Scale-free First-Passage Percolation*

Hansheng Liu, Shuzhen Zhang, Haotian Ju, Linzhe Teng
Project Leader: Aditya Gopalan, Daecheol Kim
Faculty Mentor: Partha S. Dey

December 27, 2024

Abstract

First passage percolation is a mathematical model used to study the spread of rumor or information among nodes in a random network. In the long-range model, a node passes the information to another node after a random time with a mean proportional to some power α of the distance, independent of each other. For larger α, the travel time for information will get larger. This project considers the scale-free version, where each node u is assigned a non-negative random weight W_u. One can think of W_u as a quantification of node u's importance. When the weight variables have unbounded support, nodes with very high weight will be present. These form the hubs for the network and play a crucial role in the functionality of the network. In this new model, the time needed for information to pass from node u to node v is scaled by W_uW_v. We study the effect of the tail-exponent γ of W_u and power-exponent α on the passage time between two far-away nodes and, using simulation, understand the phase transition when the base graph is a one-dimensional line graph. We also use simulation to understand the distributional limit for the passage time for specific values of α, γ.

1 Introduction

Networks are ubiquitous in our world, representing systems as diverse as social relationships, transportation infrastructure, and biological interactions. Understanding how information or entities traverse these networks over time is a fundamental challenge with implications ranging from spreading diseases to functioning communication networks. First Passage Percolation (FPP) emerges as a robust mathematical framework designed to unravel the intricate dynamics of traversal through complex networks. The primary focus of FPP lies in analyzing the first passage time (FPT), representing the duration for a signal or particle to traverse from one point to another within the network.

Scale-free first-passage percolation introduces the concept of scale-free networks, characterized by a node degree distribution following a power-law distribution. In contrast to a

*Support for this project was provided by the Illinois Geometry Lab and the Department of Mathematics at the University of Illinois Urbana–Champaign.

regular network, where most nodes have a similar degree, in a scale-free network, there are a few nodes (hubs) with an exceptionally high degree, while the majority of nodes have a much lower degree. Within the framework of first passage percolation, a scale-free network implies the presence of influential hubs that markedly impact the network's efficiency. These hubs function as efficient shortcuts, expediting the transmission of signals through the network. Incorporating scale-free features in the network's structure leads to intriguing and non-trivial behaviors regarding first-passage times.

1.1 The Model

We begin establishing the foundation of our model by introducing some essential definitions. Let N be a positive integer. We denote by $V_N := \{-N, \dots -1, 0, 1, \dots, N\}$ the set of vertices and $E_N := \{(i, j) \mid i, j \in V_N\}$ the set of edges. Consider an undirected one-dimensional line graph from $-N$ to N where all the vertices are interconnected, forming a complete graph. In this graph, each edge (i, j) has ℓ^1- distance $|i - j|$. We denote this graph by L_N.

We assume that all the vertex weights are independent and identically distributed (i.i.d.) by W_i, and the distribution is defined as

$$\mathbb{P}(W_i > x) = x^{-\gamma}, x \geq 1,$$

where $\gamma > 0$ is the **scale-free parameter**. We now define the edge weight ω_{ij} by

$$\mathbb{P}(\omega_{ij} > t \mid W_i, W_j) = \exp\left(-\frac{W_i W_j}{|i - j|^\alpha} \cdot t\right), t \geq 0,$$

where $\alpha > 0$ is the **distance cost exponent**. For instance, when $\alpha = 2$, the edge weight of $(200, 500)$ is distributed as $\frac{300^2}{W_i W_j} \cdot \text{Exp}(1)$.

Definition 1. *Let $T_{i,j}$ denote the minimal total weight of a path between vertices i and j. This is called the **first-passage time (FPT)** from i to j.*

Our primary objective is to capture the scaling properties of FPT $T_N := T_{0,\frac{N}{2}}$ in the complete graph L_N as the size of the graph N grows to infinity. We explore these scaling properties concerning the values of the scale-free parameter γ and the distance cost exponent α. Our approach involves conducting multiple simulations, varying parameters such as N, α, and γ. These simulations are carefully designed to observe diverse behaviors and properties of the graph under different conditions. The purpose of these simulations is to gain insights into how the scale-free parameter γ, the distance cost exponent α and the network size collectively influence the efficiency of information transport within the scale-free network model.

1.2 Algorithm used for Simulations

To find the shortest path, we used the Dijkstra algorithm. Specific greedy improvements of the algorithm were applied for some region (see Section 2.1 for more details). In our simulations, within $O(N^2)$ times, we were able to obtain the exact shortest path for every sample. For certain samples, a highly precise approximation could be obtained in $O(N)$.

2 Expected First Passage Time

In this section, our focus is on investigating the order of FPT T_N. We model this as the total cost of traveling from the first vertex to the vertex $\frac{N}{2}$ in the graph L_N. Through careful simulation of the FPT performance on graphs with different values of γ and α, we observed six distinct cases of FPT growth behavior. See Figure 1.

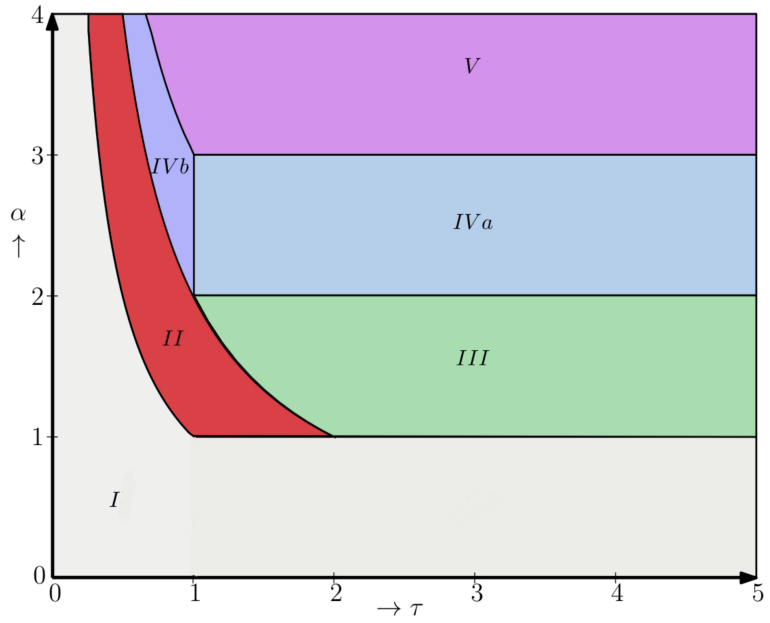


Figure 1: Phase diagram for the growth rate of T_N.

1. In Region I, The first passage time T_N follows an arbitrary small growth. That is, $T_N \to 0$ as $N \to \infty$.

2. In Region II, T_N follows a log-log growth: $T_N / \log\log n \to c$ for some constant c as $N \to \infty$.

3. In Region III, T_N follows a Ploy-log pattern: $T_N/(\log n)^{1/log_2(\alpha)+o(1)} \to c$ for some constant c as $N \to \infty$.

4. In Region IV_a, T_N has a sub-linear growth: $T_N/N^{\alpha-2} \to c$ for some constant c as $N \to \infty$. Similarly, in Region IV_b, T_N has a sub-linear growth $T_N/N^{\alpha-2/\gamma} \to c'$ for some constant c' as $N \to \infty$.

5. In Region V, T_N follows a linear growth: $T_N/N \to c$ for some constant c as $N \to \infty$.

All the constants above are distinct and need to be determined later. We have figured out their growth order, but the exact limits have not been determined. Therefore, We suppose all our simulation results follow the expectation pattern below in Section 3.

2.1 Greedy Algorithms for some Regions

A greedy algorithm is any algorithm that follows the problem-solving heuristic of making the locally optimal choice at each stage. Although a greedy strategy does not guarantee an optimal solution, a greedy heuristic can yield locally optimal solutions that approximate a globally optimal solution. By employing this approach, we can significanlty reduce our simulation time within a reasonable timeframe.

We have simulated the following four regions both by using Dijkstra and by using the greedy algorithm. We were able to obtain similar results from both. A rough description of the greedy algorithm in each region is as follows.

- **Region** II. Pick $\theta > 1$. Find the node with the largest weight among nodes from the starting node to $\frac{N}{\theta}$. Recursively find maximum weight nodes from starting node to $\frac{N}{\theta^2}$, $\frac{N}{\theta^3}, \ldots$ until $\frac{N}{\theta^k}$ is close enough to the starting node for some k. Do the same thing from the right side with the target node. Connect all these nodes to construct the path.

- **Region** III. Pick $0 < \theta < 1$. Find the shortest edge from a random node n_1 in $[0, N^\theta]$ to another random node n_2 in $[N - N^\theta, N]$. Recursively apply this algorithm to the sets $[0, n_1]$, and $[n_2, N]$.

- **Region** IV_a**.** Pick $\theta > 1$. Find the shortest edge from a random node n_1 in $[0, \frac{N}{\theta}]$ to another random node n_2 in $[N - \frac{N}{\theta}, N]$. Recursively apply this algorithm to the sets $[0, n_1]$, and $[n_2, N]$.

- **Region** IV_b**.** Pick $0 < \theta < 1$. Find the node with the largest weight in $[0, N^\theta]$. Recursively find nodes from starting node to $N^{\theta^2}, N^{\theta^3}, \ldots$ until N^{θ^k} is close enough to the starting node for some k. Do the same thing from the right side with the target node. Connect all these nodes to construct the path.

3 Empirical Distribution of First Passage Time

This section is dedicated to validating the convergence of the empirical distribution of first-passage time (FPT), as illustrated in the provided phase diagram. (Figure 1)

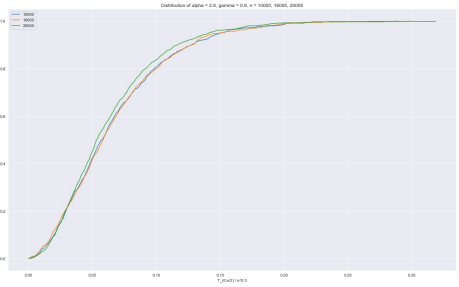
We generated a substantial amount of experimental data, from which we will present only the most representative sets. These will enable us to observe the scaling limit of FPT. For instance, in the case of $\gamma = 0.8$, we observed properties that manifest in three distinct regions, labeled as I, II, IV_b:

	$\alpha = 0.5$	$\alpha = 2.2$	$\alpha = 2.8$	$\alpha = 3.2$	$\alpha = 3.8$
$N = 8000$	7×10^{-5}	0.106	0.0719	0.0154	0.0135
$N = 16000$	4.75×10^{-5}	0.101	0.0681	0.0145	0.0131
$N = 20000$	3.82×10^{-5}	0.109	0.0643	0.0147	0.0128

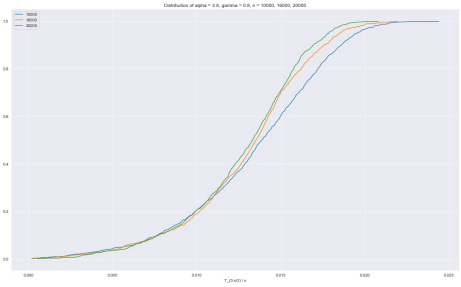
Table 1: Average of T_N when $\gamma = 0.8$.

It is observable that, although the number of nodes is increasing, the first passage time is somewhat decreasing. This phenomenon is attributable to using a scaling function in each

region. The reduction in time supports the idea that our distribution is generally convergent, with a rate corresponding to our initial presupposition. To gain better visualization, here is an example of the convergence of $\alpha = \{2.8, 3.8\}$:



CDF of scaled T_N when $\alpha = 2.8, \gamma = 0.8$ (Region IV_b)

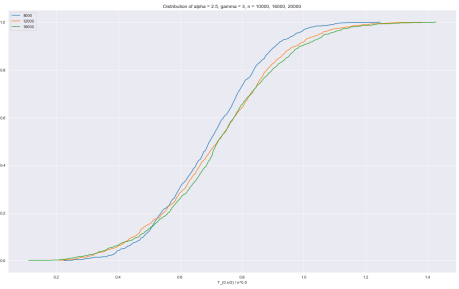


CDF of scaled T_N when $\alpha = 3.8, \gamma = 0.8$ (Region V)

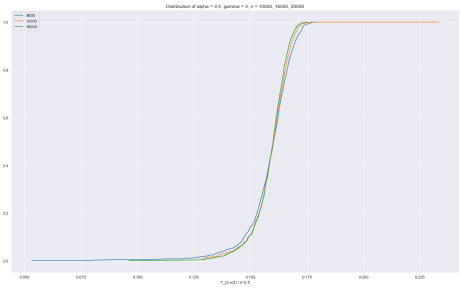
Next, we examine the case where $\gamma = 3, \alpha = 2.5, 3.5$:

	$\alpha = 0.5$	$\alpha = 1.5$	$\alpha = 2.1$	$\alpha = 2.5$	$\alpha = 2.9$	$\alpha = 3.1$	$\alpha = 3.5$
$N = 8000$	0.016	1.50	4.69	0.705	0.106	0.0799	0.1591
$N = 16000$	0.013	1.57	4.51	0.722	0.104	0.0782	0.1592
$N = 20000$	0.012	1.62	5.48	0.733	0.103	0.0774	0.1590

Table 2: Average of T_N after scaling when $\gamma = 3$.



CDF of scaled T_N when $\alpha = 2.5, \gamma = 3$ (Region IV_a)



CDF of scaled T_N when $\alpha = 3.5, \gamma = 3$ (Region V)

In contrast, by examining the graphs for $(\alpha = 3.5, \gamma = 0.8)$ and $(\alpha = 3.8, \gamma - 3)$ we can discern a phase transition between the two regions. The graph for $(\alpha = 3.8, \gamma = 3)$ exhibits a linear growth rate, which converges significantly faster than the sublinear case. Additionally, the slope in this scenario is much steeper.

4 Path Structures

We conducted simulations to analyze the path structure of each region depicted in Figure 1 in order to understand the underlying reasons for the observed order of the first passage time T_N. The following provides a summary of the path structures along with their representative images.

- Region I: This region is characterized by only a few jumps, as illustrated in Figure 2.
- Region II Both Figure 3 (run by the original algorithm) and Figure 4 (run by the greedy algorithm) display several large jumps.
- Region III: Similar to Region II, this region features several large jumps. See Figure 5.
- Region IV_a: Figures 6 exist one large jump and consisted with a lot of small jumps
- Region IV_b: Figures 7 (run by the original algorithm) and 8 (run by the greedy algorithm) both exhibit exactly one large jump along with several small jumps.
- Region V: There are many small jumps with no significant large jumps. See Figure 9

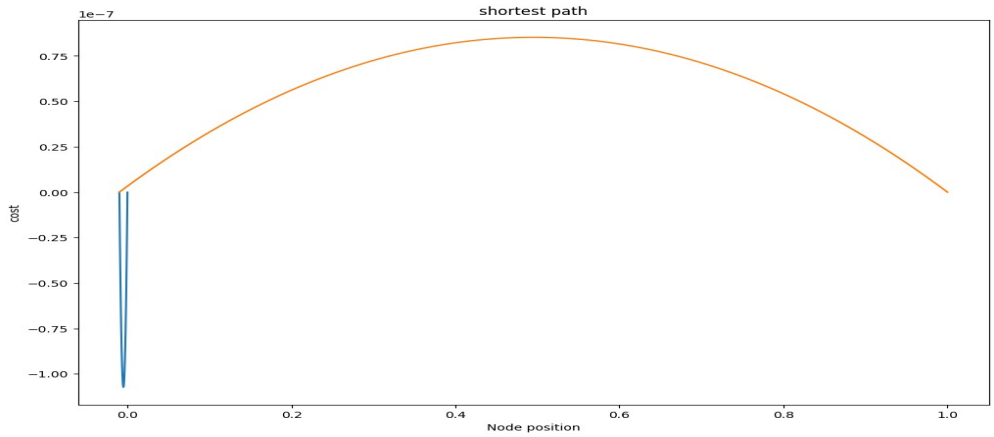


Figure 2: Region I ($\alpha = 0.5$, $\gamma = 0.8$)

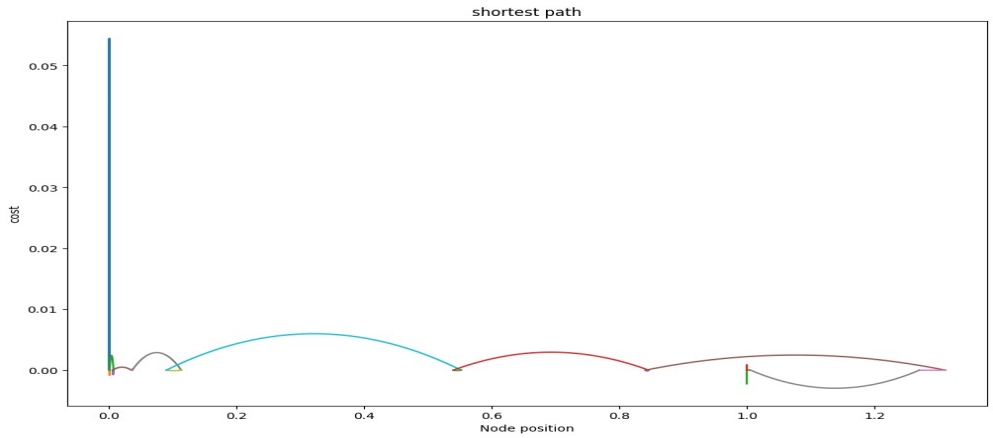


Figure 3: Region II ($\alpha = 2.2$, $\gamma = 0.8$)

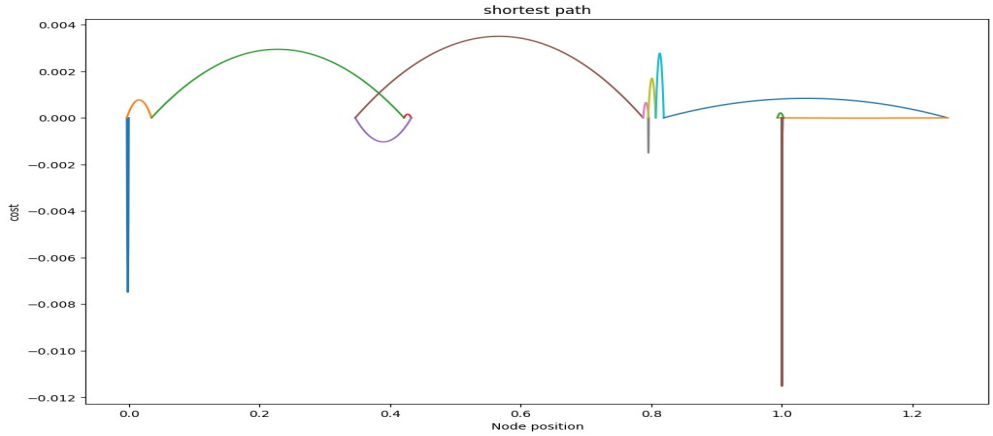


Figure 4: Region II using greedy algorithm

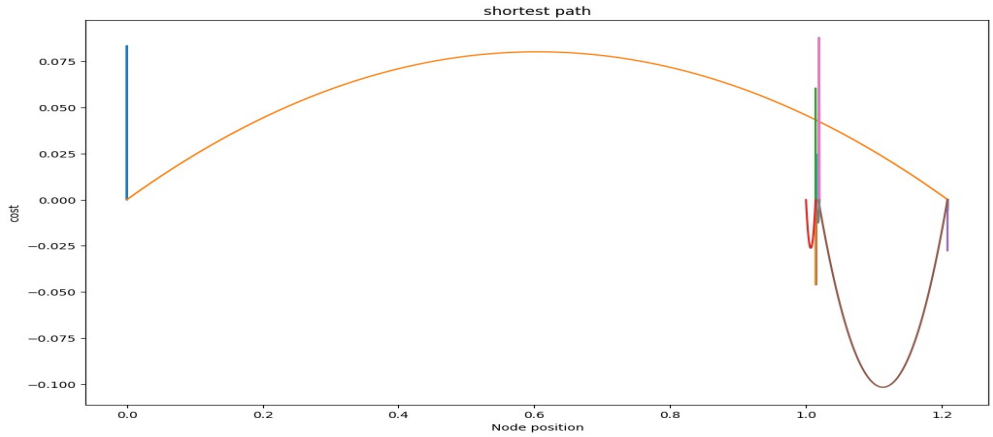


Figure 5: Region III ($\alpha = 1.5$, $\gamma = 3.0$)

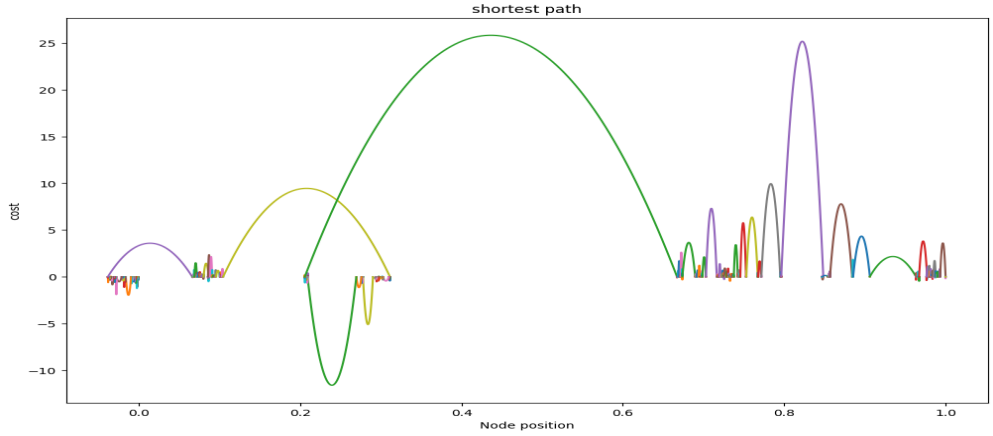


Figure 6: Region IV_a ($\alpha = 2.8$, $\gamma = 3.0$)

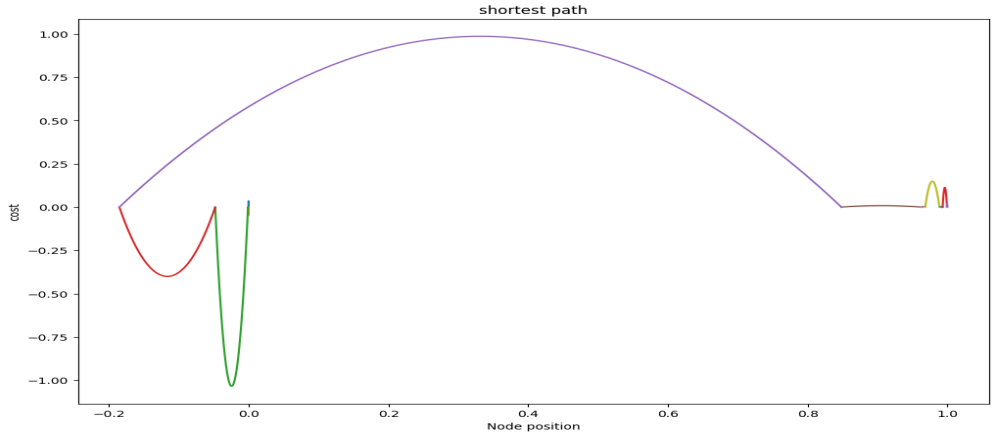


Figure 7: Region IV_b ($\alpha = 2.8$, $\gamma = 0.8$)

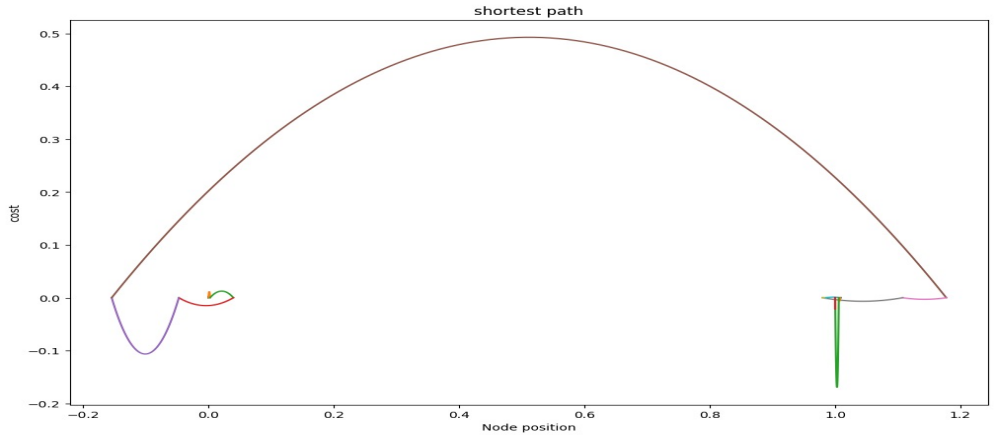


Figure 8: Region IV_b using greedy algorithm

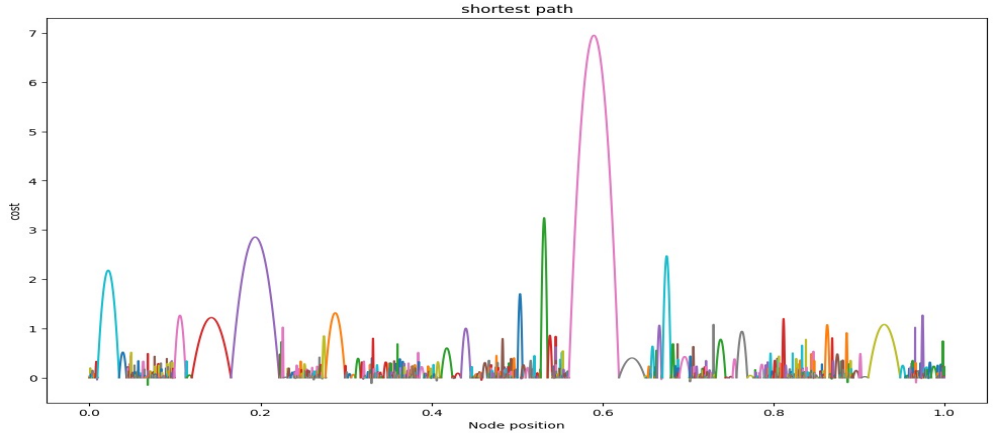


Figure 9: Region V ($\alpha = 3.2$, $\gamma = 0.8$)

5 Conclusion and Future Works

By exploring the first passage time (FPT) in scale-free networks, we have identified distinct regions with unique path structures, emphasizing the need for region-specific strategies in the analysis. The utilization of simulations has provided a comprehensive view of FPT behaviors across different graph conditions.

Our study has identified six distinct cases of FPT behavior, as summarized in the phase diagram (Figure 1). These cases, ranging from arbitrary small growth to linear growth, offer a nuanced understanding of how network parameters such as the scale-free parameter (γ) and the distance cost exponent (α) influence information transport efficiency. Additionally, the incorporation of a greedy algorithm in specific regions has demonstrated its effectiveness in reducing simulation time without compromising the accuracy of results. The comparative analysis of the greedy algorithm against the original Dijkstra algorithm in various regions provides valuable insights into trade-offs between efficiency and accuracy.

Looking forward, there are several promising avenues for further investigation. One key focus area is the augmentation of the parameter N in simulations. This is anticipated to offer deeper insights into how convergence trends are impacted by graph size, thereby enriching our understanding of the scaling property of first passage time. Another crucial aspect of future research involves a comparative analysis between the greedy algorithm and the original approach utilized in the study. This comparison aims to evaluate the trade-offs between efficiency and accuracy, particularly in terms of how the greedy algorithm performs under varying graph conditions.

Furthermore, there is a plan to explore new algorithms that could potentially reduce the computational demands for large graphs characterized by a high N value. The goal here is to strike a balance between efficiency and the integrity of results, ensuring that computational resources are optimized without compromising the accuracy and relevance of the findings. These future endeavors hold the promise of expanding our understanding of graph dynamics and providing more sophisticated tools for analyzing complex network structures.

6 References

[1] Auffinger, A., Damron, M., & Hanson, J. (2016). 50 years of first passage percolation, *arXiv*:1511.03262.

[2] Chatterjee, S., & Dey, P. S. (2016). Multiple Phase Transitions in Long-Range First-Passage Percolation on Square Lattices. *CPAM*, 69(2), 203–256.

[3] Deijfen, M., Van der Hofstad, R., & Hooghiemstra, G. (2013). Scale-free percolation. *AIHP*, Vol. 49, No. 3, pp. 817–838.