Finding the Highest Weighted (K-Core*Clustering_Coefficient) Node in a Protein-Protein Interaction Network

In this assignment your goal is to find the k-core values and clustering coefficients of all proteins in a given protein-protein interaction network. The interaction network is given as a set of undirected edges that comprise a graph. A k-core in a graph is a subgraph in which all the nodes in that subgraph have at least degree k. In other words each node in a k-core has at least k immediate neighbors. Use the O(n^3) algorithm described in class to find the highest k-core a node is a member of (the value of k will be the k-core value of that node). Use your implementation to find the k-core values of the nodes in the following human protein-protein interaction network.

http://www.ceng.metu.edu.tr/~tcanceng465_f1415/Assignments/HPRD_PPI.txt

Compute the clustering coefficient of each node in the network. Use the following formula to compute the clustering coefficient:

\[ cc = \frac{2e}{d(d-1)} \]

where e is the total number of edges between the neighbors of the node and and d is the number of neighbors of the node.

For each node, compute the product of k-core value and the clustering coefficient, and report the top 10 nodes in the graph with respect to this measure along with its k-core value and clustering coefficient.

Example output:
ABCD 7 0.76
EFGH 7 0.74
IJKL 6 0.8
.....
.....

Hint: Using an adjacency matrix/list representation for the PPI network will make things easier.

Notes: The network data is given as a list of undirected edges. Each line of the file represents an edge that connects a pair of tab separated protein ids.
**Deliverables:** A short document which contains the output and your source code.

**Submission:** Submit the deliverables via ODTU-CLASS.

**Late Submission Policy:** Penalty: 20 points per day.