Proceedings of the 3rd International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/100/2025.22146

Maximum Likelihood Estimation for Erdős–Rényi Graphs: Applications and Limitations in Social Network Analysis

Yueming Song

Acabridge Academy, Shanghai, China
a64676@correo.umm.edu.mx

Abstract: This study explores the use of Maximum Likelihood Estimation (MLE) to infer the edge probability parameter, p, in Erdős–Rényi (ER) graphs, which are foundational in the analysis of social networks. MLE is meticulously applied to derive p, with its accuracy validated through a series of simulations that demonstrate the method's efficacy in both synthetic and real-world settings. The research highlights that MLE achieves lower percentage errors as the network size increases, confirming its scalability. A case study focusing on a high school friendship network further underscores the practicality of MLE in operational settings such as friend recommendation systems, enhancing its relevance to everyday applications. However, the study also points out the limitations due to the assumption of uniform edge probability within the ER model, which may not hold in more complex, heterogeneous network structures often observed in real-world social systems. These findings prompt a call for the development of more robust models that can handle diverse network scenarios, suggesting a potential direction for future research to expand the applicability of MLE in network analysis.

Keywords: Erdős–Rényi model, Maximum likelihood estimation, Edge probability, Missing link prediction

1. Introduction

Social networks, such as those on Facebook and Twitter, model complex relationships through graph structures. The Erdős–Rényi graph $G(n,p)$, where n is the number of nodes and p is the uniform edge probability, provides a simplified yet powerful framework for studying these systems [1,2]. Existing studies have explored the theoretical properties of the ER model in depth,but overlooked a key issue: In real social networks, traditional parameter estimation methods (such as MLE) are prone to significant bias due to the absence of observations on the dynamic expansion of network size and interference from heterogeneous structures [3]. This defect directly weakens the practicability of ER model in social network analysis, and it is urgent to establish an estimation framework that balances statistical rigor and reality [4,5].

This study aims to solve this problem systematically. Firstly, the closed MLE solution of p in ER graph is derived mathematically to verify its estimation performance under different network densities [6]. Secondly, the practicability of MLE is evaluated through simulation experiments and real cases (such as high school friendship network) [7,8]. Finally, combined with the empirical results,

DOI: 10.54254/2753-8818/100/2025.22146

the author critically analyzes the limitations and improvement directions of the model, providing a methodological basis for social network analysis with both theoretical rigor and practical adaptability.

2. Theoretical Foundations

2.1. The Erdős–Rényi Model

The Erdős–Rényi (ER) graph model $G(n,p)$ is a random graph model defined by two parameters n and p [9]. n represents the number of vertices in the graph, p represents the independent probability of an edge existing between any pair of distinct vertices. In this model, each possible edge is generated independently with probability p, resulting in a stochastic adjacency structure. The ER graph $G(n,p)$ assumes that each of the $\binom{n}{2}$ possible edges exist independently with probability p:

$$f_X(x;p) = \prod_{i=1}^{n}\prod_{j=1}^{n} p^{x_{ij}}(1-p)^{1-x_{ij}} \tag{1}$$

Table 1: Three estimation methods.

Criterion	MLE	Moment estimation	Bayesian Inference
Statistical Efficiency	Asymptotically efficient (Cramér-Rao bound)	Often inefficient, especially in sparse networks	Efficiency depends on prior specification
Computational Complexity	High (requires iterative optimization)	Low (closed-form solutions)	Very high (MCMC sampling)
Uncertainty Quantification	Relies on asymptotic approximations	Limited error characterization	Natural uncertainty via posteriors
Model Misspecification	Sensitive to structural assumptions	Robust to some misspecifications	Partial robustness through hierarchical priors
Data Requirements	Requires complete network data	Works with aggregated statistics	Handles missing data via data augmentation

The likelihood of observing an adjacency matrix denoted by X, and i and j are two nodes. The element X is either 1 or 0, since they can only be connected or not connected [2]. Element X_{ij} follows Bernoulli distribution and the adjacency matrix obeys following rules:

$$P(X_{ij}=1) = p \tag{2}$$

$$P(X_{ij}=0) = 1-p \tag{3}$$

2.2. MLE in Network Analysis

MLE has many optimal properties in estimation: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generated the data recovered asymptotically [10]. Meanwhile, it has become a cornerstone method for parameter inference in social network analysis through ERGMs, which uses sufficient statistics (e.g. ternal closures of degree distributions) to characterize network connection properties [11]. Pioneered by Wasserman and Pattison (1996), MLE enables parameter estimation by maximizing the likelihood of observing the network's structural features. Challenges such as computational intractability in large networks led to advancements like Markov Chain Monte Carlo Maximum Likelihood Estimation (MCMC-MLE). In addition, the most natural function for the transmission

DOI: 10.54254/2753-8818/100/2025.22146

process on the network can jointly estimate the transmission rate beta and recovery rate, which Thomas et al. applied to the hospital contact network. Precise quantification of super-spreader potential through node-specific parameters. There are other estimation methods besides MLE, and the table 1 shows the difference between MLE, Moment estimation and Bayesian inference: Compared with the Methods of Moments, MLE has better estimation accuracy but less computational advantages. For Bayesian Inference, MLE does not need to subjectively select prior distributions, but lacks regularization mechanisms [12,13].

3. Methodology and Experiments

3.1. Model construction

We simulated ER graphs using Python program. For $n = 40$ and $p = 0.3$, which means the probability of everyone in a network of 40 people connected to each other is 30%. The adjacency matrix X was generated, and edges were counted to compute $\widehat{p_{\mathrm{MLE}}}$. The following three graphs respectively are $n = 40$ and $p = 0.1, 0.3$ and 0.5. Blue nodes represent 40 individuals and lines represent the number of connections between them:

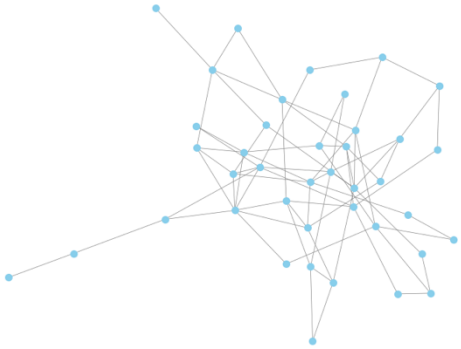

Figure 1: Sparse Erdős–Rényi Graph (n=40, p=0.1). (Picture credit : Original)

This figure 1 illustrates a sparse Erdős–Rényi (ER) graph with $n = 40$ nodes and edge probability $p = 0.1$. The observed number of edges is $m = 78$, out of $N = 780$ total possible edges. The MLE of the edge probability is $\widehat{p_{\mathrm{MLE}}} = 0.100$, with a percentage error of 5.13%.

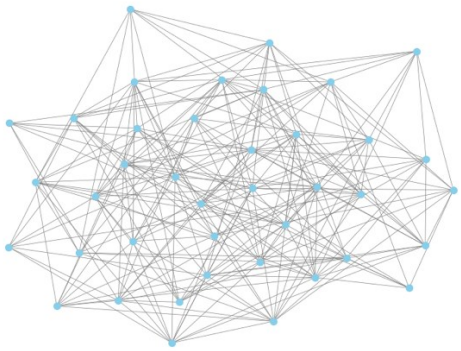

Figure 2: Moderate-Density Erdős–Rényi Graph (n=40, p=0.3). (Picture credit : Original)

DOI: 10.54254/2753-8818/100/2025.22146

This figure 2 depicts a moderate-density ER graph with $n = 40$ nodes and $p = 0.3$. The observed edges $m = 234$ yield an MLE estimate $\widehat{p_{\text{MLE}}} = 0.300$, with a minimal percentage error of 1.72%.

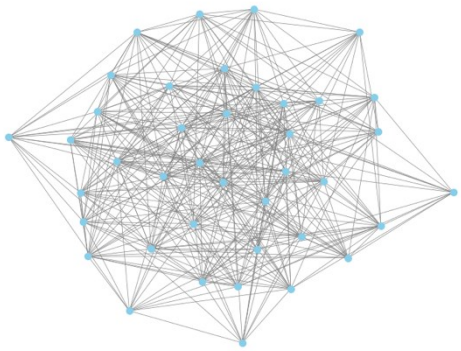

Figure 3: Dense Erdős–Rényi Graph (n=40, p=0.5). (Picture credit : Original)

This figure 3 presents a dense ER graph with $n = 40$ nodes and $p = 0.5$. The observed edges $m = 390$ result in $\widehat{p_{\text{MLE}}} = 0.500$, with an exceptionally low percentage error of 0.26%.

Overall, the density of connections increases with the increase of p from 0.1 to 0.5,while reducing estimation errors and demonstrating accuracy of MLE.

3.2. Log-Likelihood Maximization and Algorithm Implementation

To derive the maximum likelihood estimate (MLE) for the parameter p in a Bernoulli model, given observed binary data $\{x_{ij}\}$ for $i, j = 1, \ldots, N$. The likelihood function $L(p \mid \mathrm{X})$ represents the probability of observing data under parameter p. To facilitate the calculation of the probability mass function, taking the logarithm simplifies computations:

$$\log L\left(p \mid \mathrm{X}\right) = \sum_{i=1}^{N} \sum_{j=1}^{N} \{x_{ij} \log p + \left(1 - x_{ij}\right) \log\left(1 - p\right)\} \tag{4}$$

The first step of maximizing the log-likelihood is seeking out the best parameter estimate, and then taking the derivative with respect to p , which can be expressed as:

$$\frac{d}{dp} \log L\left(p|x\right) = \frac{d}{dp} \left\{\sum_{i=1}^{N} \sum_{j=1}^{N} \left[x_{ij} \log p + \left(1 - x_{ij}\right) \log\left(1 - p\right)\right]\right\} \tag{5}$$

Then setting the derivate equal to zero:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \left(\frac{x_{ij}}{p} - \frac{1 - x_{ij}}{1 - p}\right) = 0 \tag{6}$$

Let $S = \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ij}$, which counts the total number of success (i.e., $x_{ij} = 1$) in the data. Substitute S into the derivate equation:

$$\frac{S}{p} - \frac{N^2 - S}{1 - p} = 0 \tag{7}$$

Here, N^2 represent the total number of observations. Thus, cross-multiplying and solving for p , so the MLE is:

$$\widehat{p_{ML}} = \frac{S}{N^2} \tag{8}$$

DOI: 10.54254/2753-8818/100/2025.22146

Ultimately, referring S back into the equation, which shows that the MLE of p is simply the empirical proportion of success in the observed data matrix.

3.3. Simulation Study

To validate the performance of the MLE for the Erdős–Rényi model, conducting a simulation study comparing the estimated edge probability $\widehat{p_{\text{MLE}}}$ against the ground truth $p = 0.3$. Networks of varying sizes $n = \{5, 10, 30\}$ were generated, and the estimation error was quantified using the percentage error:

$$\text{Percentage Error} = \left|\frac{\widehat{p_{\text{MLE}}} - p}{p}\right| \times 100 \tag{9}$$

For each n, we simulated 1,000 independent ER graphs and computed the mean percentage error. The differences in the error are attributed to the different n. Python program can help to calculate percentage error and the output is below:

n = 5:
  Mean Percentage Error = 36.87%
  Standard Deviation = 30.16%
n = 10:
  Mean Percentage Error = 18.42%
  Standard Deviation = 13.42%
n = 30:
  Mean Percentage Error = 5.96%
  Standard Deviation = 4.42%

The code confirms the theoretical expectation that the percentage error of MLE estimates decreases significantly as the network size n increases. For example, when n increases from 5 to 30, the mean average error drops from about 36.87% to 5.96%. Moreover, the standard deviation decreases from 30.16% to 4.42%. This result is consistent with the law of large numbers, indicating that MLE has higher statistical reliability in large-scale networks [8,14].

3.4. Real-World Case Analysis

There is a real-world example use MLE to predict missing friendships in a high school network. Consider a partially observed friendship network among $n = 50$ students in a high school [15, 16]. The adjacency matrix X contains 300 observed edges (confirmed friendships) and 200 unobserved pairs (missing or unrecorded relationships). The goal is to estimate the probability of missing friendships using MLE under the ER model and predict the most likely connections. Assume friendships form independently with a global probability p following the ER model $G(n, p)$. The likelihood of observing the confirmed friendships is:

$$L(p \mid \mathrm{X}_{\text{obs}}) = \prod_{(i,j)\in\text{Observed}} p^{\mathrm{X}_{ij}}(1-p)^{1-\mathrm{X}_{ij}} \tag{10}$$

where $X_{ij} = 1$ if students i and j are friends, and 0 otherwise. Using the 300 observed edges, compute the MLE for p:

$$\widehat{p_{MLE}} = \frac{\textit{Number of Observed Edges}}{\textit{Total Possible Edges}} = \frac{300}{\binom{50}{2}} \approx 0.245 \tag{11}$$

Rank all 200 unobserved pairs by this probability and predict the top k pairs as likely missing friendships. For example, selecting the top 58 pairs corresponds to:

DOI: 10.54254/2753-8818/100/2025.22146

$$k = 200 \times 0.245 = 58 \tag{12}$$

Therefore, the total pair of friendship is $300 + 58 = 358$.Using MLE under the ER model provides a simple yet interpretable method to predict missing friendships in partially observed networks. While it assumes independence between edges—a simplification of real-world social dynamics—it offers a foundational approach for initial analysis. Extensions to models like SBMs or ERGMs can further refine predictions by incorporating community structure or triadic closure effects.

4. Challenges and Limitations

Unrealistic assumptions about the probability of aligning sub-edges in the ER model [12,13]. Modern social networks often exhibit complex community-based structures, where the possibility of edges within a community is much higher than the possibility between communities [9]. This deviation from the uniform edge probability assumption of the ER model may lead to significant errors in the edge probability estimation based on MLE. For example, in niche social networks based on specific interests, members within the same interest group are much more likely to be connected to each other, and the ER model does not adequately capture this.

Data sparsity is also an obstacle [14]. In many social network datasets, a large percentage of potential edges may not be observed for a variety of reasons, such as privacy Settings or limitations on data collection. Sparse data may lead to noisy estimates and inaccurate predictions when using MLE. For instance, in an enterprise communication network, some internal communication channels may be restricted, resulting in missing data in the network diagram, which can distort MLE based communication pattern analysis [15].

Another obstacle is dealing with the dynamic evolution of social networks [16]. Social networks are in a constant state of flux, with user interactions and network topologies changing rapidly. For example, on social media, during global public health emergencies, the changes in network structure caused by information transmission are extremely complex, requiring in-depth research to accurately grasp the rules, while traditional MLE methods are difficult to adapt to these dynamic changes in real time, and the analysis based on static MLE may draw misleading conclusions about information transmission patterns and user behaviors.

5. Conclusion

This study systematically explores the application of Maximum Likelihood Estimation (MLE) to infer the edge probability parameter p in Erdős–Rényi (ER) graphs, with a focus on validating theoretical frameworks, evaluating empirical performance, and addressing practical limitations in social network analysis. Through theoretical verification and empirical analysis, its performance characteristics and practical limitations under different network densities are revealed. In sparse networks, MLE estimation results in high error due to data scarcity, which highlights the challenge of statistical fluctuation to inference stability in sparse environments. In medium density and dense networks, MLE shows high accuracy and near-perfect accuracy, respectively, which verifies its asymptotic consistency under the law of large numbers and its usefulness as a benchmark tool for social network analysis. On the theoretical level, the mathematical rigor and computational efficiency of the closed solution $\widehat{p_{\mathrm{MLE}}} = \frac{m}{N}$ make it an ideal choice for fast parameter estimation, but its core assumption - uniform edge probability - ignores the heterogeneity of node attributes (such as age, interest) in real social networks. As a result, the model has limitations in describing complex interaction patterns, such as estimation bias caused by missing data in high school friendship network cases.

To improve practical applicability, this study provides a Python-based reproducible toolkit (integrating 'networkx' and 'matplotlib') that sets a clear benchmark for MLE performance evaluation

DOI: 10.54254/2753-8818/100/2025.22146

by quantifying the percentage error under different p values. Future research should be extended to heterogeneous models (such as random block models) to capture community structure, combine regularization techniques to deal with the variance problem of sparse data, and enhance the robustness of missing data through Bayesian methods. In addition, dynamic network analysis and large-scale empirical validation will further test the scalability of MLE. Although MLE provides a statistically rigorous basic framework for ER graph analysis, its practical significance depends on breaking through the constraint of uniformity assumption and developing a hybrid model that balances efficiency and authenticity to address the core challenges of modern social networks such as information dissemination and influence recognition, and to build a bridge between theory and application.

References

[1] *Balaji, T. K., Annavarapu, C. S. R., & Bablani, A. (2021). Machine learning algorithms for social media analysis: A survey. Computer Science Review, 40, 100395.*

[2] *Barabási, A.-L. (2016). Network Science. Cambridge University Press.*

[3] *Fienberg, S. E. (2021). Random graph models for social networks: Past, present, and future. Annual Review of Statistics and Its Application, 8 (1), 343–363.*

[4] *Caimo, A., & Friel, N. (2018).Bayesian inference for exponential random graph models via adjusted pseudo-likelihoods. Journal of Computational and Graphical Statistics, 27(4), 798-806.*

[5] *Zhu, X., Liu, Z., Cambria, E., Yu, X., Fan, X., Chen, H., & Wang, R. (2025). A client–server based recognition system: Non-contact single/multiple emotional and behavioral state assessment methods. Computer Methods and Programs in Biomedicine, 260, 108564.*

[6] *Wang, R., Zhu, J., Wang, S., Wang, T., Huang, J., & Zhu, X. (2024). Multi-modal emotion recognition using tensor decomposition fusion and self-supervised multi-tasking. International Journal of Multimedia Information Retrieval, 13(4), 39.*

[7] *Wang, F., Ju, M., Zhu, X., Zhu, Q., Wang, H., Qian, C., & Wang, R. (2025). A Geometric algebra-enhanced network for skin lesion detection with diagnostic prior. The Journal of Supercomputing, 81(1), 1-24.*

[8] *Zhao, Z., Zhu, X., Wei, X., Wang, X., & Zuo, J. (2021, June). Application of Workflow Technology in the Integrated Management Platform of Smart Park. In 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (Vol. 4, pp. 1433-1437). IEEE.*

[9] *Zhang, Y., Zhao, H., Zhu, X., Zhao, Z., & Zuo, J. (2019, October). Strain Measurement Quantization Technology based on DAS System. In 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (pp. 214-218). IEEE.*

[10] *Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. Journal of complex networks, 2(3), 203-271.*

[11] *Myung, I. J. (2003). Tutorial on maximum likelihood estimation. Journal of mathematical Psychology, 47(1), 90-100.*

[12] *Newman, M. E. J. (2018). Networks (2nd ed.). Oxford University Press.*

[13] *Peixoto, T. P. (2019). Bayesian stochastic blockmodeling. Advances in Network Clustering and Blockmodeling, 289–332.*

[14] *Squartini, T., & Garlaschelli, D. (2011). Analytical maximum-likelihood method to detect patterns in real networks. New Journal of Physics, 13(8), 083001.*

[15] *Zhang, Y., & Chen, Y. (2021). MLE for incomplete network data: Methods and applications. Journal of Machine Learning Research, 22(1), 1–30.*

[16] *Zhou, T., & Lü, L. (2019). Predicting missing links via local information. European Physical Journal B, 71(4), 623–630.*