# Construction of random graphs with given clustering

Pieter Trapman

University Medical Center Utrecht

Utrecht University

Edinburgh, 16 May 2007

**Two ways to use random networks in epidemiology:**

1. Find an epidemic induced random network in a randomly mixing population or on a given network.

2. Replace given network by random network with desired properties and run epidemic on it.

**Two problems:**

1. Which properties to desire?

2. How to construct a random network with the desired properties?

**Suggestions for desirable properties:**

1. “Dimension” of the network ($\rightarrow$ infinite)

2. Degree distribution ($\rightarrow$ power law?)

3. Clustering (i.e. number of triangles in the network)

4. Degree correlation (i.e. the correlation of degrees of neighbouring individuals)

5. . . .

6. Susceptible to analysis

Possible way to construct network with desired degree distribution (configuration model by Molloy and Reed):
$\mathcal{N}$ vertices, $D$ is r.v. with desired distribution.

- Let there be $\mathcal{N}$ individuals and assign a random number of half-edges (edges with only one endpoint assigned to a vertex) to each individual, where the number of half-edges assigned to the individuals are i.i.d. and distributed as $D$.
- If the total number of half-edges is odd repeat the first step until you get an even number of half-edges.
- Pair the half edges at random.

Clustering in the network Solution 1:

**Definition:** The 3 vertices $v_i v_j v_k$ form a triple if $v_j$ is connected to both $v_i$ and $v_k$. If $v_i$ is also connected $v_k$ then $v_i v_j v_k$ is a triangle.
Let $\phi$ be the fraction of triples that is also a triangle.

Replace individuals by super-individuals (households and bachelors)
Individuals in households have only one neighbour outside household.

Use configuration model to connect super individuals.
Choose household sizes and “bachelor degrees” clever.

Problems:
High positive degree correlation.
“Bachelors have all the fun”.
Network looks unrealistic.

**Other way to construct graphs with clustering:**

**Bipartite graphs**: Use configuration model with two types of individuals: ordinary individuals and temporary individuals.
The total degree of both types should be equal, pair half edges such that every edge has two different types of end-vertices.
Delete the temporary individuals and connect alle individuals that were connected to the same temporary individual.

**Problem:** No easy control on both degree distribution and clustering.

**Important result by Kuulasmaa (1982)**

Consider $SIR$-epidemic on a given network, Let $\mathcal{I}$ be a random variable distributed as the infectious period. Let $\mathbb{E}(1 - e^{-\tau\mathcal{I}}) = p$.

Compare infectious period distributions with fixed $p$.

Worst case: $\mathbb{P}(\mathcal{I} = c) = 1$

Best case: $\mathbb{P}(\mathcal{I} = \infty) = 1 - \mathbb{P}(\mathcal{I} = 0) = p$.

It is sensible to combine the two ways of using random graphs in epidemiology:

Replace network by random network.

Construct the induced network of worst (best) case epidemic by randomly deleting edges (vertices).

**Questions:**

- How to control both clustering and degree correlation?
- Is it possible to define a graph growth model (like preferential attachment) with given clustering and degree distribution?
- Is it possible to make realistic ”finite dimensional” graphs, on which epidemics are susceptible to analysis?
- How to analyse dynamical models (Pair formation in sexual networks).