Outline

- Overview
  What is a random graph?

- Historical Context
  The Erdös-Rényi random graph and its properties
  Random regular graph

- Application to Networks
  Wattz-Strogatz random graph model
  Random graphs as network models

- Small-world Experiment
  Milgram letters
  Complex networks and decentralized search
What is a Random Graph?

Random graphs may be described as a probability distribution over the set of all possible graphs on \( n \) vertices.

- Starting with a set of \( n \) isolated vertices, add edges between vertices based on a set of rules or a probability function.

- Different random graph models produce different probability distributions on graphs.

**Example: G(n, p)**

G is sampled from a distribution with \( n = 6 \) and \( p = 1/6 \)

![Graph Example](image)
The Erdös-Rényi (ER) Random Graph

**Definition:** Given a positive integer \( n \) and a probability value \( 0 \leq p \leq 1 \), define the graph \( G(n, p) \) to be the undirected graph on \( n \) vertices whose edges are chosen as follows: for all pairs of vertices \( v, w \) there is an edge \( (v, w) \) with probability \( p \).

The theory of random graphs began with a paper by Erdös and Rényi in the late 1950s. It was originally devised as a mathematical tool for existence proofs.
Theorem: For a constant $p$, $G(n, p)$ has a diameter of 2 (w.h.p.).
(the diameter is the longest shortest path between two vertices)

Let $X$ = the number of vertex pairs with no common neighbour
Show that $X = 0$ (w.h.p.)

Let $X_{u,v} = \begin{cases} 
1 & \text{if } u \text{ and } v \text{ do not have a common neighbour} \\
0 & \text{if they do}
\end{cases}$

$Pr(X_{u,v} = 1) = (1 - p^2)^{n-2}$
$\leftarrow (1 - p^2) = (1 - Pr(u \text{ and } v \text{ have a common neighbour}))$

$E(X) = \sum_{u,v} E(X_{u,v}) = \left( \begin{array}{c} n \\ 2 \end{array} \right) (1-p^2)^{n-2}$

$\lim_{n \to \infty} E(X) = 0$

$Pr(X \geq 1) \leq E(X) / 1 = 0 \quad \leftarrow \text{Markov’s Inequality}$

Therefore $Pr(X = 0) = 1$
Properties of the ER Random Graph (2)

What is the expected number of triangles in a random graph with $p$ as a function of $n$?

Let $p = c/n$
Let $X = \text{the number of triangles in } G(n, c/n)$

$$E(X) = \binom{n}{3} (c/n)^3 \quad \leftarrow \text{(vertex triple combinations) } \times \text{ (prob. of an edge between them)}$$

$$\lim_{n \to \infty} E(X) = \frac{c^3}{6}$$

The number of triangles is independent of $n$. 
**The Random Regular Graph**

**Definition:** Given positive integers $n$ and $r$ where $nr$ is even, define the graph $G(n, r)$ to be the undirected graph on $n$ vertices where every vertex has degree $r$.

In this model the existence of different edges is not independent.

**Example: $G(n, r)$**

$G$ is sampled from a distribution with $n = 10$ and $r = 3$
Watts and Strogatz Random Graph Model

**Construction:** Construct a ring lattice with \( n \) nodes each connected to \( k \) neighbours. For each \( i^{th} \) node, with probability \( p \) rewire each edge \((i, j)\) so that \((i, j)\) becomes \((i, l)\) where \( l \) is chosen with uniform probability from all possible values that avoid self-loops \((l \neq i)\) and link duplication (there is no edge \((i, l')\) with \( l' = l \)).

\[ n = 12, \ k = 4, \ p = 0.1 \]
Watts and Strogatz Random Graph Model

- A standard random graph $G(n, p(n))$ is locally very sparse but in real network data, many of a node’s neighbours are joined to each other by edges.

- The Watts and Strogatz model demonstrates this with structured local clustering in addition to random links.
Complex Networks

Real networks are complex

Example:

- The brain consists of $10^{11}$ neurons (vertices) connected to one another by axons (edges) interacting in an intricate way.

- The range of factors that contribute to the observed structure will be too intricate to be captured in a simple model.
Random Graphs as Network Models

Recent cross-disciplinary work has sought to develop random graph models that capture the qualitative properties found in large social, technological, and information networks.

- Using local and probabilistic rules we can describe the complexity of the network.

- Random graph models that capture some of the qualitative properties observed in network data have the potential to help us reason, at a general level, about the ways in which real-world networks are organized.
An Example: Small-world Experiment

Stanley Milgram, a social psychologist conducted this experiment in the 1960s.

- Letters were mailed randomly to individuals living in Nebraska.
- They were told to forward the letter to a target individual in Boston using only their social contacts.
- Most letters arrived after only 6 steps.

Considering the population of the United States, how could this happen…

Stanley Milgram
In his paper *Complex networks and decentralized search algorithms* Klienberg describes a random graph that models a social network of acquaintances based on the Milgram experiment.

A decentralized algorithm is proposed that will find short paths through the network without knowledge of the complete network.

**Jon Kleinberg** is a computer scientist at Cornell University known for his work in algorithms and networks.
The Random Graph Model

Start with a $n \times n$ lattice of nodes

**distance:** $d((i, j), (k, l))$ is the number of lattice steps between node $(i, j)$ and node $(k, l)$

$$d(u, v) = 4$$
Local Contacts

**p**: for a universal constant $p \geq 1$, the node $u$ has a directed edge to every other node within a distance $p$.

$p = 2$
Long Range Contacts

$q$: for universal constants $q \geq 1$ and $r \geq 1$, the node $u$ has directed edges to $q$ other nodes using independent random trials.

$p = 1, \ q = 2$
Long Range Contacts

inverse $r^{th}$ power distribution:

the $i^{th}$ directed edge from $u$ has endpoint $v$ with probability:

$$\Pr = \frac{[d(u, v)]^{-r}}{\sum_v [d(u, v)]^{-r}}$$

$p = 1$, $q = 1$, $r = 2$
Long Range Contacts

choose \( r = 2 \)

if \( r = 0 \) we have a uniform distribution over long range contacts

if \( r \) is large then the long range contacts become clustered around \( u \)

\[
Pr = \frac{[d(u, v)]^{-r}}{\sum_v [d(u, v)]^{-r}}
\]
Pass a message from s to t

The message holder $u$ has knowledge of:

i) local contacts of all nodes
ii) target $t$
iii) the location of all nodes that have come in contact with the message

$p = 1, q = 1, r = 2$
Algorithm

Decentralized Navigation \((G, s, t)\)

\(u \leftarrow s\)

while \((u \neq t)\) do
    \(u \leftarrow\) the node \(v\) that is as close to the target as possible and there is an edge \((u, v)\)

\(O(\log^2 n)\)
d(v, t) = 11
$d(v, \ t) = 10$
\[ d(v, t) = 7 \]
\[ d(v, t) = 1 \]
\[ d(v, t) = 0 \]
Analysis

**Theorem:** The algorithm with \( r = 2 \) and \( p = q = 1 \) finds the target node in \( O(\log^2 n) \) time when run on the described random graph model.

The complete proof can be read in *The Small-World Phenomenon: An Algorithmic Perspective* by Jon Kleinberg.

It is based on the following properties:

- long range contacts are nearly uniformly distributed over distance scales
- define **phase** \( j \): \( d(u, t) \) is between \( 2^j \) and \( 2^{j+1} \)
- time spent in each phase \( j \) is bounded by \( O(\log n) \)
- at most \( 1 + \log(n) \) phases

Time to find target = \( O(\log^2 n) \)
References


Jon Klienberg, *Complex Networks and Decentralized Search Algorithms.*

Joel Spencer, *Nine Lectures on Random Graphs.*

Matthias Grossglauser, Patrick Thiran, *Networks out of Control: Models and Methods for Random Networks.*


http://jeremykun.com/2013/08/22/the-erdos-renyi-random-graph/
We are being controlled by the random outcomes of a complex system.