COLLEGE OF ENGINEERING
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

MEDICAL SCHOOL
UNIVERSITY OF MICHIGAN

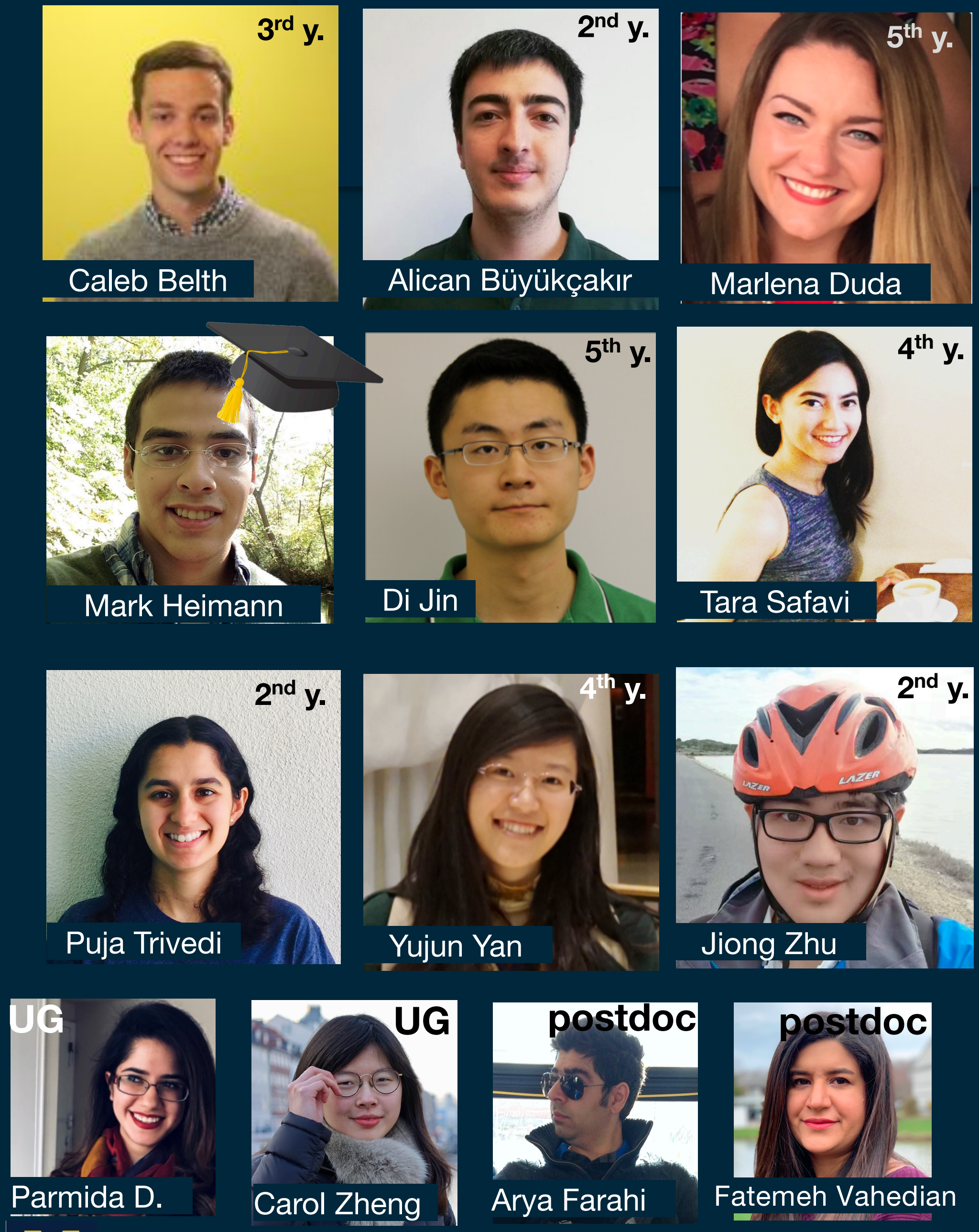
GEMS LAB

Representation Learning Beyond Homophily & Proximity

Danai Koutra
Morris Wellman Assistant Professor, CSE
Computational Medicine and Bioinformatics (courtesy)

Statistical Inference for Network Models Symposium, NetSci – Sep 20, 2020

Joint work with: Leman Akoglu, Mark Heimann, Di Jin, Ryan Rossi, Tara Safavi, Yujun Yan, Lingxiao Zhao, Jiong Zhou …



M CSE GEMS LAB

++

GEMS Lab @ University of Michigan — Home | People | News | Research | Data and code | Lab photos

GEMS LAB

Welcome!

We are the **Graph Exploration and Mining at Scale (GEMS)** lab at the University of Michigan, founded and led by Danai Koutra. Our team researches important data mining and machine learning problems involving interconnected data: in other words, *graphs or networks*.

From airline flights to traffic routing to neuronal interactions in the brain, graphs are ubiquitous in the real world. Their properties and complexities have long been studied in fields ranging from mathematics to the social sciences. However, many pressing problems involving graph data are still open. One well-known problem is *scalability*. With continual advances in data generation and storage capabilities, the size of graph datasets has dramatically increased, making scalable graph methods indispensable. Another is the changing nature of data. Real graphs are almost always *dynamic*, evolving over time. Finally, many important problems in the social and biological sciences involve analyzing not one but *multiple* networks.

So, what do we do?

The problems described above call for **principled, practical, and highly scalable graph mining methods**, both theoretical and application-oriented. As such, our work connects to fields like linear algebra, distributed systems, deep learning, and even neuroscience. Some of our ongoing projects include:

- Algorithms for multi-network tasks, like matching nodes across networks
- Learning low-dimensional representations of networks in metric spaces
- Abstracting or "summarizing" a graph with a smaller network
- Analyzing network models of the brain derived from fMRI scans
- Distributed graph methods for iteratively solving linear systems
- Network-theoretical user modeling for various data science applications

We're grateful for funding from Adobe, Amazon, the Army Research Lab, the Michigan
for Data Science (MIDAS), Microsoft Azure, the National Science Foundation (NSF), and

Interested?

If you're interested in joining our group, send an email with your interests and CV to g
opportunities@umich.edu.

COMPUTER SCIENCE & ENGINEERING UNIVERSITY OF MICHIGAN — MIDAS MICHIGAN INSTITUTE FOR DATA SCIENCE UNIVERSITY OF MICHIGAN

News

May 2020
1 paper accepted to KDD'20!

April 2020
Caleb receives an NDSEG Fellowship!

March 2020
Caleb receives an NSF GRFP!

February 2020
Danai receives a Google Faculty Research Award!

February 2020
Danai was recognised as an Outstanding Senior PC Member at WSDM'20!

January 2020
1 paper accepted to WebConf

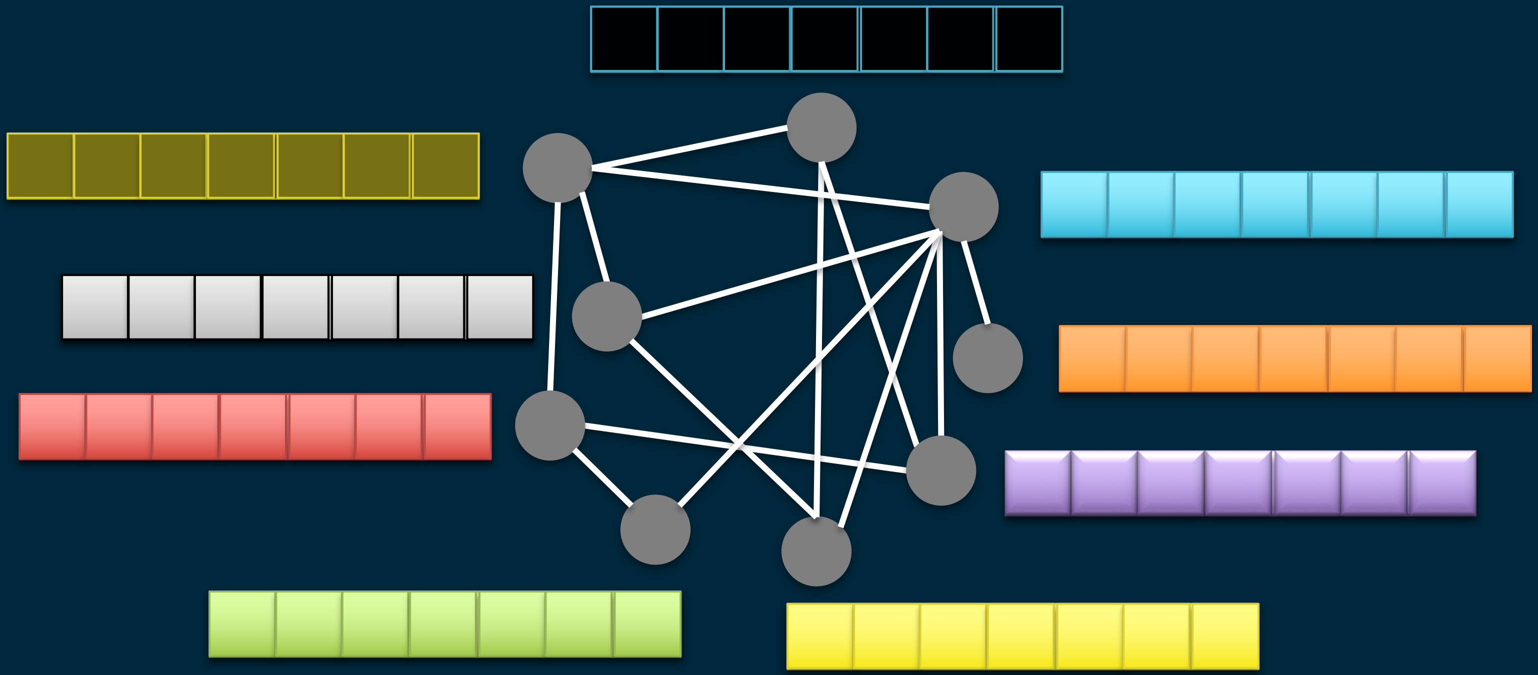
January 2020
Danai named Morris Wellman Professor!

January 2020
Research Fellow Fatemeh Vahedian

NSF · ARL · P&G · MIDAS · Adobe · amazon · Microsoft Azure

Network Representation Learning: Goal

- Given a graph G
- Automatically learn a feature vector representation for network objects (e.g., nodes, subgraphs)



"Must read papers in Network Representation Learning"

https://github.com/thunlp/NRLPapers

+ many more

MORGAN & CLAYPOOL PUBLISHERS

Graph Representation Learning

William L. Hamilton

McGill University

2020

Pre-publication draft of a book to be published by Morgan & Claypool publishers.
Unedited version released with permission.
All relevant copyrights held by the author and publisher extend to this pre-publication draft.

Citation: William L. Hamilton. (2020). Graph Representation Learning. Morgan and Claypool, *forthcoming*.

A Comprehensive Survey on Graph Neural Networks

Zonghan Wu, Shirui Pan, *Member, IEEE*, Fengwen Chen, Guodong Long,
Chengqi Zhang, *Senior Member, IEEE*, Philip S. Yu, *Fellow, IEEE*

CAMBRIDGE UNIVERSITY PRESS

Deep Learning on Graphs

by Yao Ma and Jiliang Tang

Heterogeneous Network Representation Learning

Yuxiao Dong[1], **Ziniu Hu**[2], **Kuansan Wang**[1], **Yizhou Sun**[2] and **Jie Tang**[3]
[1]Microsoft Research, Redmond
[2]University of California, Los Angeles
[3]Tsinghua University, Beijing
{yuxdong, kuansanw}@microsoft.com, {bull, yzsun}@cs.ucla.edu, jietang@tsinghua.edu.cn

Abstract

Representation learning has offered a revolution-ary learning paradigm for various AI domains. In his survey, we examine and review the problem of epresentation learning with the focus on heteroge-neous networks, which consists of different types of vertices and relations. The goal of this problem s to automatically project objects, most commonly,

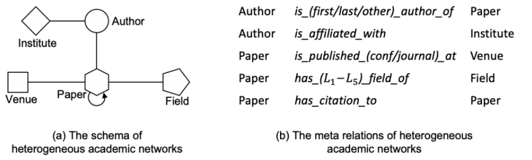


Figure 1: The schema and meta relations of Open Academic Graph

Journal of Machine Learning Research 21 (2020) 1-73 Submitted 6/19; Revised 2/20; Published 3/20

Representation Learning for Dynamic Graphs: A Survey

Seyed Mehran Kazemi MEHRAN.KAZEMI@BOREALISAI.COM
Rishab Goel RISHAB.GOEL@BOREALISAI.COM
Borealis AI, 310-6666 Saint Urbain, Montreal, QC, Canada

Kshitij Jain KSHITIJ.JAIN@BOREALISAI.COM
Ivan Kobyzev IVAN.KOBYZEV@BOREALISAI.COM
Akshay Sethi AKSHAY.SETHI@BOREALISAI.COM
Peter Forsyth PETER.FORSYTH@BOREALISAI.COM
Pascal Poupart PASCAL.POUPART@BOREALISAI.COM
Borealis AI, 301-420 West Graham Way, Waterloo, ON, Canada

M CSE GEMS LAB

“Must read papers in Network Representation Learning”
https://github.com/thunlp/NRLPapers

MORGAN & CLAYPOOL PUBLISHERS

GRAPH REPRESENTATION LEARNING

WILLIAM L. HAMILTON

PRE-PUBLICATION DRAFT OF A BOOK TO BE PUBLISHED BY MORGAN & CLAYPOOL PUBLISHERS. UNEDITED VERSION RELEASED WITH PERMISSION. ALL RELEVANT COPYRIGHTS HELD BY THE AUTHOR AND PUBLISHER EXTEND TO THIS PRE-PUBLICATION DRAFT.

Citation: William L. Hamilton. (2020). Graph Representation Learning. Morgan and Claypool, *forthcoming*.

JOURNAL OF LaTeX CLASS FILES, VOL. XX, NO. XX, AUGUST 2019 1

A Comprehensive Survey on Graph Neural Networks

Zonghan Wu, Shirui Pan, *Member, IEEE*, Fengwen Chen, Guodong Long, Chengqi Zhang, *Senior Member, IEEE*, Philip S. Yu, *Fellow, IEEE*

+ many more

Heterogeneous Network Representation Learning

Jie Tang[3]

@tsinghua.edu.cn

Most work focuses on two typical settings:

homophily & proximity

by Yao Ma and Jiliang Tang

Journal of Machine Learning Research 21 (2020) 1-73 Submitted 6/19; Revised 2/20; Published 3/20

Representation Learning for Dynamic Graphs: A Survey

Seyed Mehran Kazemi MEHRAN.KAZEMI@BOREALISAI.COM
Rishab Goel RISHAB.GOEL@BOREALISAI.COM
Borealis AI, 310-6666 Saint Urbain, Montreal, QC, Canada

Kshitij Jain KSHITIJ.JAIN@BOREALISAI.COM
Ivan Kobyzev IVAN.KOBYZEV@BOREALISAI.COM
Akshay Sethi AKSHAY.SETHI@BOREALISAI.COM
Peter Forsyth PETER.FORSYTH@BOREALISAI.COM
Pascal Poupart PASCAL.POUPART@BOREALISAI.COM
Borealis AI, 301-420 West Graham Way, Waterloo, ON, Canada

M CSE GEMS LAB

This talk

- Generalizing GNNs beyond homophily [Arxiv’20]

- Node embeddings: beyond proximity [ACM TKDD’20 +]

This talk

- Generalizing GNNs beyond homophily [Arxiv'20]

- Node embeddings: beyond proximity [ACM TKDD'20 +]

Based on the following paper

https://arxiv.org/abs/2006.11468

Generalizing Graph Neural Networks Beyond Homophily



Jiong Zhu
University of Michigan
jiongzhu@umich.edu

Yujun Yan
University of Michigan
yujunyan@umich.edu

Lingxiao Zhao
Carnegie Mellon University
lingxia1@andrew.cmu.edu

Mark Heimann
University of Michigan
heimann@umich.ed

Leman Akoglu
Carnegie Mellon University
oglu@andrew.cmu.edu

Danai Koutra
University of Michigan
dkoutra@umich.edu



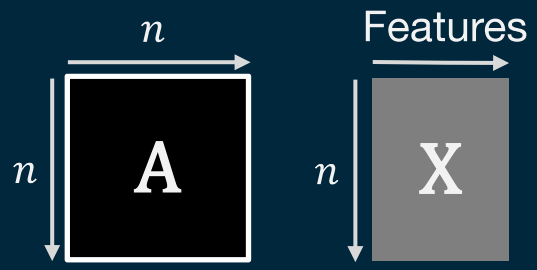
Abstract

We investigate the representation power of graph neural networks in the semi-supervised node classification task under *heterophily* or *low homophily*, i.e., in networks where connected nodes may have *different* class labels and *dissimilar* features. Most existing GNNs fail to generalize to this setting, and are even outperformed by models that ignore the graph structure (e.g., multilayer perceptrons). Motivated by this limitation, we identify a set of key designs—ego- and neighbor-embedding separation, higher-order neighborhoods, and combination of intermediate representations—that boost learning from the graph structure under heterophily, and combine them into a new graph convolutional neural network, H_2GCN. Going beyond the traditional benchmarks with strong homophily, our empirical analysis on synthetic and real networks shows that, thanks to the identified designs, H_2GCN has *consistently* strong performance across the full spectrum of low-to-high homophily, unlike competitive prior models without them.

1 Introduction

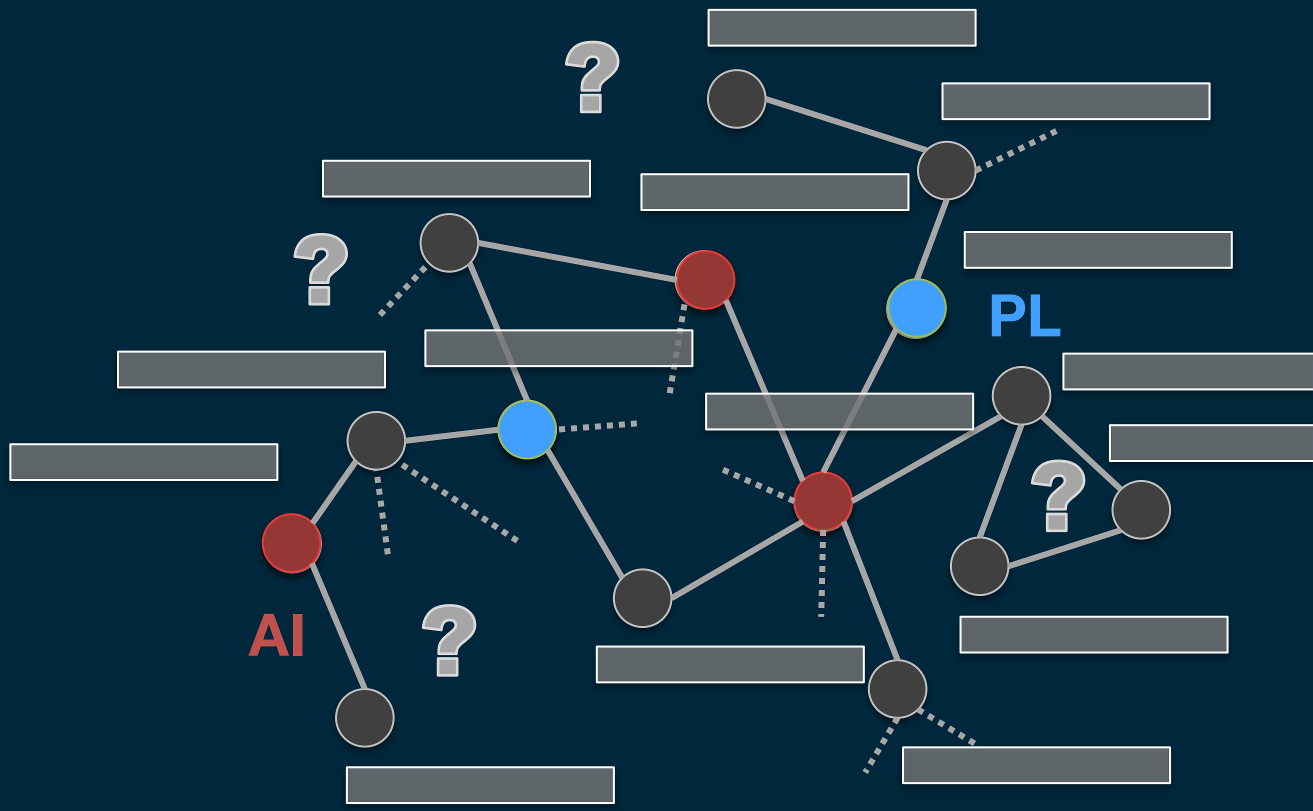
M CSE GEMS LAB

Semi-supervised Node Classification



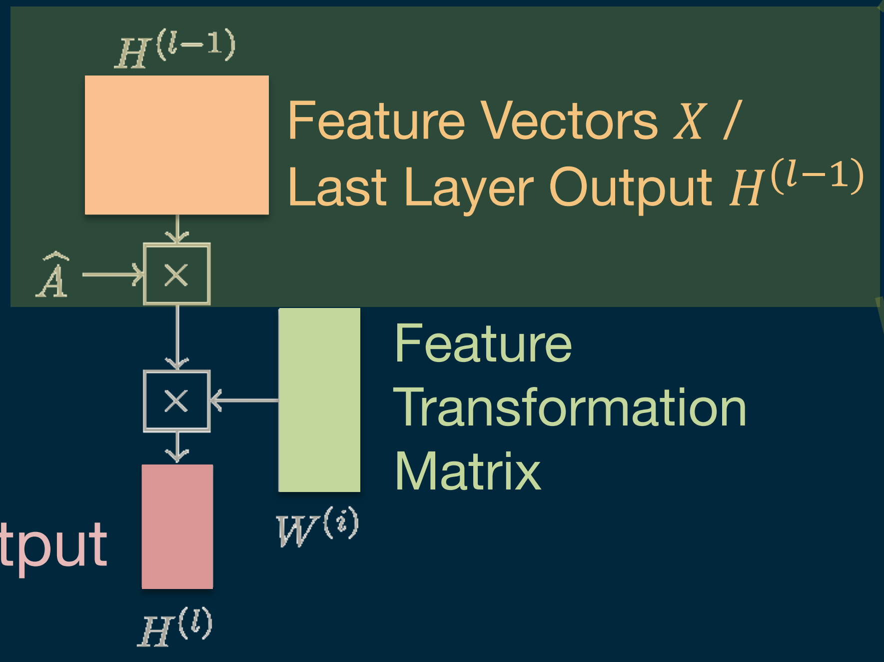
- Given a graph G with adjacency matrix **A**
  node feature matrix **X**
  a few labeled nodes (e.g., red/blue)

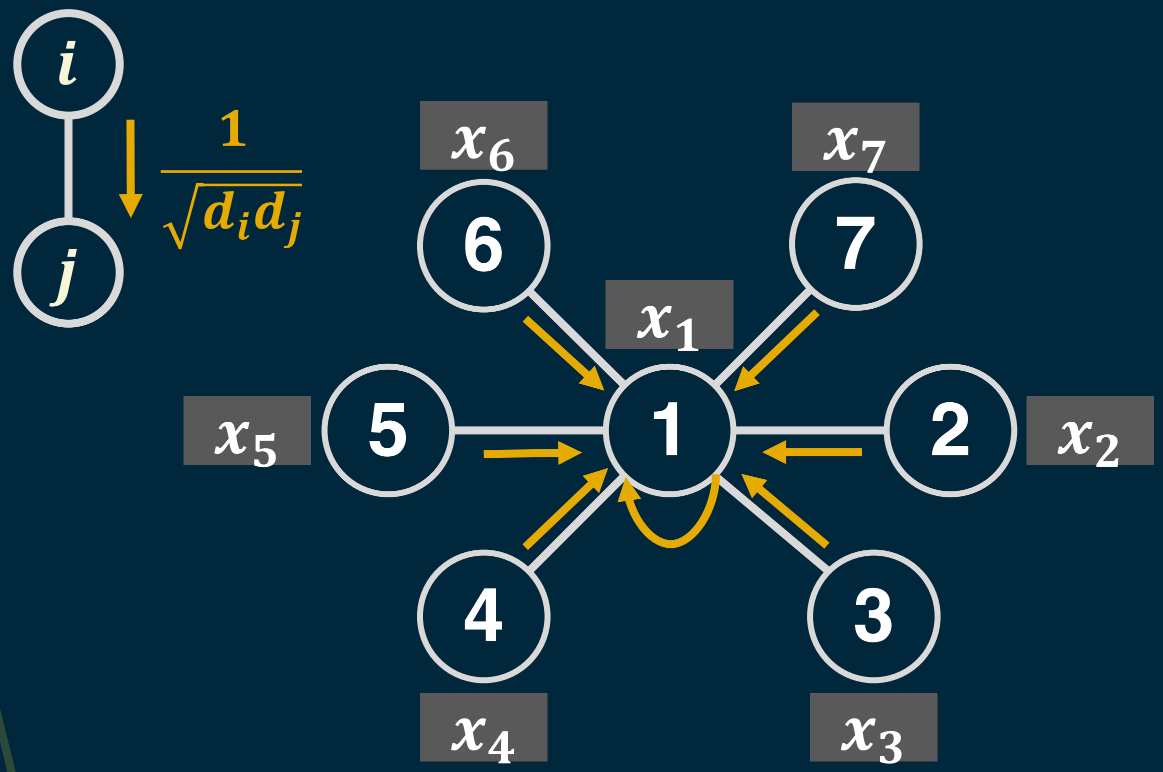
- Find the class label of each of the remaining nodes.



Graph Neural Networks

GCN [Kipf+ *ICLR17*]





Classification Result

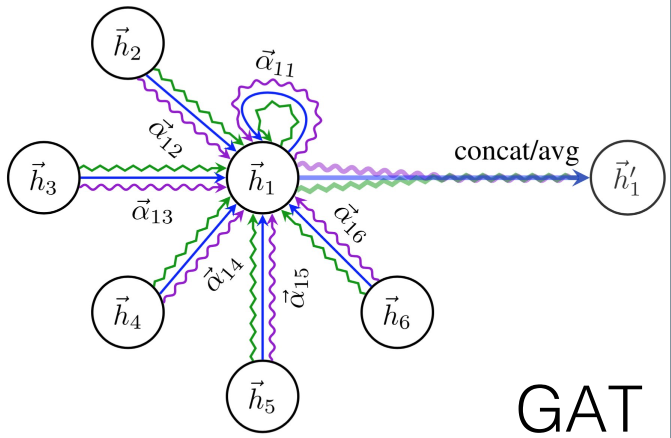
$$Z = f(X, A) = \mathrm{softmax}\Big(\hat{A}\ \mathrm{ReLU}\Big(\hat{A}XW^{(0)}\Big)\ W^{(1)}\Big)$$

$$\hat{A}X = \text{weighted\_avg}\ (x_1\ x_2\ x_3\ x_5\ x_6\ x_4\ x_7)$$

 [Block Diagram: Abu-El-Haija ICML'18]

Many architectures improving upon GCN

- Using different aggregators
  - GraphSAGE [NeurIPS17], …
- Adding an edge-level attention mechanism
  - GAT [ICLR17]
  - AGNN [arXiv18], …
- Aggregating beyond immediate neighborhood
  - MixHop [ICML19]
  - GDC [NeurIPS19]
  - Geom-GCN [ICLR20], …
- …



However, most existing GNN models are effective on graphs with strong homophily.

["Must-read papers on GNN" - https://github.com/thunlp/GNNPapers]

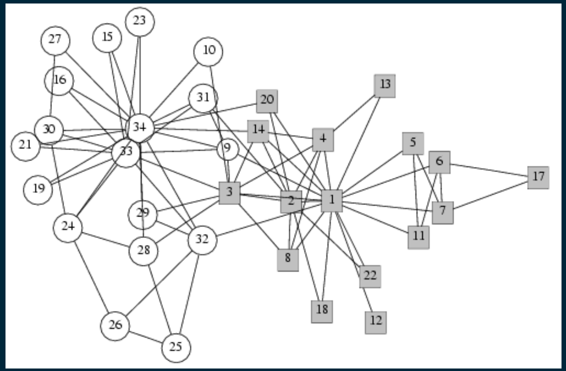
Homophily and Heterophily

Largely overlooked

Homophily

"Birds of a feather flock together"
Most of linked nodes are similar

- Social Networks (wrt. political beliefs, age)
- Citation Networks (wrt. research area)



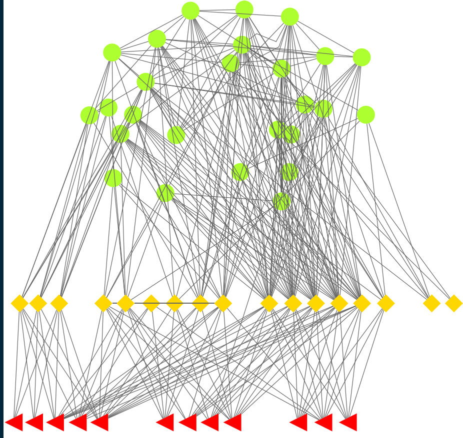
Zachary's Karate club

Heterophily

"Opposites Attract"
Most of linked nodes are different

- Friend network (e.g., talkative / silent friends)
- Protein structures (wrt. amino acid types)
- E-commerce (wrt. fraudsters / accomplices)

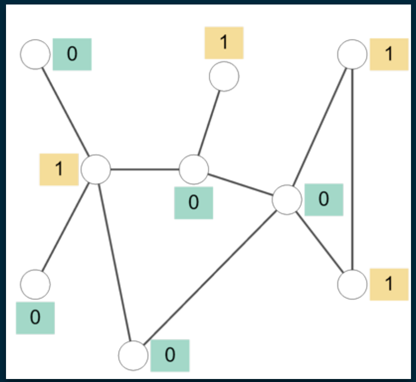
Honest Users
Accomplices
Fraudsters

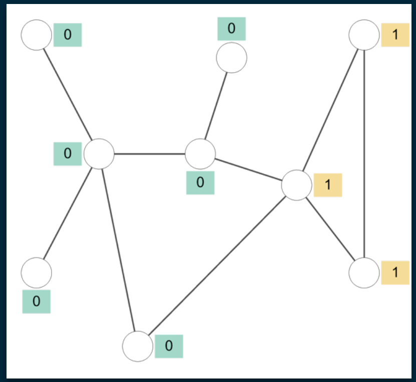


Measuring Homophily / Heterophily

- Edge homophily ratio h: fraction of intra-class edges (i.e., total edges which link nodes with the same class)

$$h = \frac{|\{(u,v): (u,v) \epsilon E \wedge y_u = y_v\}|}{|E|}$$





Strong Heterophily

$h = 0$

h

Strong Homophily

$h = 1$

$h = 0.3$

$h = 0.8$

M CSE GEMS LAB

Our Contributions

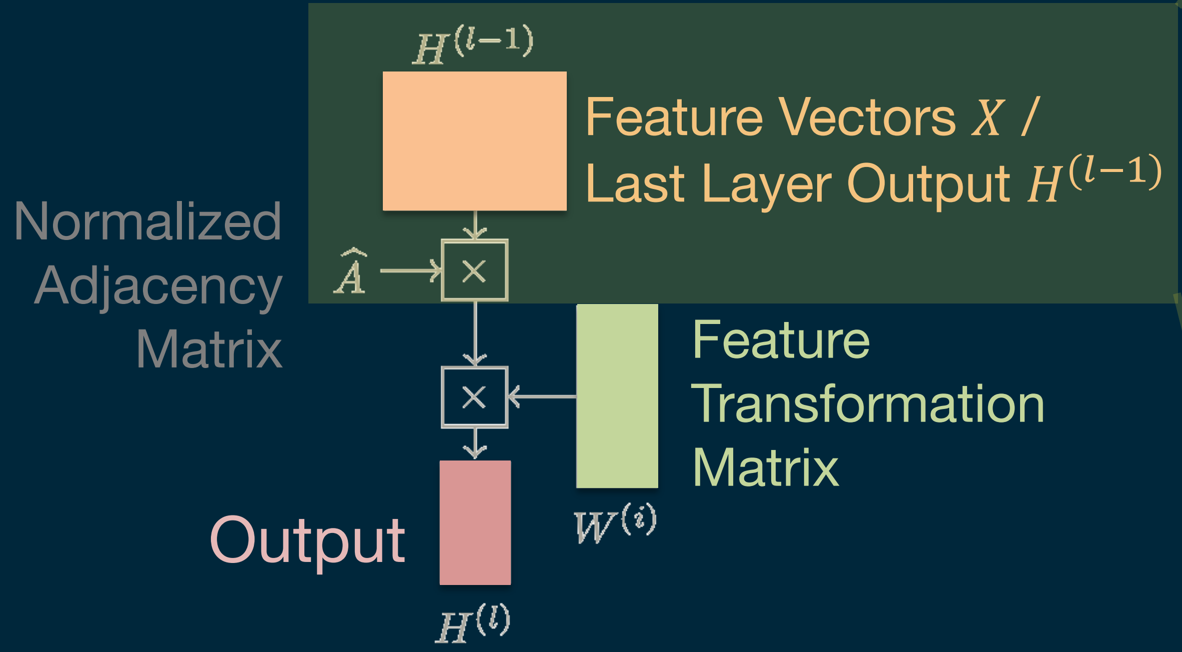
Reveal current limitations of GNNs in heterophily settings

Identify key design choices that boost learning in heterophily, without trading off accuracy in homophily

Conduct an extensive empirical evaluation

Revisiting GCN & Homophily Assumption

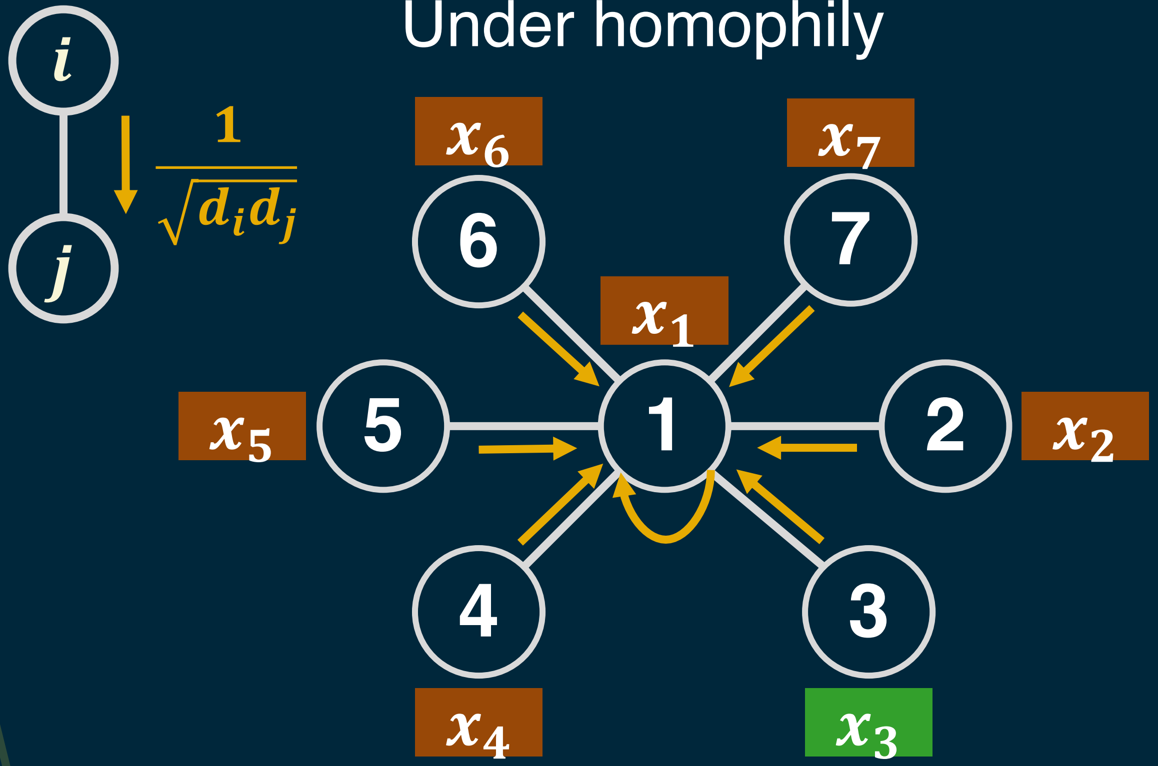
GCN [Kipf+ *ICLR17*]



Classification Result

$$Z = f(X, A) = \text{softmax}\Big(\hat{A}\ \text{ReLU}\Big(\hat{A}XW^{(0)}\Big)\ W^{(1)}\Big)$$

Under homophily

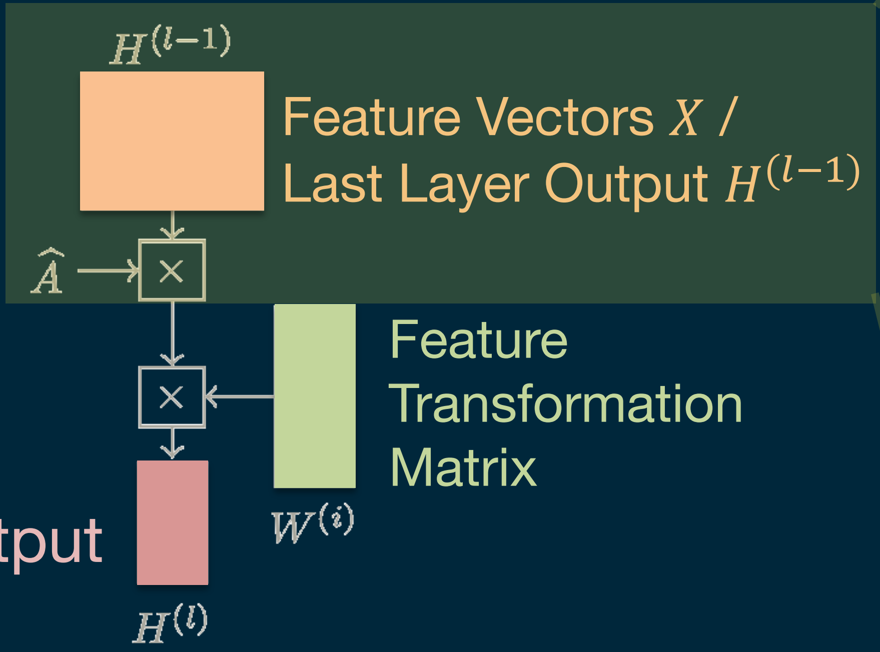


$$\hat{A}X = \text{weighted\_avg}\,(x_1\ x_2\ x_3\ x_5\ x_6\ x_4\ x_7) =$$

In homophily cases, the GCN aggregator will
help with denoising and generalization.

When GCN meets Heterophily…

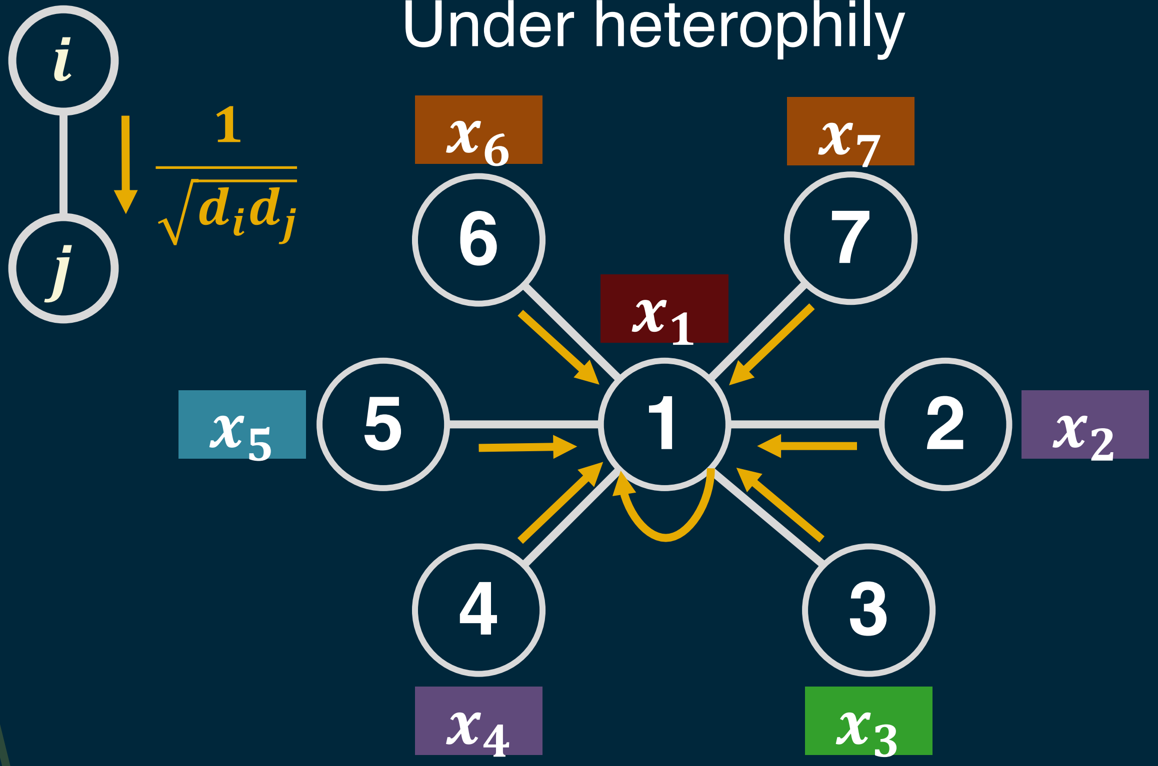
GCN [Kipf+ *ICLR17*]



Classification Result

$$Z = f(X, A) = \mathrm{softmax}\Big(\hat{A}\ \mathrm{ReLU}\big(\hat{A}XW^{(0)}\big)\ W^{(1)}\Big)$$

Under heterophily

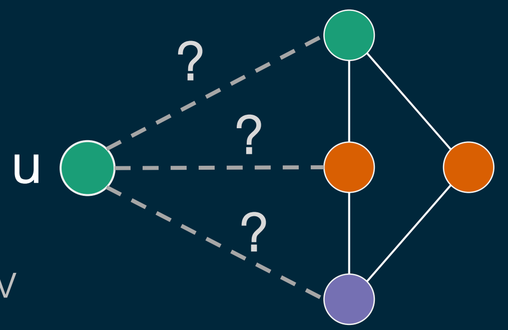


$$\hat{X} = \text{weighted\_avg}\,(x_1\ x_2\ x_3\ x_5\ x_6\ x_4\ x_7) = \ ??$$

In heterophily cases, the GCN aggregator will blur the features, making them indistinguishable.

Heterophily: Empirical Study Setup

- Synthetic graphs:
  - ✧ Control the edge homophily ratio h
  - ✧ Modified preferential attachment model
    - P[new node u links to existing node v] $\propto h_{ij} \cdot d_v$



degree of v

Class Compatibility Matrix

i & j: class of u & v

dst node class

src node class

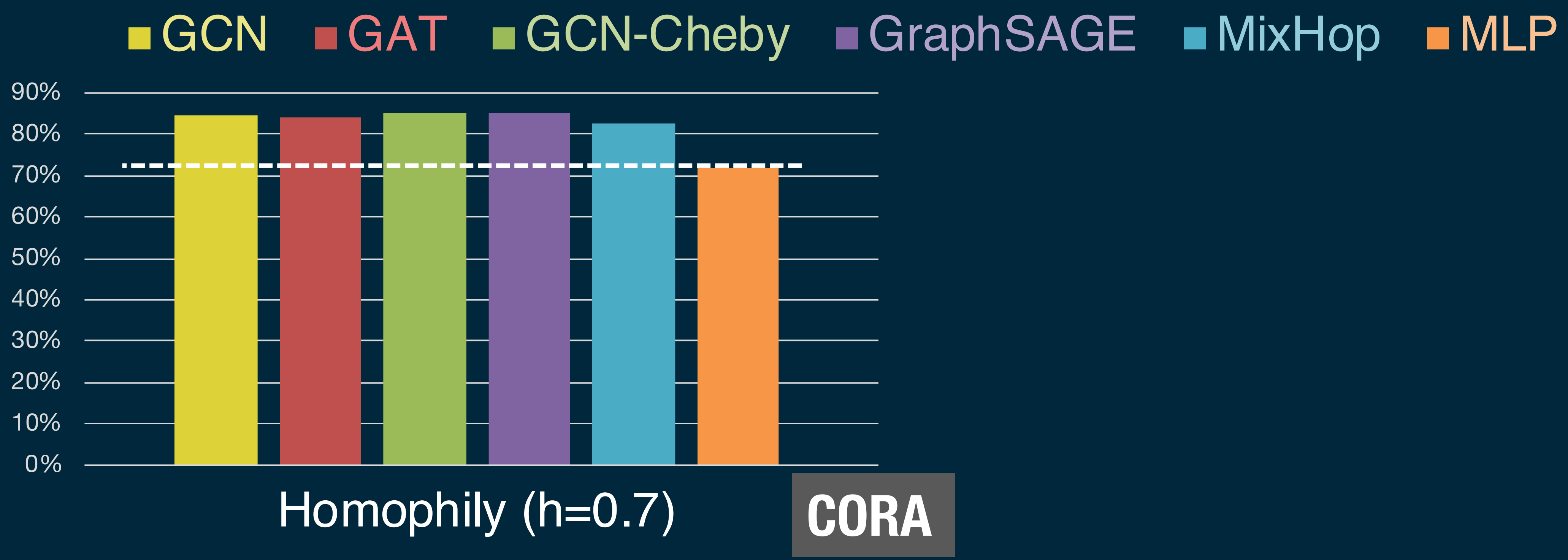
0.20	0.40	0.40
0.40	0.20	0.40
0.40	0.40	0.20

edge homophily ratio h

  - ✧ Node feature vectors sampled from real graphs (e.g., Cora)

GEMS LAB [Karimi et al. Scientific Reports '18] [Abu-El-Haija et al. ICML'19]

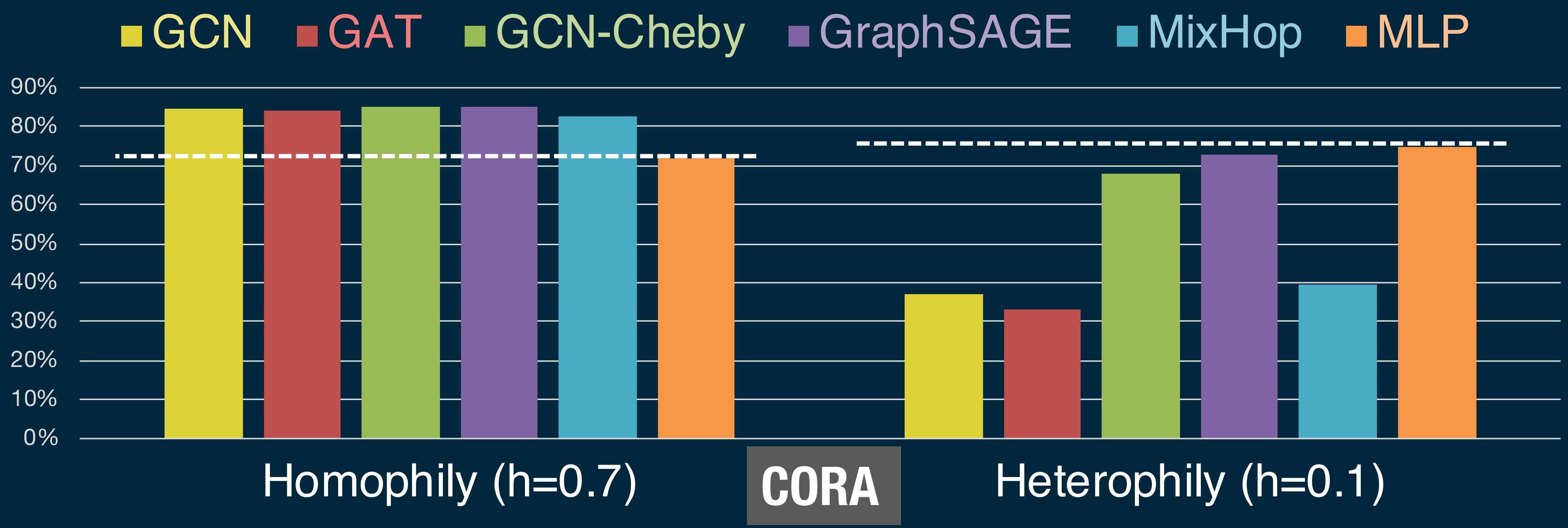
Heterophily: Empirical Study



[Kipf & Welling. ICLR’17] [Veličković et al. ICLR’18] [Defferrard et al. NeurIPS’16] [Hamilton et al. NeurIPS’17] [Abu-El-Haija et al. ICML’19]

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

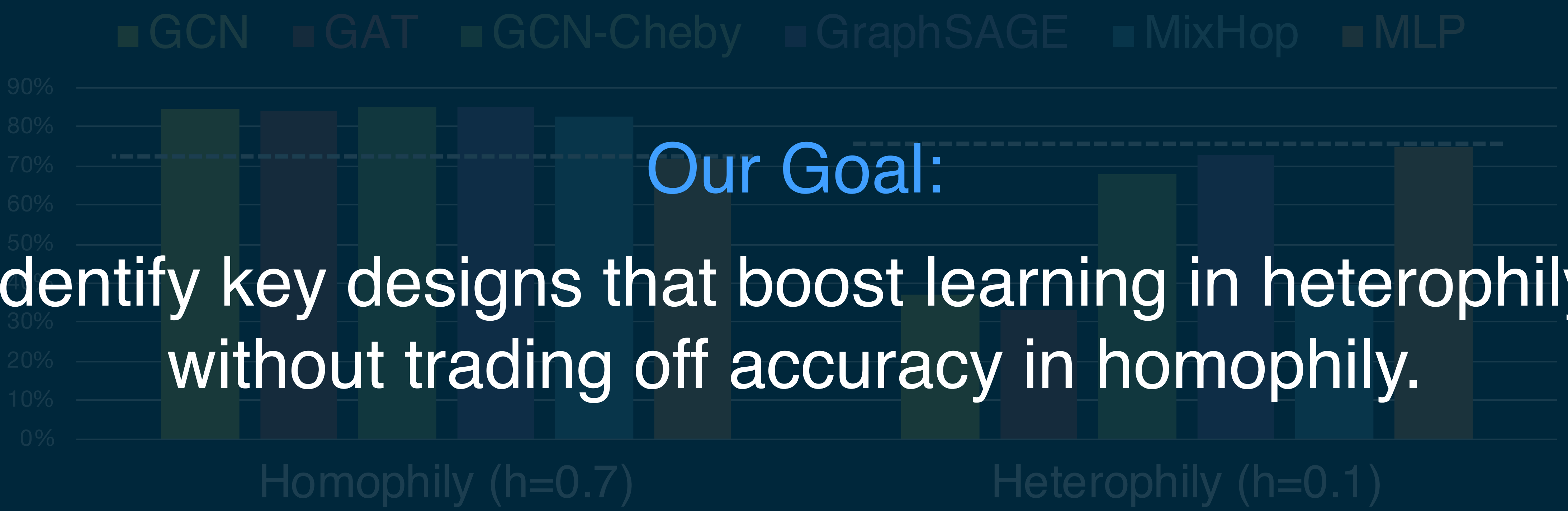
Heterophily: Empirical Study



Under heterophily, Multilayer Perceptron (MLP), which is graph agnostic, performs better than GNN variants.

M CSE GEMS LAB [Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

Heterophily: Empirical Study



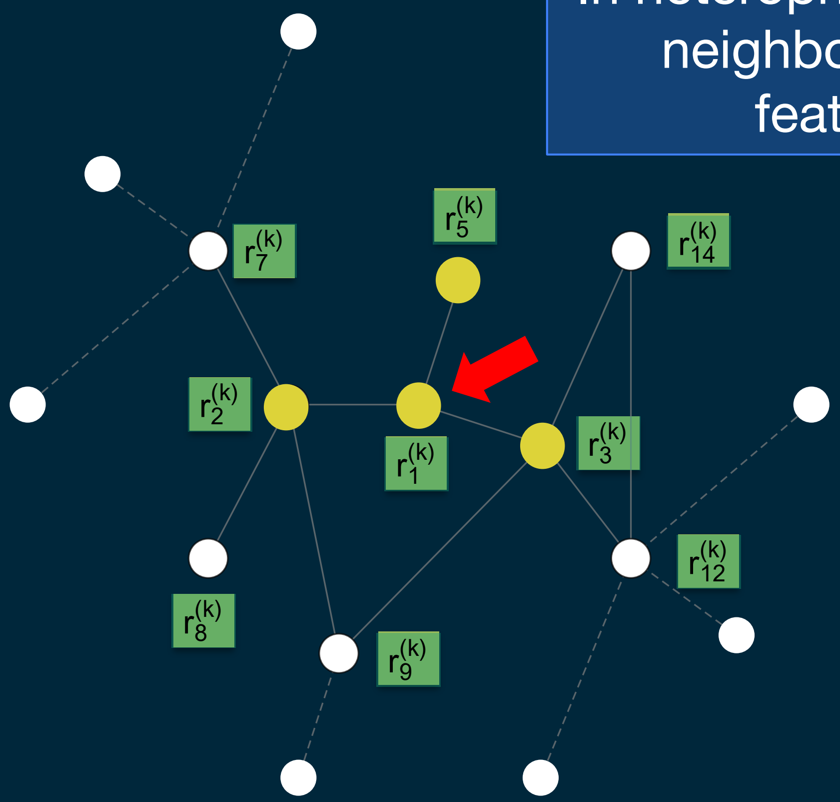
Our Goal:

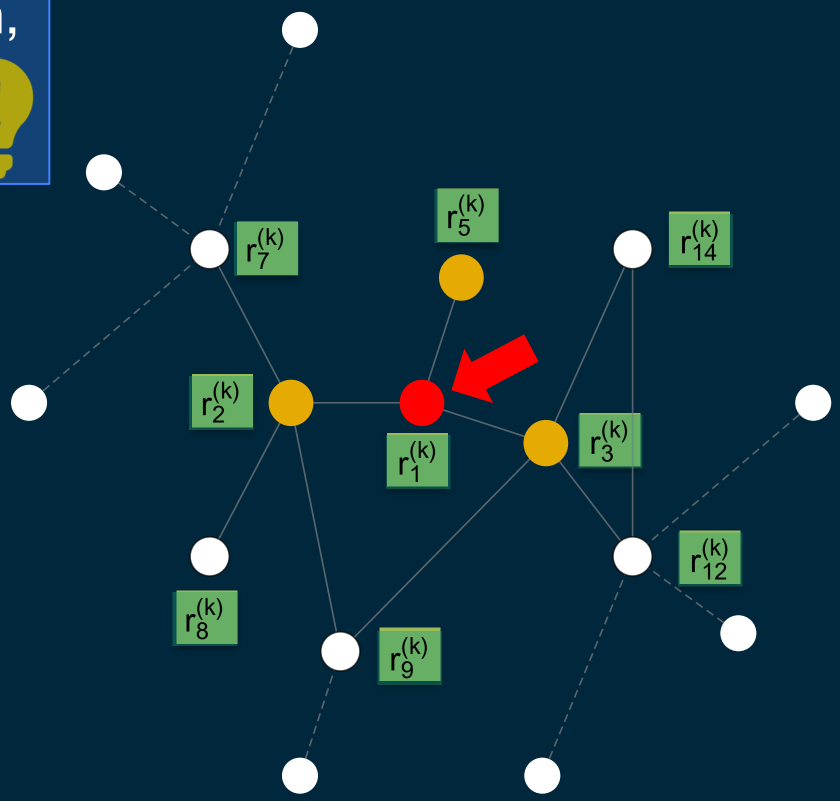
Identify key designs that boost learning in heterophily, without trading off accuracy in homophily.

Under heterophily, Multilayer Perceptron (MLP), which is graph agnostic, performs better than GNN variants.

D1: Ego- & Neighbor-embedding Separation

In heterophily settings, by definition, neighbors may have different features and classes.





$$r_1^{(k+1)} = \mathrm{AGGR}(\{\bullet\})$$

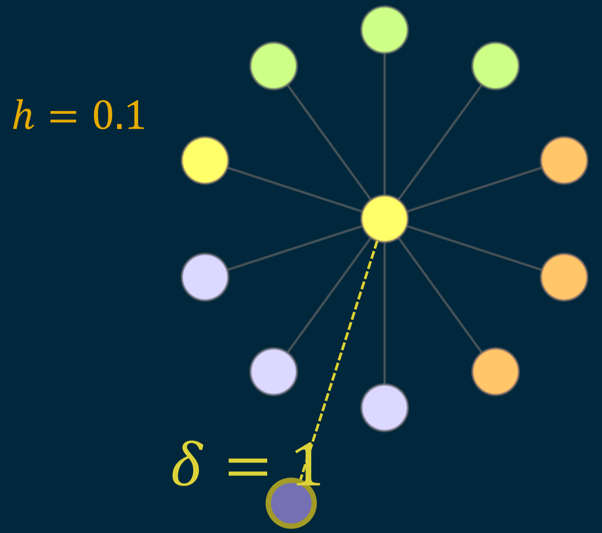
$$r_1^{(k+1)} = \mathrm{COMBINE}(\bullet, \mathrm{AGGR}(\{\bullet\}))$$

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

D1: Theoretical Justification

Goal. Compare generalization ability of two GCN layer formulations:

$\mathbf{AXW}$ and $(\mathbf{A}+\mathbf{I})\mathbf{XW}$ (without separation).



Theorem 1. In heterophily settings, a GCN layer formulated as $(\mathrm{A}+\mathrm{I})\mathrm{XW}$, which does *not* separate ego- and neighbor-embeddings, misclassifies under a less amount of deviation $|\delta|$ and therefore generalizes less than a AXW layer.

Sketch of Proof

1. Derive closed form solutions for W under certain conditions (e.g., same h).
2. Add / remove δ neighbors with class labels different than the ego-class.
3. Compare the absolute amount of deviation $|\delta|$ needed for each formulation to misclassify.

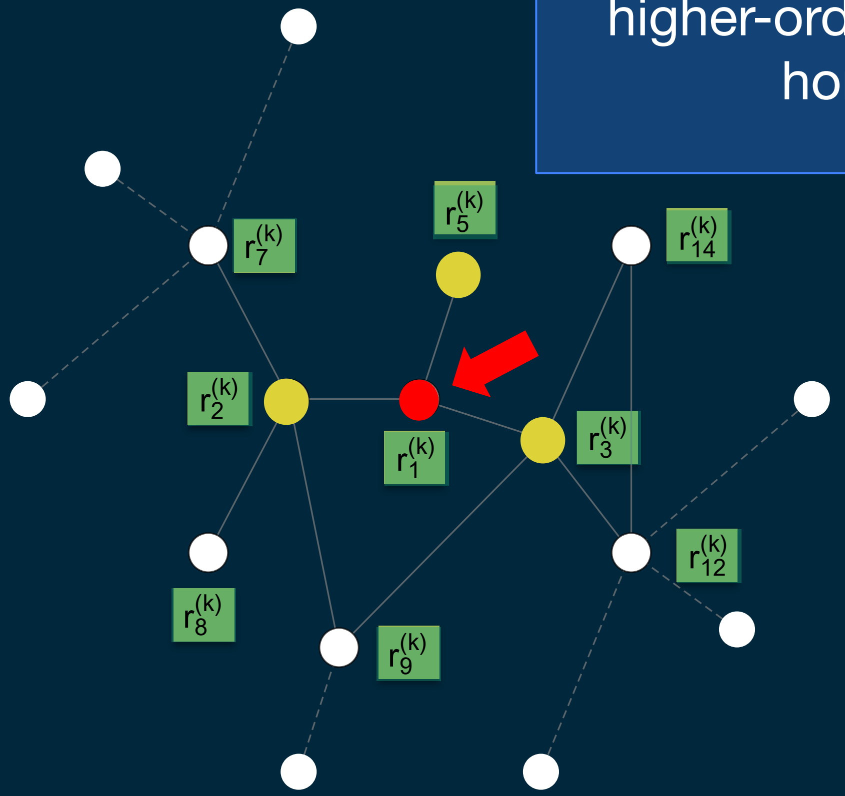
Reminder: 2-layer GCN

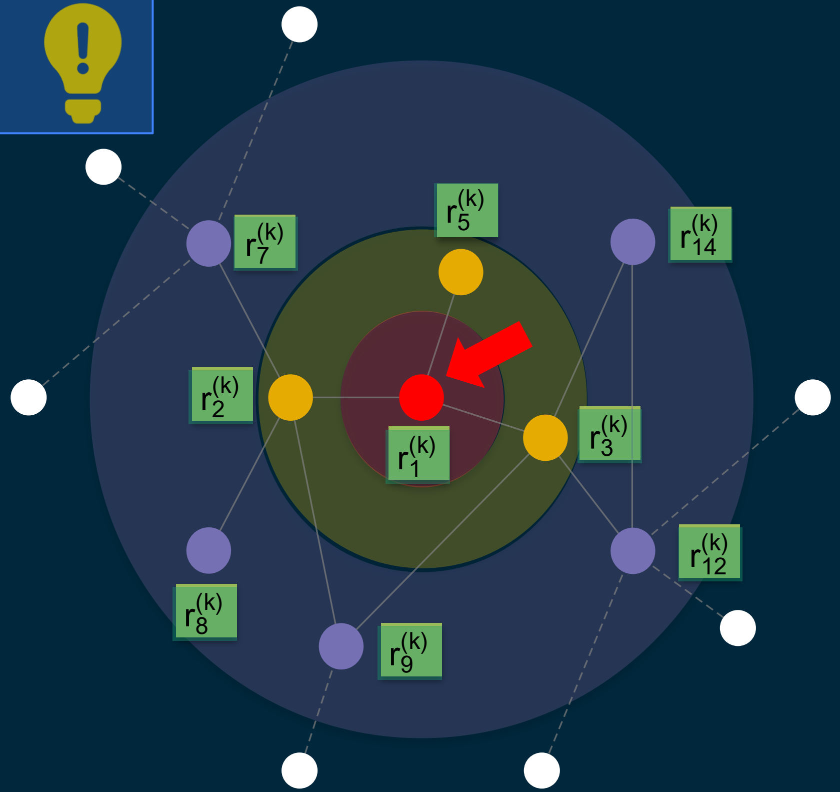
$$\mathrm{softmax}\left(\hat{A}\ \mathrm{ReLU}\left(\hat{A}XW^{(0)}\right)W^{(1)}\right)$$

A: adjacency matrix
X: node feature matrix
W: learnable weight matrix

D2: Higher-order Neighborhoods

In heterophily settings, in expectation, higher-order neighborhoods may be homophily-dominant.





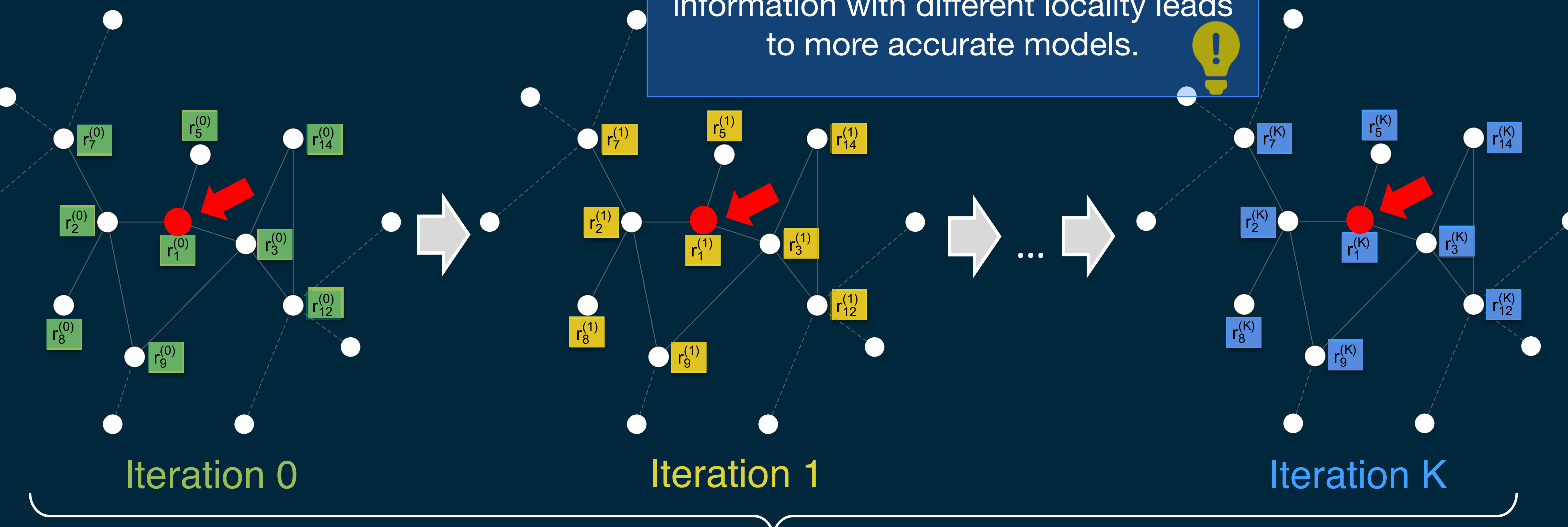
$$r_1^{(k+1)} = \mathrm{COMBINE}(\bullet, \mathrm{AGGR}(\{\bullet\}))$$

$$r_1^{(k+1)} = \mathrm{COMBINE}(\bullet, \mathrm{AGGR}(\{\bullet\}), \mathrm{AGGR}(\{\bullet\}))$$

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

D3: Combination of Intermediate Representations

In heterophily settings, collecting information with different locality leads to more accurate models.



$$r_1^{(\text{final})} = \text{COMBINE}(\mathbf{r}_1^{(0)}, \mathbf{r}_1^{(1)}, \dots, \mathbf{r}_1^{(K)})$$

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020] [Xu, Li, et al. ICML '18]

Overview of Designs

- Design D1 models (at each layer)
  - ✧ the ego- and neighbor-representations *distinctly*
- Design D2 leverages (at each layer)
  - ✧ representations of neighbors at different distances *distinctly*
- Design D3 leverages (at the final layer)
  - ✧ the learned ego-representations at previous layers *distinctly*

H_2GCN

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

Overview of Designs

- Design D1 models (at each layer)
  - ✧ the ego- and neighbor-representations *distinctly*
- Design D2 leverages (at each layer)
  - ✧ representations of neighbors at different distances *distinctly*
- Design D3 leverages (at the final layer)
  - ✧ the learned ego-representations at previous layers *distinctly*

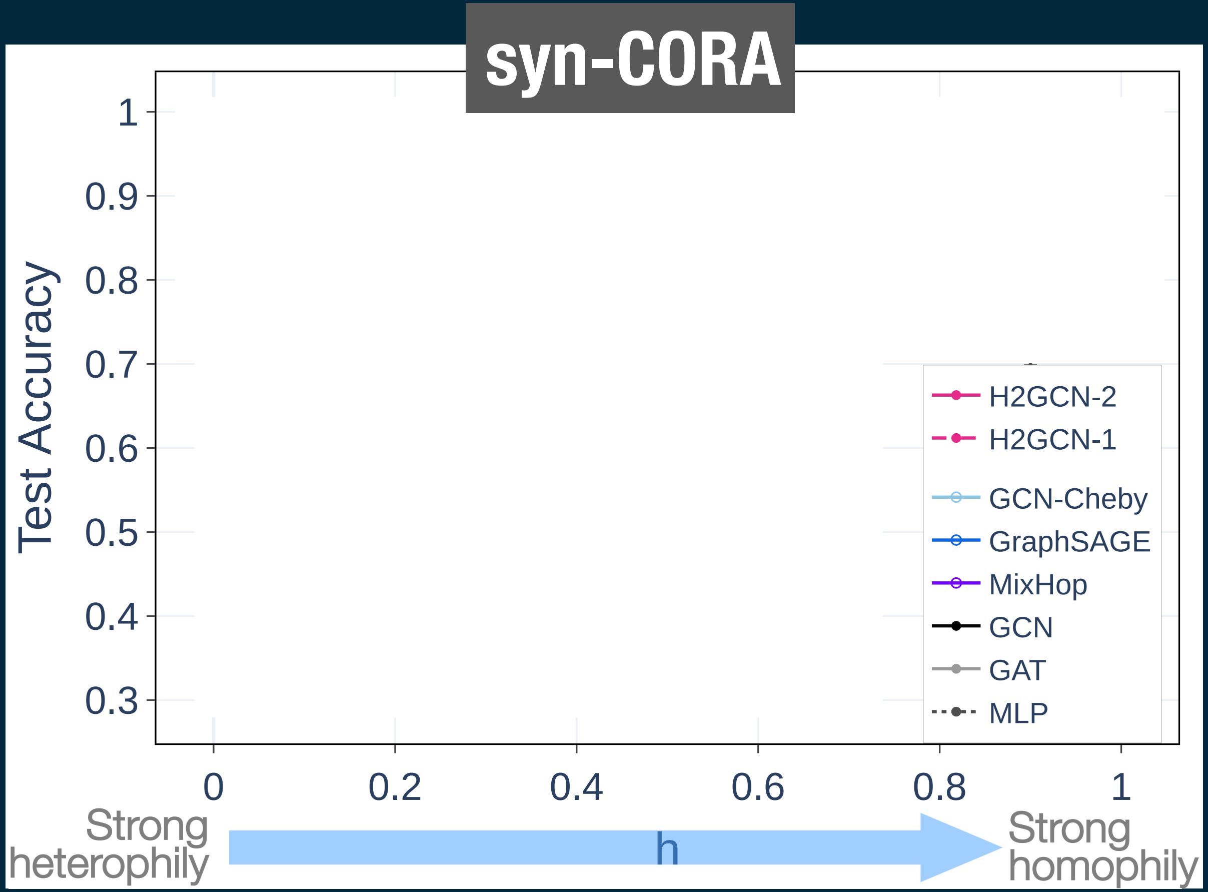
$\}$ H_2GCN

Existing works have used some subsets of these designs, but not in heterophily settings, and do not provide in-depth theoretical and empirical evaluations.

Method	D1	D2	D3
GCN [12]	✗	✗	✗
GAT [31]	✗	✗	✗
GCN-Cheby [5]	✗	✓	✗
GraphSAGE [8]	✓	✗	✗
MixHop [1]	✗	✓	✗
H_2GCN (proposed)	✓	✓	✓

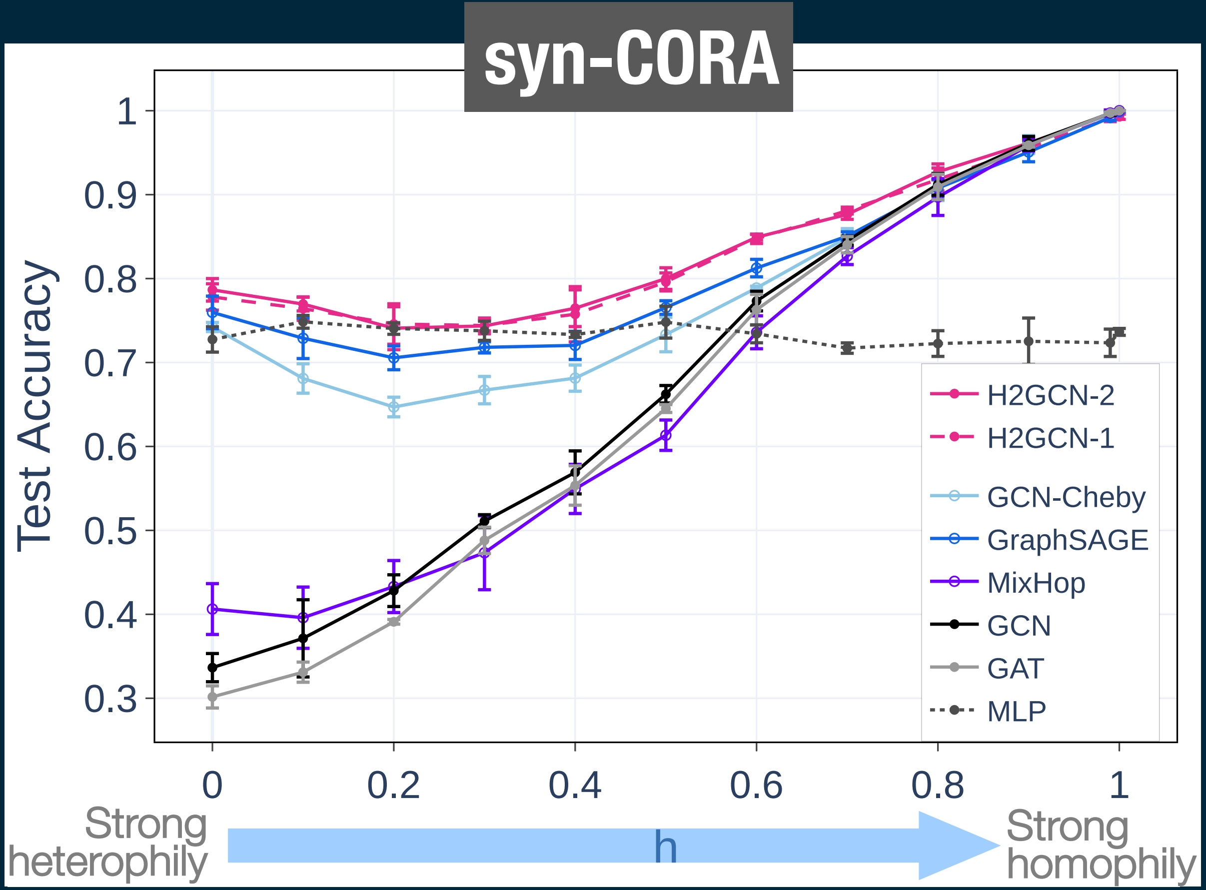
M CSE GEMS LAB

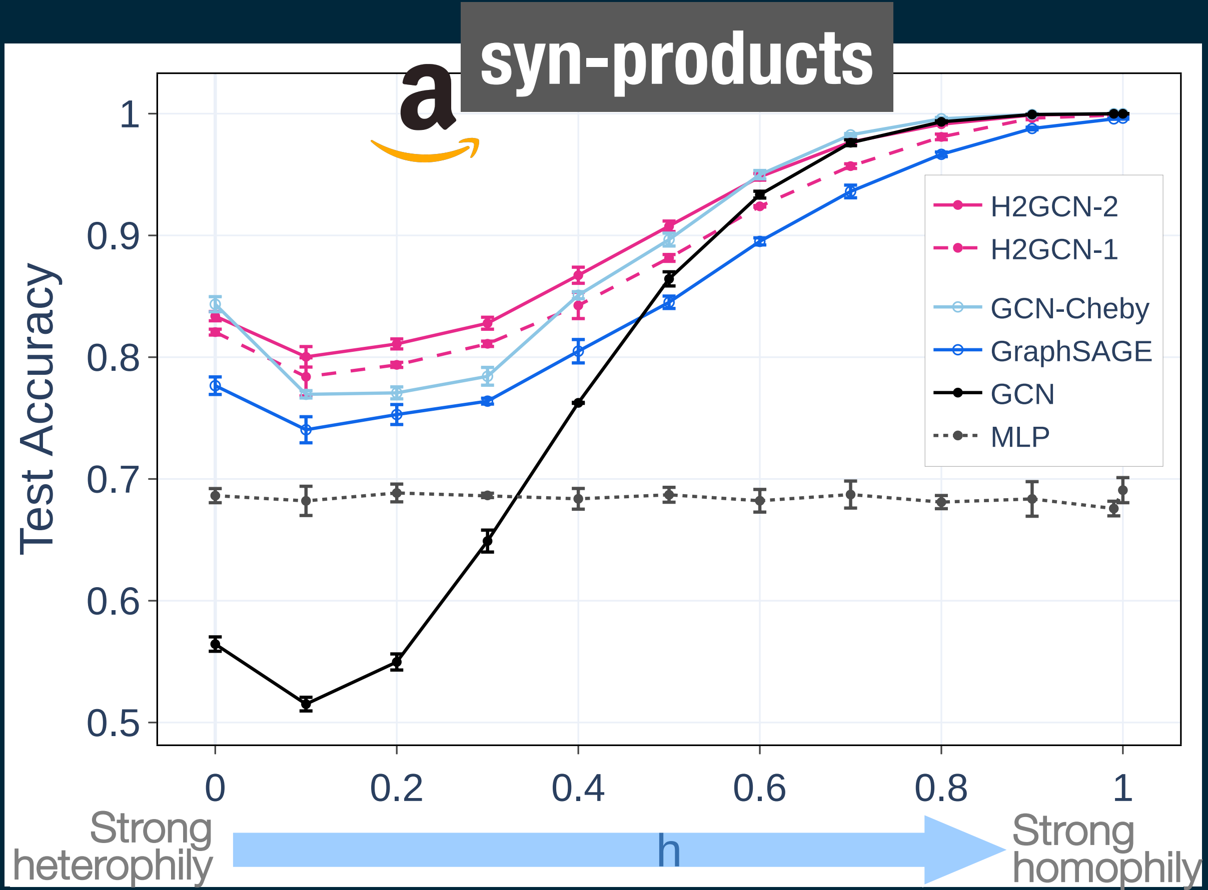
Results on Synthetic Benchmarks



[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

Results on Synthetic Benchmarks



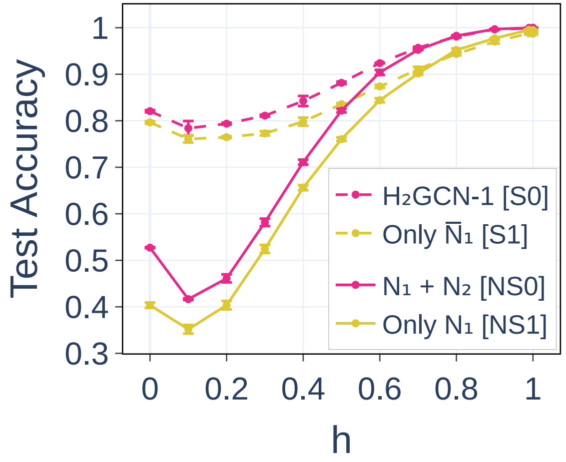


H_2GCN has the best trend overall, outperforming the baseline models in most heterophily settings, while tying with other models in homophily.

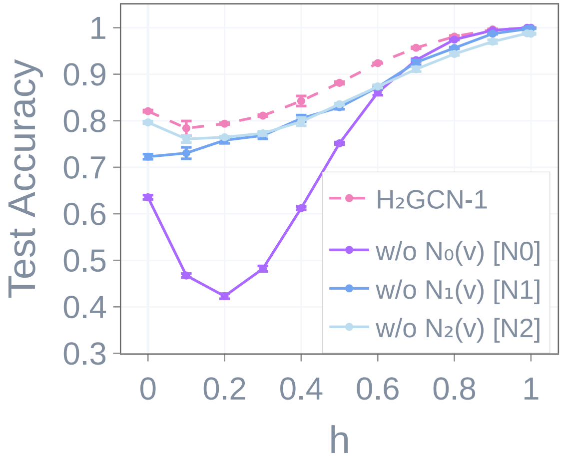
[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020] [Hu, Fey, et al. OGB '20]

Significance of Designs D1-D3

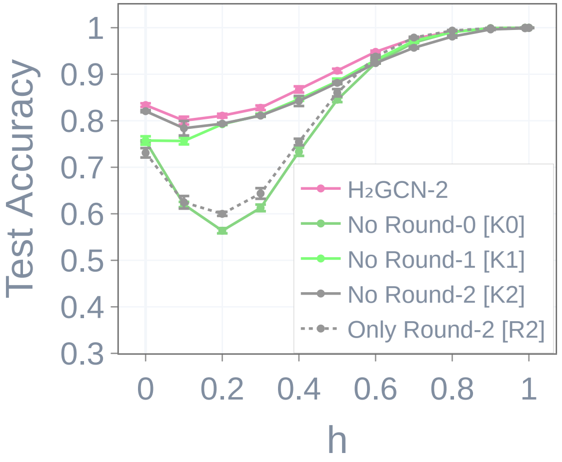
syn-products



Design **D1**:
embedding separation



Design **D2**:
Higher-order neighborhoods



Design **D3**:
Intermediate representations

Separating the embeddings leads to +40% acc for heterophily.

The H$_2$GCN variants that incorporate the designs D1-D3 significantly outperform the other variants, especially for low homophily settings.

M CSE GEMS LAB [Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

Results on Real Benchmarks

	Texas	Wisconsin	Actor	Squirrel	Chameleon	Cornell	Cora Full	Citeseer	Pubmed	Cora	Avg Rank
Hom. ratio h	**0.11**	**0.21**	**0.22**	**0.22**	**0.23**	**0.3**	**0.57**	**0.74**	**0.8**	**0.81**	
#Nodes $\lvert\mathcal{V}\rvert$	183	251	7,600	5,201	2,277	183	19,793	3,327	19,717	2,708	
#Edges $\lvert\mathcal{E}\rvert$	295	466	26,752	198,493	31,421	280	63,421	4,676	44,327	5,278	
#Classes $\lvert\mathcal{Y}\rvert$	5	5	5	5	5	5	70	7	3	6	
H_2GCN-1											
H_2GCN-2											
GraphSAGE											
GCN-Cheby											
MixHop											
GCN											
GAT*											
GEOM-GCN*											
MLP											

Strong heterophily

Strong homophily

[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

Results on Real Benchmarks

	Texas	Wisconsin	Actor	Squirrel	Chameleon	Cornell	Cora Full	Citeseer	Pubmed	Cora	Avg Rank
Hom. ratio h	**0.11**	**0.21**	**0.22**	**0.22**	**0.23**	**0.3**	**0.57**	**0.74**	**0.8**	**0.81**	
#Nodes $\lvert\mathcal{V}\rvert$	183	251	7,600	5,201	2,277	183	19,793	3,327	19,717	2,708	
#Edges $\lvert\mathcal{E}\rvert$	295	466	26,752	198,493	31,421	280	63,421	4,676	44,327	5,278	
#Classes $\lvert\mathcal{Y}\rvert$	5	5	5	5	5	5	70	7	3	6	
H_2GCN-1	83.24±7.07	84.31±3.70	34.31±1.31	28.98±1.97	52.96±2.09	78.11±6.68	67.49±0.78	76.72±1.50	88.50±0.64	86.34±1.56	3.7
H_2GCN-2	80.00±6.77	83.14±4.26	34.49±1.63	32.33±1.94	58.38±1.76	79.46±4.80	68.58±0.34	76.67±1.39	88.34±0.68	87.67±1.42	2.9
GraphSAGE	82.70±5.87	81.76±5.55	34.37±1.30	41.05±1.08	58.71±2.30	75.95±5.17	65.80±0.59	75.61±1.57	88.01±0.77	86.60±1.82	3.8
GCN-Cheby	78.65±5.76	77.45±4.83	33.80±0.83	40.86±1.49	63.38±1.37	71.35±9.89	67.14±0.58	76.25±1.76	88.08±0.52	86.86±0.96	3.9
MixHop	74.59±8.94	71.96±3.70	25.43±1.93	29.08±3.76	46.10±4.71	67.84±9.40	58.77±0.60	70.75±2.95	80.75±2.29	83.10±2.03	7.5
GCN	59.46±5.25	59.80±6.99	30.09±1.00	36.68±1.65	60.26±2.42	57.03±4.67	67.81±0.50	76.41±1.63	87.30±0.68	87.24±1.24	5.3
GAT*	58.38	49.41	28.45	30.03	42.93	54.32	N/A	74.32	87.62	86.37	7.6
GEOM-GCN*	67.57	64.12	31.63	38.14	60.90	60.81	N/A	77.99	90.05	85.27	4.6
MLP	81.08±5.41	84.12±2.69	**35.53±1.23**	29.29±1.40	46.51±2.53	**80.81±6.91**	58.53±0.46	72.36±2.01	86.63±0.38	74.61±1.97	5.3

- H_2GCN variants have consistently strong performance across the full spectrum.
- Other models that use some of the designs D1-D3 (e.g., GraphSAGE, GCN-Cheby) also perform significantly better than models that lack these designs.

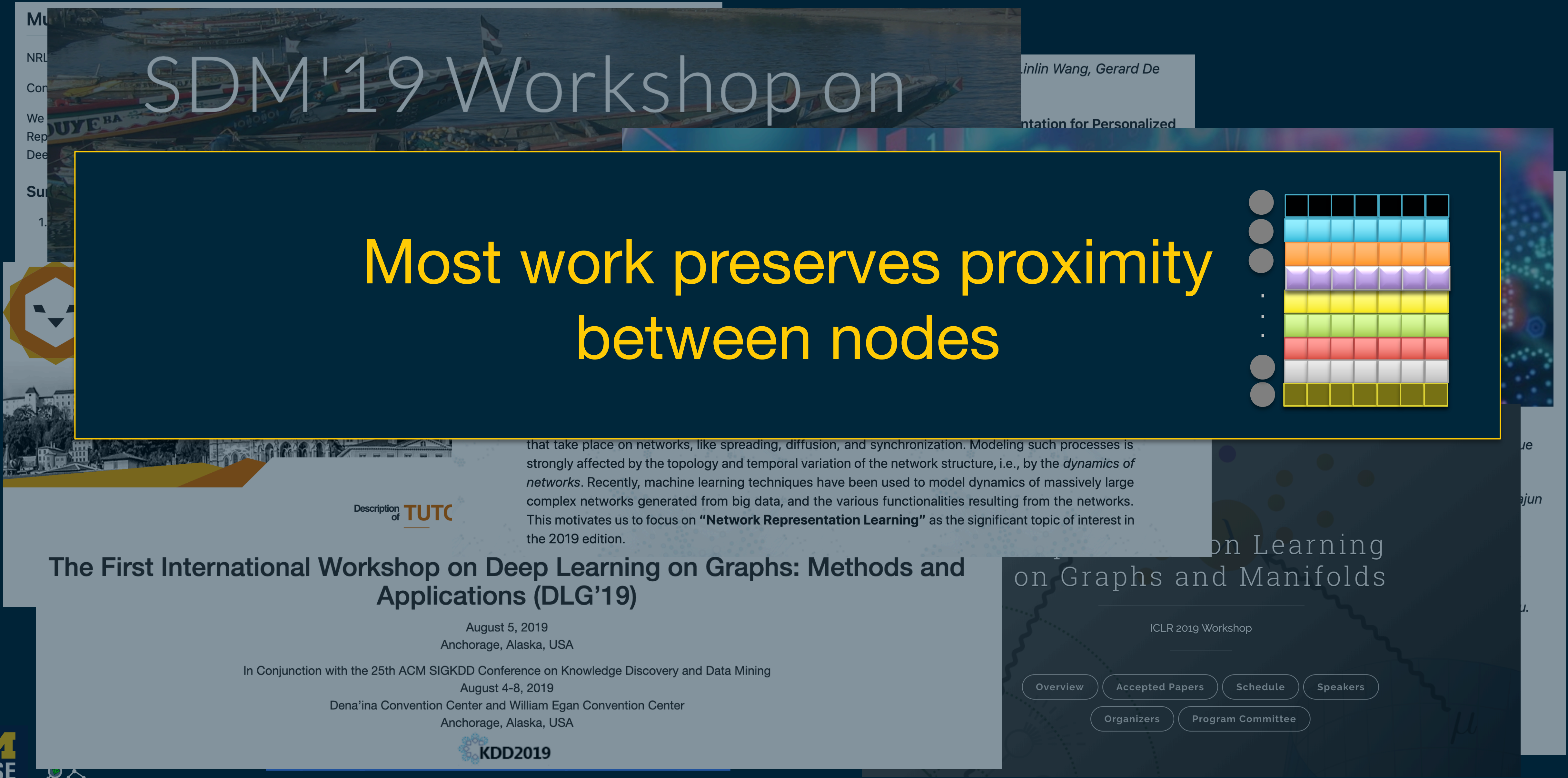
[Jiong Zhu, Yujun Yan, et al. arxiv: 2006.11468, 2020]

This talk

- Generalizing GNNs beyond homophily [Arxiv'20]

- Node embeddings: beyond proximity [ACM TKDD'20 +]

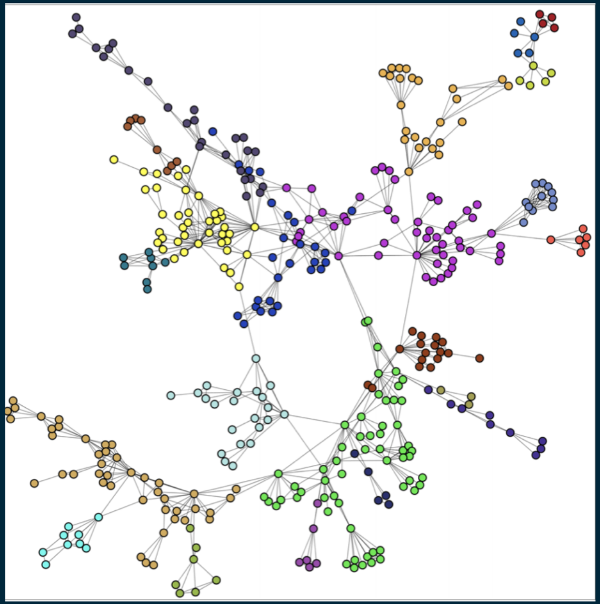
A lot of work on network representation learning!

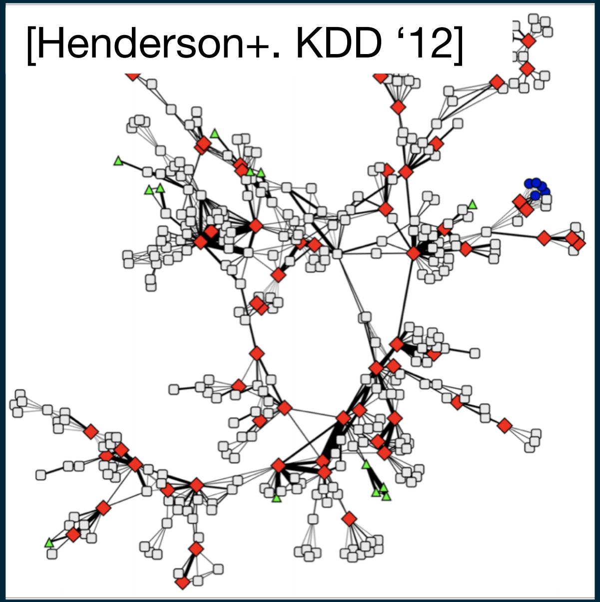


Most work preserves proximity between nodes

M CSE

Proximity vs. Structural Similarity





Find similar nodes in the same part of the network (communities)

Useful for link prediction, clustering, classification assuming homophily

[Perozzi+ ’14; Grover+ ’16; Tang+ ’15; …]

Find nodes with similar roles all over the network

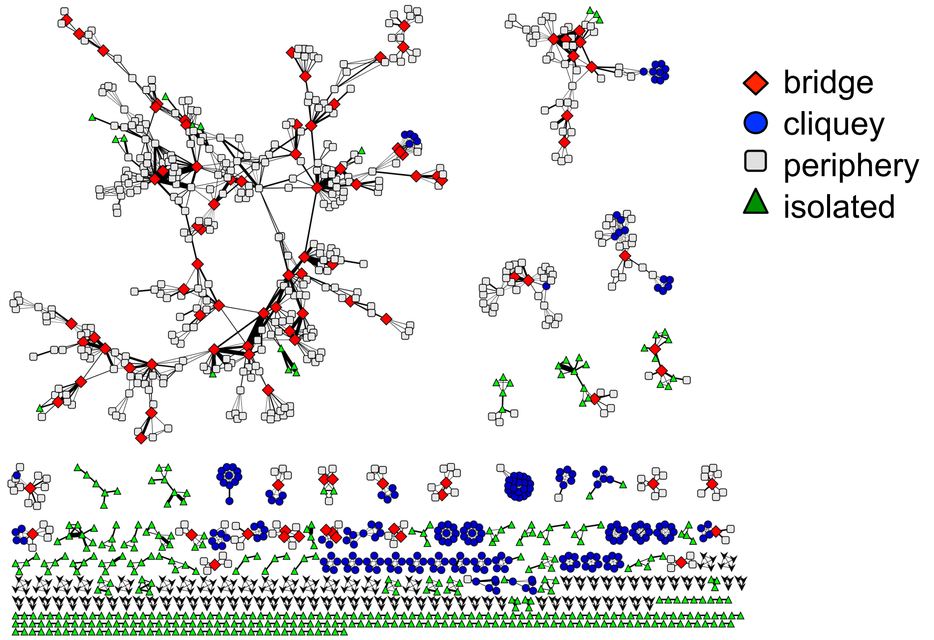
Useful for role-based classification, transfer learning, …

[Ribeiro+ ‘17; Donnat+ ’18, ..]

M CSE GEMS LAB

What are roles?

- The ways in which nodes / entities / actors relate to each other
- "The behavior expected of a node occupying a specific position" [Homans '67]
  - ✧ e.g., centers of stars
  - ✧ members of cliques
  - ✧ peripheral nodes
- Equivalence class: collection of nodes with the same role



Network Science Co-authorship Graph
[Newman 2006]

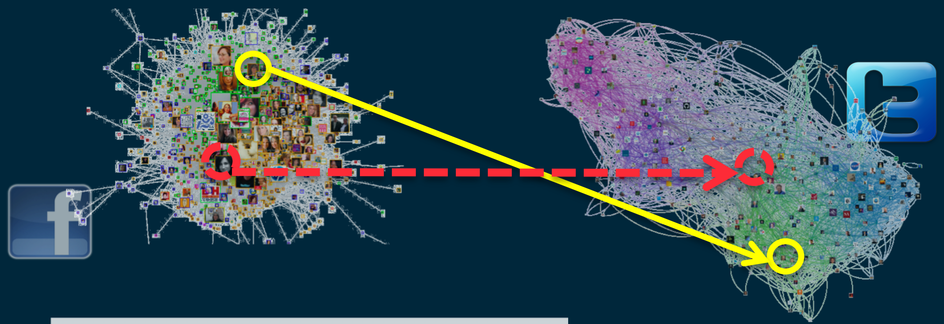
Relevant Sociology Literature

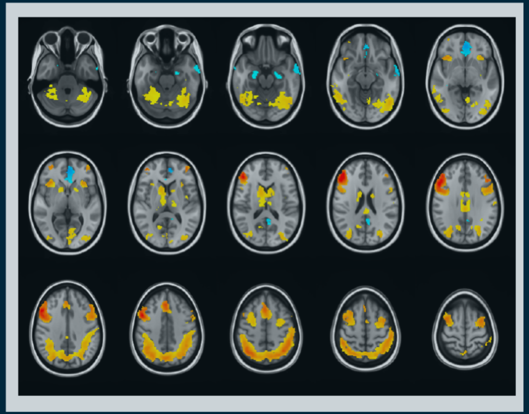
- S.P. Borgatti and M.G. Everett. 1992. Notions of position in social network analysis. Sociological methodology22, 1 (1992)
- Stephen P Borgatti, Martin G Everett, and Jeffrey C Johnson. 2018. Analyzing social networks. Sage
- F. Lorrain and H.C. White. 1971. Structural equivalence of individuals in social networks. Journal of Mathematical Sociology
- S. Boorman, H.C. White: Social Structure from Multiple Networks: II. Role Structures. American Journal of Sociology, 81:1384-1446, 1976.
- R.S. Burt: Positions in Networks. Social Forces, 55:93-122, 1976.
- M.G. Everett, S. P. Borgatti: Regular Equivalence: General Theory. Journal of Mathematical Sociology, 19(1):29-52, 1994.
- K. Faust, A.K. Romney: Does Structure Find Structure? A critique of Burt's Use of Distance as a Measure of Structural Equivalence. Social Networks, 7:77-103, 1985.
- K. Faust, S. Wasserman: Blockmodels: Interpretation and Evaluation. Social Networks, 14:5–61. 1992.
- R.A. Hanneman, M. Riddle: Introduction to Social Network Methods. University of California, Riverside, 2005.
- L.D. Sailer: Structural Equivalence: Meaning and Definition, Computation, and Applications. Social Networks, 1:73-90, 1978.
- M.K. Sparrow: A Linear Algorithm for Computing Automorphic Equivalence Classes: The Numerical Signatures Approach. Social Networks, 15:151-170, 1993.
- S. Wasserman, K. Faust: Social Network Analysis: Methods and Applications. Cambridge University Press, 1994.
- H.C. White, S. A. Boorman, R. L. Breiger: Social Structure from Multiple Networks I. Blockmodels of Roles and Positions. American Journal of Sociology, 81:730-780, 1976.
- D.R. White, K. Reitz: Graph and Semi-Group Homomorphism on Networks and Relations. Social Networks, 5:143-234, 1983.
- …

Sometimes structural similarity is more appropriate than proximity

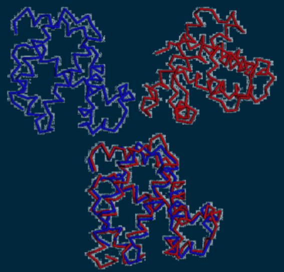
Multiple networks

Alignment or matching [CIKM’18]





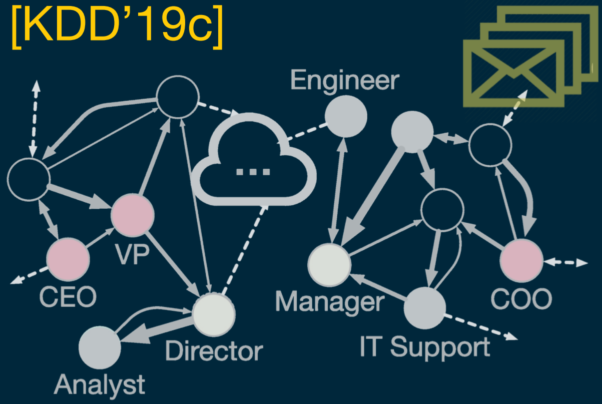
Transfer learning



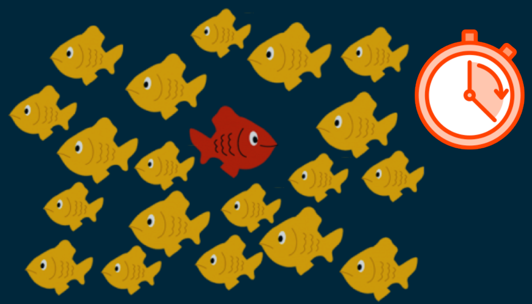
Graph comparison / classification [KDD’19a; ICDM’19a]

Single network

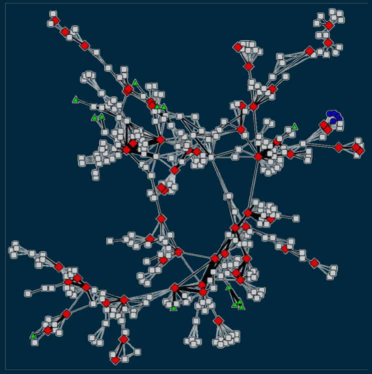
Node classification [KDD’19c]



Anomaly detection [KDD’19b]

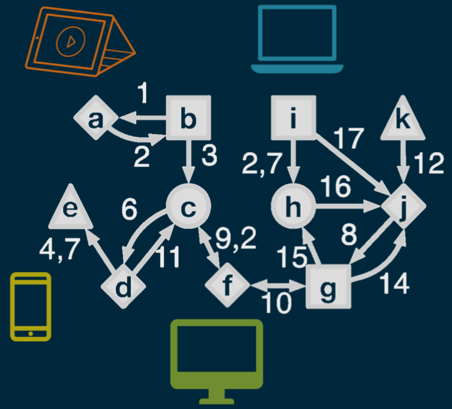


Role query



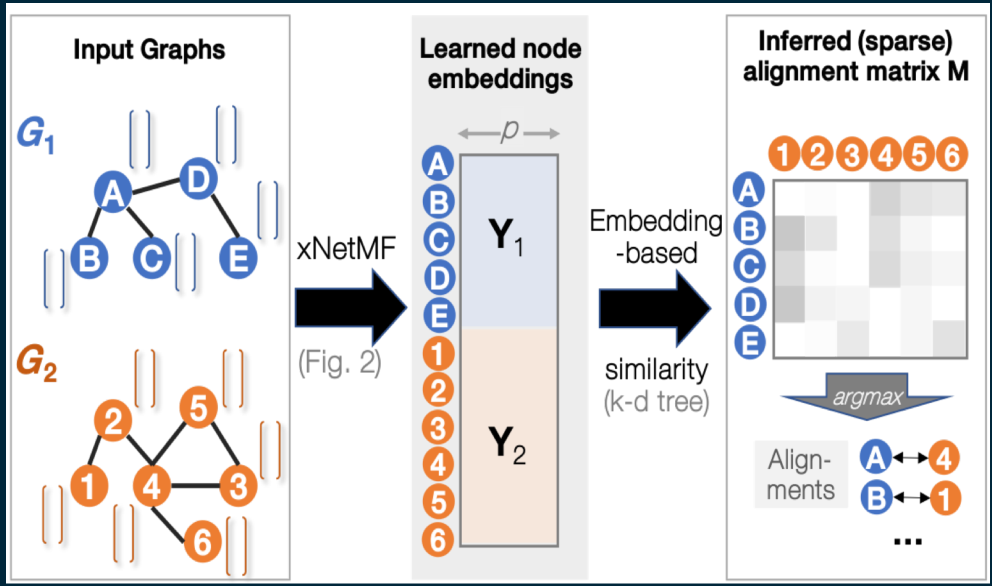
[PKDD’19]

Identity resolution



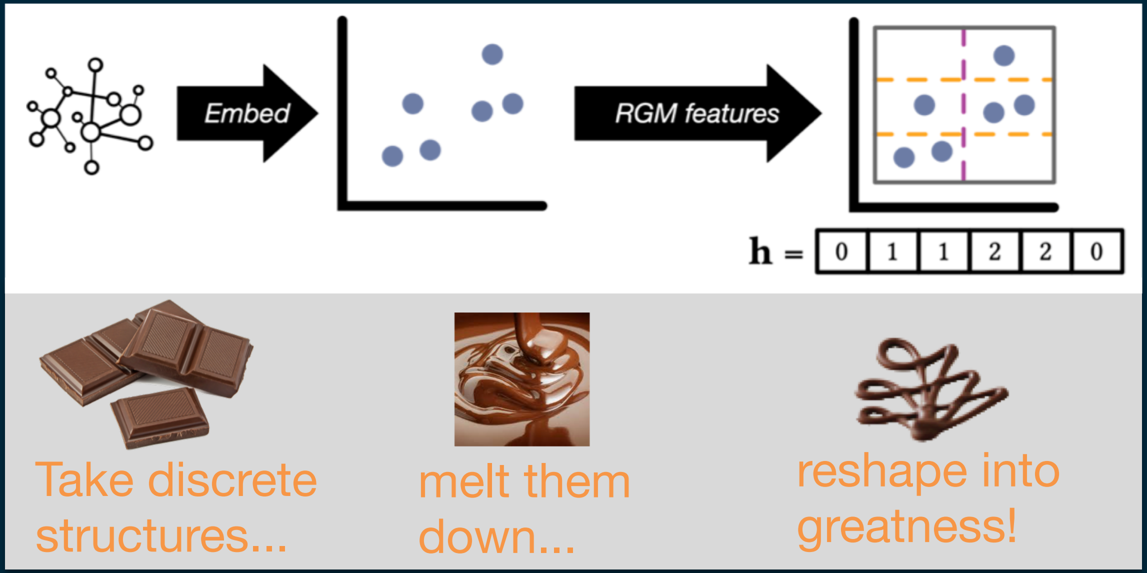
Embedding-based Collective Network Mining


Structural embeddings for network alignment



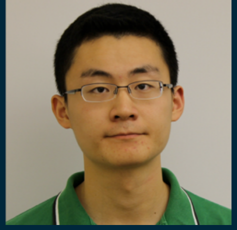
[Mark Heimann, Haoming Shen, Tara Safavi, Danai Koutra. ACM CIKM'18]

Distribution of node embeddings as multiresolution features for graph classification



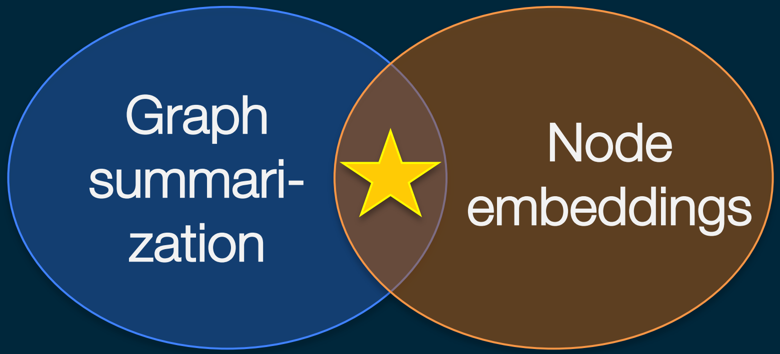
[Mark Heimann, Tara Safavi, Danai Koutra. IEEE ICDM'19]

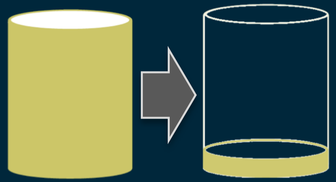
Embedding-based Single Network Mining



Latent network summarization

- Find a compressed representation that captures the key structural information:
  - ✧ independent of graph size (|V|, |E|), and
  - ✧ capable of deriving node representations on the fly

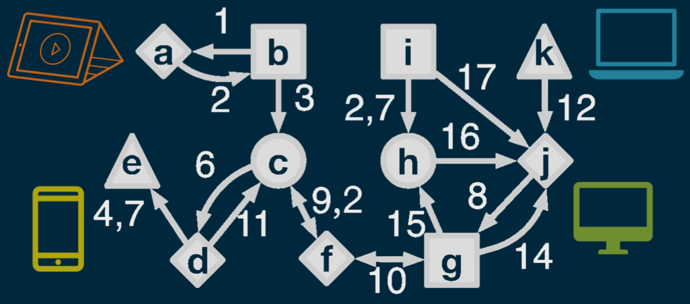




https://github.com/GemsLab/MultiLENS

Sparse hash-based embeddings

- Learn a function $\chi: V \rightarrow \{0,1\}^d$ s.t. the derived d-dim embeddings
  - ✧ preserve similarities in interactions
  - ✧ accurately capture temporal information in the input heterogeneous network G(V, E)



https://github.com/GemsLab/node2bits

[Di Jin, Rossi et al. ACM KDD'19]

[Di Jin, Mark Heimann, et al. PKDD'19]

ACM TKDD'20

http://tinyurl.com/proximity-role-emb

On Proximity and Structural Role-based Embeddings in Networks: Misconceptions, Techniques, and Applications

RYAN A. ROSSI, Adobe Research, USA
DI JIN, University of Michigan, USA
SUNGCHUL KIM, Adobe Research, USA
NESREEN K. AHMED, Intel Labs, USA
DANAI KOUTRA, University of Michigan, USA
JOHN BOAZ LEE, Worcester Polytechnic Institute, USA

Structural roles define sets of structurally similar nodes that are more similar to nodes inside the set than outside, whereas communities define sets of nodes with more connections inside the set than outside. Roles based on structural similarity and communities based on proximity are fundamentally different but important complementary notions. Recently, the notion of structural roles has become increasingly important and has gained a lot of attention due to the proliferation of work on learning representations (node/edge embeddings) from graphs that preserve the notion of roles. Unfortunately, recent work has sometimes confused the notion of structural roles and communities (based on proximity) leading to misleading or incorrect claims about the capabilities of network embedding methods. As such, this paper seeks to clarify the misconceptions and key differences between structural roles and communities, and formalize the general mechanisms (*e.g.*, random walks, feature diffusion) that give rise to community or role-based structural embeddings. We theoretically prove that embedding methods based on these mechanisms result in either community or role-based structural embeddings. These mechanisms are typically easy to identify and can help researchers quickly determine whether a method preserves community or role-based embeddings. Furthermore, they also serve as a basis for developing new and improved methods for community or role-based structural embeddings. Finally, we analyze and discuss applications and data characteristics where community or role-based embeddings are most appropriate.

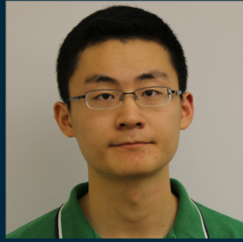
Mechanisms that lead to proximity- and structural role-based embeddings

Embedding Type	General Mechanism	Examples of Methods
COMMUNITY-BASED (Section 4)	**Random Walks** (Sec. 4.1)	Spectral embedding [Chung 1997] deepwalk [Perozzi et al. 2014] node2vec [Grover and Leskovec 2016] LINE [Tang et al. 2015] GraRep [Cao et al. 2015] ComE+ [Cavallari et al. 2019]
	Feature Prop./Diffusion (Sec. 4.2)	GCN [Kipf and Welling 2017] GraphSage [Hamilton et al. 2017] MultiLENS [Jin et al. 2019c]
ROLE-BASED (Section 5)	**Graphlets** (Sec. 5.1)	deepGL [Rossi et al. 2017] MCN [Lee et al. 2018b] HONE [Rossi et al. 2018b]
	Feature-based Walks (Sec. 5.2)	role2vec [Ahmed et al. 2018] node2bits [Jin et al. 2019a] SimSum [Liu et al. 2018b, 2019]
	Feature-based MF (Sec. 5.3)	rolX [Henderson et al. 2012] HERO [Ahmed et al. 2017b] EMBER [Jin et al. 2019b]

[Rossi, Di Jin, et al. ACM TKDD'20]

Empirical Study of Role-based Embedding Methods





STRUCTURAL Equivalence

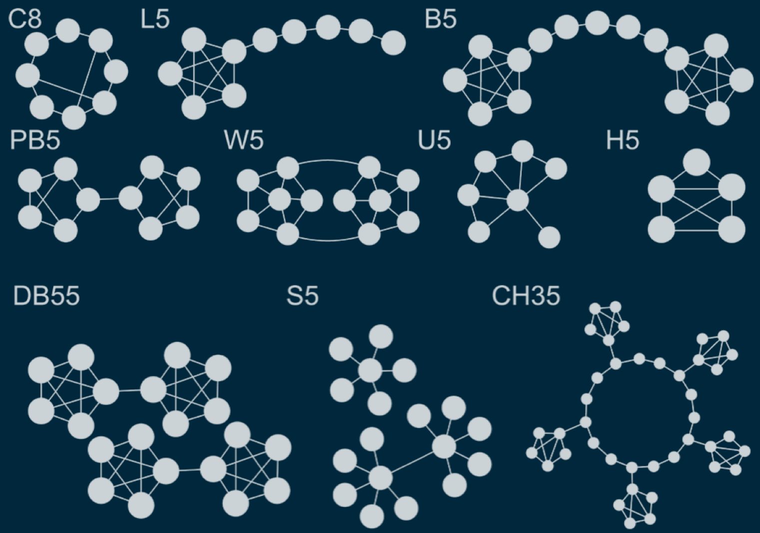
Identical relationships to all other nodes

AUTOMORPHIC Equivalence

Structure-preserving mapping between nodes

REGULAR Equivalence

Equivalent relationships to equivalent other nodes



Synthetic Datasets

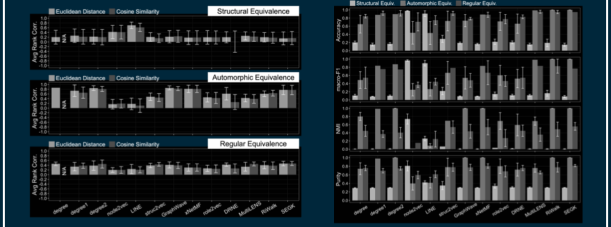
Air Traffic
Protein
Blog
Facebook
Email
...

Real Datasets

- node2vec
- LINE
- struc2vec
- GraphWave
- xNetMF
- role2vec
- DRNE
- MultiLENS
- RiWalk
- SEGK

Structural Embedding Methods

INTRINSIC EXTRINSIC



Evaluation

Empirical Study of Role-based Embedding Methods

STRUCTURAL Equivalence
Identical relationsh...
nodes

AUTOMOR... Equivalence
Structure-preservi...
between nodes

REGULAR Equivalence
Equivalent relationships to equivalent other nodes

C8 L5 B5

node2vec
LINE
GraphWave
role2vec
MultiLENS
SEGK
...Embedding ...hods

Coming Soon!

Python package for structural role-based embeddings + evaluation routines

https://github.com/GemsLab

Air Traffic
Protein
Blog
Facebook
Email
...
Real Datasets

EXTRINSIC

Evaluation

[Mark Jin, Mark Heimann, Di Jin, Danai Koutra. KDD MLG'20]

Take-away messages

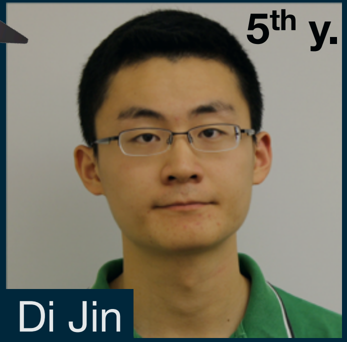
- Leveraging distinct representations (at different levels) in GNNs can help handle challenging heterophily settings [Arxiv '20]
  - ✧ Many future directions to be explored
  - ✧ Need for larger, more diverse datasets with heterophily (OGB effort?)
- Structural embeddings are less studied, but are more appropriate than proximity-based embeddings in several tasks [TKDD '20; MLG '20; …]
  - ✧ Different embedding mechanisms give rise to communities and roles
  - ✧ There are some misconceptions in the literature about the types of equivalences that structural embeddings capture

Talk based on the following papers

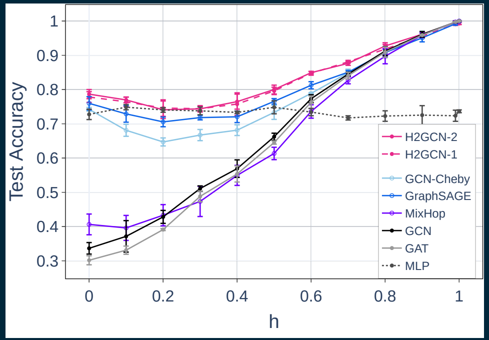
- Mark Heimann, Haoming Shen, Tara Safavi, Danai Koutra. REGAL: Representation Learning-based Graph Alignment. ACM CIKM'18.
- Yujun Yan, J. Zhu, Marlena Duda, Eric Solarz, Chandra Sripada, Danai Koutra. GroupINN: Grouping-based Interpretable Neural Network-based Classification of Limited, Noisy Brain Data. ACM KDD'19a.
- Di Jin, R. Rossi, Eunyee Koh, Sungchul Kim, Anup. Rao, Danai Koutra. Latent Network Summarization: Bridging Network Embedding and Summarization. ACM KDD'19b.
- D. Jin*, Mark Heimann*, Tara Safavi, Mengdi Wang, Wei Lee, Lindsay Snider, Danai Koutra. Smart Roles: Inferring Professional Roles in Email Networks. ACM KDD'19c.
- D. Jin, Mark Heimann, Ryan Rossi, Danai Koutra. node2bits: Compact Time- and Attribute-aware Node Representations for User Stitching. ECML/PKDD'19.
- Mark Heimann, Tara Safavi, Danai Koutra. Distribution of Node Embeddings as Multiresolution Features for Graphs. IEEE ICDM 2019. [best student paper award]
- Ryan A. Rossi, Di Jin, Sungchul Kim, Nesreen K. Ahmed, Danai Koutra, John Boaz Lee. On Proximity and Structural Role-based Embeddings in Networks: Misconceptions, Techniques, and Applications. ACM TKDD 2020.
- Mark Jin, Mark Heimann, Di Jin, Danai Koutra. Understanding and Evaluating Structural Node Embeddings. ACM KDD MLG workshop 2020.
- Jiong Zhu, Yujun Yan, Lingxiao Zhao, Mark Heimann, Leman Akoglu, Danai Koutra. Generalizing Graph Neural Networks Beyond Homophily. arxiv.org/abs/2006.11468, 2020.

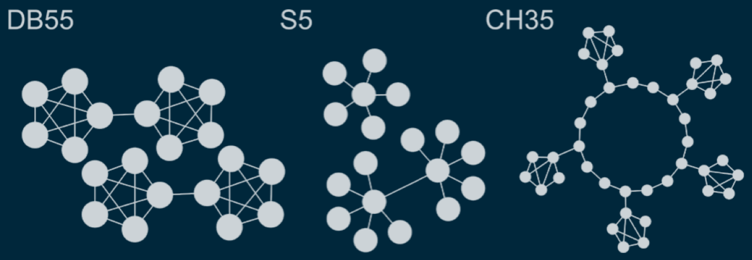
Thank you! Questions?

Danai Koutra
dkoutra@umich.edu
GEMS LAB

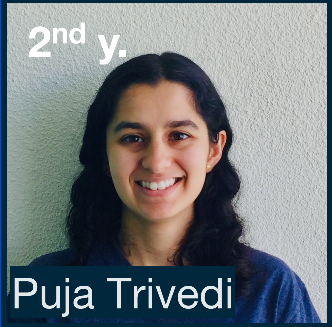


Representation Learning Beyond Homophily & Proximity





</> https://github.com/GemsLab



NSF ARL P&G MIDAS amazon Adobe Microsoft Azure