Graph Representation

Networks are

Large → Use computational resources for analysis

Sparse → Efficient ways of storing the information

Objects to analyze → Need representations amenable to efficient algorithms
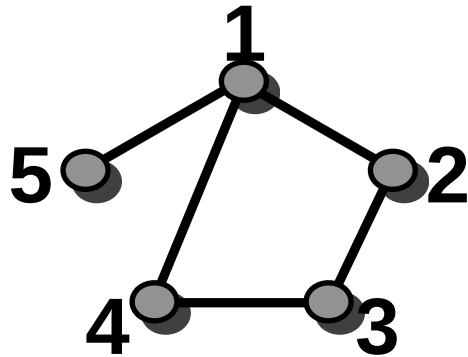
Matrix Representation

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1N} \\ A_{21} & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ A_{N1} & \ldots & \ldots & A_{NN} \end{pmatrix}$$

$A_{ij} = 1$ if there is a link pointing from node j to node i

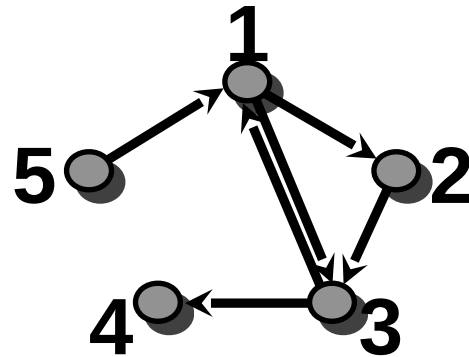$A_{ij} = 0$ if nodes i and j are not connected to each other

Adjacency Matrix

Undirected

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Directed

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Adjacency Matrix

Advantages

→ Convenient for analytical calculations

→ Easy to remove/add an edge (changing the value of an element in the matrix is $O(1)$)

Disadvantages

→ Needs a lot of memory - $O(N^2)$ space

→ Inconvenient for numerical calculations

Adjacency Matrix

O(N^2) doesn't sound like a lot, but consider:

➤ Adjacency matrix has N^2 entries, where N is the number of nodes in the network

➤ If an integer = 4 bytes, and RAM =10 GB
$4N^2 = 10\text{Gb}$, then $N \sim 50,000$

M.E.J. Newman. (2010). *Networks: An Introduction.* Oxford University Press.

Edge list (ij-form)

$$I = \begin{pmatrix} \cdots \\ i \\ \cdots \end{pmatrix} \quad J = \begin{pmatrix} \cdots \\ j \\ \cdots \end{pmatrix} \quad \text{Only if } A_{ij} = 1$$

➡ Order of rows not important
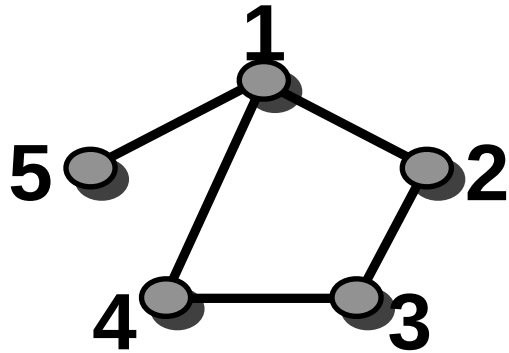
➡ If undirected:

only report link with i<j (L rows)
report both i j, and row j i ($2L$ rows, redundancy)

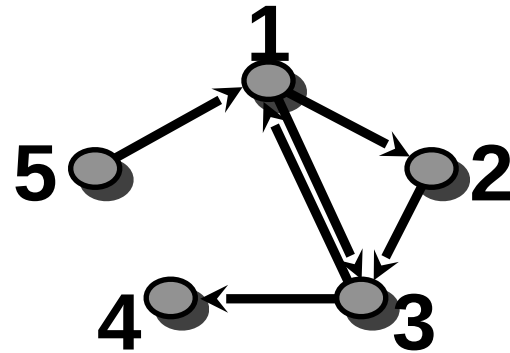➡ If directed:

report each arc (L arcs)

Edge list (ij-form)

Undirected

Directed

$$I = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 3 \end{pmatrix} \quad J = \begin{pmatrix} 2 \\ 4 \\ 5 \\ 3 \\ 4 \end{pmatrix}$$

$$I = \begin{pmatrix} 5 \\ 1 \\ 3 \\ 1 \\ 2 \\ 3 \end{pmatrix} \quad J = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \\ 3 \\ 4 \end{pmatrix}$$

Edge list (ij-form)
Weighted networks

$$I = \begin{pmatrix} \cdots \\ i \\ \cdots \end{pmatrix} \quad J = \begin{pmatrix} \cdots \\ j \\ \cdots \end{pmatrix} \quad W = \begin{pmatrix} \cdots \\ w_{ij} \\ \cdots \end{pmatrix}$$

Only if $\quad w_{ij} \neq 0$

Edge list (ij-form)

Advantages

→ Most used in data repositories

→ Convenient for data collection

Disadvantages

→ Not easy to find elements
(e.g. check the existence of an edge is O(L)

→ Not useful for many analytical calculations