Celine Hajjar
celinehajjar98@gmail.com

Lab session # 5
ALTEGRAD 2021

01/24/22

Note: In order to run the code provided with this report, please change the directory specified in the `os.chdir()` function at the beginning of every code file.

Question 1

The DeepWalk algorithm is supposed to give insight on the localized structures of the graph. Hence, the node representations generated by the algorithm in this case should reflect the clear separation of the connected components. We would expect the cosine similarities to be high for nodes within a connected component, and low for nodes in different connected components.

Question 2

If we plot the points represented by these two embeddings in a 2D space, we find that the two sets of points are related by a 90-degree rotation (mathematically, $X_2^T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} X_1^T$).
Given that rotations preserve the distances among points in the same set, we can say that the two embedding matrices convey the same structural information.

Question 3

After the first message-passing layer, the feature vector of each node is the sum of the feature vectors of its direct neighbors. Similarly, after the second layer, every feature vector is the sum of the feature vectors of the direct neighbors as well as their neighbors. As such, using more than one message-passing layer allows us to take into account the extended neighborhood of a node when calculating its feature vector. However, if the aforementioned neighborhood were too large (if for example the number of message-passing layers exceeds the diameter of the graph), the network would fail to capture local structures.

Question 4

Let G be a P_4 path graph. Its adjacency matrix would be: $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ and in its normalized form:

$$\hat{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{2} \end{pmatrix}$$

If $X = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, we have: $\hat{A}XW^0 = \begin{pmatrix} 0.4541 & -0.1816 \\ 0.5374 & -0.215 \\ 0.5374 & -0.215 \\ 0.4541 & -0.1816 \end{pmatrix}$ and $Z^0 = \begin{pmatrix} 0.4541 & 0 \\ 0.5374 & 0 \\ 0.5374 & 0 \\ 0.4541 & 0 \end{pmatrix}$.
So:

$$Z^1 = f(\hat{A}Z^0W^1) = \begin{pmatrix} 0.1339 & 0 & 0.3572 & 0.2232 \\ 0.1631 & 0 & 0.435 & 0.2719 \\ 0.1631 & 0 & 0.435 & 0.2719 \\ 0.1339 & 0 & 0.3572 & 0.2232 \end{pmatrix}$$

We notice that the first and last node have the same feature vector, and the same goes for the two center nodes. This reflects the fact that the feature vectors can separate inner nodes from outer nodes in the path graph.