Lecture 3
CDS-292

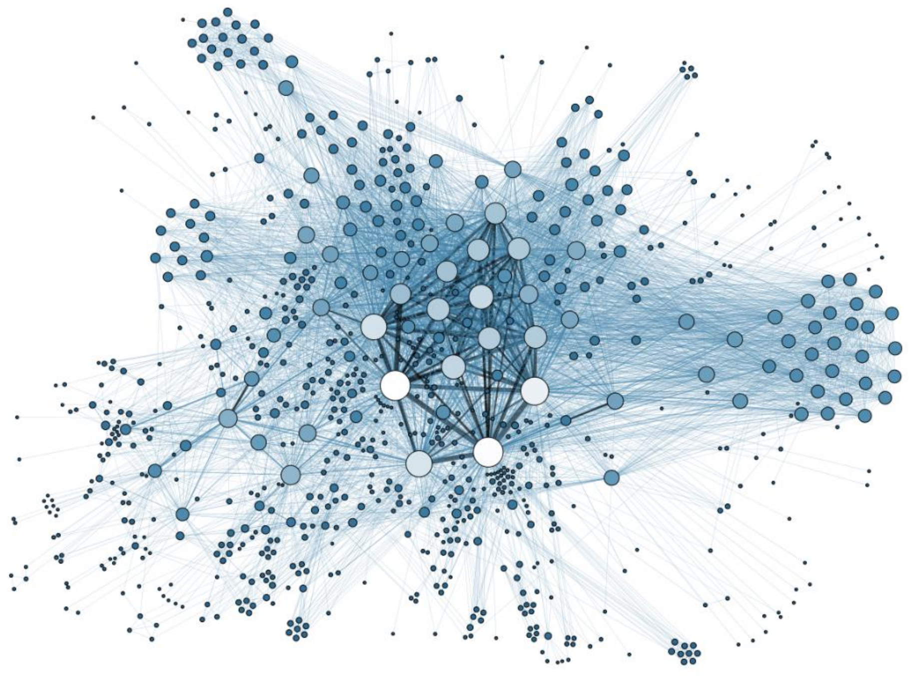
Joseph Shaheen

History

- Moreno (1934)
- Bott (1957)
- Erdos & Renyi (1959)
- Traverse & Milgram (1969)
- Granovetter (1973)
- Freeman (1970s)
- Holland & Reinhard (1981); Krackhardt (1987)
- Barabasi, Wattz and Strogatz (1990s)
- Newman, Snijders (2000s)

Background

- A Network: ***Nodes***, ***Actors***, ***Vertices*** connected by ***edges***, ***ties***, ***relations***
- Can be single mode, or n-partite
- Can be measured and analyzed through structural/mathematical metrics acting on multiple levels: ***Node-level***, ***Dyad-level***, ***Network*** as a whole



Degree of a Node

- Different nodes will have different degrees. To track the number of connections each node has we introduce the term Node Degree, or as it is more commonly known, Degree

k_i is the number of total links connected to node i

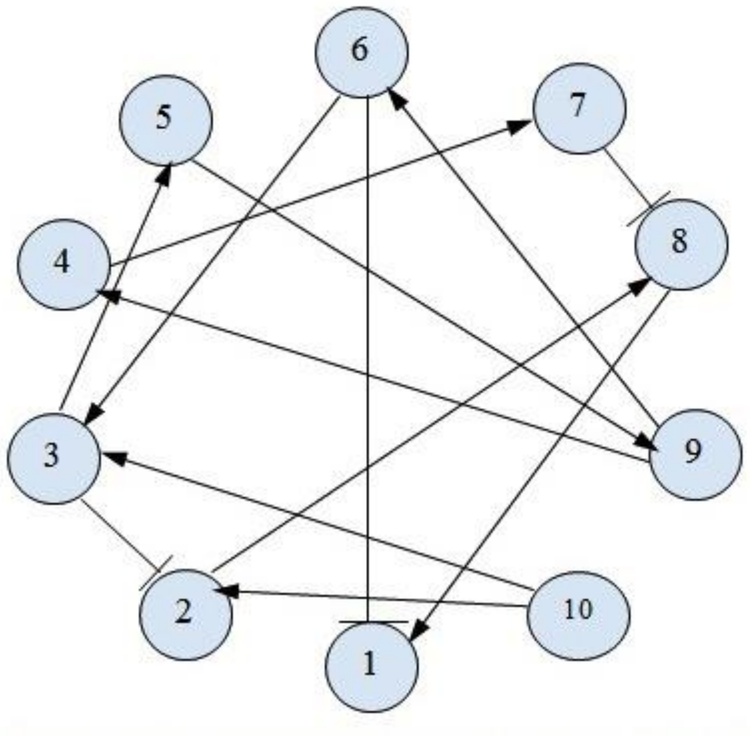
- To keep track of the number of links between 2 nodes, say node i and node j, mathematically, we introduce a simple function that we call the link indicator. It’s basically an on or off switch.

$$a_{i,j} = \begin{cases} 1 \; if \; i \; and \; j \; are \; connected \\ 0 \; otherwise \end{cases}$$

To Calculate the Node Degree

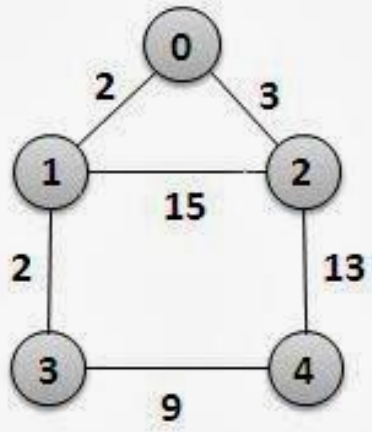
$$k_i = \sum_{j=1}^{n} a_{ij}$$

- Let’s look at an example:



Source: https://www3.beacon-center.org/blog/2014/03/03/beacon-researchers-at-work-constructing-models-for-gene-regulatory-networks/

Adjacency Matrix Example

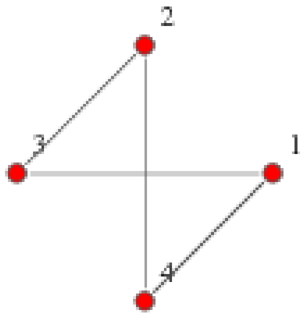


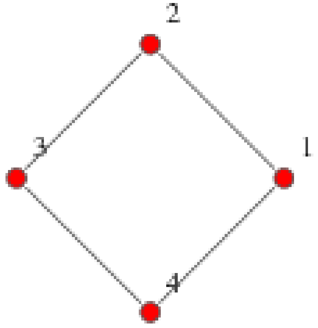
	0	1	2	3	4
0	0	2	3	0	0
1	2	0	15	2	0
2	3	15	0	0	13
3	0	2	0	0	9
4	0	0	13	9	0

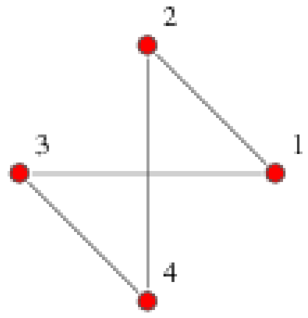
Adjacency Matrix Representation of Weighted Graph

https://www.thecrazyprogrammer.com/2014/03/representation-of-graphs-adjacency-matrix-and-adjacency-list.html

Simple Adjacency Matrix







$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

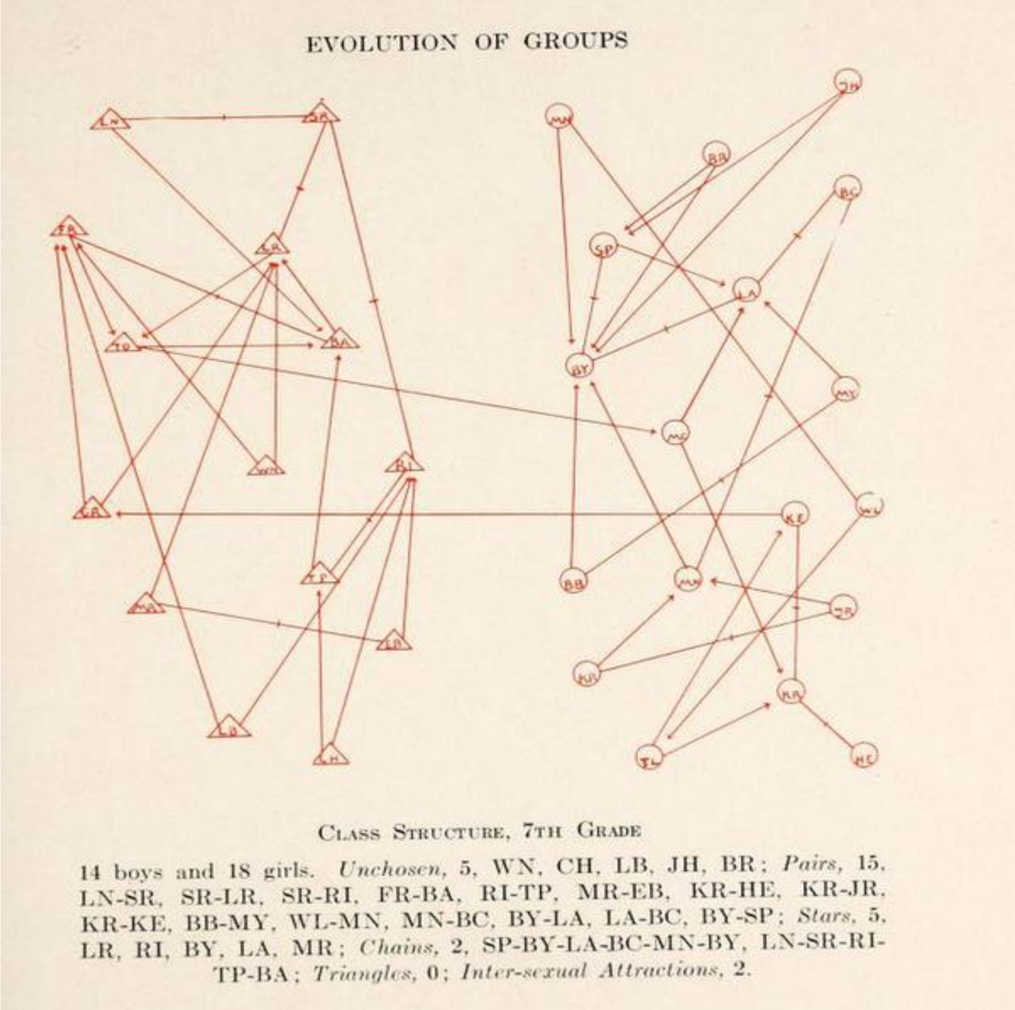
Introducing Density

- Density is a very useful network-level metric property
- It tell us how sparse or full a network is and this has interesting sociological and physical interpretations

$We\ define\ the\ density\ \rho$ as

$$\rho = \frac{m}{n\,(n-1)/2} = \frac{2m}{n(n-1)} \quad where\ m\ is\ number\ of\ links\ and\ n\ is\ number\ of\ nodes$$

Moreno (1934)



EVOLUTION OF GROUPS

CLASS STRUCTURE, 7TH GRADE

14 boys and 18 girls. *Unchosen*, 5, WN, CH, LB, JH, BR; *Pairs*, 15, LN-SR, SR-LR, SR-RI, FR-BA, RI-TP, MR-EB, KR-HE, KR-JR, KR-KE, BB-MY, WL-MN, MN-BC, BY-LA, LA-BC, BY-SP; *Stars*, 5, LR, RI, BY, LA, MR; *Chains*, 2, SP-BY-LA-BC-MN-BY, LN-SR-RI-TP-BA; *Triangles*, 0; *Inter-sexual Attractions*, 2.

Moreno Data Set

- Organized as a list of edges

BA AB

BA BR1

DE EP

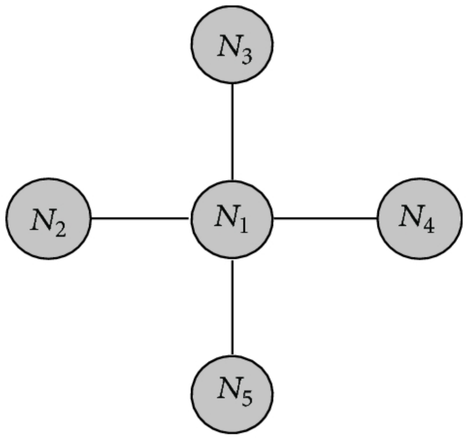
Node1 Node2

Finding Path Length (all types through matrix algebra)

$A_{ij}^{l} = \sum_{g=1}^{n} a_{ig}\, a_{gj}$ where l is the path length

In other words, to find out how many paths of length 2 exist between every node i and every node j,
we simply multiply the adjacency matrix by itself. For path length 3, we multiply the adjacency matrix by itself 3 times..etc..

Example

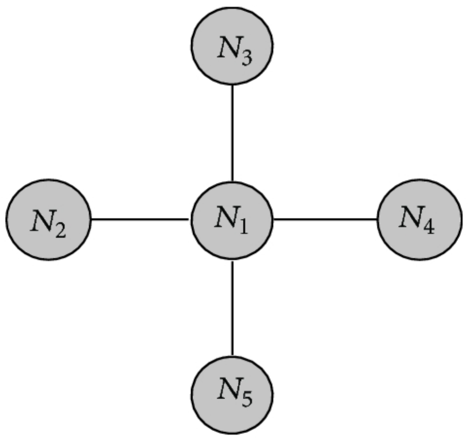


$$
\begin{matrix}
0 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0
\end{matrix}
$$

How many paths exist between every node of distance/path length 2?

Finding Path Length (Cont)

$$\begin{matrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{matrix} \times \begin{matrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{matrix} \coloneqq \begin{matrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{matrix}$$



There is at least 1, distance 2 path length path between every node except N1

Small World vs. Large World

- Small versus large worlds is a concept heavily dependent on the idea of path length.
- In a small world the average path length between nodes tends to be small, and in a large world, the distance can be large, even enormous.
- What Traverse and Milgram (69) showed was that the modern world's communication infrastructure allows us to model it as a small world. That was a breakthrough!

Small World vs. Large World

- …and we know a little about how the average path length behaves for typical small world and large world networks.

$$\langle l \rangle \propto \begin{cases} n^{\alpha} \\ \log n \end{cases} where\ \ 0 < \ \alpha\ \leq 1, n\ is\ the\ number\ of\ nodes$$

$$(though\ is\ isn't\ mentioned, but\ alpha\ is\ a\ fractal\ dimension)$$

Example: Let’s say we have 10 billion people in a single network. If we treat is as a large world then living on a 2 dimensional space ($\propto = 1/2$), then the average path length can be estimated to be 100,000 (huge).

If we estimate it to be a small world, then log(10 billion) = 10