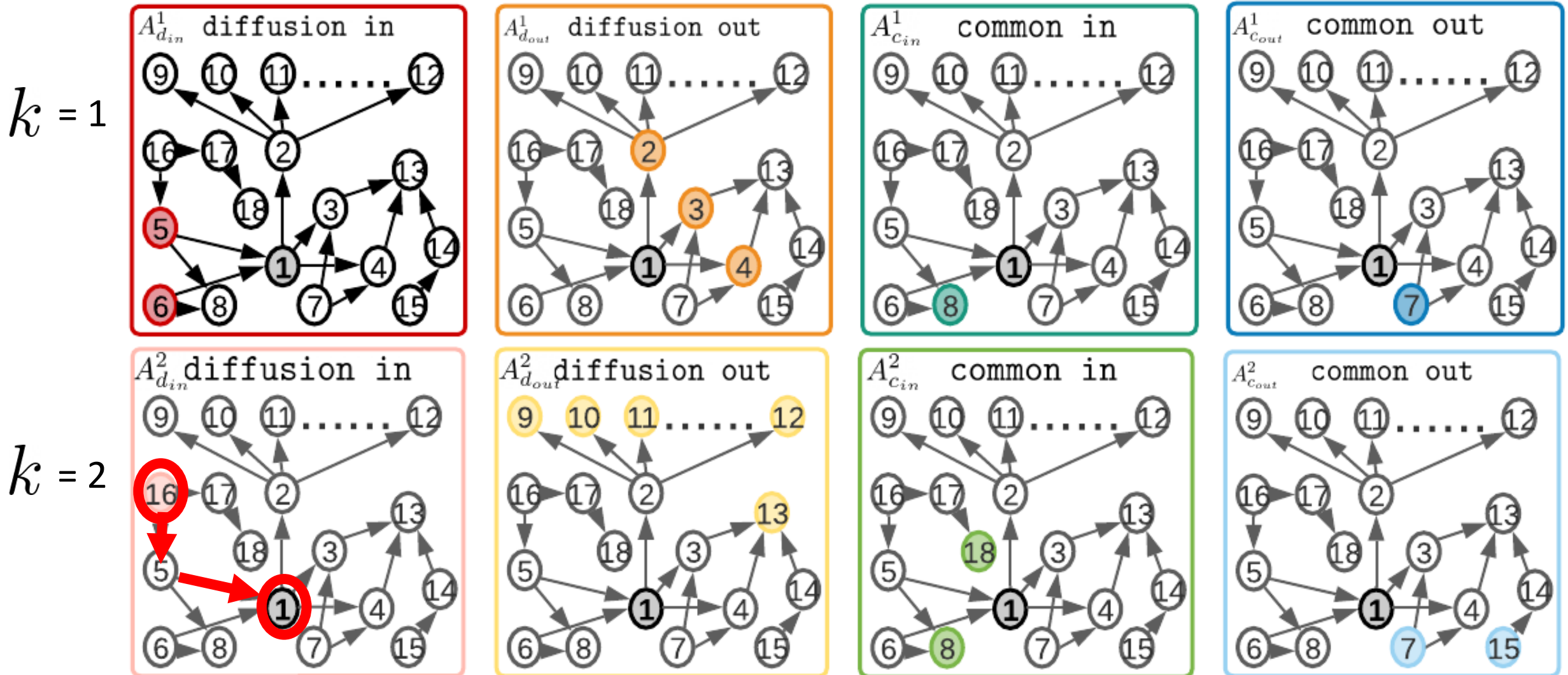
**The real-world inductive bias: neighbors of a node can be partitioned into groups based on the semantics.**



**The real-world inductive bias: neighbors of a node can be partitioned into groups based on the semantics.**

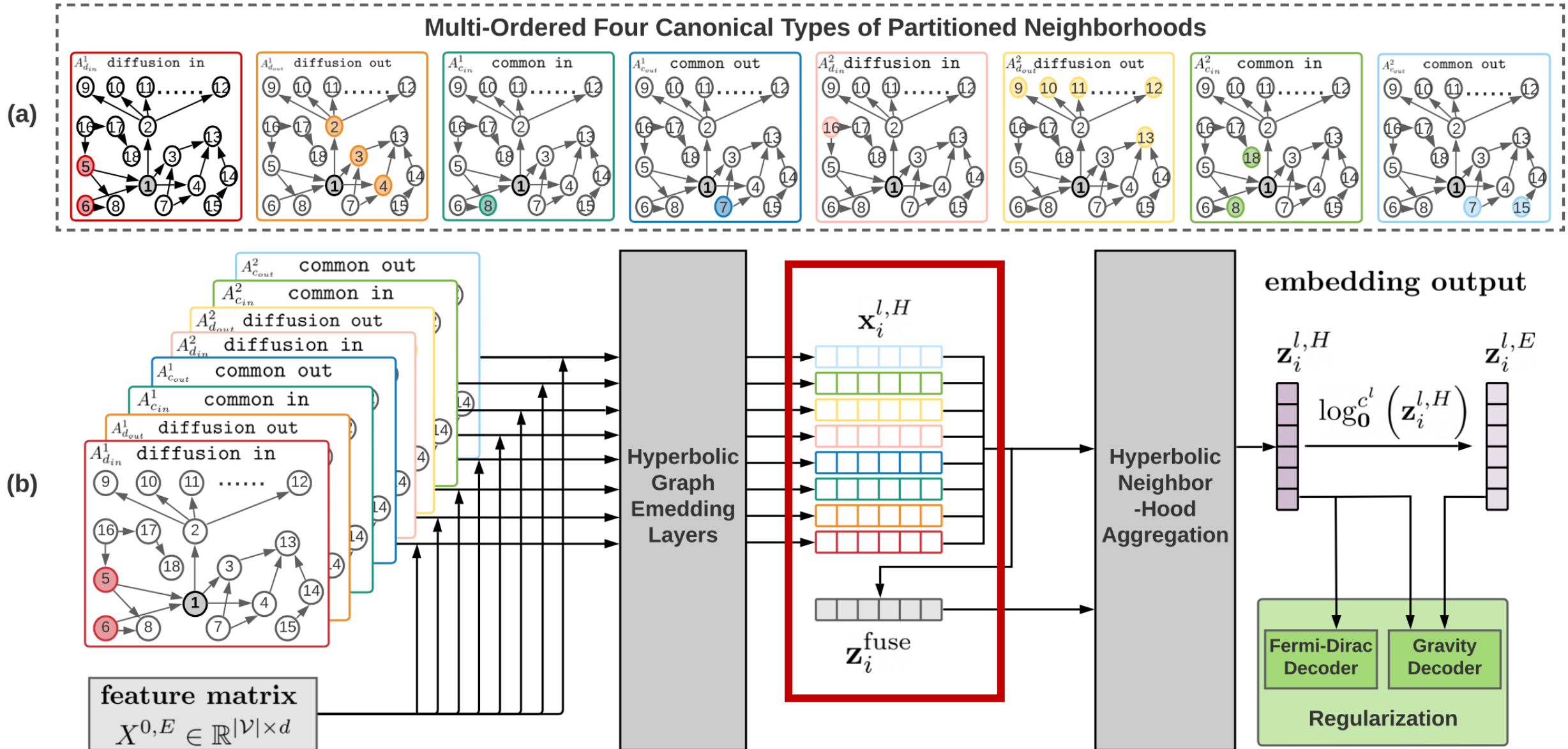


# Multi-Ordered 4 Canonical Types of Partitioned Neighborhoods



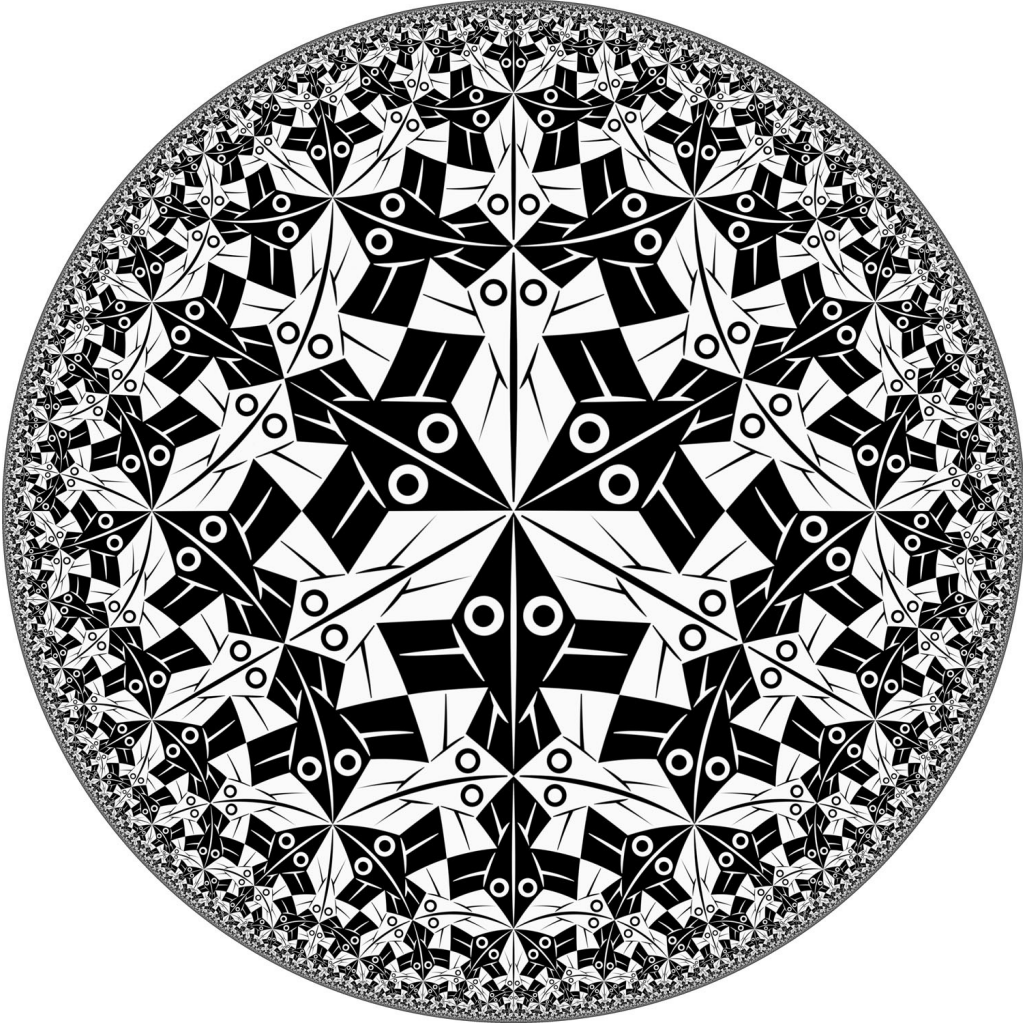
**D-HYPR: Neighborhood Modeling with Partitioned and Larger Receptive Fields**

# Our solution: D-HYPR



We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.

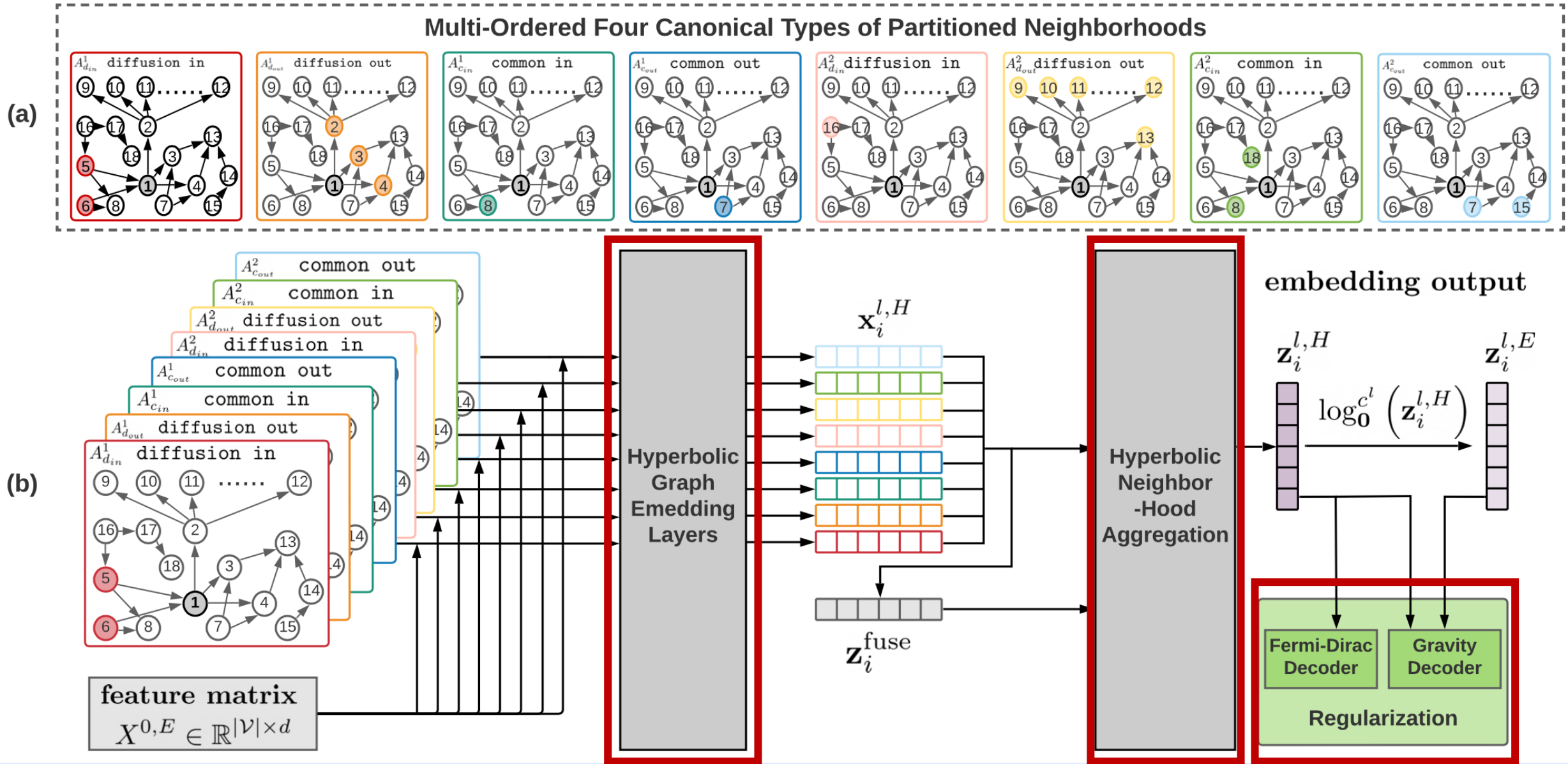
# Hyperbolic embeddings can incur smaller data distortion for real-world digraphs.



“Circle Limit 1” by M.C. Escher illustrates the Poincaré disc model of hyperbolic space. Each tile is of constant area in hyperbolic space, but vanishes in Euclidean space at the boundary <sup>[1]</sup>.

[1] Benjamin Paul Chamberlain, James Clough, and Marc Peter Deisenroth. Neural Embeddings of Graphs in Hyperbolic Space. arXiv preprint arXiv:1705.10359 (2017).

# Our solution: D-HYPR



**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Homophily and preferential attachment are two driving forces of link formation

**Homophily**



**Node Similarity**

↓ modelled by

**Fermi-Dirac  
Decoder [1]**

$$p(i, j)_f = \frac{1}{e^{\left( d_{\mathbb{D}}^{cl} \left( z_i^{l,H}, z_j^{l,H} \right)^2 - r \right) / t} + 1}$$

**Preferential Attachment**



**Node Connectivity**

↓ modelled by

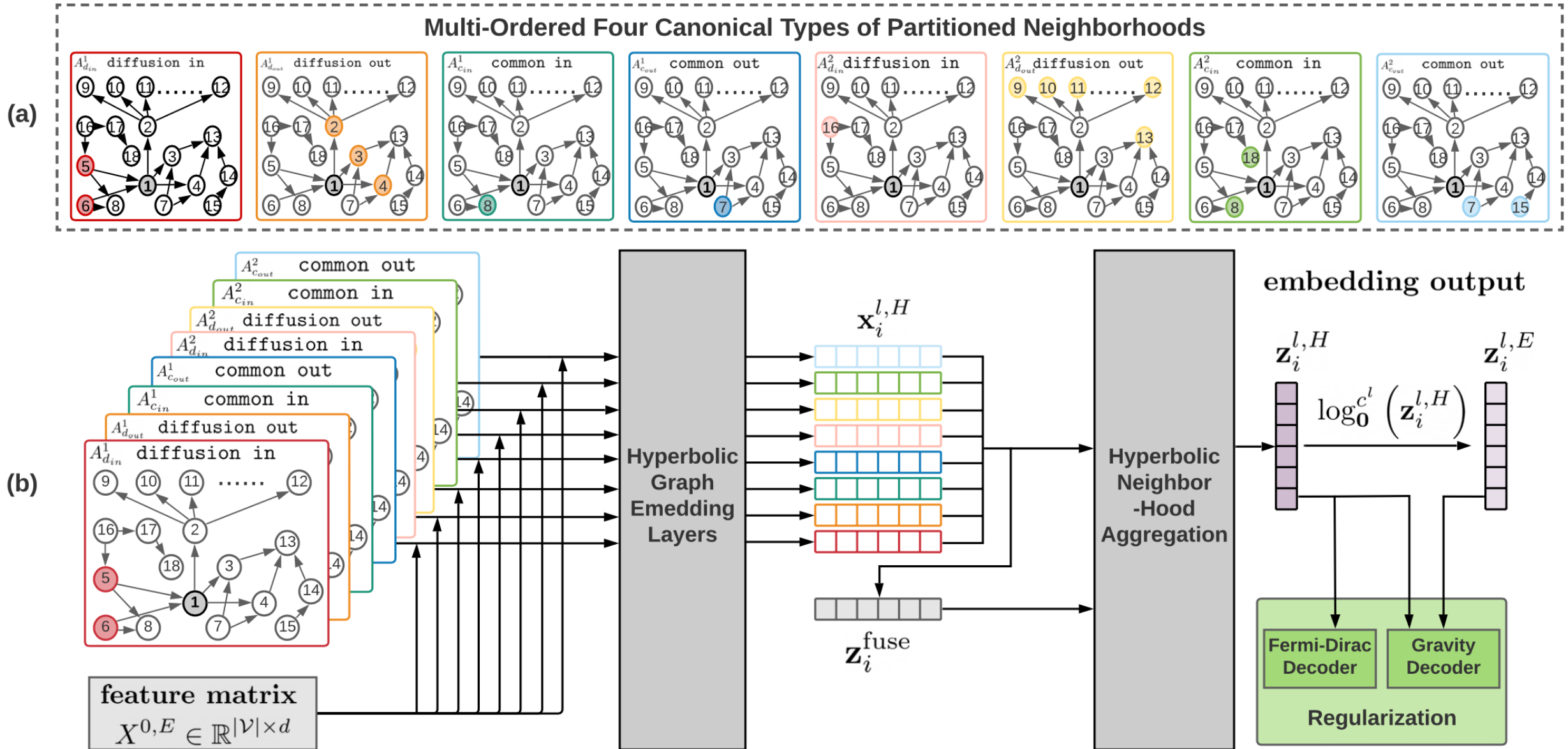
**Gravity  
Decoder [2]**

$$p(i, j)_g = \gamma \left( m_j - \lambda \log \left( d_{\mathbb{D}}^{cl} \left( z_i^{l,H}, z_j^{l,H} \right)^2 \right) \right)$$

[1] Dmitri Krioukov, et al. Hyperbolic Geometry of Complex Networks. Physical Review E 82, 3 (2010), 036106.

[2] Guillaume Salha, et al. Gravity-Inspired Graph Autoencoders for Directed Link Prediction. CIKM 2019.

# Our solution: D-HYPR



We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.

# Results: Link Presence Prediction

Best, Second best, *Third best*

Model (4/8-Dim)	Air				Cora			
	4-Dim		8-Dim		4-Dim		8-Dim	
	AUC	AP	AUC	AP	AUC	AP	AUC	AP
GCN [22]	67.88 (61.73)	67.88 (60.51)	69.21 (64.05)	69.68 (63.48)	65.92 (61.00)	65.92 (59.97)	70.89 (65.67)	71.26 (65.28)
VGAE [23]	69.77 (62.86)	70.73 (62.55)	73.49 (66.87)	74.04 (66.95)	63.86 (56.90)	63.86 (55.39)	66.60 (60.33)	66.60 (58.75)
GAT [51]	69.02 (63.48)	69.02 (62.86)	71.31 (67.03)	71.31 (67.01)	68.18 (64.73)	68.18 (64.31)	72.70 (68.70)	73.93 (69.08)
Gravity GCN <sup>†</sup> [40]	65.20 (59.41)	67.73 (60.98)	74.00 (68.91)	75.43 (69.14)	70.37 (65.80)	70.37 (64.65)	75.29 (71.85)	77.17 (72.50)
Gravity VGAE <sup>†</sup> [40]	62.24 (55.48)	62.24 (54.97)	68.00 (60.23)	68.00 (59.57)	66.74 (61.79)	66.74 (60.61)	71.04 (65.45)	71.04 (64.15)
DGCN <sup>†</sup> [49]	74.36 (65.75)	71.42 (63.27)	77.23 (70.60)	75.86 (70.27)	75.33 ( <i>71.88</i> )	71.95 ( <i>68.58</i> )	79.01 (75.30)	79.01 (74.28)
DiGCN <sup>†</sup> [48]	72.59 (64.37)	70.01 (61.66)	74.65 (69.27)	75.40 (68.29)	70.61 (65.81)	67.11 (61.57)	74.63 (70.65)	74.88 (69.86)
MagNet <sup>†</sup> [58]	72.26 (58.44)	71.10 (57.92)	76.64 (64.26)	78.62 (64.69)	<i>77.45</i> (55.93)	<i>79.32</i> (56.84)	77.46 (66.82)	76.59 (63.96)
HAT <sup>§</sup> [59]	<i>76.11</i> ( <u>71.24</u> )	<i>73.72</i> ( <u>69.35</u> )	<i>80.52</i> ( <i>75.13</i> )	<i>79.73</i> ( <i>74.05</i> )	76.25 ( <u>72.84</u> )	74.38 ( <u>70.27</u> )	<i>82.58</i> ( <i>77.82</i> )	<i>82.05</i> ( <i>77.39</i> )
HGCN <sup>§</sup> [8]	<u>80.90</u> ( <i>66.63</i> )	<u>80.90</u> ( <i>65.95</i> )	<u>84.67</u> ( <u>77.65</u> )	<u>85.97</u> ( <u>78.14</u> )	<u>80.02</u> (67.37)	<u>82.16</u> (66.66)	<u>85.05</u> ( <u>83.07</u> )	<u>88.04</u> ( <u>84.63</u> )
D-HYPR (ours) <sup>†§</sup>	<b>85.79</b> (* <b>81.69</b> )	<b>85.92</b> (* <b>81.93</b> )	<b>88.46</b> (* <b>84.26</b> )	<b>88.46</b> (* <b>84.82</b> )	<b>86.08</b> (* <b>83.99</b> )	<b>88.74</b> (* <b>85.33</b> )	<b>88.88</b> (* <b>86.31</b> )	<b>91.13</b> (* <b>87.76</b> )
Relative Gains (%)	6.04 (14.67)	6.21 (18.14)	4.48 (8.51)	2.90 (8.55)	7.57 (15.31)	8.01 ( <u>21.43</u> )	4.5 (3.9)	3.51 (3.7)

Metrics:

- AUC (Area under the ROC Curve)
- AP (Average Precision)

<sup>†</sup> : DRL method

<sup>§</sup> : hyperbolic space used

\* : statistically superior

**Best score (Average score)**

**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Results: Link Presence Prediction

Best, Second best, *Third best*

Model (32-Dim)										
	Air		Cora		Blog		Survey		DBLP	
	AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP
MLP	81.29 (76.52)	83.53 (78.18)	84.47 (81.67)	87.70 (83.69)	93.31 (92.48)	93.31 (92.45)	91.21 (89.98)	92.46 (90.75)	51.22 (49.98)	51.22 (49.99)
NERD <sup>†</sup> [19]	60.62 (56.39)	67.37 (60.19)	65.62 (62.02)	71.68 (65.66)	95.03 (94.00)	95.03 (93.47)	77.12 (69.30)	79.60 (70.80)	95.78 (95.37)	95.93 (95.41)
ATP <sup>†</sup> [44]	68.99 (66.40)	68.99 (64.99)	88.47 ( <i>86.44</i> )	88.47 (86.04)	85.05 (83.46)	85.05 (79.30)	73.53 (71.47)	73.53 (70.64)	60.43 (59.21)	60.43 (57.37)
APP <sup>†</sup> [61]	85.08 (82.72)	86.35 (84.58)	86.65 (85.50)	89.80 (87.22)	92.33 (91.65)	92.33 (90.55)	91.16 (90.34)	92.77 (91.14)	95.58 (95.33)	95.73 (95.41)
GCN [22]	76.71 (72.27)	80.95 (75.13)	80.77 (78.73)	85.67 (81.21)	91.87 (90.18)	92.16 (90.54)	89.29 (87.98)	91.78 (89.42)	92.98 (92.34)	94.37 (93.15)
VGAE [23]	77.79 (73.75)	82.73 (76.75)	80.80 (79.24)	85.47 (81.57)	92.25 (91.39)	92.80 (91.85)	90.07 (88.78)	92.39 (90.14)	93.36 (92.64)	94.85 (93.45)
GAT [51]	84.21 (80.24)	84.79 (81.46)	85.40 (82.58)	88.53 (84.60)	92.69 (89.95)	92.69 (89.83)	92.01 (91.05)	93.09 (91.65)	95.94 (95.62)	96.28 (95.80)
Gravity GCN <sup>†</sup> [40]	85.16 (82.22)	86.86 (83.50)	85.62 (83.87)	88.73 (85.62)	95.11 (94.46)	95.11 (94.31)	91.63 (90.86)	93.11 (91.76)	96.89 (96.78)	<i>97.46 (97.34)</i>
Gravity VGAE <sup>†</sup> [40]	83.98 (80.06)	85.67 (81.61)	87.17 (84.46)	89.51 (86.22)	<u>96.15</u> ( <u>95.59</u> )	<u>96.15</u> ( <u>95.42</u> )	91.64 (90.96)	93.23 (91.82)	95.98 (95.57)	96.24 (95.81)
DGCN <sup>†</sup> [49]	77.83 (73.68)	80.79 (75.64)	83.57 (81.34)	85.48 (83.00)	87.74 (86.74)	88.13 (86.75)	90.47 (89.49)	91.27 (89.94)	92.26 (91.83)	90.16 (89.52)
DiGCN <sup>†</sup> [48]	75.35 (71.27)	77.64 (73.97)	81.80 (78.90)	83.03 (79.92)	91.98 (90.50)	89.34 (87.36)	89.85 (88.17)	89.80 (88.08)	89.99 (89.72)	89.93 (89.60)
MagNet <sup>†</sup> [58]	79.32 (75.58)	80.66 (76.34)	82.77 (71.90)	81.63 (69.84)	91.83 (90.81)	90.46 (89.29)	86.65 (84.81)	87.76 (85.71)	81.89 (80.57)	81.68 (81.50)
HNN <sup>§</sup> [13]	<u>88.42</u> ( <i>85.79</i> )	<u>88.95</u> ( <i>86.40</i> )	<i>88.75</i> (86.33)	<i>90.81</i> ( <i>87.81</i> )	<i>95.80</i> ( <i>95.39</i> )	<i>95.80</i> ( <i>95.16</i> )	<i>92.07</i> ( <i>91.39</i> )	<u>93.40</u> ( <i>92.04</i> )	<i>97.43</i> ( <i>97.14</i> )	97.43 (97.13)
HGCN <sup>§</sup> [8]	<i>88.26</i> ( <u>86.12</u> )	<i>88.88</i> ( <u>86.64</u> )	<u>89.24</u> ( <u>87.68</u> )	<u>91.54</u> ( <u>88.97</u> )	95.64 (95.23)	95.64 (95.00)	<u>92.15</u> ( <u>91.50</u> )	<i>93.38</i> ( <u>92.08</u> )	<u>97.54</u> ( <u>97.33</u> )	<u>97.62</u> ( <u>97.37</u> )
D-HYPR (ours) <sup>†§</sup>	<b>89.07</b> ( <b>86.33</b> )	<b>89.21</b> (* <b>86.86</b> )	<b>89.50</b> (* <b>88.22</b> )	<b>91.62</b> (* <b>89.47</b> )	<b>96.19</b> ( <b>95.62</b> )	<b>96.18</b> (* <b>95.48</b> )	<b>92.56</b> (* <b>91.96</b> )	<b>93.63</b> (* <b>92.46</b> )	<b>97.66</b> (* <b>97.38</b> )	<b>97.75</b> (* <b>97.44</b> )
Relative Gains (%)	0.74 (0.24)	0.29 (0.25)	0.29 (0.62)	0.09 (0.56)	0.04 (0.03)	0.03 (0.06)	0.44 (0.50)	0.25 (0.41)	0.12 (0.05)	0.13 (0.07)

Metrics:

- AUC (Area under the ROC Curve)
- AP (Average Precision)

<sup>†</sup> : DRL method

<sup>§</sup> : hyperbolic space used

\* : statistically superior

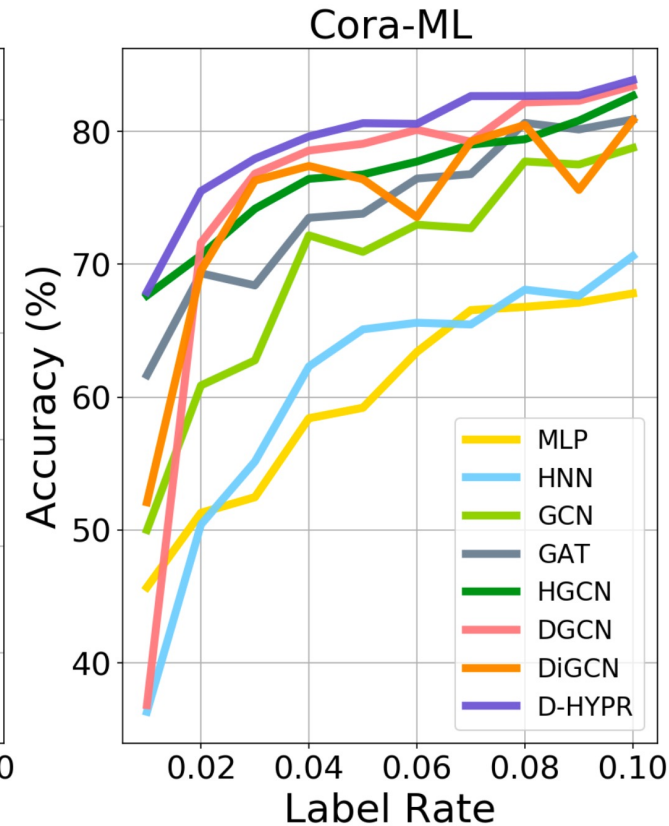
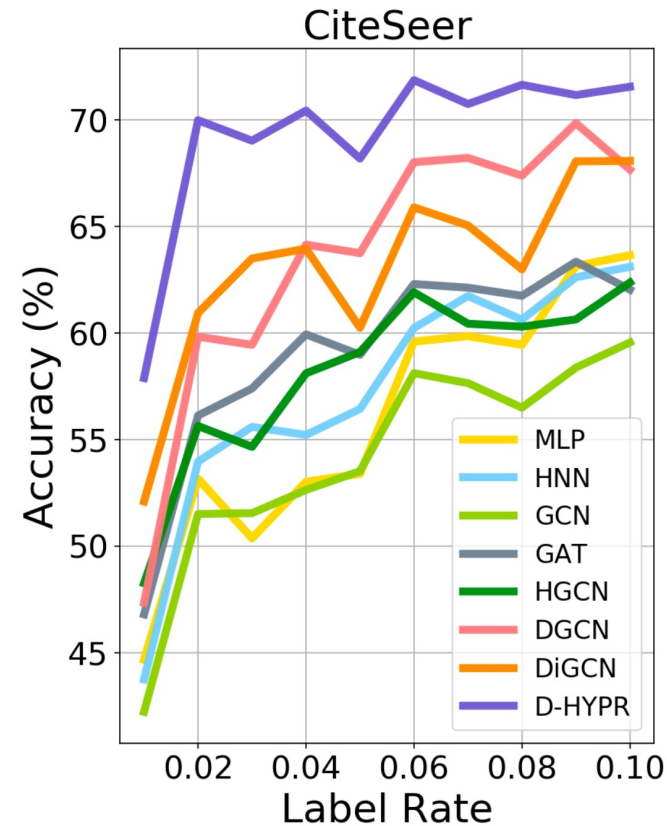
**Best score (Average score)**

**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Results: Node Classification

**Best**, Second best, *Third best*

	Model	CiteSeer	Cora-ML
4-Dim	MLP	37.68 $\pm$ 3.0	51.19 $\pm$ 6.3
	GCN [22]	32.82 $\pm$ 7.9	60.56 $\pm$ 9.8
	GAT [51]	<i>51.97<math>\pm</math>4.2</i>	<i>68.38<math>\pm</math>3.4</i>
	DGCN [49]	38.67 $\pm$ 10.0	53.44 $\pm$ 11.1
	DiGCN [48]	<u>53.43<math>\pm</math>10.3</u>	<u>71.35<math>\pm</math>2.3</u>
	HNN [13]	47.44 $\pm$ 2.9	52.76 $\pm$ 4.9
	HGCN [8]	42.24 $\pm$ 3.6	52.17 $\pm$ 5.9
	D-HYPR (ours)	<b>*65.72<math>\pm</math>2.9</b>	<b>*74.63<math>\pm</math>1.2</b>
	Relative Gains (%)	23.00	4.60
8-Dim	MLP	51.70 $\pm$ 2.6	60.48 $\pm$ 1.8
	GCN [22]	36.26 $\pm$ 6.5	67.62 $\pm$ 10.8
	GAT [51]	50.81 $\pm$ 3.9	74.87 $\pm$ 1.8
	DGCN [49]	<i>57.27<math>\pm</math>2.4</i>	<i>77.16<math>\pm</math>4.4</i>
	DiGCN [48]	<u>60.37<math>\pm</math>2.6</u>	<u>78.38<math>\pm</math>1.2</u>
	HNN [13]	50.73 $\pm$ 3.1	61.54 $\pm$ 2.1
	HGCN [8]	52.57 $\pm$ 2.3	73.44 $\pm$ 2.3
	D-HYPR (ours)	<b>*67.96<math>\pm</math>1.6</b>	<b>*81.55<math>\pm</math>1.6</b>
	Relative Gains (%)	12.57	4.04
...			
256-Dim	MLP	57.26 $\pm$ 2.2	64.86 $\pm$ 3.1
	GCN [22]	55.82 $\pm$ 3.2	75.20 $\pm$ 1.9
	GAT [51]	57.66 $\pm$ 2.4	74.19 $\pm$ 1.5
	DGCN [49]	<u>65.90<math>\pm</math>1.5</u>	<u>81.29<math>\pm</math> 1.4</u>
	DiGCN [48]	46.36 $\pm$ 13.75	<i>79.46<math>\pm</math>1.2</i>
	HNN [13]	54.64 $\pm$ 2.4	66.09 $\pm$ 2.0
	HGCN [8]	<i>58.23<math>\pm</math> 2.3</i>	76.91 $\pm$ 1.7
	D-HYPR (ours)	<b>*71.10<math>\pm</math>1.2</b>	<b>*81.80<math>\pm</math> 1.4</b>
	Relative Gains (%)	7.89	0.63



\* : statistically superior

**Average score  $\pm$  Standard Deviation**

**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Results: Link Property Prediction

**Best**, Second best, *Third best*

	Model	Node Classification	Link Sign Prediction
4-Dim	GCN [22]	$17.01 \pm 0.1$	$78.96 \pm 0.4$
	GAT [51]	$40.75 \pm 10.7$	$79.38 \pm 0.2$
	HGCN [8]	$36.07 \pm 5.3$	$78.72 \pm 0.0$
	D-HYPR (ours)	<b>*<math>71.27 \pm 0.79</math></b>	<b>*<math>79.83 \pm 0.0</math></b>
	Relative Gains (%)	74.90	0.57
8-Dim	GCN [22]	$39.26 \pm 9.5$	$78.76 \pm 0.1$
	GAT [51]	$46.78 \pm 10.5$	$79.41 \pm 0.2$
	HGCN [8]	$58.40 \pm 10.9$	$79.23 \pm 0.2$
	D-HYPR (ours)	<b>*<math>70.53 \pm 1.6</math></b>	<b>*<math>79.47 \pm 0.3</math></b>
	Relative Gains (%)	20.77	0.08
32-Dim	GCN [22]	$37.77 \pm 6.7$	$79.39 \pm 0.1$
	GAT [51]	$46.12 \pm 8.5$	$79.66 \pm 0.1$
	HGCN [8]	$52.63 \pm 5.8$	$79.21 \pm 0.2$
	D-HYPR (ours)	<b>*<math>71.65 \pm 1.0</math></b>	<b>*<math>79.73 \pm 0.2</math></b>
	Relative Gains (%)	36.14	0.09

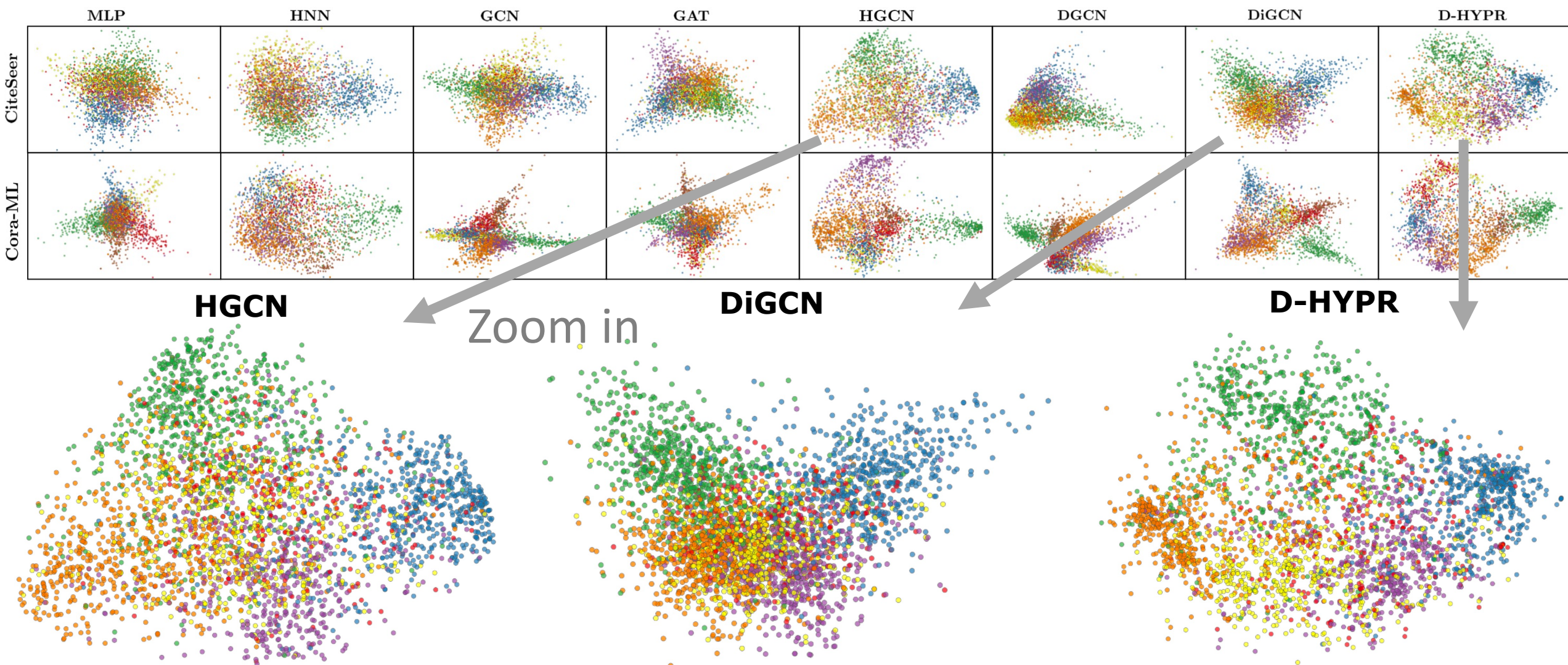
\* : statistically superior

Average score  $\pm$  Standard Deviation

**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Results: Embedding Visualization

Each dot represents a node, and colors reflect the ground truth class labels of nodes.



We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.

# Parameter Sensitivity

Best, Second best, *Third best*

## Hyper-parameters:

- (1)  $\lambda \rightarrow$  a smaller  $\lambda$  emphasizes the asymmetric node connectivity, used by the Gravity decoder.
- (2)  $K \rightarrow$  the maximal order of the  $k$ -order proximity matrix, larger means a wider receptive field and more scale information.

Parameter sensitivity analysis in terms of  $K$

$K$	1	2	3
CiteSeer	69.23 ± 1.5	<b>70.66 ± 1.2</b>	69.76 ± 1.5
Cora-ML	82.16 ± 1.3	82.16 ± 1.3	<b>82.19 ± 1.3</b>

Compare D-HYPR with SOTA

	Model	CiteSeer	Cora-ML
32-Dim	MLP	53.18 ± 1.6	61.63 ± 1.8
	GCN [22]	53.20 ± 3.1	69.51 ± 8.5
	GAT [51]	63.03 ± 0.6	71.91 ± 0.9
	DGCN [49]	<i>64.17±2.4</i>	<u>81.29±1.6</u>
	DiGCN [48]	<u>65.83±1.8</u>	<i>78.08±1.9</i>
	HNN [13]	56.10 ± 2.2	62.49 ± 2.6
	HGCN [8]	59.02 ± 2.3	76.48 ± 1.5
	D-HYPR (ours)	<b>*70.66 ± 1.2</b>	<b>*82.19 ± 1.3</b>
	Relative Gains (%)	7.34	1.11

(32-Dim, the Node Classification task).

Parameter sensitivity analysis in terms of  $\lambda$

$\lambda$	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00	4.25	4.50	4.75	5.00	10.0
CiteSeer	69.74 ±1.6	<b>70.66 ±1.2</b>	70.46 ±1.3	70.44 ±1.4	70.30 ±1.1	70.34 ±1.3	69.99 ±1.4	69.79 ±1.6	69.24 ±1.6	69.61 ±1.2	68.13 ±1.4	68.05 ±1.3	68.12 ±1.9	67.64 ±1.8	67.85 ±1.8	67.67 ±1.9	67.69 ±1.9	67.34 ±2.3	67.18 ±2.1	67.12 ±1.7	66.85 ±1.5
Cora-ML	81.29 ±1.3	81.18 ±1.2	81.59 ±1.2	81.68 ±1.4	81.83 ±1.1	81.97 ±1.0	<b>82.16 ±1.3</b>	81.65 ±1.2	81.10 ±1.0	81.17 ±1.4	81.59 ±1.0	81.66 ±1.2	81.32 ±1.1	81.93 ±1.1	80.19 ±1.5	80.31 ±1.3	79.13 ±2.3	79.51 ±1.7	80.18 ±1.9	79.78 ±1.4	77.73 ±2.0

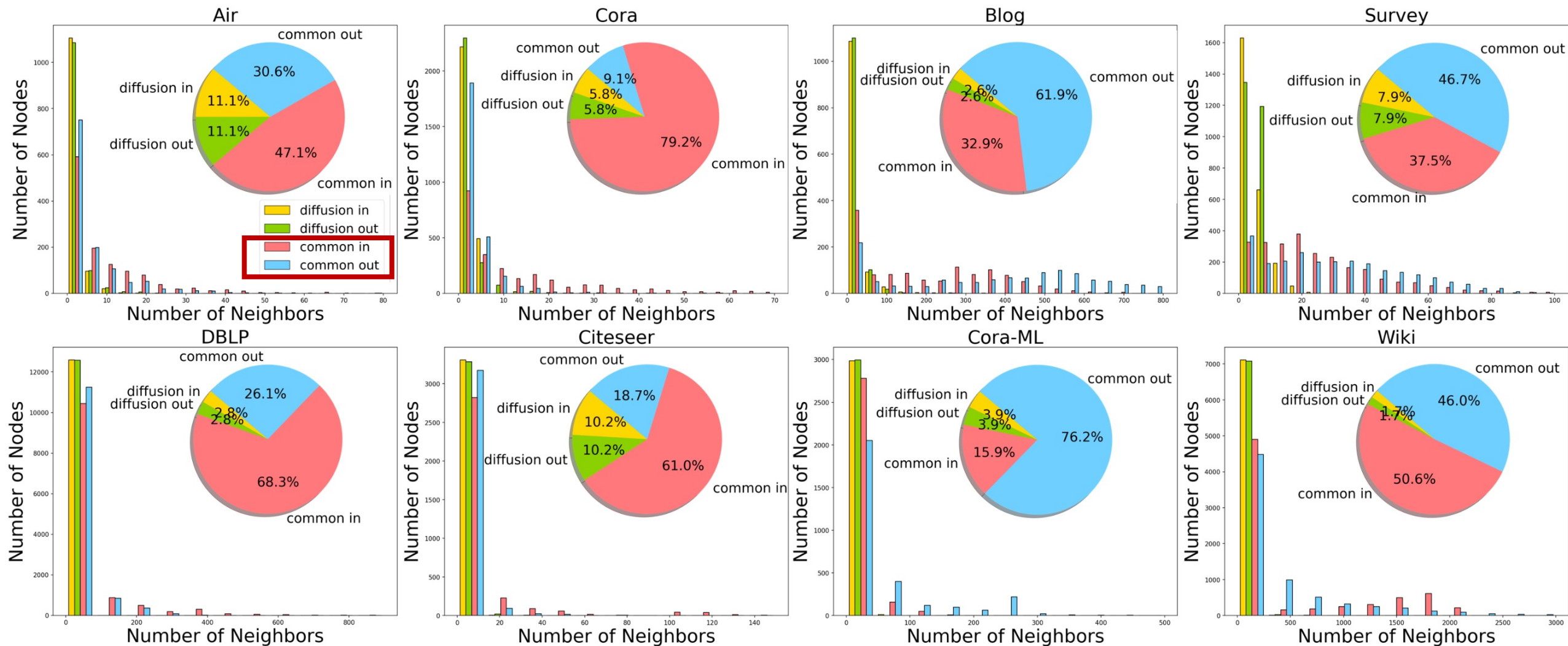
We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.

# Ablation Study

Method	CiteSeer	Cora-ML
D-HYPR (Our Full Design)	$70.66 \pm 1.2$	$82.19 \pm 1.3$
No $A_{d_{in}}^k$	$68.72 \pm 1.2$	$82.11 \pm 1.2$
No $A_{d_{out}}^k$	$69.10 \pm 0.9$	$81.33 \pm 1.4$
No $A_{c_{in}}^k$	$69.98 \pm 1.0$	$81.86 \pm 1.6$
No $A_{c_{out}}^k$	$69.84 \pm 1.3$	$81.74 \pm 1.8$
No Hyperbolic Neighborhood Collaboration	$70.13 \pm 1.5$	$82.03 \pm 1.1$
No Gravity	$68.58 \pm 1.3$	$79.21 \pm 1.5$
No Fermi-Dirac	$70.03 \pm 1.2$	$82.05 \pm 1.3$
No Self-Supervision	$67.85 \pm 1.9$	$78.15 \pm 2.1$
Euclidean	$61.86 \pm 5.4$	$73.38 \pm 6.7$
Euclidean and No Neighborhood Collaboration	$51.01 \pm 6.2$	$65.46 \pm 12.1$
A + Three Learnable Matrices	$60.97 \pm 12.7$	$78.92 \pm 2.9$

We propose **Digraph HYPERbolic Networks (D-HYPR)** to address the problem.

# Neighborhood Analysis of Datasets



**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Conclusion

- ❑ We propose **D-HYPR**: the Digraph HYPERbolic Network, as a novel GNN-based formalism for **Digraph Representation Learning** (DRL) by addressing **Neighborhood Modeling** and **Asymmetry Preservation**.
- ❑ Through extensive and rigorous evaluation involving **21** prior techniques, we empirically demonstrate the superiority of D-HYPR.
- ❑ D-HYPR retains effectiveness given a **low budget of embedding dimensionality or labeled training samples**, which is desirable for real-world applications.

**Limitations:** increased number of parameters, due to the use of multiple neighborhoods.

## **Future work:**

- Automatic and dynamic neighborhood partitioning
- Parameter-sharing mechanism
- Theoretical analyses
- Novel large-scale applications

**We propose Digraph HYPERbolic Networks (D-HYPR) to address the problem.**

# Thank you!

Code and data:

<https://github.com/hongluzhou/dhypr>



## Honglu Zhou

- ❖ CS PhD Student
- ❖ IVI Lab, Rutgers University
- ❖ [honglu.zhou@rutgers.edu](mailto:honglu.zhou@rutgers.edu)
- ❖ Website: <https://sites.google.com/view/hongluzhou/>



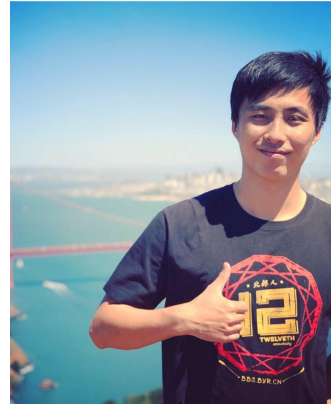
## Advith Chegus

- ❖ CS Master Student
- ❖ Rutgers University
- ❖ [ac1771@rutgers.edu](mailto:ac1771@rutgers.edu)
- ❖ Website: <https://github.com/advil64>



## Samuel S. Sohn

- ❖ CS PhD Student
- ❖ IVI Lab, Rutgers University
- ❖ [sss286@cs.rutgers.edu](mailto:sss286@cs.rutgers.edu)
- ❖ Website: <https://www.sohn.tech/>



## Zuohui Fu

- ❖ CS PhD Student
- ❖ Rutgers University
- ❖ [zuohui.fu@rutgers.edu](mailto:zuohui.fu@rutgers.edu)
- ❖ Website: <https://zuohuif.github.io/>



## Gerard de Melo

- ❖ Professor
- ❖ HPI / University of Potsdam
- ❖ [gerard.demelo@hpi.de](mailto:gerard.demelo@hpi.de)
- ❖ Website: <http://gerard.demelo.org>



## Mubbasir Kapadia

- ❖ Associate Professor
- ❖ IVI Lab, Rutgers University
- ❖ [mubbasir.kapadia@rutgers.edu](mailto:mubbasir.kapadia@rutgers.edu)
- ❖ Website: <https://ivi.cs.rutgers.edu/>