The giant component in a given graph

Fan Chung[1], Paul Horn [1], Linyuan Lu [2]

[1] Department of Mathematics
University of California, San Diego

[2] Department of Mathematics
University of South Carolina

February 12, 2009

Motivation

For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

Goal: Study these problems in the most general (i.e. non-model specific) way we can

Motivation

For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

Goal: Study these problems in the most general (i.e. non-model specific) way we can

Motivation

For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

Goal: Study these problems in the most general (i.e. non-model specific) way we can

Motivation

For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

Goal: Study these problems in the most general (i.e. non-model specific) way we can

Motivation

For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

Goal: Study these problems in the most general (i.e. non-model specific) way we can

Motivation

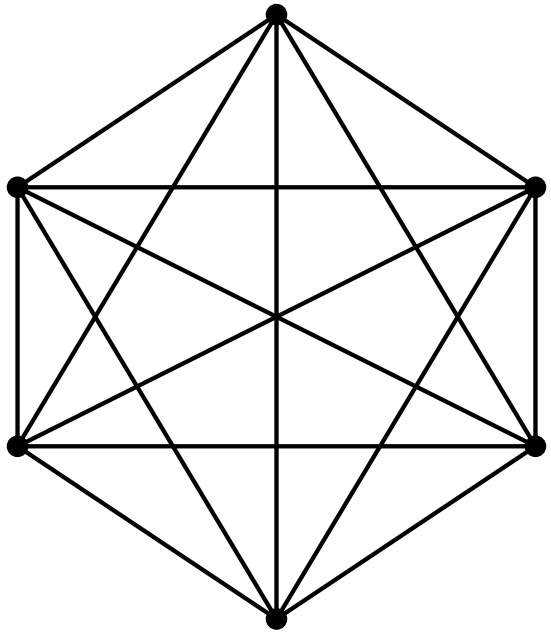
For real world graphs:

- Often incomplete data on large networks
- Disease spreads over contact networks
- Computer viruses propagate in computer networks
- Many graphs we observe are subgraphs of host graphs.

How do random subgraphs relate to their underlying host graph?

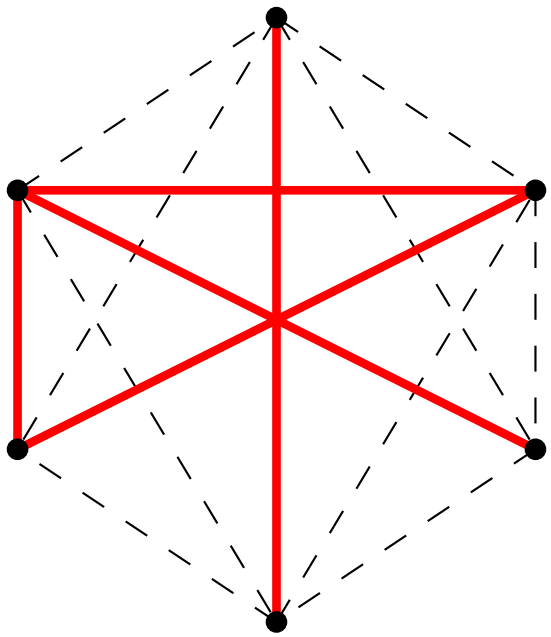
Goal: Study these problems in the most general (i.e. non-model specific) way we can

A Model of Random Subgraphs



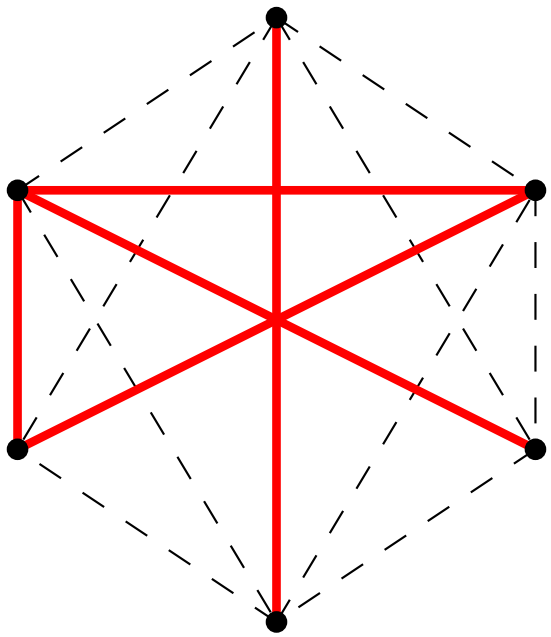
- Suppose $G = K_n$
- Take edges with probability p, discard with probability $1 - p$
- Erdős-Rényi $\mathcal{G}_{n,p}$ model.

A Model of Random Subgraphs



- Suppose $G = K_n$
- Take edges with probability p, discard with probability $1 - p$
- Erdős-Rényi $\mathcal{G}_{n,p}$ model.

A Model of Random Subgraphs



- Suppose $G = K_n$
- Take edges with probability p, discard with probability $1 - p$
- Erdős-Rényi $\mathcal{G}_{n,p}$ model.

Asymptotic Notation

Throughout the talk we use the following asymptotic notation. If $f(n), g(n)$ are functions:

- $f(n) = O(g(n))$ if $\limsup_{n\to\infty} \frac{|f(n)|}{|g(n)|} < \infty$.
- $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{|f(n)|}{|g(n)|} = 0$.
- $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

Properties of random graphs

Property $\mathcal{A}$ occurs a.a.s. $\Rightarrow \mathbb{P}(A) \to 1$ as $n \to \infty$.

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components have size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$ a.a.s.

and

- $p < (1-\epsilon)\frac{\log n}{n}$: G disconnected a.a.s.
- $p > (1+\epsilon)\frac{\log n}{n}$: G connected a.a.s.

Properties of random graphs

Property $\mathcal{A}$ occurs a.a.s. $\Rightarrow \mathbb{P}(A) \to 1$ as $n \to \infty$.

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components have size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$ a.a.s.

and

- $p < (1-\epsilon)\frac{\log n}{n}$: G disconnected a.a.s.
- $p > (1+\epsilon)\frac{\log n}{n}$: G connected a.a.s.

Properties of random graphs

Property $\mathcal{A}$ occurs a.a.s. $\Rightarrow \mathbb{P}(A) \to 1$ as $n \to \infty$.

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components have size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$ a.a.s.

and

- $p < (1-\epsilon)\frac{\log n}{n}$: G disconnected a.a.s.
- $p > (1+\epsilon)\frac{\log n}{n}$: G connected a.a.s.

Properties of random graphs

Property $\mathcal{A}$ occurs a.a.s. $\Rightarrow \mathbb{P}(A) \to 1$ as $n \to \infty$.

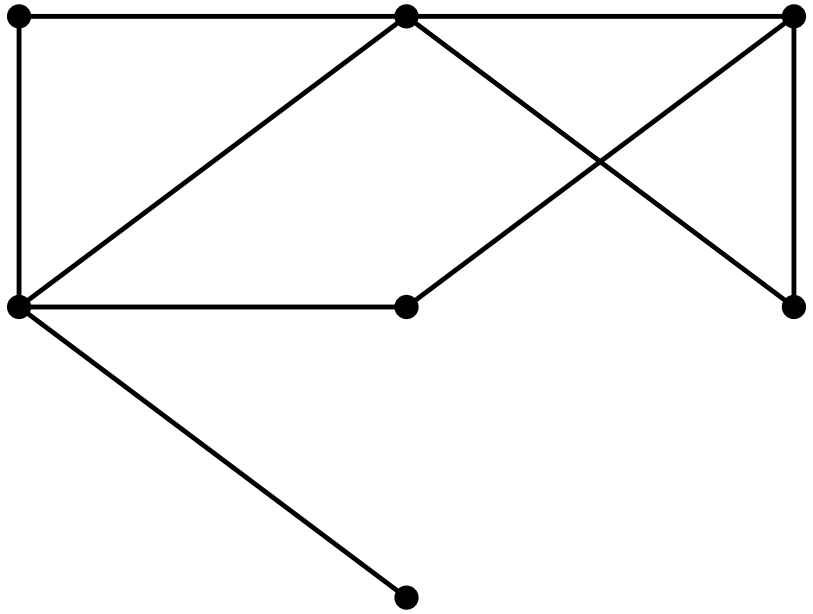
Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components have size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$ a.a.s.

and

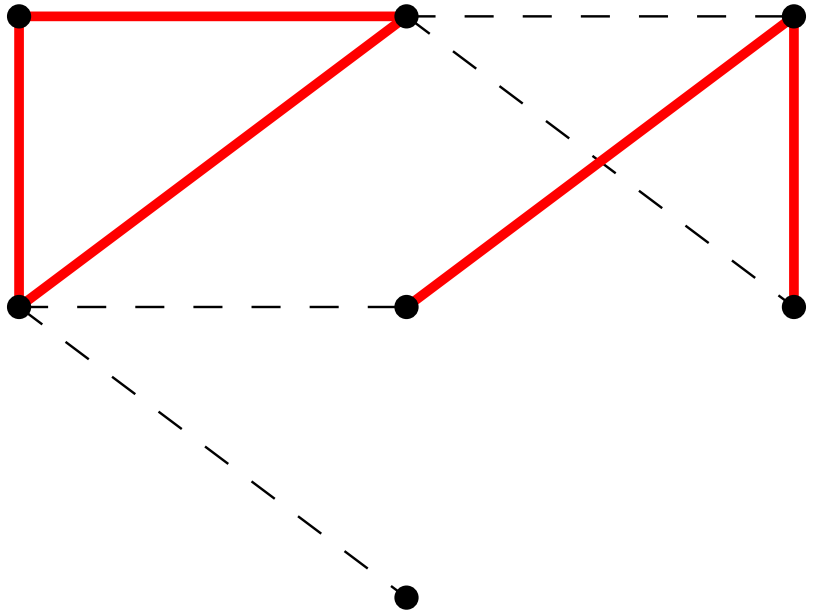
- $p < (1-\epsilon)\frac{\log n}{n}$: G disconnected a.a.s.
- $p > (1+\epsilon)\frac{\log n}{n}$: G connected a.a.s.

A model of random subgraph



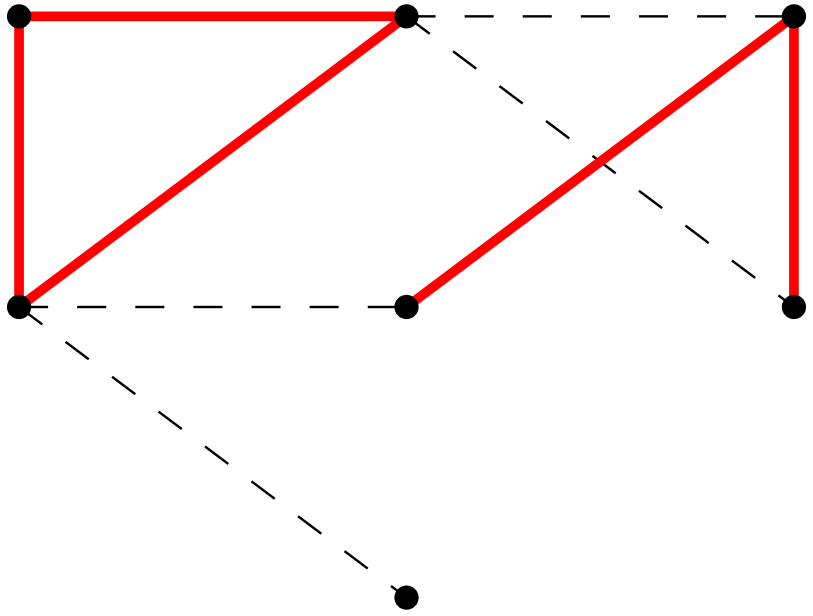
- Start with a general graph
- Take edges with probability p, discard with probability $1 - p$
- Known as bond-percolation in probability literature.

A model of random subgraph



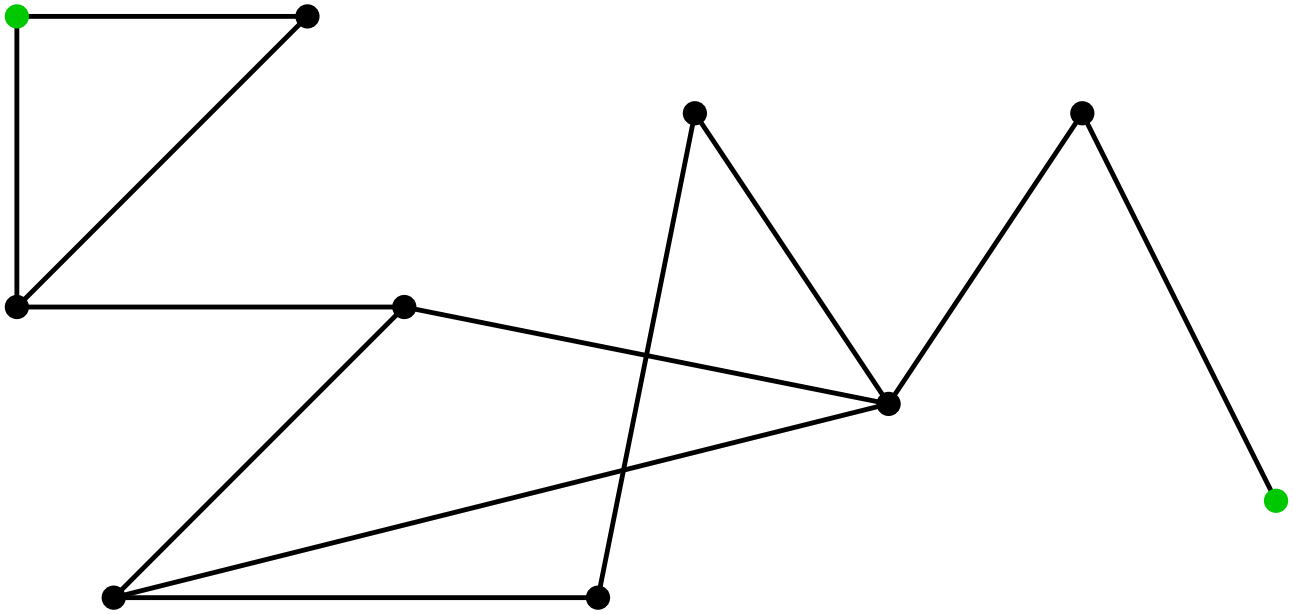
- Start with a general graph
- Take edges with probability p, discard with probability $1 - p$
- Known as bond-percolation in probability literature.

A model of random subgraph



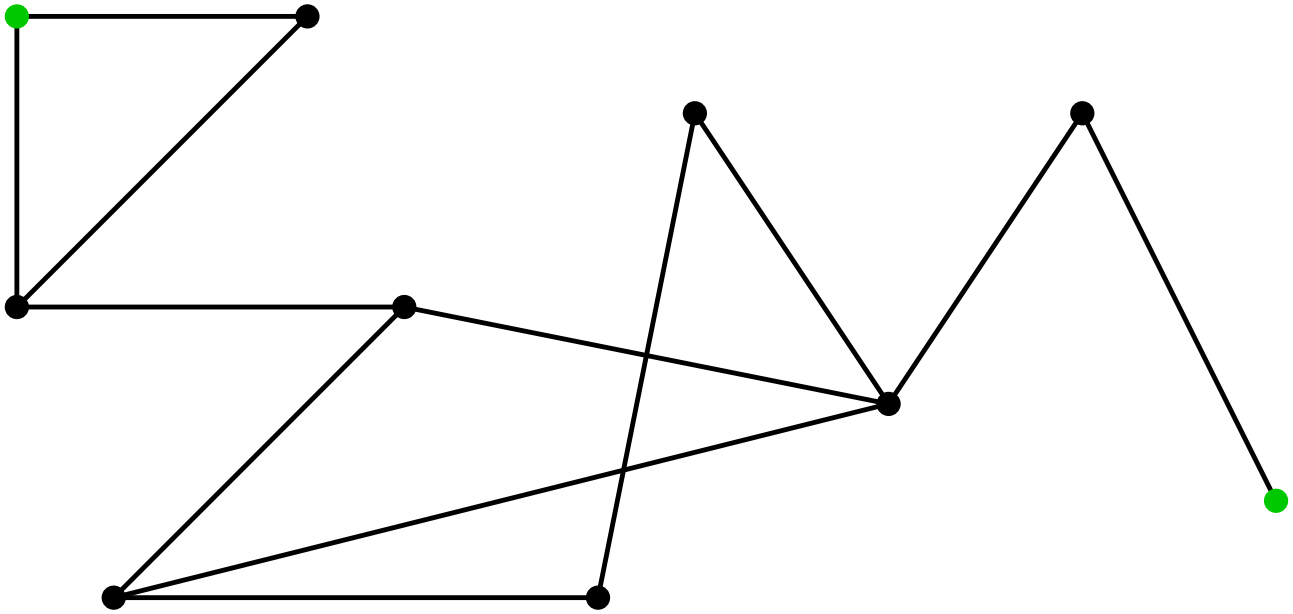
- Start with a general graph
- Take edges with probability p, discard with probability $1 - p$
- Known as bond-percolation in probability literature.

Epidemiological question



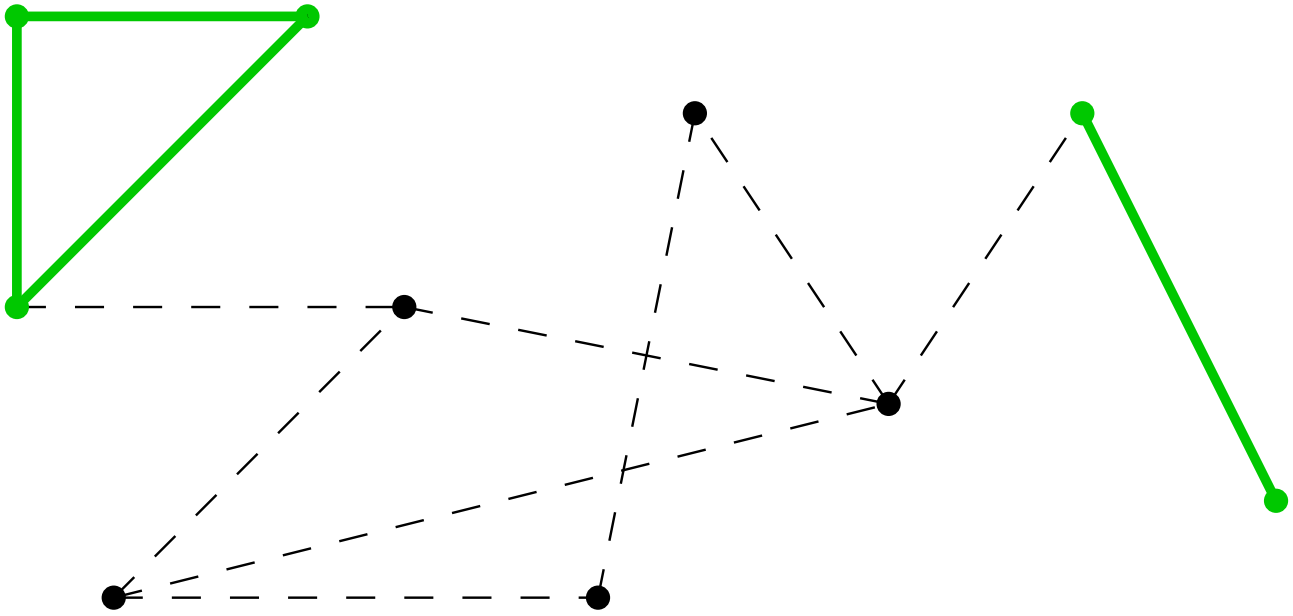
- Take host graph to be contact graph, some initially infected vertices.
- p is transmission rate of disease

Epidemiological question



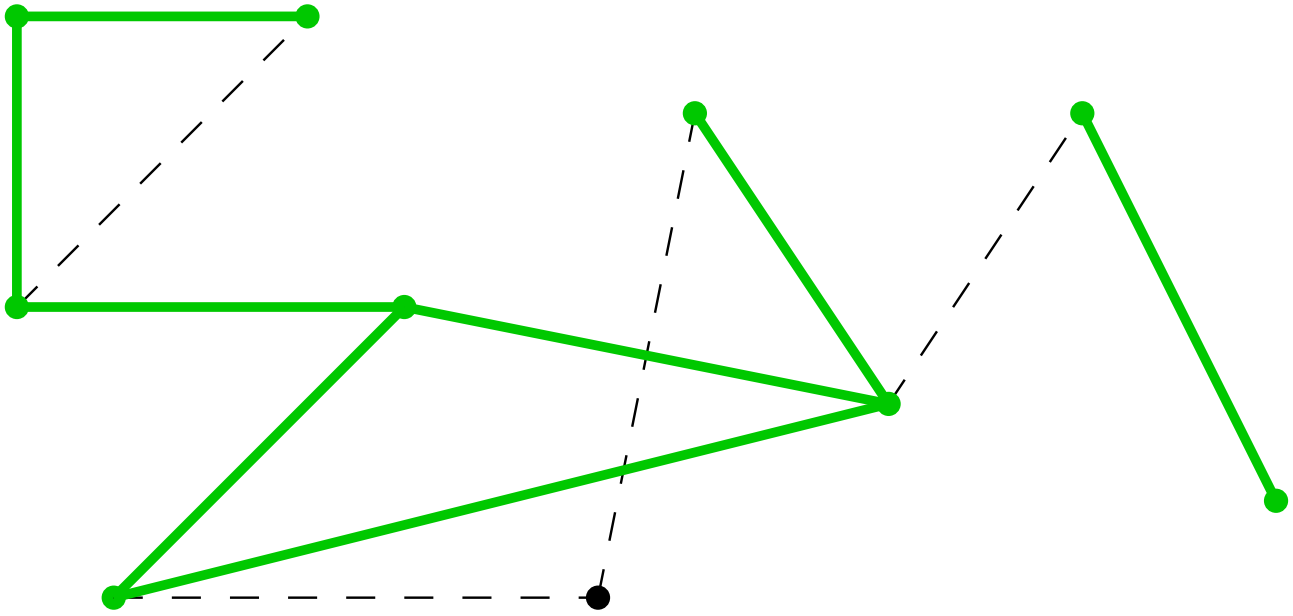
- Take host graph to be contact graph, some initially infected vertices.
- p is transmission rate of disease

Epidemiological question



- Take host graph to be contact graph, some initially infected vertices.
- p is transmission rate of disease
- Does the disease die out quickly? (no large component)

Epidemiological question



- Take host graph to be contact graph, some initially infected vertices.
- p is transmission rate of disease
- Or is there a chance that it grows into an epidemic? (giant component)

Questions

- Threshold in this model?
- How can we get the control we need to prove rigorous theorems?

Questions

- Threshold in this model?
- How can we get the control we need to prove rigorous theorems?

Notation

- For a vertex v, d_v denotes its degree.
- Let $X \subseteq V(G)$. Then

$$\mathrm{vol}_k(X) = \sum_{v \in X} d_v^k$$

  We let $\mathrm{vol}(X) = \mathrm{vol}_1(X)$.
- We let

$$d = \frac{\mathrm{vol}(G)}{n} = \frac{\mathrm{vol}_1(G)}{\mathrm{vol}_0(G)} \qquad \tilde{d} = \frac{\mathrm{vol}_2(G)}{\mathrm{vol}(G)}$$

  denote the first and second order average degrees of G.
- If G is regular, $d = \tilde{d}$.

How to understand general graphs?

Spectral Graph Theory!

- Associate a matrix with a graph
- Use the eigenvalues of the matrix to understand properties of graphs

How to understand general graphs?

Spectral Graph Theory!

- Associate a matrix with a graph
- Use the eigenvalues of the matrix to understand properties of graphs

How to understand general graphs?

Spectral Graph Theory!

- Associate a matrix with a graph
- Use the eigenvalues of the matrix to understand properties of graphs

How to understand general graphs?

Spectral Graph Theory!

- Associate a matrix with a graph
- Use the eigenvalues of the matrix to understand properties of graphs

The Adjacency Matrix

There are many ways of associating a matrix with G. Perhaps the simplest is the adjacency matrix:

$$A = [a_{i,j}]; a_{i,j} = \begin{cases} 1 & \text{if } v_i \sim v_j \\ 0 & \text{otherwise.} \end{cases}$$

Pros/Cons

- G d-regular $\Rightarrow$ eigenvalues and properties closely related.
- G not regular $\Rightarrow$ eigenvalues can be dominated by degrees.

The Adjacency Matrix

There are many ways of associating a matrix with G. Perhaps the simplest is the adjacency matrix:

$$A = [a_{i,j}]; a_{i,j} = \begin{cases} 1 & \text{if } v_i \sim v_j \\ 0 & \text{otherwise.} \end{cases}$$

Pros/Cons

- G d-regular $\Rightarrow$ eigenvalues and properties closely related.
- G not regular $\Rightarrow$ eigenvalues can be dominated by degrees.

The Adjacency Matrix

There are many ways of associating a matrix with G. Perhaps the simplest is the adjacency matrix:

$$A = [a_{i,j}]; a_{i,j} = \begin{cases} 1 & \text{if } v_i \sim v_j \\ 0 & \text{otherwise.} \end{cases}$$

Pros/Cons

- G d-regular $\Rightarrow$ eigenvalues and properties closely related.
- G not regular $\Rightarrow$ eigenvalues can be dominated by degrees.

The (normalized) Laplacian

- Normalized Laplacian:

$$\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}$$

$A =$ adjacency matrix, $D =$ diagonal degree matrix.

- Eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

- The *spectral gap* of G is $\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}$.
- If G is connected, and not bipartite $\sigma < 1$.

The (normalized) Laplacian

- Normalized Laplacian:

$$\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}$$

$A =$ adjacency matrix, $D =$ diagonal degree matrix.

- Eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

- The *spectral gap* of G is $\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}$.
- If G is connected, and not bipartite $\sigma < 1$.

The (normalized) Laplacian

- Normalized Laplacian:

$$\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}$$

  $A =$ adjacency matrix, $D =$ diagonal degree matrix.
- Eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

- The *spectral gap* of G is $\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}$.
- If G is connected, and not bipartite $\sigma < 1$.

The (normalized) Laplacian

- Normalized Laplacian:

$$\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}$$

  $A =$ adjacency matrix, $D =$ diagonal degree matrix.
- Eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

- The *spectral gap* of G is $\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}$.
- If G is connected, and not bipartite $\sigma < 1$.

The Spectral Gap

If $\sigma < 1$, G has many nice properties:

- Discrepancy properties:

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma\sqrt{\mathrm{vol}X\mathrm{vol}Y}$$

- Fast mixing of random walks:

$$\Delta(t) \leq e^{-t(1-\sigma)}\frac{\mathrm{vol}(G)}{d_{min}}$$

  $\Delta(t)$ is relative pointwise distance from stationary distribution.

- Expansion properties:

$$\frac{\mathrm{vol}(\Gamma(S))}{\mathrm{vol}(S)} \geq (1-\sigma)(1+\sigma)\frac{\mathrm{vol}(\overline{S})}{\mathrm{vol}(G)}.$$

The Spectral Gap

If $\sigma < 1$, G has many nice properties:

- Discrepancy properties:

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma\sqrt{\mathrm{vol}X\mathrm{vol}Y}$$

- Fast mixing of random walks:

$$\Delta(t) \leq e^{-t(1-\sigma)}\frac{\mathrm{vol}(G)}{d_{min}}$$

  $\Delta(t)$ is relative pointwise distance from stationary distribution.

- Expansion properties:

$$\frac{\mathrm{vol}(\Gamma(S))}{\mathrm{vol}(S)} \geq (1-\sigma)(1+\sigma)\frac{\mathrm{vol}(\overline{S})}{\mathrm{vol}(G)}.$$

The Spectral Gap

If $\sigma < 1$, G has many nice properties:

- Discrepancy properties:

$$\left| e(X, Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma \sqrt{\mathrm{vol} X \mathrm{vol}\, Y}$$

- Fast mixing of random walks:

$$\Delta(t) \leq e^{-t(1-\sigma)} \frac{\mathrm{vol}(G)}{d_{min}}$$

  $\Delta(t)$ is relative pointwise distance from stationary distribution.

- Expansion properties:

$$\frac{\mathrm{vol}(\Gamma(S))}{\mathrm{vol}(S)} \geq (1-\sigma)(1+\sigma)\frac{\mathrm{vol}(\overline{S})}{\mathrm{vol}(G)}.$$

Recall: For $G = K_n$

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$.

Call $\frac{1}{n}$ the percolation threshold or giant component threshold. Denote by p_c.

Recall: For $G = K_n$

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$.

Call $\frac{1}{n}$ the percolation threshold or giant component threshold. Denote by p_c.

Recall: For $G = K_n$

Erdős and Rényi (1959): $G \in \mathcal{G}_{n,p}$ satisfies

- $p < \frac{1-\epsilon}{n}$: All components size $O(\log n)$ a.a.s.
- $p > \frac{1+\epsilon}{n}$: Unique component of size $\Theta(n)$.

Call $\frac{1}{n}$ the percolation threshold or giant component threshold. Denote by p_c.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Special Host Graphs

- $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$; Kesten (1980)
- $p_c = \frac{1}{n}$ for hypercube Q_n; Ajtai, Komlós, Szemerédi (1982) and Bollobás, Kohayakawa and Łuczak (1992)
- $p_c = \frac{1}{d}$ for (n, d, λ)-graphs. (d-regular, $\sigma = o(1)$); Frieze, Krivelevich and Martin (2004)
- Percolation on expanders of bounded degree; Alon, Benjamini, Stacey (2004)
- $p_c = \frac{1}{\rho}$, ρ largest eigenvalue of A for very dense graphs (average degree $= \Theta(n)$); Bollobás, et. al, (2008).

Real world graphs: not regular, and not dense.

Results

- Recall:

$$\tilde{d} = \frac{\sum d_i^2}{\sum d_i} \qquad \mathrm{vol}_k(G) = \sum d_i^k.$$

Theorem (Chung, H., Lu)

For a general graph G, $p_c = \frac{1}{\tilde{d}}$ is the percolation threshold so long as G satisfies

(i) $\Delta = o(\frac{\tilde{d}}{\sigma})$, $\Delta = o(\frac{d\sqrt{n}}{\log n})$

(ii) $\sigma = o(n^{-\kappa})$ *for some* $\kappa > 0$

(iii) $vol_3(G) = O(d vol_2(G))$

where Δ is the maximum degree.

Notes:

- $\Delta = o(\frac{\tilde{d}}{\sigma}) \quad \Rightarrow \quad \rho = (1 + o(1))\tilde{d}$ where $\rho = ||A||$.
- Thus fits the pattern observed in other cases.
- Parts can be strengthened in several ways, e.g.

Theorem (Chung, H., Lu)

For a general graph G, if $p < \frac{1-\epsilon}{\tilde{d}}$ then there is a.a.s. no giant component in a percolated subgraph so long as G satisfies

(i) $\Delta = o(\frac{\tilde{d}}{\sigma})$.

In this case the largest component has volume at most $O(\sqrt{vol_2(G)}g(n))$ where $g(n) \to \infty$ arbitrarily slowly.

Notes:

- $\Delta = o(\frac{\tilde{d}}{\sigma}) \quad \Rightarrow \quad \rho = (1 + o(1))\tilde{d}$ where $\rho = ||A||$.
- Thus fits the pattern observed in other cases.
- Parts can be strengthened in several ways, e.g.

Theorem (Chung, H., Lu)

For a general graph G, if $p < \frac{1-\epsilon}{\tilde{d}}$ then there is a.a.s. no giant component in a percolated subgraph so long as G satisfies

(i) $\Delta = o(\frac{\tilde{d}}{\sigma})$.

In this case the largest component has volume at most $O(\sqrt{vol_2(G)}g(n))$ where $g(n) \to \infty$ arbitrarily slowly.

Notes:

- $\Delta = o(\frac{\tilde{d}}{\sigma}) \quad \Rightarrow \quad \rho = (1 + o(1))\tilde{d}$ where $\rho = ||A||$.
- Thus fits the pattern observed in other cases.
- Parts can be strengthened in several ways, e.g.

Theorem (Chung, H., Lu)

For a general graph G, if $p < \frac{1-\epsilon}{\tilde{d}}$ then there is a.a.s. no giant component in a percolated subgraph so long as G satisfies

(i) $\Delta = o(\frac{\tilde{d}}{\sigma})$.

In this case the largest component has volume at most $O(\sqrt{vol_2(G)}g(n))$ where $g(n) \to \infty$ arbitrarily slowly.

Through a lens

- What happens when $p = (1 + o(1))\frac{1}{n}$ for $\mathcal{G}_{n,p}$? This is the 'double jump'
- Actual window $= \Theta(n^{-4/3})$; Bollobás (1984), Łuczak (1990)
- Window $= \Theta(d^{-1}n^{-1/3})$ for certain regular graphs; Borgs, et. al. (2005,2006)
- Open: the window in our case.

Through a lens

- What happens when $p = (1 + o(1))\frac{1}{n}$ for $\mathcal{G}_{n,p}$? This is the 'double jump'
- Actual window $= \Theta(n^{-4/3})$; Bollobás (1984), Łuczak (1990)
- Window $= \Theta(d^{-1}n^{-1/3})$ for certain regular graphs; Borgs, et. al. (2005,2006)
- Open: the window in our case.

Through a lens

- What happens when $p = (1 + o(1))\frac{1}{n}$ for $\mathcal{G}_{n,p}$? This is the 'double jump'
- Actual window $= \Theta(n^{-4/3})$; Bollobás (1984), Łuczak (1990)
- Window $= \Theta(d^{-1}n^{-1/3})$ for certain regular graphs; Borgs, et. al. (2005,2006)
- Open: the window in our case.

Through a lens

- What happens when $p = (1 + o(1))\frac{1}{n}$ for $\mathcal{G}_{n,p}$? This is the 'double jump'
- Actual window $= \Theta(n^{-4/3})$; Bollobás (1984), Łuczak (1990)
- Window $= \Theta(d^{-1}n^{-1/3})$ for certain regular graphs; Borgs, et. al. (2005,2006)
- Open: the window in our case.

Traditional Methods

- Original method of Erdős-Renyi for $p > \frac{1+\epsilon}{n}$: counting components.
- Show that expected # of components of size $c \log n \leq |X| \leq \epsilon n \to 0$
- Show that the number of vertices in components of size $c \log n$ is $< n$.
- $\Rightarrow$ GCC!

How to count such components on a general, irregular graph?

Traditional Methods

- Original method of Erdős-Renyi for $p > \frac{1+\epsilon}{n}$: counting components.
- Show that expected # of components of size $c \log n \leq |X| \leq \epsilon n \to 0$
- Show that the number of vertices in components of size $c \log n$ is $< n$.
- $\Rightarrow$ GCC!

How to count such components on a general, irregular graph?

Traditional Methods

- Original method of Erdős-Renyi for $p > \frac{1+\epsilon}{n}$: counting components.
- Show that expected # of components of size $c \log n \leq |X| \leq \epsilon n \to 0$
- Show that the number of vertices in components of size $c \log n$ is $< n$.
- $\Rightarrow$ GCC!

How to count such components on a general, irregular graph?

Traditional Methods

- Original method of Erdős-Renyi for $p > \frac{1+\epsilon}{n}$: counting components.
- Show that expected # of components of size $c \log n \leq |X| \leq \epsilon n \to 0$
- Show that the number of vertices in components of size $c \log n$ is $< n$.
- $\Rightarrow$ GCC!

How to count such components on a general, irregular graph?

Traditional Methods

- Original method of Erdős-Renyi for $p > \frac{1+\epsilon}{n}$: counting components.
- Show that expected # of components of size $c \log n \leq |X| \leq \epsilon n \to 0$
- Show that the number of vertices in components of size $c \log n$ is $< n$.
- $\Rightarrow$ GCC!

How to count such components on a general, irregular graph?

Another Method

- 'More modern' method: branching processes:
- Branching process on $n-1$-regular trees
  - If $p < \frac{1-\epsilon}{n-1}$ process dies out
  - If $p > \frac{1+\epsilon}{n-1}$ process survives forever with positive probability
- Tie this to existence giant component.

Irregular graphs: different vertices have different degrees. How do we show the growth now?

Another Method

- 'More modern' method: branching processes:
- Branching process on $n-1$-regular trees
  - If $p < \frac{1-\epsilon}{n-1}$ process dies out
  - If $p > \frac{1+\epsilon}{n-1}$ process survives forever with positive probability
- Tie this to existence giant component.

Irregular graphs: different vertices have different degrees. How do we show the growth now?

Another Method

- 'More modern' method: branching processes:
- Branching process on $n-1$-regular trees
  - If $p < \frac{1-\epsilon}{n-1}$ process dies out
  - If $p > \frac{1+\epsilon}{n-1}$ process survives forever with positive probability
- Tie this to existence giant component.

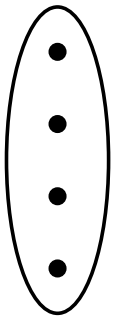
Irregular graphs: different vertices have different degrees. How do we show the growth now?

Another Method

- 'More modern' method: branching processes:
- Branching process on $n-1$-regular trees
  - If $p < \frac{1-\epsilon}{n-1}$ process dies out
  - If $p > \frac{1+\epsilon}{n-1}$ process survives forever with positive probability
- Tie this to existence giant component.

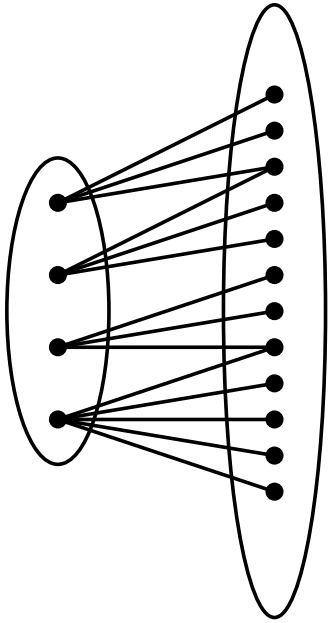
Irregular graphs: different vertices have different degrees. How do we show the growth now?

Growing a giant component



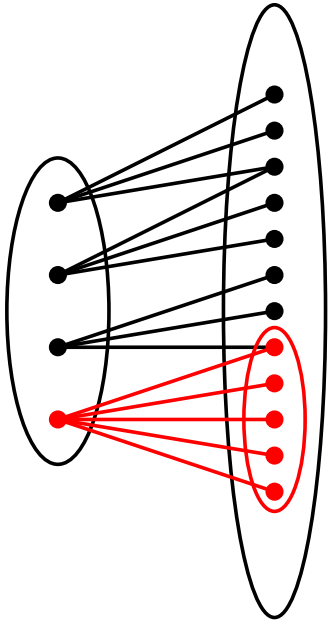
- Start with a (not necessarily connected) seed set and iteratively take neighborhoods
- Once we have a large enough component, grow from that
- Use spectral, probabilistic tools to make sure growth happens

Growing a giant component



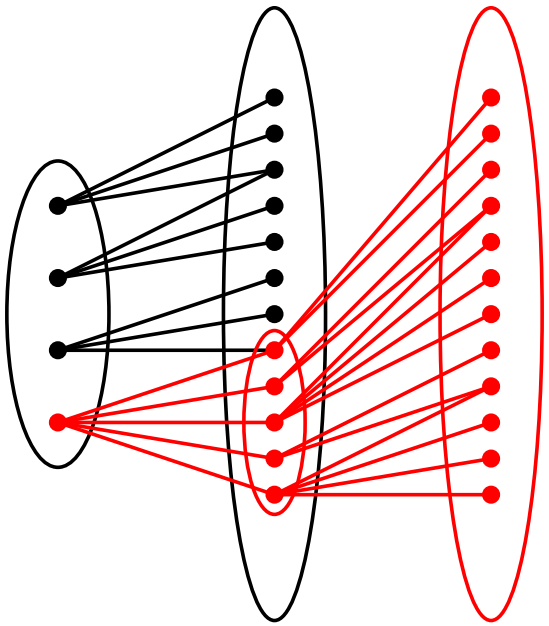
- Start with a (not necessarily connected) seed set and iteratively take neighborhoods
- Once we have a large enough component, grow from that
- Use spectral, probabilistic tools to make sure growth happens

Growing a giant component



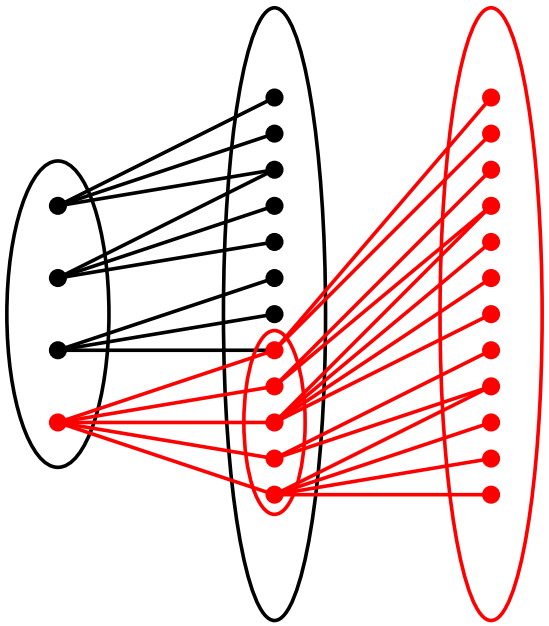
- Start with a (not necessarily connected) seed set and iteratively take neighborhoods
- Once we have a large enough component, grow from that
- Use spectral, probabilistic tools to make sure growth happens

Growing a giant component



- Start with a (not necessarily connected) seed set and iteratively take neighborhoods
- Once we have a large enough component, grow from that
- Use spectral, probabilistic tools to make sure growth happens

Growing a giant component



- Start with a (not necessarily connected) seed set and iteratively take neighborhoods
- Once we have a large enough component, grow from that
- Use spectral, probabilistic tools to make sure growth happens

Tools

- Recall the classical spectral bound:

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma\sqrt{\mathrm{vol}X\,\mathrm{vol}Y}$$

- Says many edges must leave a set with high volume.
- Not enough: need to know the volume of neighborhoods!

Tools

- Recall the classical spectral bound:

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma\sqrt{\mathrm{vol}X\mathrm{vol}Y}$$

- Says many edges must leave a set with high volume.
- Not enough: need to know the volume of neighborhoods!

Tools

- Recall the classical spectral bound:

$$\left| e(X,Y) - \frac{\mathrm{vol}(X)\mathrm{vol}(Y)}{\mathrm{vol}(G)} \right| \leq \sigma\sqrt{\mathrm{vol}X\,\mathrm{vol}Y}$$

- Says many edges must leave a set with high volume.
- Not enough: need to know the volume of neighborhoods!

Tools

Stronger spectral results are necessary

Lemma

For two disjoint sets S and T, we have

$$\left| \sum_{v \in T} d_v |\Gamma(v) \cap S| - \frac{vol(S) vol_2(T)}{vol(G)} \right| \leq \sigma \sqrt{vol(S) vol_3(T)}.$$

$$\left| \sum_{v \in T} d_v |\Gamma(v) \cap S|^2 - \frac{vol(S)^2 vol_3(T)}{vol(G)^2} \right|$$

$$\leq \quad \sigma^2 vol(S) \max_{v \in T} \{d_v^2\} + 2\sigma \frac{\sqrt{vol(S)^3 vol_5(T)}}{vol(G)}.$$

Wrapping Up

- Lemma implies: $\mathrm{vol}(S)$ in some range $\Rightarrow \mathbb{E}[\mathrm{vol}(\Gamma(S))] \geq (1+\epsilon)\mathrm{vol}(S)$
- Chernoff Bounds: $\mathrm{vol}(\Gamma(S)) \geq (1+\epsilon/10)\mathrm{vol}(S)$ with high probability
- Range allows us to grow until $\mathrm{vol}(S) = \epsilon\mathrm{vol}(G)$.
- Check: sum of error probabilities $\to 0$.

Wrapping Up

- Lemma implies: $\mathrm{vol}(S)$ in some range $\Rightarrow \mathbb{E}[\mathrm{vol}(\Gamma(S))] \geq (1+\epsilon)\mathrm{vol}(S)$
- Chernoff Bounds: $\mathrm{vol}(\Gamma(S)) \geq (1+\epsilon/10)\mathrm{vol}(S)$ with high probability
- Range allows us to grow until $\mathrm{vol}(S) = \epsilon\mathrm{vol}(G)$.
- Check: sum of error probabilities $\to 0$.

Wrapping Up

- Lemma implies: $\mathrm{vol}(S)$ in some range $\Rightarrow \mathbb{E}[\mathrm{vol}(\Gamma(S))] \geq (1+\epsilon)\mathrm{vol}(S)$
- Chernoff Bounds: $\mathrm{vol}(\Gamma(S)) \geq (1+\epsilon/10)\mathrm{vol}(S)$ with high probability
- Range allows us to grow until $\mathrm{vol}(S) = \epsilon\mathrm{vol}(G)$.
- Check: sum of error probabilities $\to 0$.

Wrapping Up

- Lemma implies: $\mathrm{vol}(S)$ in some range $\Rightarrow \mathbb{E}[\mathrm{vol}(\Gamma(S))] \geq (1+\epsilon)\mathrm{vol}(S)$
- Chernoff Bounds: $\mathrm{vol}(\Gamma(S)) \geq (1+\epsilon/10)\mathrm{vol}(S)$ with high probability
- Range allows us to grow until $\mathrm{vol}(S) = \epsilon\mathrm{vol}(G)$.
- Check: sum of error probabilities $\to 0$.

Sketch of Other Direction

- Suppose volume of largest component is x.
- Choose two vertices at random proportionally to their degrees.
- Get lower bound that they are in same component in terms of x.
- Upper bound: a walk between them must survive

$$\mathbb{P}(\ldots) \leq \sum_{k=0}^{n} \frac{1}{\mathrm{vol}(G)^2} p^k \mathbf{d}^* A^k \mathbf{d}.$$

- If x too large; contradiction!

Sketch of Other Direction

- Suppose volume of largest component is x.
- Choose two vertices at random proportionally to their degrees.
- Get lower bound that they are in same component in terms of x.
- Upper bound: a walk between them must survive

$$\mathbb{P}(\ldots) \leq \sum_{k=0}^{n} \frac{1}{\mathrm{vol}(G)^2} p^k \mathbf{d}^* A^k \mathbf{d}.$$

- If x too large; contradiction!

Sketch of Other Direction

- Suppose volume of largest component is x.
- Choose two vertices at random proportionally to their degrees.
- Get lower bound that they are in same component in terms of x.
- Upper bound: a walk between them must survive

$$\mathbb{P}(\ldots) \leq \sum_{k=0}^{n} \frac{1}{\mathrm{vol}(G)^2} p^k \mathbf{d}^* A^k \mathbf{d}.$$

- If x too large; contradiction!

Sketch of Other Direction

- Suppose volume of largest component is x.
- Choose two vertices at random proportionally to their degrees.
- Get lower bound that they are in same component in terms of x.
- Upper bound: a walk between them must survive

$$\mathbb{P}(\ldots) \leq \sum_{k=0}^{n} \frac{1}{\mathrm{vol}(G)^2} p^k \mathbf{d}^* A^k \mathbf{d}.$$

- If x too large; contradiction!

Sketch of Other Direction

- Suppose volume of largest component is x.
- Choose two vertices at random proportionally to their degrees.
- Get lower bound that they are in same component in terms of x.
- Upper bound: a walk between them must survive

$$\mathbb{P}(\ldots) \leq \sum_{k=0}^{n} \frac{1}{\mathrm{vol}(G)^2} p^k \mathbf{d}^* A^k \mathbf{d}.$$

- If x too large; contradiction!

Other Questions

- Weaken $\mathrm{vol}_2(G) = O(d\mathrm{vol}_3(G))$ condition for giant components.
  - Under what conditions is $\frac{1}{\rho}$ the threshold?
- Investigate scaling window about $p_c = \frac{1}{\tilde{d}}$.
- Structural properties of giant component in GCC?

Thank You!

Other Questions

- Weaken $\mathrm{vol}_2(G) = O(d\mathrm{vol}_3(G))$ condition for giant components.
  - Under what conditions is $\frac{1}{\rho}$ the threshold?
- Investigate scaling window about $p_c = \frac{1}{\tilde{d}}$.
- Structural properties of giant component in GCC?

Thank You!

Other Questions

- Weaken $\mathrm{vol}_2(G) = O(d\mathrm{vol}_3(G))$ condition for giant components.
  - Under what conditions is $\frac{1}{\rho}$ the threshold?
- Investigate scaling window about $p_c = \frac{1}{\tilde{d}}$.
- Structural properties of giant component in GCC?

Thank You!

Other Questions

- Weaken $\mathrm{vol}_2(G) = O(d\mathrm{vol}_3(G))$ condition for giant components.
  - Under what conditions is $\frac{1}{\rho}$ the threshold?
- Investigate scaling window about $p_c = \frac{1}{\tilde{d}}$.
- Structural properties of giant component in GCC?

Thank You!