Ben KABONGO B.
ben.kabongo_buzangu@ens-paris-saclay.fr

Lab session # 2
ALTEGRAD 2023

10/22/23

1 Question 1

The maximum number of edges (triangles) in an n-node undirected graph without self-loops is the number of edges (triangles) in a complete graph with the same number of nodes. Each node is connected to the $n-1$ other nodes by an edge.

1.1 Number of edges

There are n ways to choose the first node, $n-1$ for the second node and $2! = 2$ ways to swap the two nodes to have the same edge. So there are $\frac{n(n-1)}{2}$ edges.

1.2 Number of triangles

There are n ways to choose the first node, $n-1$ for the second, $n-2$ for the third and $3! = 6$ ways of swapping the three nodes to get the same triangle. So there are $\frac{n(n-1)(n-2)}{6}$ triangles.

2 Question 2

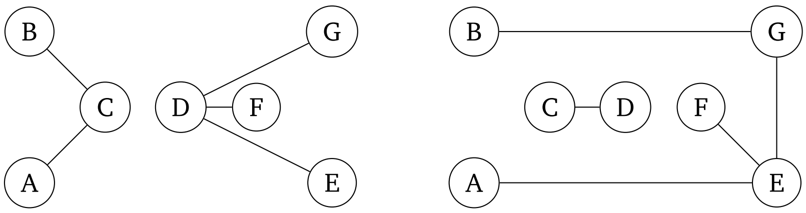

Figure 1: Graphs with the same degree distribution. Graph G1 on the left, G2 on the right.

Graphs G1 and G2 in Figure 1 have the same degree distribution. They have 5, 1 and 1 nodes with a degree of 1, 2 and 3 respectively. However, there is clearly no bijection that can associate nodes in such a way as to achieve the same node connectivity in both graphs. As the related components of the two graphs are completely different, no such bijection exists.
Consequently, two graphs with the same distribution of node degrees are not necessarily isomorphic to each other.

3 Question 3

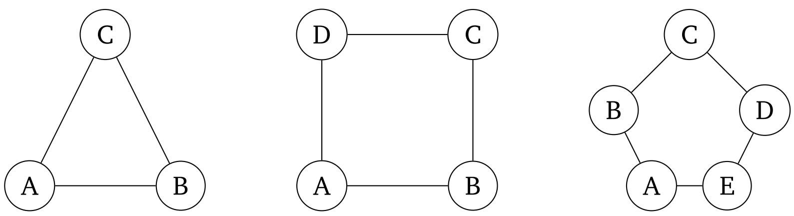

Figure 2: Cycle graphs. C_3 (triangle) on the left, C_4 (square) in the center, C_5 (pentagon) on the right.

- **For** $n = 3$, there are no open triplets. The global clustering coefficient is 1.
- **For** $n > 3$, there are no closed triplets. The global clustering coefficient is 0.

4 Question 4

The smallest eigenvalue of L_{rw} is 0, and the vector associated with this value is $u1 = \mathbf{1}$, with all components equal to 1. Hence,

$$\sum_i \sum_j A_{ij} \left([u1]_i - [u1]_j\right)^2 = \sum_i \sum_j A_{ij} \left(1-1\right)^2 = 0$$

5 Question 5

Let G_1 be the left-hand graph and G_2 the right-hand graph. We have that $m_1 = m_2 = m = 14$.
Two clusters in G_1: blue (1) and orange (2).

- $l_1 = 6$. The degrees of v_1, v_2, v_3 and v_4 in the cluster subgraph are all equal to 3. Then $d_1 = 3+3+3+3 = 12$.

$$q_1 = \frac{l_1}{m} - (\frac{d_1}{2*m})^2 = \frac{6}{14} - (\frac{12}{2\times 14})^2 = 0.245$$

- For the second cluster we also have, $l_2 = 6$, $d_2 = 12$, $q_2 = q_1$

The modularity of G_1 is therefore:

$$Q_1 = q_1 + q_2 = 0.490$$

Two clusters in G_2: blue (1) and orange (2).

- $l_1 = 5$. The degrees of v_1, v_2, v_4, v_6 and v_8 in the cluster subgraph are 2, 2, 3, 2 and 1. Then $d_1 = 2+2+3+2+1 = 10$.

$$q_1 = \frac{l_1}{m} - (\frac{d_1}{2*m})^2 = \frac{5}{14} - (\frac{10}{2\times 14})^2 = 0.230$$

- $l_2 = 2$. The degrees of v_3, v_5 and v_7 in the cluster subgraph are 1, 2 and 1. Then $d_2 = 1+2+1 = 4$.

$$q_1 = \frac{l_2}{m} - (\frac{d_2}{2*m})^2 = \frac{2}{14} - (\frac{4}{2\times 14})^2 = 0.122$$

The modularity of G_2 is therefore:

$$Q_2 = q_1 + q_2 = 0.352$$

6 Question 6

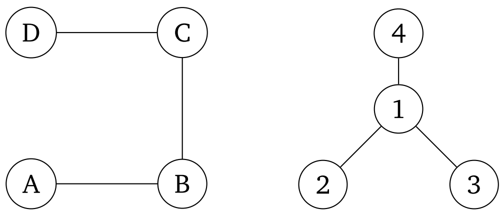

Figure 3: $n = 4$ nodes. Path graph on right, star graph on left.

Consider the graphs P_4 and S_4, as shown in Figure 3.
Let's calculate the shortest paths between each pair of nodes for both graphs.

	A	**B**	**C**	**D**
A	0	1	2	3
B	1	0	1	2
C	2	1	0	1
D	3	2	1	0

Table 1: Table of shortest paths of graph P_4.

	1	**2**	**3**	**4**
1	0	1	1	1
2	1	0	2	2
3	1	2	0	2
4	1	2	2	0

Table 2: Table of shortest paths of graph S_4.

The length of the shortest paths varies between 0 and 3 inclusive, for both graphs.
Let's count the occurrence of each length for each graph. We obtain the following tables:

Length	**0**	**1**	**2**	**3**
Occurence	4	6	4	2

Table 3: Path length occurrences for P_4.

Length	**0**	**1**	**2**	**3**
Occurence	4	6	6	0

Table 4: Path length occurrences for S_4.

We conclude that $f_{P_4} = \begin{bmatrix}4 & 6 & 4 & 2\end{bmatrix}$ and $f_{S_4} = \begin{bmatrix}4 & 6 & 6 & 0\end{bmatrix}$

- $k(P_4, P_4) = {f_{P_4}}^T f_{P_4} = 72$
- $k(P_4, S_4) = {f_{P_4}}^T f_{S_4} = 76$
- $k(S_4, S_4) = {f_{S_4}}^T f_{S_4} = 88$

7 Question 7

Each component of the vector f_G being a number of occurrences is ≥ 0.

$$k(G, G^{'}) = f_G^T f_{G'} = \sum_i f_{G\,i}^T f_{G'\,i} = 0 \iff \forall i, f_{G\,i}^T f_{G'\,i} = 0 \iff \forall i, f_{G\,i}^T = 0 \text{ or } f_{G'\,i} = 0$$

$k(G, G^{'}) = 0$ implies that the graphs G, $G^{'}$ are not made up of the same graphlets. In other words, if a graphlet is recognized in one of the graphs, it is not recognized in the other. The two graphs are then not similar. Graphlets (for $n = 3$) are graphs that are not similar to each other.

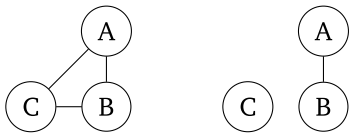

Figure 4: Examples of two graphs G_1 and G_2 such that $k(G_1, G_2) = 0$.

We have that $f_{G1} = \begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}$ and $f_{G2} = \begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix}$ and therefore $k(G_1, G_2) = 0$.