Clustering Coefficient

Prof. O-Joun Lee

Dept. of Artificial Intelligence,
The Catholic University of Korea
ojlee@catholic.ac.kr

네트워크 과학 연구실
NETWORK SCIENCE LAB

가톨릭대학교
THE CATHOLIC UNIVERSITY OF KOREA

Contents

- The clustering coefficient measures how connected a vertex's neighbors are to one another.
- The range is from 0 to 1 (from non-neighbor are connected to each other to all neighbors are fully connected).
- There are three types of clustering coefficient:
  - Local clustering coefficient.
  - Average clustering coefficient.
  - Global clustering coefficient.

Local Clustering Coefficient

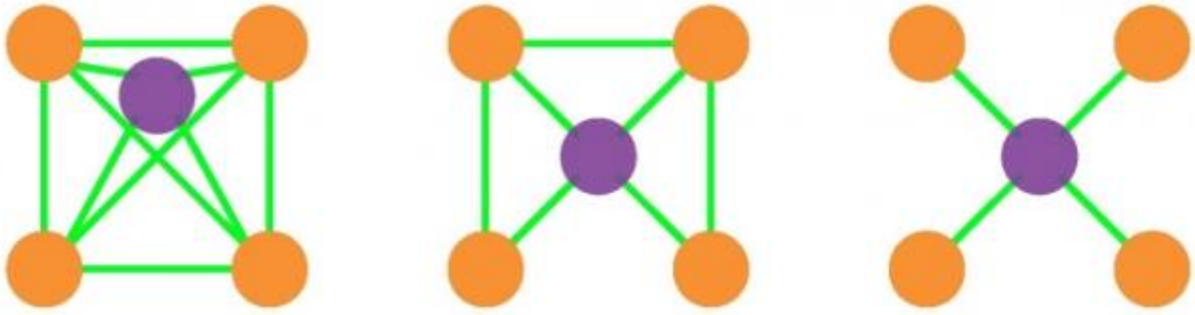
- How close its neighbours are to being a clique (complete graph).
- For a node i with degree d_i and L_i represents the number of edges between neighbors of node i.

  The local clustering coefficient C_i for a node i is defined as:

$$C_i = \frac{2L_i}{d_i(d_i - 1)}$$

50% chance that two neighbors of a node i are linked

neighbors of node i form a complete graph

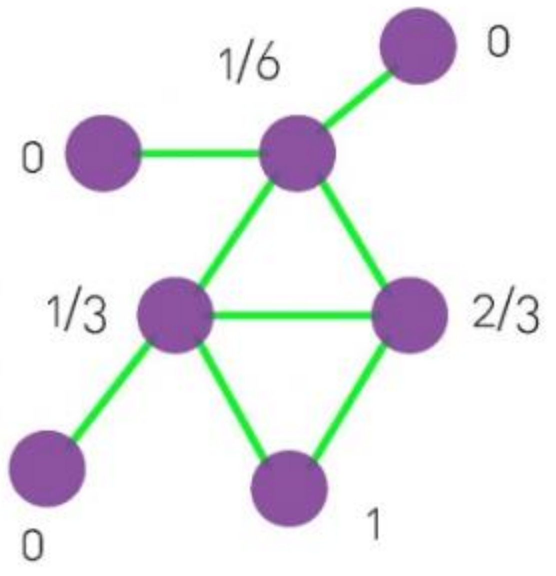


None of neighbors of node i link to each other

$C_i=1$ $C_i=1/2$ $C_i=0$

- The degree of clustering of a whole network is captured by the average clustering coefficient, namely $\langle C \rangle$, representing the average of all the local clustering coefficient C_i over all nodes $i = 1, \ldots, N$.

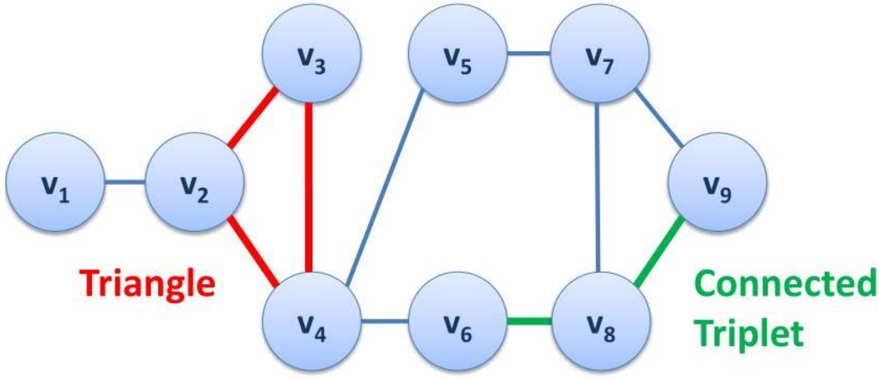
$$\langle C \rangle = \frac{1}{N} \sum_{i=0}^{N} C_i$$



$$\langle C \rangle = \frac{1}{7} * \left(0 + \frac{1}{6} + \frac{1}{3} + \frac{2}{3} + 1 + 0 + 0 \right) = 0.333$$

Global Clustering Coefficient

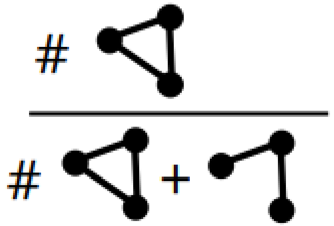
- The global clustering coefficient is based on triplets of nodes.
- A triplet consists of three connected nodes. A triangle therefore includes three closed triplets, one centered on each of the nodes.



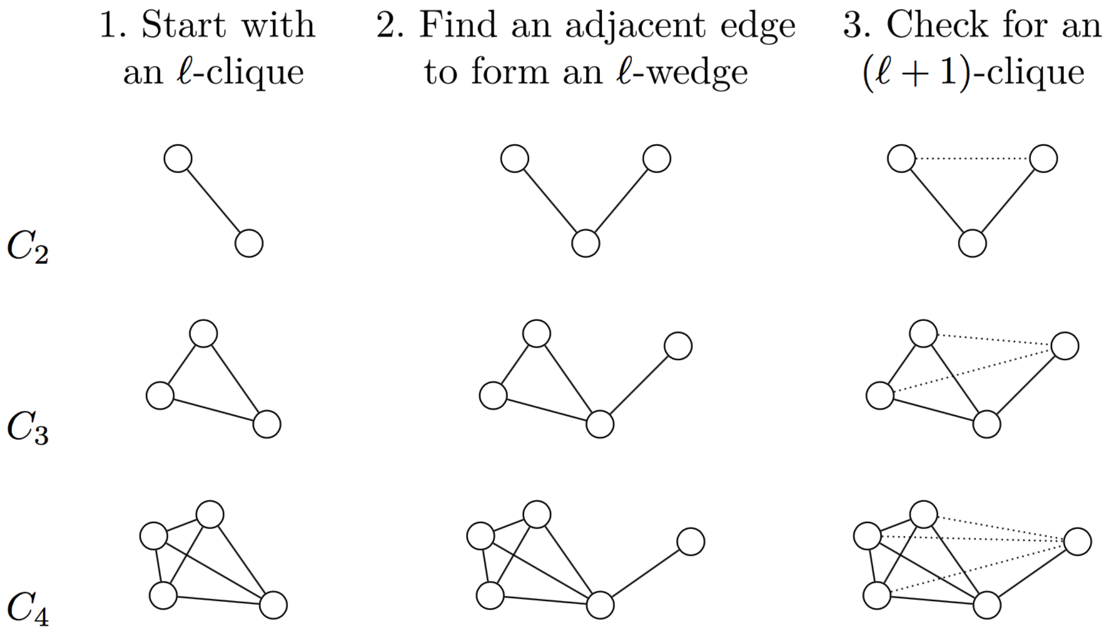
- The global clustering coefficient is the number of closed triplets over the total number of triplets (both open and closed)

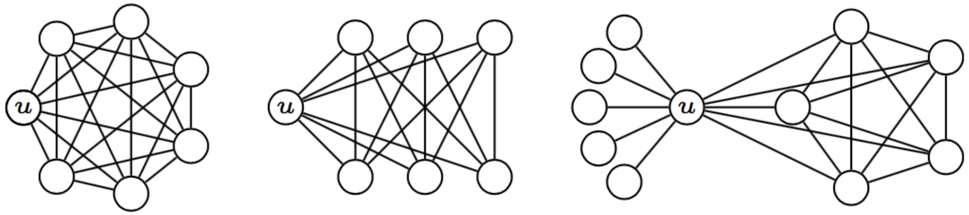
Closed triplets: $(\mathrm{v}_2, \mathrm{v}_3, \mathrm{v}_4)$, $(\mathrm{v}_7, \mathrm{v}_8, \mathrm{v}_9)$
Connected triplets: $(\mathrm{v}_6, \mathrm{v}_8, \mathrm{v}_9), \cdots$

$$C(G) = \frac{\#of\ closed\ triplets}{\#\ of\ connected\ triplets}$$



- The clustering coefficient can be extended to higher order structures with k-cliques.





$C_2(u)$	1	$\frac{d}{2(d-1)} \approx \frac{1}{2}$	$\frac{d-2}{4d-4} \approx \frac{1}{4}$
$C_3(u)$	1	0	$\frac{d-4}{2d-4} \approx \frac{1}{2}$
$C_4(u)$	1	0	$\frac{d-6}{2d-6} \approx \frac{1}{2}$

네트워크 과학 연구실
NETWORK SCIENCE LAB

가톨릭대학교
THE CATHOLIC UNIVERSITY OF KOREA