1.022 Introduction to Network Models

Amir Ajorlou

Laboratory for Information and Decision Systems
Institute for Data, Systems, and Society
Massachusetts Institute of Technology

Lecture 9

- ER graphs exhibit phase transitions

  $\Rightarrow$ Sharp transitions between behaviors as $n \to \infty$

ER connectivity theorem

- A threshold function for the connectivity of $G_{n,p(n)}$ is $p(n) = \frac{\ln(n)}{n}$
- Let $p(n) = \lambda \frac{\ln(n)}{n}$ then

  $\Rightarrow$ If $\lambda < 1 \;\Rightarrow \mathbb{P}(connected) \to 0$ as $n \to \infty$

  $\Rightarrow$ If $\lambda > 1 \;\Rightarrow \mathbb{P}(connected) \to 1$ as $n \to \infty$

Threshold for Connectivity

- To show disconnectedness, it is sufficient to show that the probability that there exists at least one isolated node goes to 1.
- Let I_i be a Bernoulli random variable defined as

$$I_i = \begin{cases} 1 & \text{if node } i \text{ is isolated,} \\ 0 & \text{otherwise.} \end{cases}$$

- We can write the probability that an individual node is isolated as

$$q = \mathbb{P}(I_i = 1) = (1-p)^{n-1} \approx e^{-pn} = e^{-\lambda \log(n)} = n^{-\lambda},$$

  where we use $\lim_{n\to\infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}$ to get the approximation.
- Let $X = \sum_{i=1}^{n} I_i$ denote the total number of isolated nodes. Then, we have

$$\mathbb{E}[X] = n \cdot n^{-\lambda}.$$

Sketch of the Proof

- For $\lambda < 1$, we have $\mathbb{E}[X] \to \infty$. We want to show that this implies $\mathbb{P}(X = 0) \to 0$.
- In general, this is not true. But, here it holds.
- We can show that the variance of X is of the same order as its mean:

$$\text{var}(X) \sim \mathbb{E}[X],$$

where $a(n) \sim b(n)$ denotes $\frac{a(n)}{b(n)} \to 1$ as $n \to \infty$.
- This implies that

$$\mathbb{E}[X] \sim \text{var}(X) \geq (0 - \mathbb{E}[X])^2 \mathbb{P}(X = 0),$$

and therefore,

$$\mathbb{P}(X = 0) \leq \frac{\mathbb{E}[X]}{\mathbb{E}[X]^2} = \frac{1}{\mathbb{E}[X]} \to 0.$$

- It follows that $\mathbb{P}(\text{at least one isolated node}) \to 1$ and therefore, $\mathbb{P}(\text{disconnected}) \to 1$ as $n \to \infty$, completing the proof.

Converse Sketch (Optional)

- If $p(n) = \lambda \frac{\log(n)}{n}$ with $\lambda > 1$, then $\mathbb{P}(\text{disconnected}) \to 0$.
  - $\mathbb{E}[X] = n^{1-\lambda} \to 0$ for $\lambda > 1$. Almost surely no isolated node.
  - We need more to establish connectivity.
  - The event "graph is disconnected" is equivalent to the existence of k nodes without an edge to the remaining nodes, for some $k \leq n/2$.
  - We have

$$\mathbb{P}(\{1,\ldots,k\} \text{ not connected to the rest}) = (1-p)^{k(n-k)} \Rightarrow$$

$$\mathbb{P}(\exists \text{ k nodes not connected to the rest}) = \binom{n}{k}(1-p)^{k(n-k)} \Rightarrow$$

$$\mathbb{P}(\text{disconnected graph}) \leq \sum_{k=1}^{n/2} \binom{n}{k}(1-p)^{k(n-k)}.$$

  - bounding RHS and some algebraic manipulation yields

$$\mathbb{P}(\text{disconnected graph}) \leq Cn^{-1+\lambda} \stackrel{\lambda>1}{\to} 0.$$

Phase Transitions — Connectivity Threshold

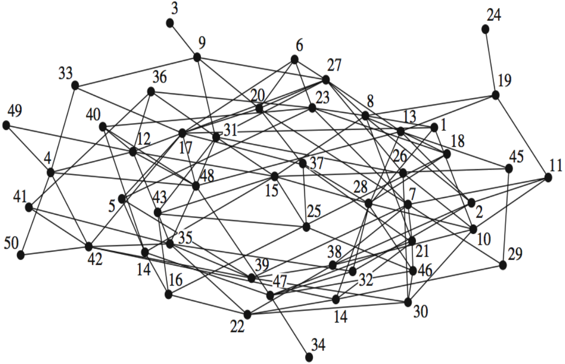

Figure: Emergence of connectedness: a random network on 50 nodes with $p = 0.10$.

ER giant component theorem

- A threshold function for the emergence of a giant component in $G_{n,p(n)}$ is $p(n) = \frac{1}{n}$
- Let $p(n) = \frac{\lambda}{n}$ then

  $\Rightarrow$ If $\lambda < 1$ $\Rightarrow$ Size of largest component $\sim \ln(n)$ as $n \to \infty$

  $\Rightarrow$ If $\lambda > 1$ $\Rightarrow$ Size of largest component $\sim n$ as $n \to \infty$

- In fact, the size of giant component satisfies

$$1 - q = e^{-\lambda q}$$

Giant Component Size

- q: giant component size, the probability of a randomly chosen node is in giant component
- Consider a vertex not in the giant component: For every other vertex j either
  - i is not connected to j by an edge, or
  - i is connected to j but j is not in the giant component.
- This gives

$$1 - q = (1 - p + p(1 - q))^{n-1} = (1 - pq)^{n-1},$$

RHS can be approximated as $e^{-p(n-1)q} \sim e^{-\lambda q}$ when $n \to \infty$.

- $q = 0$ is always a solution of $1 - q = e^{-\lambda q}$.
- looking at the derivative of both sides at $q = 0$, we can show the existence of a nonzero solution if and only if $\lambda > 1$.