Deep Graph Generation for Microcanonical Graph Ensemble

graph generation, microcanonical graph ensemble, deep reinforcement leaning, graph neural network, data augmentation

Extended Abstract

Under the maximum-entropy principle, Exponential Random Graph Models (ERGM) [1] have been proposed to generate canonical graph ensembles following the distribution $p_\theta(G) \propto \exp(-\theta s(G))$. Here, sufficient statistics $s(G)$ encode desired structural features (e.g., assortativity or global clustering coefficient), and the control parameter θ tunes their expected values. In practice, graphs from such ensembles are typically sampled via Markov Chain Monte Carlo (MCMC) methods. However, determining the value of θ that yields a target $s(G)$ is generally nontrivial, requiring a full parameter sweep. This becomes computationally prohibitive when multiple constraints are imposed, as the search space grows exponentially.

Moreover, ERGMs impose *soft constraints* on $s(G)$, controlling only its expectation but not its variance. Consequently, the sampled graphs may exhibit a wide distribution of $s(G)$, leading to biased statistical analyses, especially in finite-size ensembles or nonequivalent ensemble systems [2, 3].

In this work, we propose an alternative framework that constructs graph ensembles by iteratively rewiring edges starting from an Erdős–Rényi network, guided by a graph neural network (GNN) agent. The agent observes the current graph structure and selects edge pairs to rewire, gradually transforming the network toward satisfying desired structural properties. The agent is trained using Proximal Policy Optimization (PPO) [4], enabling it to efficiently navigate the combinatorial search space.

Our approach directly enforces *hard constraints* on $s(G)$, enabling the generation of microcanonical graph ensembles without the need to sweep over θ-space. This property makes the method scalable to multiple constraints and allows statistically rigorous analysis of networked systems. Furthermore, the method facilitates the generation of synthetic graphs that match selected structural properties of rare real-world networks (e.g., Internet, power grid), offering promising applications in data augmentation for machine learning.

Ethical Considerations

The proposed methodology does not involve human subjects, sensitive personal data, or potentially harmful deployment scenarios. Generated graph ensembles are synthetic and used solely for scientific and methodological purposes.

References

[1] Juyong Park and M. E. J. Newman. "Statistical mechanics of networks". In: *Phys. Rev. E* 70 (6 2004), p. 066117. URL: https://link.aps.org/doi/10.1103/PhysRevE.70.066117.

[2] Tiziano Squartini et al. “Breaking of Ensemble Equivalence in Networks”. In: *Phys. Rev. Lett.* 115 (26 2015), p. 268701. URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.268701.

[3] Frank Den Hollander and Maarten Markering. “Breaking of ensemble equivalence for dense random graphs under a single constraint”. In: *Journal of Applied Probability* 60.4 (2023), pp. 1181–1200.

[4] John Schulman et al. *Proximal Policy Optimization Algorithms*. 2017. URL: https://arxiv.org/abs/1707.06347.

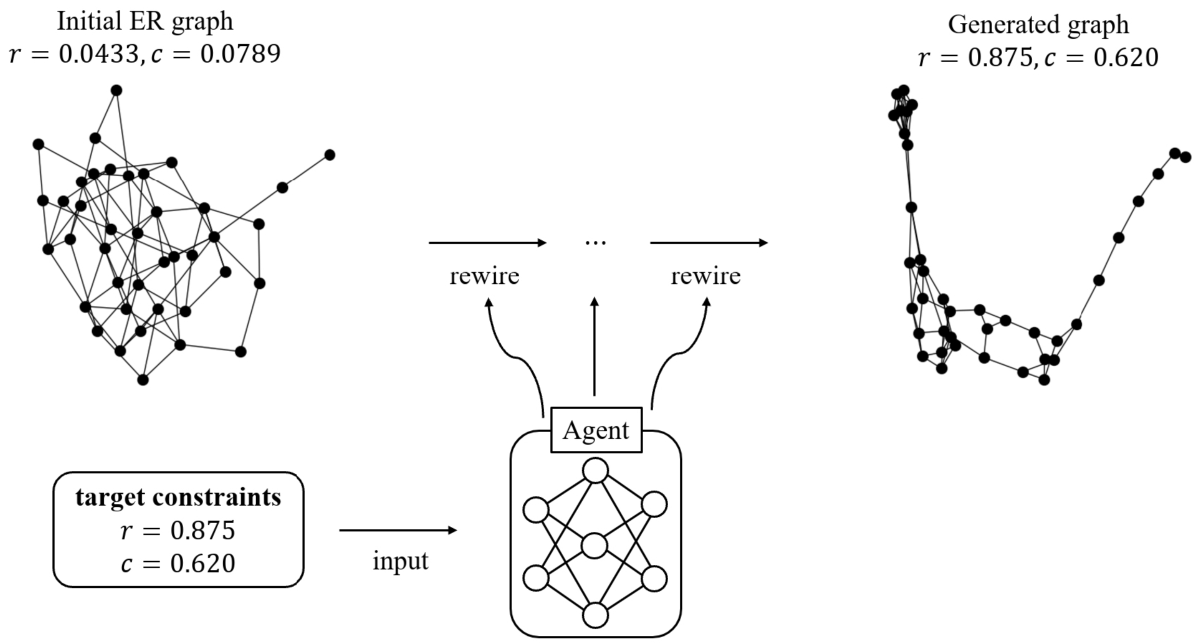

Figure 1: **Deep graph generation** Starting from the initial ER random network, the graph neural network agent gradually rewires the network to satisfy desired structural properties.