Undergraduate Research Fellowship Programme (URFP) 2020-21
– Summer Research Internship

Poster No.: B7

Name: Zheng Yahuan
University No.: 3035334967
Supervisor: Prof. Jeff Jianfeng Yao

On the Critical Behavior of Erdős–Rényi Random Graphs

Zheng Yahuan

Department of Statistics and Actuarial Science, HKU

ABSTRACT

In this study, we examine the behavior of the Erdős-Rényi random graphs with a focus on the giant component size. A preliminary literature review is conducted, followed by some simulation experiments to visualize the random networks. In addition, a real-world network, the collaboration network of mathematicians, is investigated using real dataset together with simulations. It has been found that the network is highly connected with a significant giant component and is scale-free with a degree sequence that follows a power law with exponential cutoff.

INTRODUCTION

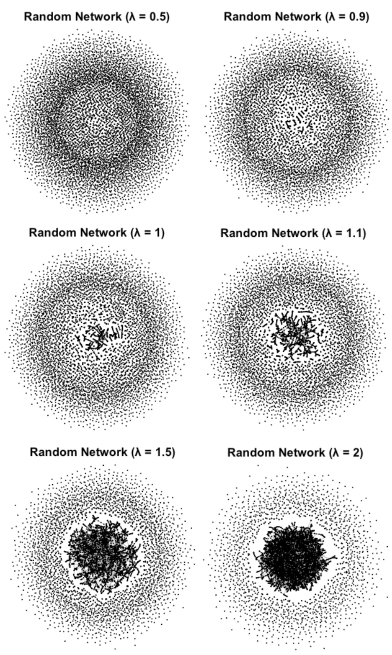
Random graphs serve as powerful tools for modelling large complex networks in multiple application fields. The Erdős-Rényi random graph is the simplest model for random graphs. An Erdős-Rényi graph with size n and edge probability p is defined as the graph generated by randomly and independently drawing edges between n vertices with probability p. For a fixed $\lambda > 0$, the random graph $ER_n(\lambda/n)$ is *subcritical* when $\lambda < 1$, *critical* when $\lambda = 1$, and *supercritical* when $\lambda > 1$.

The Erdős-Rényi model is internally connected with the random walk representation of a branching process. Denote by ζ_λ and η_λ the survival and extinction probability for a population following a branching process, respectively. When $\lambda \le 1$, we have $\zeta_\lambda = 0$ and thus $\eta_\lambda = 1$. When $\lambda > 1$, there is a positive $\zeta_\lambda > 0$ that the population survives forever. Accordingly, the law of large numbers and central limit theorem for the giant component size $\mathcal{C}_{max}$ could be derived for the supercritical regime.

METHODOLOGY

We employ some simulation techniques to visualize the random graph $ER_n(\lambda/n)$. Particularly we are interested in investigating the behavior of the giant component for various choices of λ. Our selected values of λ include $\lambda = 0.5$, $\lambda = 0.9$, $\lambda = 1$, $\lambda = 1.1$, $\lambda = 1.5$ and $\lambda = 2$. We fix the number of vertices to be $n = 10000$.

Simulated ER(10000, λ/10000)



SOME QUICK FACTS FROM LITERATURE REVIEW

> Real-world Networks

- Small-world effect
- Scale-free degree sequence

> $ER_n(\lambda/n)$

- $\lambda < 1$: No giant component
- $\lambda = 1$: $\mathcal{C}_{max} = O(n^{2/3})$
- $\lambda > 1$: $\mathcal{C}_{max} = O(n\zeta_\lambda)$

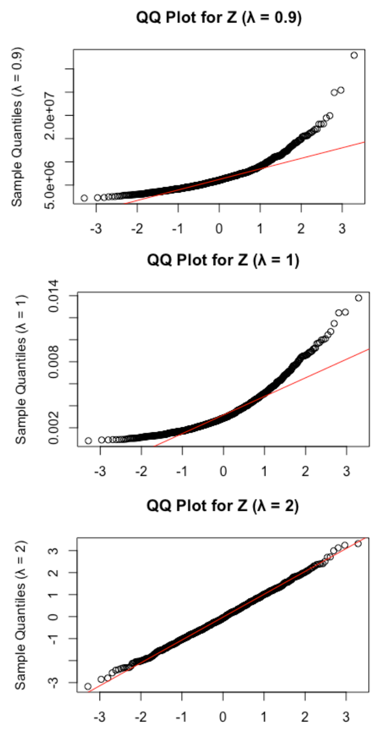
> Central Limit Theorem

- $\frac{\mathcal{C}_{max}-\zeta_\lambda n}{\sqrt{n}} \xrightarrow{d} N(0,\sigma_\lambda^2)$
- $\sigma_\lambda^2 = \frac{\zeta_\lambda(1-\zeta_\lambda)}{(1-\lambda+\lambda\zeta_\lambda)^2}$

When $\lambda < 1$, there are no observable giant components. When $\lambda = 1$, the giant component size is of the order $n^{2/3} \approx 464.1489$. When $\lambda > 1$, the graphs become more and more connected as λ increases and is highly connected when $\lambda = 2$. Additionally, the giant component size in the experiment is 7994 when $\lambda = 2$, close to $\zeta_\lambda n = 7698.12$. This verifies the Law of Large Numbers.

Fix n = 10000 again and for each choice of λ, we simulate 1000 random graphs. Denote by $x_i = \frac{\mathcal{C}_{max}-\zeta_\lambda n}{\sqrt{n}}$ and $z_i = \frac{x_i}{\sqrt{\sigma_\lambda}}$. We hence have 1000 observations of the standardized random variable $Z = \frac{x}{\sqrt{\sigma_\lambda}}$ The Central Limit Theorem states that Z has a limiting normal distribution with mean 0 and variance σ_λ^2 when $\lambda > 1$. The QQ plots of Z under different choices of λ show that no normality is observed when $\lambda = 0.9$ and $\lambda = 1$, while strong normality is observed for $\lambda = 2$. This illustrates the Central Limit Theorem for the giant component size when $\lambda > 1$.

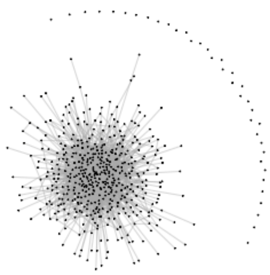
Normality Test for $\mathcal{C}_{max}$



RESULTS

In this section, we consider the subgraph induced by people with Erdős number 1 (512 people as of August 2020). With some effort we derive the degree sequence of the subgraph, which we then feed into a R program to carry out 1000 simulations of the graph. We then study the behavior of the giant component, degree distribution etc.

Collaboration Graph of Mathematicians

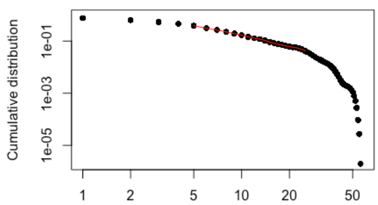


We observe a significant cluster in the subgraph generated by people with Erdős number 1. The average number of people in the giant component is 472.816 and the average cluster coefficient is 0.059, which indicates that the graph is strongly connected and highly clustered.

The average distance between two mathematicians with Erdős number 1 is 7.336 with standard deviation 0.6305. This leads us to believe that the network of mathematicians is indeed a "small-world", where two mathematicians are typically separated by $7 - 1 = 6$ other mathematicians.

We also observe the power-law pattern for the distribution of degree sequence. More specifically, the log-log plot of the complementary cumulative distribution function (CCDF) of the degree sequence reveals that a power-law relationship is likely to be present for degrees between 5 and 25, as the corresponding observations fall on an approximately straight line, while the faster decay thereafter implies an exponential or Gaussian cutoff.

Log-log Plot of CCDF



REFERENCES

Alon, N., & Spencer, J. H. (2004). The probabilistic method. John Wiley & Sons.

Amaral, L. A. N., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. Proceedings of the National Academy of Sciences, 97(21), 11149-11152.

Newman, M. E., Watts, D. J., & Strogatz, S. H. (2002). Random graph models of social networks. Proceedings of the National Academy of Sciences, 99 (suppl 1), 2566-2572.

Pittel, B. (1990). On tree census and the giant component in sparse random graphs. Random Structures & Algorithms, 1(3), 311-342.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. Social networks, 29(2), 173-191.

Van Der Hofstad, R. (2016). Random graphs and complex networks (Vol. 1). Cambridge University Press.