Algorithmic and Economic Aspects of Networks

Nicole Immorlica

Random Graphs

What is a random graph?

Erdos-Renyi Random Graphs

Specify

number of vertices n

edge probability p

For each pair of vertices i < j,

create edge (i,j) w/prob. p

G(n,p)

Erdos-Renyi Random Graphs

What does random graph $G(n,p)$ look like?

(as a function of p)

Random Graph Demo

http://ccl.northwestern.edu/netlogo/models/GiantComponent

Properties of G(n,p)

$p < 1/n$	disconnected, small tree-like components
$p > 1/n$	a giant component emerges containing const. frac. of nodes

Proof Sketch

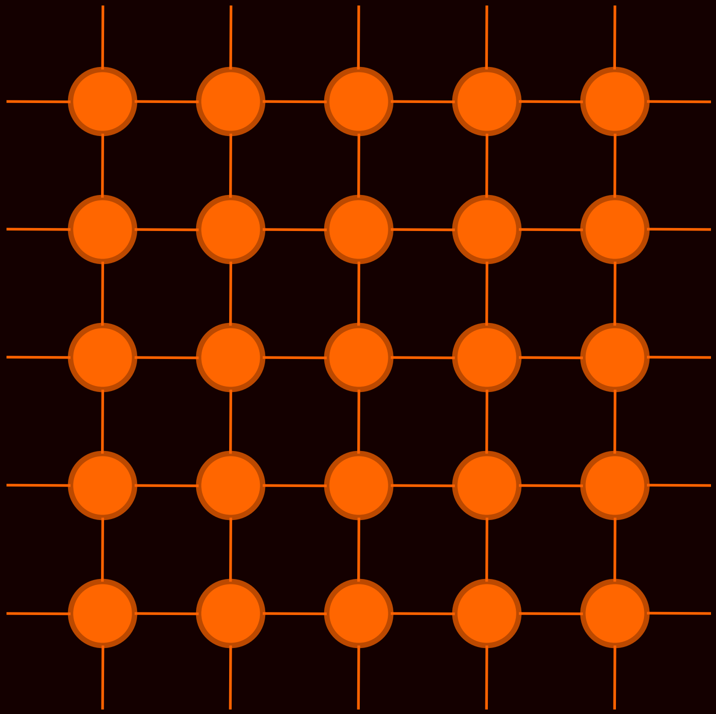
1. Percolation
2. Branching processes
3. Growing spanning trees

Percolation

1. Infinite graph
2. Distinguished node i
3. Probability p

Each link gets ``open'' with probability p

Q. What is size of component of i?



Percolation Demo

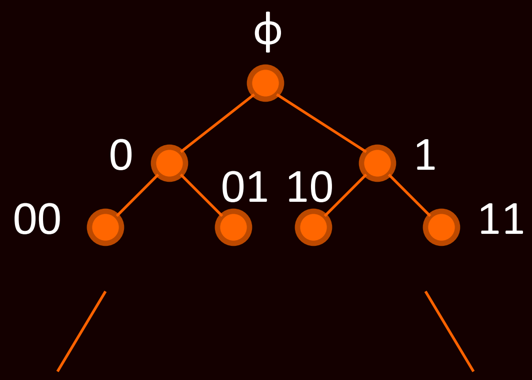
http://ccl.northwestern.edu/netlogo/models/Percolation

Percolation on Binary Trees

$V = \{0,1\}^*$

$E = (x,y)$ s.t. $y = x0$ or $y = x1$

distinguished node ϕ



Def. Let $\Theta(p) = \Pr[\text{comp}(\phi) \text{ is infinite}]$. The critical threshold is $p_c = \sup \{ p \mid \Theta(p) = 0\}$.

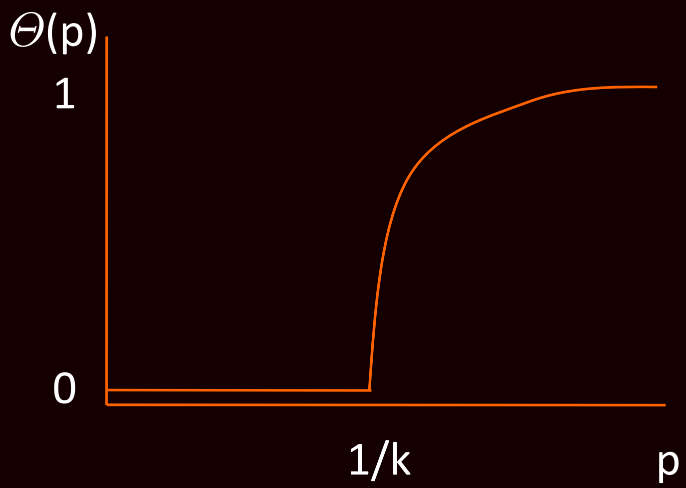
Critical Threshold

Def. Let $\Theta(p)$ = Pr[comp(ϕ) is infinite]. The critical threshold is $p_c = \sup \{ p \mid \Theta(p) = 0\}$.

Thm. Critical threshold of binary trees is $p_c = ½$.

Prf. On board.

Critical Threshold



Thm. Critical threshold of k-ary trees is $p_c = 1/k$.

Branching Processes

Node i has X_i children distributed as B(n,p):

$$\Pr[X_i = k] = (n \text{ choose } k)\, p^k (1-p)^{(n-k)}$$

Q. What is probability species goes extinct?

A. By percolation, if $p > (1+\epsilon)/n$, live forever.

Note extinction ←→ Exists i, $X_1 + \ldots + X_i < i$.

Erdos-Renyi Random Graphs

We will prove (on board)

(1) If $p = (1-\epsilon)/n$, then there exists c_1 s.t. Pr[G(n,p) has comp > c_1 log n] goes to zero

(2) If $p = (1+2\epsilon)/n$, then there exists c_2 s.t. Pr[G(n,p) has comp > c_2 n] goes to one

First show (on board)

(3) If $p = (1+2\epsilon)/n$, then there exists c_2, c_3 s.t. Pr[G(n,p) has comp > c_2 n] > c_3

Emergence of Giant Component

Theorem. Let $np = c < 1$. For $G \in G(n, p)$, w.h.p. the size of the largest connected component is $O(\log n)$.

Theorem. Let $np = c > 1$. For $G \in G(n, p)$, w.h.p. G has a giant connected component of size $(\beta + o(n))n$ for constant $\beta = \beta c$; w.h.p, the remaining components have size $O(\log n)$.

Application

Suppose …

the world is connected by G(n,p)

someone gets sick with a deadly disease

all neighbors get infected unless immune

a person is immune with probability q

Q. How many people will die?

Analysis

1. Generate G(n,p)
2. Delete qn nodes uniformly at random
3. Identify component of initially infected individual

Analysis

Equivalently,

1. Generate G((1-q)n, p)
2. Identify component of initially infected individual

Analysis

By giant component threshold,

- $p(1-q)n < 1$ → disease dies
- $p(1-q)n > 1$ → we die

E.g., if everyone has 50 friends on average, need prob. of immunity = 49/50 to survive!

Summary

Random graphs G(n, c/n) for c > 1 have …

- ✓ unique giant component
- ✓ small (logarithmic) diameter
- ✗ low clustering coefficient (= p)
- ✗ Bernoulli degree distribution

A model that better mimics reality?

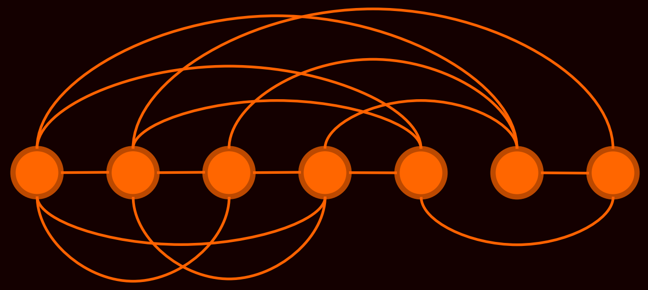
In real life

Friends come and go over time.

Growing Random Graphs

On the first day, God created

m+1 nodes who were all friends



And on the (m+i)'th day, He created

a new node (m+i) with m random friends

Mean Field Approximation

Estimate distribution of random variables by distribution of *expectations*.

E.g., degree dist. of growing random graph?

Degree Distribution

$$F_t(d) = 1 - \exp[\,-(d - m)/m\,]$$

(on board)

This is exponential, but social networks tend to look more like power-law deg. distributions…

In real life

The rich get richer

… much faster than the poor.

Preferential Attachment

Start: m+1 nodes all connected

Time t > m: a new node t with m friends distributed according to degree

$$\Pr[\text{link to } j] = m \times \deg(j) / \sum \deg(.)$$

$$= m \times \deg(j) / (2mt)$$

$$= \deg(j) / (2t)$$

Degree Distribution

Cumulative dist.: $F_t(d) = 1 - m^2/d^2$

Density function: $f_t(d) = 2m^2/d^3$

(heuristic analysis on board,

for precise analysis, see Bollobas et al)

A power-law!

Assignment:

- Readings:
  - Social and Economic Networks, Chapters 4 & 5
  - M. Mitzenmacher. *A brief history of generative models for power law and lognormal distributions*. Internet Mathematics 1, No 2, 226-251, 2005.
  - D.J. Watts, and S.H. Strogatz. *Collective dynamics of small-world networks*. Nature 393, 440-442, 1998.
- Reactions:
  - Reaction paper to one of research papers, or a research paper of your choice
- Presentation volunteer?