Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

PENM: A PARAMETRIC EVOLUTIONARY NETWORK MODEL FOR SCHOLAR COLLABORATION NETWORK SIMULATION

Jiadong Liu[1], Keqin Guan[1], Siyu Chen[1], Wei Chen[1,2], and Wai Kin (Victor) Chan[1]

[1]Institute of Data and Information, Tsinghua Shenzhen International Graduate School (SIGS), Tsinghua University, Shenzhen, P.R. CHINA.
[2]Tencent Technology (Shenzhen) Co. LTD., Shenzhen, P.R. CHINA.

ABSTRACT

Identifying suitable collaborators has become an important challenge in research management, where insights into the structure and evolution of scholar collaboration networks are essential. However, existing studies often adopt static or locally dynamic views, limiting their ability to capture long-term network evolution. To address these issues, this paper introduces PENM (Parametric Evolutionary Network Model), a simulation framework designed to model the evolution of scholar collaboration networks through parametric mechanisms. The PENM simulates node and edge evolution through probabilistic rules and tunable parameters, reflecting realistic academic behaviors like cumulative collaboration and co-author expansion. We provide a theoretical analysis of the model's growth patterns under varying parameters and verify these findings through simulation. Evaluations on real-world datasets demonstrate that PENM evolves networks with degree distributions closely aligned with actual scholar networks. PENM offers a versatile simulation-based approach for modeling academic collaboration dynamics, enabling applications and simulation of future academic ecosystems.

1 INTRODUCTION

In modern scientific research, the ability to identify and recommend appropriate collaborators is becoming increasingly critical, particularly in the context of high-impact projects. Central to this process is the scholar collaboration network, which simulates a dynamic structure of academic cooperation. Despite extensive research on scholar collaboration networks, most studies rely on static representations, limiting simulations to local dynamics. Consequently, the broader evolutionary nature of academic interactions has received limited attention.

To address this gap, we propose the Parametric Evolutionary Network Model, which simulates scholar collaboration as a dynamic process driven by time-dependent interaction rules. Unlike traditional models, PENM leverages simulation to model the evolutionary growth and transformation of collaboration networks. This approach not only reflects the real-world dynamics of academic partnerships but also provides predictive capability for future network structures.

Our contributions are twofold. First, we develop a time-stepped, simulation-based model to generate evolving collaboration networks within a finite time horizon. Second, we validate the model through simulation-driven analysis and real-world collaboration datasets. The results demonstrate that PENM can evolve networks whose degree distributions closely match those of actual scholar collaboration networks, thus supporting its applicability in practical contexts such as network forecasting and academic recommendation.

The remainder of this paper is structured as follows. Section 2 reviews related work. Section 3 presents the formulation of the PENM model, followed by theoretical analysis and simulation experiments. Section 4 describes validation on real-world datasets, and Section 5 concludes the paper.

2 RELATED WORK

Understanding the evolution of scholar collaboration networks is a long-standing problem at the intersection of social network analysis and generative modeling. There are some research on empirical analyses of scholar collaboration networks. Early studies primarily focused on characterizing the structural properties of co-authorship networks. Franceschet (2011) analyzed large-scale computer science collaboration networks and reported properties such as small-world structure and clustering. Kong, Shi, Wang, Ma, Wan, and Xia (2019) explored the evolution of elite research collaborations over decades, highlighting long-term trends and densification patterns. More recent studies, such as Xiao and Prompanyo (2023) and Chuang and Chen (2022), investigated how local collaboration structures relate to academic influence and international cooperation. While these studies provide valuable descriptive insights, they typically adopt static or cross-sectional network observations of networks. As a result, they fall short of modeling how collaboration networks dynamically evolve over time.

With the rise of machine learning, predictive models for collaborator recommendation have gained prominence. Embedding-based approaches like ACNE and Scholar2vec (Wang et al. 2020) integrate node attributes and collaboration histories to support relationship mining. Reinforcement learning and topic-aware embeddings have also been used (Zhang et al. 2017; Kong et al. 2018) to recommend collaborators based on evolving research interests. Despite their predictive power, these models primarily aim at individual-level recommendations and often operate as black boxes. They lack the global, interpretable structure modeling needed to understand macro-level network dynamics, which is a central focus of our PENM framework.

In parallel, modeling how networks grow and evolve is a well-studied topic in complex systems. The Barabási-Albert model (Barabási and Albert 1999) introduced preferential attachment to explain scale-free networks, inspiring many follow-ups. For example, Zhao, Lai, Park, and Ye (2005) and Simas and Rocha (2008) extended such models to handle finite-size effects and stochastic processes. Others incorporated small-world or community structures (Pasta et al. 2013; Gyarmati and Trinh 2010), or focused on improving generative efficiency (Hadian et al. 2016). More recently, domain-specific extensions have emerged, such as ECNA for epidemiological contact networks (Eden et al. 2021) and evolutionary level design in games (Ruela and Delgado 2018). However, these models often lack mechanisms tailored to academic collaboration dynamics, such as cumulative co-authorship or triadic closure tendencies.

While existing studies have enriched our understanding of academic collaboration networks through empirical analysis, collaborator recommendation, and generative modeling, there remains a gap in interpretable models that explicitly simulate the dynamic evolution of such networks. Descriptive analyses often lack simulation-based predictive power, recommendation models prioritize individual-level accuracy over global network simulation, and general-purpose generative models rarely simulate the behavioral mechanisms specific to scholarly collaboration. To address these limitations, we introduce the Parametric Evolutionary Network Model, a simulation-driven framework that captures node and edge dynamics via probabilistic, parameter-tuned rules. PENM simulates realistic features like cumulative co-authorship and triadic closure, and its adaptable structure enables alignment with observed network properties, making it a robust tool for simulating and forecasting academic collaboration network evolution.

3 PARAMETRIC EVOLUTIONARY NETWORK MODEL

In our study, we represent scholars as nodes and model their collaborations as edges between these nodes. Scholars with longer tenures in a research field tend to form more collaborations. Moreover, they often establish connections not only with new researchers but also with collaborators of their existing co-authors, forming triadic closures. Our model PENM aims to simulate the temporal growth and structural dynamics of scholar collaboration networks. Table 1 summarizes all notations used in the PENM model, including variables and parameters that govern the network's evolution.

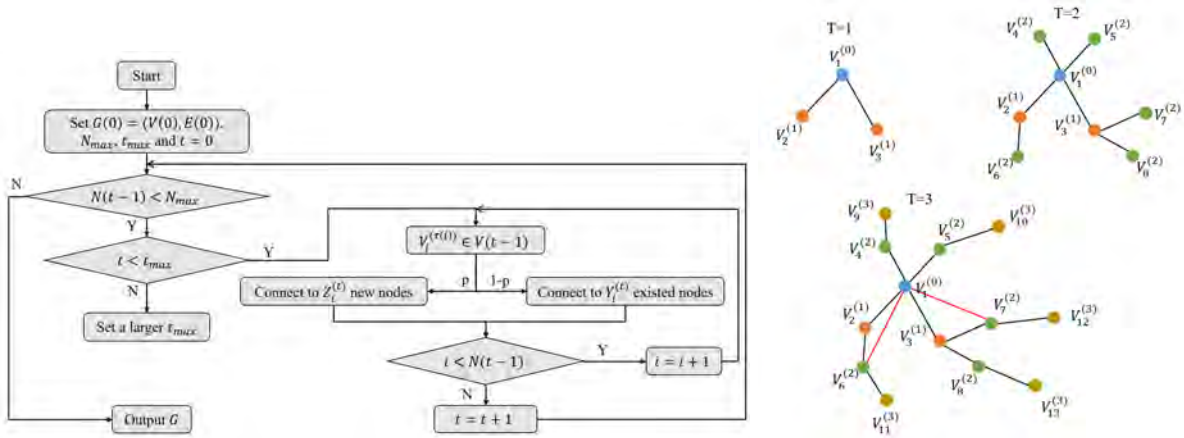
The model begins with an initial network at $t = 0$. From $t = 1$, the network evolves by iterating through each existing node. At each time step, node $i \in V(t-1)$ undergoes two probabilistic processes:

Table 1: All of the notations in the model.

Variables	
t	The term t is a time step. It does not represent a real time interval, but rather indicates a time unit in the process of network evolution.
t_{max}	The upper bound of time step. Set as 50 in this paper.
$\tau(i)$	The first time node i added to the network at time t.
$G(t)$	The growth network at time t. $G(t) = (V(t), E(t))$.
$V(t)$	The node set of growth network at time t.
$V_i^{(\tau(i))}$	The i-th node in $V(t)$, first added at time $\tau(i)$.
$N(t)$	The total number of nodes in the network at time t.
N_{max}	The upper bound for number of nodes.
$Z_i^{(t)}$	The number of new nodes added to node i at time t.
$E(t)$	The edge set of growth network at time t.
$M(t)$	The total number of edges in the network at time t.
$Y_i^{(t)}$	The number of existed nodes added to node i.
$neb(i)$	The neighbor of node i in the growth network.
$d_i^{(t)}$	The degree of node i in the growth network at time t.
$d_{max}^{(t)}$	The maximum degree in the growth network at time t.
$d_{avg}^{(t)}$	The average degree in the growth network at time t.
Parameters	
p	The probability governs the overall balance between node growth and edge formation.
$degree_threshold$	The minimum degree condition for nodes to initiate connections with existing nodes.
α	The coefficient controls the rate at which a node attracts new nodes.
β	The coefficient affects the likelihood of forming new edges with 2 step neighbors.
$\lambda_i^{(t)}$	A Poisson parameter. $Z_i^{(t)}$ obeys Poisson distribution with parameter $\lambda_i^{(t)}$.
$\mu_{ij}^{(t)}$	The probability that node i connects to a neighbor of its j-th neighbor at time t.

- **Node growth (with probability p)**: Node i connects to a set of newly introduced nodes. The number of such connections follows a Poisson distribution with parameter $\lambda_i^{(t)} = (1 - \frac{\tau(i)}{t})\alpha$, where $\tau(i)$ is the time node i first appeared. This reflects the intuition that nodes with longer tenure in the network attract more newcomers.
- **Edge growth (with probability $1-p$)**: Node i forms new edges with nodes that are two steps away — neighbors of their neighbors. The probability of connecting to a second-degree neighbor j is given by: $\mu_{ij}^{(t)} = \frac{d_i^{(t-1)} \cdot d_j^{(t-1)}}{d_{max}^{(t-1)} \cdot d_{max}^{(t-1)}}\beta$. This mechanism encourages triadic closure, where high-degree nodes are more likely to facilitate indirect connections.

This generative process emulates the growth patterns commonly observed in academic collaboration networks. The network starts with a single node representing an initial scholar. As time progresses, new researchers enter the field and initiate collaborations—a process modeled through the node growth mechanism. The Poisson distribution captures the stochastic nature of incoming collaborations. The parameter $\lambda_i^{(t)}$ is controlled by α, which determines how quickly a node increases with time steps. The edge growth process captures indirect collaborations through existing connections, governed by the parameter β. Together, p, α, β and $degree_threshold$ control the balance between exploratory (new nodes) and consolidative (existing network) collaboration dynamics. The simulation algorithm of network evolution and an example are illustrated in Figure 1.



(a) (b)

Figure 1: (a) The simulation algorithm of Parametric Evolutionary Network Model. (b) An illustrative example of network evolution over time. The black edges represent connections formed via the node growth process, while the red edges represent those formed through edge growth. As time progresses, the network expands by iteratively applying these two probabilistic processes, capturing both exploratory and consolidative collaboration behaviors.

3.1 Node Growth Process Analysis and Simulation

In this section, we theoretically analyze the node growth mechanism in PENM and examine how parameters influence the expansion of the network in terms of node counts. This analysis provides insight into how the network scales over time and how different settings of model parameters affect growth dynamics.

The node growth process refers to the introduction of new nodes into the network. At each time step, existing nodes may establish connections with new nodes, and the number of such connections follows a Poisson distribution with parameter $\lambda_i^{(t)} = (1 - \frac{\tau(i)}{t})\alpha$. This formulation reflects the intuition that the longer a node has been in the network, the more likely it is to attract new connections. To quantify the effect of parameters on the growth trajectory, we analyze the expected increase in the number of nodes:

$$E(N(t) - N(t-1)) = p\alpha \sum_{i=1}^{N(t-1)} (1 - \frac{\tau(i)}{t}) = p\alpha \sum_{i=1}^{N(t-1)} (1 - \frac{i(N(i) - N(i-1)}{t}) = \frac{p\alpha}{t} \sum_{i=1}^{t-1} N(i) \tag{1}$$

Then $N(t)$ can be written as (from Appendix.A **Lemma 1**):

$$N(t) \approx N(0) \cdot \frac{p\alpha}{e^{p\alpha}} \sum_{k=0}^{\infty} \frac{(p\alpha)^k}{k!} \binom{n+k}{n} \sim e^{p\alpha\sqrt{n}}. \tag{2}$$

The theoretical result shows that the network size is mainly governed by $p\alpha$. Different combinations of p and α can lead to similar node growth patterns if $p\alpha$ remains constant. o validate this conclusion, we perform iterative simulations, comparing simulated outcomes with theoretical predictions. We performed 100 iterations of the node growth process and computed the average of the outcomes for $N(t)$, yielding the simulated mean N(t). Concurrently, the calculated N(t) was derived using the formula presented in Equation (2).

As shown in Figure 2, the simulation results align closely with the theoretical curve, confirming that the growth rate of $N(t)$ is primarily determined by the value of $p\alpha$. Increasing either p or α accelerates the introduction of new nodes into the network. These simulation results confirm the model's robustness in simulating realistic node growth dynamics.

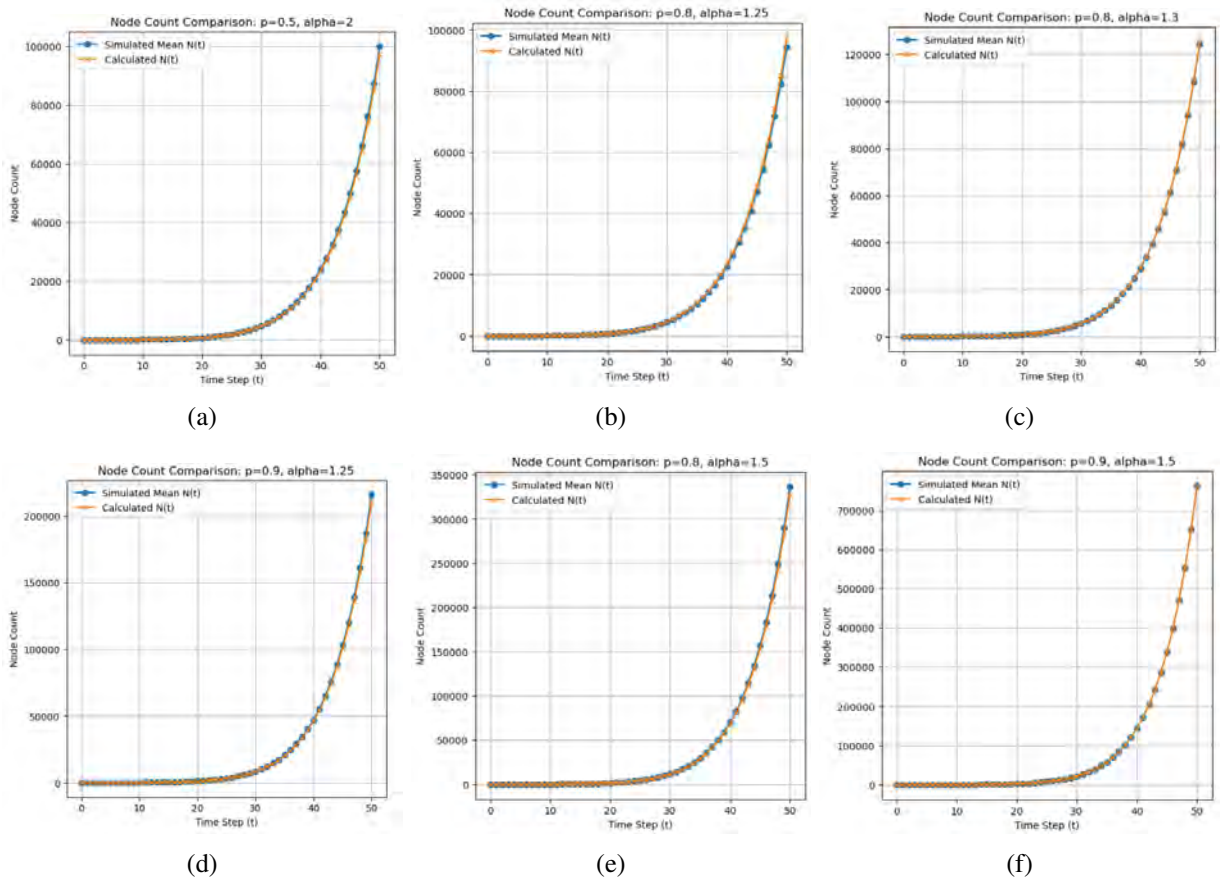


Figure 2: Comparison of theoretical number of nodes and simulated number of nodes over time under different parameter values. Calculated $N(t)$ is from (2), Simulated Mean $N(t)$ is averaged over 100 simulations.

3.2 Edge Growth Process Analysis and Simulation

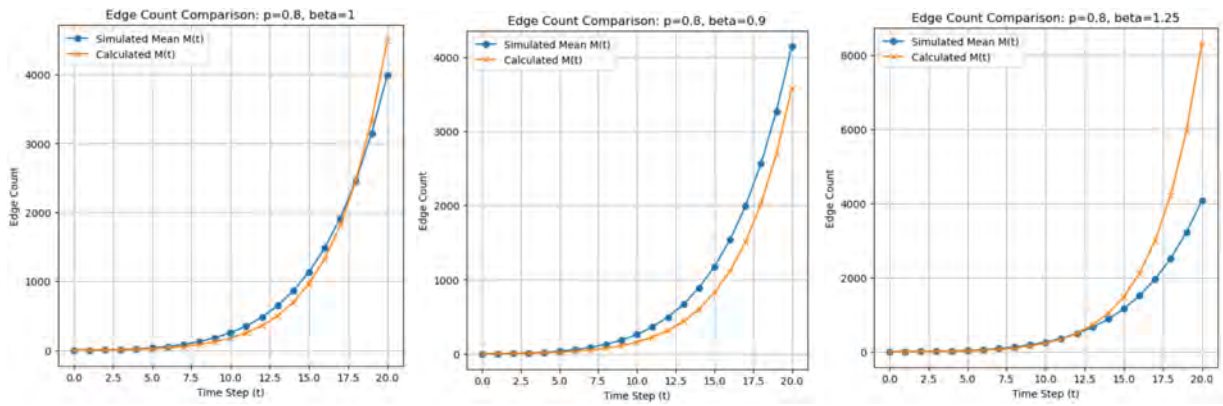
This section explores the edge growth mechanism in PENM and examines how parameters influence the total number of edges. Understanding edge growth is essential for capturing the structural evolution and densification of collaboration networks. In PENM, the growth of edges occurs through two channels: (1) Node-induced edges: When a node forms connections with new nodes, the number of added edges equals the number of new collaborators. (2) Internal edge formation: Existing nodes can also form new edges with other existing nodes—specifically, with neighbors of their current neighbors. The probability of connecting to a two-step neighbor is defined by $\mu_{ij}^{(t)}$. Then the expectation of number of added edges is:

$$
\begin{aligned}
E(M(t)-M(t-1)) &= p\sum_{i=1}^{N(t-1)}\lambda_i^{(t)} + (1-p)\sum_{i=1}^{N(t-1)}\sum_{j\in neb(i)}\mu_{ij}^{(t)}. \\
&= \frac{p\alpha}{t}\sum_{i=1}^{t-1}N(i) + \frac{(1-p)\beta}{(d_{max}^{(t-1)})^2}\sum_{i=1}^{N(t-1)}\sum_{j\in neb(i)}d_i^{(t-1)}d_j^{(t-1)}. \\
&\approx \frac{p\alpha}{t}\sum_{i=1}^{t-1}N(i) + \frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2}(1-p)\beta N(t-1) \qquad (3)
\end{aligned}
$$

From Appendix.A **Lemma 2**, when the coefficient $(1-p)\beta$ is large, the growth rate of $M(t)$ will significantly outdistance that of $N(t)$. The approximation of $M(t)$ is:

$$M(t) = \frac{1}{c+1}[c\sum_{k=0}^{t}(c+1)^{n-k}N(k) - N(t)], \text{ where } c = \frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2}(1-p)\beta. \tag{4}$$

To support this analysis, we conducted extensive simulations of edge growth dynamics This process is based on node growth process, we set unrelated parameter $\alpha = 1.3$. However, the coefficient $\frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2}$ is not constant. In the subsequent analysis, we fix the time step at 20, keeping the varying coefficient within a relatively small range. We proceed by fixing the parameter $\frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2} = 1.5$, and then analyze the impact of varying β.



(a) (b) (c)

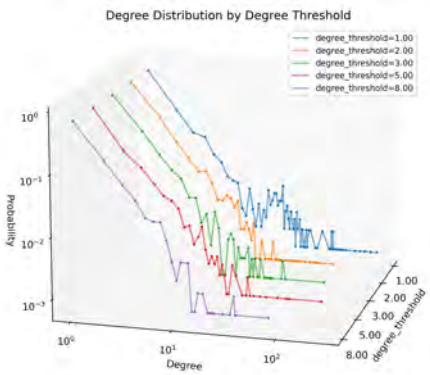
Figure 3: Edge growth trajectories under different values of parameter. Calculated $M(t)$ is from (4), Simulated Mean $M(t)$ is averaged over 100 simulations.

As we transition from Figure 3.(a) to Figure 3.(b), the parameter β decreases from 1 to 0.9. The probability of nodes connecting with existing nodes becomes higher, resulting in a larger maximum degree $d_{max}^{(t-1)}$. With the coefficient $\frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2}$ reduced, the simulation results indicate a slower rate of node growth. For Figure 3.(c), β is increased from 1 to 1.25, leading to a larger $\frac{(d_{avg}^{(t-1)})^3}{(d_{max}^{(t-1)})^2}$ and, consequently, a faster rate of node growth. These simulation outcomes demonstrate the model's capability to simulate varying connectivity levels in collaboration networks and to capture the structural complexity of real academic collaborations.

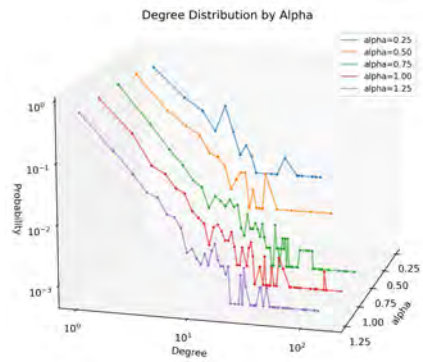
3.3 Parameter Sensitivity Analysis

In this section, we conduct a sensitivity analysis to examine how key parameters affect the degree distribution of networks simulated by the PENM model. This analysis provides guidance for selecting appropriate parameter values when simulating real-world collaboration networks with different structural characteristics. The analysis contains four parameters: α, β, p, and *degree_threshold*. To analyze the effects of each parameter, we systematically vary their values and observe the resulting changes in the degree distribution. Figure 4 presents four subplots, each corresponding to one parameter, illustrating how the distribution shifts under different parameter values.

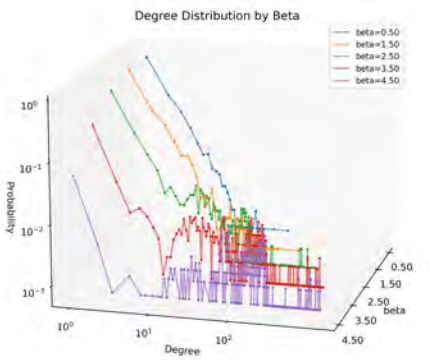
Figure 4 illustrates the sensitivity of PENM's degree distribution to key parameters: a higher *degree_threshold* reduces the maximum degree and concentrates nodes at lower degrees, lower α slows node introduction



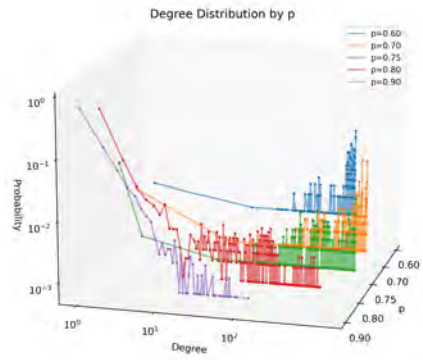
(a) Varying *degree_threshold* affects the maximum degree and low-degree node concentration.



(b) Changes in α influence the pace of new node introduction, impacting degree spread.



(c) Higher β values increase the density of edges, shifting degree distribution to the right.



(d) Parameter p, balances node and edge growth. Lower p values lead to higher average degrees and heavier tails.

Figure 4: Sensitivity analysis for parameters *degree_threshold*, α, β, and p.

for a more uniform degree spread, increasing β enhances edge density to shift the distribution toward higher degrees ($> 10^2$), and decreasing p, which balances node and edge growth, significantly boosts edge formation, yielding higher average degrees and heavier-tailed distributions.

Overall, the sensitivity analysis demonstrates that the PENM model can flexibly reproduce various types of degree distributions by adjusting its parameters. This tunability enhances its applicability across a wide range of collaboration networks, enabling users to tailor the model according to observed real-world properties.

4 FITTING WITH REAL WORLD NETWORKS

In this section, we evaluate the performance of the Parametric Evolutionary Network Model on real-world scholar collaboration datasets. Specifically, we use three datasets: ca-CSphd (Dohleman 2006), ca-Erdos992 (Batagelj and Mrvar 2000) and cit-DBLP (Rossi and Ahmed 2015). For comparison, we also include two established network generation models: the Network Reconstruction Model (NRM) (Pandey and Adhikari 2017), the Evolving Network Model (ENM) (Li et al. 2023).

4.1 Evaluation Metrics

In this study, we evaluate the model's performance by focusing on the degree distribution similarity. The degree distribution serves as a pivotal indicator of network structure, delineating the prominence of well-connected nodes, or hubs, and quantifying the inequality in collaboration patterns. These hubs play a pivotal role in knowledge dissemination and network robustness. The degree distribution's clear interpretability and its established use in network analysis make it an ideal benchmark for assessing how closely our simulated networks mirror actual academic networks.

We use three widely adopted metrics to assess distributional similarity: the Pearson correlation coefficient (ρ), the Kullback-Leibler (KL) divergence, and the Jensen-Shannon (JS) divergence (Nielsen 2021). The Pearson correlation coefficient ρ serves as a metric for evaluating the correlation between distributions and is calculated as follows:

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

where x_i and y_i are the values from the two distributions, and $\bar{x}$and $\bar{y}$ are their respective means. The range of ρ is $[-1, 1]$.

The Kullback–Leibler (KL) divergence measures how one probability distribution diverges from a reference distribution, and is defined as:

$$KL(P||Q) = \sum_{i=1}^{n} P(x_i) log(\frac{P(x_i)}{Q(x_i)}),$$

where P and Q represent the true and generated degree distributions, respectively. Its range is $[0, +\infty)$.

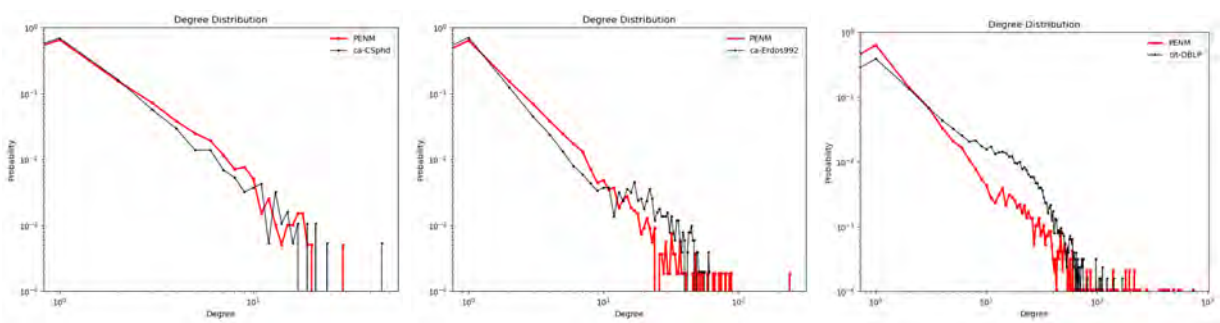
The Jensen–Shannon divergence is a symmetrized and smoothed version of the KL divergence, based on the KL divergence, and is defined as:

$$JS(P||Q) = \frac{1}{2}KL(P||Q) + \frac{1}{2}KL(Q||P),$$

The JS divergence ranges between 0 and 1 and is often preferred due to its symmetry and stability.

4.2 Comparison of PENM, NRM, ERM with Real Data

This section evaluates the degree distributions simulated by PENM against real-world collaboration networks. The detail of 3 network dataset: ca-CSphd has 1882 nodes and 1740 edge; ca-Erdos992 has 5094 nodes and 7515 edges; cit-DBLP has 12591 nodes and 49635 edges.



(a) (b) (c)

Figure 5: Degree distribution comparisons between real-world scholar collaboration networks (black) and networks generated by PENM (red): (a) ca-CSphd, (b) ca-Erdos992, (c) cit-DBLP.

The degree distributions simulated by PENM closely align with real-world data, which is displayed in Figure 5. The red lines represent the degree distributions produced by PENM, while the black lines correspond to the ground-truth networks. Across all three datasets, the degree distribution generated by PENM closely aligns with the real data, demonstrating the model's effectiveness in reproducing real-world structural patterns.

Table 2: Optimal parameter for different dataset.

Network	*degree_threshold*	p	α	β
ca-CSphd	5	0.95	0.78	0.05
ca-Erdos992	2	0.9	1.25	0.65
cit-DBLP	3	0.91	3	4

Table 2 presents the optimal parameters for PENM simulations across each dataset. The acquisition of these optimal parameters was achieved through a grid search approach. The initial parameters were estimated based on the analysis presented in Section 3.3. The variation in parameter values across datasets highlights the flexibility of PENM in adapting to networks with different characteristics.

Table 3: Distribution similarity.

		Evaluation Metrics		
Network	Models	ρ	KL	JS
ca-CSphd	ERM	0.993	**0.036**	0.096
	NRM	0.993	0.089	0.106
	PENM(ours)	**0.998**	0.057	**0.085**
ca-Erdos992	ERM	0.984	0.114	0.179
	NRM	0.981	0.113	0.17
	PENM(ours)	**0.996**	**0.089**	**0.113**
cit-DBLP	ERM	0.978	0.448	0.372
	NRM	0.933	**0.24**	0.255
	PENM(ours)	**0.98**	0.299	**0.229**

Our PENM model outperforms the other two models in some metrics across the datasets, which is shown in Table 3. In most cases, PENM achieves higher Pearson correlation and lower divergence values. Specifically, PENM achieves the highest ρ for ca-CSphd and ca-Erdos992, suggesting a stronger correlation with the actual data distributions. Additionally, PENM demonstrates lower KL and JS values for ca-CSphd and cit-DBLP, indicating a closer match to the real data distributions in terms of these divergence metrics. These simulation results confirm PENM's ability to accurately simulate the degree distributions of real scholarly networks. The model's flexibility allows it to adjust to varying network sizes and structures—making it a practical tool for modeling academic collaboration dynamics.

5 CONCLUSION AND OUTLOOK

In this paper we present the PENM model, a novel framework for simulating the evolutional characteristics of collaboration networks. Through rigorous theoretical analysis and simulation-based validation on real-world datasets (ca-CSphd and ca-Erdos992 and cit-DBLP), we demonstrated the model's capability to accurately simulate scholar collaboration networks. Its parameterized structure enables tailored simulations to diverse scenarios, addressing the limitations of existing static or locally dynamic approaches.

While PENM currently simulates structural network evolution of collaboration networks, it simplifies the real-world scenario by considering only topological evolution and ignoring node-level attributes such as research fields, affiliations, and seniority. Incorporating such attributes alongside more node-level attributes

in future extensions could enrich the model and enhance its realism and explanatory power. Given its parametric nature, PENM is particularly suited for integrating node-centric data and supporting more robust simulations. In future work, we plan to validate the model using additional structural metrics, such as clustering coefficient and average path length, to provide a more comprehensive assessment beyond degree distributions. This will further enhance PENM as a flexible and practical tool for studying academic collaborations across diverse contexts.

We propose that simulation-driven models like PENM hold potential for forecasting future trends in scholar collaborations. In addition, future versions of PENM could incorporate mechanisms to simulate node inactivity, such as scholars exiting academia, to better reflect the lifecycle of real-world academic participants. Future work could leverage PENM's simulation capabilities to develop predictive analytics tools, enhancing the strategic formation of research partnerships. We plan to extend the application of this model, using simulations to deepen insights into network dynamics across various domains, providing new perspectives for collaboration network forecasting and recommendation.

ACKNOWLEDGMENT

This research was funded by the Tencent (Shenzhen) Co., Ltd. (Tencent University Cooperation Project No. 20239680152) and the Guangdong Pearl River Plan (2019QN01X890).

A APPENDIX

All experiments presented in this study were conducted using Python. The computational environment was configured with an AMD Ryzen 9 7940HX with Radeon Graphics processor, operating at 2.40 GHz, and equipped with 15.2 GB of RAM.

Lemma 1. For recursive sequence:

$$a_n = a_{n-1} + \frac{m}{n}\sum_{i=0}^{n-1} a_i.$$

With $a_0 > 0,\ a_1 = 2a_0$. The analytical expression of a_n is:

$$a_n = \frac{ma_0}{e^m} \cdot \sum_{k=0}^{\infty} \frac{m^k}{k!}\binom{n+k}{n} + (1-m)a_0.$$

Proof. Let $f(x) = \sum\limits_{n=1}^{\infty} a_n x^n$. Then $f^{'}(x) = \sum\limits_{n=1}^{\infty} n a_n x^{n-1}$. The recursive sequence can be written as:

$$na_n = na_{n-1} + m\sum_{k=0}^{n-1} a_k = (n-1)a_{n-1} + a_{n-1} + m\sum_{k=0}^{n-1} a_k.$$

Then we have:

$$\begin{aligned} f^{'}(x) - a_1 = \sum_{n=2}^{\infty} n a_n x^{n-1} &= \sum_{n=2}^{\infty}(n-1)a_{n-1}x^{n-1} + \sum_{n=2}^{\infty} a_{n-1}x^{n-1} + m\sum_{n=2}^{\infty}\sum_{k=0}^{n-1} a_k x^{n-1} \\ &= xf^{'}(x) + f(x) - (m+1)a_0 + \frac{m}{1-x}f(x). \end{aligned}$$

The differential equation is

$$f^{'}(x) + \frac{x-1-m}{(1-x)^2}f(x) + \frac{m-1}{1-x}a_0 = 0.$$

Then

$$f(x) = \frac{ma_0}{e^m} \cdot \frac{1}{1-x} \cdot e^{\frac{m}{1-x}} + \frac{1-m}{1-x} a_0$$
$$= \sum_{n=0}^{\infty} [\frac{ma_0}{e^m} \cdot \sum_{k=0}^{\infty} \frac{m^k}{k!} \binom{n+k}{n} + (1-m)a_0] x^n.$$

Thus,

$$a_n = \frac{ma_0}{e^m} \cdot \sum_{k=0}^{\infty} \frac{m^k}{k!} \binom{n+k}{n} + (1-m)a_0 \sim e^{m\sqrt{n}}. \quad \square$$

Lemma 2. For recursive sequence:

$$b_n = a_n - a_{n-1} + (c+1)b_{n-1}.$$

With $b_0 = a_0$. The analytical expression of b_n is:

$$b_n = \frac{1}{c+1}[c \sum_{k=0}^{n} (c+1)^{n-k} a_k - a_n].$$

Proof. Let $g(x) = \sum_{n=1}^{\infty} b_n x^n$. Then we have:

$$g(x) - b_0 = \sum_{n=1}^{\infty} b_n x^n = \sum_{n=1}^{\infty} (a_n - a_{n-1}) x^n + (c+1) \sum_{n=1}^{\infty} b_{n-1} x^n = (1-x) f(x) + (c+1) x g(x) - a_0.$$

Therefore,

$$g(x) = \frac{1-x}{1-(c+1)x} f(x) = \frac{ma_0}{e^m} \cdot \frac{1}{1-(c+1)x} \cdot e^{\frac{m}{1-x}} + \frac{1-m}{1-(c+1)x} a_0$$
$$= \frac{1}{c+1} \sum_{n=0}^{\infty} [c \sum_{k+l=n} a_k (c+1)^l - a_n] x^n.$$

Thus,

$$b_n = \frac{1}{c+1}[c \sum_{k=0}^{n} (c+1)^{n-k} a_k - a_n]. \quad \square$$

REFERENCES

Barabási, A.-L., and R. Albert. 1999. "Emergence of Scaling in Random Networks". *Science* 286(5439):509–512.

Batagelj, V., and A. Mrvar. 2000. "Some Analyses of Erdos Collaboration Graph". *Social Networks* 22(2):173–186.

Chuang, Y.-T., and Y.-H. Chen. 2022. "Social Network Analysis and Data Visualization of MIS International Collaboration in Taiwan". *Library Hi Tech* 40(5):1422–1458.

Dohleman, B. S. 2006. "Exploratory Social Network Analysis with Pajek". *Psychometrika* 71(3):605.

Eden, M., R. Castonguay, B. Munkhbat, H. Balasubramanian, and C. Gopalappa. 2021. "Agent-based Evolving Network Modeling: a New Simulation Method for Modeling Low Prevalence Infectious Diseases". *Health Care Management Science* 24(3):623–639.

Franceschet, M. 2011. "Collaboration in Computer Science: A Network Science Approach". *Journal of The American Society for Information Science and Technology* 62(10):1992–2012.

Gyarmati, L., and T. A. Trinh. 2010. "Scafida: A scale-free Network Inspired Data Center Architecture". *ACM SIGCOMM Computer Communication Review* 40(5):4–12.

Hadian, A., S. Nobari, B. Minaei-Bidgoli, and Q. Qu. 2016. "Roll: Fast in-memory Generation of Gigantic Scale-free Networks". In *Proceedings of The 2016 International Conference on Management of Data*, June 26th – July 1st, San Francisco, USA, 1829–1842.

Kong, X., M. Mao, J. Liu, B. Xu, R. Huang, and Q. Jin. 2018. "TNERec: Topic-aware Network Embedding for Scientific Collaborator Recommendation". In *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Octvber 8th – 12nd, Guangzhou, China, 1007–1014.

Kong, X., Y. Shi, W. Wang, K. Ma, L. Wan, and F. Xia. 2019. "The Evolution of Turing Award Collaboration Network: Bibliometric-level and Network-level Metrics". *IEEE Transactions on Computational Social Systems* 6(6):1318–1328.

Li, Y., M. Feng, and J. Kurths. 2023. "Evolving Network Modeling Driven by The Degree Increase and Decrease Mechanism". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53(9):5369–5380.

Nielsen, F. 2021. "On A Variational Definition for The Jensen-Shannon Symmetrization of Distances Based on The Information Radius". *Entropy* 23(4):464.

Pandey, P. K., and B. Adhikari. 2017. "A Parametric Model Approach for Structural Reconstruction of Scale-free Networks". *IEEE Transactions on Knowledge and Data Engineering* 29(10):2072–2085.

Pasta, M. Q., Z. Jan, A. Sallaberry, and F. Zaidi. 2013. "Tunable and Growing Network Generation Model with Community Structures". In *2013 International Conference on Cloud and Green Computing*, September 30th – October 2th, Washington DC, United States, 233–240.

Rossi, R., and N. Ahmed. 2015. "The Network Data Repository with Interactive Graph Analytics and Visualization". In *Proceedings of The AAAI Conference on Artificial Intelligence*, Volume 29, January 25th – 30th, Austin, USA.

Ruela, A. S., and K. V. Delgado. 2018. "Scale-free Evolutionary Level Generation". In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, August 14th – 17th, Maastricht, Netherlands, 1–8.

Simas, T., and L. M. Rocha. 2008. "Stochastic Model for Scale-free Networks with Cutoffs". *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 78(6):066116.

Wang, W., J. Liu, T. Tang, S. Tuarob, F. Xia, Z. Gong *et al*. 2020. "Attributed Collaboration Network Embedding for Academic Relationship Mining". *ACM Transactions on The Web (TWEB)* 15(1):1–20.

Xiao, X., and M. Prompanyo. 2023. "Influence of Scientific Collaboration Network on Academic Performance". *Int. J. Bus. Manag* 16:1–68.

Zhang, Y., C. Zhang, and X. Liu. 2017. "Dynamic Scholarly Collaborator Recommendation via Competitive Multi-agent Reinforcement Learning". In *Proceedings of The Eleventh ACM Conference on Recommender Systems*, August 27th – 31st, Como, Italy, 331–335.

Zhao, L., Y.-C. Lai, K. Park, and N. Ye. 2005. "Onset of Traffic Congestion in Complex Networks". *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 71(2):026125.

AUTHOR BIOGRAPHIES

JIADONG LIU is a master student at the Shenzhen international Graduate School, Tsinghua University, China. His email address ljd23@mails.tsinghua.edu.cn.

KEQIN GUAN is a Ph.D. candidate at the Shenzhen international Graduate School, Tsinghua University, China. His email address keqinoly@163.com.

SIYU CHEN is a master student of the Shenzhen International Graduation School (SIGS), Tsinghua University, China. Her email address is chen-sy23@mails.tsinghua.edu.cn.

WEI CHEN, corresponding author. He is a D.Eng candidate at Shenzhen International Graduate School, Tsinghua University, China. He currently serves as a Director at Tencent. Additionally, Wei is the founder and director of Wizsci (wizsci.com), a scientific research platform dedicated to advancing academic-industry collaboration. His email address is cw.chen@163.com.

WAI KIN (VICTOR) CHAN, corresponding author. He is Professor of Shenzhen International Graduation School (SIGS), Tsinghua University, China. His research interests include discrete-event simulation, agent-based simulation, big-data analytics and their applications in social networks, business big-data, service systems, healthcare, transportation, energy markets, and manufacturing. His email address is chanw@sz.tsinghua.edu.cn.