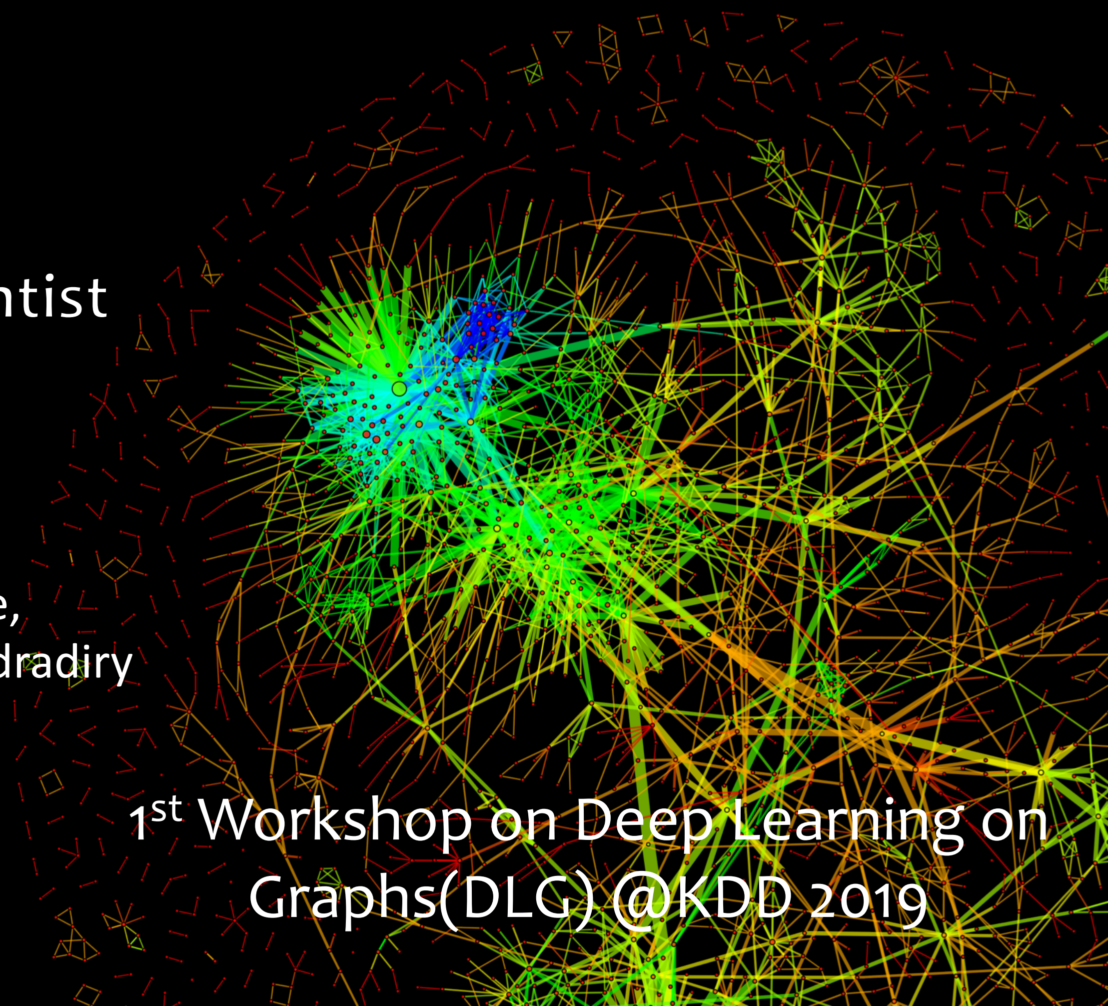
role2vec: Role-based Network Embeddings

Nesreen K. Ahmed

Senior Staff Research Scientist
Intel Labs, California

Ryan Rossi, John Boaz Lee, Ted Willke,
Rong Zhou, Xiangnan Kong, Hoda Eldradiry



1st Workshop on Deep Learning on Graphs(DLG) @KDD 2019

Representation Learning in Graphs

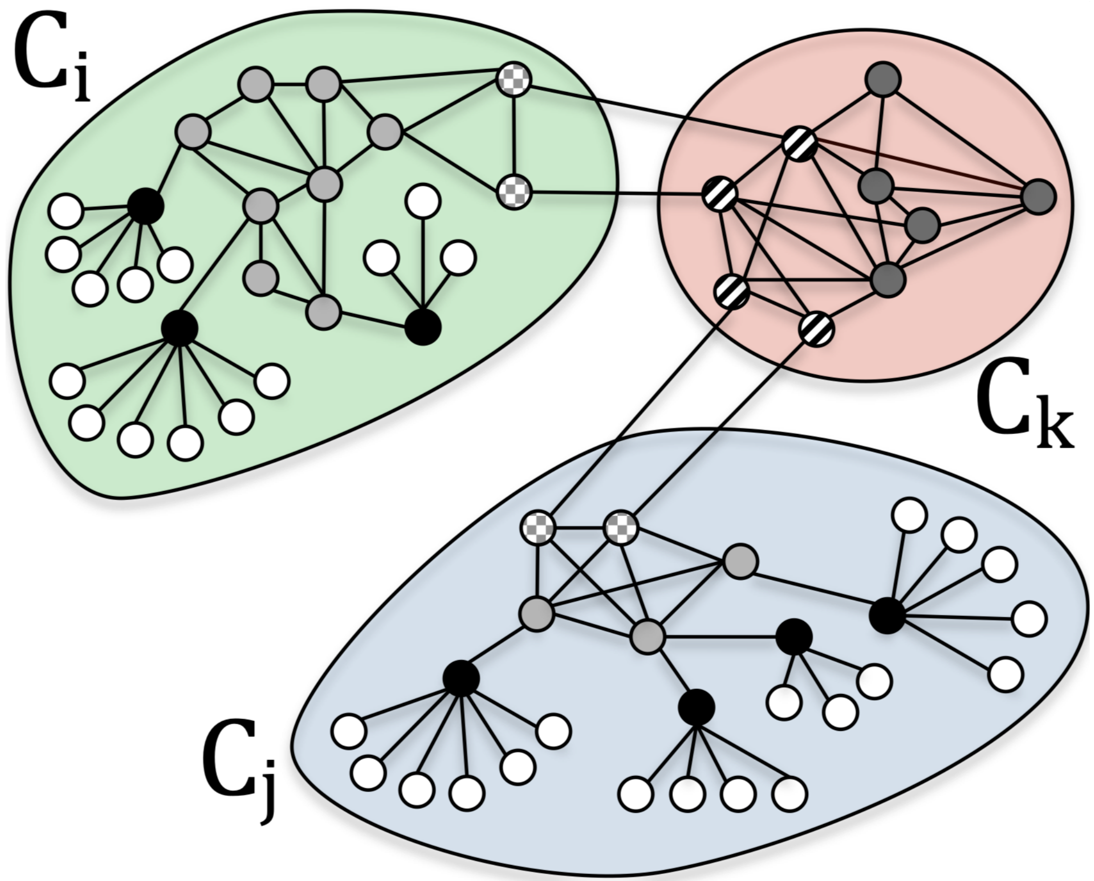
- **Goal:** Learn representation (features) for a set of graph elements (nodes, edges, etc.)

$$\text{Given } G = (V, E)$$
$$\text{Learn a function } f : V \rightarrow \mathbb{R}^d$$

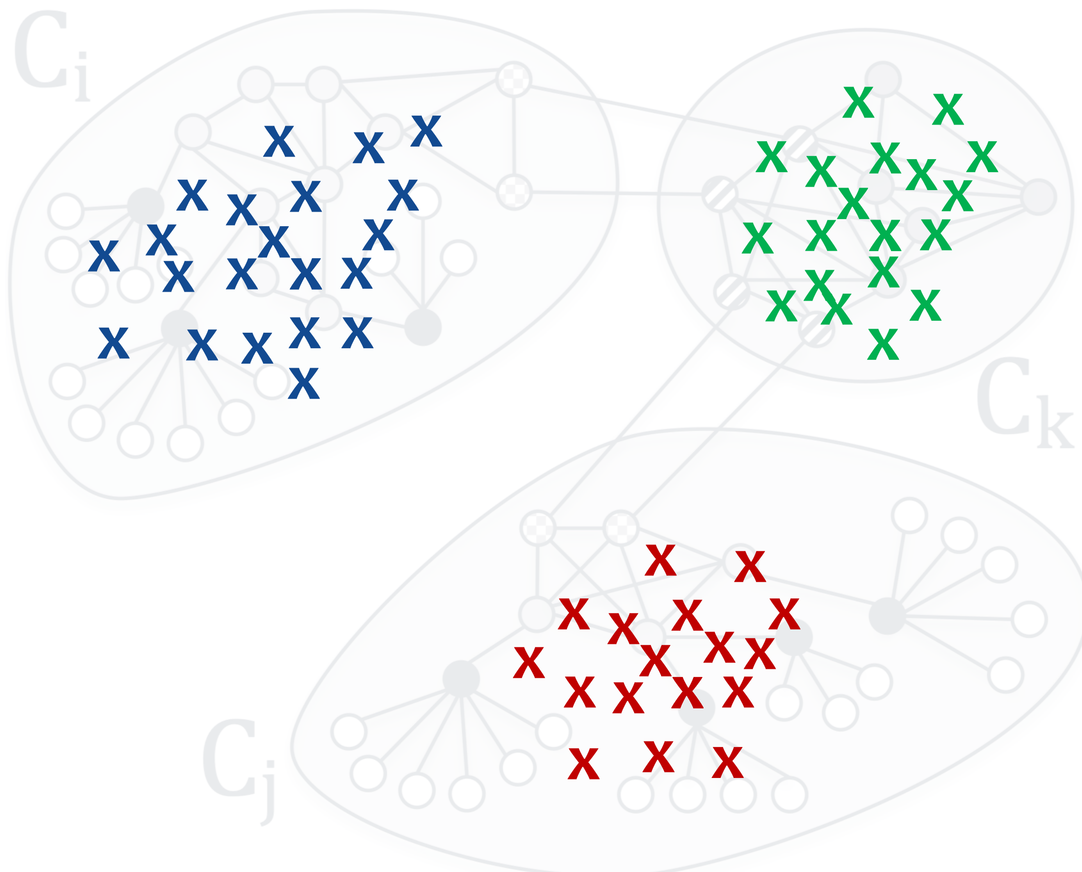
- **Key intuition**: Map the graph elements (e.g., nodes) to the d-dimension space
- Use the features for any downstream prediction task

Many examples: deepWalk, node2vec, GCN … etc

Two Complimentary Notions in Graphs: Proximity vs. Structural Similarity



Most Existing Work Focus on Modeling Graph Proximity



No guarantee that nearby vertices are ***structurally similar***

e.g., Deepwalk, GraRep, node2vec, Line, etc.

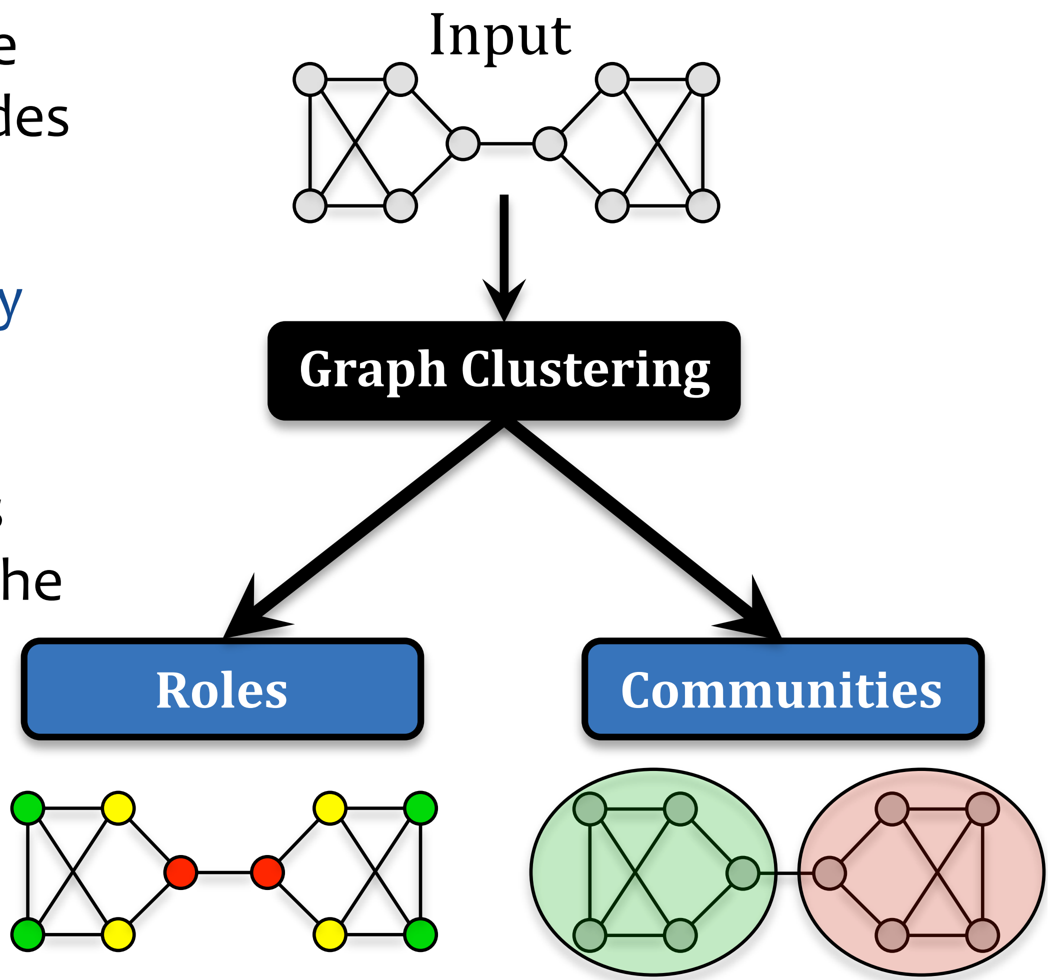
Communities and Roles

Roles are sets of nodes that are more structurally similar to nodes inside the set than outside

→ Based on structural similarity

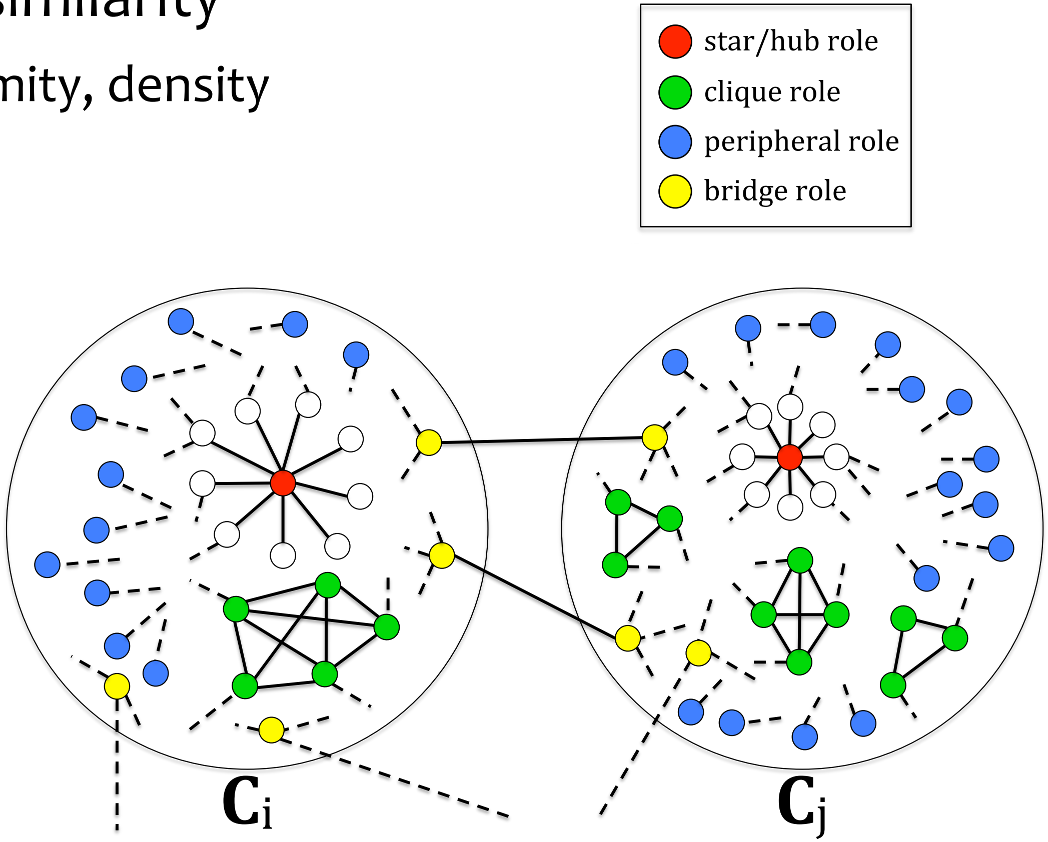
Communities are sets of nodes with more connections inside the set than outside.

→ Based on proximity, density



Roles and Communities are *Complimentary*

- Roles based on structural similarity
  - Communities based on proximity, density
- Roles globally distributed
  - Communities are local
- Roles generalize
  - Communities do not generalize across graphs (for graph transfer learning tasks)



Problem:
Learn Role-based Embeddings

Goal: Find d-dimensional embeddings of nodes that preserve *structural* similarity

Based on structural properties of nodes + attributes (if any)

Properties warranted by approach:

- General & unifying framework
- Methods generalized via framework are representationally more powerful
- Space-efficient

role2vec: Learning Role-based Graph Embeddings

- **Step 1:** Mapping vertices to vertex-roles.

$$\Phi \; : \; \mathbf{x} \; \rightarrow \; w$$

- **Step 2:** Sample Feature-based/Attributed Random walks.

A feature-based/attributed walk of length *L* is a sequence of adjacent vertex-roles

$$\Phi(\mathbf{x}_{v_0}), \ldots, \Phi(\mathbf{x}_{v_t}), \Phi(\mathbf{x}_{v_{t+1}}), \ldots, \Phi(\mathbf{x}_{v_{L-1}})$$

induced by a randomly chosen sequence of indices

$$(v_t : t = 0, 1, ..., L-1)$$

generated by a random walk of length *L* and a function Φ

- **Step 3:** Model the conditional probability that relate each vertex-role to the roles of its context.

$$\mathbb{P}\Big[\Phi\langle\mathbf{x}_{c_i}\rangle|\Phi\langle\mathbf{x}_i\rangle\Big] = \prod_{j\in c_i} \mathbb{P}(\Phi\langle\mathbf{x}_j\rangle \; |\Phi\langle\mathbf{x}_i\rangle)$$

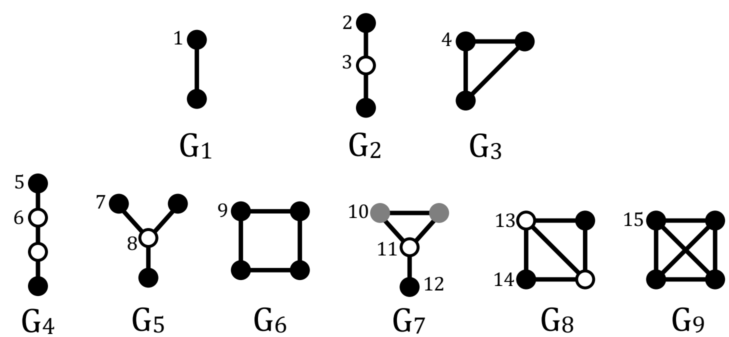
Generalize Existing Methods via Framework

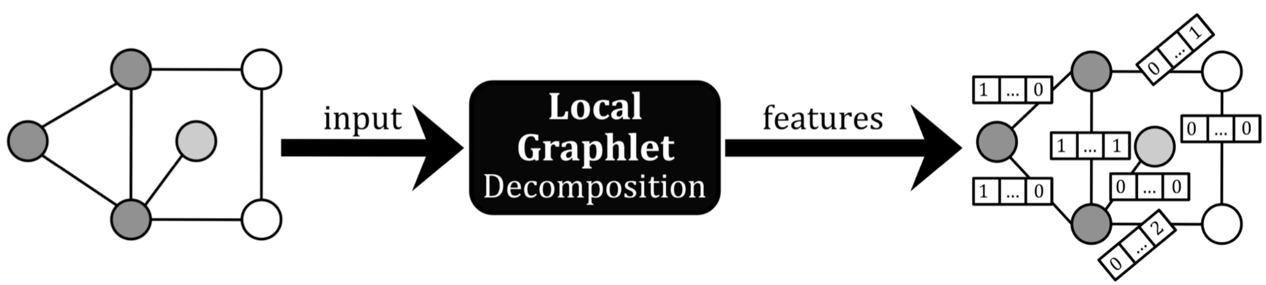
- Embeddings can be learned using basic Skip-gram model
- Other existing or future RW-based embedding methods can be easily generalized via the proposed framework

Examples: DeepWalk, node2vec, metapath2vec, struct2vec, and deep graph models, e.g., GRAM

We use small subgraphs called motifs as structural features

$$\Phi(\mathbf{x}) = x_1 \circ x_2 \circ \cdots \circ x_K$$





Prediction

Average improvement in AUC of 16.5%

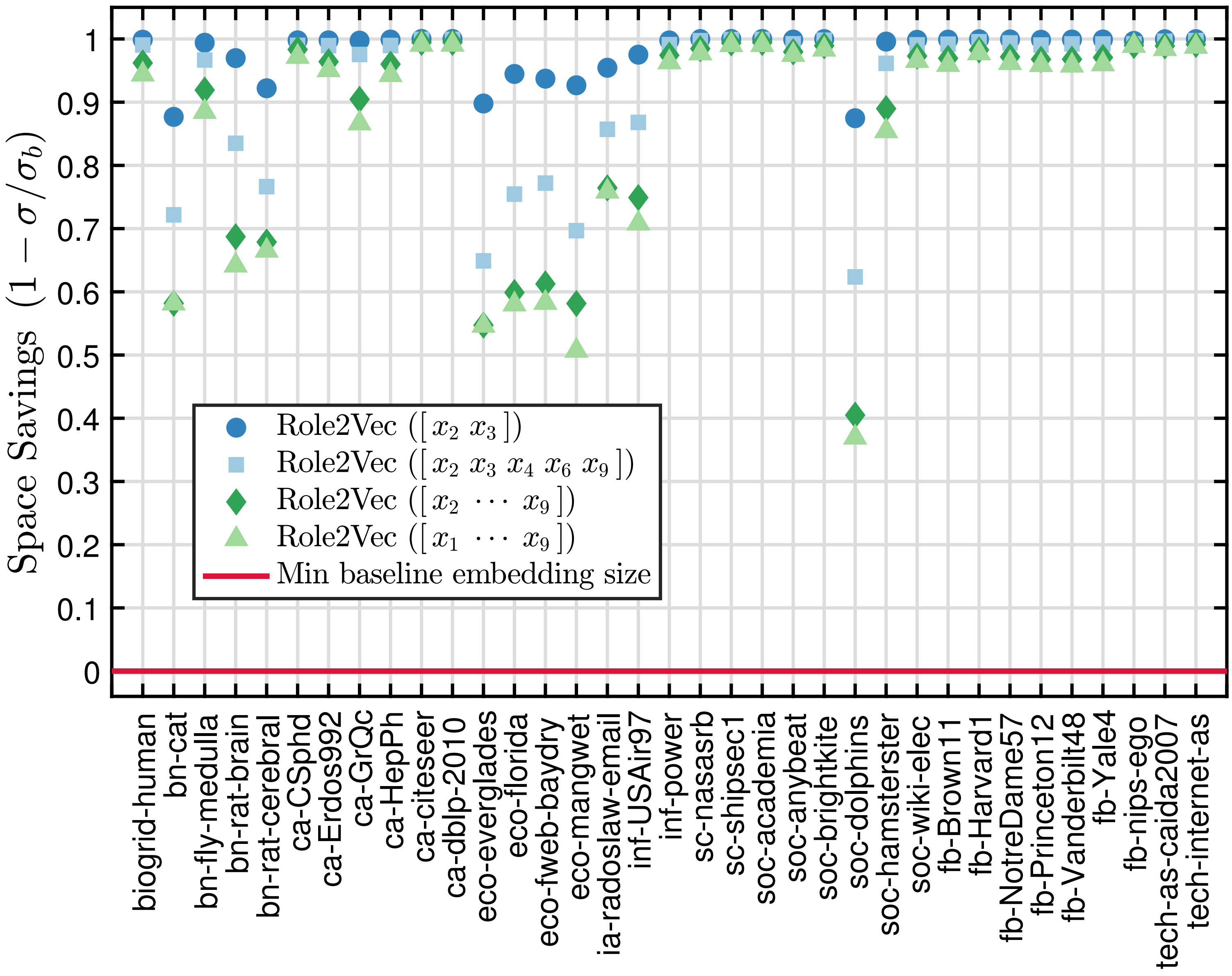
Experimental setup: 10-fold cross-validation, repeated for 10 random trials, D=128, Predict link existence via Logistic Regression

Table 1: AUC scores for various methods using $(\alpha_i + \alpha_j)/2$. **N2V**=node2vec, **DW**=DeepWalk and **S2V**=struc2vec.

GRAPH	R2V	R2V-DW	N2V	DW	LINE	S2V
bn–cat	**0.710**	0.688	0.627	0.627	0.672	0.669
bn–rat–brain	**0.748**	0.731	0.716	0.716	0.691	0.729
bn–rat–cerebral	**0.867**	0.846	0.813	0.811	0.709	0.858
ca–CSphd	**0.838**	**0.838**	0.768	0.735	0.620	0.791
eco–fweb–baydry	**0.681**	0.656	0.655	0.627	0.660	0.623
ia–radoslaw–email	**0.867**	0.847	0.756	0.745	0.769	0.857
soc–anybeat	**0.961**	0.960	0.854	0.848	0.850	0.883
soc–dolphins	**0.656**	0.597	0.580	0.498	0.551	0.590
fb–Yale4	**0.793**	**0.793**	0.742	0.728	0.763	0.758
web–EPA	**0.926**	0.925	0.804	0.738	0.768	0.861

Space-Efficiency

role2vec requires on average 850x less space



Conclusion

- Introduced notion of feature-based/attributed walks
- Proposed a generic framework for learning role-based embeddings based on this notion
- Learns universal functions that can generalize across networks/graphs → Useful for inductive & graph-based transfer learning
- role2vec achieves a mean gain in AUC of 16.5% while requiring 850x less space than existing methods

References

- Efficient estimation of word representations in vector space. ICLR 2013 [Mikolov et. al]
- A Framework for Generalizing Graph-based Representation Learning Methods. arXiv:1709.04596 2017 [Ahmed et. al]
- Role Discovery in Networks. TKDE 2015 [Rossi & Ahmed]
- A Higher-order Latent Space Network Model. AAAI 2017 [Ahmed, Rossi, Willke, Zhou]
- node2vec: Scalable Feature Learning for Networks. KDD 2016 [Grover, Leskovec]
- DeepWalk: online learning of social representations. KDD 2014 [Perozzi, Al-Rafou, Skiena]
- Efficient Graphlet Counting for Large Networks. ICDM 2015, [Ahmed et al.]
- Graphlet Decomposition: Framework, Algorithms, and Applications. J. Know. & Info. 2016 [Ahmed et al.]
- Network Motifs: Simple Building Blocks of Complex Networks. Science 2002, [Milo et al.]
- Uncovering Biological Network Function via Graphlet Degree Signatures. Cancer Informatics 2008 [Milenković-Pržulj]
- Graph Kernels. JMLR 2010, [Vishwanathan et al.]
- The Structure and Function of Complex Networks. SIAM Review 2003, [Newman]
- Biological network comparison using graphlet degree distribution. Bioinformatics 2007 [Pržulj]
- Efficient Graphlet Kernels for Large Graph Comparison. AISTAT 2009 [Shervashidze et al.]
- Local structure in social networks. Sociological methodology 1976, [Holland-Leinhardt]
- The strength of weak ties: A network theory revisited. Sociological theory 1983 [Granovetter]

Thank You!
Questions?

nesreen.k.ahmed@intel.com
http://nesreenahmed.com

1st Workshop on Deep Learning on Graphs(DLG) @KDD 2019