WEEK 11 Lecture 1

CHAPTER 7
SMALL-WORLD NETWORKS

Many real-world networks have at the same time

$\boxed{A}$ small distances between their nodes (S.W.D.P.) in W3

$\boxed{B}$ large number of triangles

- How to measure this ?
- It is easy to produce A + B in networks with "many" links

PUZZLE : real networks are SPARSE

Introduction of WS SMALL-WORLD network model

Shortest-path distance matrix

$$\mathbb{d} = \{ d_{ij} \}$$

DIAMETER $D = \max_{i,j} \{ d_{ij} \}$

CHARACTERISTIC PATH LENGTH $\ell = \langle d_{ij} \rangle$

Watts and Strogatz (W/S) in Nature 1998 paper

7.2 CLUSTERING COEFFICIENT

DEF

The NODE CLUSTERING COEFFICIENT C_i of node i is

$$C_i = \begin{cases} \dfrac{T_i}{\frac{k_i(k_i-1)}{2}} & \text{if } k_i > 1 \\ 0 & \text{if } k_i = 0, 1 \end{cases}$$

where k_i is the degree of node i, and

T_i is the # of triangles passing through node i

The NETWORK CLUSTERING COEFFICIENT C_{WS}

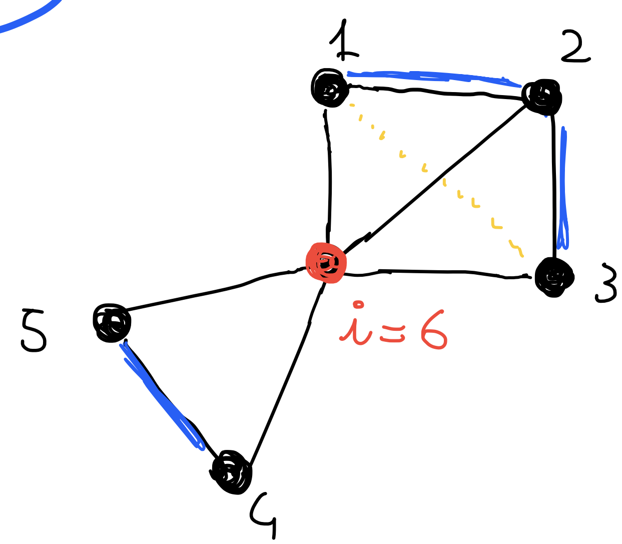
is $C_{WS} = \frac{1}{N} \sum_{i=1}^{N} C_i$ ← average over all the nodes in the network

Notice that $\frac{k_i(k_i-1)}{2} = \binom{k_i}{2}$ is the max # of triangles

that can pass through a node of degree k_i

Hence $C_i \in [0,1] \ \forall i \Rightarrow C \in [0,1]$

EX

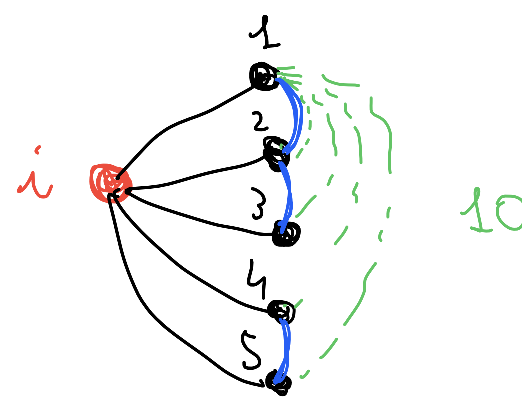


$N = 6$

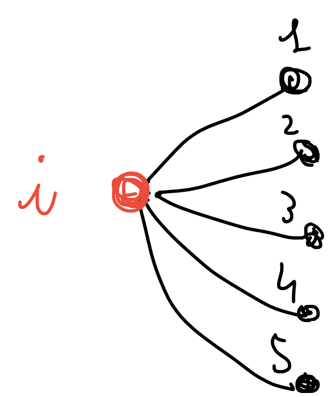
$i=6$

$k_i = 5$

$$\binom{k_i}{2} = \frac{k_i(k_i-1)}{2} = \frac{5\cdot 4}{2} = 10$$



max # of triangles that can pass through node $i=6$



Finally for $i=6$

$$C_i = \frac{T_i}{\binom{k_i}{2}} = \frac{3}{10}$$

$$C_1 = \frac{1}{1} \qquad C_2 = \frac{2}{3} \qquad C_3 = \frac{1}{1} \qquad C_4 = \frac{1}{1} \qquad C_5 = \frac{1}{1} \qquad C_6 = \frac{3}{10}$$

$\binom{2}{2} = 1$ $\qquad$ $\frac{3 \cdot 2}{2}$

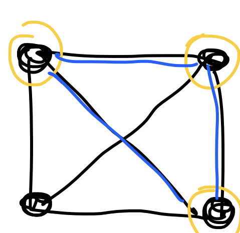
$$C_{ws} = \frac{1}{N} \sum_{i=1}^{N} C_i = \frac{1}{6} \left[1 + \frac{2}{3} + 1 + 1 + 1 + \frac{3}{10} \right] = \frac{1}{6} \cdot \frac{149}{30} \simeq$$

$$\simeq 0.83\ldots$$

Ex COMPLETE NETWORKS

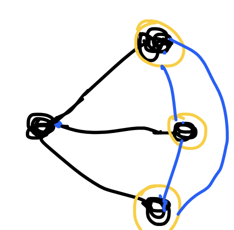
All the nodes have clustering coefficient equal to 1 : $C_i = 1 \ \forall i$

$\mathbb{K}_4$



$N = 4$

$k_i = 3 \quad \forall i$

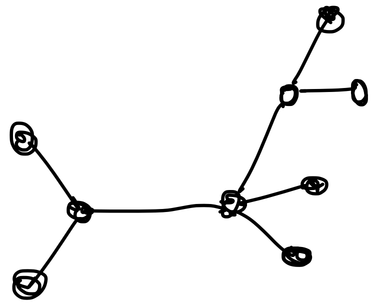


$$C_i = \frac{3}{\frac{3 \cdot 2}{2}} = 1 \quad \forall i$$

$\Downarrow$

$$C_{ws}^{\text{complete netw}} = 1$$

Ex TREES



$C_i = 0 \quad \forall i \quad \Rightarrow \quad C_{ws}^{tree} = 0$

because a tree has no cycles of any size (so no triangles)

linear chain

There are alternative measures to C_{ws}

DEF

a.k.a. **TRANSITIVITY**

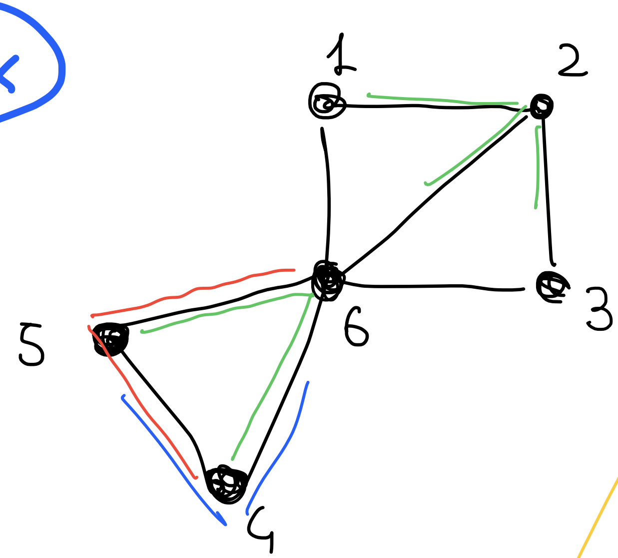
introduced in the social network community

The GLOBAL CLUSTERING COEFFICIENT C

$$C = \frac{3 \cdot T}{\# \text{ of distinct paths of length } 2}$$

where T is the # of triangles in the network

Ex



$N = 6$

$T = 3$

\# of distinct paths of length 2

1	2	3	4	5	6	
1	3	1	1	1	10	= 17

$$C = \frac{3 \cdot 3}{17} = \frac{9}{17} \simeq 0.53$$

7.3 SMALL-WORLD NETWORKS

taken from W3

DEF SMALL-WORLD (DISTANCE) PROPERTY (S.W.D.P.)

A network has the SWDP if

$$D \simeq O(\ln N)$$

the diameter is of the same order of magnitude as $\ln N$

big O

or if

$D \simeq o(\ln N)$ is of a smaller order of magnitude

small o

$D \simeq O(\ln N) \Rightarrow \lim_{N\to\infty} \frac{D}{\ln N} = c < \infty$ with $c > 0$

$D \simeq o(\ln N) \Rightarrow \lim_{N\to\infty} \frac{D}{\ln N} = 0$

Basically SWDP $\Longleftrightarrow \lim_{N\to\infty} \frac{D}{\ln N} = \text{constant} < \infty$

PROPOSITION

If a network has the S.W.D.P. then either

$$\ell \simeq O(\ln N) \quad \text{or} \quad \ell \simeq o(\ln N)$$

characteristic path length (average distance)

Hence $\lim_{N\to\infty} \frac{\ell}{\ln N} < \infty$

Proof: $\ell \le D$

REMARK 1

We will later have that $\boxed{\ell^{Rand} = \frac{\ln N}{\ln \langle k \rangle}}$

Hence SWDP means that ℓ of a network is of the same order of magnitude of the average distance ℓ^{rand} in a random network having the same average degree $\langle k \rangle$

DEF

A network has a HIGH CLUSTERING COEFFICIENT C_{ws} if

$$C_{ws} \gg \frac{\langle k \rangle}{N}$$

REMARK 2 We will later have that $\boxed{C^{Rand}_{ws} = \frac{\langle k \rangle}{N}}$

DEF **SMALL-WORLD NETWORKS**

Networks with [A] the S.W.D.P.

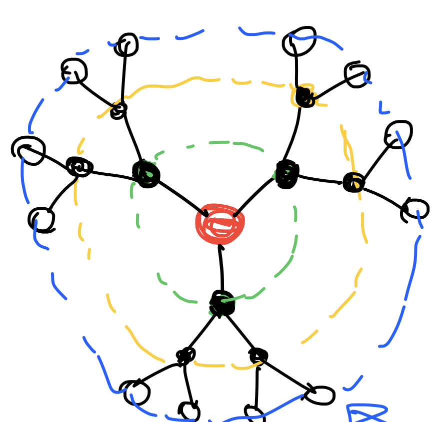
[B] high clustering coefficient

7.5 CAYLEY TREES

DEF **CAYLEY TREE**

A symmetric regular tree constructed starting from a given node (the origin) of degree **K**. Every node at distance $0 < d < \mathbf{P}$ from the origin has degree K, while the nodes at distance $d = P$ have degree 1. The quantity $\mathbf{b = K - 1}$ is called the **BRANCHING RATIO**

EX Cayley tree with $K = 3$ $P = 3$



branching ratio

$$b = K - 1 = 2$$

PROPOSITION

Cayley trees have the SWDP as:

$$D \simeq 2\,\frac{\ln N}{\ln b} \quad \text{for } N \gg 1$$

Proof

(a) We first prove that:

# of nodes at distance d from the origin

$$N_d = \begin{cases} 1 & d = 0 \\ K\, b^{d-1} & 1 \le d \le P \end{cases}$$

For $d = 0$ $\qquad N_0 = 1$ (this is the origin)

For $d > 0$ $\qquad$ Proof of $N_d = K\, b^{d-1}$ by induction

- $\boxed{d = 1} \to N_1 = K\, b^{1-1} = K$ which is true because the origin has K neighbours

- Then we assume $N_d = K\, b^{d-1}$ holds true for $1 \le d \le P$ and we prove $N_{d+1} = K\, b^{d}$

$$N_{d+1} = N_d \cdot b = k b^{d-1} \cdot b = k b^d$$

each node at distance d leads to $b = k-1$ nodes

(b) We then calculate N (the total # of nodes in the network) as

$$N = \underbrace{\sum_{d=0}^{P} N_d}_{\substack{\text{\# of nodes at distance}\\ d \le P \text{ from the origin}}} = N_0 + \sum_{d=1}^{P} N_d = 1 + k \sum_{d=1}^{P} b^{d-1} = 1 + k \sum_{n=0}^{P-1} b^n$$

$N_d = k b^{d-1}$ $\quad 1 \le d \le P$

Using the GEOMETRIC SUM FORMULA

$$\sum_{n=0}^{M} r^n = \frac{1 - r^{M+1}}{1 - r}$$

with $M = P-1$ and $r = b$ we get:

$$N = 1 + k \frac{1 - b^{P\cancel{-1}\cancel{+1}}}{1-b} = 1 + k \frac{1 - b^P}{1-b} = 1 + k \frac{(k-1)^P - 1}{k-2}$$

$b = k-1$

(c) We can now calculate the diameter D as

$$D = 2P$$

$$N = 1 + k\,\frac{(k-1)^{\frac{D}{2}} - 1}{k-2} \qquad (N-1)\,\frac{k-2}{k} = (k-1)^{\frac{D}{2}} - 1$$

$$\ln\left[(N-1)\,\frac{k-2}{k} + 1\right] = \frac{D}{2}\,\ln(k-1)$$

$$\boxed{D = 2\,\frac{\ln\left[(N-1)\,\frac{k-2}{k} + 1\right]}{\ln(k-1)}}$$

For $N \gg 1$ $\quad \ln\left[(N-1)\,\frac{k-2}{k} + 1\right] \simeq \ln\left[(N-1)\,\frac{k-2}{k}\right] \simeq \ln\left[N\,\frac{k-2}{k}\right] =$

$$= \ln N - \ln k + \ln(k-2) \simeq \ln N$$

$$D \simeq 2\,\frac{\ln N}{\ln(k-1)} = 2\,\frac{\ln N}{\ln b}$$

PROPOSITION

In a Cayley tree $C_i = 0 \ \forall i$. Hence $C_{WS} = 0$

being trees, Cayley trees do not have triangles

7.6 POISSON NETWORKS

$$P(N) = \frac{c}{N} \qquad c = \langle k \rangle$$

In Poisson networks the # of cycles of any size is finite when $N \to \infty$.

So such networks are "locally tree-like"

This is the reason why we expect similar results to those obtained for Cayley trees

PROPOSITION

The average distance $\ell = \langle d_{ij} \rangle$ of a Poisson network with average degree $\langle k \rangle = c$ is

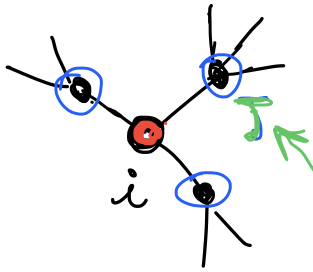
$$\boxed{\ell \simeq \frac{\ln N}{\ln c} \quad \text{for } N \gg 1}$$

Proof

(a)

# of nodes at distance d from a node i

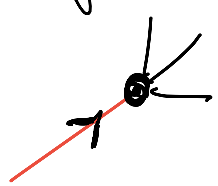
$$N_d(i) \simeq \begin{cases} 1 & d = 0 \\ k_i\, c^{d-1} & d > 0 \end{cases}$$



- In 1 step from i we can reach k_i nodes
- Then the # of nodes that we can reach from this branch (from node i) is

$b_j = k_j - 1$ branching ratio of node i

average branching ratio of a node reached by following a link



$$\boxed{\bar{b} = \frac{\langle k(k-1)\rangle}{\langle k\rangle}} = \frac{c^2}{c} = c$$

See W12

for POISSON NETWORK

$$N_d(i) \simeq k_i \, c^{d-1}$$

Hence the average # of nodes at distance d from a RANDOMLY CHOSEN node is

$$N_d \simeq \begin{cases} 1 & d = 0 \\ \langle k \rangle c^{d-1} = c^d & d > 0 \end{cases}$$

(b) $N \simeq N_{d=\ell}$ $\qquad N = e^{\ell}$ $\qquad \ln N = \ell \ln e$

characteristic path length

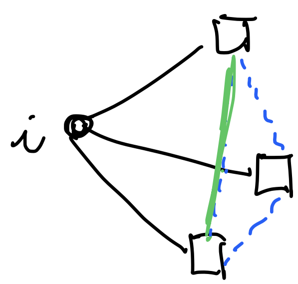
$$\ell \simeq \frac{\ln N}{\ln e} = \frac{\ln N}{\ln \langle k \rangle}$$

PROPOSITION

The clustering coefficient C_{ws} of a POISSON NETWORK is

$$\boxed{C_{ws} \simeq \frac{c}{N} = \frac{\langle k \rangle}{N} = p \quad \text{for } N \gg 1}$$

Proof



A node i with k_i will be part of

$$T_i = \frac{k_i(k_i-1)}{2} \cdot p \quad \text{triangles}$$

$$C_i = \frac{T_i}{\frac{k_i(k_i-1)}{2}} = p$$

$$C = p = \frac{c}{N}$$