CONNECTIVITY PROPERTIES AND STRUCTURE OF RANDOM GRAPHS OBTAINED BY VERTEX PERCOLATION ON ERDŐS-RÉNYI BINOMIAL RANDOM GRAPH

Bhaskar Ray & Dibyendu Saha

GUIDE: ANTAR BANDYOPADHYAY

JUNE 18, 2020



The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn

The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn



The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn



 \rightarrow

The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn

Edge Deletion



 \rightarrow

The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn

Edge Deletion







The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn

Edge Deletion







The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph,Kn



Vertex Percolation







 \rightarrow

The Problem

There is a large state with several cities with a road between every pair of cities. Suddenly, the state decides to block a few randomly chosen roads for maintenance. Also, some of the cities randomly decide not to allow travel through roads that have one end in that city, independently of the other cities. Now, a traveller standing in one of the remaining cities wants to know the chances that he can travel to all of the remaining cities without having to break any rule.

Complete Graph, Kn



 \rightarrow

Edge Deletion



Vertex Percolation



 \rightarrow

MATHEMATICAL FORMULATION OF THE PROBLEM

Let, the following be defined,

$$\mathcal{V}_n := V(K_n) = [n]$$

$$\mathcal{E}_n := E(K_n) = \{\{u, v\} | u \neq v ; u, v \in [n]\}$$

Let

 $\begin{aligned} &(X_v)_{v \in \mathcal{V}_n} \sim \text{i.i.d. Bernoulli} (q_n) \\ &(Y_e)_{e \in \mathcal{E}_n} \sim \text{i.i.d. Bernoulli} (p_n) \\ &X_v \And Y_e \text{ are independent } \forall \ v \in \mathcal{V}_n \And e \in \mathcal{E}_n \end{aligned}$

Let G be random graph such that

$$V(G) := \{ v | v \in \mathcal{V}_n \& X_v = 1 \}$$

$$E(G) := \{ e := \{ u, v \} | e \in \mathcal{E}_n \& X_v = 1, X_u = 1, Y_e = 1 \}$$

Such a G is our graph of interest. We will denote such a graph as

 $G \sim \mathcal{G}(n, p_n, q_n)$

RELATION WITH ERDŐS-RÉNYI BINOMIAL RANDOM GRAPH

For such a $G \sim \mathcal{G}(n, p_n, q_n)$ we have the following results.

RELATION WITH ERDŐS-RÉNYI BINOMIAL RANDOM GRAPH

For such a $G \sim \mathcal{G}(n, p_n, q_n)$ we have the following results.

Result

 $\mathbb{P}\left(e \in E(G)\right) = q_n^2 p_n \; \forall e \in \mathcal{E}_n$

Result

 $\mathbb{P}\left(e = \{u, v\} \in E(G) | X_u = 1, X_v = 1\right) = p_n$

For such a $G \sim \mathcal{G}(n, p_n, q_n)$ we have the following results.

Result

 $\mathbb{P}\left(e \in E(G)\right) = q_n^2 p_n \; \forall e \in \mathcal{E}_n$

Result

 $\mathbb{P}\left(e = \{u, v\} \in E(G) | X_u = 1, X_v = 1\right) = p_n$

Result

If vertex percolation is done first, the graph obtained is an Erdős-Rényi Binomial Random Graph on the remaining set of vertices i.e., $G|(X_v)_{v \in \mathcal{V}_n} \sim \mathcal{G}(|V(G)|, p_n)$

For such a $G \sim \mathcal{G}(n, p_n, q_n)$ we have the following results.

Result

 $\mathbb{P}\left(e \in E(G)\right) = q_n^2 p_n \; \forall e \in \mathcal{E}_n$

Result

 $\mathbb{P}\left(e = \{u, v\} \in E(G) | X_u = 1, X_v = 1\right) = p_n$

Result

If vertex percolation is done first, the graph obtained is an Erdős-Rényi Binomial Random Graph on the remaining set of vertices i.e., $G|(X_v)_{v \in \mathcal{V}_n} \sim \mathcal{G}(|V(G)|, p_n)$

Result

 $\mathbb{E}\left[|V(G)|\right] = nq_n$









Connectivity in Erdős-Rényi Binomial Random Graph

Theorem (Connectivity Threshold)

Let $G \sim \mathcal{G}(n, p_n)$. Then,

$$\lim_{n \to \infty} \mathbb{P}_{\mathcal{G}(n,p_n)} \left(G \text{ is connected} \right) = \begin{cases} \mathsf{O} & ; \frac{p_n}{\log(n)/n} \to \mathsf{O} \\ 1 & ; \frac{p_n}{\log(n)/n} \to \infty \end{cases}$$

Theorem (Critical Window)

Fix $t \in \mathbb{R}$ and $\lambda_n = \log n + t$;

$$\lim_{n\to\infty}\mathbb{P}_{\mathcal{G}\left(n,\frac{\lambda_n}{n}\right)}(G \text{ is connected}) = \exp\left\{-\exp\left\{-t\right\}\right\}$$

CONNECTIVITY IN $\mathcal{G}(n, p_n, q_n)$

In this direction, we have done some simulations, with the following objectives,

Estimating the connectivity probability for different values constant sequence q_n and p_n .

- Estimating the connectivity probability for different values constant sequence q_n and p_n .
- Fitting a functional form of *p* on estimated probability of connectivity.

- Estimating the connectivity probability for different values constant sequence q_n and p_n .
- Fitting a functional form of *p* on estimated probability of connectivity.
- Estimating the parameters.

- Estimating the connectivity probability for different values constant sequence q_n and p_n .
- Fitting a functional form of *p* on estimated probability of connectivity.
- Estimating the parameters.
- Checking the relation between the estimated parameters and *n*.

- Estimating the connectivity probability for different values constant sequence q_n and p_n .
- Fitting a functional form of *p* on estimated probability of connectivity.
- Estimating the parameters.
- Checking the relation between the estimated parameters and *n*.
- Compare those relations with our guess.

FITTING OF FUNCTIONAL FORM

In Erdős-Rényi Binomial Random Graph, we also have an another version of the theorem related to Critical window,

Theorem

Fix $p \in (0, 1)$, then $\mathbb{P}_{\mathcal{G}(n,p)}(G \text{ is connected}) = \exp \{-\exp \{-n(p - \log n/n)\}\}(1 + o(1))$

Our guess is that, in case of connectivity it should behave like Erdős-Rényi Binomial Random Graph but with number of vertices as its expected number of vertices, *nq_n*. Hence, we have done various simulations, to

- Plot $\hat{p}_{con.}$ vs p for wide range of q_n .
- Fit the functional form, $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$
- **Estimate** δ_n and c_n

Our observations from these simulations are,

- The rapid increase in connectivity probability, are mostly around $\log(nq_n)/nq_n$ for several q_n values.
- Fit of the functional form, $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$ produces a very high value of multiple correlation.

ESTIMATED PARAMETERS AND THEIR RELATION WITH n and q_n

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

Previous simulation suggests, that \hat{c}_n is "around" $\log(nq_n)/nq_n$ as c_n is the parameter which represents the value around which the rapid increase from 0 to 1 takes place.

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

- Previous simulation suggests, that \hat{c}_n is "around" $\log(nq_n)/nq_n$ as c_n is the parameter which represents the value around which the rapid increase from 0 to 1 takes place.
- Whereas, δ_n is the parameter which represents the rate of that rapid increase. It is clear from the functional from, that if we increase δ_n then the rapidity will be faster.

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

- Previous simulation suggests, that \hat{c}_n is "around" $\log(nq_n)/nq_n$ as c_n is the parameter which represents the value around which the rapid increase from 0 to 1 takes place.
- Whereas, δ_n is the parameter which represents the rate of that rapid increase. It is clear from the functional from, that if we increase δ_n then the rapidity will be faster.

To check the change of δ_n w.r.t. q_n as well as n we have done the following simulations,

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

- Previous simulation suggests, that \hat{c}_n is "around" $\log(nq_n)/nq_n$ as c_n is the parameter which represents the value around which the rapid increase from 0 to 1 takes place.
- Whereas, δ_n is the parameter which represents the rate of that rapid increase. It is clear from the functional from, that if we increase δ_n then the rapidity will be faster.

To check the change of δ_n w.r.t. q_n as well as n we have done the following simulations,

Simulation 1[Change in $\hat{\delta}_n$ with q_n]

- Fix a 'large enough' value of *n*.
- Fix a wide range of value for *q_n* distributed over the region (0, 1).
- For each value of q_n , plot $\hat{p}_{con.}$ vs $p^* := p \hat{c}_n$ on same plot.

Let, the estimated value of c_n be \hat{c}_n and that of δ_n be $\hat{\delta}_n$.

- Previous simulation suggests, that \hat{c}_n is "around" $\log(nq_n)/nq_n$ as c_n is the parameter which represents the value around which the rapid increase from 0 to 1 takes place.
- Whereas, δ_n is the parameter which represents the rate of that rapid increase. It is clear from the functional from, that if we increase δ_n then the rapidity will be faster.

To check the change of δ_n w.r.t. q_n as well as n we have done the following simulations,

Simulation 1[Change in $\hat{\delta}_n$ with q_n]	Simulation 2[Change in $\hat{\delta}_n$ with n]
 Fix a 'large enough' value of n. Fix a wide range of value for q_n distributed over the region (0, 1). For each value of q_n, plot p̂_{con}. vs p* := p - ĉ_n on same plot. 	 Fix a constant value of q_n. Fix a set of value for large enough n.[In our case {1000, 1001, · · · , 1500}] For each value of n, plot p̂_{con}. vs p* on same plot.



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. q_n



Figure: Change in $\hat{\delta}_n$ w.r.t. *n*



Figure: Change in $\hat{\delta}_n$ w.r.t. *n*



Figure: Change in $\hat{\delta}_n$ w.r.t. *n*

In view of the connection between $\mathcal{G}(n, p_n, q_n)$ and $\mathcal{G}(nq_n, p_n)$ and connectivity results in Erdős-Rényi Binomial Random Graph, our guess is that, when we fit $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$ there,

In view of the connection between $\mathcal{G}(n, p_n, q_n)$ and $\mathcal{G}(nq_n, p_n)$ and connectivity results in Erdős-Rényi Binomial Random Graph, our guess is that, when we fit $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$ there,

$$c_n = rac{\log(nq_n)}{nq_n}$$
 and $\delta_n = nq_n$

In view of the connection between $\mathcal{G}(n, p_n, q_n)$ and $\mathcal{G}(nq_n, p_n)$ and connectivity results in Erdős-Rényi Binomial Random Graph, our guess is that, when we fit $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$ there,

$$c_n = rac{\log(nq_n)}{nq_n}$$
 and $\delta_n = nq_n$

To check the relation between the estimated parameters and *n* and compare that with our guess, we perform the following simulation,

In view of the connection between $\mathcal{G}(n, p_n, q_n)$ and $\mathcal{G}(nq_n, p_n)$ and connectivity results in Erdős-Rényi Binomial Random Graph, our guess is that, when we fit $\hat{p}_{con.} \approx \exp\{-\exp\{-\delta_n(p-c_n)\}\}$ there,

$$\mathcal{L}_n = rac{\log(nq_n)}{nq_n}$$
 and $\delta_n = nq_n$

To check the relation between the estimated parameters and *n* and compare that with our guess, we perform the following simulation,

Simulation 3

- Fix an $n q_n$ relationship.
- For $n \in \{1000, 1001, \cdots, 1500\}$,
 - Simulate values for $\hat{p}_{con.}(p)$ for different values of p.
 - Fit $\hat{p}_{con.}(p) \approx \exp\{-\exp\{-\delta_n (p c_n)\}\}$ and estimate \hat{c}_n and $\hat{\delta}_n$
- Plot *n* vs \hat{c}_n and *n* vs $\log nq_n/nq_n$ in same plot and plot *n* vs $\hat{\delta}_n$ and *n* vs nq_n in another same plot.
- Calculate $Cor(\{\hat{c}_n\}, \{\log nq_n/nq_n\})$ and $Cor(\{\hat{\delta}_n\}, \{nq_n\})$.

8

ESTIMATING THE PARAMETERS AS FUNCTION OF $n [q_n = 1/n^{\frac{1}{8}}]$

Estimating the Parameters as Function of $n \left[q_n = 1/n^{rac{1}{8}} ight]$

Figure: Plot of *n* vs \hat{c}_n and *n* vs $\log nq_n/nq_n$

Figure: Plot of *n* vs $\hat{\delta}_n$ and *n* vs nq_n



Estimating the Parameters as Function of $n [q_n = 0.3]$

Estimating the Parameters as Function of $n [q_n = 0.3]$

Figure: Plot of *n* vs \hat{c}_n and *n* vs $\log nq_n/nq_n$

Figure: Plot of *n* vs $\hat{\delta}_n$ and *n* vs nq_n



ESTIMATING THE PARAMETERS AS FUNCTION OF $n \left[q_n = (n+1)^{\frac{7}{8}} - n^{\frac{7}{8}}\right]$

ESTIMATING THE PARAMETERS AS FUNCTION OF $n \left[q_n = (n+1)^{\frac{7}{8}} - n^{\frac{7}{8}}\right]$

Figure: Plot of *n* vs \hat{c}_n and *n* vs $\log nq_n/nq_n$

Figure: Plot of *n* vs $\hat{\delta}_n$ and *n* vs nq_n



ESTIMATING THE PARAMETERS AS FUNCTION OF $n [q_n = (n + 1)^{rac{3}{4}} - n^{rac{3}{4}}]$

Estimating the Parameters as Function of $n \left[q_n = (n+1)^{\frac{3}{4}} - n^{\frac{3}{4}}\right]$

Figure: Plot of *n* vs \hat{c}_n and *n* vs $\log nq_n/nq_n$

Figure: Plot of *n* vs $\hat{\delta}_n$ and *n* vs nq_n



In view of the simulations and observations, we propose the following conjectures.

In view of the simulations and observations, we propose the following conjectures.

Conjecture (Connectivity 'Threshold' after Percolation)

Let $G \sim \mathcal{G}(n, p_n, q_n)$ such that, $nq_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\lim_{n \to \infty} \mathbb{P}(G \text{ is connected}) = \begin{cases} 0 & ; \frac{p_n}{\log nq_n/nq_n} \to 0\\ 1 & ; \frac{p_n}{\log nq_n/nq_n} \to \infty \end{cases}$$

In view of the simulations and observations, we propose the following conjectures.

Conjecture (Connectivity 'Threshold' after Percolation)

Let $G \sim \mathcal{G}(n, p_n, q_n)$ such that, $nq_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\lim_{n \to \infty} \mathbb{P}(G \text{ is connected}) = \begin{cases} 0 & ; \frac{p_n}{\log nq_n/nq_n} \to 0\\ 1 & ; \frac{p_n}{\log nq_n/nq_n} \to \infty \end{cases}$$

Conjecture ('Critical Window' for $\mathcal{G}(n, q_n, p_n)$)

Fix $t \in \mathbb{R}$ and $\lambda_n = \log nq_n + t$. Provided that $nq_n \to \infty$ as $n \to \infty$,

$$\lim_{n \to \infty} \mathbb{P}_{\mathcal{G}\left(n, \frac{\lambda_n}{nq_n}, q_n\right)}(G \text{ is connected}) = \exp\left\{-\exp\left\{-t\right\}\right\}$$

ΤΗΑΝΚ ΥΟυ