GLEE: Geometric Laplacian Eigenmap Embedding

Northeastern University Network Science Institute

Leo Torres — *leo@leotrs.com*

Kevin Chan — *kevin.s.chan.civ@mail.mil*

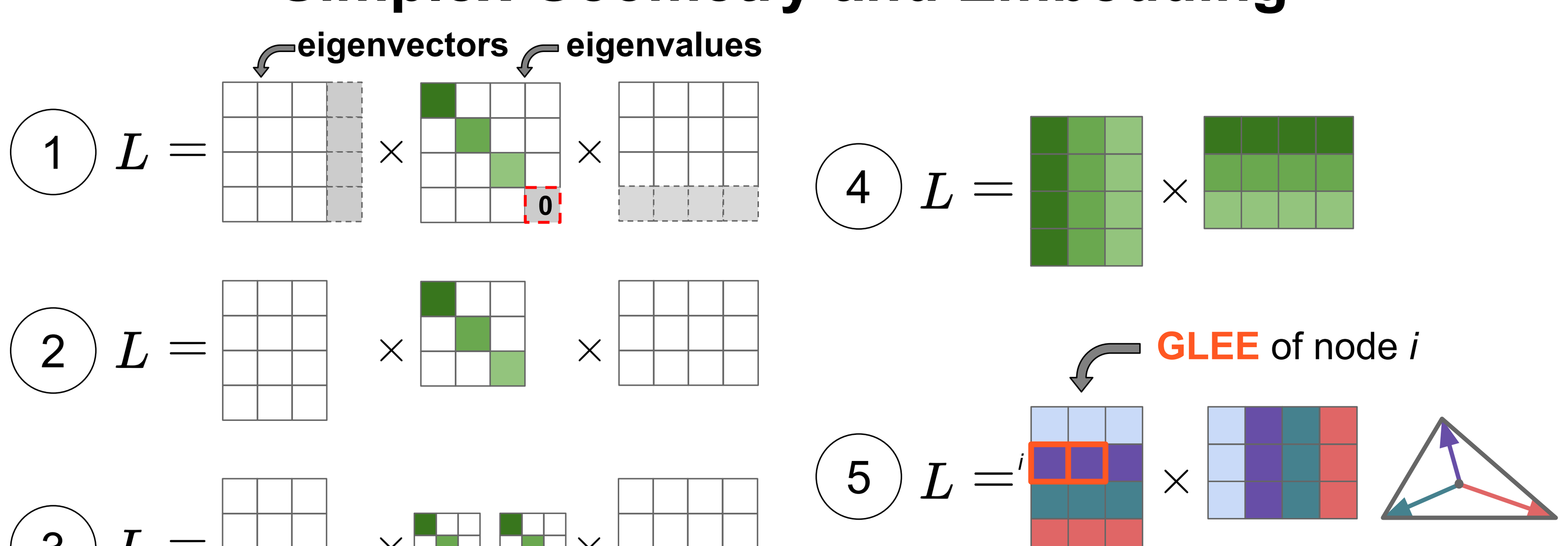
Tina Eliassi-Rad — *tina@eliassi.org*

DEVCOM Army Research Laboratory

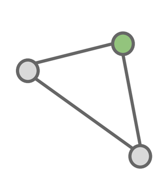
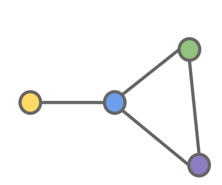
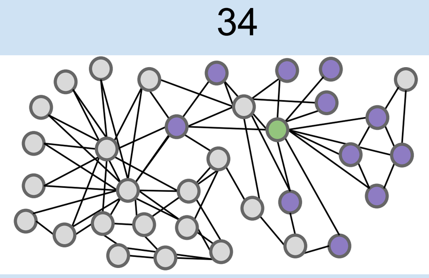
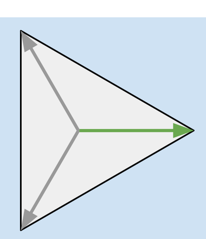
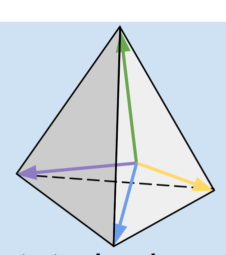
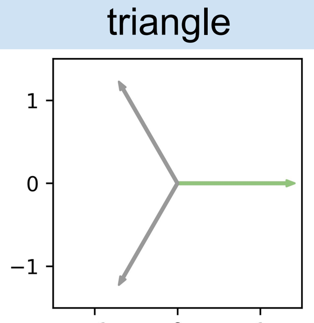
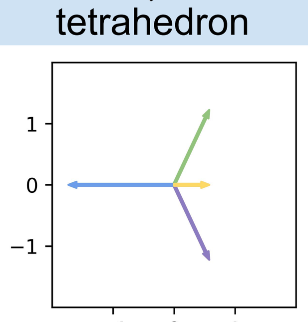
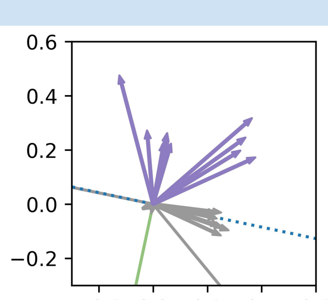
Abstract

- **Graph embedding** builds a low-dimensional representation of a graph.
- Popular in the literature is the **distance-minimization** assumption: if two nodes are close (in the graph), their embeddings must be close (in embedding space).
- We dispose of the distance-minimization assumption. Instead, our new method **Geometric Laplacian Eigenmap Embedding (GLEE)** builds **an embedding with geometric properties** by leveraging the so-called **simplex geometry** of graphs.
- Benefits of GLEE:
  - **Deterministic** and **interpretable**.
  - Great performance, especially in the case of **low clustering.**
  - **Robust to noise**: it can recover graph structure in the presence of a high percentage of noisy edges.

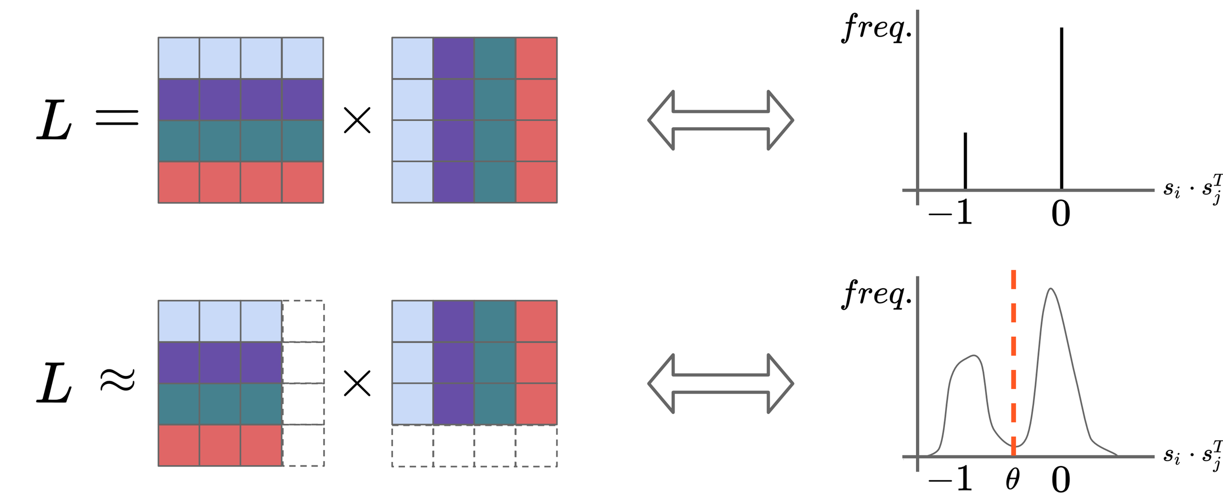
Simplex Geometry and Embedding



Devriendt, Van Mieghem, **The simplex geometry of graphs**, *Journal of Complex Networks*, https://doi.org/10.1093/comnet/cny036

Nodes	3	4	...	34
Example graph G			...	
Simplex of G	 triangle	 tetrahedron	...	33-dimensional simplex
GLEE of G			...	

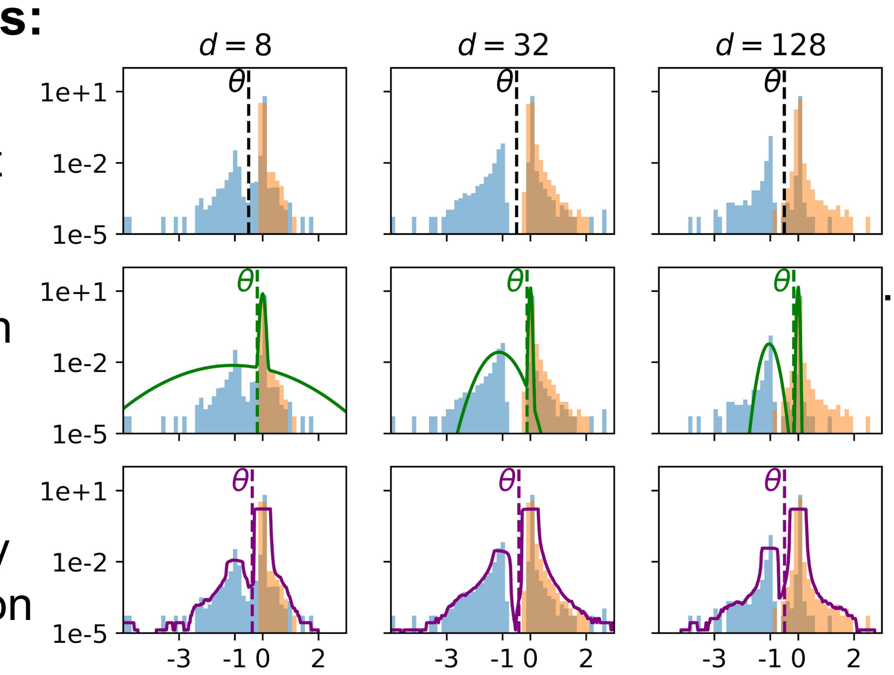
Graph reconstruction: a classification problem with extreme class imbalance



Problem: find optimal value of θ.

Three solutions:

1. Constant
2. Gaussian Mixtures
3. Density Estimation



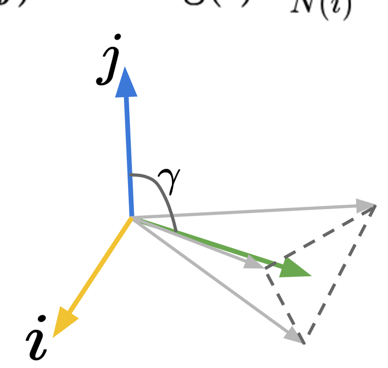
...plus **subsampling** to deal with class imbalance.

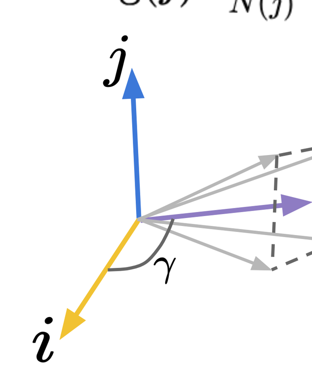
Link Prediction: interpreting the geometry of GLEE

GLEE (number of common neighbors, CN):

$$CN(i,j) = -\deg(i) C_{N(i)}^T \cdot s_j = -\deg(j) C_{N(j)}^T \cdot s_i$$

$C_{N(i)}$ is the center of mass of neighbors of i

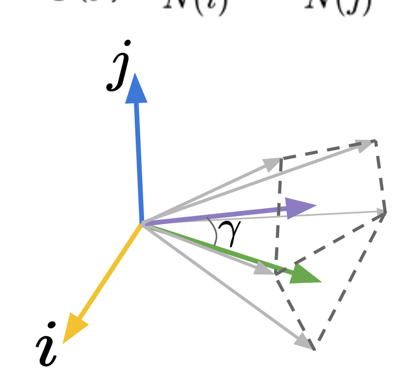




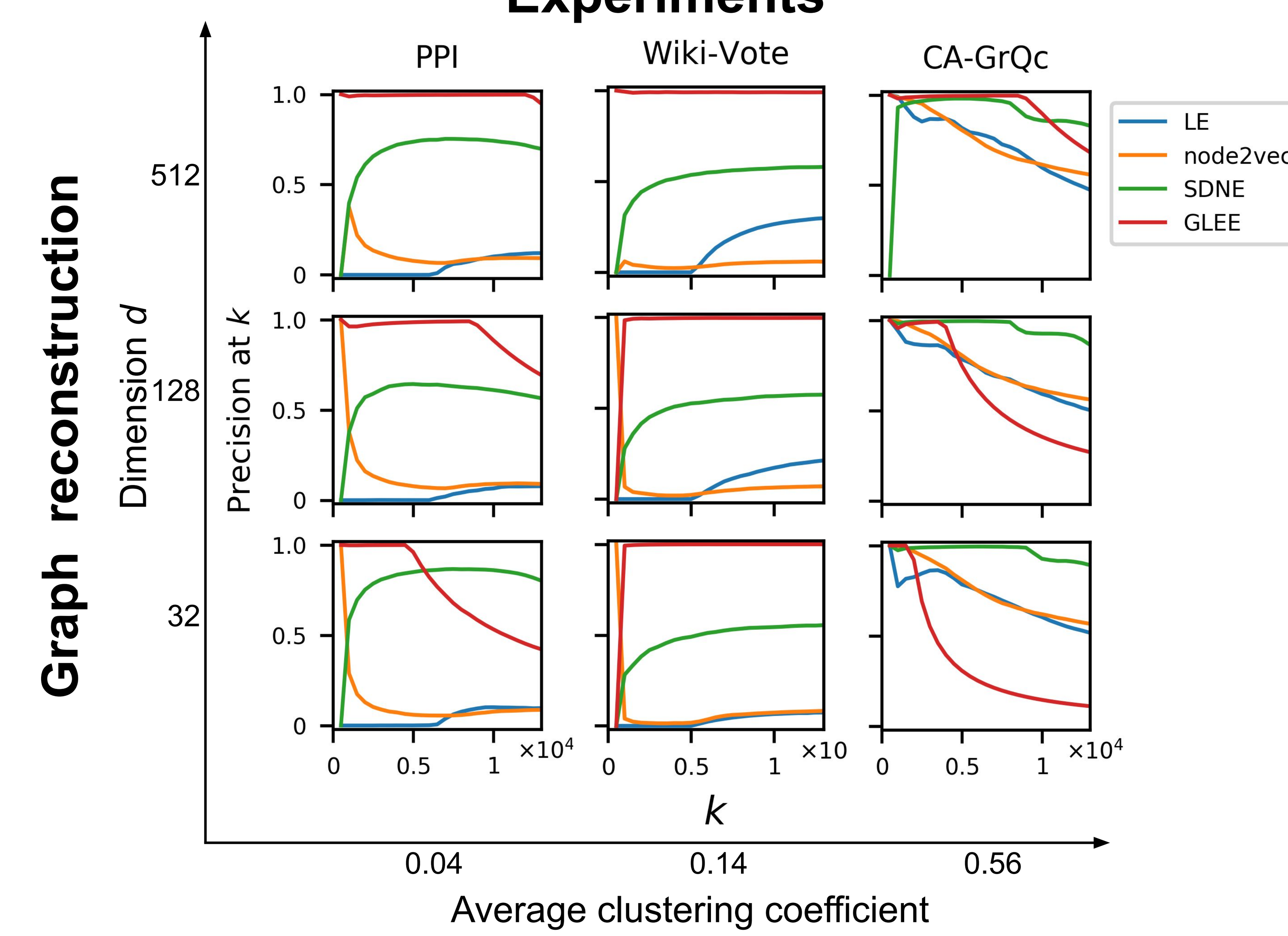
$C_{N(j)}$ is the center of mass of neighbors of j

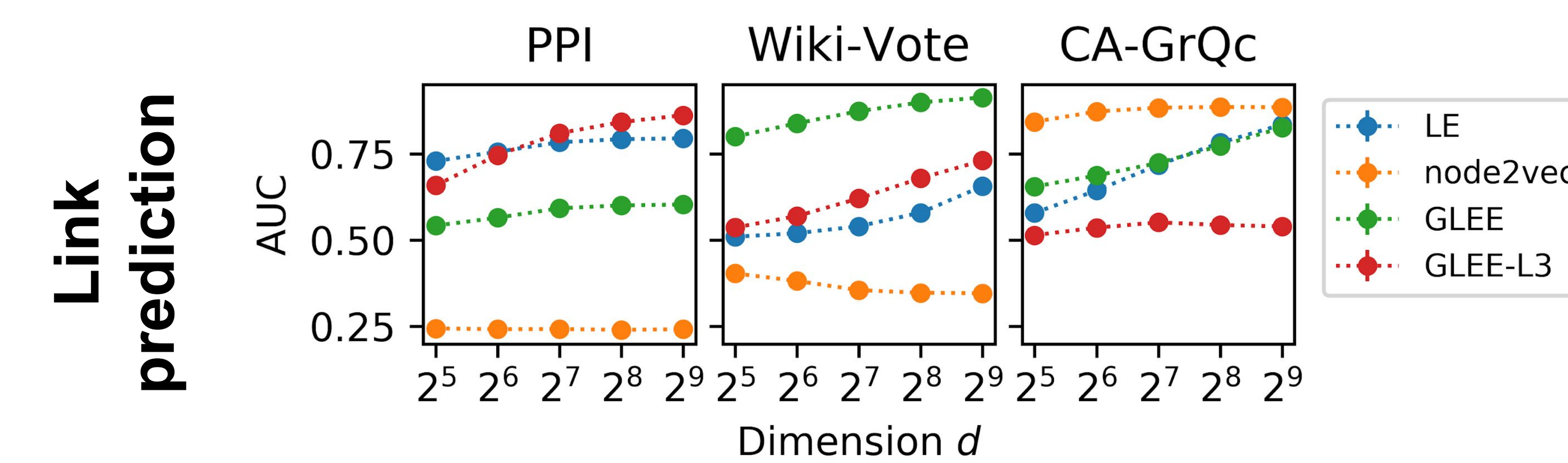
GLEE-L3 (number of paths of length 3, L3):

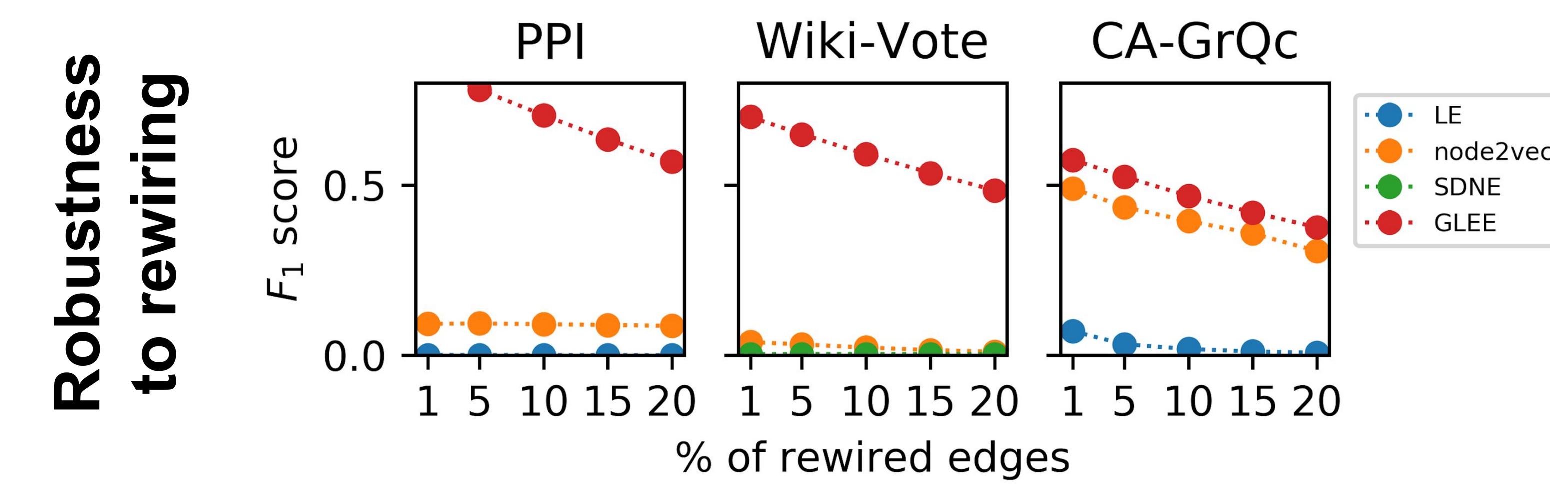
$$L3(i,j) = -\deg(i)\deg(j) C_{N(i)}^T \cdot C_{N(j)}^T + \sum_{k \in N(i) \cap N(j)} \|s_k\|^2$$



Experiments







Conclusions and Future Work

1. GLEE replaces distance-minimization with the direct encoding of graph structure in the geometry of the embedding space.
2. GLEE performs best when the graph has low clustering coefficient, and performance increases as the embedding dimension increases.
3. **What other geometric properties of embeddings can we utilize?**

Work supported by **NSF** CNS-1314603, **NSF** IIS-1741197, Army Research Laboratory **Cooperative Agreement** W911NF-13-2-0045.