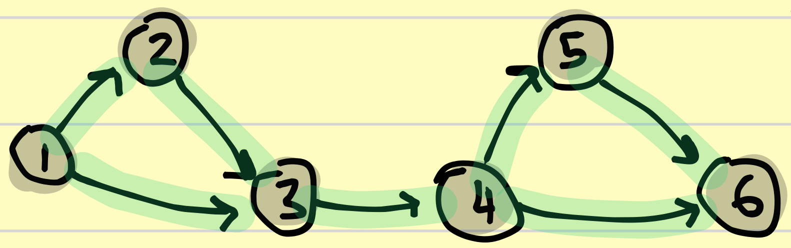
20251009_Business_Intelligence_WIN6

NETWORKS

A network is defined by a group of nodes and edges called GRAPH.

Network = GRAPH = $G(N \cup E)$ [N. Nodes / E. Edges]

Example:



E $\{(1\to 2), (2\to 3), (1,3), (3\to 4), (4\to 5), (5,6), (4\to 6)\}$

N $\{1, 2, 3, 4, 5, 6\}$

How to quantify & compare networks?
- Average Path Length
- Clustering Coefficient
- Degree Distribution
- LAPLACIAN MATRIX of GRAPH

1. APL. Average distance between nodes.

$$APL = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij} \qquad N: \text{n° Nodes}$$

$$APL = \frac{1}{6(6-1)} \cdot \Big[{}^{1}\big(\overset{d_{12}}{1} + \overset{d_{13}}{1} + \overset{d_{14}}{2} + \overset{d_{15}}{3} + \overset{d_{16}}{3}\big) + {}^{2}\big(\overset{d_{21}}{0} + \overset{d_{23}}{1} + \overset{d_{24}}{2} + \overset{d_{25}}{3} + \overset{d_{26}}{3}\big) + {}^{3}\big(\overset{d_{31}}{0} + \overset{d_{32}}{0} + \overset{d_{34}}{1} + \overset{d_{35}}{2} + \overset{d_{36}}{2}\big) + {}^{4}\big(\overset{d_{41}}{0} + \overset{d_{42}}{0} + \overset{d_{43}}{0} + \overset{d_{45}}{1} + \overset{d_{46}}{1}\big) + {}^{5}\big(\overset{d_{51}}{0} + \overset{d_{52}}{0} + \overset{d_{53}}{0} + \overset{d_{54}}{0} + \overset{d_{56}}{1}\big) + {}^{6}\big(0\big) \Big] = \ldots$$

- The best APL is the shortest.
- When you compare 2 or more networks, the one with shortest APL would be usually faster/more effective.

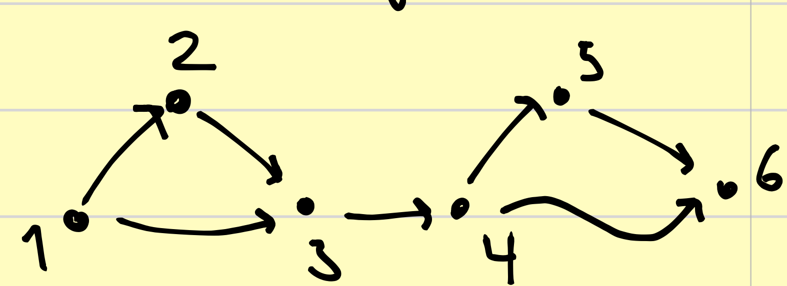
2. CC. Describes how good groups are created in the network.

$$\boxed{CC = \frac{1}{N}\sum_{i=1}^{N}\frac{2L_i}{k_i(k_i-1)}}$$

L_i: Number of relationships between the neighbours of node "i" (how friends of "i" are friends to each other)

k_i: Number of nodes

$$CC = \frac{1}{6}\left[\overset{1}{\left[\frac{2\cdot 1}{2(2-1)}\right]} + \overset{2}{[0]} + \overset{3}{[0]} + \overset{4}{\left[\frac{2\cdot 1}{2(2-1)}\right]} + \overset{5}{[0]} + \overset{6}{[0]}\right] = \frac{1}{3}$$



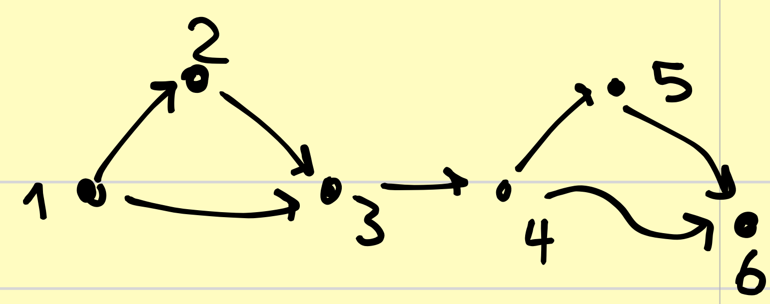
The higher the clustering coefficient, the better the communication between the elements of the network.

3. Laplacian Matrix. It contains all relevant info. of the network.

Is defined as $\boxed{L = D - A}$

$L \equiv$ Laplacian Matrix
$D \equiv$ Degree "
$A \equiv$ Adjacency "

$$D = \begin{cases} k_i & i = j \\ 0 & i \neq j \end{cases}$$



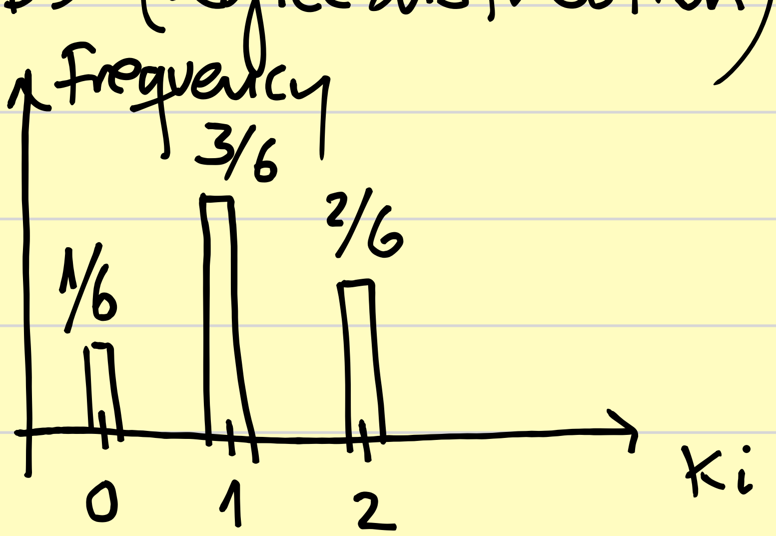
$$A = \begin{cases} 1 & \text{if connection btw } i \to j \\ 0 & \text{otherwise} \end{cases}$$

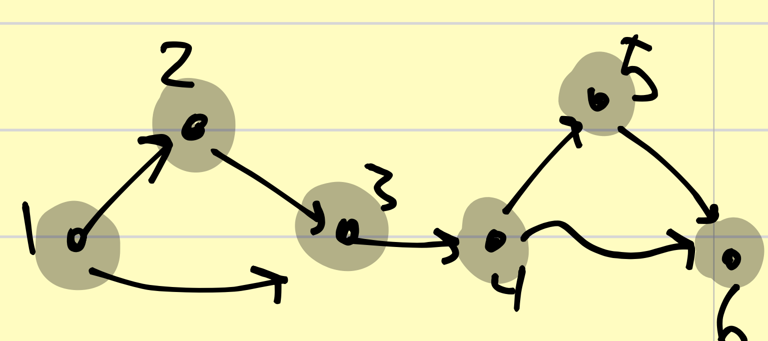
$$L = \begin{array}{c|cccccc} D & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 2 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 1 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} - \begin{array}{c|cccccc} A & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline & 0 & 1 & 1 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 1 & 1 \\ & 0 & 0 & 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 & 0 & 0 & 0 \end{array} =$$

$$L = \begin{array}{c|cccccc} 1 & 2 & -1 & -1 & 0 & 0 & 0 \\ 2 & 0 & 1 & -1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 & -1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 2 & -1 & -1 \\ 5 & 0 & 0 & 0 & 0 & 1 & -1 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

✓

4. DD (Degree distribution)





Network topology.

- The probability that a node with k neighbours attaches a new node in a random network is given by a Poisson distribution with parameter λ:

$$P_{Random}(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- The prob. that a node is connected to another with k neighbours in a real network is given by:

$$P_{Real}(X=k) = k^{-\gamma} \quad \gamma = \text{Exponent Degree}$$

$\boxed{\text{Network Science (Baràbasi, 2016)}}$