How To Determine If A Random Graph With A Fixed Degree Sequence Has A Giant Component

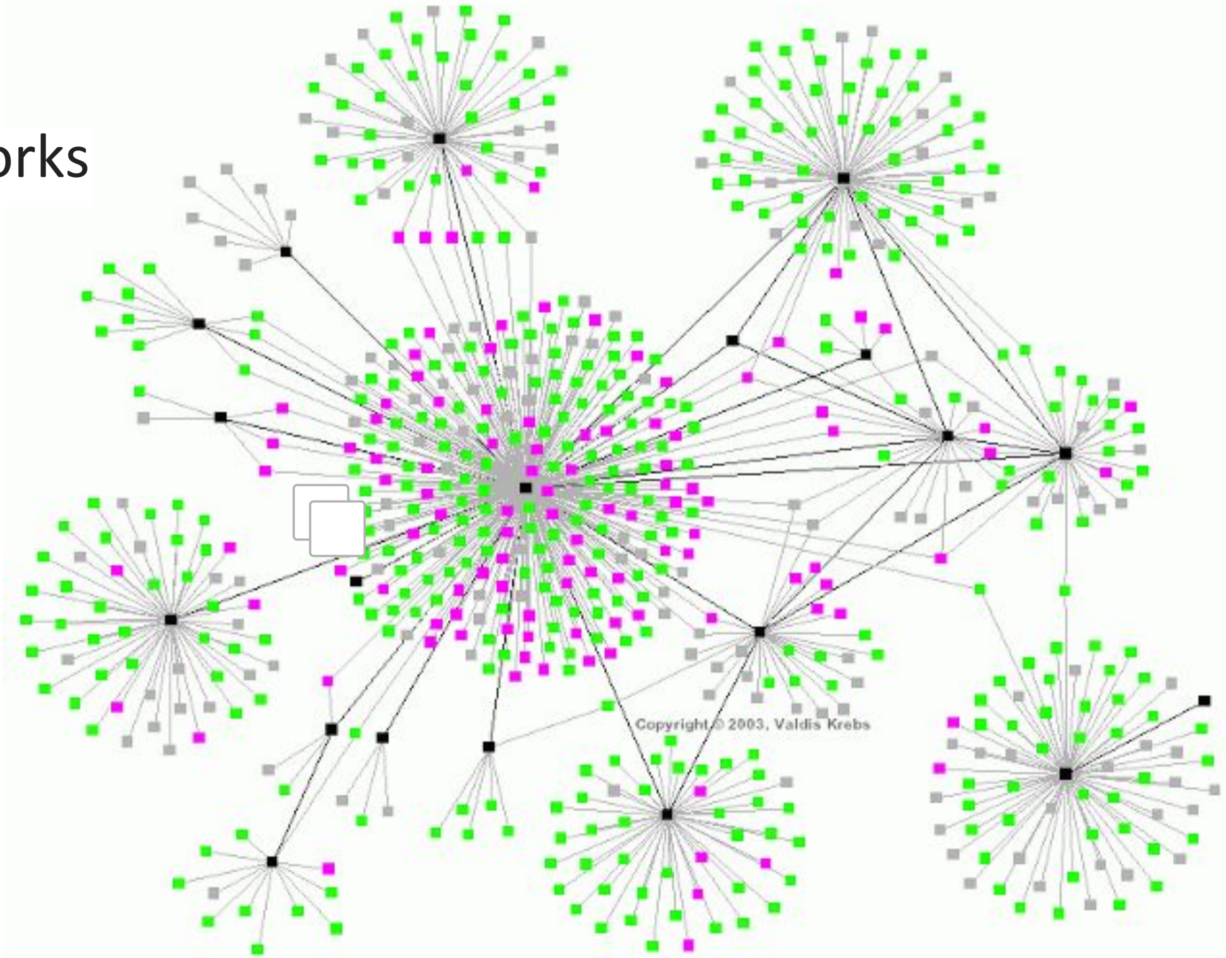
Bruce Reed

Nice
July 2019

Some Applications of Clustering in Real World Netwroks

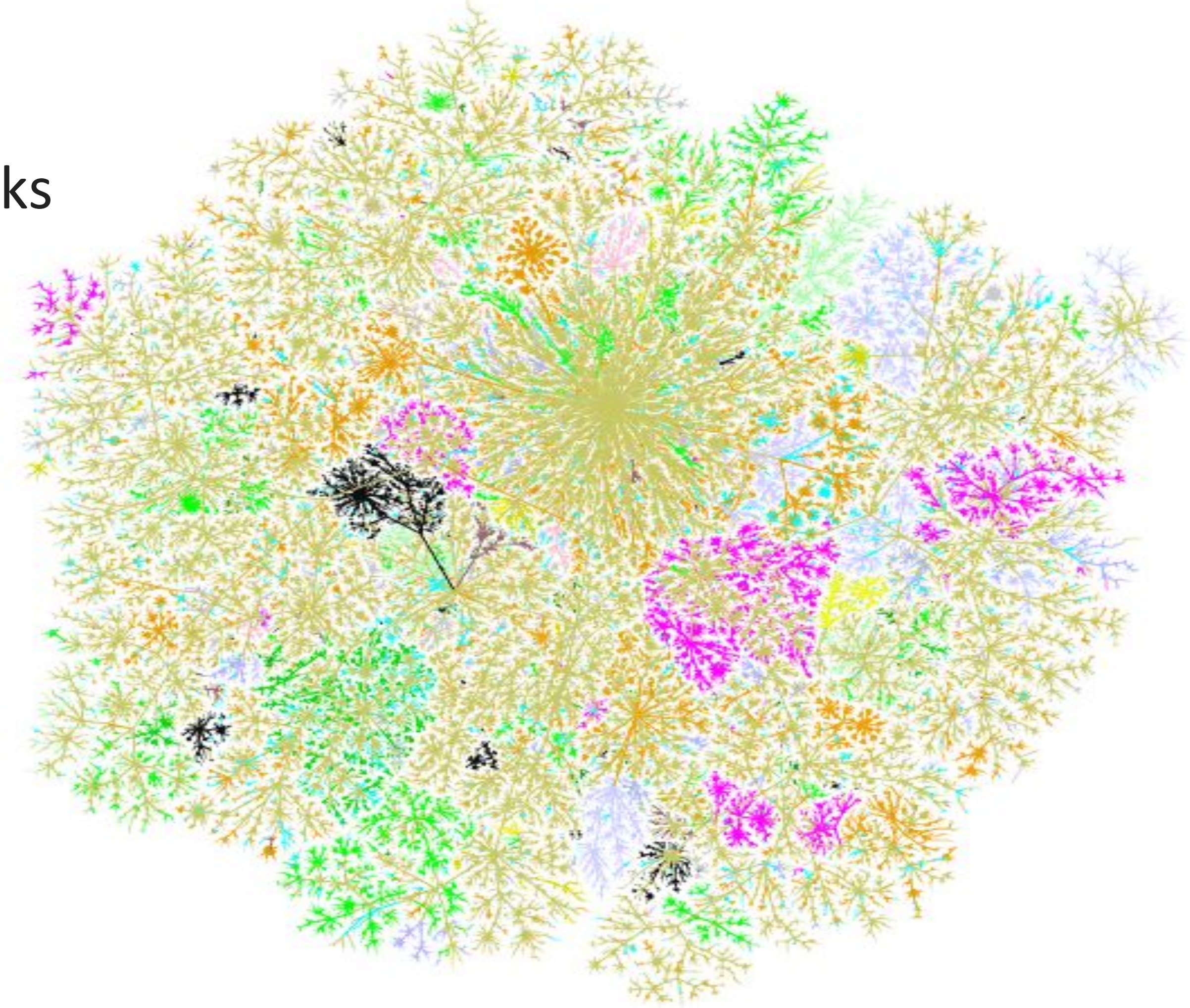
Looking for Clusters I: Epidemiological Networks

Transmission Network Analysis
to complement routine TB
contact analysis
McKenzie et al. AJPH 2007



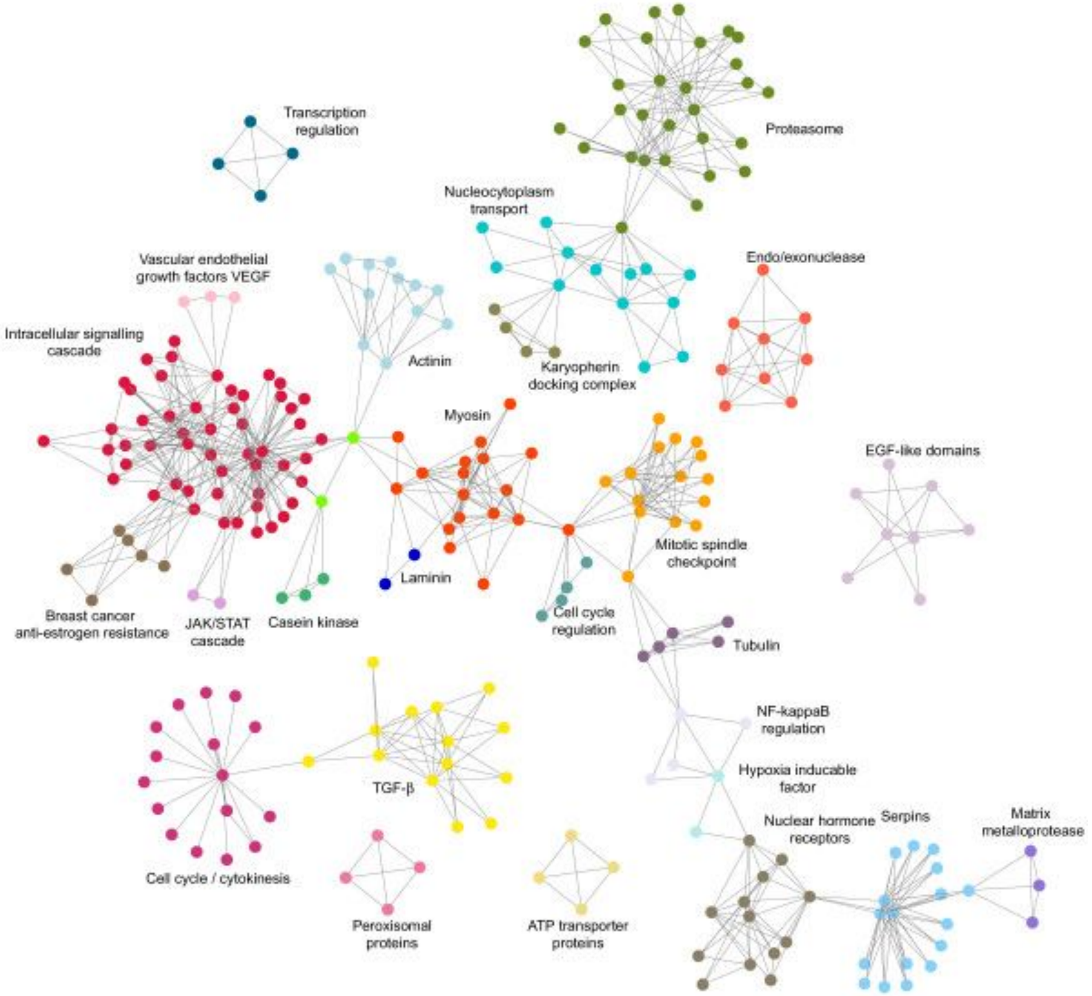
Looking for Clusters II: Communication Networks

Internet Mapping Project
Bell Laboratories
May 3 1999



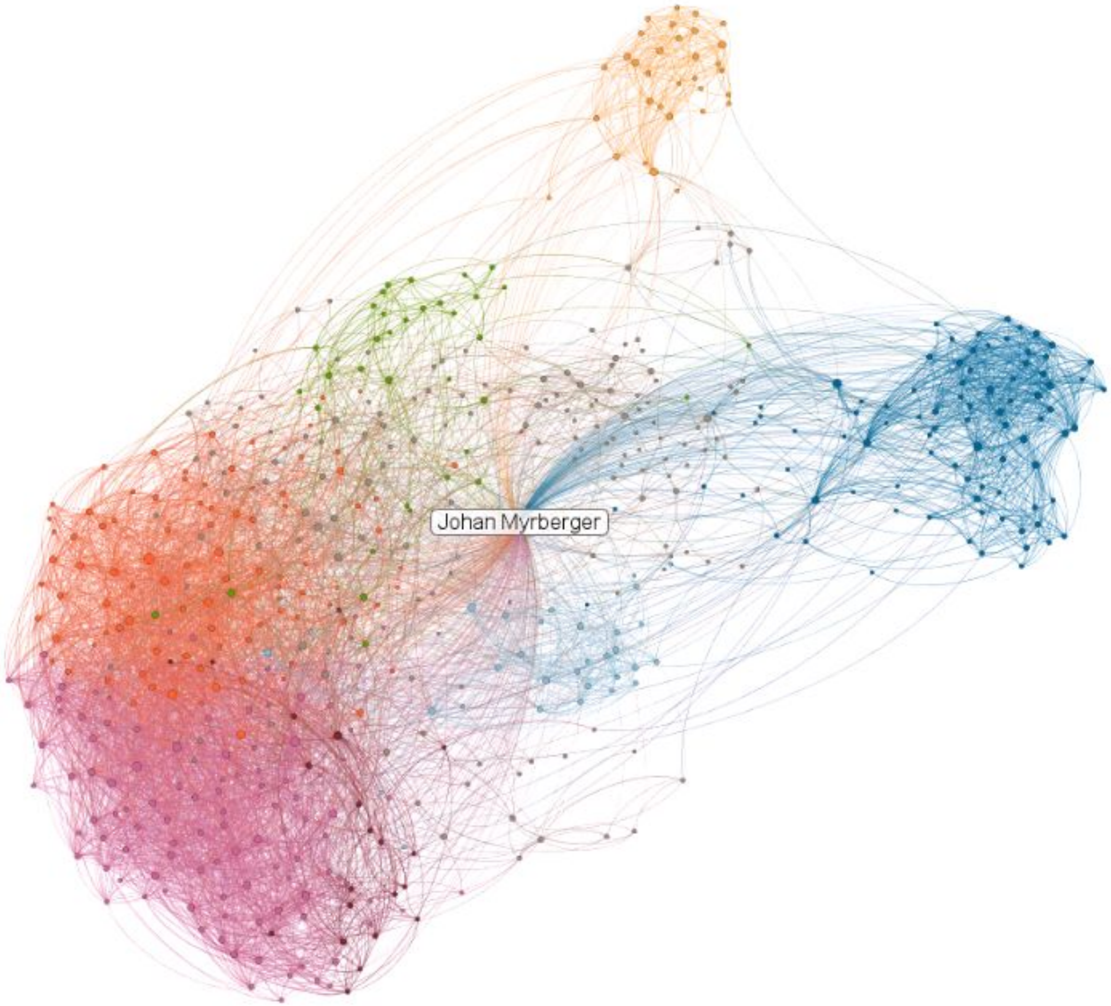
Looking for Clusters III: Biological Networks

Jonsson et al. BMC Bioinformatics 2006



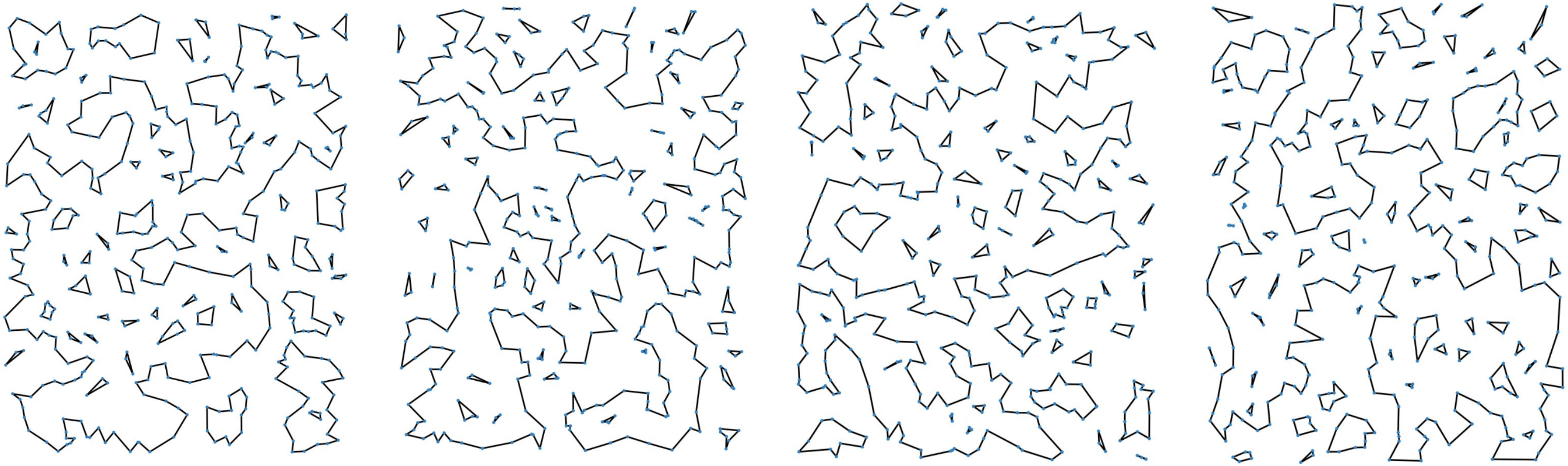
Looking for Clusters IV: Social Networks

Linked in Network of Johann Myrberger

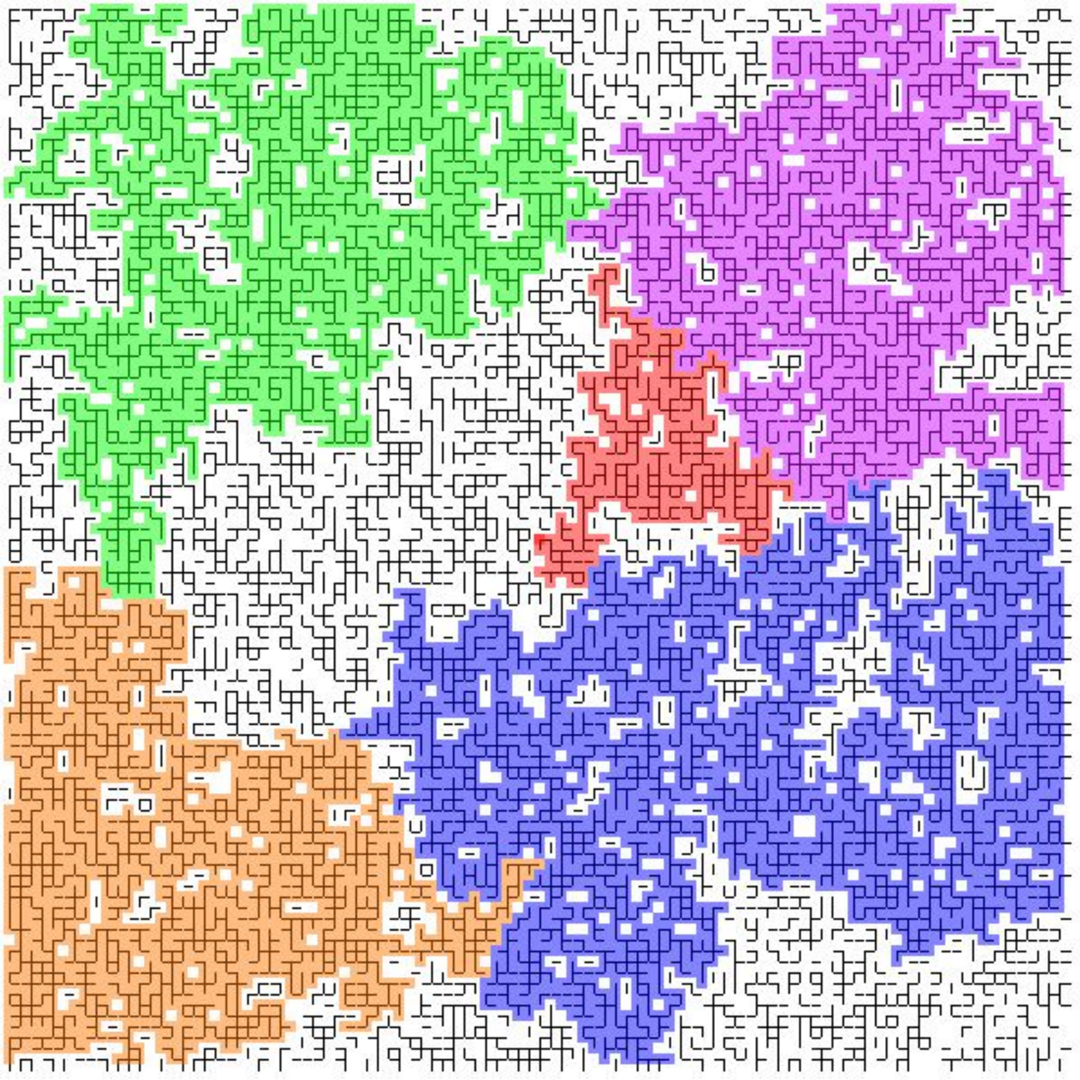


Looking for Clusters V: Euclidean 2-factors

Is there a c>0 s.t. the minimum cost 2-Factor for n uniformly chosen points almost surely contains a component with cn points, Bill Cook, Private Communication 2014



Looking for Clusters VI: Percolation



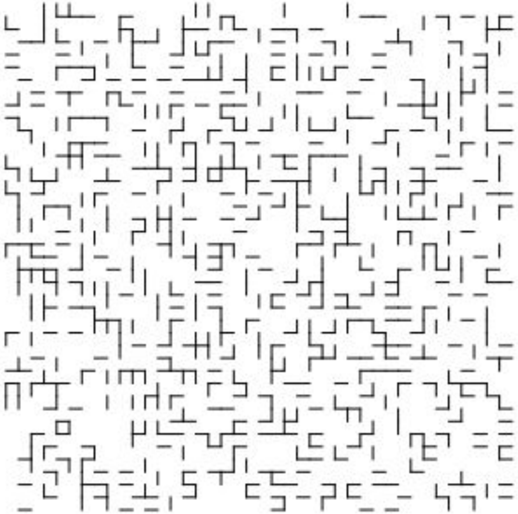
Random Networks as Controls

A common technique to analyze the properties of a single network is to use statistical randomization methods to create a reference network which is used for comparison purposes.

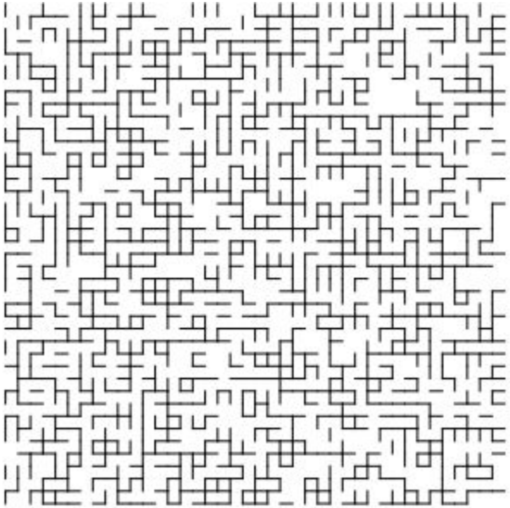
Mondragon and Zhou, 2012.

Factors Determining
How Much Clustering Occurs

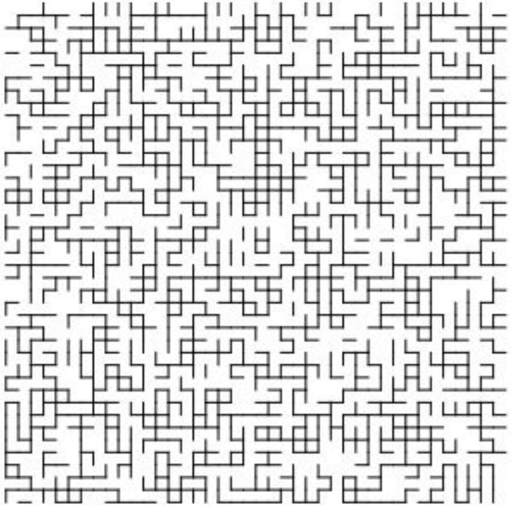
More Edges Means More Clustering



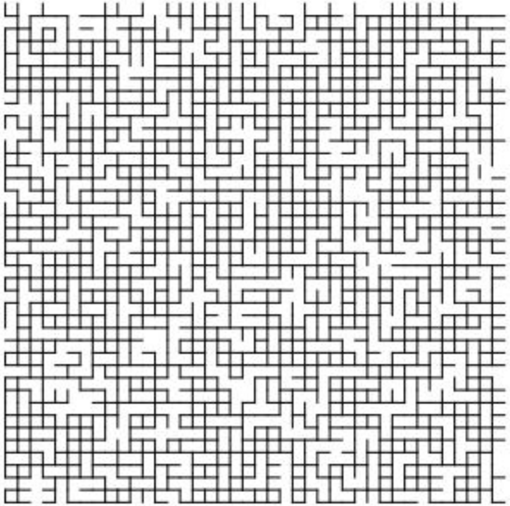
p=0.25



p=0.48

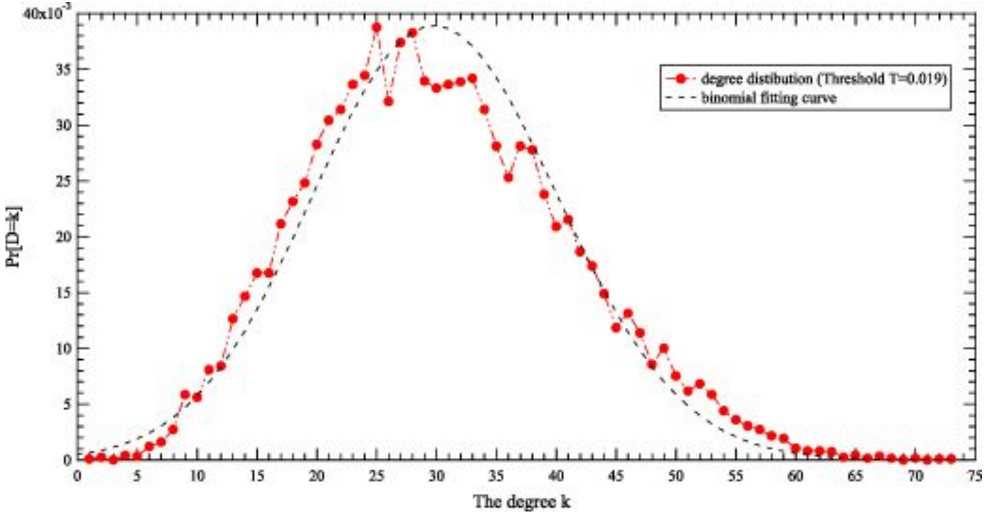


p=0.52



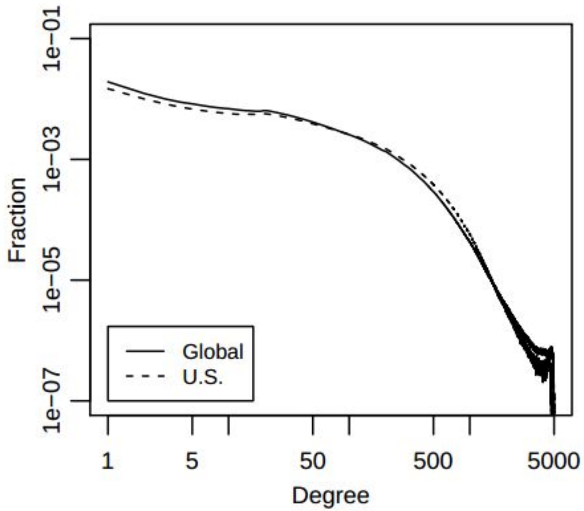
p=0.75

Degree Distributions Differ

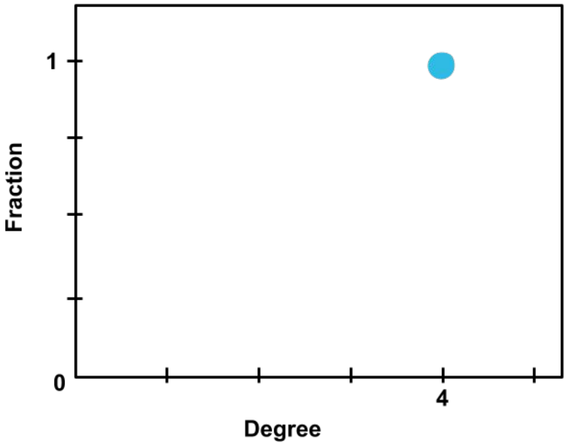


Classic Erdős-Renyi Model

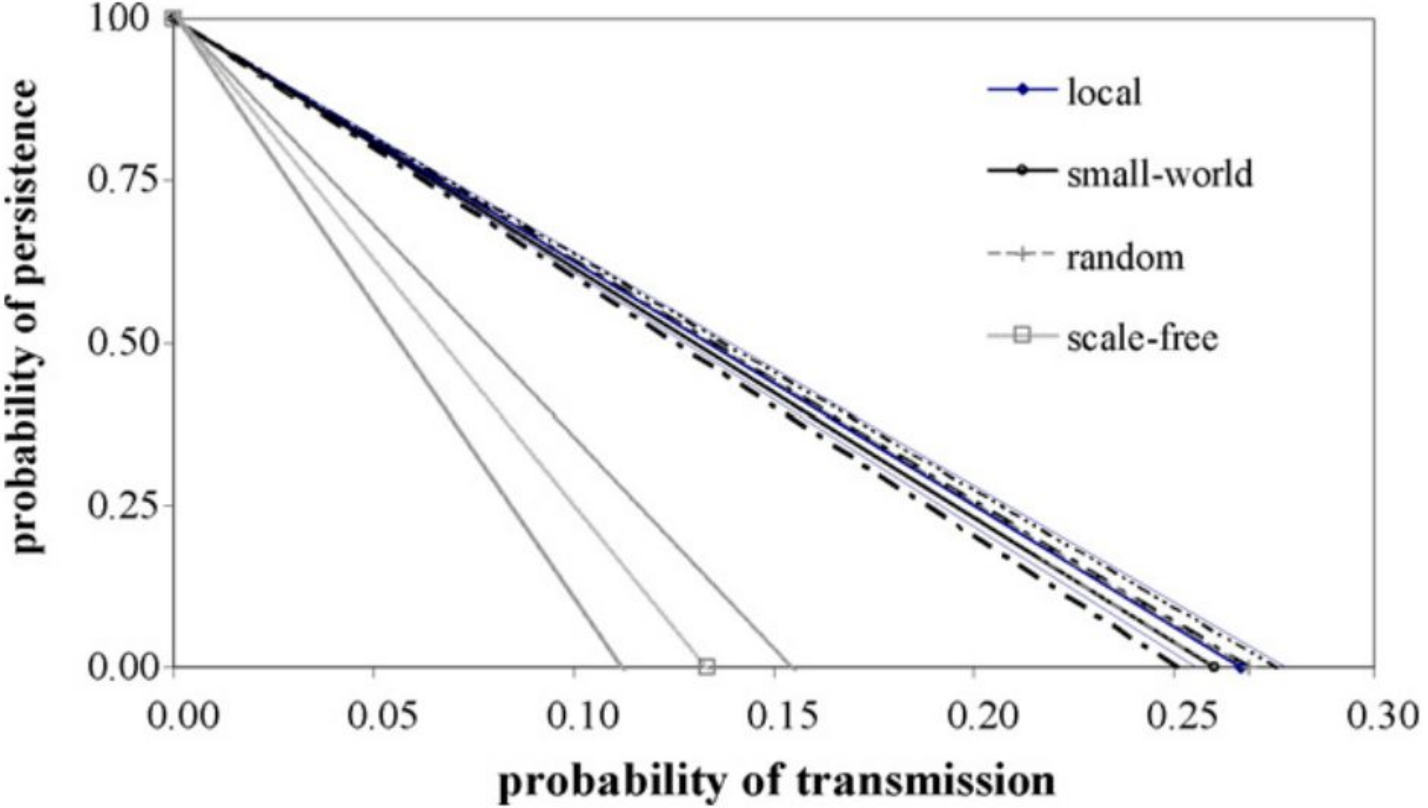
Facebook Friends



Lattice



Network Structure Affects Cluster Size



Our Focus:
Giant Components

Does a uniformly chosen graph on a given degree sequence have a giant component?

Does a uniformly chosen graph on a given degree sequence have a giant component?

For a sequence D of nonzero degrees, G(D) is a uniformly chosen graph with degree sequence D.

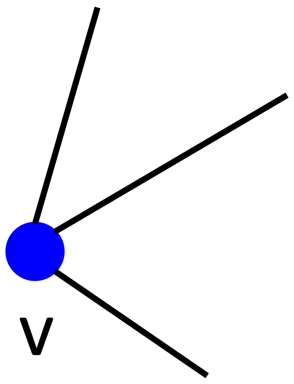
Does a uniformly chosen graph on a given degree sequence have a giant component?

For a sequence D of nonzero degrees, G(D) is a uniformly chosen graph with degree sequence D.

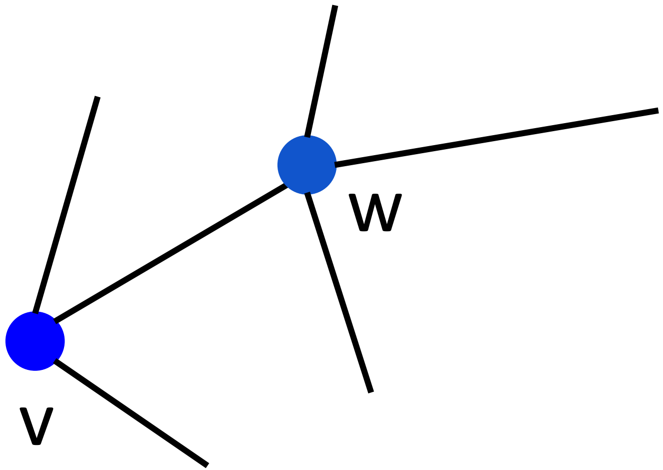
Will assume D is non-decreasing and all degrees are positive.

The First Answer

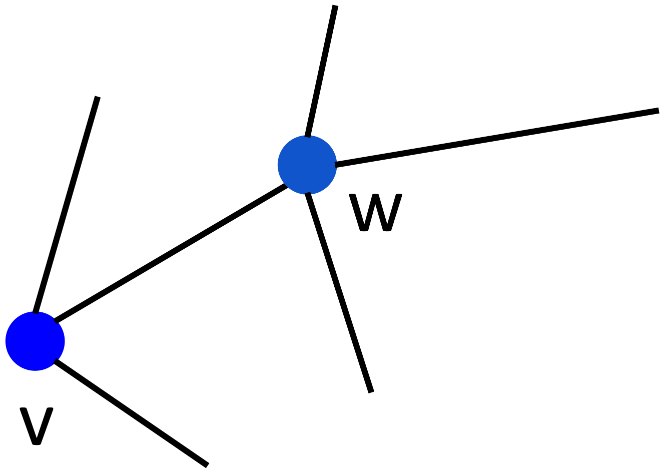
A Heuristic Argument



A Heuristic Argument



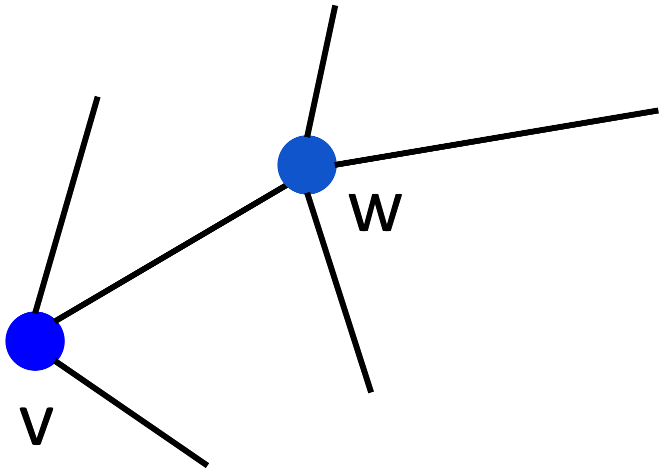
A Heuristic Argument



Change in number of open edges:

$d(w) - 2$

A Heuristic Argument



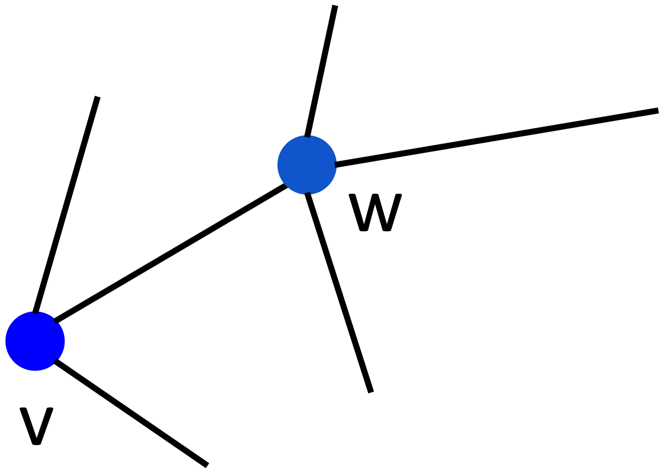
Change in number of open edges:

$$d(w) - 2$$

Probability pick w:

$$d(w) / \sum_{u} d(u)$$

A Heuristic Argument



Change in number of open edges:

$$d(w) - 2$$

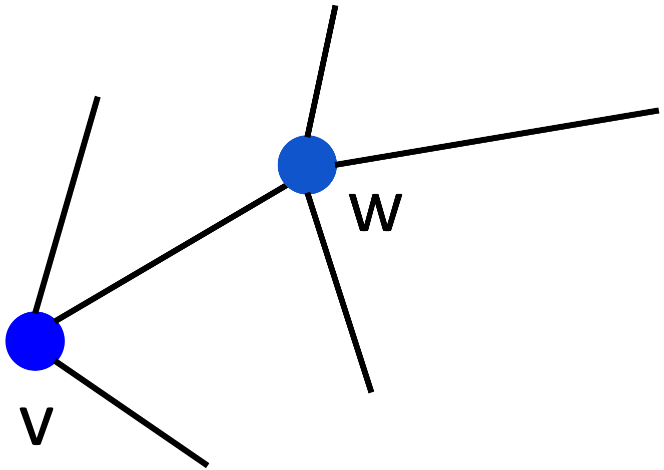
Probability pick w:

$$d(w) / \sum_u d(u)$$

Expected change:

$$\sum_u d(u)(d(u) - 2) / \sum_u d(u)$$

A Heuristic Argument



Giant Component if and only if

$\sum_u d(u)(d(u)-2)$ is positive??

Change in number of open edges:

$$d(w) - 2$$

Probability pick w:

$$d(w) / \sum_u d(u)$$

Expected change:

$$\sum_u d(u)(d(u) - 2) / \sum_u d(u)$$

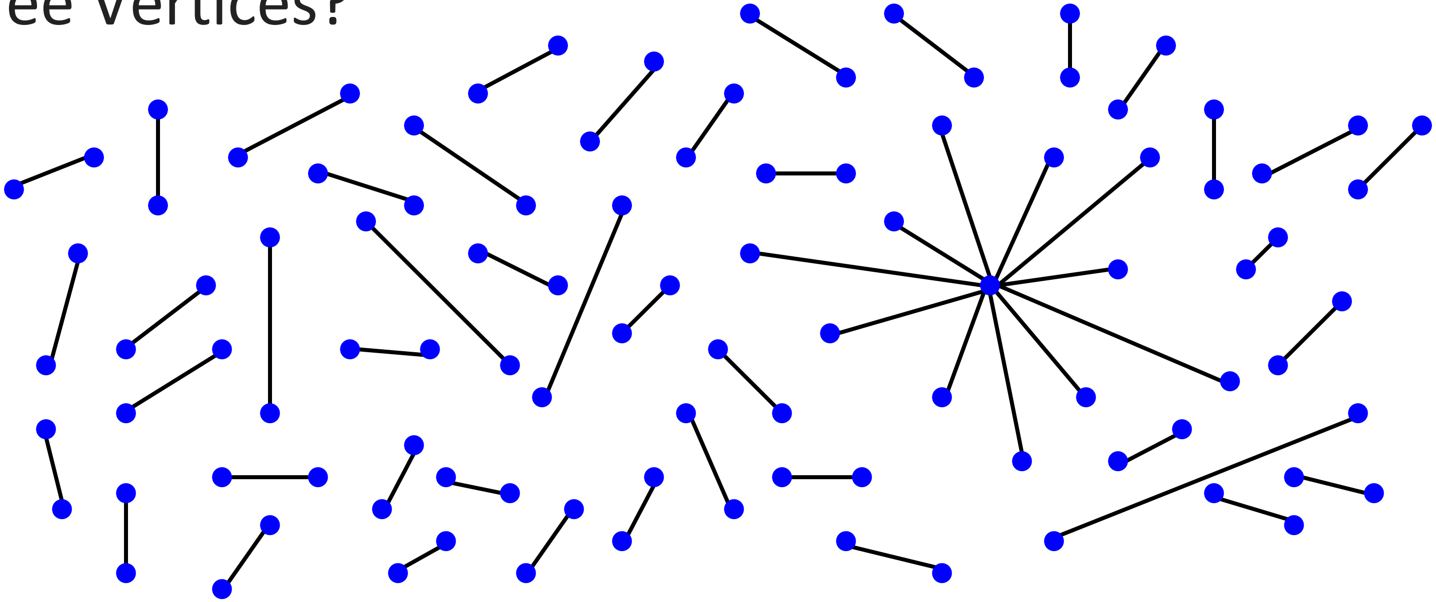
Molloy-Reed(1995) Result

Under considerable technical conditions including maximum degree at most $n^{1/8}$:

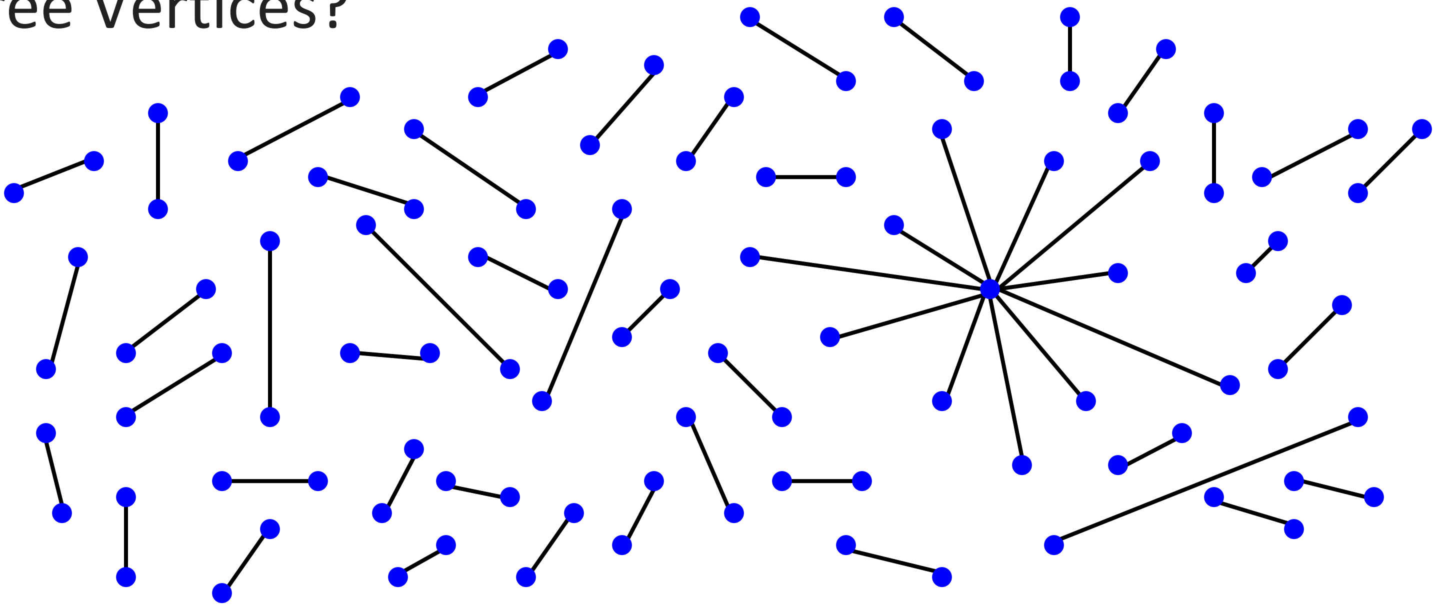
$$\sum_u d(u)(d(u) - 2) > \varepsilon n$$ implies a giant component exists.

$$\sum_u d(u)(d(u) - 2) < -\varepsilon n$$ implies no giant component exists.

Why Can't We Prove The Result For Graphs With High Degree Vertices?

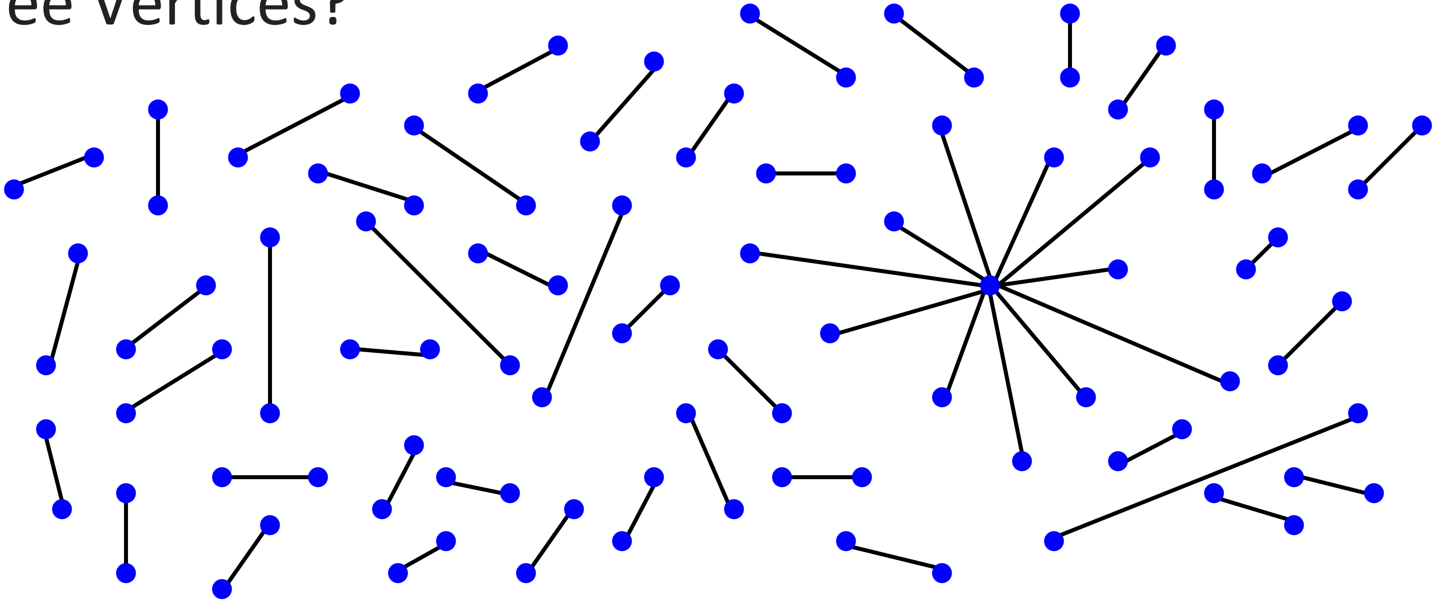


Why Can't We Prove The Result For Graphs With High Degree Vertices?



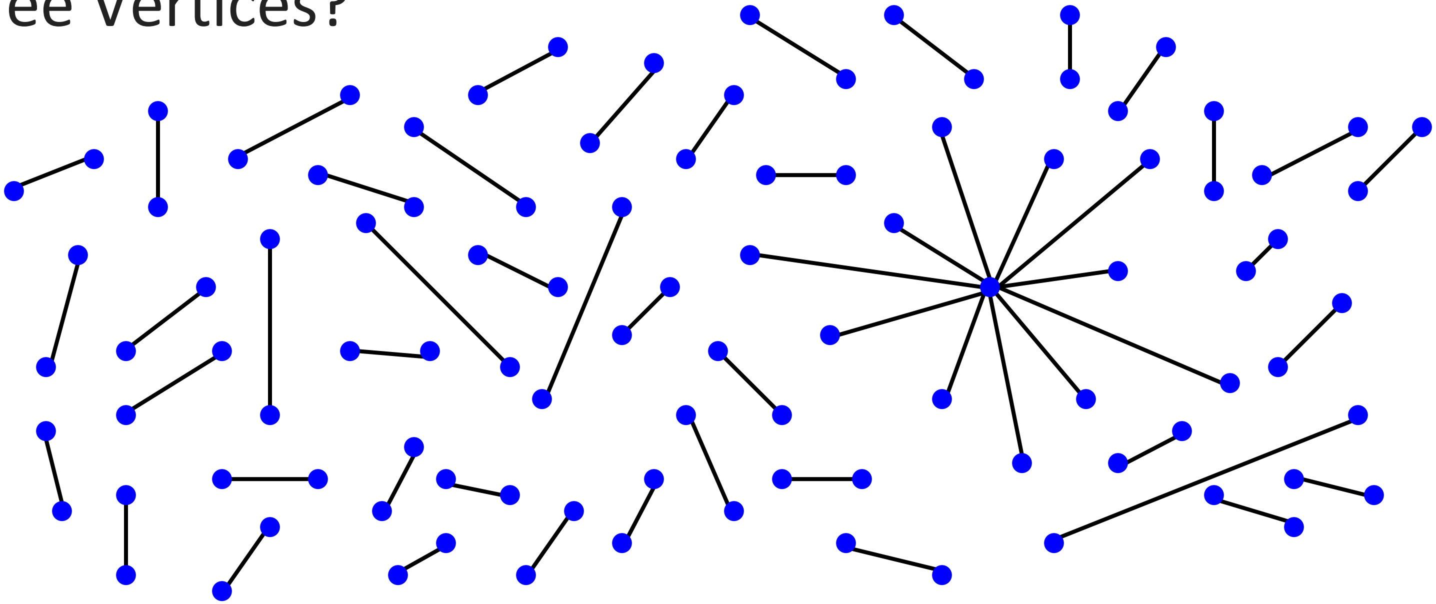
Because it is false.

Why Can't We Prove The Result For Graphs With High Degree Vertices?



Cannot translate results from the non-simple case.

Why Can't We Prove The Result For Graphs With High Degree Vertices?



Cannot translate results from the non-simple case.

Hard to prove concentration results.

A Fuller Answer

OUR QUESTION REVISITED

Does a uniformly chosen graph on a given degree sequence have a giant component?

For a sequence D of nonzero degrees, G(D) is a uniformly chosen graph with degree sequence D.

Will assume D is non-decreasing and all degrees are positive.

Four Definitions

M is the sum of the degrees in D which are not 2.

D is f-well behaved if M is at least $f(n)$.

$$j_D = \min\left(i \text{ s.t. } \sum_{j=1}^{i} d_j(d_j - 2) > 0, n\right)$$

$$R_D = \sum_{j_D}^{n} d_j$$

One Crucial Observation

$\sum_{j=1}^{n}$ d(u)(d(u) − 2) is at least R_D

One Crucial Observation

$$\sum_{j=1}^{n} d(u)(d(u)-2) \text{ is at least } R_D$$

and for some $\gamma > 0$ remains above $R_D/2$ until the sum of the degrees of the vertices explored is at least γR_D.

One Crucial Observation

$$\sum_{j=1}^{n} d(u)(d(u)-2) \text{ is at least } R_D$$

and for some $\gamma > 0$ remains above $R_D/2$ until the sum of the degrees of the vertices explored is at least γR_D.

But goes negative once all the vertices with index $> j_D$ are explored.

Two Theorems

Theorem 1: For any $f \to \infty$ and $b \to 0$, if a well behaved degree distribution D satisfies $R_D \leq b(n)M$ then G(D) has no giant component

.

Two Theorems

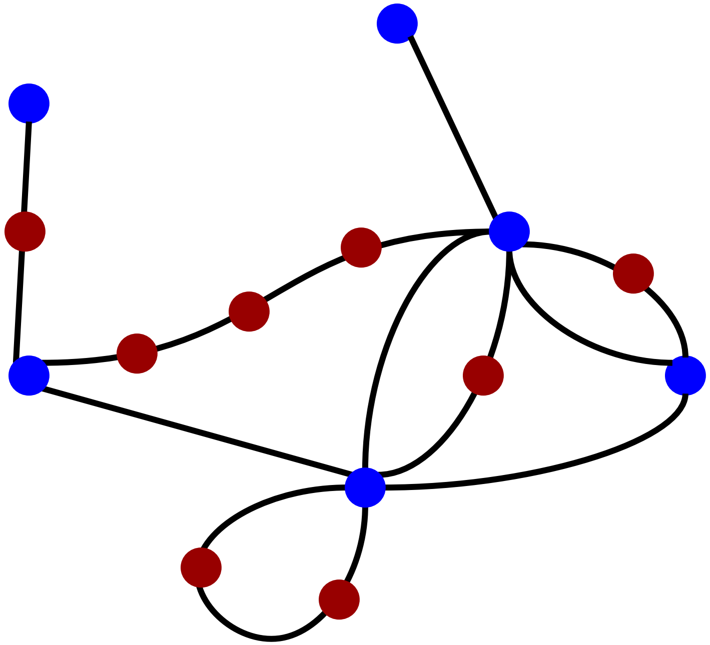
Theorem 1: For any $f \to \infty$ and $b \to 0$, if a well behaved degree distribution D satisfies $R_D \leq b(n)M$ then G(D) has no giant component.

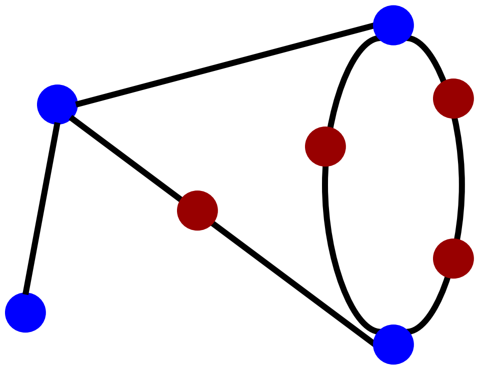
Theorem 2: For any $f \to \infty$ and $\varepsilon > 0$ if a well behaved degree distribution D satisfies $R_D \geq \varepsilon M$ then G(D) has a giant component

(Joos, Perarnau-Llobet, Rautenbach, Reed 2015)

Why we focus on M and not n

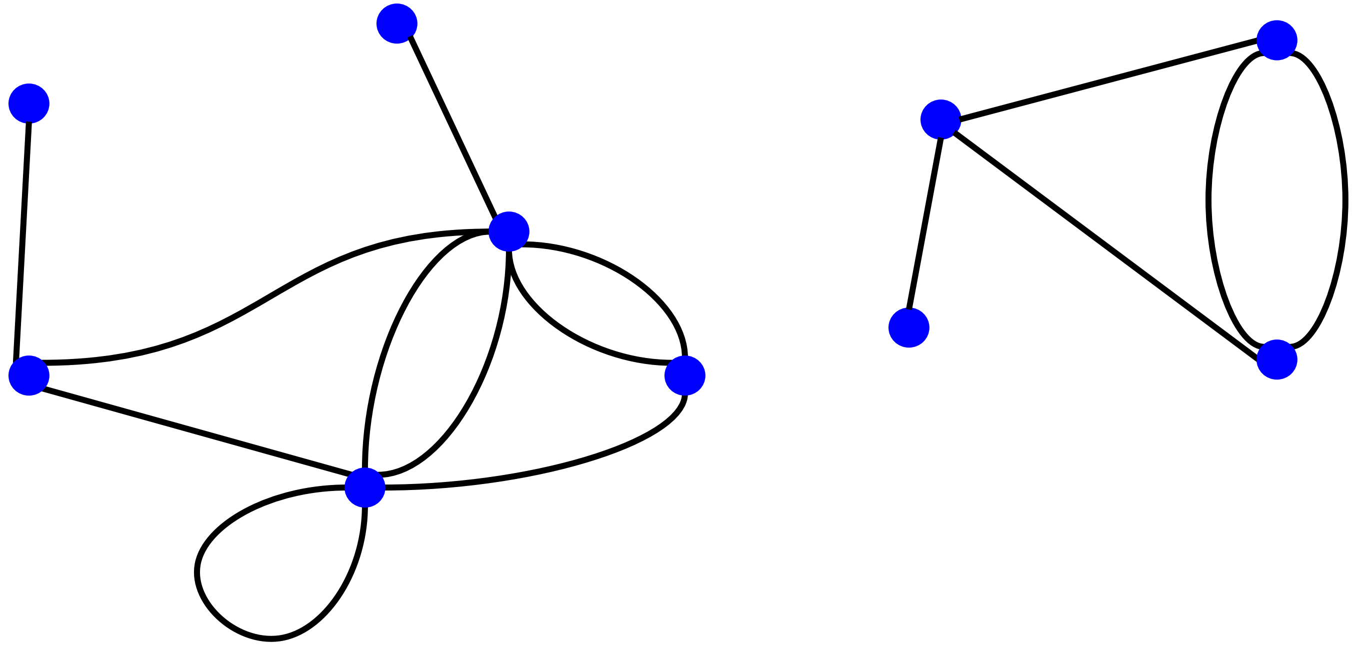
And edges not vertices





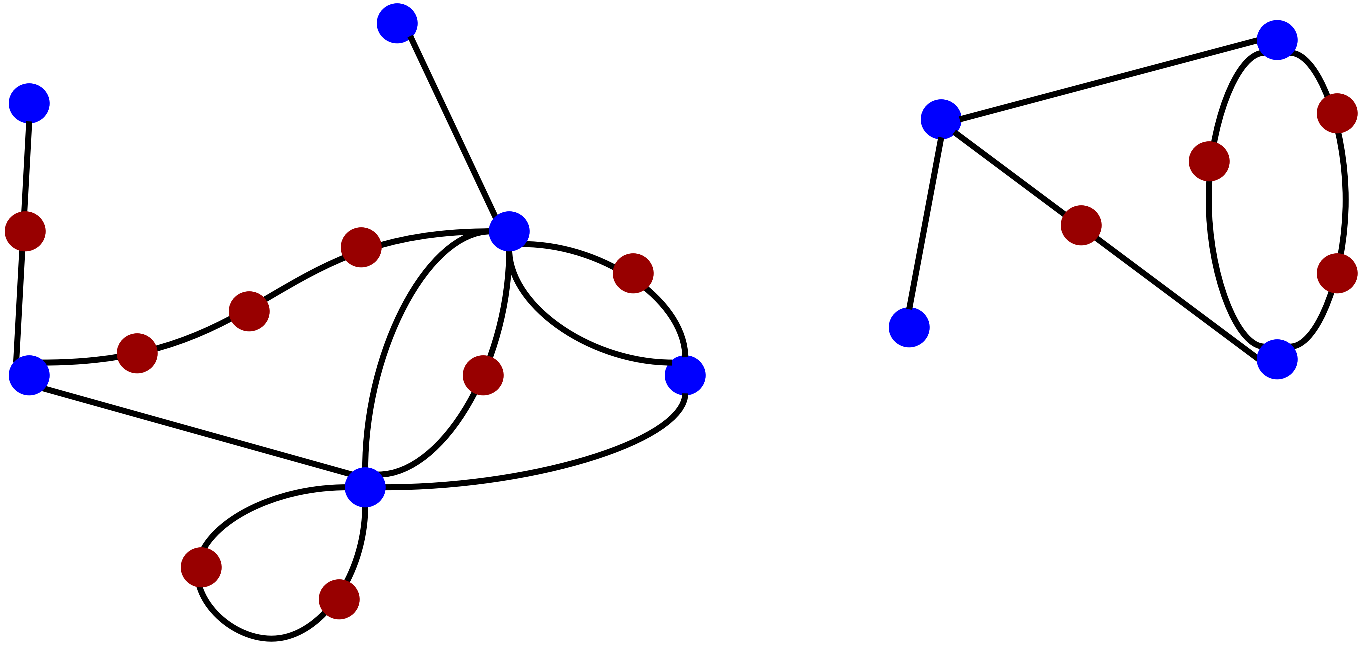
Why we focus on M and not n

And edges not vertices



Why we focus on M and not n

And edges not vertices



What About Badly Behaved Graphs?

Badly Behaved graphs do not have 0-1 Behaviour

Badly Behaved graphs do not have 0-1 Behaviour

For all $0<\varepsilon<1$, the probability of a component of size at least εn lies between c and $1-c$ for some constant c between 0 and 1.

Badly Behaved graphs do not have 0-1 Behaviour

For all $0<\varepsilon<1$, the probability of a component of size at least εn lies between c and 1-c for some constant c between 0 and 1.

If all vertices of degree 2 just taking a random 2-factor.

Badly Behaved graphs do not have 0-1 Behaviour

For all $0<\varepsilon<1$, the probability of a component of size at least εn lies between c and 1-c for some constant c between 0 and 1.

If all vertices of degree 2 just taking a random 2-factor.

If M is at most some constant b, with probability $p(b)>0$ all but $\varepsilon n/2$ of the vertices lie in cyclic components.

Two Theorems

Theorem 1: For any $f \to \infty$ and $b \to 0$, if a well behaved degree distribution D satisfies $R_D \leq b(n)M$ then G(D) has no giant component.

Theorem 2: For any $f \to \infty$ and $\varepsilon > 0$ if a well behaved degree distribution D satisfies $R_D \geq \varepsilon M$ then G(D) has a giant component

(Joos, Perarnau-Llobet, Rautenbach, Reed 2015)

Differences in the Proof

Determine if there is a component K of the multigraph obtained by suppressing degree 2 vertices satisfying:

(*) $|E(K)| > \varepsilon' M$.

Use a combinatorial switching argument to obtain bounds on edge probabilities in this multigraph.

Differences in the Proof - When No Giant Component Exists

Begin the random process with a large enough set of high degree vertices that our process has negative drift.

Differences in the Proof - When No Giant Component Exists

Begin the random process with a large enough set of high degree vertices that our process has negative drift.

Show drift becomes more and more negative over time, if the process does not die out.

Differences in the Proof - When A Giant Component Exists

Focus on the set H = {v | d(v) > $\sqrt{M}$)/log(M)}

Differences in the Proof - When A Giant Component Exists

Focus on the set $H = \{v \mid d(v) > \sqrt{M})/\log(M)\}$

We can show, using our combinatorial switching argument, that depending on the sum of the sizes of the components intersecting H, either

(a) there is a giant component containing all of H , or
(b) we can reduce to a problem with H empty.

For which the conditions ensuring that a giant component exists hold.

Thank you for your attention!