ECON5152 — Network Analysis

Week 6: Structure in Random Graphs

Santiago Montoya-Blandón

February 19, 2026

Reading for this week

- Necessary:
  - Newman, *Networks*: Chapter 12
  - Jackson, *Social and Economic Networks*: Sections 4.1.4,–4.1.6, 4.2.6–4.2.7, 4.3
- Suggested:
  - Jackson, *Social and Economic Networks*: Chapter 3
  - Newman, *Networks*: Chapter 4

Outline

1. Motivation
2. Configuration model
    - Degree sequences
    - Adjacency matrix and graph construction
    - Potential issues
    - Correspondence to Erdős–Rényi random graphs
    - Link probabilities and phase transitions
3. Extensions of structured random graphs
    - Directed networks
    - Bipartite networks
    - Trees and hierarchical structures
    - Clustering and subgraphs
    - **Stochastic Block Model**
    - **“Small-world” networks**
    - Applications: social and citation networks

UofG

Motivation

- Introduced random graphs and how to easily generate them
- Useful comparative benchmark
- Cannot characterize individual realizations
- Can characterize "typical" or average network
- Achieve some structure by choosing configuration

From Erdős–Rényi to Structured Randomness

- ER graphs processes:
  - $\mathcal{G}(n,p)$: each edge included independently with probability p
  - $\mathcal{G}(n,m)$: graph chosen uniformly with exactly m edges
  - *Degree distribution*: probability that a node i has degree $k \in \{0,\dots,n-1\}$

$$\Pr(d_i = k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

  - Binomial degree distribution for small n
  - Can approximate as Poisson for large n

$$\Pr(d_i = k) \approx e^{-(n-1)\cdot p} \frac{[(n-1)\cdot p]^k}{k!}$$

Motivation: Poisson vs Real-world

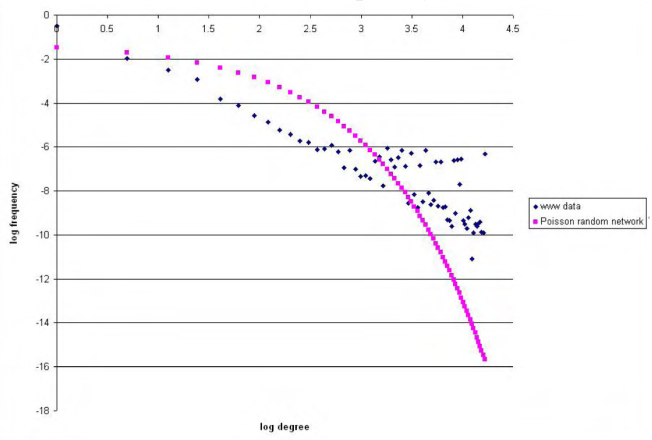


Figure: Degree of Internet (WWW) and Poisson networks

From Erdős–Rényi to Structured Randomness

- ER graphs predict degrees of nodes in large graphs look Poisson
- Identical degree distribution for all nodes
- Clustering is small compared to size
- Real networks exhibit
  - Heterogeneous degrees
  - High clustering
  - Community structure
  - “Small world” nature
- We now introduce models that preserve randomness but add structure

Configuration Model

- Structured generation method
- Generate a random network with given degrees

Configuration Model

- Structured generation method
- Generate a random network with given degrees
- Fix a degree sequence: $d_1, d_2, \ldots, d_n$
- Write nodes as many times as their degrees

$$\underbrace{111111}_{d_1}\,\underbrace{2222}_{d_1}\cdots\underbrace{nnnnnnnnn}_{d_n}$$

- Construct a random graph by:
  1. Choosing any two nodes at random from the list
  2. Create a connection between them
  3. Delete them from list and start again
- **Result:** random graph with fixed degrees!

Configuration model: Adjacency Matrix example

Degree sequence:

$$d_1 = 3,\ d_2 = 2,\ d_3 = 2,\ d_4 = 1$$

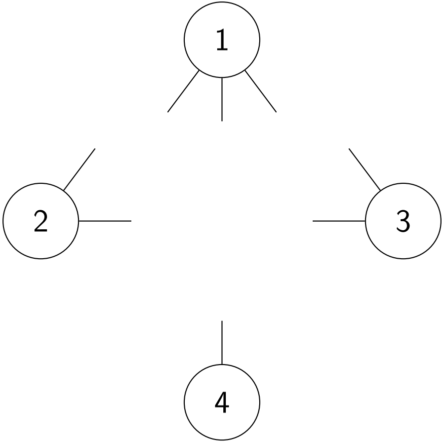
Edges:

$$(1,2),(1,3),(1,4),(2,3)$$

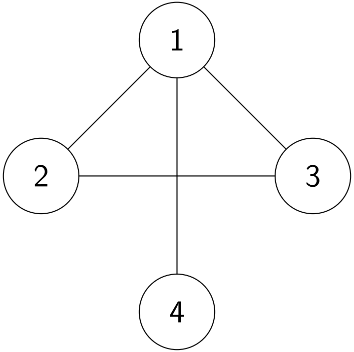
Adjacency matrix:

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

Configuration model: Graph example with "stubs"



Configuration model: Graph example



Issues with Configuration Model

Possible issues:

- Multi-edges:

$$(4, 2), (2, 1), (1, 3), (1, 3)$$

- Self-links:

$$(4, 2), (2, 3), (1, 3), (1, 1)$$

- Degree sum must be even or there are leftover nodes

Correspondence with ER Graphs

- With ER graphs, we have degree distribution (not sequence)
- Specifically, $d_i \sim \text{Binomial}(n-1, p)$
- Fix network size n and linking probability p
- Draw a sequence of degrees at random from this distribution
- **Result:** configuration model with Binomial ($\approx$ Poisson) degrees

Probability of Linking

- Total number of edges m?
- List of nodes to connect has size $\sum_{i=1}^{n} d_i$
- Each step connects two nodes from the list $\implies m = (1/2)\sum_{i=1}^{n} d_i$

Probability of Linking

- Total number of edges m?
- List of nodes to connect has size $\sum_{i=1}^{n} d_i$
- Each step connects two nodes from the list $\implies m = (1/2)\sum_{i=1}^{n} d_i$
- In terms of graph "stubs", each node has d_i of them
- Total $2m$ stubs in the graph before connections happen
- Each 1 stub (from $i \in \mathcal{N}$) can connect to remaining $2m - 1$ (any $j \in \mathcal{N}$)
- Probability any two nodes i and j link in configuration model:

$$\Pr(A_{i,j} = 1) := p_{ij} = \frac{d_i d_j}{2m - 1}$$

- Higher degree nodes more likely to connect

Configuration model: Visualizations

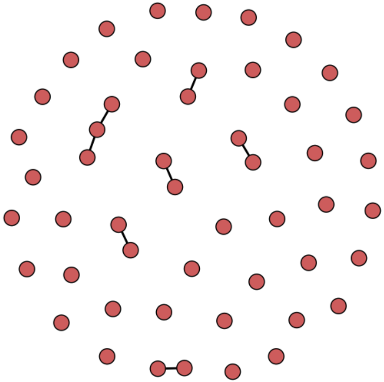


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0002$

Configuration model: Visualizations

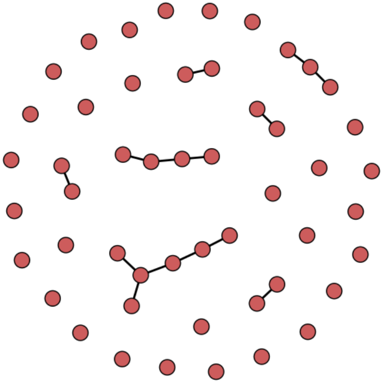


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0005$

Configuration model: Visualizations

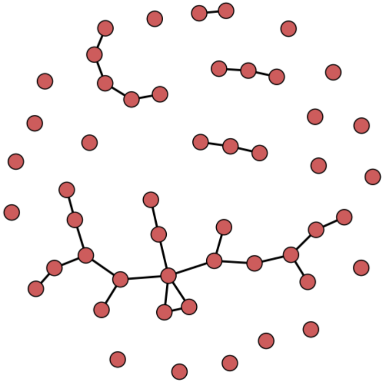


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0010$

Configuration model: Visualizations

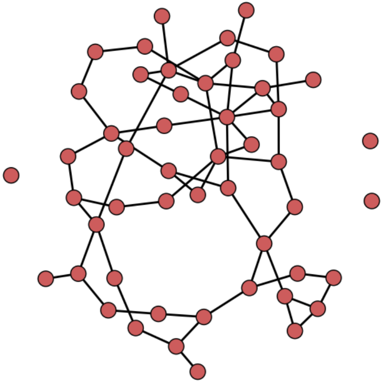


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0020$

Configuration model: Visualizations

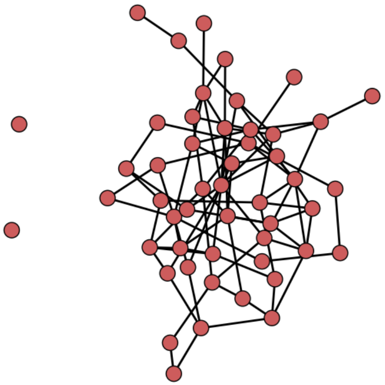


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0030$

Configuration model: Visualizations

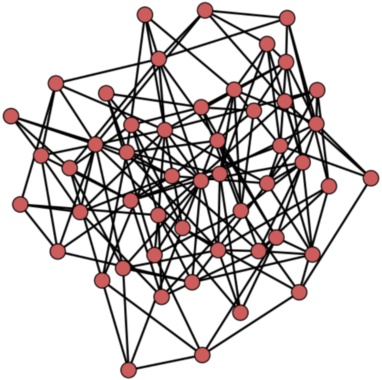


Figure: Configuration model with $n = 50$ nodes, Poisson degrees with $p = 0.0050$

Phase Transitions

- Similar to ER graphs, configuration model exhibits phase transitions
- Increasing degree sequences creates structure
- Emergence of important properties
  - Link formation and small components
  - Cycles and trees
  - Giant component
  - Connectedness
- In ER, thresholds depend on p/n
- In configuration model, depend on

$$\frac{c_2}{c_1} := \frac{\mathbb{E}[d_i^2]}{\mathbb{E}[d_i]}$$

- How likely to make paths of length 2 compared to length 1

Directed Networks

Two approaches:

1. Assign directed links randomly
2. Specify in-degree and out-degree sequences

Must satisfy:

$$\sum_i d_i^{in} = \sum_i d_i^{out}$$

Bipartite Networks

- Nodes divided into two sets $\mathcal{U}$ and $\mathcal{V}$
- Edges only across sets
- Degree sequences must satisfy:

$$\sum_{i \in \mathcal{U}} d_i = \sum_{j \in \mathcal{V}} d_j$$

- Applications:
  - Firms and workers
  - Authors and papers
  - Migrants and countries
  - Users and content

Trees

Tree construction methods:

1. Fix number of children per node
2. Random branching process
3. Layer-by-layer expansion

Used for:

- Organizational hierarchies
- Information diffusion
- Scientific article citations
- Temporal (dynamic) graphs

Clustering and Transitivity

- ER and configuration models exhibit low clustering
- Directly increase clustering by including larger substructures
- Extensions (assign probabilities to):
  - Triangles
  - Cliques (complete subgroups)
  - Core-periphery
- Simply decide on properties and assign them probabilities when generating

Community Structure

- Configuration or ER models do not allow for *communities*
- Groups with dense connections *within* the group
- Sparse (weak) connections *between* groups
- Introduce a model that can reflect such features
- Use block structures for modelling

Stochastic Block Model (SBM)

Steps:

- Choose nodes n and number of communities K $(g_1, \ldots, g_K)$
- Assign nodes to communities (either at random or by hand)
- Simplest case:
  1. Probability of connecting to others in the group: p_{within}
  2. Probability of connecting outside group: p_{between}
- More generally, define for each two groups g_i and g_j a probability $p(i, j)$
- Connect links at random using these two probabilities

Positive Assortativity

If:

$$p_{within} > p_{between}$$

Then:

- Strong community structure
- Positive assortativity
- Common in most real-world networks of mutual benefits

Negative Assortativity

If:

$$p_{within} < p_{between}$$

Then:

- Opposites attract
- Cross-group mixing dominates
- Common in some matching markets
- For example: dating or obtaining employment

Small-World Phenomenon

Empirical facts: networks exhibit

- Small average path length
- Small diameters
- Small degrees
- High clustering
- Long indirect chains

How to understand and reflect such features in random networks?

Milgram's Small-World Experiment

Question: How many social steps separate two random people?

Psychologist Stanley Milgram in 1960s conducts a famous field experiment:

- Participants in the U.S. were asked to send a letter to a target person in Boston
- They could only forward it to someone they personally knew
- Letters traveled through chains of acquaintances
- Milgram asked participants to add a new stamp to letter after sending to keep track of path lengths

Result:

- Successful chains had average length ≈ 6
- Leads to famous concept of "six degrees of separation"
- Most people are 6 steps away from each other (not really!)

Neighborhood Expansion

If average degree ≈ 100:

$$100^1 = 100$$

$$100^2 = 10,000$$

$$100^3 = 1,000,000$$

$$100^4 = 100,000,000$$

Rapid coverage of population.

But Links Are Not Random

Real networks show:

- Clusters (local pockets)
- Long-range ties (tendrils)

This combination generates small worlds.

Examples

- Actor networks (Bacon number)
- Publication networks (Erdős number)
- Citation networks
- Early MySpace (Tom!)

Key Takeaways

- Configuration model preserves degree sequence
- SBM introduces community structure
- Small-world networks combine clustering and short paths
- Relate structured randomness to real network features