Component structure of the vacant set induced by a random walk on a random graph

Colin Cooper (KCL)
Alan Frieze (CMU)
Vacant set definition and results
Introduction: random walks and cover time
$G_{n,p}$ vacant set
Random $r$-regular graphs, vacant set
Random walk on $600 \times 600$ toroidal grid. Black visited, white unvisited.
What is the component structure of vacant set?
Notation

Finite graph $G = (V, E)$.

$\mathcal{W}_u$ Simple random walk on $G$, starting at $u \in V$.

The vacant set
$\mathcal{R}(t)$ Set of vertices unvisited by $\mathcal{W}_u$ up to time $t$
$\Gamma(t)$ Sub-graph of $G$ induced by vacant set $\mathcal{R}(t)$

Can think of vacant set $\mathcal{R}(t)$ as coloured red, and visited vertices $\mathcal{B}(t)$ as colored blue.

How large is $\mathcal{R}(t)$?
What is the likely component structure of $\Gamma(t)$?
Evolution of vacant set

As the walk progresses $\Gamma(t)$ is reduced from the whole graph $G$ to a graph with no vertices

In the context of sparse random graphs, as $R(t)$ gets smaller, $\Gamma(t)$ will get sparser and sparser. (Small sets don’t induce many edges)

One might expect that at some time $\Gamma(t)$ will break up into small components
This is basically what we prove. It is a sort of random graph process in reverse
We say that $\Gamma(t)$ is \textit{sub-critical} at step $t$, if all of its components are of size $O(\log n)$

We say that $\Gamma(t)$ is \textit{super-critical} at step $t$, if it has a unique giant component, (of size $\Theta(\mathcal{R}(t))$ ) and all other components are of size $O(\log n)$

In the cases we consider there is a $t^*$, which is a (\textit{whp}) threshold for transition from super-criticality to sub-criticality
Vacant set of $G_{n,p}$. We assume that

$$p = \frac{c \log n}{n}$$

where $(c - 1) \log n \to \infty$ with $n$, and $c = n^{o(1)}$. Let

$$t(\epsilon) = n (\log \log n + (1 + \epsilon) \log c)$$

**Theorem**

*Let $\epsilon > 0$ be a small constant. Then whp we have*

(i) $\Gamma(t)$ is super-critical for $t \leq t(-\epsilon)$

(ii) $\Gamma(t)$ is sub-critical for $t \geq t(\epsilon)$

Giant component of $\mathcal{R}(t)$ until $t > n \log \log n$

Cover time $T_{cov} \sim n \log n$ when $c > 1$ constant
Random graphs $G_{n,r}$

For $r \geq 3$, constant, let

$$t^* = \frac{r(r - 1) \log(r - 1)}{(r - 2)^2} n$$

Theorem

Let $\epsilon > 0$ be a small constant. Then whp we have

(i) $\Gamma(t)$ is super-critical for $t \leq (1 - \epsilon)t^*$

(ii) For $t \leq (1 - \epsilon)t^*$, size of giant component is $\Omega(n)$

(iii) $\Gamma(t)$ is sub-critical for $t \geq (1 + \epsilon)t^*$

e.g. for 3-regular random graphs $r = 3$, and $t^* = (6 \log 2) n$

Giant component until $t^*(6 \log 2)n$

Cover time $T_{cov} \sim 2n \log n$
Previous Work

Benjamini and Sznitman; Windisch:
Considered the infinite $d$-dimensional torus $d \geq 3$, and discrete torus for large $d$

Černy, Teixeira and Windisch:
Considered random $r$-regular graphs $G_{n,r}$
They show sub-criticality for $t \geq (1 + \epsilon)t^*$
and existence of a unique giant component for $t \leq (1 - \epsilon)t^*$
These proofs use the concept of random interlacements of continuous time random walks
Our proof: Discrete time

- Simple. Based on established random graph results
- Gives results for $G_{n,p}$
- Completely characterizes the component structure
- Proves that in the super-critical phase $t \leq t^*$, the second largest component of $G_{n,r}$ has size $O(\log n)$ whp
  - Gives the small tree structure of $\Gamma(t)$

Subsequent Work: Černy, Teixeira and Windisch:
Consider random $r$-regular graphs $G_{n,r}$
Investigate scaling window around $t^*$ using annealed model
Proof technique: $r$-regular r.g’s

- Use walk to reveal the graph: Annealed model
- Estimate un-visit probability of vertices by walk and hence size and degree sequence $d$ of vacant set $R(t)$
- Graph $\Gamma(t)$ induced by vacant set $R(t)$ is random
- Given degree sequence $d$ of $\Gamma(t)$, use Molloy-Reed condition for existence and size of giant component
- Count small tree components
Cover time $T_{cov}$ of random walk on graph $G$

$T_{cov}$ is the maximum expected time, over all start vertices $u$, for a random walk $\mathcal{W}_u$ to visit all vertices of $G$. 
Cover time $T_{cov}$ of random walk on graph $G$

$T_{cov}$ is the maximum expected time, over all start vertices $u$, for a random walk $\mathcal{W}_u$ to visit all vertices of $G$.

We studied \textit{whp} cover time of random walks on random graphs.
Cover time $T_{cov}$ of random walk on graph $G$

$T_{cov}$ is the maximum expected time, over all start vertices $u$, for a random walk $\mathcal{W}_u$ to visit all vertices of $G$.

We studied \texttt{whp} cover time of random walks on random graphs

1. Erdös-Renyi random graphs $G_{n,p}$
   Let $np = c \log n$ and $(c - 1) \log n \to \infty$ then

   $$T_{cov} \sim c \log \left( \frac{c}{c - 1} \right) n \log n.$$
Cover time $T_{cov}$ of random walk on graph $G$

$T_{cov}$ is the maximum expected time, over all start vertices $u$, for a random walk $\mathcal{W}_u$ to visit all vertices of $G$.

We studied **whp** cover time of random walks on random graphs

1. Erdös-Renyi random graphs $G_{n,p}$
   Let $np = c \log n$ and $(c - 1) \log n \to \infty$ then
   
   $$T_{cov} \sim c \log \left( \frac{c}{c - 1} \right) n \log n.$$

2. Random regular graphs, where $3 \leq r = O(1)$ then
   
   $$T_{cov} \sim \frac{r - 1}{r - 2} n \log n.$$
Cover time $T_{cov}$ of random walk on graph $G$

$T_{cov}$ is the maximum expected time, over all start vertices $u$, for a random walk $W_u$ to visit all vertices of $G$.

We studied **whp** cover time of random walks on random graphs

1. Erdös-Renyi random graphs $G_{n,p}$
   Let $np = c \log n$ and $(c - 1) \log n \to \infty$ then
   $$T_{cov} \sim c \log \left( \frac{c}{c - 1} \right) n \log n.$$

2. Random regular graphs, where $3 \leq r = O(1)$ then
   $$T_{cov} \sim \frac{r - 1}{r - 2} n \log n.$$

3. Web-graphs $G(m, t)$ where $m \geq 2$
   $$T_{cov} \sim \frac{2m}{m - 1} t \log t.$$
Directed graphs: random digraphs $\mathcal{D}_{n,p}$

The main challenge for $\mathcal{D}_{n,p}$, was to obtain the stationary distribution

**Theorem**

Let $np = d \log n$ where $d = d(n)$, and let $m = n(n-1)p$

Let $\gamma = np - \log n$, and assume $\gamma = \omega(\log \log n)$

Then whp, for all $v \in V$,

$$\pi_v \sim \frac{\deg^-(v)}{m},$$

and

$$T_{cov} \sim d \log \left( \frac{d}{d-1} \right) n \log n$$
Basic idea to estimate $T_{cov}$. Rapidly mixing graphs

Estimate the un-visit probability of states. The probability a given state has not been visited after $t$ steps of the process.

For rapidly mixing processes this is (at most)

The probability a given state has not been visited after $t$ steps of the process, starting from stationarity.

The expected hitting time of state $v$ from stationarity can be approximated by $E_{\pi}H_v \sim R_v/\pi_v$ where $R_v$ is expected number of returns to $v$ during a suitable mixing time.

Waiting time of first visit to $v$ tends to geometric distn, success probability $p_v \sim \pi_v/R_v$. 

Basic idea to estimate $T_{cov}$. Rapidly mixing graphs

- Estimate the **un-visit probability** of states
  The probability a given state has has not been visited after $t$ steps of the process
Basic idea to estimate $T_{cov}$. Rapidly mixing graphs

- Estimate the un-visit probability of states
  The probability a given state has has not been visited after $t$ steps of the process
- For rapidly mixing processes this is (at most)
  The probability a given state has has not been visited after $t$ steps of the process, starting from stationarity

$$E_{\pi} \sim R_v / \pi_v$$ where $R_v$ is expected number of returns to $v$ during a suitable mixing time

The waiting time of first visit to $v$ tends to geometric distn, success probability $p_v \sim \pi_v / R_v$
Basic idea to estimate $T_{cov}$. Rapidly mixing graphs

- Estimate the un-visit probability of states
  The probability a given state has not been visited after $t$ steps of the process
- For rapidly mixing processes this is (at most)
  The probability a given state has not been visited after $t$ steps of the process, starting from stationarity
- The expected hitting time of state $v$ from stationarity can be approximated by

$$E_{\pi} H_v \sim R_v / \pi_v$$

where $R_v$ is expected number of returns to $v$ during a suitable mixing time
Basic idea to estimate $T_{cov}$. Rapidly mixing graphs

- Estimate the un-visit probability of states
  The probability a given state has has not been visited after $t$ steps of the process
- For rapidly mixing processes this is (at most)
  The probability a given state has has not been visited after $t$ steps of the process, starting from stationarity
- The expected hitting time of state $v$ from stationarity can be approximated by

\[
E_{\pi} H_v \sim R_v / \pi_v
\]

where $R_v$ is expected number of returns to $v$ during a suitable mixing time
- Waiting time of first visit to $v$ tends to geometric distn, success probability $p_v \sim \pi_v / R_v$
Summary: Unvisit Probability

Let $T_{\text{mix}}$ be a suitable mixing time of the walk.
Let $\pi_v$ denote the stationary distribution of $v$.
Let $R_v$ denote the expect number of returns to $v$ by the walk $\mathcal{W}_v$ in the time $T_{\text{mix}}$.

Then Unvisit Probability

$$\Pr(\mathcal{W}_u(\tau) \neq v : \tau = T_{\text{mix}}, \ldots, t) \sim e^{-t\pi_v/R_v}$$

True under assumptions that hold for many random graph models

For random graphs we can estimate $R_v$ accurately from the graph structure for most vertices, and bound it suitably for all vertices.
Example: $r$-regular random graphs

How to calculate $R_v$ for random $r$-regular graphs?

If $v$ is tree-like (not near any short cycles) then $R_v \sim \frac{r-1}{r-2}$

Same as: biased random walk on the half line $(0, 1, 2, ...)$

$Pr(\text{go left}) = \frac{1}{r}$, \hspace{1cm} $Pr(\text{go right}) = \frac{r-1}{r}$
Example: $r$-regular random graphs

- $\pi_v = 1/n$
- $T_{mix}$ the mixing time $O(\log n)$
- Most vertices are locally tree-like
  For such vertices $R_v \sim (r - 1)/(r - 2)$, expected number of returns to start in infinite $r$-regular tree
Example: $r$-regular random graphs

- $\pi_v = 1/n$
- $T_{\text{mix}}$ the mixing time $O(\log n)$
- Most vertices are locally tree-like
  
  For such vertices $R_v \sim (r - 1)/(r - 2)$, expected number of returns to start in infinite $r$-regular tree

\[
\Pr(\text{v unvisited in } T_{\text{mix}}, \ldots, t) \sim e^{-t\pi_v/R_v} \sim e^{-t(r-2)/(r-1)n}
\]
Example: $r$-regular random graphs

- $\pi_v = 1/n$
- $T_{mix}$ the mixing time $O(\log n)$
- Most vertices are locally tree-like
  For such vertices $R_v \sim (r-1)/(r-2)$, expected number of returns to start in infinite $r$-regular tree

$$\Pr(\text{v unvisited in } T_{mix}, \ldots, t) \sim e^{-t\pi_v/R_v}$$

$$\sim e^{-t(r-2)/(r-1)n}$$

- $\mathcal{R}(t)$ set of vertices not visited by walk at step $t$
- Size of set of unvisited vertices $\mathcal{R}(t) \sim ne^{-t(r-2)/(r-1)n}$
Example: $r$-regular random graphs

- $\pi_v = 1/n$
- $T_{mix}$ the mixing time $O(\log n)$
- Most vertices are locally tree-like
  - For such vertices $R_v \sim (r - 1)/(r - 2)$, expected number of returns to start in infinite $r$-regular tree

\[
\Pr(v \text{ unvisited in } T_{mix}, \ldots, t) \sim e^{-t\pi_v/R_v} \\
\sim e^{-t(r-2)/(r-1)n}
\]

- $R(t)$ set of vertices not visited by walk at step $t$
- Size of set of unvisited vertices $R(t) \sim ne^{-(r-2)/(r-1)n}$
- We know the size of $R(t)$, the vacant set
Example: $r$-regular random graphs

- $\pi_v = 1/n$
- $T_{mix}$ the mixing time $O(\log n)$
- Most vertices are locally tree-like
  For such vertices $R_v \sim (r - 1)/(r - 2)$, expected number of returns to start in infinite $r$-regular tree

$$
\Pr(v \text{ unvisited in } T_{mix}, \ldots, t) \sim e^{-t\pi_v/R_v} \\
\sim e^{-t(r-2)/(r-1)n}
$$

- $\mathcal{R}(t)$ set of vertices not visited by walk at step $t$
- Size of set of unvisited vertices $\mathcal{R}(t) \sim ne^{-t(r-2)/(r-1)n}$
- We know the size of $\mathcal{R}(t)$, the vacant set
- Now we need to find the structure of $\mathcal{R}(t)$
Component structure of vacant set of $G_{n,p}$
Distribution of edges in $\Gamma(t)$

**Lemma**

Consider a random walk on $G_{n,p}$

Conditional on $N = |\mathcal{R}(t)|$, $\Gamma(t)$ is distributed as $G_{N,p}$.

**Proof**
This follows easily from the principle of deferred decisions. We do not have to expose the existence or absence of edges between the unvisited vertices of $\mathcal{R}(t)$  

Thus to find the super-critical/ sub-critical phases, we only need high probability estimates of $|\mathcal{R}(t)|$ as $t$ varies

This, we know how to do, from our work on cover time of random graphs
Size of vacant set $\mathcal{R}(t)$ in $G_{n,p}$

Analysis of $G_{n,p}$ is for $np = c \log n$

whp

1. $E(|\mathcal{R}(t)|) \sim \sum_v e^{-rt\pi_v/R_v}$
Size of vacant set $\mathcal{R}(t)$ in $G_{n,p}$

Analysis of $G_{n,p}$ is for $np = c \log n$

whp

1. $\mathbb{E}(|\mathcal{R}(t)|) \sim \sum_v e^{-t \pi_v/R_v}$

2. Almost all vertices have $\sim$ average degree $c \log n$
   Thus $\pi_v \sim 1/n$
Size of vacant set $\mathcal{R}(t)$ in $G_{n,p}$

Analysis of $G_{n,p}$ is for $np = c \log n$

whp

1. $\mathbb{E}(|\mathcal{R}(t)|) \sim \sum_v e^{-t \pi_v / R_v}$
2. Almost all vertices have $\sim$ average degree $c \log n$
   Thus $\pi_v \sim 1/n$
3. $R_v = 1 + o(1)$ for all $v \in V$

Size of vacant set

$$\mathbb{E}(|\mathcal{R}(t)|) \sim n e^{-(1+o(1))t/n}.$$ 

We use Chebyshev to show that $|\mathcal{R}(t)|$ is concentrated.
Size of 'giant' component

- Threshold criteria for random graph $G_{N,p}$ is $Np \sim 1$
- Recall that $t_\theta = n(\log \log n + (1 + \theta) \log c)$ So, at $t_\theta$,
  \[
  \mathbb{E}(|R(t_\theta)|^p) \sim \frac{1}{c^\theta}
  \]

- When $\theta = 0$, then $\mathbb{E}(|R(t_\theta)|^p) \sim 1$
- The threshold $t^*$ occurs at around
  \[
  t^* \sim n(\log \log n + \log c)
  \]

- Size of giant is order $|R(t_\theta)|$. As $t \rightarrow t^*$ from below, Size of 'giant' is order $1/p = n/(c \log n)$. i.e. $|R(t^*)| \sim 1/p$
- Above $t^*$ max component size collapses to $O(\log n)$
Random regular graphs

Component structure of vacant set of random graphs $G_{n,r}$ for $r \geq 3$, constant.
Reminder: Vacant set of $r$-regular random graphs

- Most vertices are locally tree-like
  For such vertices $R_v \sim (r - 1)/(r - 2)$, expected number of returns to start in infinite $r$-regular tree

  \[ \Pr(v \text{ unvisited in } T_{mix}, \ldots, t) \sim e^{-t(r-2)/(r-1)n} \]

- A similar upper bound can be obtained for the non-tree-like vertices
- Size of vacant set $\mathcal{R}(t) \sim ne^{-t(r-2)/(r-1)n}$
Let
\[ t^* = \frac{r(r - 1) \log(r - 1)}{(r - 2)^2} n. \]

**Theorem**

Let \( \epsilon > 0 \) be a small constant. Then whp we have

(i) \( \Gamma(t) \) is super-critical for \( t \leq (1 - \epsilon)t^* \),

(ii) For \( t \leq (1 - \epsilon)t^* \), size of giant component is \( \Omega(n) \)

(iii) \( \Gamma(t) \) is sub-critical for \( t \geq (1 + \epsilon)t^* \) and
Proof outline for $r$-regular random graph

- Generate the graph in the configuration model using the random walk
- Graph $\Gamma(t)$ induced by vacant set $\mathcal{R}(t)$ is random
- Estimate un-visit probability of vertices to find size of $\mathcal{R}(t)$
- Estimate degree sequence $d$ of $\Gamma(t)$ in the configuration model, using size of vacant set $\mathcal{R}(t)$, and number of unvisited edges $\mathcal{U}(t)$
- Given the degree sequence $d$ of $\Gamma(t)$, we can use Molloy-Reed condition for existence of giant component in a random graph with fixed degree sequence
- Estimate number of small trees in configuration model
Degree sequence of $\Gamma(t)$

Vacant set size. $|R(t)| = (1+o(1))N_t$ where $N_t = ne^{-\frac{(r-2)t}{(r-1)n}}$

Vertex degree. Let $D_s(t)$ the number of unvisited vertices of $\Gamma(t)$ of degree $s$ in $\Gamma(t)$ (ie with $r-s$ visited neighbours)
For $0 \leq s \leq r$, and for ranges of $t$ given below, whp

$$D_s(t) \sim N_t \binom{r}{s} p_t^s (1-p_t)^{r-s}$$

where $p_t = e^{-\frac{(r-2)^2}{(r-1)r} \frac{t}{n}}$

Range of validity. $\tau_{r-s} \ll t \leq (1-\epsilon)t_s$ where $\tau_0 = 0$,

$$\tau_{r-s} = n^{1-1/(r-s)}, \quad t_s = \frac{(r-1)r}{(r-2)(s(r-2)+r)} \cdot n \log n.$$
Uniformity

Lemma

Consider a random walk on $G_r$. Conditional on $N = |\mathcal{R}(t)|$ and degree sequence $d = d_{\Gamma(t)}(v), v \in \mathcal{R}(t)$, then $\Gamma(t)$ is distributed as $G_{N,d}$, the random graph with vertex set $[N]$ and degree sequence $d$.

Proof   Basic idea: Reveal $G_r$ using the random walk. Suppose that we condition on $\mathcal{R}(t)$ and the history of the walk, $\mathcal{H} = (W_u(0), W_u(1), \ldots, W_u(t))$. If $G_1, G_2$ are graphs with vertex set $\mathcal{R}(t)$ and if they have the same degree sequence then substituting $G_2$ for $G_1$ will not conflict with $\mathcal{H}$. Every extension of $G_1$ is an extension of $G_2$ and vice-versa. □

Thus we only need:
Good model of component structure of $G_{N,d}$
High probability estimates of the degree sequence $D_s(t)$ of $\Gamma(t)$. 


Main variables

By calculating un-visit probabilities in various ways, we can estimate the size at step \( t \) of

- \( \mathcal{R}(t) \) the set of unvisited vertices
- \( \mathcal{U}(t) \) the set of unvisited edges
- \( D_s(t) \) the number of unvisited vertices of degree \( s \) in \( \Gamma(t) \)
- ie number of unvisited vertices with \( r - s \) edges incident with visited vertices \( \mathcal{B}(t) \)
Annealed process

We use the random walk to generate the graph in the configuration model as a random pairing $F$

- $B_t$ blue config. points at step $t$
  which form discovered pairing $F_t$
- $R_t$ red config. points at step $t$
  This will form un-generated pairing $F - F_t$
- Visited vertices may have config. points in $R_t$, corresponding to unexplored edges
Next configuration pairing

Example: Move to an unvisited vertex
Walk at current vertex $X_t \in B(t)$
Given the walk selects a red config. point of $X_t$ (if any), the probability this is paired with an config. point in $\mathcal{R}(t)$ is $\frac{r|\mathcal{R}(t)|}{|R_t|-1}$
Shrinking Vertices: First visit to a set of vertices $S$

$S$ subset of vertices of $G$. $\gamma(S)$ is $S$ shrunk to a vertex $\Gamma(G)$ is $G$ with $S$ shrunk to $\gamma(S)$

$$\Pr_G(S \text{ unvisited at step } t) \sim \Pr_{\Gamma(G)}(\gamma(S) \text{ unvisited at step } t)$$

Note: Notation overloaded $\Gamma(t)$ and $\Gamma(G)$—apologies
Degree of unvisited vertex

Vertex $v$ has 3 unvisited neighbours $x, y, z$ and 2 visited neighbours $a, b$, so $s = 3$, $r - s = 2$

Calculate probability that exactly $\{v, x, y, z\}$ are unvisited, and $a, b$ visited from probability that $\{v, x, y, z\}$ are unvisited, $\{v, x, y, z, a\}$ are unvisited etc. Contract e.g. $\{v, x, y, z\}$ to a single vertex $\gamma$ of degree 20 with 3 loops
The degree sequence of $\mathcal{R}(t)$

Unvisit probability

$$\Pr(v \in \mathcal{R}(t)) \sim e^{-t(r-2)/(r-1)n}.$$ 

To analyse the degree sequence of $\Gamma(t)$ we prove

Lemma

If the neighbours of $v$ in $G$ are $w_1, w_2, \ldots, w_r$ then

$$\Pr(v, w_1, \ldots, w_s \in \mathcal{R}_t, w_{s+1}, \ldots, w_r \in \mathcal{B}(t))$$

$$\sim e^{-(r-2)t/(r-1)n} p^s_t (1 - p_t)^{r-s}$$

where $p_t = e^{-t(r-2)^2/n(r-1)}$.
We write

\[ \Pr_{\mathcal{W}}(\{v, w_1, \ldots, w_s\} \subseteq \mathcal{R}(t) \text{ and } \{w_{s+1}, \ldots, w_r\} \subseteq \mathcal{B}(t)) \]

\[ = \sum_{X \subseteq [s+1,r]} (-1)^{|X|} \Pr_{\mathcal{W}}((\{v, w_1, \ldots, w_s\} \cup X) \subseteq \mathcal{R}(t)) \]

\[ \sim \sum_{X \subseteq [s+1,r]} (-1)^{|X|} e^{-tp_{\gamma_X}}, \]

where

\[ p_{\gamma_X} \sim \frac{((r - 2)(s + |X|) + r)(r - 2)}{r(r - 1)n}. \]

To prove this we contract \( \{v, w_1, \ldots, w_s\} \cup X \) to a single vertex \( \gamma_X \) creating \( \Gamma_X(t) \).

We then estimate the probability that \( \gamma_X \) hasn’t been visited by a random walk on \( \Gamma_X(t) \). (Unvisit probability)
For this we argue that $|\{v, w_1, \ldots, w_s\} \cup X| = s + |X| + 1$

$$\pi_{\gamma X} = \frac{r(s + |X| + 1)}{rn}$$

and

$$R_{\gamma X} \sim \frac{(s + |X| + 1)r(r - 1)}{((r - 2)(s + |X|) + r)(r - 2)}$$

Expression for $R_{\gamma X}$ is obtained by considering the expected number of returns to the origin in an infinite tree with branching factor $r - 1$ at each non-root vertex. At the root there are $s + |X|$ loops and $(r - 2)(s + |X|) + r$ branching edges..
Reminder: $R_v$ for random $r$-regular graphs

A transition on the loops returns to $\gamma_X$ immediately, and a transition on any other edge is (usually) like a walk in a tree.

If $v$ is tree-like (not near any short cycles) then $R_v \sim \frac{r-1}{r-2}$

Same as: random walk on the line $(0, 1, 2, \ldots)$

$\Pr( \text{go left} ) = \frac{1}{r}$, \hspace{1cm} $\Pr( \text{go right} ) = \frac{r-1}{r}$
Degree sequence of $\Gamma(t)$. Molloy-Reed

Unvisit probability

$$\text{Pr}(v \in R(t)) \sim e^{-t(r-2)/(r-1)n}$$

and the degree of a vertex in $\Gamma(t)$ is (approximately) binomial $Bin(r, p_t)$ where $p_t = e^{-t(r-2)^2/(n(r-1)^r)}$

Once we know the degree sequence we can use the Molloy-Reed criterion to see whether or not there is a giant component. $G$ has a giant component iff $S > 0$, where

$$S = \sum_v d_v(d_v - 2).$$

Direct calculation gives $t^* = \frac{r(r-1)\log(r-1)}{(r-2)^2} n$ as the critical value
Heuristically, \( t^* = \frac{r(r-1)\log(r-1)}{(r-2)^2} n \) can be obtained from the degree sequence of unvisited vertices.

Branching outward from an unvisited vertex
The probability an edge goes to another unvisited vertex:

\[
p_t = e^{-\frac{(r-2)^2 t}{(r-1)n}}
\]

We need branching factor \((r - 1)p_t > 1\), to have a chance to get a large component.

At \( t^* = \frac{r(r-1)\log(r-1)}{(r-2)^2} n \)

\[
(r - 1)p_t = (r - 1)e^{-\frac{(r-2)^2 t}{(r-1)n}}
= (r - 1)e^{-\log(r-1)}
= 1
\]
Rooted subtrees of the infinite regular tree

Number of rooted $k$-subtrees of the infinite $r$-regular tree

$$\frac{r}{((r - 2)k + 2)} \binom{(r - 1)k}{k - 1}$$
Number of small components in $\Gamma(t)$

$N_t = \mathbb{E}|\mathcal{R}(t)|$. Expected size of vacant set

$p_t$ probability of a red edge

$N(k, t)$: Number of unvisited tree components of $\Gamma(t)$ with $k$ vertices

Theorem

Let $\epsilon$ be a small positive constant. Let $1 \leq k \leq \epsilon \log n$ and $\epsilon n \leq t \leq (1 - \epsilon) t_{k-1}$. Then whp:

$$N(k, t) \sim r \frac{r}{k((r-2)k+2)} \binom{(r-1)k}{k-1} N_t p_t^{k-1}(1 - p_t)^{k(r-2)+2}$$
Vertices on small components of vacant set

Let

\[ t^* = n \frac{r(r - 1)}{(r - 2)^2} \log(r - 1). \]

Theorem

Let \( \mu(t) \) be the expected proportion of vertices on small trees. The function \( \mu(t) \) increases from 0 at \( t = 0 \), to a maximum value \( \mu^* = \frac{1}{(r - 1)^{r/(r-2)}} \) at \( t \to t^* \), and decreases to 0 as \( t \to \frac{(r - 1)}{(r - 2)} n \log n \).
Example: $r = 3$. Vacant set as a function of $\tau = t/n$

Proportion of vertices in vacant set $N(t)/n \sim e^{-t/n((r-2)/(r-1))}$

Proportion of vertices in unvisited tree components
Threshold: \( r = 3, \ t^* = 6 \log 2 \)

\[
t^* = \frac{r(r - 1) \log(r - 1)}{(r - 2)^2} n
\]

The cusp is at \( t^* = 6 \log 2 \sim 4 \), with \( \mu^* = 1/8 \).

Propn. of vertices in vacant set, and on small tree components
Closing observations

- Both classes of graphs \((G(n, p), G(n, r))\) exhibit threshold behavior.
- The size of the giant can be estimated in the super-critical range.
- The number of small components of a given size can be estimated.
- The technique is simple, but seems restricted to random graphs.
THANK YOU

QUESTIONS