# Networks: An Introduction (Newman) Chapter 8.1 - 8.4 discussion

Kwame Okrah

April 3, 2014

# The large-scale structure of real world networks

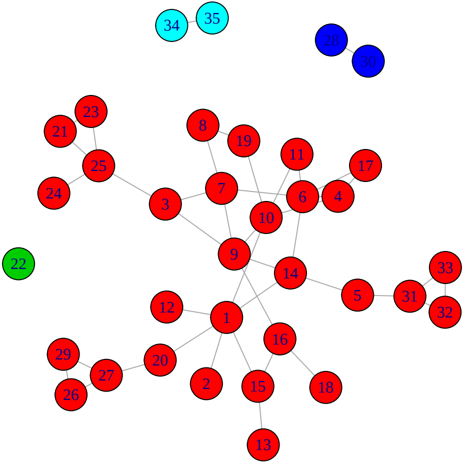
We will apply the ideas of chapter 6 and 7 to real world networks to get a sense of how they behave in general (mostly).

For today we will look at:

1. **Components**
   - Undirected and
   - Directed networks
2. **Shortest Path and the Small-World Effect**
3. **Degree Distribution**
4. **Power Laws and Scale-Free Networks**

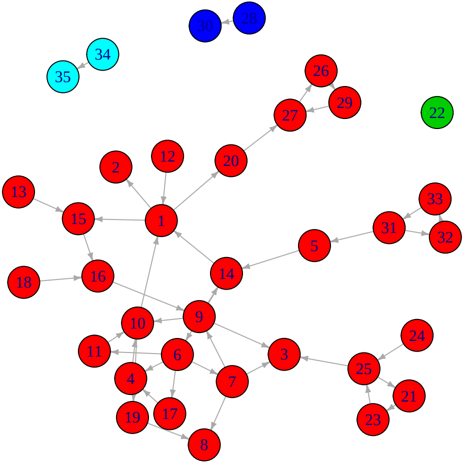
# Components (Undirected Networks)

Any two vertices in a component must be connected.



# Weakly Connected Components (Directed Networks)

Any two vertices in a component must be connected without regards to



# Some Examples of Large Components and Large Weakly Connected Components

We typically find that large components are usually more than half of vertex size ($n$) and not infrequently over 90% .

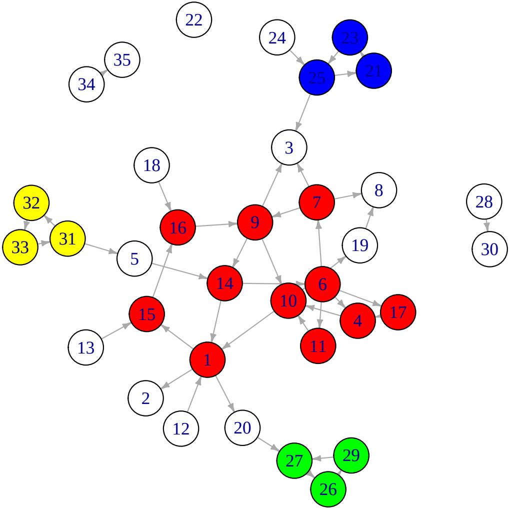
Table : (Table 18.1) Basic stats for some selected networks. Where n=no. of vertices, m=no. of edges, c=mean degree, S=% of largest component.

| Network | Type | n | m | c | S |
|---|---|---|---|---|---|
| Marine food web | Directed | 134 | 598 | 4.46 | 1.000 |
| Software pack. | Directed | 1,439 | 1,723 | 1.20 | 0.998 |
| Metabolic | Undirected | 765 | 3,686 | 9.64 | 2.56 |
| Bio. co-auth. | Undirected | 1,520,251 | 11,803,064 | 15.53 | 4.92 |
| Phy. co-auth. | Undirected | 52, 909 | 245,300 | 9.27 | 6.19 |
| Protein int. | Undirected | 2,115 | 2,240 | 2.12 | 6.80 |
| Student dating | Undirected | 573 | 477 | 1.66 | 16.01 |

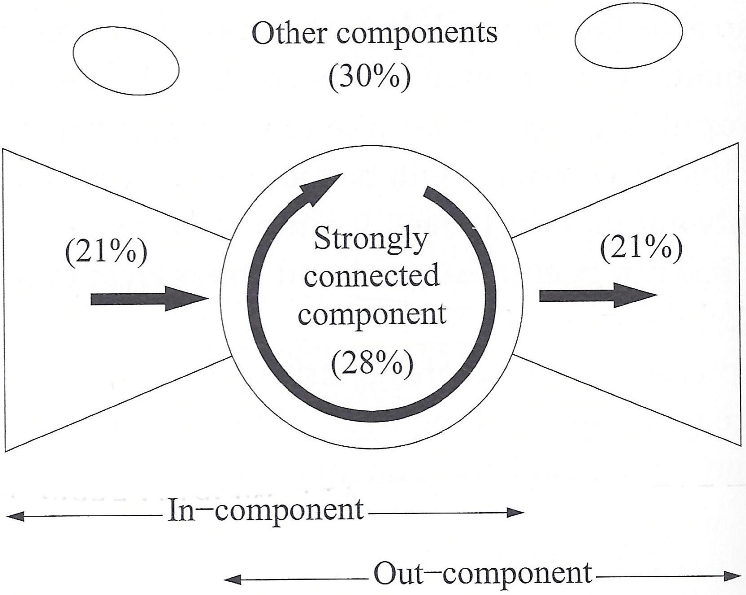
# Strongly Connected Components (Directed Networks)

- Any two vertices in a component must be connected (directed).
- Every strongly connected component generate an **in-component** and an **out-component**.

# Example: Strongly Connected Components



# The Bowtie Effect



# The Shortest Path and the Small-World Effect

- **Small-world effect**: the observation that in most real world networks the distances between vertices are small.
- Recall: $l_i = \frac{1}{n}\sum_j d_{ij}$, where $d_{ij}$ is the shortest path between $i$ and $j$; is the mean geodesic distance for vertex $i$.

  The mean geodesic distance of the network is defined as:
  $l = \frac{1}{n}\sum_i l_i$.
- For random graphs (Erdos-Reyni graphs): $l \propto log(n)$
- For scale free networks: $l \propto log(log(n))$

  (The definition of a scale-free network is below)

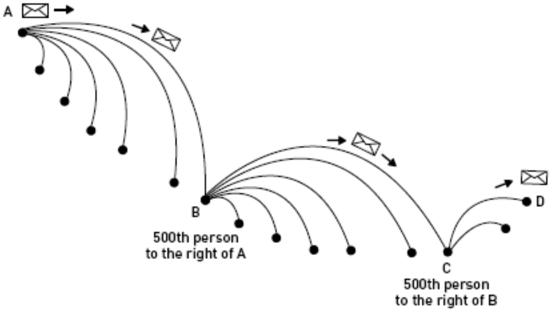
# The Shortest Path and the Small-World Effect

Table : (Table 18.1) Basic stats for some selected networks. Where l=mean geodesic distance connected vertex pairs.

| Network | Type | n | m | c | l |
|---|---|---|---|---|---|
| Metabolic | Undirected | 765 | 3,686 | 9.64 | 2.56 |
| Bio. co-auth. | Undirected | 1,520,251 | 11,803,064 | 15.53 | 4.92 |
| Phy. co-auth. | Undirected | 52, 909 | 245,300 | 9.27 | 6.19 |
| Protein int. | Undirected | 2,115 | 2,240 | 2.12 | 6.80 |
| Student dating | Undirected | 573 | 477 | 1.66 | 16.01 |

# Funneling and the Sociometric Superstars

- (Suggested by Milgram) **Funneling**: The idea that a given vertex has one or two connecting vertices that are well connected (i.e. have a high degree).



# Degree Distributions

- The **distribution of degrees** of a given network is a *KEY* fundamental property.
- Consider a simple undirected network with vertices, $\{1, 2, \ldots, n\}$.
- If we pick a vertex $i \in \{1, 2, \ldots, n\}$ at random, can we guess what it's degree $k_i$ will be ?

  No: But we know that it can be any counting number between 0 and $n - 1$.

  Also depending on what natural phenomena the network is describing, we can say something sensible (on average).

# Ways to Characterize A Degree Distribution

Table : Degree Distribution in table form

| degree: | 0 | 1 | 2 | ... | n-1 |
|---|---|---|---|---|---|
| probability: | $p_0$ | $p_1$ | $p_2$ | ... | $p_{n-1}$ |

Or

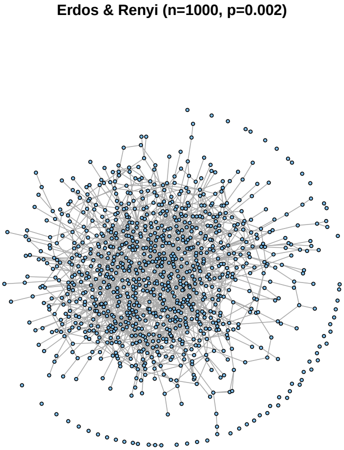
- Mathematical Function (aka. probability mass function, pmf):

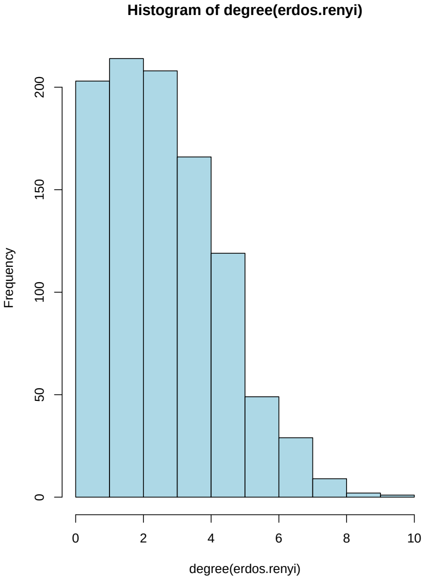
  $$p(k), \quad k = 0, 1, 2, \ldots, n-1.$$

When $p(k)$ follows the Power Law we say that we have a **scale-network**.

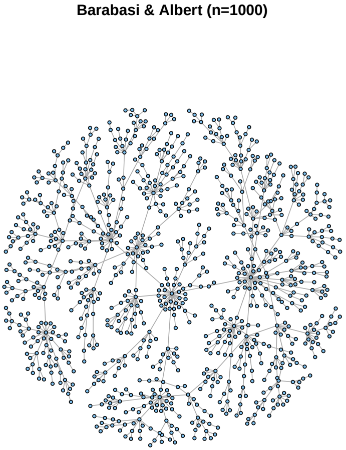
When a network has random connections (with a fixed probability of connection) then we have the **Erdos-Renyi network**. In this case $p(k)$ will follow a Binomial Law (usually approximated by a Poisson Law).

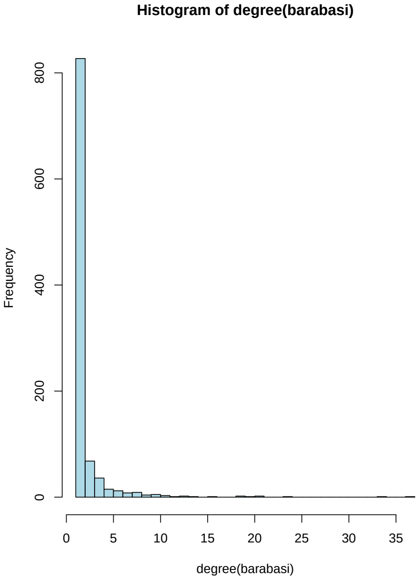
# Example: Erdos-Renyi (random)



# Example: Barabasi and Albert (scale free)



# Directed Networks

For directed networks we can talk about the in-degree distribution and the out-degree distribution, separately.

Ideally we would like to know the joint distribution instead.
This is an on-going research topic.

Next time we will look more closely at the Power Law and some of its consequences.