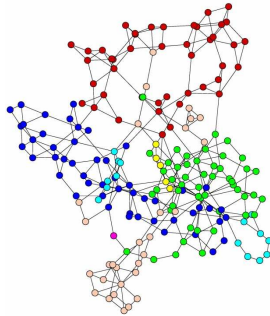




Universiteit Leiden



Metabolic Network Analysis

Lecture 2:
Network statistics

dr. Sander Hille

shille@math.leidenuniv.nl

<http://www.math.leidenuniv.nl/~shille>

Snellius, Niels Bohrweg 1, room 401



Mathematisch
Instituut

Spring semester 2019



Universiteit Leiden

Network Analysis

-- some elementary statistics --



Mathematisch
Instituut

N : total number of nodes (vertices)

Nodes degree or connectivity

k : number of neighbours of a particular node

$P(k)$: degree distribution
(frequency of nodes of nodes degree k)

Mean path length

l_{ij} : length of shortest path from node i to node j

$\langle l \rangle$: mean path length:

$$\langle l \rangle := \frac{2}{N(N-1)} \sum_{i < j} l_{ij}$$

'small world network': $\langle l \rangle$ depends logarithmically on N

$$\langle l \rangle \sim \log N$$

(2) Sander Hille

Spring semester 2019

Metabolic Nw. An. L2



Network Analysis

-- some elementary statistics --



● Clustering coefficient

k_i : number of neighbours of node i
 n_i : number of edges connecting the k_i neighbours of node i to each other

C_i : clustering coefficient of node i

$$C_i := \frac{2n_i}{k_i(k_i - 1)} \quad \text{'connectedness among the neighbours of node i'}$$

$\langle C \rangle$: average clustering coefficient

$$\langle C \rangle := \frac{1}{N} \sum_i C_i$$

$C(k)$: average clustering coefficient of all nodes that are connected to k neighbours (i.e. of nodes degree k)



Network Analysis

-- some elementary statistics --



● Clustering coefficient (continued)

When

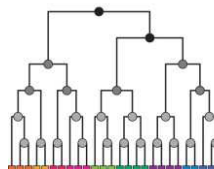
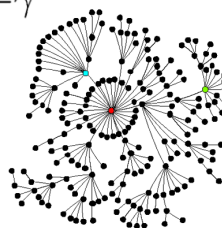
$$C(k) \sim k^{-\beta} \quad \text{'power law'}$$

then the network has a *hierarchical* structure

When

$$P(k) \sim k^{-\gamma}$$

then the network is '*scale-free*'





Universiteit Leiden

Network Analysis

-- some elementary statistics --



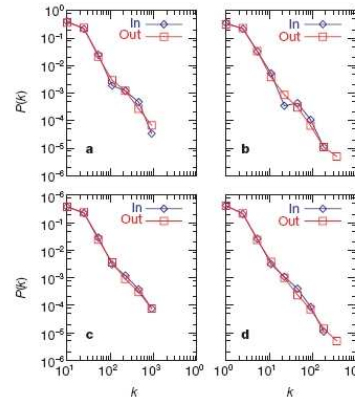
Statistics from some real life networks

(for undirected substrate graphs with currency metabolites removed ranging over 6 Archeae, 32 Bacteria, 5 Eukaryotes)

$$P(k) \sim k^{-2.2}$$

'Metabolic networks are scale-free'

- a: *Archaeoglobus fulgidus* (Archeae)
- b: *Escherichia coli* (Bacteria)
- c: *Caenorhabditis elegans* (Eukaryotes)
- d: Average over all 43 researched species



(Jeong ea. Nature 407 (2000), pp.651—654)

(5) Sander Hille

Spring semester 2019

Metabolic Nw. An. L2



Universiteit Leiden

Network Analysis

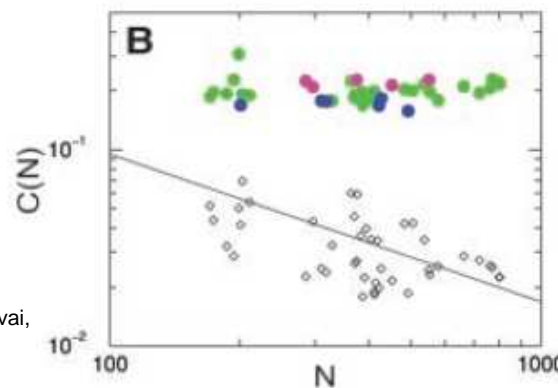
-- some elementary statistics --



Statistics from some real life networks

(for undirected substrate graphs with currency metabolites removed ranging over 6 Archeae, 32 Bacteria, 5 Eukaryotes)

- Archeae (6)
- Bacteria (32)
- Eukaryotes (5)



(Ravasz, Somera, Mongru, Oltvai, Barabasi, Science 297 (2002), pp.1551-1555)

(6) Sander Hille

Spring semester 2019

Metabolic Nw. An. L2