Random Graphs

Social and Technological Networks

Rik Sarkar

University of Edinburgh, 2019.

Today

- What are random graphs?
  - How can we define “random graphs”?
- Some properties of random graphs

Erdos – Renyi Random graphs

Erdos – Renyi Random graphs

$$\mathcal{G}(n, p)$$

- n: number of vertices
- p: probability that any particular edge exists

  - Take V with n vertices
  - Consider each possible edge. Add it to E with probability p

Expected number of edges in an ER graph

- Expected total number of edges
- Expected number of edges at any vertex

Expected number of edges

- Expected total number of edges $\binom{n}{2}p$

- Expected number of edges at any vertex

$$(n-1)p$$

Expected number of edges

- For $p = \dfrac{c}{n-1}$
- The expected degree of a node is : ?

Isolated vertices

- How likely is it that the graph has isolated vertices?

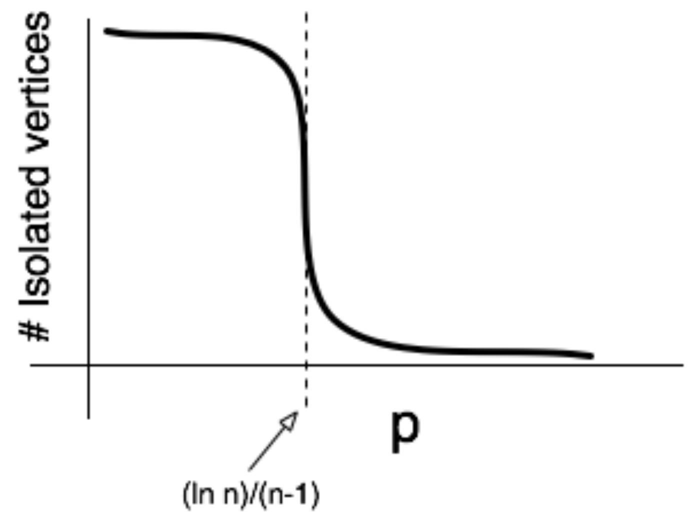
Isolated vertices

- How likely is it that the graph has isolated vertices?

- What happens to the number of isolated vertices as p increases?

Threshold phenomenon: Probability or number of isolated vertices

- The tipping point, phase transition



- Common in many real systems

Probability of Isolated vertices

- Isolated vertices are

- Likely when: $p < \frac{\ln n}{n}$

- Unlikely when: $p > \frac{\ln n}{n}$

- Let's deduce

Terminology of high probability

- Poly(n) means a polynomial in n
- A polynomial in n is considered reasonably ‘large’
  - Whereas something like constant, or log n is considered ‘small’

- Something happens with high probability if

$$\Pr[event] \geq \left(1 - \frac{1}{\mathrm{poly}(n)}\right)$$

- Thus for large n, w.h.p there is no isolated vertex
- Expected number of isolated vertices is miniscule

Useful inequalities

$$\left(1 + \frac{1}{x}\right)^x \leq e$$

$$\left(1 - \frac{1}{x}\right)^x \leq \frac{1}{e}$$

Union bound

- For events A, B, C …

- Pr[A or B or C …] ≤ Pr[A] + Pr[B] + Pr[C] + …

- Theorem 1:
- If $$p = (1+\epsilon)\frac{\ln n}{n-1}$$
- Then the probability that there exists an isolated vertex $$\leq \frac{1}{n^{\epsilon}}$$

- Theorem 2
- For $p = (1-\epsilon)\frac{\ln n}{n-1}$
- Probability that vertex v is isolated $\geq \frac{1}{(2n)^{1-\epsilon}}$

- Theorem 2
- For $p = (1-\epsilon)\frac{\ln n}{n-1}$
- Probability that vertex v is isolated $\geq \frac{1}{(2n)^{1-\epsilon}}$
- Expected number of isolated vertices:

$$\geq \frac{n}{(2n)^{1-\epsilon}} = \frac{n^{\epsilon}}{2}$$

Polynomial in n

Global CC in ER graphs

- What happens when p is very small (almost 0)?
- What happens when p is very large (almost 1)?

Global CC in ER graphs

- What happens at the tipping point?

Theorem

- For $p = c\frac{\ln n}{n}$

- Global cc in ER graphs is vanishingly small

$$\lim_{n\to\infty} cc(G) = \lim_{n\to\infty} \frac{\#\text{ closed triads}}{\#\text{ all triads}} = 0$$

- In other words, there is no constant c
  - Such that cc(ER-graph) > c
  - At the tipping point

Random graphs: Emergence of giant component

- Suppose N_G is the size of the largest connected component in an ER graph

- How does N_G/N change with p?

- When is N_G/N at least a constant?
  - (giant component: at least a constant fraction of nodes)

Giant component

- When $p = (1-\varepsilon)/n$
  - W.h.p no GC, components of size $O(\log n)$
- When $p = (1+\varepsilon)/n$
  - W.h.p GC exists, where $N_G/N \sim \varepsilon$
- When $p = 1/n$
  - W.h.p Largest component has size $n^{2/3}$

Project

- Topics will be out today.
- Check email/piazza

- You can work/discuss in groups 2 of 3
  - But final submission is individual.

- Select Project by Friday Oct 11
  - Don’t rush. Think carefully what you might do in that topic. Look at the dataset.
  - Once selected. Don’t delay. Start working immediately.

- Final submission due: Thursday November 14.

Project

- October 21:
  - Due date for midway project plan.
- Submit a 1 page description of topic and problem statement, and how you plan to approach it. And get comments.
- No marks. Feedback only.
- Office hours: Tuesdays 2pm. IF 3.45.