# ER Network and Giant Component

**Chenkai Wang**

**Abstract**

In this essay, we first introduce the ER network. After that, we use simulation and numerical calculation methods to work out $P_S$. Besides that, we plot giant component for Poisson and power law distribution. Finally, we explore the phase transition in ER network briefly.

**Keywords: ER Network, $P_S$, Phase Transition**

SUSTech Southern University of Science and Technology

# Contents

# 1 Part I: ER Network

## 1.1 A Brief History from Seven Bridges of Königsberg to ER Network

The research method of using graphs to represent problems can be traced back to the study of the Seven Bridges of Königsberg by the great mathematician Euler in 1736. Königsberg is a town in East Prussia that is bisected by a river that cuts through the city, with seven bridges connecting the two banks of the river. The question that people are interested in is, can all bridges be traversed under the premise that all bridges are traversed only once? Euler gave sufficient and necessary conditions to be able to traverse and made a negative answer to this question. Euler's research on this problem created an essential branch of mathematics - graph theory.

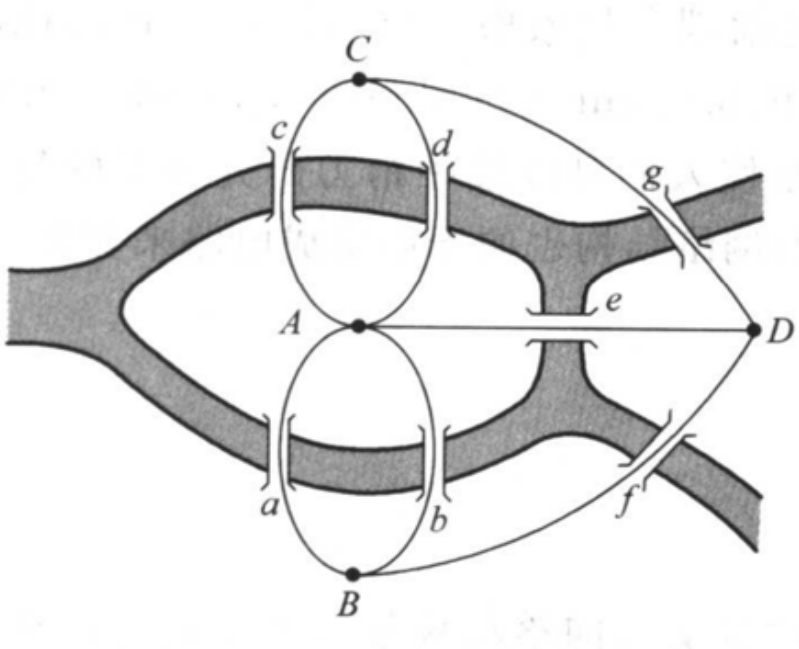


Figure 1: An illustration of Seven Bridges of Königsberg

After solving the Seven Bridges problem, graph theory did not gain rapid development immediately. Graph theory and random networks have come a long way since Paul Erdős and Alfréd Rényi published their famous paper[3] in the 1960s. For a long time after that, the random network model proposed by Erdős and Rényi in [3] was an essential tool for studying real networks. At

the same time, people have also designed social experiments to reveal some characteristics of social networks, among which Milgram's social survey[6] and game of Bacon[1] are widely known.

## 1.2 Basic Structure of ER Network

For ER network, also known as random network, there are two main popular definitions, which can be denoted as $G(n, l)$ model and $G(n, p)$ model, respectively.

- $G(n, l)$ model consists of $n$ nodes connected by $l$ randomly place edges[3];
- $G(n, p)$ model consists of $n$ nodes and edges linking each pair of nodes with probability $p$[4].

Figure 2: An example of ER network with $n = 50$ and $p = 0.6$

This essay will use the $G(n, p)$ model. This is not only because the model is easier to calculate and display some properties of random networks but also because it is rare for the number of links to remain constant for real networks.

ER network is undirected, and the average degree $\langle k \rangle$ is

$$\langle k \rangle = \frac{1}{n} \sum_{i=1}^{n} k_i = p(n-1)$$

Namely,$\langle k \rangle \approx np$ as $n \to \infty$. Therefore, we can use appropriate $n$ and $p$ to generate a random network with desired $\langle k \rangle$.

## 1.3 Comparison of Two Different Calculation Methods of $P_S$

To make it more precise, we decide to generate the ER network from $G(100, \frac{0.5}{99})$.

### 1.3.1 $P_s$ from Simulation

Generate the ER network 1000 times and work out the average $P_s$ as the simulation result, which is shown in section 1.3.3.

The following graph is one of the generated networks during the simulation. To make it easier to distinguish, connected clusters with different sizes have different colors.

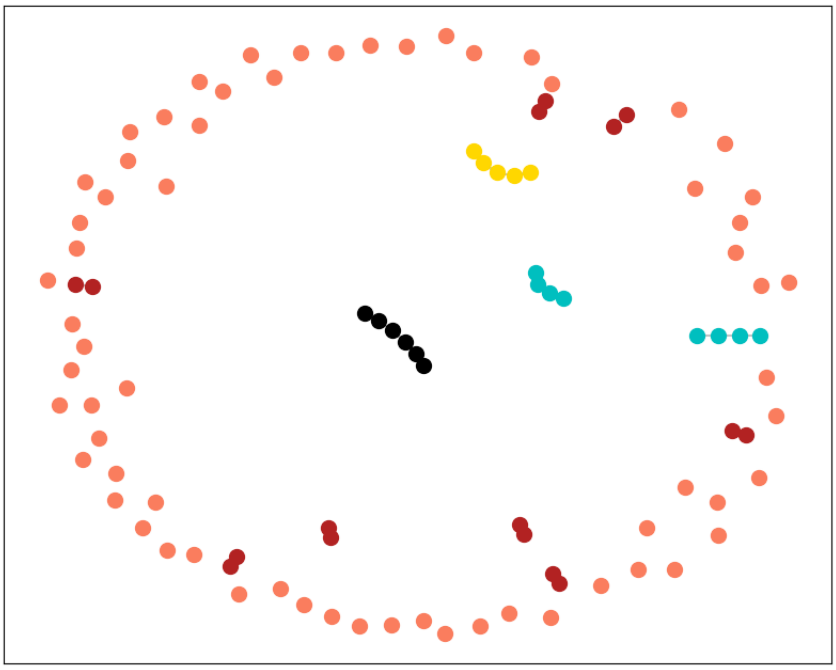

Figure 3: An example of the generated ER network during simulation

### 1.3.2 $P_s$ from Numerical Calculation

According to the slides, $P_s$ can be calculated by the following formula

$$P_s = \frac{1}{2\pi i} \oint \frac{H_0(z)}{z^{s+1}} \, \mathrm{d}z$$

and $H_0(x)$ will be introduced pretty soon.

Set $z = e^{i\theta}$ we have

$$P_s = \frac{1}{2\pi} \int_0^{2\pi} \frac{H_0\left(\mathrm{e}^{i\theta}\right)}{\mathrm{e}^{i\theta s}} \, \mathrm{d}\theta$$

We can divide the interval into $N$ equal sub-intervals to apply the numerical method. Then $P_s$ can be approximated by

$$P_s \approx \frac{1}{2\pi} \sum_{k=0}^{N-1} \frac{H_0\left(\mathrm{e}^{i\frac{2k\pi}{N}}\right)}{\mathrm{e}^{i\frac{2k\pi s}{N}}} \times \frac{2\pi}{N}$$

We first introduce some crucial formulas. Let $G_0$ be the generative function of a network. In our example, we have

$$G_0(x) = \sum_k p_k x_k = (1 - p + px)^{N-1}.$$

Also, according to the slides, we have

$$G_1(x) = \frac{G_0'(x)}{G_0'(1)} = (1 - p + px)^{N-2}.$$

Meanwhile,

$$\begin{cases} H_1(x) = xG_1(H_1(x)) \\ H_0(x) = xG_0(H_1(x)) \end{cases}$$

Use the above Now we can calculate $H_0(\bullet)$ and get simulated $P_s$ eventually. The result is shown in section 1.3.3.

### 1.3.3 The Result and Comparison

| Size $s$ | $P_s$ from simulation | $P_s$ from numerical calculation |
|---|---|---|
| 0 | 0.6020 | 0.6088 |
| 1 | 0.1921 | 0.1853 |
| 2 | 0.0882 | 0.0838 |
| 3 | 0.0460 | 0.0448 |
| 4 | 0.0235 | 0.0262 |
| 5 | 0.0171 | 0.0163 |
| 6 | 0.0102 | 0.0106 |
| 7 | 0.0076 | 0.0071 |
| 8 | 0.0045 | 0.0048 |
| 9 | 0.0035 | 0.0034 |
| 10 | 0.0028 | 0.0024 |
| 11 | 0.0006 | 0.0017 |
| 12 | 0.0013 | 0.0012 |
| 13 | 0.0000 | 0.0009 |
| 14 | 0.0000 | 0.0007 |
| 15 | 0.0000 | 0.0005 |

Table 1: The result and comparison table of two calculation methods

It can be drawn from the table that

- The results from two different methods are consistent;
- The results are closer When size $s$ is small;
- When $s$ is big, the numerical calculation method will give a non-zero result while the simulation will give 0.

# 2 Part II: Giant Component

From the definition, we have

$$S(T) = 1 - H_0(1, T)$$

To calculate $H_0(1,T)$ we need the following formula from slides

$$\begin{cases} G_0(x;T) = G_0(1+(x-1)T) \\ G_1(x;T) = xG_1(1+(x-1)T) \\ H_0(x;T) = xG_0(H_1(x;T);T) \end{cases}$$

Set $x = 1$ and do iteration, we can get $H_0(1,T)$ and therefore $S(T)$ eventually.

## 2.1 Giant Component for Poisson Distribution

The degree distribution of Poisson is

$$p_k = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$

Combined with the formula mentioned at the beginning of section 2. we can plot the giant component size $S$ vs. $T$. Also, from the slides we know for Poisson distribution $T_c = \frac{1}{\langle k \rangle}$ where $\langle k \rangle = G_0'(1)$ represents the average degree. Meanwhile, the plot shows consistency with our theory.

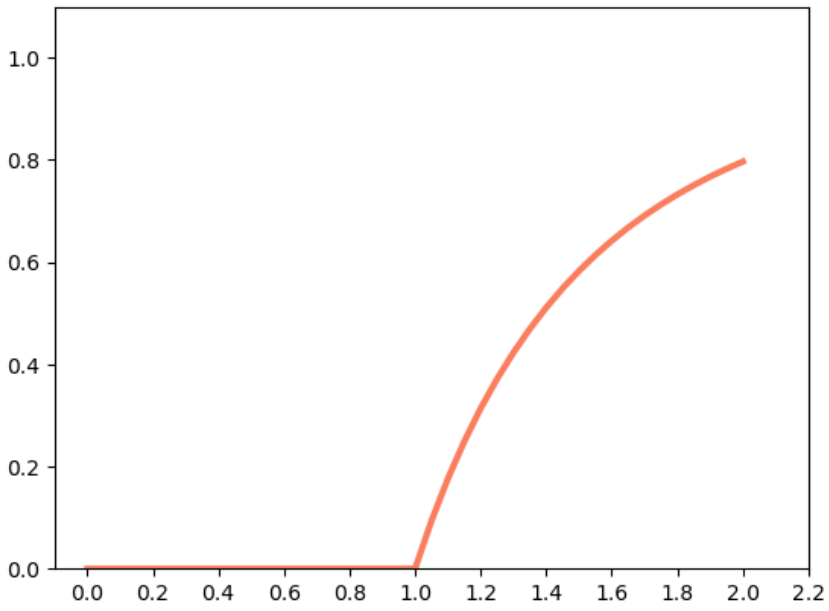


Figure 4: $S$ vs $T$ with k=1

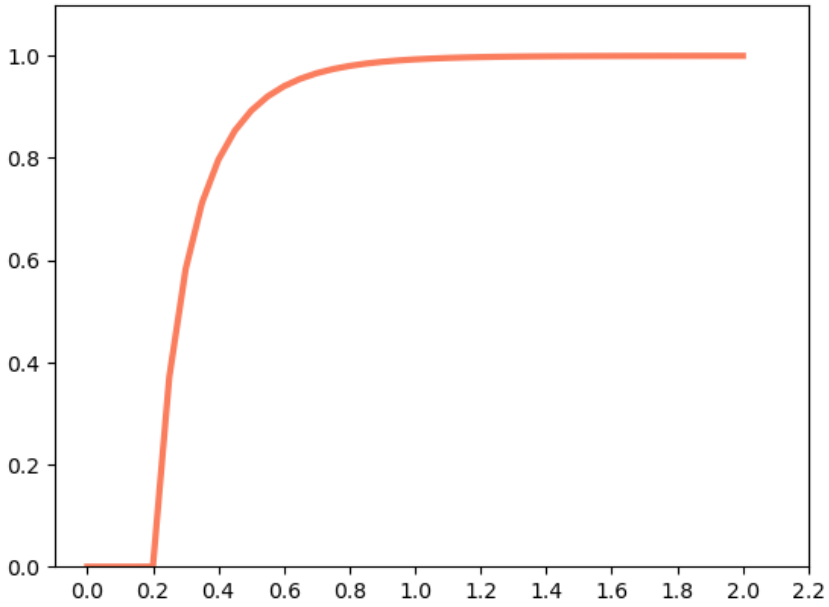


Figure 5: $S$ vs $T$ with k=5

## 2.2 Power Law Degree Distribution

According to the given restriction, we can write the degree distribution $p_k$

$$p_k = \sum_{k=2}^{\sqrt{N}} k^{-2.5}$$

Similarly, following the step in section 2.1, we can get the following graphs. Also from the slides we know $T_c = \frac{\langle k \rangle}{\langle k(k-1) \rangle}$.

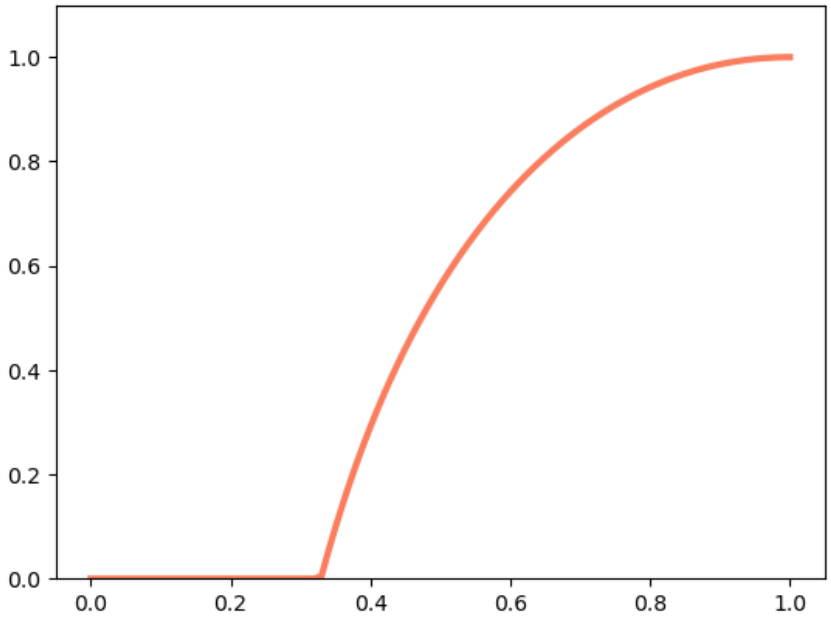


Figure 6: $S$ vs $T$ with N = 100

In this case, we have $\langle k \rangle = 3.09$ and $T_c = 0.329$.

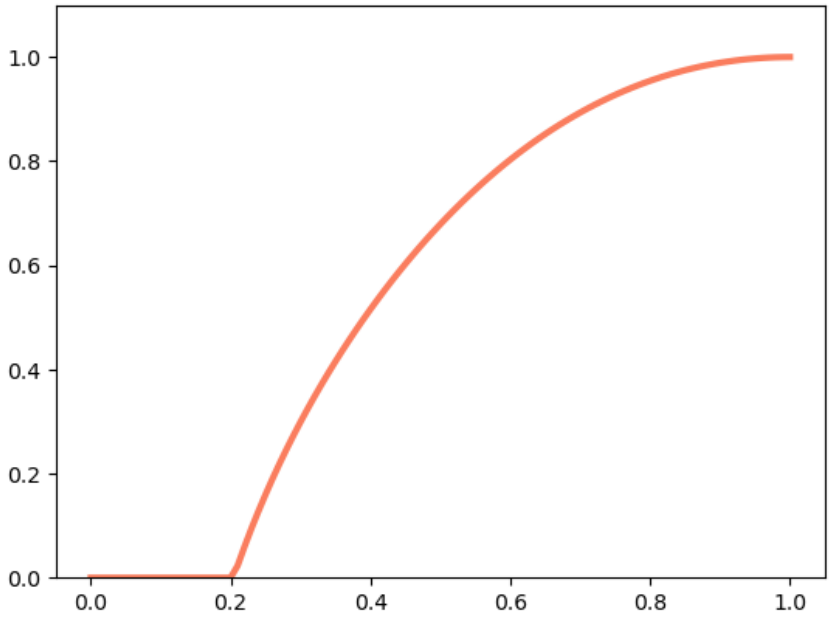


Figure 7: $S$ vs $T$ with N = 500

In this case, we have $\langle k \rangle = 3.55$ and $T_c = 0.204$.

# 3 Part III: Extra Part about ER Network

If we define giant component as the connected clusters with most nodes, no matter whether the network is sparse or not, a problem emerges. When generating figure 2, we find that the size of a giant component does not increase linearly with increasing $p$.

We consulted some textbook[2] and papers[5] and finally found the corresponding explanation: there exists a phase transition in ER network when $p$ is increasing. More specifically[5],

- Let$p = \frac{1-\epsilon}{n}$, where $\epsilon > 0$ is a small enough constant and let $G \sim G(n, p)$. Then **w.h.p.** all connected components of G are of size at most $\frac{7}{\epsilon^2} lnn$.
- Let $p = \frac{1+\epsilon}{n} n$. Then **w.h.p.** $G$ has a connected component with at least $\frac{\epsilon n}{2}$ nodes.

where **w.h.p.** denote with high probability[5].

Meanwhile, we generate an ER network with $n = 150$ and increasing $p$ to illustrate the phase transition.

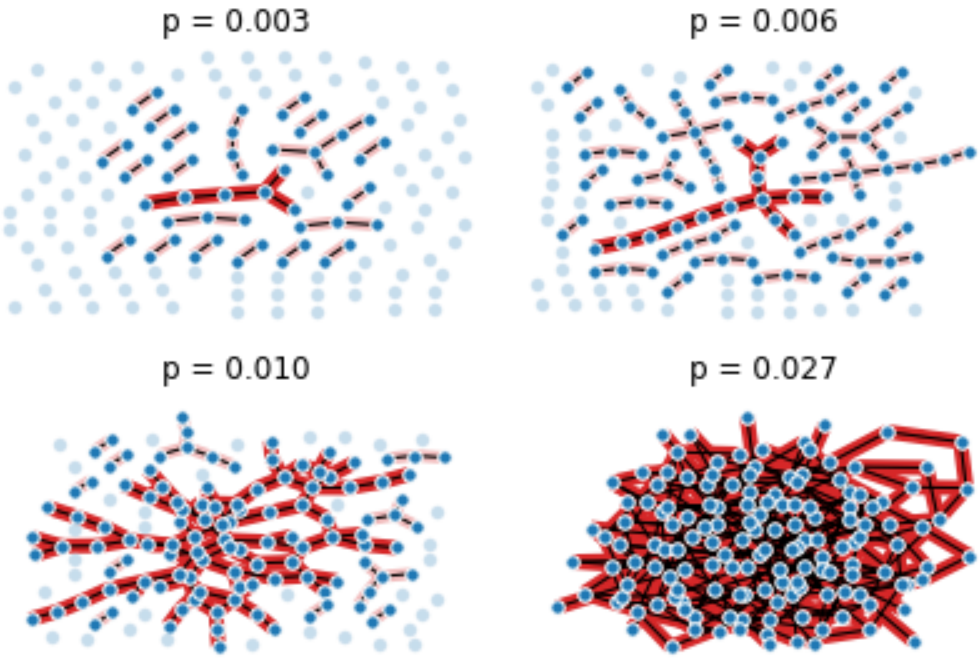


Figure 8: Phase transition in ER network with $n = 150$

# References

[1] https://oracleofbacon.org/.

[2] http://networksciencebook.com/.

[3] Paul Erdős, Alfréd Rényi, et al. "On the evolution of random graphs". In: *Publ. Math. Inst. Hung. Acad. Sci* 5.1 (1960), pp. 17–60.

[4] Edgar N Gilbert. "Random graphs". In: *The Annals of Mathematical Statistics* 30.4 (1959), pp. 1141–1144.

[5] Michael Krivelevich and Benny Sudakov. "The phase transition in random graphs: A simple proof". In: *Random Structures & Algorithms* 43.2 (2013), pp. 131–138.

[6] Stanley Milgram. "The small world problem". In: *Psychology today* 2.1 (1967), pp. 60–67.