COLLEGE OF ENGINEERING
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

MEDICAL SCHOOL
UNIVERSITY OF MICHIGAN

GEMS LAB

The Power of Summarization in Network Representation Learning

Danai Koutra

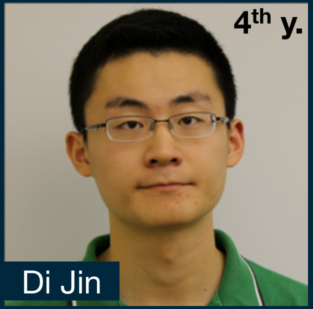
Assistant Professor, CSE

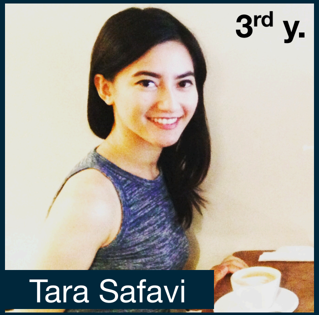
Computational Medicine and Bioinformatics (courtesy)

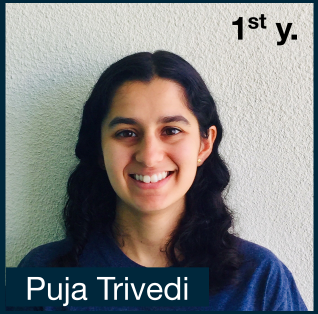
Joint work with: Caleb Belth, Christos Faloutsos, Brian Gallagher, Aram Galstyan, Mark Heimann, Di Jin, Yike Liu, Ryan Rossi, Tara Safavi, Neil Shah, Chandra Sripada, Pedro Szekely, Yujun Yan, Linhong Zhu, …

Great Lakes Workshop on Data Science – September 20-22, 2019

Slides at: https://bit.ly/2m2tIJo







++

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 indispensible. 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 c
fields like linear algebra, distributed systems, deep learning, and even neuroscie
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 Mich
for Data Science (MIDAS), Microsoft Azure, the National Science Foundation (NSF

Interested?

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

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

News

August 2019
2 papers accepted at ICDM

June 2019
1 paper accepted at ASONAM

June 2019
1 paper accepted at PKDD

May 2019
Welcome new PhDs!

May 2019
Tara passes her prelim

April 2019
3 papers accepted to KDD 2019

March 2019
Danai receives NSF CAREER award

January 2019
Danai awarded an Amazon research grant

December 2018

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

A lot of work on network representation learning!

Must-read papers on NRL/NE.

NRL: network representation learning. NE: network embedding.

Contributed by Cunchao Tu, Yuan Yao and Zhengyan Zhang.

We release OpenNE, an open source toolkit for NE/NRL. This repository p
Representation Learning) training and testing framework. Currently, the
DeepWalk, LINE, node2vec, GraRep, TADW and GCN.

Survey papers:

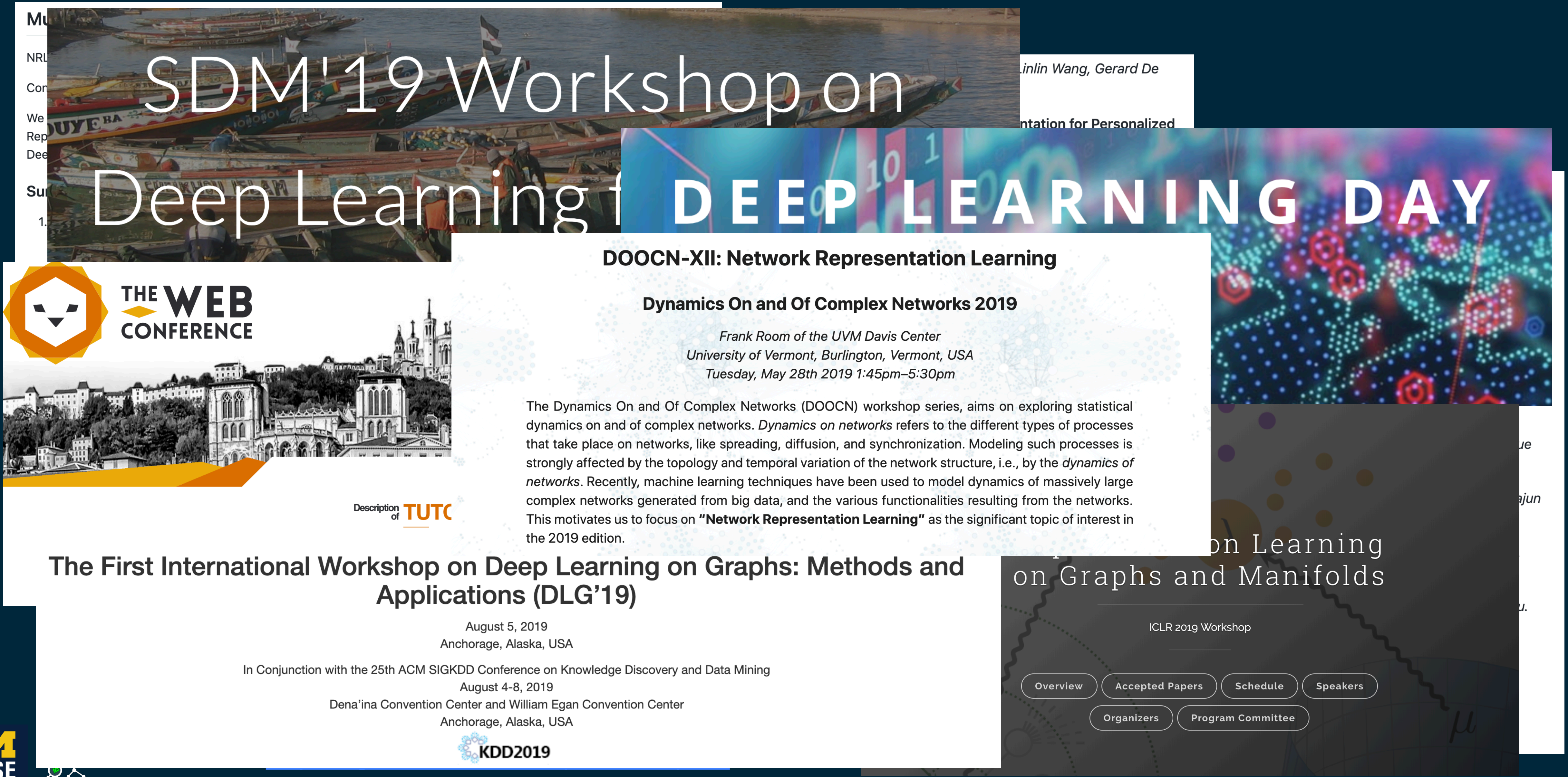
1. **Representation Learning on Graphs: Methods and Applications.** *W*
   2017. paper
2. **Graph Embedding Techniques, Applications, and Performance: A**
3. **A Comprehensive Survey of Graph Embedding: Problems, Technic**
   Zheng, Kevin Chen-Chuan Chang. 2017. paper
4. **Network Representation Learning: A Survey.** *Daokun Zhang, Jie Yi*
5. **A Tutorial on Network Embeddings.** *Haochen Chen, Bryan Perozzi,*
6. **Network Representation Learning: An Overview.(In Chinese)** *Cunc*
   2017. paper
7. ***Relational inductive biases, deep learning, and graph networks.*** *P*
   Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malino
   Santoro, Ryan Faulkner, Caglar Gulcehre, Francis Song, Andrew Balla
   Kelsey Allen, Charles Nash, Victoria Langston, Chris Dyer, Nicolas He
   Botvinick, Oriol Vinyals, Yujia Li, Razvan Pascanu. 2018. paper

54. **Link Prediction via Subgraph Embedding-Based Convex Matrix Completion.** *Zhu Cao, Linlin Wang, Gerard De melo.* AAAI 2018.
55. **Generative Adversarial Network based Heterogeneous Bibliographic Network Representation for Personalized Citation Recommendation.** *J. Han, Xiaoyan Cai, Libin Yang.* AAAI 2018.
56. **DepthLGP: Learning Embeddings of Out-of-Sample**
    Zhu. AAAI 2018. paper
57. **Structural Deep Embedding for Hyper-Networks.** *Ke*
    paper
58. **TIMERS: Error-Bounded SVD Restart on Dynamic Ne**
    Zhu. AAAI 2018. paper
59. **Community Detection in Attributed Graphs: An Emb**
    Zhang. AAAI 2018.
60. **Bernoulli Embeddings for Graphs.** *Vinith Misra, Sumi*
61. **Distance-aware DAG Embedding for Proximity Searc**
    Zhou Zhao, Fanwei Zhu, Kevin Chen-Chuan Chang, M
62. **GraphGAN: Graph Representation Learning with Ge**
    Wang, MIAO ZHAO, Weinan Zhang, Fuzheng Zhang, X
63. **HARP: Hierarchical Representation Learning for Net**
    AAAI 2018. paper code
64. **Representation Learning for Scale-free Networks.** *R*
    2018. paper
65. **Social Rank Regulated Large-scale Network Embedd**
    2018. paper

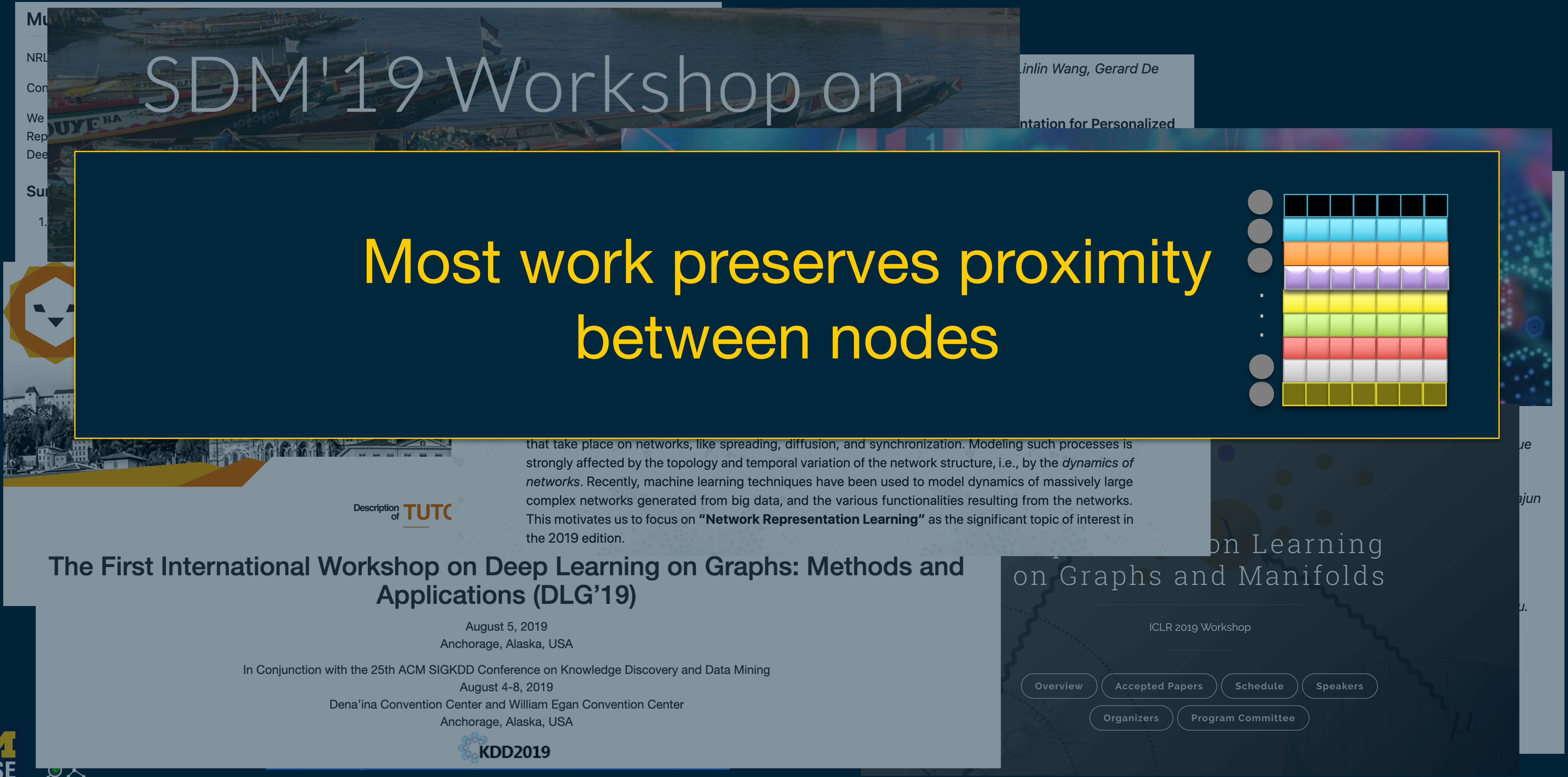
101. **Integrative Network Embedding via Deep Joint Reconstruction.** *Di Jin, Meng Ge, Liang Yang, Dongxiao He, Longbiao Wang, Weixiong Zhang.* IJCAI 2018.
102. **Scalable Multiplex Network Embedding.** *Hongming Zhang, Liwei Qiu, Lingling Yi, Yangqiu Song.* IJCAI 2018. paper
103. **Adversarially Regularized Graph Autoencoder for Graph Embedding.** *Shirui Pan, Ruiqi Hu, Guodong Long, Jing Jiang, Lina Yao, Chengqi Zhang.* IJCAI 2018.
104. **Dynamic Network Embedding : An Extended Approach for Skip-gram based Network Embedding.** *Lun Du, Yun Wang, Guojie Song, Zhicong Lu, Junshan Wang.* IJCAI 2018.
105. **Discrete Network Embedding.** *Xiaobo Shen, Shirui Pan, Weiwei Liu, Yew-Soon Ong, Quan-Sen Sun.* IJCAI 2018.
106. **Deep Attributed Network Embedding.** *Hongchang Gao, Heng Huang.* IJCAI 2018.
107. **Active Discriminative Network Representation Learning.** *Li Gao, Hong Yang, Chuan Zhou, Jia Wu, Shirui Pan, Yue Hu.* IJCAI 2018.
108. **ANRL: Attributed Network Representation Learning via Deep Neural Networks.** *Zhen Zhang, Hongxia Yang, Jiajun Bu, Sheng Zhou, Pinggang Yu, Jianwei Zhang, Martin Ester, Can Wang.* IJCAI 2018.
109. **Feature Hashing for Network Representation Learning.** *Qixiang Wang, Shanfeng Wang, Maoguo Gong, Yue Wu.* IJCAI 2018.
110. **Constructing Narrative Event Evolutionary Graph for Script Event Prediction.** *Zhongyang Li, Xiao Ding, Ting Liu.* IJCAI 2018. paper code
111. **Deep Inductive Network Representation Learning.** *Ryan A. Rossi, Rong Zhou, Nesreen K. Ahmed.* WWW 2018. paper
112. **A Unified Framework for Community Detection and Network Representation Learning.** *Cunchao Tu, Xiangkai Zeng, Hao Wang, Zhengyan Zhang, Zhiyuan Liu, Maosong Sun, Bo Zhang, Leyu Lin.* TKDE 2018. paper

https://github.com/thunlp/NRLPapers

A lot of work on network representation learning!

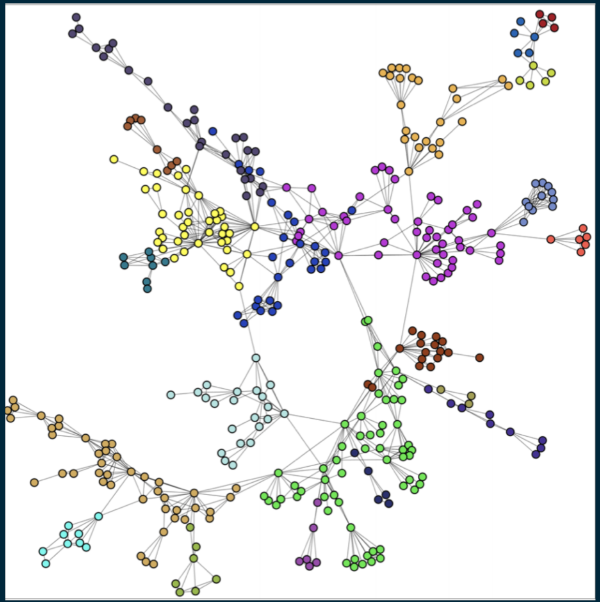


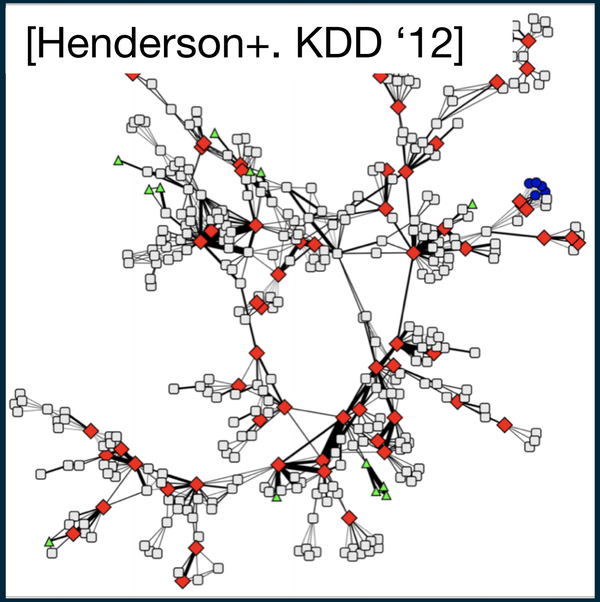
A lot of work on network representation learning!



Most work preserves proximity between nodes

Proximity vs. Structural Similarity





Find similar nodes in the same part of the network

Useful for link prediction, clustering, classification assuming homophily

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

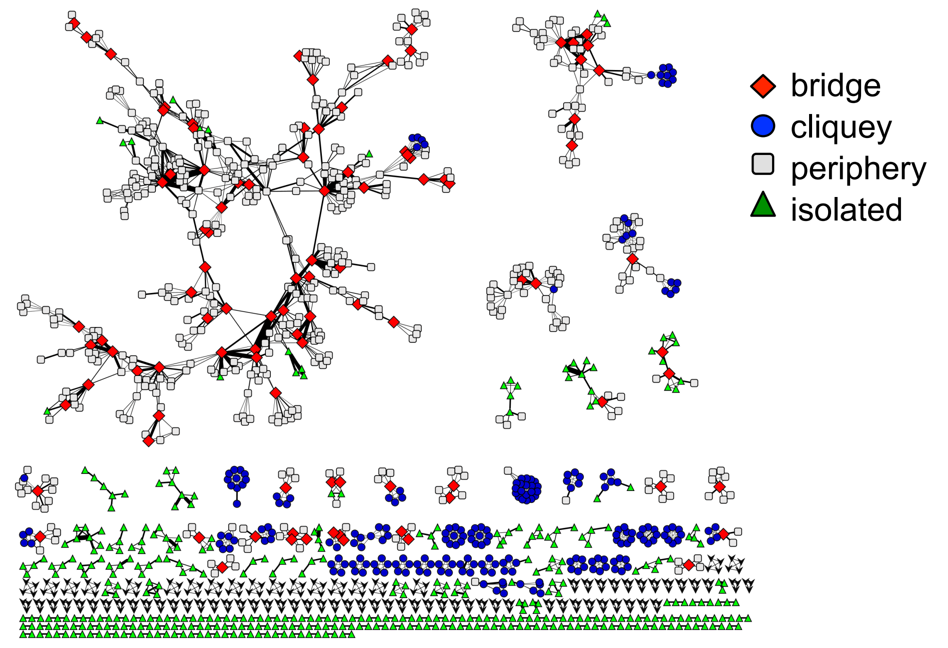
Find nodes with similar roles all over the network

Useful for role-based classification, transfer learning, …

[Ribeiro+ ‘17; Donnat+ ’18, ..]

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



Network Science Co-authorship Graph
[Newman 2006]

- Position or equivalence class:
  - ✧ collection of nodes with the same role

[Lorrain & White '71] [Borgatti & Everett '92] [Wasserman & Faust. '94] [Henderson et al. KDD'12]

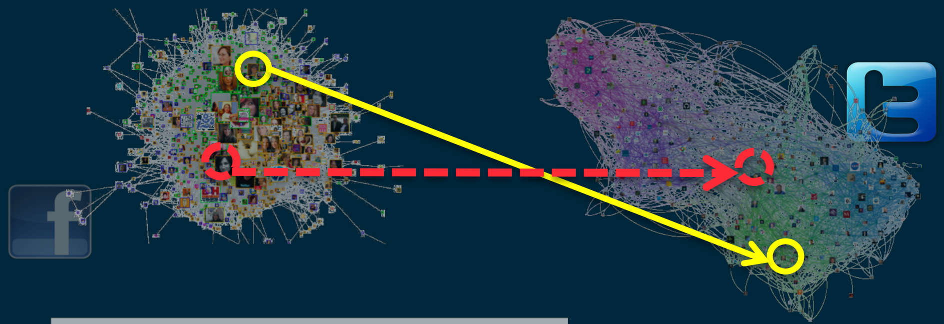
Relevant Sociology Literature

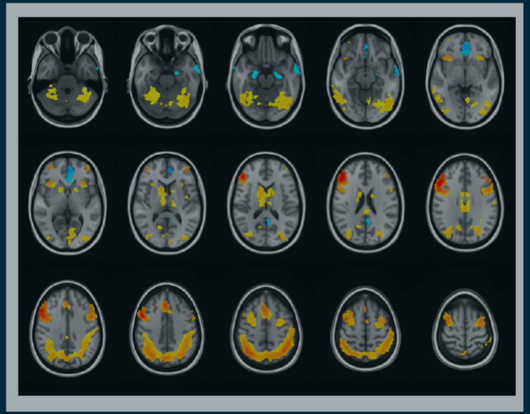
- S.P. Borgatti and M.G. Everett. 1992. Notions of position in social network analysis. Sociological methodology22, 1 (1992)
- Stephen P Borgatti, Martin G Everett, and Jeffrey C Johnson. 2018. Analyzing social networks. Sage
- F. Lorrain and H.C. White. 1971. Structural equivalence of individuals in 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

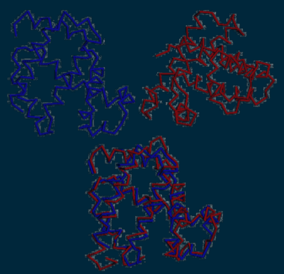
Multiple networks

Alignment or matching [CIKM'18]





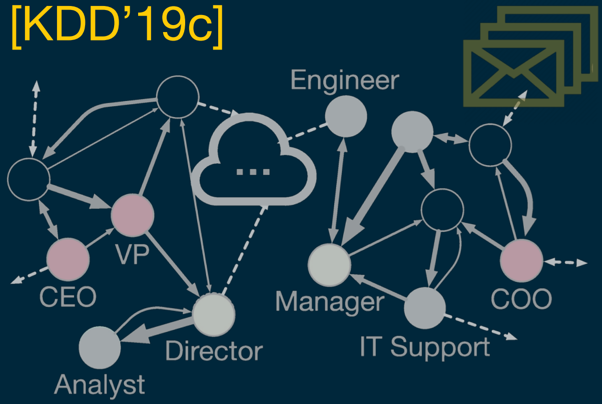
Transfer learning



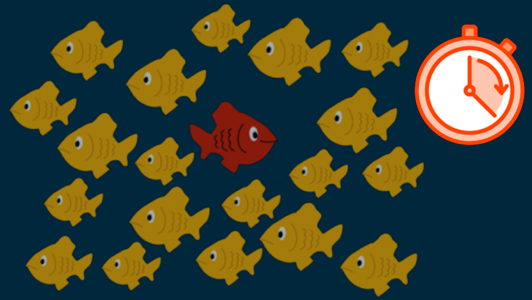
Graph comparison / classification [KDD'19a; ICDM'19a]

Single network

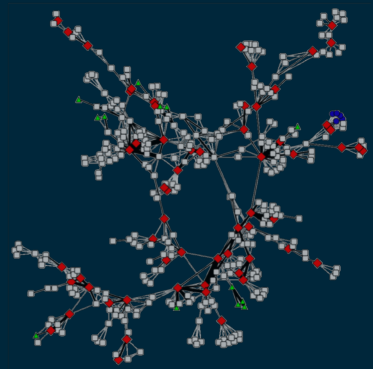
Node classification [KDD'19c]



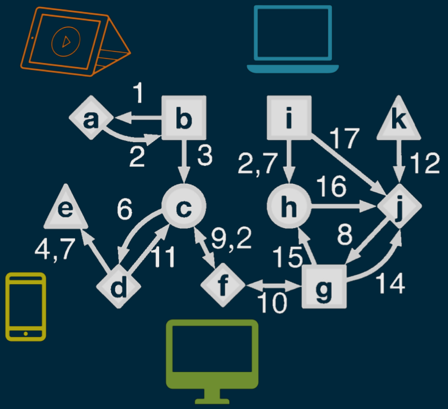
Anomaly detection [KDD'19b]



Role query



[PKDD'19]
Identity resolution



Sometimes structural similarity is more appropriate than proximity

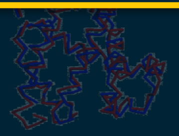
Alignment or matching

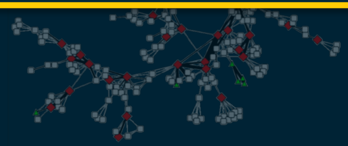
Node classification

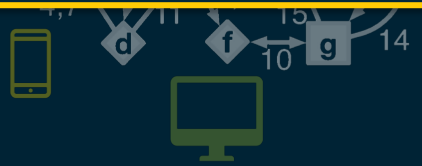
Anomaly detection

Multiple networks

Graph comparison / classification







From Community to Role-based Graph Embeddings

Ryan A. Rossi, Di Jin, Sungchul Kim, Nesreen K. Ahmed, Danai Koutra, John Boaz Lee

arXiv

https://arxiv.org/abs/1908.08572

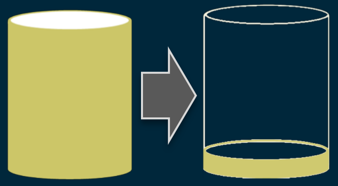
M CSE GEMS LAB

[Henderson+, RolX; KDD'12]

This talk: Summarization in Network Representation Learning

- Summarization within a GCN for faster training, data denoising and interpretability [ACM KDD’19a]

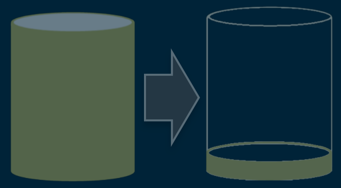
- Embedding summarization for compression and on-the-fly computation [ACM KDD’19b; PKDD’19]



This talk: Summarization in Network Representation Learning

- Summarization within a GCN for faster training, data denoising and interpretability [ACM KDD’19a]

- Embedding summarization for compression and on-the-fly computation [ACM KDD’19b; PKDD’19]



Interpretable NN-based Classification

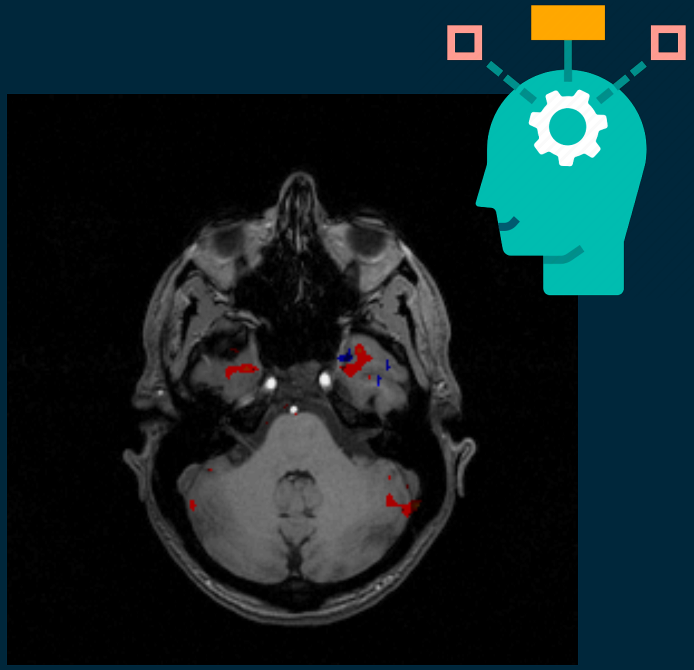
- Given a set of networks
  - ✧ each associated with a label

- Devise an efficient, interpretable, and parsimonious model
  - ✧ that can accurately predict and
  - ✧ explain each label

[Yujun Yan, Jiong Zhu, et al. ACM KDD '19]

Interpretable NN-based Classification

- Given a set of networks ⟹ Subjects with fMRI brain graphs
  - ✧ each associated with a label ⟹ phenotype



- Devise an efficient, interpretable, and parsimonious model
  - ✧ that can accurately predict and
  - ✧ explain each label ⟹ phenotype

Related Work

- Linear models (PCA, ICA, matrix factorization)
  + Denoising
  – Fail to capture non-linear interactions

- Neural-network models (different variants of GCN)
  + Able to model non-linear interactions
  – Need many training samples
  – Need many parameters
  – Long time for training
  – "Black" box

	Fast	Parsimonious	Interpretable
CNN (KDD'17), GraphCNN (NIPS'16)	✗	✗	✗
GCN (ICLR'17), DGCNN (AAAI'18)	✓	✗	✗
Diffpool (NIPS'18)	✓	✗	inadequate
GroupINN (proposed)	✓	✓	✓

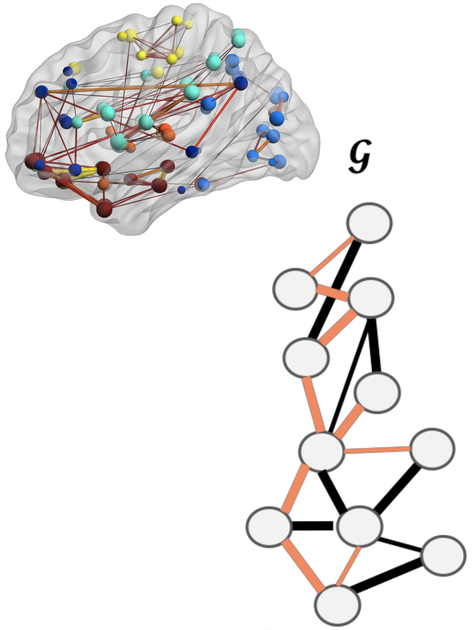
Related Work

- Linear models (PCA, ICA, matrix factorization)
  + Denoising
  –
- N
  +
  –
  –
  – Long time for training
  – "Black" box

			pretable
GraphCNN (NIPS'16)	X	X	X
GCN (ICLR'17), DGCNN (AAAI'18)	✓	X	X
Diffpool (NIPS'18)	✓	X	inadequate
GroupINN (proposed)	✓	✓	✓

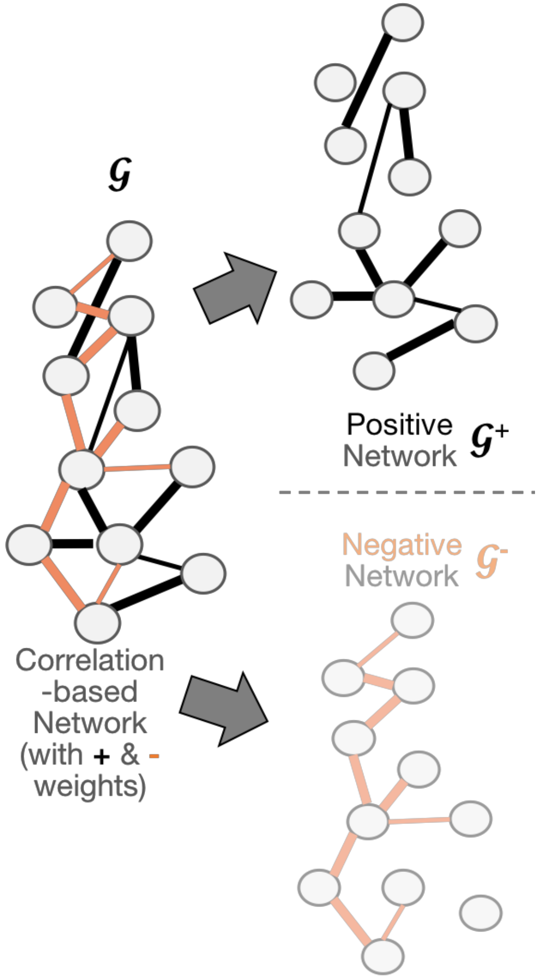
Can we build an interpretable NN-based model that is insensitive to noise, parsimonious and able to capture nonlinearities in the prediction task?

GroupINN Architecture

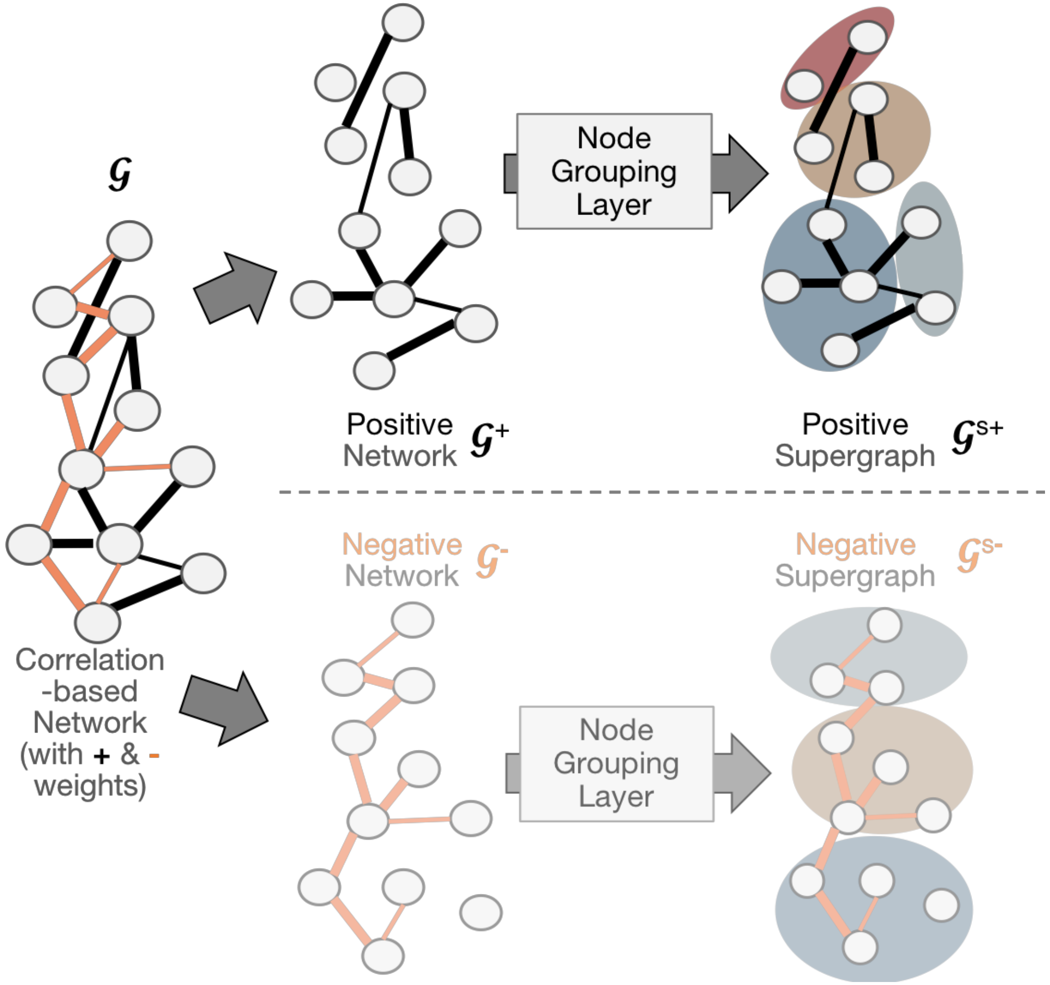


Correlation
-based
Network
(with + & -
weights)

GroupINN Architecture



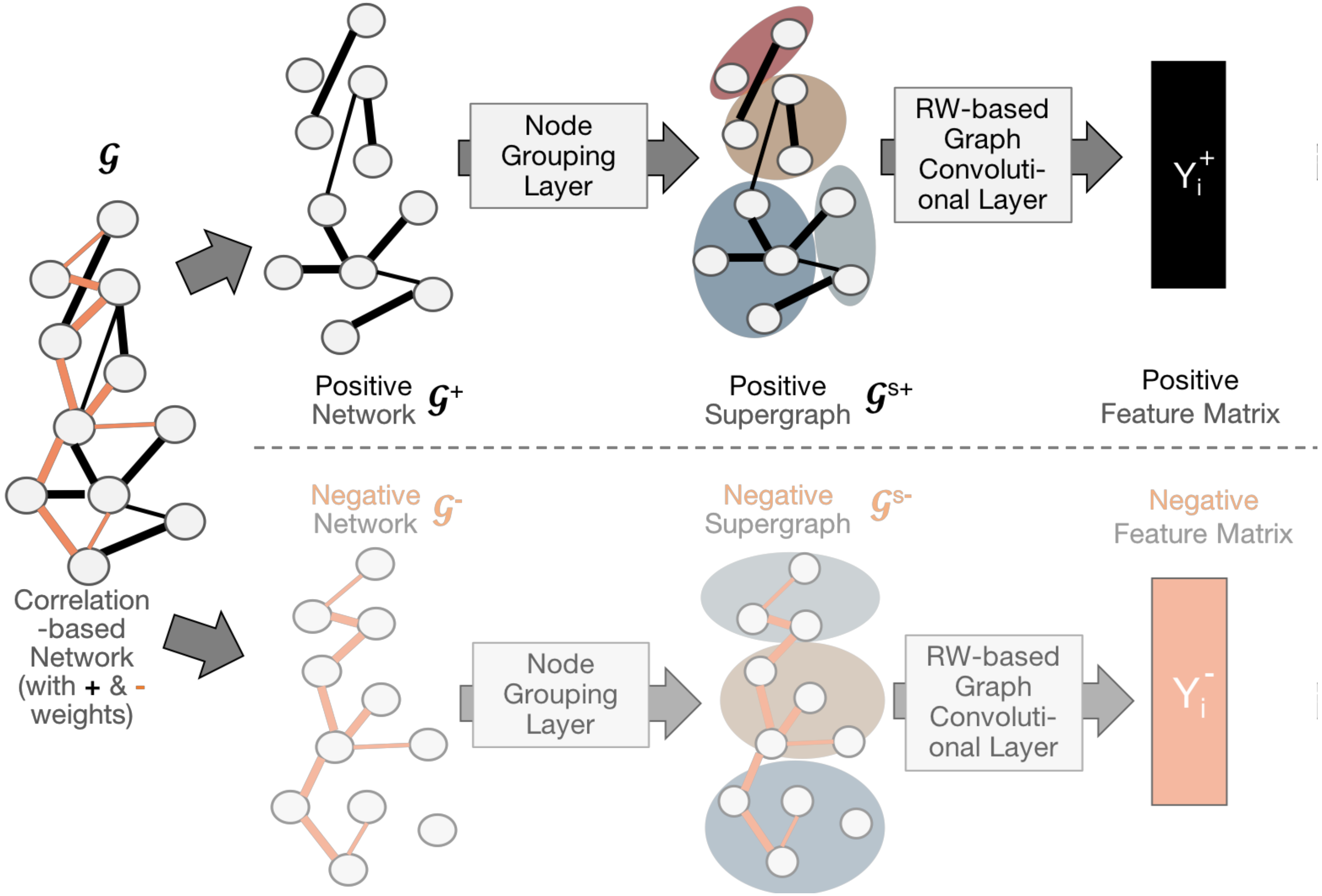
GroupINN Architecture



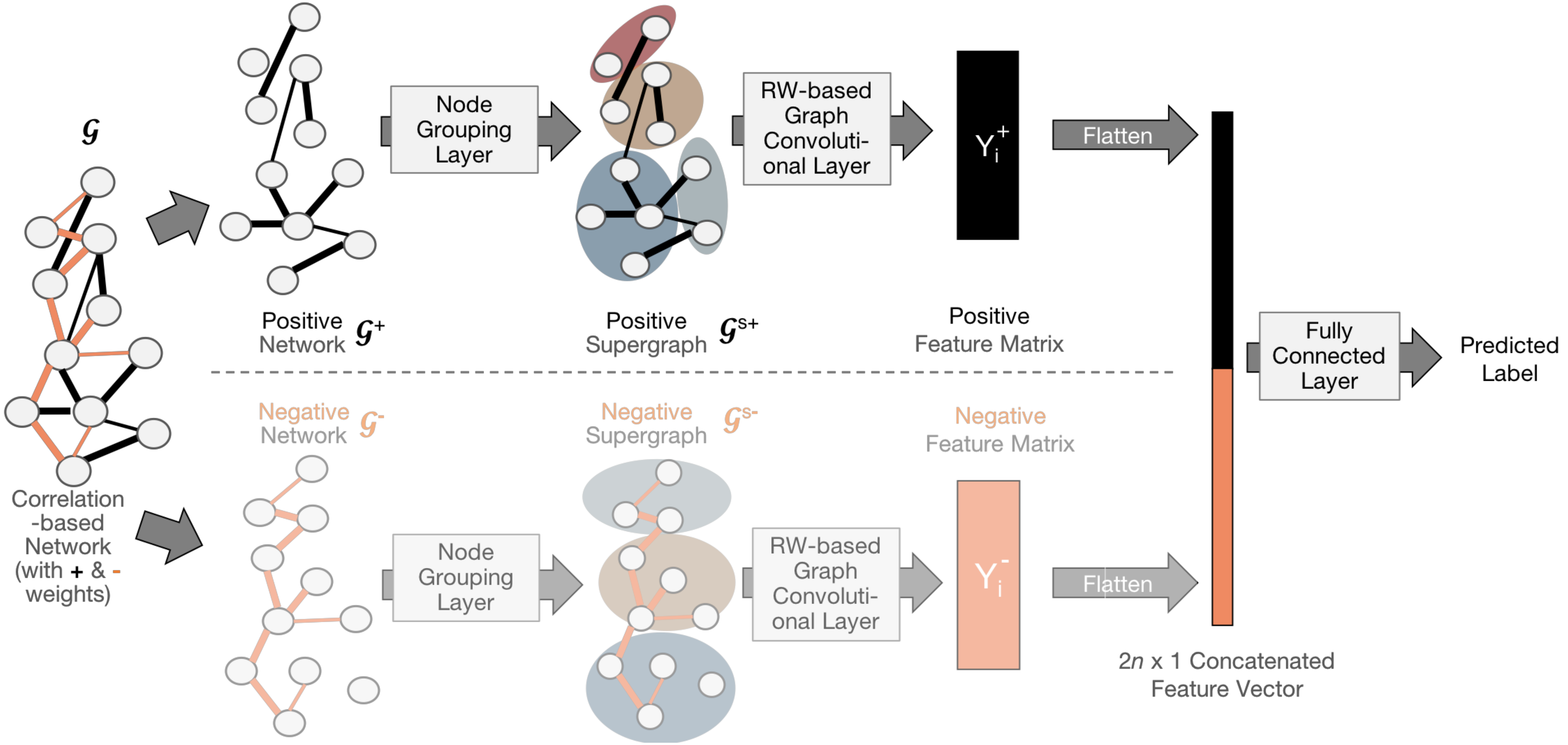
Graph Summarization to

- handle noisy data
- train from small samples of high-dim data
- support interpretability

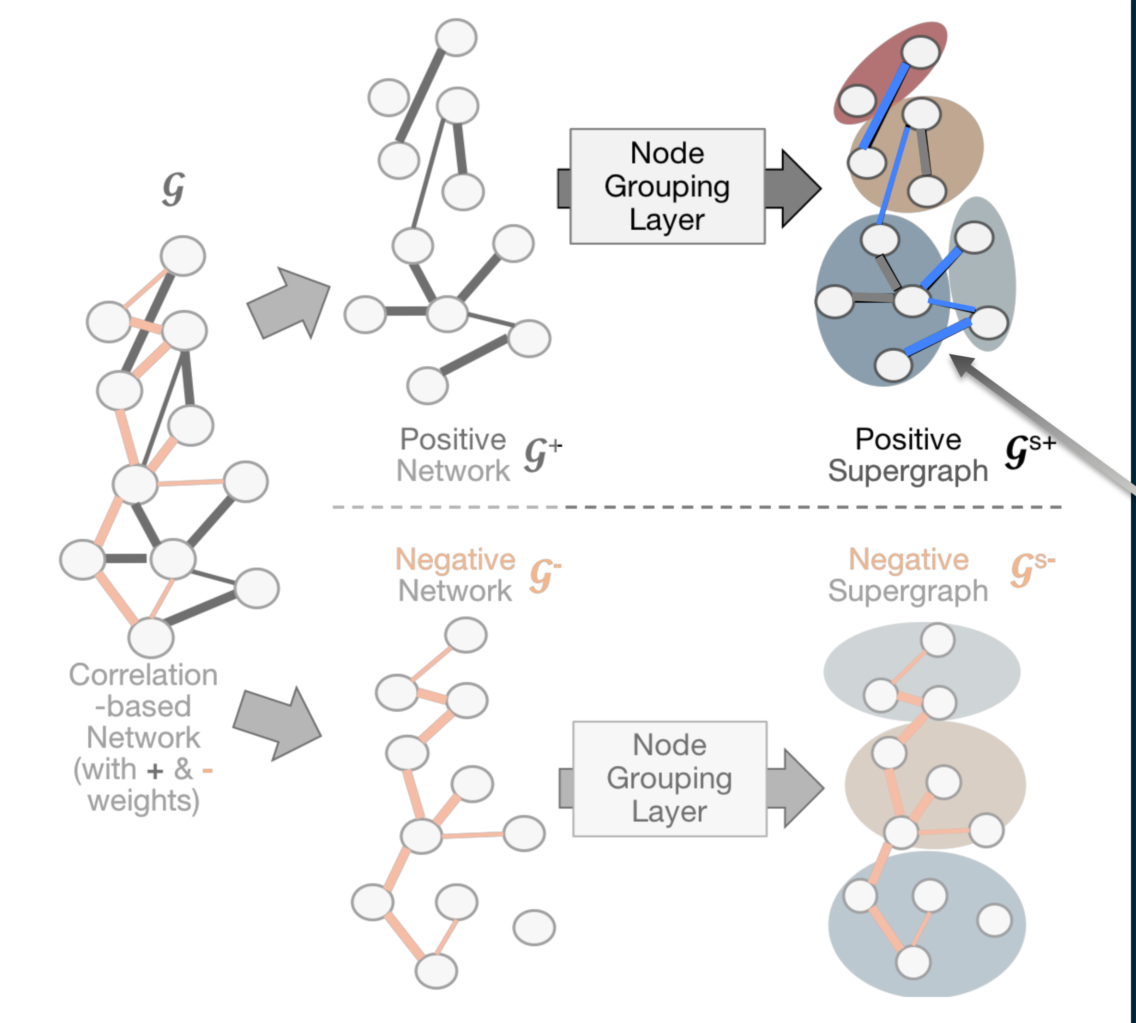
GroupINN Architecture



GroupINN Architecture



1. Node Grouping / *Summarization* Layer



- Recent findings have shown that some nodes (ROIs) are most related to the phenotype of interest
  → some edges are expected to be more indicative

  [Cohen+J Neurosci '16] [Cole+ NeuroImage '07]

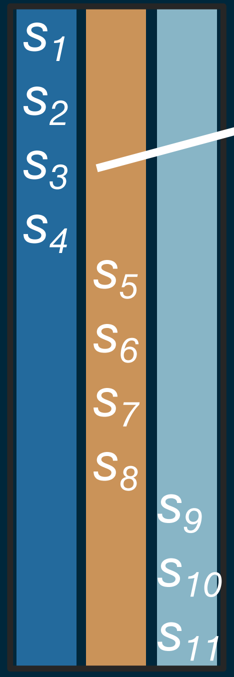
- Node grouping layer:
  - "hides" the non-indicative edges into a *supernode* and
  - highlights the indicative edges

[Yujun Yan, Jiong Zhu, et al. ACM KDD '19] https://github.com/GemsLab/GroupINN

1. Node Grouping / *Summarization* Layer

Details

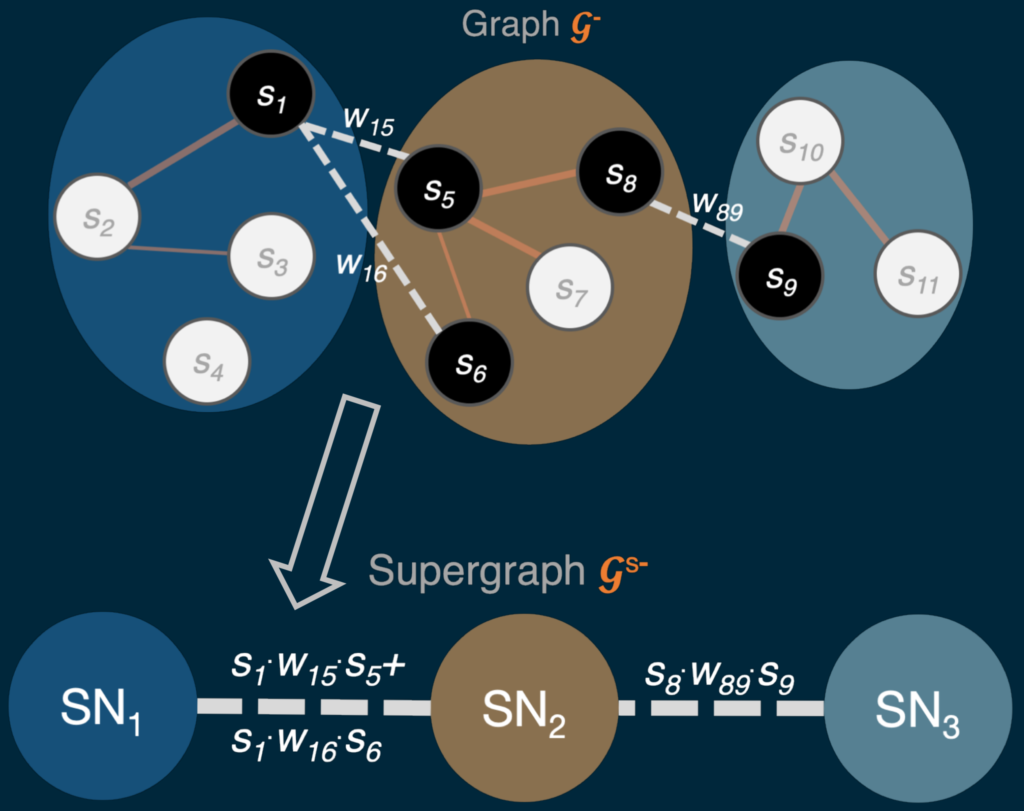
- **F**: learnable common membership matrix



Real valued importance score of node i in the prediction task

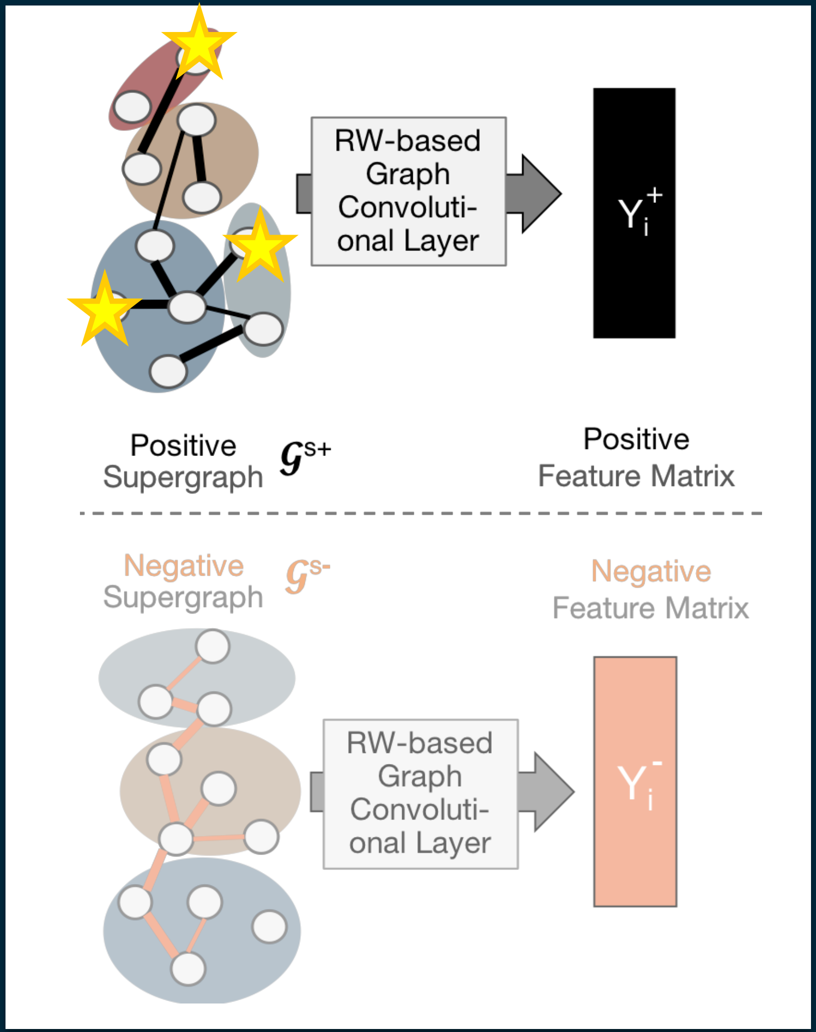
Interpretability

- Nonnegative
- Orthogonal (ideally)
- Nodes in supernode **not** required to be similar / well-connected



Adj. of supergraph $\mathbf{W}^s = \mathbf{F}^T \mathbf{W} \mathbf{F}$

2. RWR-based Graph Conv Layer: Intuition

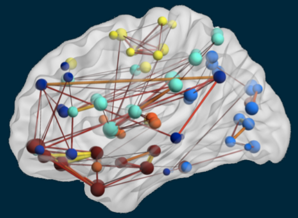


- Random walks:
  - useful tool to sample graph structure
  - the RWR scores quantify the similarities of other nodes to the seed nodes
- Design: The output $\mathbf{Y}_i$ of layer i is:

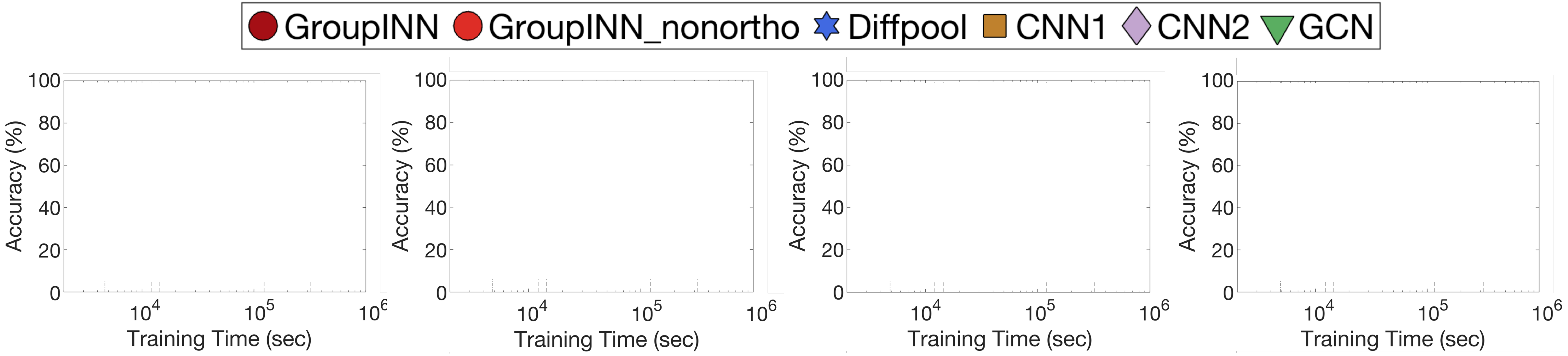
$$\mathbf{Y}_\mathrm{i} = \sigma(\mathrm{c}\mathbf{W}^{s}\mathbf{Y}_{\mathrm{i}-1}\mathbf{Q}_\mathrm{i} + \mathbf{I})$$

For more structure, multiple $\tilde{\mathbf{Q}}$ at diff. distances q_1 q_2 q_3

Q1. Comparison with *NN*-based methods



w/o orthogonality [Ying, NeurIPS' 18] [Wang, KDD' 17] [Kipf, ICLR' 17]



Emotion Gambling Social Working memory

thanks to the summarization layer

GroupINN models are up to 69× faster at training than all the baseline methods, while achieving same or higher accuracy in a variety of prediction tasks.

M CSE GEMS LAB [Yujun Yan, Jiong Zhu, et al. ACM KDD '19] https://github.com/GemsLab/GroupINN

Q2. Parsimony of GroupINN

Less is better!

Methods	# parameters	Normalized wrt GroupINN
CNN-1		
CNN-2		
GCN		
Diffpool		
GroupINN	2,892	1×

thanks to the summarization layer

GroupINN can use 15% or much fewer model parameters to achieve comparable or better performance of the baseline methods.

Q3. Interpretability

Acronyms of brain subnetworks. AN: auditory; CBLN: cerebellar; CON: cingulo-opercular; DAN: dorsal attention; FPN: frontoparietal; MRN: memory retrieval; SN: salience ; VAN: ventral attention ; VN: vision; SM.M: sensory/somatomotor mouth; SM.H: sensory/somatomotor hand

Tasks	Within subnetworks								
	GroupINN			PCA			Diffpool		
Working Memory	MRN	FPN	SN	SM.M	SM.H	AN	SM.M	MRN	CBLN
Gambling	VAN	VN	DAN	SM.H	AN	SM.M	SM.M	FPN	MRN
Emotion	SN	CON	VAN	SM.M	SM.H	AN	DMN	MRN	SM.M
Social	FPN	SN	VAN	SM.M	CBLN	AN	DAN	SM.M	FPN

- GroupINN finds the most task-positive sub-networks.
- PCA and Diffpool are misled by strong noisy signals from mouth and hand motion.

thanks to the summarization layer

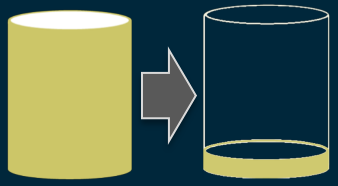
[Cohen, et al. , J Neurosci 2016];
[Cole, et al. Neuron 2014];
[Davison, et al. PLOS Comp Bio 2015]

M CSE GEMS LAB [Yujun Yan, Jiong Zhu, et al. ACM KDD '19]

This talk: Summarization in Network Representation Learning

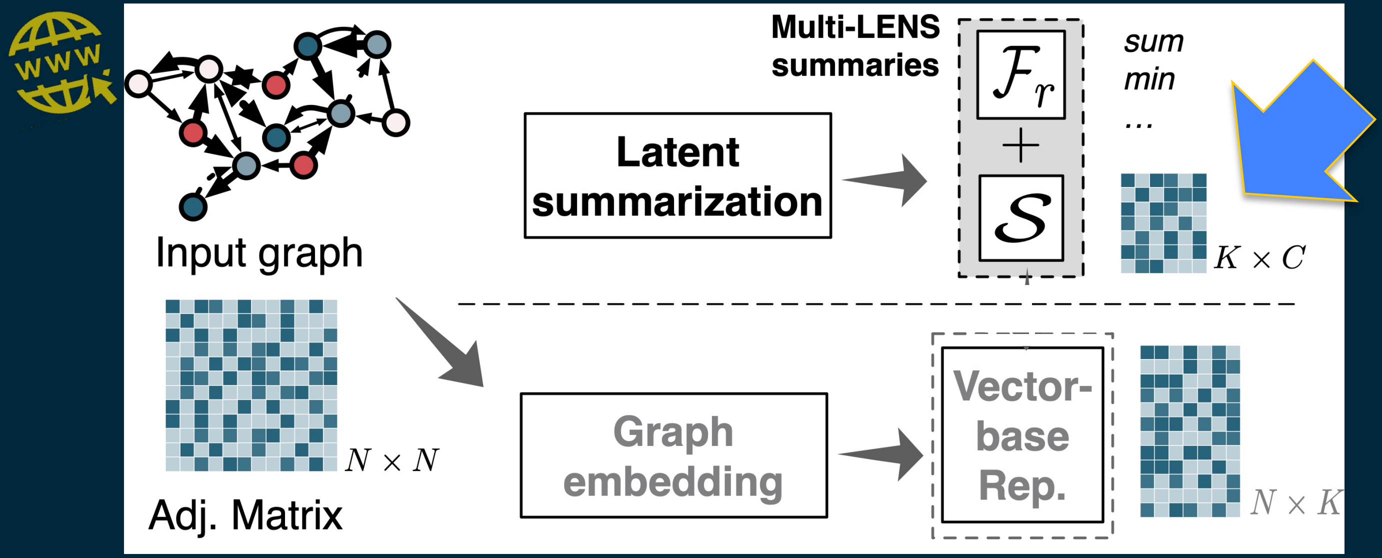
- Summarization within a GCN for faster training, data denoising and interpretability [ACM KDD'19a]

- Embedding summarization for compression and on-the-fly computation [ACM KDD'19b; PKDD'19]

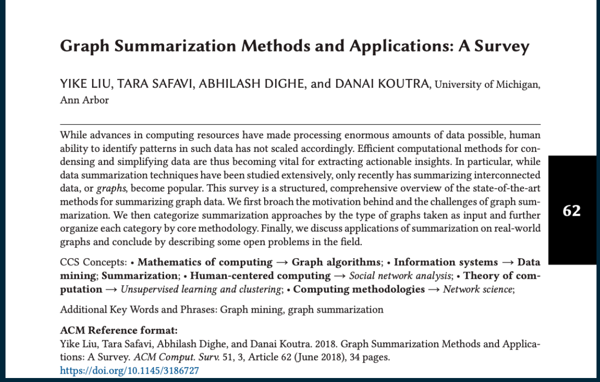


Embeddings are powerful, but can take up a lot of space!

- For 1B nodes and K=128 → 1TB to store the embeddings!
- Can we summarize them?



Graph Summarization Survey



Graph Summarization Methods and Applications: A Survey

YIKE LIU, TARA SAFAVI, ABHILASH DIGHE, and DANAI KOUTRA, University of Michigan, Ann Arbor

While advances in computing resources have made processing enormous amounts of data possible, human ability to identify patterns in such data has not scaled accordingly. Efficient computational methods for condensing and simplifying data are thus becoming vital for extracting actionable insights. In particular, while data summarization techniques have been studied extensively, only recently has summarizing interconnected data, or *graphs*, become popular. This survey is a structured, comprehensive overview of the state-of-the-art methods for summarizing graph data. We first broach the motivation behind and the challenges of graph summarization. We then categorize summarization approaches by the type of graphs taken as input and further organize each category by core methodology. Finally, we discuss applications of summarization on real-world graphs and conclude by describing some open problems in the field.

62

CCS Concepts: • **Mathematics of computing** → **Graph algorithms**; • **Information systems** → **Data mining**; **Summarization**; • **Human-centered computing** → *Social network analysis*; • **Theory of computation** → *Unsupervised learning and clustering*; • **Computing methodologies** → *Network science*;

Additional Key Words and Phrases: Graph mining, graph summarization

ACM Reference format:
Yike Liu, Tara Safavi, Abhilash Dighe, and Danai Koutra. 2018. Graph Summarization Methods and Applications: A Survey. *ACM Comput. Surv.* 51, 3, Article 62 (June 2018), 34 pages.
https://doi.org/10.1145/3186727

Graph Summarization

Network type

- Static
  - Plain
  - Labeled
- Dynamic
  - Plain

Core techniques employed

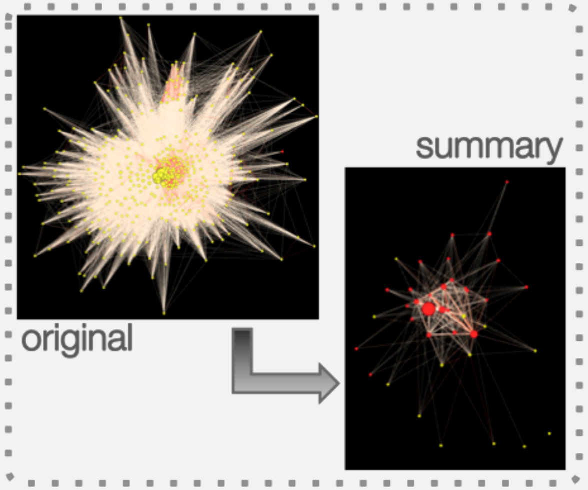
- Structure only
  - Grouping
  - Compression
  - Simplification
  - Influence
- Structure + labels
  - Grouping
  - Compression
  - Influence
- Temporal structure
  - Grouping
  - Compression
  - Influence

Applications

- Query Efficiency
- Compression
- Visualization
- Pattern Discovery
- Influence Analysis

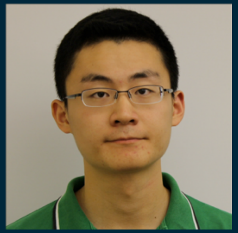
Open Problems

- Richer Data
- Streaming / Incremental Summarization
- Automated Insight Extraction
- Evaluation

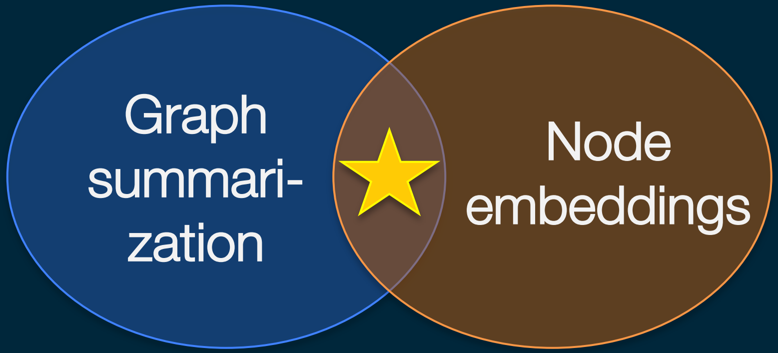


Related Research Areas:
graph clustering, partitioning, community detection, sampling, sparsification, sketches, compression

Latent Network Summarization



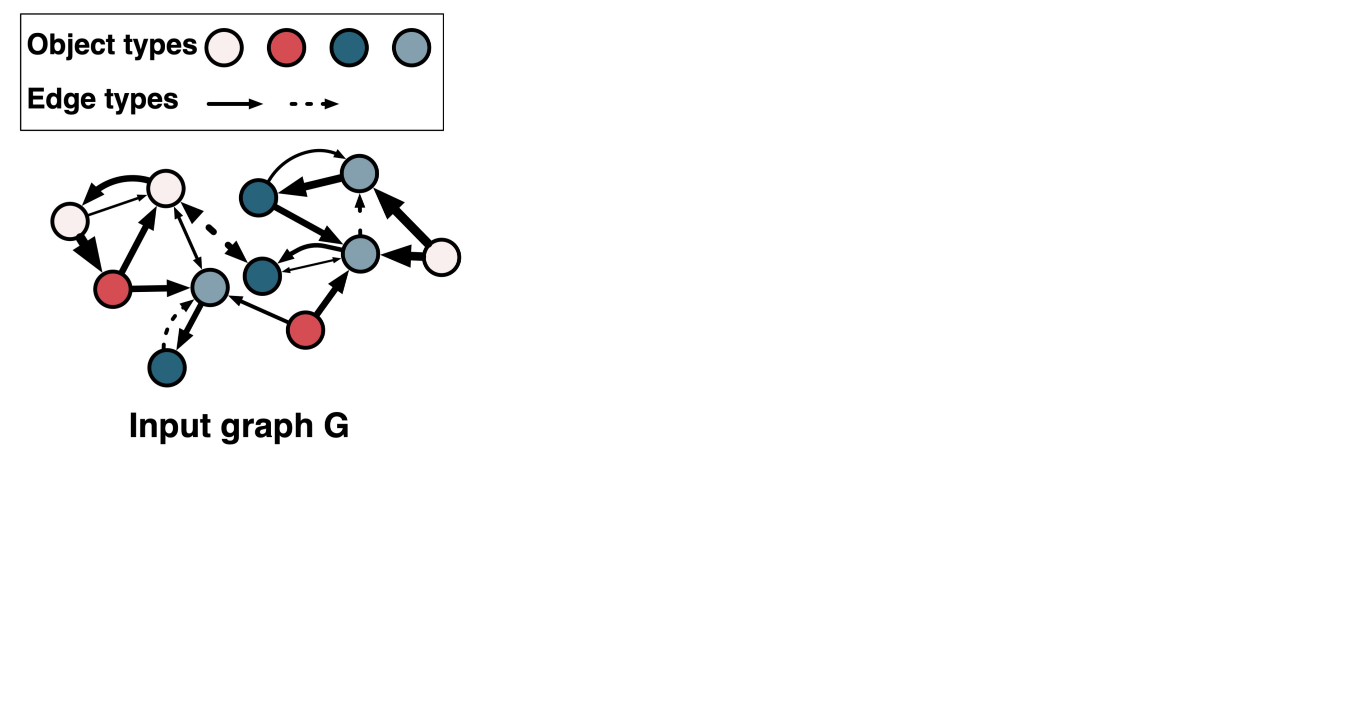
- Given: a graph G(V, E)
- Find: a compressed representation that captures the key structural information and is
  - ✧ independent of graph size (|V|, |E|), and
  - ✧ capable of deriving node representations on the fly



Comparison to Related Work

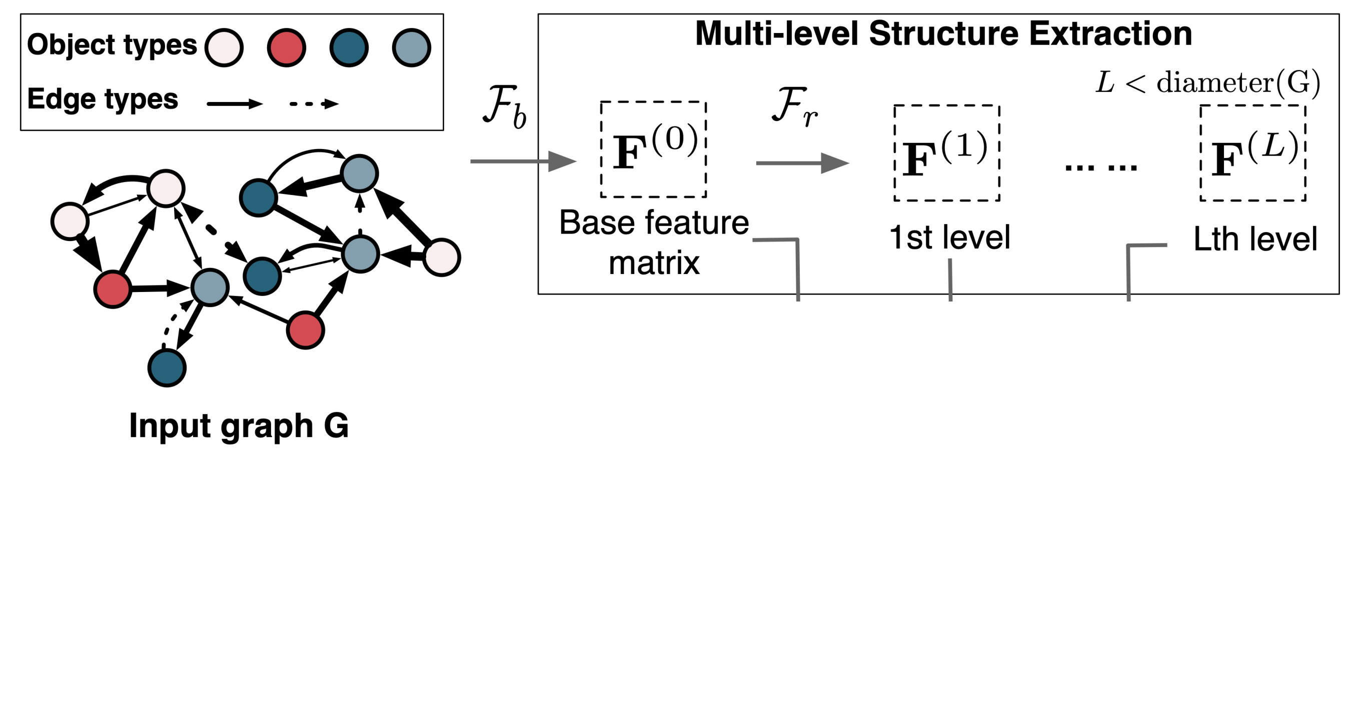
	Input	Representations / Output			Method	
	Hetero-geneity	Size indep.	Node specific	Proxim. indep.	Scalable	Induc.
Aggregation [2]	✓	✗	✗	✗	✓	✗
Cosum [34]	✗	✗	✗	✓	✗	✗
AspEm [31]	✓	✗	✓	✗	✓	✗
metapath2vec [8]	✓	✗	✓	✗	✓	✗
n2vec [11], LINE [32]	✗	✗	✓	✗	✓	✗
struc2vec [26]	✗	✗	✓	✓	✗	✗
DNGR [6]	✗	✗	✓	✗	✗	✗
GraphSAGE [12]	✓	✗	✓	✓	✓	✓
Multi-LENS	✓	✓	✓	✓	✓	✓

Latent Network Summarization: Overview



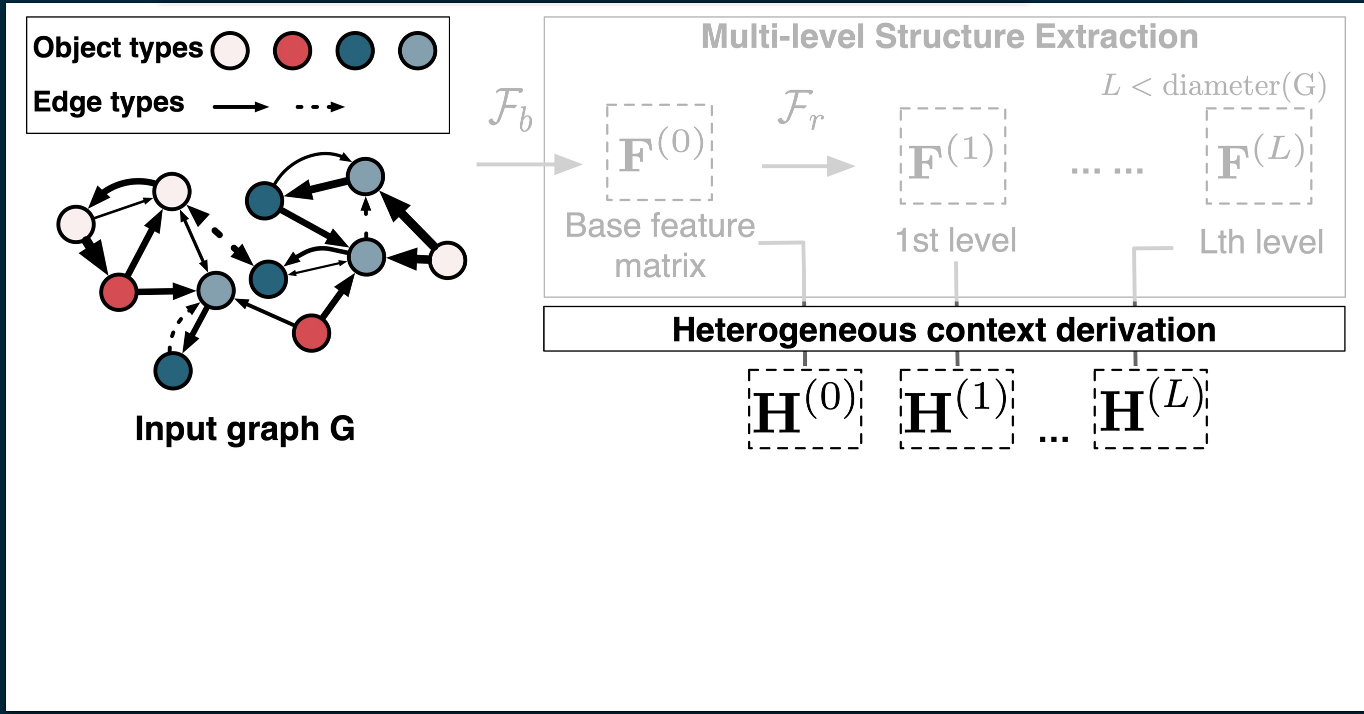
Latent Network Summarization: Overview

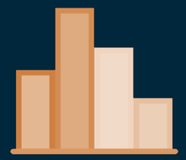
1. Relational functions to aggregate nodewise structural features automatically



Latent Network Summarization: Overview

1. Relational functions to aggregate nodewise structural features automatically

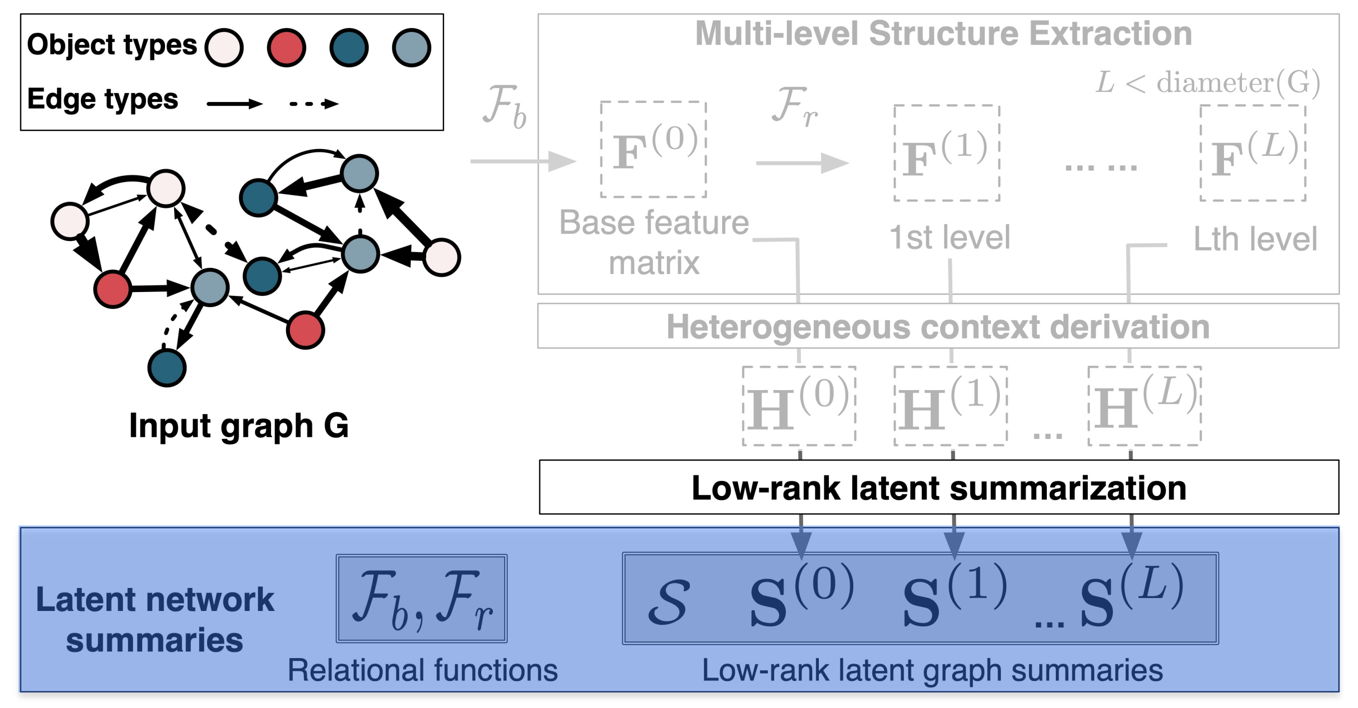




2. Histogram-based heterogeneous contexts for nodes

Latent Network Summarization: Overview

1. Relational functions to aggregate nodewise structural features automatically



2. Histogram-based heterogeneous contexts for nodes

3. Subspace vectors from which we can derive the embeddings

Space comparison

Data	SE	LINE	n2vec	DW	m2vec	AspEm	G2G	ML (MB)
facebook	8.13x	8.48x	12.79x	12.84x	3.82x	8.50x	9.17x	**0.58**
yahoo	187.1x	180.0x	242.2x	231.0x	79.8x	197.4x	195.8x	**0.62**
dbpedia	710.0x	714.2x	996.4x	996.2x	-	749.2x	743.6x	**0.81**
digg	608.2x	612.8x	848.9x	830.3x	259.9x	641.7x	635.2x	**0.54**
bibson.	1512.1x	1523.0x	2152.5x	2152.5x	-	1595.8x	-	**0.75**

Multi-LENS requires 4-2152x less output storage space than the other embedding methods.

Data	#Nodes	#Edges	#Node Types	Graph Type
facebook	4 039	88 234	1	unweighted
yahoo-msg	100 058	1 057 050	2	weighted
dbpedia	495 936	921 710	4	unweighted
digg	283 183	4 742 055	2	unweighted
bibsonomy	977 914	3 754 828	3	weighted

Link Prediction

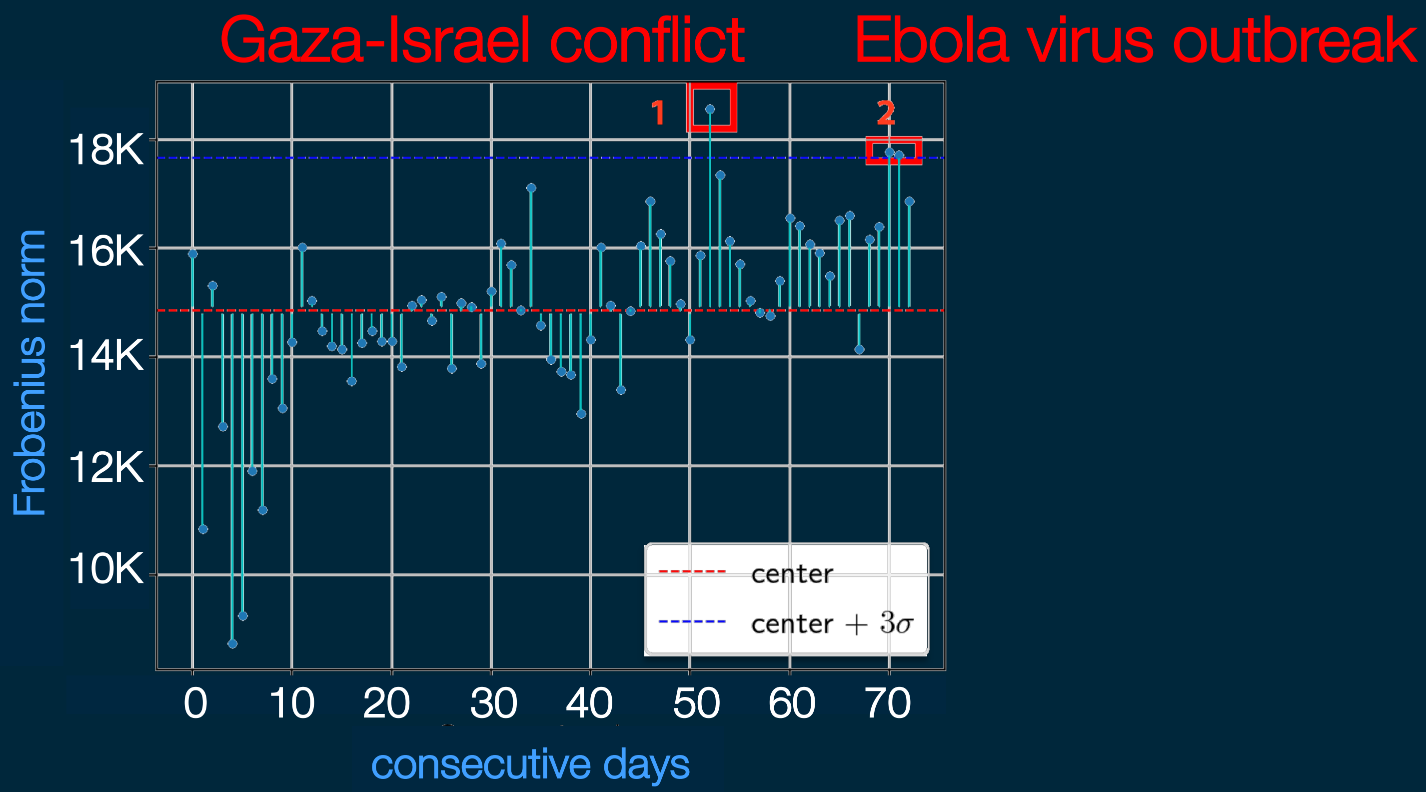
Data	Metric	NA	SE	LINE	DW	n2vec	GR	s2vec	DNGR	m2vec	AspEm	G2G	**ML**(L = 1)	**ML**(L = 2)
facebook	**AUC**	0.6213	0.6717	0.7948	0.7396	0.7428	0.8157	0.8155	0.7894	0.7495	0.5886	0.7968	0.8703	**0.8709***
	ACC	0.5545	0.5995	0.7210	0.6460	0.6544	0.7368	0.7388	0.7062	0.7051	0.5628	0.7274	**0.7920***	0.7904
	F1 macro	0.5544	0.5716	0.7210	0.6296	0.6478	0.7367	0.7387	0.7060	0.7041	0.5628	0.7273	**0.7920***	0.7905
yahoo-msg	**AUC**	0.7189	0.5375	0.6745	0.7715	0.7830	0.7535			0.6708	0.5587	0.6988	0.8443	**0.8446***
	ACC	0.2811	0.5224	0.6269	0.6927	0.7036	0.6825	OOT	OOM	0.6164	0.5379	0.6564	**0.7587***	**0.7587***
	F1 macro	0.2343	0.5221	0.6265	0.6897	0.7016	0.6821			0.6145	0.5377	0.6562	**0.7577***	**0.7577***
dbpedia	**AUC**	0.6002	0.5211	0.9632	0.8739	0.8774					0.6364	0.7384	**0.9820***	0.9809
	ACC	0.3998	0.5399	0.9111	0.8436	0.8436	OOM	OOT	OOM	OOT	0.5869	0.6625	**0.9186**	0.9151
	F1 macro	0.2968	0.4539	0.9110	0.8402	0.8402					0.5860	0.6613	**0.9186**	0.9150
digg	**AUC**	0.7199	0.6625	0.9405	0.9664	0.9681				0.9552	0.5644	0.8978	**0.9894***	0.9893
	ACC	0.2801	0.6512	0.8709	0.9023	0.9049	OOM	OOT	OOM	0.8891	0.5459	0.8492	**0.9596***	0.9590
	F1 macro	0.2660	0.6223	0.8709	0.9019	0.9046				0.8890	0.5459	0.8492	**0.9595***	0.9590
bibsonomy	**AUC**	0.7836	0.6694	0.9750	0.6172	0.6173					0.6127		**0.9909***	0.9909
	ACC	0.2164	0.6532	0.9350	0.5814	0.5816	OOM	OOT	OOM	OOT	0.5790	OOM	**0.9485***	0.9466
	F1 macro	0.2070	0.6064	0.9349	0.5781	0.5782					0.5772		**0.9485***	0.9466

The Multi-LENS node embeddings outperform all the baselines by 3.5–34.3% in AUC.

Inductive Anomaly Detection

- Learn summary of G_{t-1}, apply to G_t
- Compute the distance between the embeddings at *t-1* and *t*

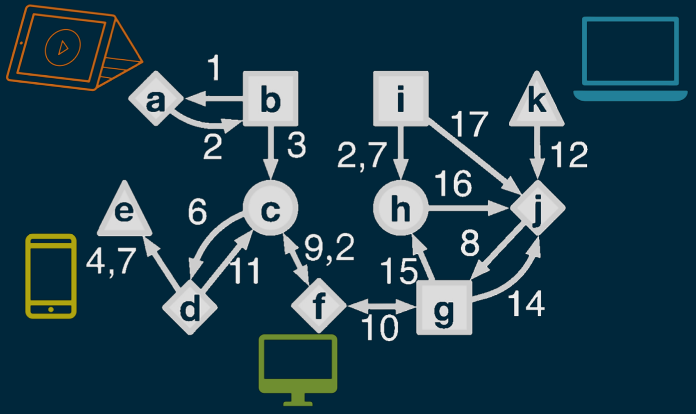
308K hashtags
2.6M edges
May -- July 2014



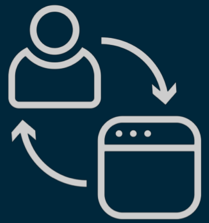
[Di Jin, Rossi et al. ACM KDD'19] https://github.com/GemsLab/MultiLENS

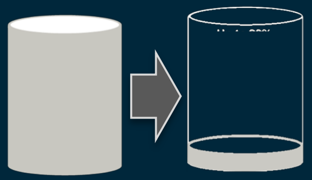
Can we summarize / compress the embeddings in a different way?

Temporal, Hash-based Node Embeddings



- Given: a time-evolving heterogeneous network $G(V, E)$
- Learn: a function $\chi: V \rightarrow \{0,1\}^d$ s.t. the derived d-dim embeddings
  1) preserve similarities in interactions in G,
  2) are space-efficient, and
  3) accurately capture temporal information and the heterogeneity of the underlying network





[Di Jin, Mark Heimann, et al. PKDD’19] https://github.com/GemsLab/node2bits

Example: Find similarities in user interactions

- User stitching:
  - ✧ The task of identifying and matching various online references to the **same user** in real-world web services.
  - ✧ Instance of **entity resolution**

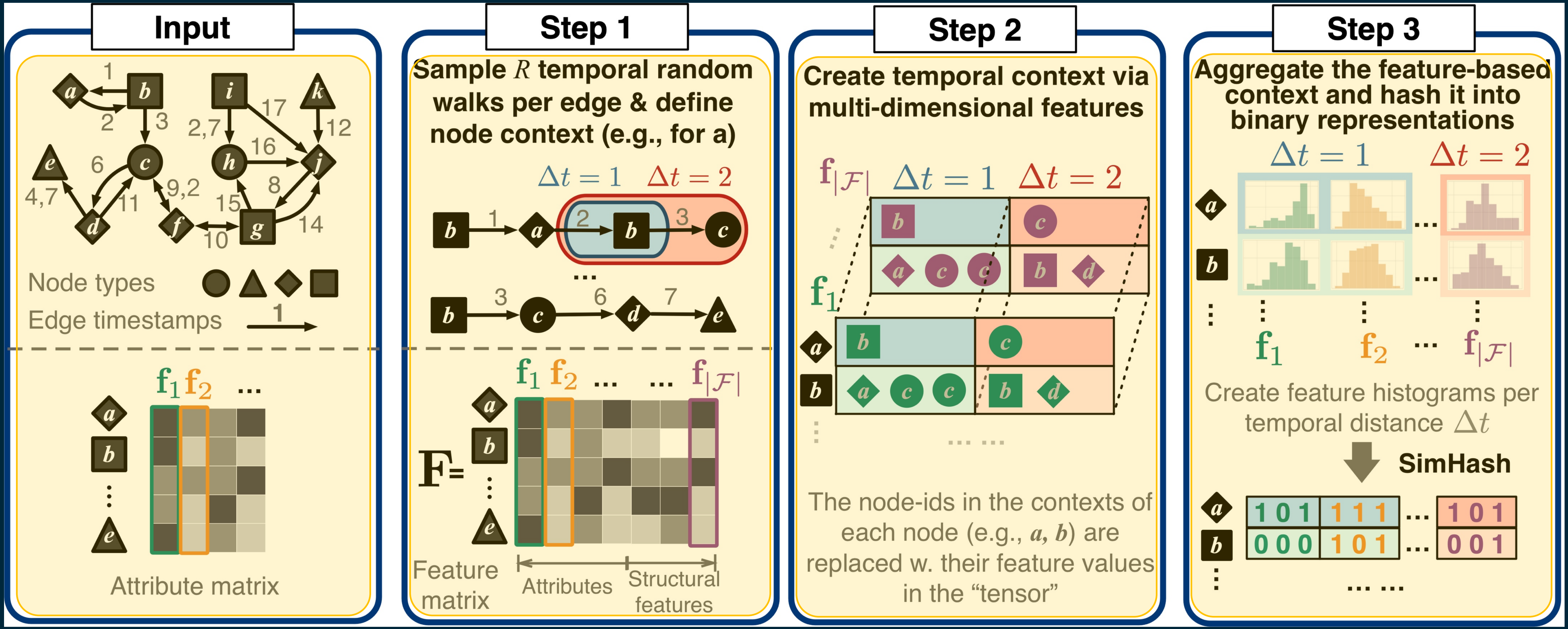
Web device log data

[Cohen, W.W., Richman, J., KDD 2012]
[Dasgupta, A. +, WSDM 2012]
[Saha Roy, R.,+ WWW'15]
[Kim, Kini, + WWW'17]
[Bhattacharya, I., Getoor, L. TKDD 2007], …

node2bits: Key ideas

- [R1] Graph heterogeneity
  - ✧ General approach that aggregates rich features + node types
- [R2] Temporal dynamics
  - ✧ Temporally valid walks to capture short- and long-term interactions
  - ✧ Functionally similar nodes are represented by multiple features (structural sim)
- [R3] Efficient similarity comparison
  - ✧ Use LSH to hash similar nodes (linear complexity)
- [R4] Low storage requirement
  - ✧ Binary hashcode with fixed length

node2bits: Workflow



[Di Jin, Mark Heimann, et al. PKDD'19] https://github.com/GemsLab/node2bits

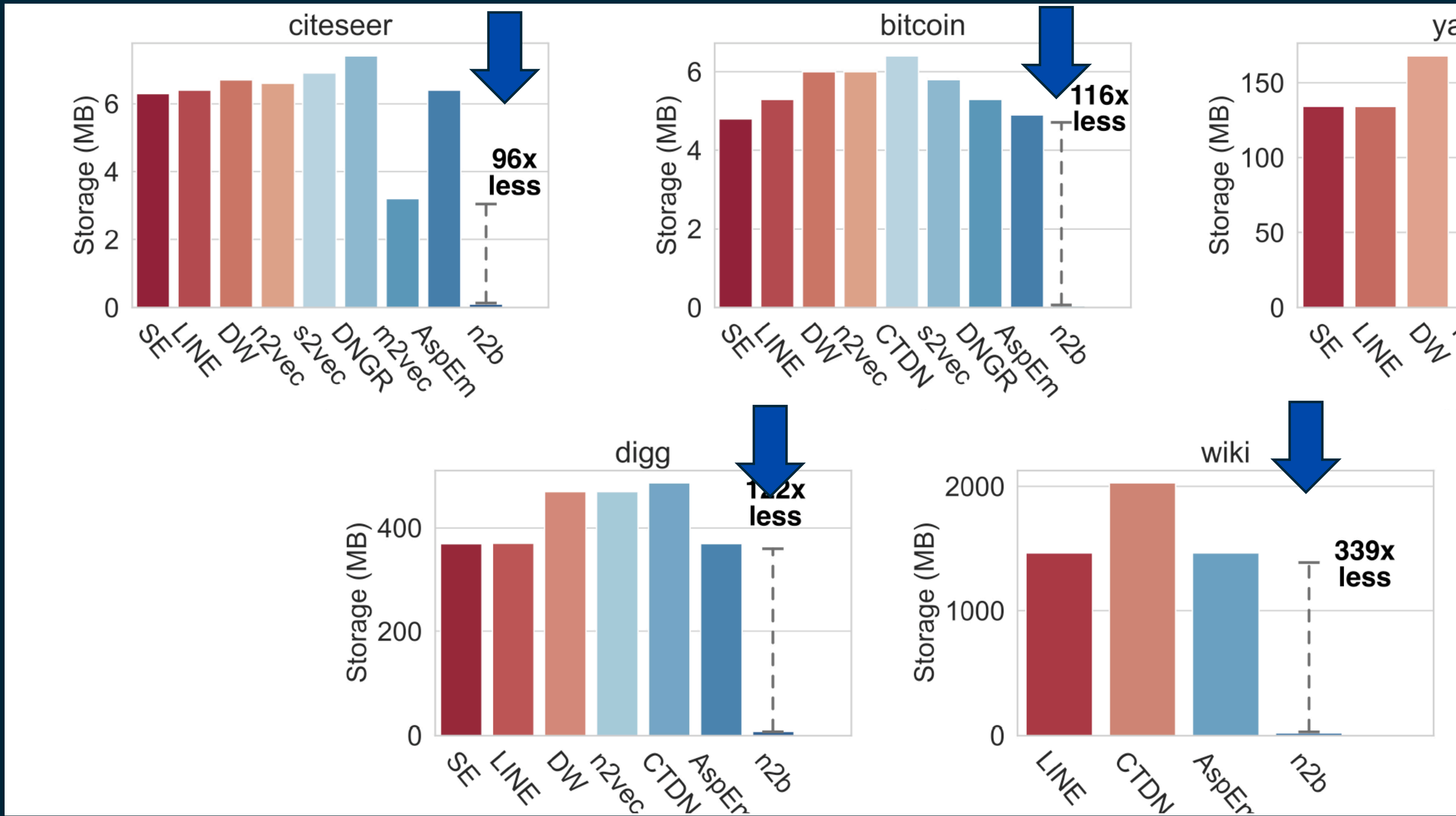
Q1: *Supervised* Identity stitching

short term   long term

	Metric	CN	SE	LINE	DW	n2vec	s2vec	DNGR	AspEm	CTDNE	N2B-0	N2B-SH	N2B-LN
bitcoin	**AUC**	0.7474	0.5828	0.6071	0.6306	0.6462	0.8025	0.5909	0.5344	0.6987	0.7584	0.7609	0.7380
	ACC	0.7174	0.5842	0.5842	0.6158	0.6158	0.7263	0.5526	0.5316	0.6000	0.7211	0.7268	0.6737
	F1	0.7001	0.5728	0.5828	0.6158	0.6157	0.7263	0.5525	0.5315	0.5964	0.7209	0.7271	0.6735
digg	**AUC**	0.6217	0.5171	0.7878	0.7398	0.7445			0.5105	0.6967	0.8185*	0.7611	0.7587
	ACC	0.6217	0.5152	0.7694	0.6971	0.7013	OOT	OOM	0.5088	0.5915	0.7982*	0.7418	0.7444
	F1	0.5585	0.3770	0.7683	0.6960	0.7003			0.5088	0.5884	0.7958*	0.7411	0.7433
wiki	**AUC**	0.6997		0.7854					0.5374	0.7707	0.8230	0.8259*	0.8214
	ACC	0.6997	OOT	0.7132	OOM	OOM	OOT	OOM	0.5141	0.6488	0.7145	0.7510*	0.7103
	F1	0.6699		0.7129					0.5141	0.6398	0.7088	0.7476*	0.7067
comp-X	**AUC**	0.5970		0.5000					0.5213		0.8095*	0.7496	0.7525
	ACC	0.5970	OOM	0.6757	OOM	OOM	OOT	OOM	0.5103	OOM	0.8414*	0.7959	0.7975
	F1	0.5189		0.4032					0.5103		0.8154*	0.7581	0.7606

Dynamic + static variants of node2bits outperform baselines by up to 5.2% in AUC and 4.9% in F1 score. Short-term tactic performs better.

Q2: Output storage efficiency



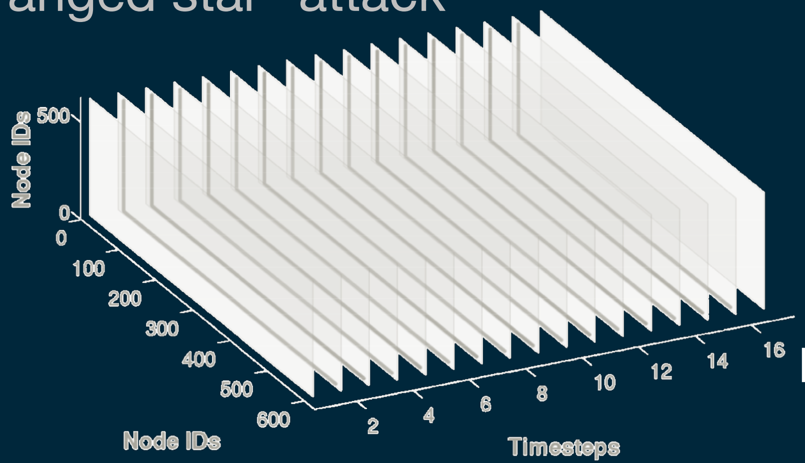
Real-valued vs. binary embeddings

node2bits uses 63-339× less space than the baselines, while achieving comparable or better stitching performance.

Summarizing Large Networks: Overview

Survey: [CSUR'18]

“ranged star” attack

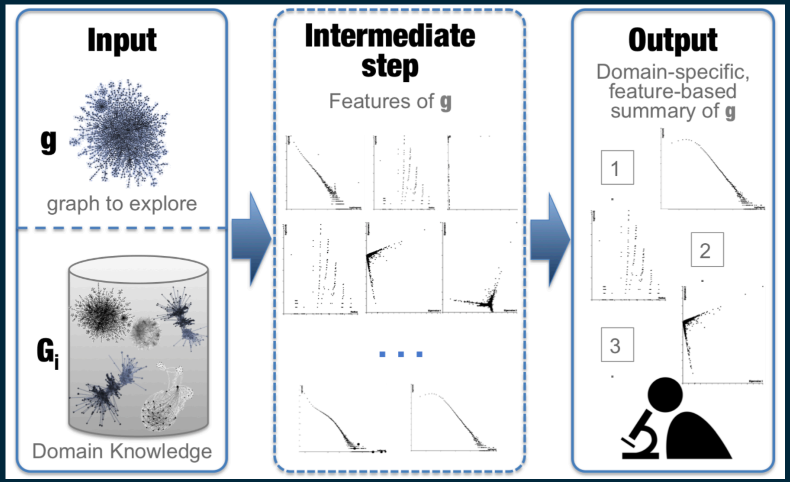


Structural Summaries

[SDM'14, KDD'15, Dat Bull Eng'17, SNAM'18, SDM'19, …]

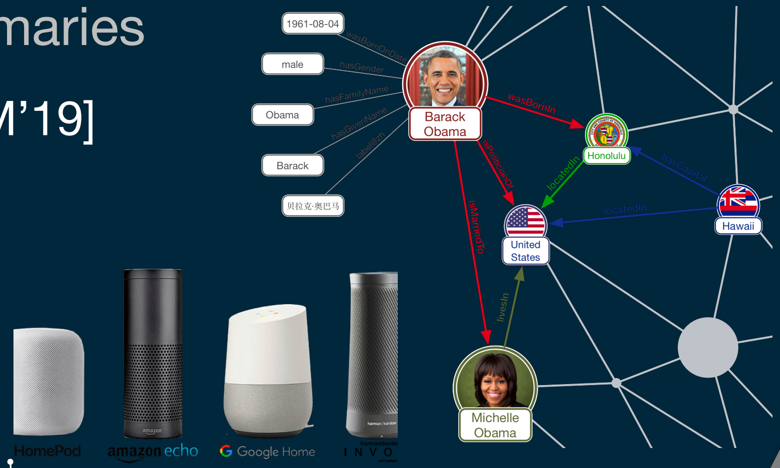
Domain-specific Summaries

[ICDM'17, KDD'19]



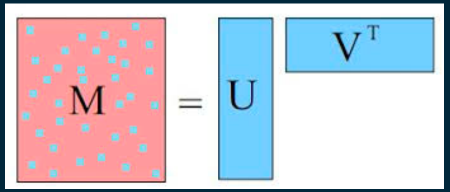
Query-on-the-edge + Rule-based Summaries

[ICDM'19]



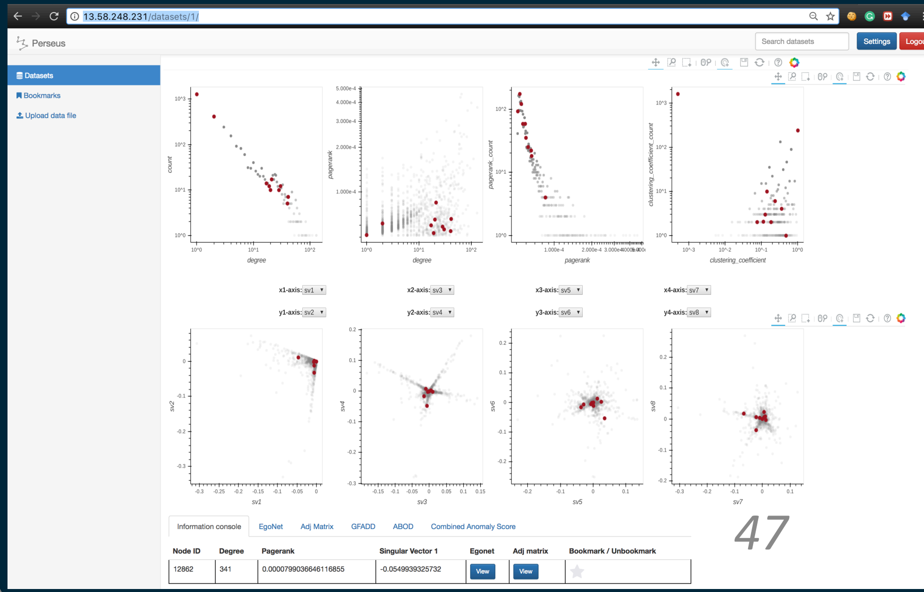
Latent Summaries

[KDD'19]



Interactive Summaries

[VLDB'15, Informatics'17]

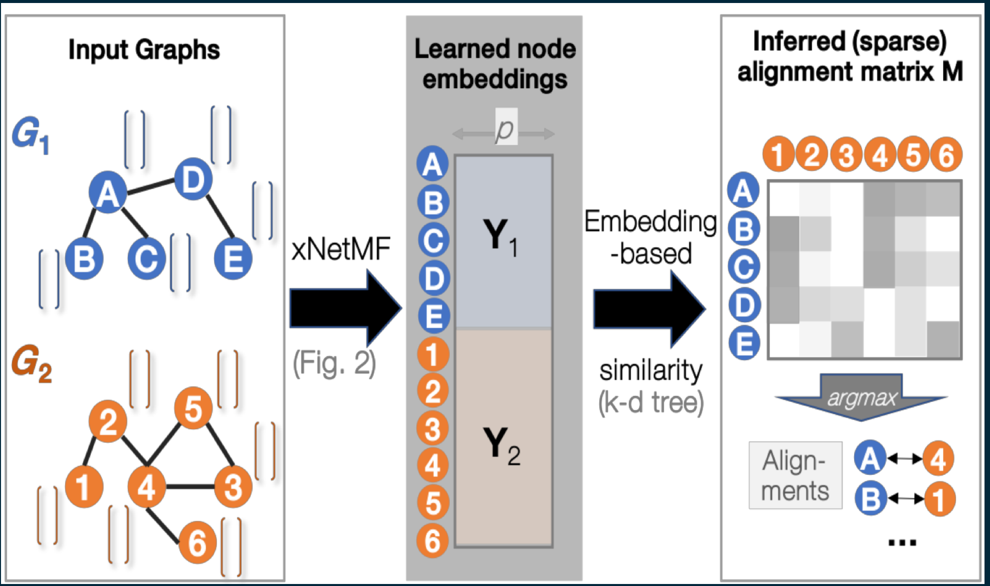


M CSE

GEMS LAB

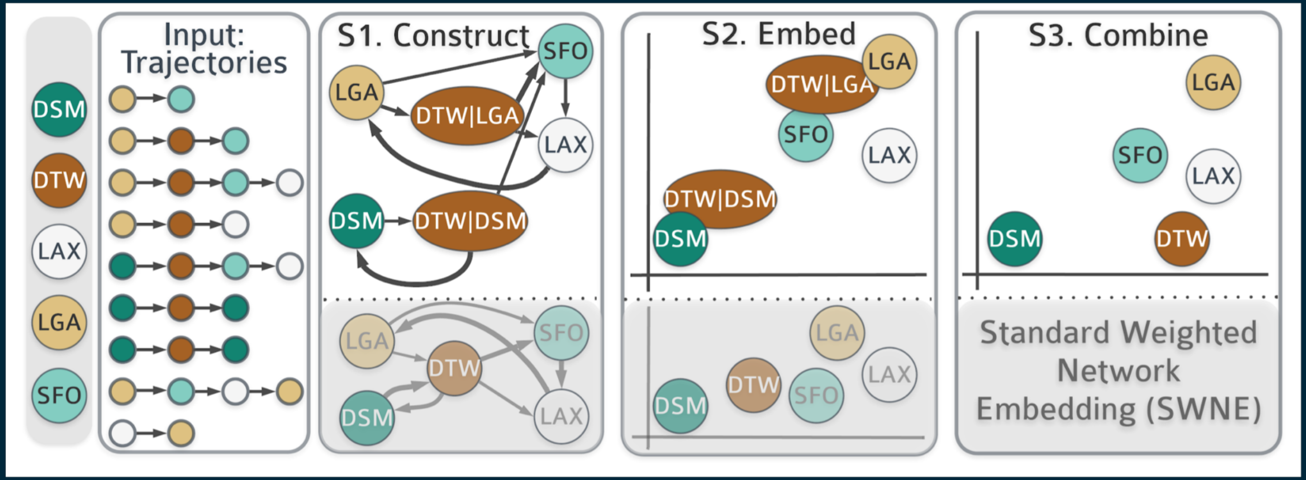
Beyond Summarization

Structural embeddings for network alignment



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

When is it useful to learn over higher-order networks, and when is it not?

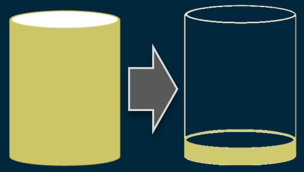


[Caleb Belth, Fahad Kamran, Donna Tjandra, Danai Koutra. IEEE/ACM ASONAM'19]

Take-away messages: Summarization in Network Representation Learning

- Structural embeddings are less studied, but are more appropriate than proximity-based ones in several tasks
- Summarization within a GCN can help with faster training, data denoising and interpretability [ACM KDD'19a]

- Embedding summarization can achieve compression and on-the-fly computation of representations [ACM KDD'19b; PKDD'19]



- Histograms are powerful at capturing the graph structure [ACM CIKM'18; ACM KDD'19b,c; ECML/PKDD'19; IEEE ICDM'19]
  - ✧ flexible, versatile (heterogeneity, attributes, directionality, weights…),
  - ✧ less information loss

Talk based on the following papers

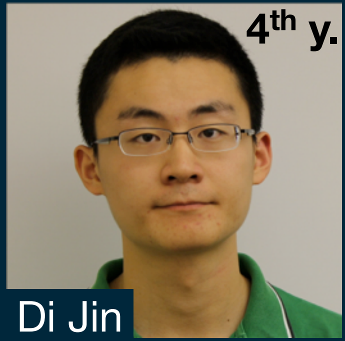
- Mark Heimann, Haoming Shen, Tara Safavi, Danai Koutra. REGAL: Representation Learning-based Graph Alignment. ACM CIKM’18.
- Y. Liu, T. Safavi, A. Dighe, D. Koutra. Graph Summarization Methods and Applications: A Survey. ACM Computing Surveys 2018.
- YujunYan, J. Zhu, Marlena Duda, Eric Solarz, Chandra 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.
- Caleb Belth, Fahad Kamran, Donna Tjandra, Danai Koutra. When to Remember Where You Came from: Node Representation Learning in Higher-order Networks. IEEE/ACM ASONAM 2019.
- Tara Safavi, Caleb Belth, Lukas Faber, Davide Mottin, Emmanuel Müller, Danai Koutra. Personalized Knowledge Graph Summarization: From the Cloud to Your Pocket. IEEE ICDM 2019.
- Mark Heimann, Tara Safavi, Danai Koutra. Distribution of Node Embeddings as Multiresolution Features for Graphs. IEEE ICDM 2019.
- Ryan A. Rossi, Di Jin, Sungchul Kim, Nesreen K. Ahmed, Danai Koutra, John Boaz Lee. From Community to Role-based Graph Embeddings. Arxiv 2019.

GEMS LAB

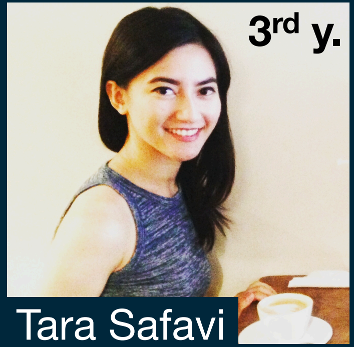
Caleb Belth

Marlena Duda

Mark Heimann



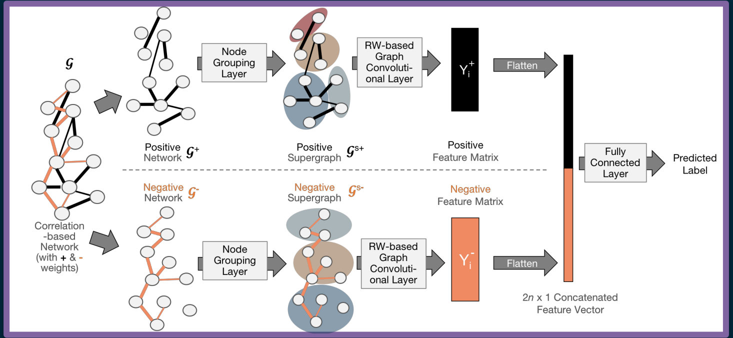
Di Jin

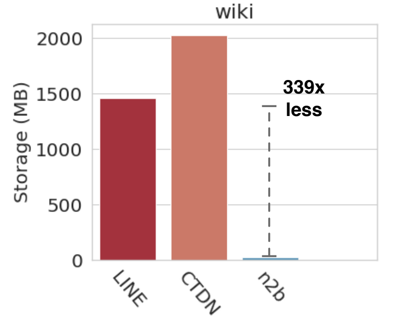


Tara Safavi

Yujun Yan

The Power of Summarization in Network Representation Learning





Thank you!
Questions?

http://danaikoutra.com

dkoutra@umich.edu

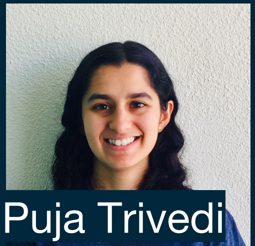
</>

https://github.com/GemsLab/GroupINN

https://github.com/GemsLab/MultiLENS

https://github.com/GemsLab/node2bits


Alican Büyükçakır


Puja Trivedi


Jiong Zhu

NSF ARL P&G MIDAS amazon Adobe Microsoft Azure