Names: Klaus Kades (UniID: fw448, MatrNr.: 3408463), Lucas-Raphael Müller (UniID: al413, Matr.Nr: 3205638), Melanie Schellenberg (UniID: gh400, Matr.Nr: 3146390), Shuhan Xiao (UniID: kg410, MatrNr: 3160697)

15.11. 2021

Complex Network Analysis

Assignment 3 "Random Graph Models and Statistical Characterizations"

Problem 3-1 Clustering Coefficient

1. total number of triangles

$$= \frac{N\cdot(N-1)\cdot(N-2)}{3}$$

Explanation: We can find each triangle in the graph by starting at each node i (N possibilities), finding its neighbouring nodes j (N-1 possibilities) and finding the neighbouring nodes k of node j (N-2 possibilities).* Since cyclic permutations don't matter here ($A\to B\to C$, $B\to C\to A$ and $C\to A\to B$ describe the same triangle), we divide by a factor 3, otherwise each ~~no~~ triangle would be counted three times, depending on the starting node. Here we also assume that the 'direction'/order matters (i.e. $A\to B\to C$ and $C\to B\to A$ describe different triangles), otherwise we would only count $\frac{N\cdot(N-1)\cdot(N-2)}{3}$ triangles.

* since we have a complete graph we know that there is an edge $j\to i$.

2. The number of triangles can be larger than the number of edges in a complete graph, for example if

$$\frac{N\cdot(N-1)\cdot(N-2)}{3} > \frac{N(N-1)}{2} \quad (\hookrightarrow \text{number of edges in a complete graph})$$

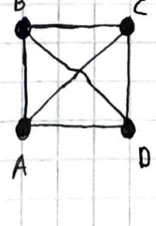
if $N>3$:

$$\Leftrightarrow \quad \frac{N-2}{3} > \frac{1}{2}$$

$$\Leftrightarrow \quad N > \frac{3}{2}+2 = \frac{7}{2} = 3.5$$

(or $N \geq 4$)

Example:



$N=4 \to$ graph has 6 edges and $\frac{4\cdot3\cdot2}{3} = 8$ triangles

(ABC, CBA, ACD, DCA, ADB, BDA, BCD, DCB)

3. To prove:

$$\#\text{connected triples in an undirected graph} = \frac{1}{2} \sum_{\substack{i,j=0 \\ i \neq j}}^{N} (A^2)_{ij} \quad \leftarrow \text{off diagonal}$$

The adjacency matrix A with its elements a_{ij} indicates whether two nodes i, j are connected by an edge. Therefore, A^2 with $(A^2)_{ij} = \sum_{k=1}^{N} A_{ik} A_{kj}$ indicates how many paths of length 2 between nodes i and j exist. $A_{ik} A_{kj} = 1$ if there's a path $i \to k$ and $k \to j$, otherwise $A_{ik} A_{kj} = 0$. By definition, (i, j, k) is a connected triple if i is connected to j and j to k, so $(A^2)_{ij}$ is the ~~sum of the~~ number of triplets between i and j. By summing up all off-diagonal elements of A^2 we get the total number of connected triples (the diagonal elements would include cycles ~~of len~~ with two edges, which we are not interested in here). Since we regard an undirected graph where $i \to k \to j$ and $j \to k \to i$ represent the same connected triplet we also need to divide by a factor $\frac{1}{2}$. Instead, it's also possible to just calculate the sum of the elements of the upper-triangular (or lower-triangular) part of matrix A^2, excluding diagonal elements.

4. global clustering coefficient:

$$C_\Delta \overset{\text{def.}}{=} \frac{3 \times \#\text{triangles}}{\#\text{connected triples}} = \frac{3 \times \frac{1}{2} \frac{1}{3} \operatorname{Tr}(A^3)}{\frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} (A^2)_{ij}} = \frac{\operatorname{Tr}(A^3)}{\cancel{\frac{1}{2}} \sum_{\substack{i,j \\ i \neq j}} (A^2)_{ij}}$$

Explanation:

According to the definition of the global clustering coefficient we need to calculate the number of connected triples, which was already shown in part 3. to be $\frac{1}{2} \sum_{i \neq j} (A^2)_{ij}$, and the

Complex Network Analysis

number of triangles. A triangle in a graph is given by 3 nodes and 3 edges connecting all 3 nodes with each other, forming a cycle. A^3 with elements $(A^3)_{ij} = \sum_{k=1}^{N} \sum_{l=1}^{N} A_{ik} A_{kl} A_{lj}$ indicates how many paths of length 3 there are between two different nodes i and j ($i \to k \to l \to j$). The diagonal elements of A^3 $(A^3)_{ii}$ therefore indicate how many paths starting and ending at a node i with length 3, so cycles with 3 nodes, exist in the graph. (Since the length is 3 we can be sure that those paths must be trails.) As the order of nodes does not ~~count~~ matter when counting triangles (the trails $i \to k \to l \to i$ and $i \to l \to k \to i$ belong to the same triangle) it follows that $\frac{1}{2}(A^3)_{ii}$ is the number of triangles containing node i. The total number of triangles is then
$\#\text{triangles} = \frac{1}{3} \cdot \frac{1}{2} \cdot \mathrm{Tr}(A^3) = \frac{1}{3} \cdot \frac{1}{2} \cdot \sum_{i=1}^{N} (A^3)_{ii}$. Here we also account for the fact ~~by~~ that the starting point of trail does not matter when counting triangles by multiplying with a prefactor $\frac{1}{3}$.