Julien Delavande
juliendelavande@gmal.com

Lab session # Lab5
ALTEGRAD 2024

11/25/24

1 Question 1

Generalization of DeepWalk to Directed and Weighted Graphs:

- **Directed Graphs**

  Modify the random walk process to respect edge directions. At each step, choose the next vertex only from the **outgoing neighbors** of the current node:

$$w_{v_i}^{(j)} \sim \text{Uniform}(\text{OutNeighbors}(w_{v_i}^{(j-1)}))$$

- **Weighted Graphs**

  Adjust the probability of transitioning to a neighbor based on edge weights. For a current node v, the probability of moving to neighbor u is:

$$P(v \to u) = \frac{w_{vu}}{\sum_{z \in \text{Neighbors}(v)} w_{vz}}$$

  where w_{vu} is the weight of the edge between v and u.

This ensures that the walks adapt to the graph's directional and weighted characteristics.

2 Question 2

The two embedding matrices X_1 and X_2 are related by a **linear transformation** in the embedding space. Specifically, X_2 is obtained by applying a **reflection** to X_1 across the line $y = -x$. This transformation can be expressed as:

$$R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Each row of X_2 is given by:

$$\mathbf{v}_i^{(X_2)} = R \cdot \mathbf{v}_i^{(X_1)}$$

The reflection preserves the relative distances and neighborhood similarities of the embeddings, ensuring that both X_1 and X_2 represent the same structural information of the graph G. This invariance to transformations like rotations or reflections arises due to the stochastic nature of DeepWalk and the lack of a fixed orientation in the embedding space.

3 Question 3

In the GCN architecture implemented in Task 10, there are **two** message passing layers. Each layer aggregates information from the immediate neighbors of a node. Therefore, the **receptive field** of each node includes all nodes within **2 hops**.

- **For a given node i:** The maximal distance of nodes considered in the calculation of $\hat{Y}_i$ is **2**.

  In general, in a GCN architecture with k message passing layers:

- **The maximal number of edges separating node i from other nodes whose features influence $\hat{Y}_i$ is k.**

This is because each message passing layer allows information to propagate one additional hop in the graph.

4 Question 4

Computation of Z_1 for K_4 and S_4.

For the Complete Graph K_4:

- 1. Normalized Adjacency Matrix $\hat{A}$:

$$\hat{A} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

- 2. First Layer Output Z_0:

$$Z_0 = \text{ReLU}\left(\hat{A}XW_0\right) = \text{ReLU}\left(\begin{bmatrix} -0.8 & 0.5 \\ -0.8 & 0.5 \\ -0.8 & 0.5 \\ -0.8 & 0.5 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix}$$

- 3. Second Layer Output Z_1:

$$Z_1 = \text{ReLU}\left(\hat{A}Z_0W_1\right) = \text{ReLU}\left(\begin{bmatrix} 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \\ 0 & 0.3 & 0.25 \end{bmatrix}$$

For the Star Graph S_4:

- 1. Normalized Adjacency Matrix $\hat{A}$:

$$\hat{A} \approx \begin{bmatrix} 0.25 & 0.35 & 0.35 & 0.35 \\ 0.35 & 0.5 & 0 & 0 \\ 0.35 & 0 & 0.5 & 0 \\ 0.35 & 0 & 0 & 0.5 \end{bmatrix}$$

- 2. First Layer Output Z_0:

$$Z_0 = \text{ReLU}\left(\hat{A}XW_0\right) = \begin{bmatrix} 0 & 0.655 \\ 0 & 0.427 \\ 0 & 0.427 \\ 0 & 0.427 \end{bmatrix}$$

- 3. Second Layer Output Z_1:

$$Z_1 = \text{ReLU}\left(\hat{A}Z_0W_1\right) = \begin{bmatrix} 0 & 0.370 & 0.309 \\ 0 & 0.267 & 0.223 \\ 0 & 0.267 & 0.223 \\ 0 & 0.267 & 0.223 \end{bmatrix}$$

Observations:

- 1. For K_4: All nodes have identical representations in Z_1 due to the symmetry of the complete graph.
- 2. For S_4: The central node (first row) has higher activation values in Z_1, reflecting its higher connectivity, while peripheral nodes have similar, lower values.

Effect of Random Features:
If the node features X were sampled from a random uniform distribution, the symmetry in K_4 would break, resulting in diverse representations in Z_1. For S_4, the representations would still reflect the graph's topology but would also incorporate variability from the random features.

References

Annexes

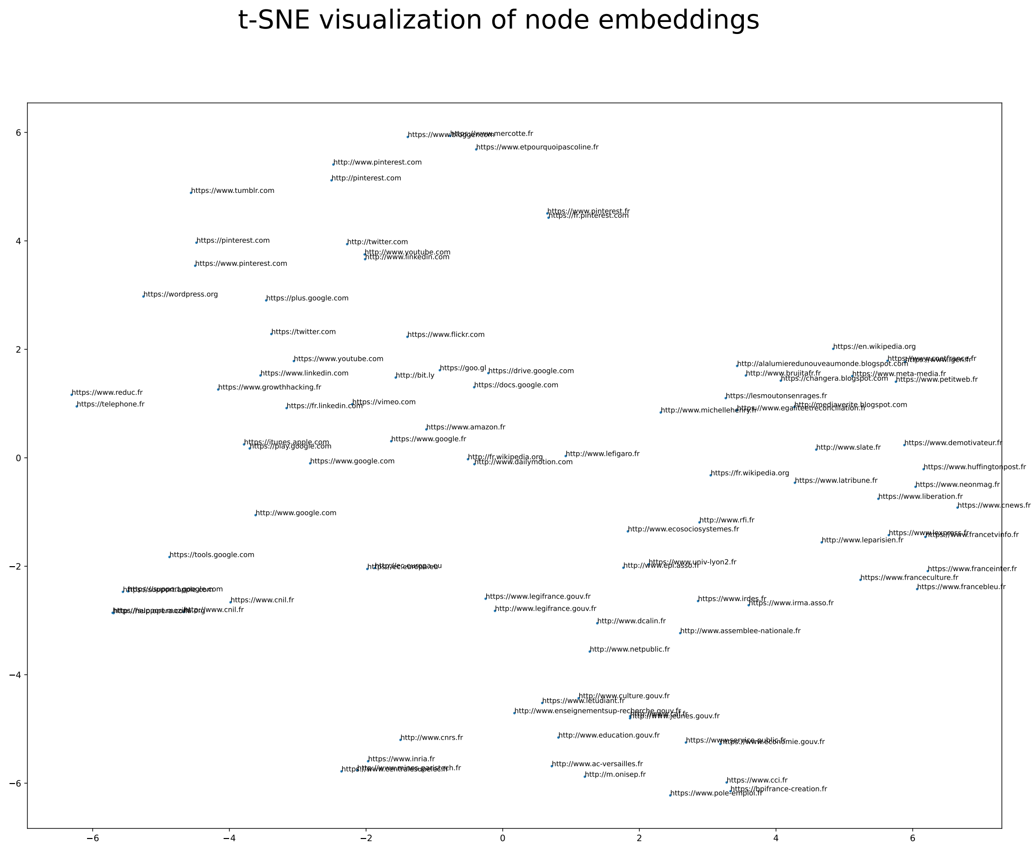

Figure 1: DeepWalk, French Web T-SNE visualization of node embeddings.

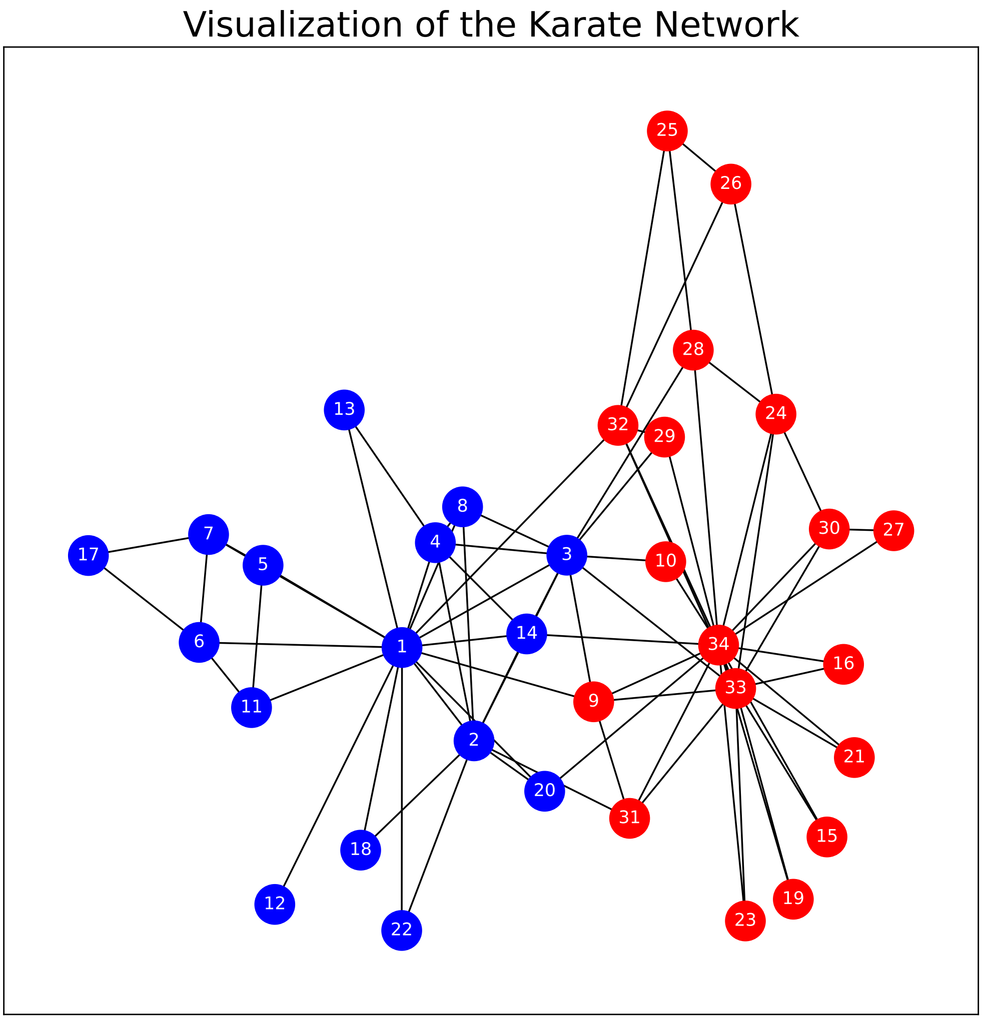

Figure 2: Karate Network visualization.

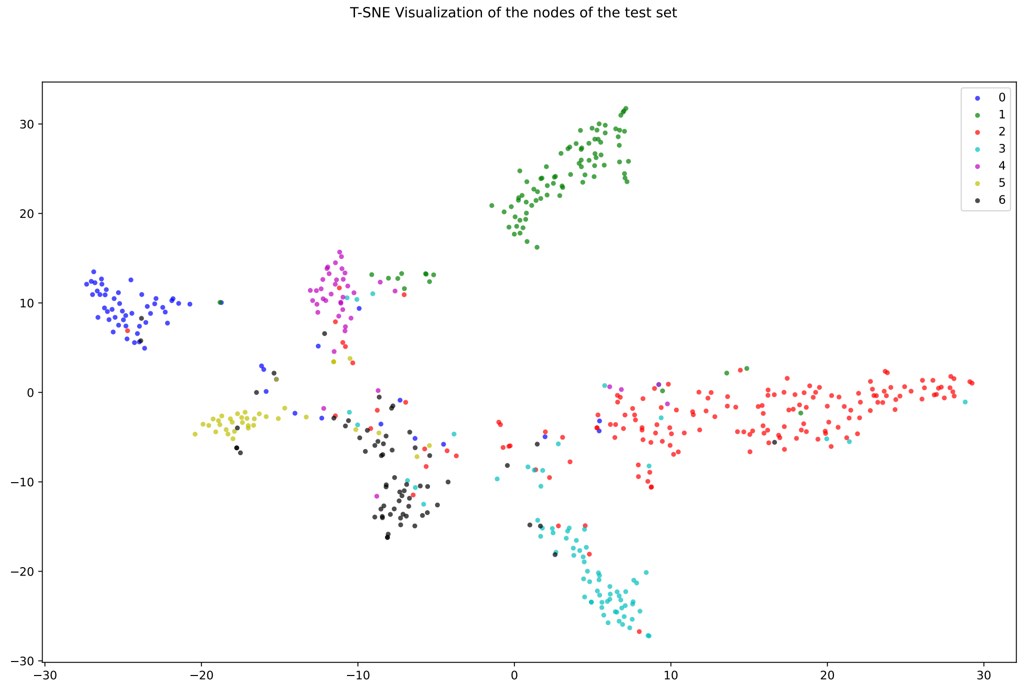

Figure 3: Cora Dataset, output of the second message passing layer, T-SNE visualization.