Complex Network Analysis: Investigating Centrality Measures and Network Disruption

Masterthesis
Mathematical Modeling, Simulation and Optimization

Meetkumar Pravinbhai Mangroliya
Matr.No. 220200921

24.12.2023

First Examiner: PD Dr. Robert Rockenfeller
Mathematical Institute, University of Koblenz, Germany
School of Biomedical Sciences, University of Queensland, Brisbane, Australia
School of Science, Technology and Engineering, University of the Sunshine Coast, Queensland, Australia

Second Examiner: Dr. Jens Dörpinghaus
Mathematical Institute, University of Koblenz, Germany
Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany

Declaration of authorship

I herewith confirm that I alone have authored this thesis, that I did not use any resources other than those I have cited — in particular no online sources not included in the bibliography section — and that I have not previously submitted this thesis in association with any other examination procedure.

Meet

Meetkumar Pravinbhai Mangroliya
Koblenz, December 24, 2023

Abstract

This thesis studies the complex behavior of complex networks, exploring the importance of nodes in network structure, the impact of node removal strategies, and the prediction of network connectivity after node removal. The study focuses on three fundamental network types: random networks, characterized by stochastic connections lacking specific underlying principles; scale-free networks, featuring a power-law degree distribution; and small-world networks, known for high clustering and short average path lengths. These networks serve as the cornerstone for comprehending the behavior of complex networks, each offering unique insights into network structure, connectivity, and robustness.

Centrality measures, such as degree, betweenness, closeness, and eigenvector centrality, are employed to identify critical nodes that play a pivotal role in network connectivity and information travel. The study examines the impact of different node and edge removal strategies, including targeted and random removal, on network resilience and identifies the critical point at which these networks experience disconnection.

Moreover, the thesis explores the ability to predict the change in degree centrality after node removal. By quantifying the predicted range, the study provides a more precise understanding of the impact of node removal on network structure.

To assess network robustness and identify the critical point at which networks become disconnected and lose the fundamental structure, the thesis employs the Kolmogorov-Smirnov (K-S) and Cramér-von Mises (CvM) tests. These statistical methods provide valuable insights into the critical transition points and help in developing resilient network designs.

The study's findings contribute to a deeper understanding of network dynamics and the factors that influence network resilience. The insights gained can be applied to various fields, including infrastructure design, social network analysis, and biological systems, to enhance network robustness and optimize network performance.

Acknowledgement

I would like to express my sincere gratitude to Dr. Jens Dörpinghaus, who generously provided me with the opportunity to explore the fascinating topic of my thesis. His insightful guidance, unwavering support, and valuable feedback have been instrumental in shaping the direction of my research.

I am also deeply thankful to PD Dr. Robert Rockenfeller, whose expertise and mentorship have played a crucial role in every stage of this academic journey.

I would also like to acknowledge the utilization of ChatGPT, a language generation model developed by OpenAI, for tasks related to paraphrasing and obtaining information on various topics.

Finally, I am grateful to my family and friends for their unwavering encouragement and understanding during the ups and downs of the research process.

Contents

1 Introduction 1

2 Background and Related Work 3
2.1 Complex Networks . . . 3
2.1.1 Random Networks . . . 3
2.1.2 Scale-free Networks . . . 4
2.1.3 Small-world Networks . . . 5
2.2 Centrality measures . . . 5
2.2.1 Degree Centrality . . . 6
2.2.2 Betweenness Centrality . . . 7
2.2.3 Closeness Centrality . . . 8
2.2.4 Eigenvector Centrality . . . 9
2.3 Node and Edge removal . . . 10
2.3.1 Impact on Network Connectivity and Structure . . . 11
2.4 Statistical Tests in Network Analysis . . . 13
2.4.1 Kolmogorov-Smirnov (K-S) test . . . 13
2.4.2 Cramér-von Mises Test . . . 14

3 Methodology 15
3.1 Generation of Complex Networks . . . 15
3.1.1 Random Networks . . . 15
3.1.2 Scale-free Networks . . . 16
3.1.3 Small-world Networks . . . 17
3.2 Selection of Centrality measures . . . 17
3.3 Simulating Node and Edge removal . . . 18
3.3.1 Node removal . . . 18
3.3.2 Edge removal . . . 19
3.4 Error measures . . . 19
3.5 Upper and Lower Bounds of Degree Centrality after Node removal . . . 20
3.6 Statistical Test procedures for Structural changes in Complex Networks . 22
3.6.1 Data Collection . . . 22
3.6.2 Hypothesis Testing . . . 23
3.6.3 Significance Level . . . 23
3.6.4 Interpretation . . . 23

4 Results and analysis **24**
4.1 Node importance in Complex Networks 24
4.2 The Impact of Node and Edge Removal Strategies on Centrality Measures of Complex Networks 28
4.2.1 High-degree node removal strategy 28
4.2.2 Low-degree node removal strategy 32
4.2.3 Random node removal strategy 35
4.2.4 Random edge removal strategy 39
4.3 Upper and Lower Bounds of Degree Centrality after removal 42
4.4 Structural changes in Complex Networks 45
4.4.1 Change in Network Connectivity 45
4.4.2 Change in Network Structure 48

5 Conclusion and Discussion **61**

Bibliography **63**

List of Figures

3.1 Erdős-Rényi Random Network. 15
3.2 Barabási-Albert Scale-Free Network. 16
3.3 Watts-Strogatz Small-World Network. 17

4.1 Scatter Plot and Correlation Matrix of an Erdős–Rényi Random Network with $n = 100$ nodes. 25
4.2 Scatter Plot and Correlation Matrix of a Barabási-Albert Scale-Free Network with $n = 100$ nodes. 26
4.3 Scatter Plot and Correlation Matrix of a Watts-Strogatz Small-World Network with $n = 100$ nodes. 27
4.4 Impact of High-degree Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes. 29
4.5 Impact of High-degree Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes. 30
4.6 Impact of High-degree Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes. 31
4.7 Impact of low-degree Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes. 32
4.8 Impact of low-degree Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes. 33
4.9 Impact of low-degree Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes. 34
4.10 Impact of Random Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes. 36
4.11 Impact of Random Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes. 37
4.12 Impact of Random Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes. 38
4.13 Impact of Random Edge Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes. 39
4.14 Impact of Random Edge Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes. 40
4.15 Impact of Random Edge Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes. 41

4.16 Upper and Lower Bounds of Degree Centrality in an Erdős–Rényi Random Network. 43
4.17 Upper and Lower Bounds of Degree Centrality in a Barabási-Albert Scale-Free Network. 43
4.18 Upper and Lower Bounds of Degree Centrality in a Watts-Strogatz Small-World Network. 44
4.19 Impact of Different Removal Strategies on Network Connectivity in an Erdős–Rényi Random Network. 46
4.20 Impact of Different Removal Strategies on Network Connectivity in a Barabási-Albert Scale-Free Network. 47
4.21 Impact of Different Removal Strategies on Network Connectivity in a Watts-Strogatz Small-World Network. 48
4.22 Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of High-Degree Node Removal. 49
4.23 Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Low-Degree Node Removal. 50
4.24 Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Random Node Removal. 51
4.25 Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Random Edge Removal. 52
4.26 Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of High-Degree Node Removal. 53
4.27 Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Low-Degree Node Removal. 54
4.28 Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Random Node Removal. 55
4.29 Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Random Edge Removal. 56
4.30 Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of High-Degree Node Removal. 57
4.31 Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Low-Degree Node Removal. 58
4.32 Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Random Node Removal. 59
4.33 Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Random Edge Removal. 60

List of Tables

2.1 Studies of Node and Edge Removal Effects in Network Connectivity . . . 12

4.1 Erdős–Rényi Random Network: Statistical Test Results after High-Degree Node Removal. . . . 50
4.2 Erdős–Rényi Random Network: Statistical Test Results after Low-Degree Node Removal. . . . 51
4.3 Erdős–Rényi Random Network: Statistical Test Results after Random Node Removal. . . . 51
4.4 Erdős–Rényi Random Network: Statistical Test Results after Random Edge Removal. . . . 52
4.5 Barabási-Albert Scale-Free Network: Statistical Test Results after High-Degree Node Removal. . . . 53
4.6 Barabási-Albert Scale-Free Network: Statistical Test Results after Low-Degree Node Removal. . . . 54
4.7 Barabási-Albert Scale-Free Network: Statistical Test Results after Random Node Removal. . . . 55
4.8 Barabási-Albert Scale-Free Network: Statistical Test Results after Random Edge Removal. . . . 56
4.9 Watts-Strogatz Small-World Network: Statistical Test Results after High-Degree Node Removal. . . . 57
4.10 Watts-Strogatz Small-World Network: Statistical Test Results after Low-Degree Node Removal. . . . 58
4.11 Watts-Strogatz Small-World Network: Statistical Test Results after Random Node Removal. . . . 59
4.12 Watts-Strogatz Small-World Network: Statistical Test Results after Random Edge Removal. . . . 60

1 Introduction

In today's world, networks are everywhere, these networks can be physical systems, such as power grids or highways, or abstract entities, such as networks of acquaintances or collaborations, we ourselves are interconnected as an individual within a network of social relationships [1].

The development of theoretical complex network models began with introduction of a random network model in 1959 [2, 3], characterized by a lack of inherent structure or organization. This model serves as a baseline for comparison with other network types. Subsequently, a small-world network model was introduced in 1998 [4], demonstrating the ability of networks to exhibit both high clustering (dense connections between nearby nodes) and a short average path lengths (efficient information propagation), a property characteristic of real-world systems. Finally, in 1999 [5], a scale-free network model emerged, highlighting the presence of a power-law degree distribution in many real-world networks, where most nodes have few connections while few nodes (hubs) are highly connected.

In network analysis, centrality measures provide crucial quantitative tools for assessing the importance of individual nodes within a network [6, 7]. Centrality measures - degree, betweenness, closeness, and eigenvector centrality - are used to identify critical nodes in many real-world networks, whether applied to analyze influential individuals in social networks, optimize essential hubs in transportation systems, or identify crucial proteins in biological networks [7]. To comprehend the resilience of networks to disruptions, node and edge removal strategies play a pivotal role. [8, 7]. These strategies involve selectively or randomly removing nodes or edges from a network to study how the network's structure and properties change. By evaluating the change in centrality measures before and after employing these strategies, we gain insights into the importance of different nodes and edges in maintaining the network's overall structure and function [7, 6].

Additionally, the research aims to identify the critical points at which anymore of a node or edge removal causes the network to be fragmented and undergo significant structural changes. This information is crucial for domains that rely on network connectivity and stability, such as telecommunications and social science, as it provides insights into how these networks might be susceptible to disruption and how their resilience can be enhanced [8, 7]. By understanding these critical points, we can proactively develop strategies to improve the networks' resilience to potential failures and improve their overall robustness.

To evaluate these changes statistically, the Kolmogorov-Smirnov (K-S) and Cramér-von Mises (CvM) tests are employed [9, 10]. These non-parametric statistical tests are used to compare the degree distribution - the distribution of node degrees, representing the number of connections each node has - of a network before and after removal and show the significant change in the structure. To our knowledge, only a few studies have employed the Kolmogorov-Smirnov (KS) test to assess the differences in the degree distributions of networks [11, 9]. However, no studies have employed the Cramér-von Mises (CvM) test for this purpose. This chapter serves as the foundation for our thesis, laying the groundwork for exploring types of complex networks and utilizing various tools to deepen our understanding.

We are interested in the following research questions:

1. The importance of nodes in complex networks: How to define the importance of nodes in a complex network?

2. The impact of node or edge removal strategies on complex networks: How do different node or edge removal strategies, such as targeted and random removal, affect a complex network?

3. Predicting the change in degree centrality: How to predict the general bounds of change in degree centrality after the removal of nodes?

4. Network connectivity and structure: How does the removal of nodes or edges influence the connectivity? and fundamental structure of a network?

2 Background and Related Work

2.1 Complex Networks

Networks are made up of nodes (which can be people, places, or things) and edges (which represent the connections between nodes). For instance, in social networks, nodes may correspond to individuals, while edges represent their relationships. In transportation networks, nodes often signify locations, and edges denote the routes connecting them.

We define a complex network as a graph (network) containing complex systems of connections that carry significant meaning across various fields of study [5]. An undirected (directed) graph $G = (V, E)$ consists of two sets V and E, such that $V \neq \emptyset$ and E is a set of unordered (ordered) pairs of elements of V [12, 1]. The elements of V ($\equiv \{v_1, v_2, \ldots, v_n\}$) are the nodes (or vertices, or points) of the graph G, while the elements of E ($\equiv \{e_1, e_2, \ldots, e_m\}$) are its links (or edges, or lines). The number of elements in V and E are denoted by n and m, respectively.

We study three types of complex networks:

1. Random networks.

2. Scale-free networks.

3. Small-world networks.

2.1.1 Random Networks

Definition 2.1 (Erdős-Rényi Random Network [2, 3])**.** *A random network, denoted as $G(N, P)$, is a graph consisting of N nodes, where any two nodes connected to each other with probability P.*

In a random network, connections between nodes are formed purely by chance, without adhering to specific rules or preferences [2]. This stochastic nature makes random networks a valuable reference point for understanding the impact of structure and connectivity on network behavior. In many scientific fields, random networks have been used as a standard to evaluate the importance of patterns and anomalies found in real-world networks [13, 7]. They provide insights into scenarios where connections established independently of node characteristics or network history. Understanding the behavior of random networks has implications for fields such as epidemiology, transportation, and social sciences, where the absence of strong organizing principles can influence system dynamics and outcomes [7].

Variations in network growth mechanisms and evolving structures can introduce deviations from the classical Erdős-Rényi model [14]. Researchers have explored extensions and modifications to the random network model to account for such nuances, which led to insights that improve the model's ability to account for complex real-world behaviors, such as the time-varying nature of network structures and processes, also known as temporal dynamics [15]. Additionally, studies [11] and [9] highlight the importance of randomness in shaping network properties, enabling a deeper understanding of their behaviors across various fields.

2.1.2 Scale-free Networks

Definition 2.2 (Barabási-Albert Scale-Free Network [5]). *A graph (network) $G = (V, E)$ is a scale-free network when the probability distribution $P(k)$ of node degrees k adheres to a power-law distribution $k^{-\gamma}$, where γ is a constant greater than 1.*

In scale-free networks, most nodes have relatively few connections, while a small number of nodes (hubs) exhibit an exceptionally high degree of connectivity [5]. The parameter γ in power-law distribution often referred to as the "fat-tail" parameter, meaning it influences the shape of the tail of the distribution [16]. A higher γ results in a slower decay of probabilities, leading to a "fatter" tail. For many real-world networks, where extreme events or values are rare but can have a meaningful influence on the network, this parameter in power-law distribution falls within the range $2 < \gamma < 3$ [11, 16].

Scale-free networks are found in diverse domains, this topology is akin to many real-world systems where a handful of entities play an outsized role [5]. For instance, think about social networks where most people have only a few friends, but a few individuals have a high number of connections. These highly connected individuals are like the 'hubs' in a scale-free network [7]. We find this pattern not only in social networks but also in various other places. On the internet, only a handful of websites, like Google or Wikipedia, get high number of visits, while most sites get just a few. The same principle applies in real life, where a small group of people might have much more money than others. This idea of a few things or people having a lot while most have only a little is what make scale-free networks interesting [7].

While many networks claim to be scale-free, not all of them genuinely confirm to this pattern when studied closely [9]. This suggests that purely relying on a power-law fit to the degree distribution to identify scale-free networks may not be wholly sufficient [11]. Instead, a more comprehensive assessment should involve determining whether a network exhibits a "fat-tailed" distribution.

2.1.3 Small-world Networks

Definition 2.3 (Watts-Strogatz Small-World Network [4])**.** *Let $G = (V, E)$ be a connected graph (network) with N nodes. Then graph G is a small-world network if the average path length (L) between any two nodes grows logarithmically with respect to the number of nodes in the graph, i.e., $L \propto \log N$.*

Small-world networks are characterized by two distinct properties: a high clustering coefficient and a short average path lengths [4]. The average path length represents the average number of steps it takes to travel between any two nodes in the network, while the clustering coefficient measures the density of connections within local neighborhoods. In small-world networks, the average path length grows logarithmically with the size of the network, indicating that information can propagate efficiently across the network [17]. Simultaneously, the high clustering coefficient provides redundancy and robustness to a network [18]. This distinctive combination of properties in small-world networks results in efficient information transfer and communication, making them a valuable model for understanding the structure of real-world networks.

2.2 Centrality measures

Centrality is a way to measure how important a node is in a network. While many centrality measures initially come from social network analysis, a field of social sciences, they have found broad applications in fields like computer science, physics and biology [7]. For instance, in computer science, centrality metrics play crucial role in understanding and optimizing network structures, aiding in tasks such as search algorithms and recommendation systems [7]. In social science, centrality metrics play a crucial role in network analysis, allowing researchers to identify and understand key individuals or entities within social networks [7]. Furthermore, in biology, these metrics prove valuable in deciphering molecular interactions and understanding the importance of specific proteins in cellular networks, contributing to advancements in fields like genomics and bioinformatics [7]. While there are several centrality measures, each capturing a different aspects of node's importance, the simplest one is degree dentrality [19].

2.2.1 Degree Centrality

Degree centrality, a key aspect of network analysis, traces its history to Freeman's pivotal work in 1978 [19]. In this seminal paper, Freeman not only introduced the concept of degree centrality but also navigated through various definitions, notably addressing the previously named "point centrality" [19]. Freeman recognized the concept of point centrality based on the degree of a point (node) from earlier studies [20, 21, 22, 23, 24, 25], but he found that the existing measures were either overly complex or overly restrictive. Instead, Freeman adopted a simpler version proposed by Nieminen in [26, 27], which counts the number of adjacencies (connections) a node has [19].

In Freeman's standardization [19], a graph denoted by $G(V, E)$ consists of a set of vertices V and a set of edges E. For a given vertex v within the network, $N(v)$ represents the set of adjacent neighbors of vertex v, and the degree of vertex v is defined as the cardinality of its neighborhood: $deg(v) = |N(v)|$. The degree centrality $(C_D(v))$ of a vertex v is then given by:

$$C_D(v) = \frac{deg(v)}{n-1} \tag{2.1}$$

Here, n represents the total number of nodes in the network. This approach to degree centrality provides a straightforward measure, quantifying the relative importance of a node based on its number of connections. Freeman [19] argued that degree centrality is a valuable measure of a node's importance because it reflects its ability to reach other nodes in the network.

Despite being the simplest centrality measure, degree centrality has a wide range of applications in various fields. It is used in social networks to identify the most connected individuals [28]. In the field of biology, degree centrality aids in pinpointing highly interacting proteins, shedding light on crucial components in biological systems [29]. Its simplicity and versatility make it a cornerstone of network analysis, facilitating insights into network structures and functions.

2.2.2 Betweenness Centrality

Freeman [30] introduced the idea of betweenness centrality, building upon the concept originated by Alex Bavelas in 1948 and Subsequent studies [31, 20, 32]. Bavelas [31] and Shaw [20] argued that a node is more central if it lies on more shortest paths between pairs of other nodes, while Cohn and Marriott [32] stated that this central nodes play important role in the network [28]. Later, Freeman [30] and Anthonisse [33] independently developed quantitative measures of betweenness centrality. Furthermore, Wasserman and Faust [28] provided a standardization of betweenness centrality.

Given a graph $G(V, E)$, where V is the set of vertices and E is the set of edges, let P_{st} represent the total number of shortest paths between two distinct vertices s and t in the graph. Further, let $P_{st}(v)$ denote the number of such paths that pass through a specific vertex v. Then, the betweenness centrality $C_B(v)$ of a vertex v is defined as follows:

$$C_B(v) = \sum_{s \neq t, s \neq v, t \neq v} \frac{P_{st}(v)}{P_{st}} . \frac{2}{(n-1)(n-2)} \tag{2.2}$$

Here, n represents the total number of vertices in the graph.

The betweenness centrality quantifies the extent to which a vertex v lies on the shortest paths between other pairs of vertices, reflecting its importance in controlling information flow within the network [30]. Nodes with high betweenness centrality act as bridges or intermediaries in the network, playing a critical role in connecting different parts of the network. This concept is particularly valuable in understanding the flow of information, influence, or resources in various networked systems.

Betweenness centrality finds widespread application across diverse domains. Newman applied betweenness centrality to identify influential actors in collaboration networks, suggesting that these individuals play a crucial role in connecting different sub-groups and facilitating the flow of information [34]. Betweenness centrality was used to identify influential individuals in egocentric networks, highlighting their role in bridging ties within their social circles [35]. The study [36] used betweenness centrality to identify critical nodes in public transport networks to assess the vulnerability of these networks against targeted attacks and develop strategies to improve their resilience. In [37] the authors identified and classified hubs in brain networks using betweenness centrality, suggesting that these hubs play a crucial role in integrating information from different brain regions and facilitating communication between them. This diverse application in fields ranging from collaboration networks to social circles, public transport systems, and even intricate brain networks, betweenness centrality stands as a versatile and invaluable metric.

2.2.3 Closeness Centrality

Freeman in [19] examined the various measures of closeness centrality proposed in the previous literatures [38, 39, 40, 24] and [41] . He concluded that the simplest and most intuitive measure was the one introduced by Sabidussi in [41], which calculates the average geodesic distance (length of shortest path length) between a node and all other nodes in the network. Freeman further introduced standardized version from [39], which is now widely used in network analysis [19].

Consider a graph $G(V, E)$, where V is the set of vertices and E is the set of edges. If $d(v_i, v_j)$ represents the shortest path distance between vertices v_i and v_j, where $i \neq j$, then according to [39] the closeness centrality ($C_C(v_i)$) of a vertex v_i is defined as follows:

$$C_C(v_i) = \frac{n-1}{\sum_{j=1}^{n} d(v_i, v_j)} \tag{2.3}$$

where n is the total number of vertices.

Closeness centrality is particularly valuable in understanding how quickly information or influence can spread through a network. It is especially important in scenarios where quick communication or access is crucial. Nodes with high closeness centrality may often found in social networks, biological networks, and transportation networks.

In social networks, a node with high closeness centrality might be a popular person who knows many other people. Newman in [34] applied closeness centrality to identify researchers who are centrally located in the collaboration network, emphasizing their proximity to other collaborators and highlighting their potential influence or efficient communication pathways within the scientific community. In biological networks, it is used to identify central nodes, such as in [42] the authors utilized closeness centrality to evaluate the connectivity of genes within gene regulatory networks, identifying central genes with potential regulatory significance. Closeness centrality is also valuable in transportation network analysis, where it identifies critical locations for optimizing travel routes and logistics. The authors in [36] used closeness centrality to determine the network's most central nodes, which in turn served as a basis for assessing the resilience of the public transport networks to targeted attacks. These diverse applications underscore the significance of closeness centrality in understanding and optimizing complex systems.

2.2.4 Eigenvector Centrality

Eigenvector centrality provides an alternative approach to assessing a node's significance in a network. Extending on the notion of degree centrality, it gives greater importance to nodes connected to other influential nodes, acknowledging that the contributions of neighbors to a node's significance in the network may vary [7]. Bonacich first introduced this concept in 1972 [43]. According to [6], eigenvector centrality is closely related to the previously proposed measure - influence measure - studied in [44, 45, 46, 47, 48].

Given a graph $G(V, E)$, where V is the set of vertices and E is the set of edges. If vertex u is the neighbor of vertex v and A_{vu} represents the elements of the network's adjacency matrix, then according to [49] the eigenvector centrality ($C_E(v)$) of a vertex v is defined as follows:

$$C_E(v) = \frac{1}{\lambda} \sum_u A_{vu} C_E(u) \tag{2.4}$$

Here, λ is a constant eigenvalue.

After rearranging we solve the following eigenvector equation:

$$\lambda v = Av \tag{2.5}$$

Where A is the adjacency matrix A_{vu} and v is the eigenvector.

Eigenvector centrality also has a wide range of applications in various domains. In social network analysis, it serves as a powerful tool for identifying influential individuals whose significance extends beyond their immediate connections [19]. This has implications for targeted marketing strategies, opinion leadership recognition, and modeling the spread of diseases within communities. Valente in [50] utilized eigenvector centrality to assess exposure scores and the flow matrix, capturing the maximum information transfer between individuals in the network. The versatility of eigenvector centrality makes it an invaluable tool for extracting key insights from complex networked systems.

Each centrality measure comes with its own set of advantages and limitations. Degree centrality, for instance, is straightforward and computationally efficient, but it might overlook significant connections within the network. Betweenness centrality excels at identifying pivotal bridges between nodes, but it doesn't directly account for regular connections. Closeness centrality prioritizes efficient communication pathways, yet it can be sensitive to the network's overall size. Eigenvector centrality considers a node's global influence, yet it might not perform optimally in networks with disconnected components.

While each centrality measure provides a unique perspective, combining multiple measures offers a more comprehensive understanding of network structure. Integrating degree, closeness, betweenness, and eigenvector centralities allows for a rigorous analysis that captures various aspects of node importance. This integrated approach overcomes limitations of individual measures and provides a richer characterization of network elements. It identifies nodes with significant connectivity, communication efficiency, bridging capacities, and global influence.

2.3 Node and Edge removal

Complex networks are both resilient and fragile, meaning that they can withstand some perturbations but are also vulnerable to others. Node and edge removal is a technique that can be used to study how these networks respond to perturbations. Node and edge removal analysis has a rich history in network science. By systematically removing nodes or edges, we gain a deeper understanding of network robustness and efficiency.

Modern research examines deeper into the intricate interplay between network structure, dynamics, and resilience under various removal scenarios. Additionally, advancements in machine learning and computational techniques have allowed for the development of predictive models to estimate the impacts of node and edge removal before they are executed, enhancing our ability to formulate proactive strategies for network optimization and management. We can use mathematical models, optimization algorithms, and simulations to study the removal of nodes and edges from a network and to predict the effects of these removal actions.

Researchers have proposed and explored diverse methods for removing nodes and edges, encompassing strategies like random removal [8, 18], as well as centrality-based removals such as degree centrality and betweenness centrality [51, 52]. A prevalent attack strategy involves pinpointing critical nodes based on metrics like degree or other centrality measures and systematically removing them in descending order of importance until the network either disconnects or loses essential properties [53]. Albert's investigation into the fragmentation of random and scale-free networks employed two distinct node removal strategies: targeted removal (attacks) and random removal (errors) [8]. In the research [54], the authors systematically eliminated genes from a biological network to examine the consequences on both network connectivity and behavior. Additionally, in the study [55], the authors performed a comparative analysis to evaluate the effects of diverse link removal strategies on various real-world complex networks. Additionally, numerous studies have examined the impact of removal strategies on complex networks across various scientific domains [56, 57, 58, 59, 60]. These methodologies provide a comprehensive understanding of how the removal of nodes and edges influences the overall structure and functionality of networks.

2.3.1 Impact on Network Connectivity and Structure

Network connectivity is essential for any network. It allows information, influence, resources, and even diseases to flow between the nodes of the network. In other words, network connectivity is what makes networks useful and powerful. Without connectivity, networks would be nothing more than a collection of isolated nodes [7]. Studying how removing nodes and edges from a network affects its ability to connect its nodes is not just a theoretical exercise rather, it can help us to understand how complex systems work and how to make them more resilient. We can identify the vulnerabilities of complex systems by identifying the nodes and edges that are essential for the network's functioning. We can also identify the types of failures that are most likely to occur in a network.

Moreover, structural changes within networks, whether through the addition or removal of nodes and edges or alterations in degree distribution, can profoundly impact their overall connectivity and efficiency [8, 61]. The effects of node and edge removal on network connectivity have been a subject of extensive research in the field of network science. Table 2.1 shows the studies on node and edge removal strategies and their implications on network connectivity.

After reviewing the work of these researchers in this field, we want to study further the importance of nodes in these networks. Additionally, we aim to explore how different node removal strategies, including targeted and random removal, influence network connectivity.

However, studying the impact of node and removal on different structures and connectivity is not without challenges. The interplay between local and global network structure presents complexities in predicting network behavior [53]. Furthermore, quantifying the impact of removal strategies on network robustness requires sophisticated algorithms and computational resources [61]. Moreover, the ethical implications of removal actions in real-world networks, such as social networks, introduce ethical considerations [62]. Our research aims to navigate these challenges and contribute to a more comprehensive understanding of network connectivity and structural changes due to node and edge removal.

Table 2.1: Studies of Node and Edge Removal Effects in Network Connectivity

Authors	**Focus**	**Findings**
Albert, Jeong, and Barabási [8]	Vulnerabilities of scale-free and random networks against high-degree (attack) and random (error) node removal.	Scale-free networks are resilient to errors but vulnerable against highly connected nodes' intentional removal.
Holme, Kim, Yoon and Han [53]	The vulnerability of complex networks to intentional attacks	Removal of highly connected nodes in complex networks, particularly scale-free networks, results in a rapid breakdown of network connectivity and functionality.
Bellingeri, Cassi and Vincenzi [63]	Evaluation of attack strategies in real-world networks	Targeted node removal based on degree and betweenness centrality disrupts a network connectivity and robustness more effectively.
Wandelt, Sun, Feng, Zanin and Havlin [51]	Comparative analysis of various network dismantling strategies, including node removal, link removal, and hybrid (node-link) removal.	node removal are generally more effective than link removal strategies, and that hybrid approaches that combine both node and link removal can be the most effective in some cases.
Callaway, Newman, Strogatz and Watts [64]	Resilience of complex networks to random failures and targeted attacks using percolation theory	Networks with a power-law distribution connections like the internet are robust to random failures but vulnerable to targeted attacks.
Chen and Li [65]	Community detection in complex networks using edge-deleting with restrictions.	Edge-deleting strategies with specific constraints can effectively strengthen communities in networks, ultimately leading to a more robust and interconnected network structure.

2.4 Statistical Tests in Network Analysis

In statistical hypothesis testing, the Cramér-von Mises (CvM) test and the Kolmogorov-Smirnov (K-S) test are two most popular non-parametric tests to evaluate the goodness-of-fit between two probability distributions [10].

2.4.1 Kolmogorov-Smirnov (K-S) test

The Kolmogorov-Smirnov (K-S) test is a statistical tool employed to evaluate whether two cumulative distribution functions (CDF) differ significantly [9, 10]. It quantifies the maximum vertical difference between the CDF of the dataset of original degree distribution and the CDF of reduced degree distribution of the network after removal of nodes or edges[9, 66, 67].

Let $F_{1,n}$ be the cumulative distribution function of original network with a number of nodes n and $F_{2,m}$ be the cumulative distribution function of the network after node or edge removal with a remaining number of nodes m, then the KS statistic, denoted as $D_{n,m}$, quantifies the maximum absolute difference between these functions, given by [10]:

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)| \tag{2.6}$$

In practice, the null hypothesis is rejected at a specified significance level α if $D_{n,m}$ exceeds a critical value or in other words, if the condition in following equation 2.7 satisfies [10].

$$D_{n,m} > c(\alpha) \cdot \sqrt{\frac{n \cdot m}{n + m}} \tag{2.7}$$

The value of $c(\alpha)$ at specified significance level is tabulated in [68]. The highest vertical difference between the two distribution functions is quantified by the K-S statistic ($D_{n,m}$). A significant difference between the original and modified degree distributions is shown by a high K-S statistic and a low related p-value in the context of network analysis, suggesting that the removal approach has significantly affected the network's structure.

The KS test is particularly suitable for comparing network degree distributions as it doesn't necessitate any assumptions regarding the underlying distribution of the data [11, 66]. The K-S test is exact and straightforward to interpret, delivering dependable outcomes even when dealing with small sample sizes [69]. This precision proves advantageous, especially in situations with constrained data [70, 7].

In the analysis of an evolving social network, Kossinets and Watts employed the Kolmogorov-Smirnov (K-S) test to compare the degree distributions of the network at different time points to identify and quantify the changes in the network's structure over time [70]. Their work highlighted the potential of the K-S test for analyzing the evolution of complex networks over time.

2.4.2 Cramér-von Mises Test

The Cramér-von Mises (CvM) test is another statistical method used for comparing the empirical cumulative distribution function (ECDF) of the sample to the cumulative distribution function (CDF) of a hypothesized distribution. The Cramér-von Mises W^2 criterion is named after Harald Cramér [71] and Richard Edler von Mises [72]. According to Anderson [73], the Cramér-von Mises W^2 criterion for testing that a sample, $x_1, x_2, \ldots$, has been drawn from a specified continuous distribution $F(x)$ is:

$$W^2 = \int_{-\infty}^{\infty} [F_N(X) - F(x)]^2 \, dF(x),$$

where $F_N(X)$ is the empirical distribution function of the sample. For a second sample, $y_1, y_2, \ldots, y_M$, a test of the hypothesis that the two samples come from the same (unspecified) continuous distribution can be based on the analogue of W^2, namely

$$T = \frac{NM}{N+M} \int_{-\infty}^{\infty} [F_N(x) - G_M(x)]^2 \, dH_{N+M}(x),$$

where $G_M(x)$ is the empirical distribution function of the second sample and $H_{N+M}(x)$ is the empirical distribution function of the two samples together.

If the computed value of T gets higher than the tabulated critical values from [73] then the null hypothesis that the two samples come from the same distribution can be rejected in the favour of alternative hypothesis. The CvM statistic (T) provides a measure of dissimilarity between the two samples.

It is general understanding that the Cramér-von Mises (CvM) test is considered a good choice for comparing distributions with heavy tails compared to the Kolmogorov-Smirnov (K-S) test due to its sensitivity to deviations in the tails of the distribution where crucial information about the network's structure is often contained [74]. There is only few studies that have employed the K-S test as the statistical tool to assess the differences in the degree distribution of the network and none to our knowledge that have employed CvM test.

3 Methodology

3.1 Generation of Complex Networks

We generated and visualized a different types of networks using the NetworkX [75] and Matplotlib [76] libraries from Python, respectively.

3.1.1 Random Networks

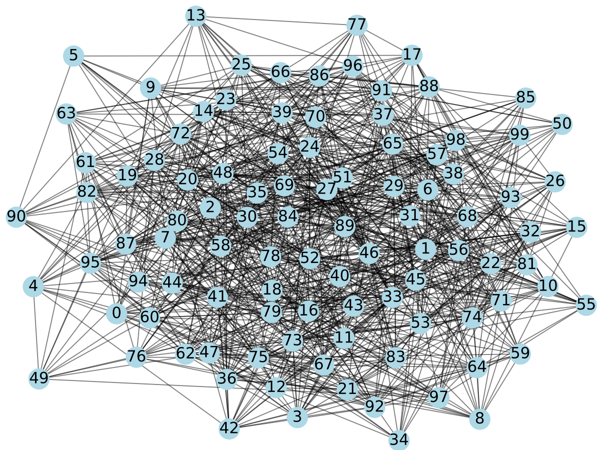


Figure 3.1: Erdős-Rényi Random Network.

We used Erdős-Rényi (ER) graph generator (`nx.erdos_renyi_graph()`) [2, 3] from NetworkX library [75] in Python to create a random network, see Figure 3.1. We define its parameter in the implementation as a number of nodes $n = 100$ for the network. Additionally, we specified the edge probability $p = 0.2$, which is the likelihood of an edge forming between any pair of nodes. Higher values of p result in denser networks, while lower values produce sparser ones. Here, Python's random state function is used to create a random state object and set a *seed* = 1 for consistent results. This step guarantees that the same network is generated each time the code is executed.

3.1.2 Scale-free Networks

The Barabási-Albert (BA) model is a well-known method for generating scale-free networks [5]. This model employs preferential attachment, where new nodes connect to existing nodes with high degrees, resulting in a power-law degree distribution. To create a scale-free network using this model, we define the number of nodes (n) and the parameter for preferential attachment (m). The parameter m in the BA model governs the number of edges added for each new node introduced, determining the network's structure. In practice, different values of m are employed to emulate diverse real-world networks, where m can signify, for instance, the number of collaborations for scientists in co-authorship networks [34], the number of followers for influencers in social networks [77], or the number of citations for scientific papers in citation networks [78].

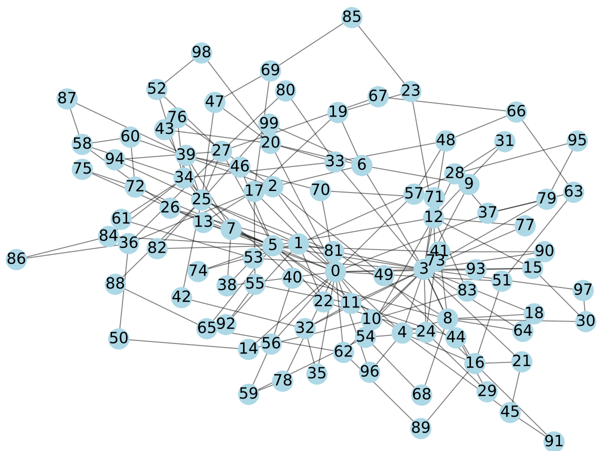


Figure 3.2: Barabási-Albert Scale-Free Network.

In our implementation, we utilized the graph generator `barabasi_albert_graph()` from Networkx library in Python to create a scale-free network [5], see Figure 3.2 and we set the parameters, number of nodes $n = 100$ and preferential attachment parameter $m = 2$.

3.1.3 Small-world Networks

The Watts-Strogatz (WS) model is a method for creating small-world networks that exhibit both the small-world characteristic and a high clustering coefficient [4]. The procedure to generate a small-world network starts with a ring with N nodes, in which each node is connected to its k nearest neighbors. Then, for each node, each link connected to a clockwise neighbor is rewired to a randomly chosen node with a probability p [4, 1]. Note that a regular lattice is obtained when $p = 0$, while a random graph with the restriction that each node has a minimum connectivity is produced when $p = 1$ [4]. The procedure generates graphs with the small-world characteristic and a non-trivial clustering coefficient for intermediate values of p[4].

In our implementation, we used the `watts_strogatz_graph()` graph generator of NetworkX library [75] in Python to create a small-world network [4] using the number of nodes $n = 100$, average degree $k = 4$, and edge rewiring probability $p = 0.2$, see Figure 3.3.

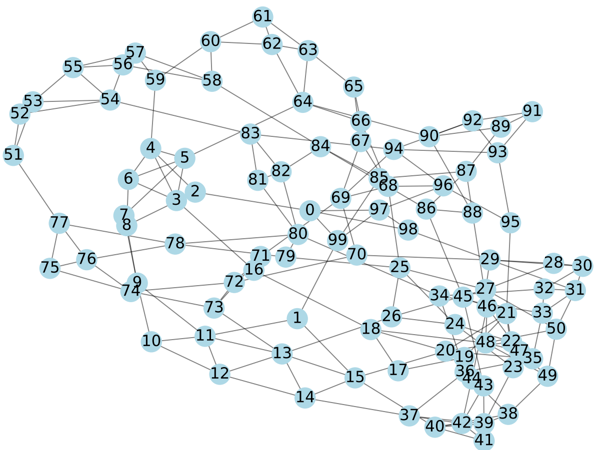


Figure 3.3: Watts-Strogatz Small-World Network.

3.2 Selection of Centrality measures

We focus on four fundamental centrality measures—degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality—in our analysis due to their distinct and complementary perspectives on node importance within networks.

In our analysis, to compute degree centrality, we utilize the NetworkX [75] library in Python, which provides a straightforward function, `nx.degree_centrality()`, to calculate the proportion of a node's connections to all possible connections [19]. Betweenness centrality is calculated using `nx.betweenness_centrality()` function, which identifies nodes acting as key intermediaries by quantifying the number of shortest paths they lie on [30]. Closeness centrality is computed using `nx.closeness_centrality()`, which measures how quickly a node can reach all others in the network [19]. Finally, eigenvector centrality is calculated using `nx.eigenvector_centrality()`, taking into account both a node's direct connections and the centrality of its neighbors, identifying nodes with global influence [43]. By applying these functions, we obtain centrality scores for each node, aiding our analysis of node importance and influence within the network.

Each of these centrality metrics offers a unique lens through which we can examine node significance, collectively providing a comprehensive analysis. Their versatility and established significance make them the ideal choice for our study of node importance in the network.

3.3 Simulating Node and Edge removal

3.3.1 Node removal

Two different approaches can be employed here to remove nodes from a network: targeted node removal and random node removal.

Targeted node removal involves strategically removing nodes based on their attributes or characteristics [8]. For instance, high-degree node removal involves removing a predefined fraction of nodes with the highest degrees from the network. This approach may help to assess the network's resilience to the intentional removal of its most connected nodes, revealing the crucial role these nodes play in maintaining network connectivity and information flow. Conversely, low-degree node removal focuses on nodes with the lowest degrees. By removing a fraction of these lowest degree nodes from the network, we evaluate a network's vulnerability to the loss of its least connected components and gain insights into their potential role in connecting different parts of the network.

By employing a combination of high-degree and low-degree node removal, this approach allows for a comprehensive exploration of the network's response to targeted disruptions, from the removal of highly central hubs to the removal of less-connected nodes, providing valuable insights into its resilience, connectivity and structure robustness.

Random node removal, another approach that we are using is a common strategy for node removal [8]. In this strategy, nodes are selected for removal entirely at random from the network. This approach does not take into consideration the specific characteristics of nodes or their roles within the network. Instead, it mimics a scenario where nodes fail or are removed without any particular pattern. Random node removal serves as a baseline for understanding how a network responds to unpredictable disruptions. It allows us to assess the network's overall robustness, as well as how its structural properties change when nodes are removed randomly. This strategy is valuable in scenarios where node failures or removals are stochastic, and it provides insights into the network's resilience against unforeseen events.

3.3.2 Edge removal

Another approach in network analysis involves the deliberate removal of edges rather than nodes. Edge removal strategies focus on the elimination of specific or random connections within the network, thereby exploring how alterations in the network's edge structure impact its overall functionality and integrity [55].

In the context of edge removal, one basic strategy random edge removal is considered. Random edge removal involves the stochastic elimination of edges throughout the network, reflecting situations where connections are lost without any predictions. This approach serves as a benchmark for assessing the network's resilience to random disturbances and may offer insights into its robustness when faced with unpredictable edge failures.

By considering random edge removal, this approach facilitates a comprehensive analysis of how changes in edge connectivity influence network dynamics, robustness, and information flow, contributing to a more profound comprehension of network structure and behavior.

3.4 Error measures

To asses the impact of node and edge removal strategies on network structures, we employ two fundamental error measures: degree distribution error and centrality error.

Degree distribution error is a key metric to gauge the impact of node and edge removal strategies on network structures. It quantifies the dissimilarity between the degree distribution of the original network and that of the network after distinct removal strategies. We rely on the Kolmogorov-Smirnov (K-S) and Cramér-von Mises (CvM) test to statistically evaluate the differences in degree distributions.

Centrality error quantifies the disparity in centrality scores (e.g., degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality) between the original network and the network following the application of removal strategies [6]. This metric may provide insights into how critical nodes or edges are affected by the removal actions, thereby shedding light on the network's resilience and functional changes. For computation, we utilize NetworkX [75], a Python library for network analysis, to calculate centrality measures. The error in centrality is quantified by taking the absolute difference in centrality scores before and after removal.

These approaches are consistent with established methodologies in network science [79, 80, 11, 8, 18]. These error measures collectively may provide insights into how different removal strategies impact network structure, resilience, and the relative importance of nodes within the network.

3.5 Upper and Lower Bounds of Degree Centrality after Node removal

Theorem 3.1 (Degree Centrality Bounds)**.** *Let $G(V, E)$ be an undirected graph, where V is the set of vertices and E is the set of edges. v be a vertex in graph G with degree centrality $C_D(v) = \frac{deg(v)}{n-1}$ (eq. 2.1), where $deg(v)$ represents the degree of vertex v (number of neighbors $|N(v)|$) and n is the total number of vertices. If r_n is the removed number of nodes from graph G, where $0 \leq r_n \leq n-1$, then (1) the upper bound coefficiant (a) is given by:*

$$a = \frac{n-1}{n-r_n-1} \tag{3.1}$$

and (2) the lower bound coefficient (b) is given by:

$$b = \max(0, \frac{deg(v)-r_n}{deg(v)} \cdot \frac{n-1}{n-r_n-1}). \tag{3.2}$$

Proof (1):
In the best-case scenario where vertex v remains unconnected to the removed nodes r_n and $deg(v)$ remains unchanged, while the total number of nodes decreases to $n - r_n - 1$. This leads to the new degree centrality of a node after removed nodes,

$$C_D(v') = a \cdot C_D(v)$$

$$\frac{deg(v)}{n - r_n - 1} = a \cdot \frac{deg(v)}{n - 1}$$

Which leads to the the upper bound coefficient a in (eq.3.1).

Using this, we can make an estimation of the upper bound for degree centrality as follows:

$$C_D(v') \leq \frac{n-1}{n - r_n - 1} \cdot C_D(v) \tag{3.3}$$

Proof (2):
In the worst-case scenario where vertex v was connected to each of the removed nodes r_n and, the vertex degree $deg(v)$ reduces to $deg(v) - r_n$, and the total number of nodes decreases to $n - r_n - 1$. Here, the updated degree centrality of vertex v after removal, denoted as $C_D(v')$, is given by:

$$C_D(v') = b \cdot C_D(v)$$
$$\frac{deg(v) - r_n}{n - r_n - 1} = b \cdot \frac{deg(v)}{n - 1}$$

This leads to the the lower bound coefficient as follows:

$$b = \frac{deg(v) - r_n}{deg(v)} \cdot \frac{n-1}{n - r_n - 1} \tag{3.4}$$

Here, the lower bound coefficient can not be accurate in the case where the number of removed nodes is higher than the node degree (number of neighbors), resulting in a negative lower bound coefficient. To prevent this we set the lower bound coefficient as the maximum between 0 and the value we get from the equation 3.4.

This leads to the final lower bound coefficient b is as in 3.2.

From this, we can estimate the lower bound of degree centrality as follows:

$$\max(0, \frac{deg(v) - r_n}{deg(v)} \cdot \frac{n-1}{n - r_n - 1}) \cdot C_D(v) \leq C_D(v') \tag{3.5}$$

Combined Inequality: The bounds of degree centrality $C_D(v')$ after node removal can be expressed using the following inequality:

$$\max(0, \frac{deg(v) - r_n}{deg(v)} \cdot \frac{n-1}{n - r_n - 1}) \cdot C_D(v) \leq C_D(v') \leq \frac{n-1}{n - r_n - 1} \cdot C_D(v) \tag{3.6}$$

This statement and the corresponding proofs provide a mathematical foundation for estimating the upper and lower bounds of degree centrality after node removal.

These bounds of degree centrality, encompassing upper and lower limits, serve as crucial indicators for assessing the resilience and vulnerability of individual nodes within a network. These bounds offer insights into how robust a node's centrality is when subjected to various forms of network disruptions, such as random or targeted node removal. A high upper bound suggests that a node plays a pivotal role in maintaining network connectivity even in adverse scenarios, while a lower bound indicates vulnerability to the network connectivity.

3.6 Statistical Test procedures for Structural changes in Complex Networks

3.6.1 Data Collection

The data collection using the Kolmogorov-Smirnov (K-S) and Cramér-von Mises (CvM) test involves several key steps. First, a random network, a scale-free network, or a small-world network is generated using the methodology from Subsections 3.1.1, 3.1.2 and, 3.1.3 respectively. The degree frequency distribution of this original network was computed, providing the reference data. This distribution provides insights into the network's degree structure and is essential for comparison with the degree structure of a network after removal of nodes or edges.

To simulate node or edge removal, we removed a fraction of nodes or edges from the different networks using different removal stratagies introduced earlier, see Section: 3.3. After the removal of nodes or edges using these strategies, the frequency distribution of node degrees in the reduced network (network after removal of nodes or edges) is calculated. This step captures the degree distribution of the network after node or edge removal, which is a critical aspect of the analysis. The K-S test and CvM test are then applied to compare the degree sequence of the original network with that of the reduced network.

3.6.2 Hypothesis Testing

The K-S and CvM test is a fundamental statistical method employed to assess whether two datasets of degree sequence before and after removal strategies have the similar degree distribution. It is defined by:

H_0 : The degree distributions of the original and reduced networks have the similar distribution.
H_1 : The degree distributions differ between the original and reduced network.

By conducting these statistical tests, we aim to either fail to reject the null hypothesis, indicating similarity in degree distributions, or reject it in favor of the alternative hypothesis, showing dissimilarity.

3.6.3 Significance Level

In these statistical tests implementation, a significance level (α) of 0.05 is chosen [13]. The selection of this significance level is a common practice in statistical hypothesis testing [81]. It represents the probability of making a Type I error, which is the rejection of the null hypothesis when it's actually true. The choice of $\alpha = 0.05$ is often considered standard and offers a balance between the risk of making a Type I error and the ability to detect meaningful differences. It was chosen based on conventions in statistical hypothesis testing and the specific research context [81].

3.6.4 Interpretation

The interpretation of the K-S and CvM test results involves assessing the p-value obtained from these tests. If the p-value is higher than the significance level ($\alpha = 0.05$), we conclude that the degree distributions of the original and reduced networks are not statistically significantly different. Conversely, if the p-value is lesser than α, we reject the null hypothesis, indicating that the degree distributions are significantly dissimilar. This interpretation guides our understanding of the structural changes in the network due to nodes or edges removal.

To visualize the results, both the original and reduced degree frequency distributions are plotted and the associated statistical results are presented in the associated tables, providing a clear understanding of the observed changes. This comprehensive methodology allows for a assessment of the impact of node and edge removal on the degree distribution of these networks, providing valuable insights into network resilience and robustness.

4 Results and analysis

This chapter is divided into four sections, each section answers one of the research questions discussed above.

Firstly, we present the results of our analysis of node importance in random, scale-free, and small-world networks. We used different centrality measures to assess node importance and how these centralities are correlated.

Second, we implemented different node and edge removal strategies to these networks and measured the change in centrality measures between before and after removal.

Third, we provided the methodology to predict the upper and lower bounds of degree centrality after node removal as discussed, see Section 3.5, and verified our predictions using the actual degree centrality values after removal.

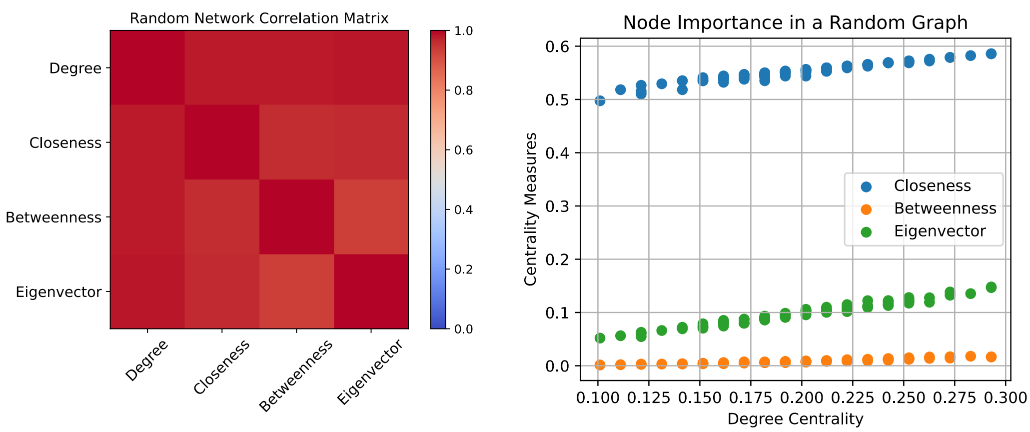
Finally, we present our findings on how network connectivity changes after node removal, and we will introduce two non-parametric tests to assess the changes in degree distribution in these networks.

4.1 Node importance in Complex Networks

The three complex networks under analysis, are a random network (see Subsection 3.1.1), a scale-free network (see Subsection 3.1.2), and a small-world network (see Subsection 3.1.3) were created following the methods that were outlined earlier. We focused on understanding the concept of node importance by examining the correlation matrix and scatter plots of various centrality metrics. Specifically, we evaluated degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality to assess the importance of individual nodes within the network.

The correlation matrix represents how the different centrality measures are related to each other. A strong correlation between two measures indicates that nodes with high values in one measure tend to have high values in the other. This can help us to understand how the different measures are interrelated and how they can be used together to assess node importance. In the scatter plot, each point corresponds to a node, and its position is determined by its degree and other centrality values. This analysis can be used to identify important nodes in the network and to understand how they contribute to the network's overall structure.

In our analysis of a random network generated using the Erdős–Rényi model [2], when we calculated the correlations between various centrality measures, the results showed very strong correlations between centrality metrics, see Fig. 4.1a, signifying the interplay among nodes in this particular network. Notably, we observed a very strong positive correlation between degree centrality and betweenness centrality (0.977), indicating that nodes with higher degrees are pivotal in facilitating efficient communication pathways within the network. The correlation between degree centrality and eigenvector centrality was equally substantial (0.984), highlighting that nodes with a high number of connections tend to hold influential positions with a broad reach in the network. Additionally, the strong correlation of closeness centrality with eigenvector centrality (0.964) and betweenness centrality (0.959) suggests that nodes with high closeness centrality are often positioned at central locations and play a crucial role in connecting different communities in the network. These results show the relationships between these centrality measures in a random network, offering a deeper understanding of how nodes interact and influence network dynamics.



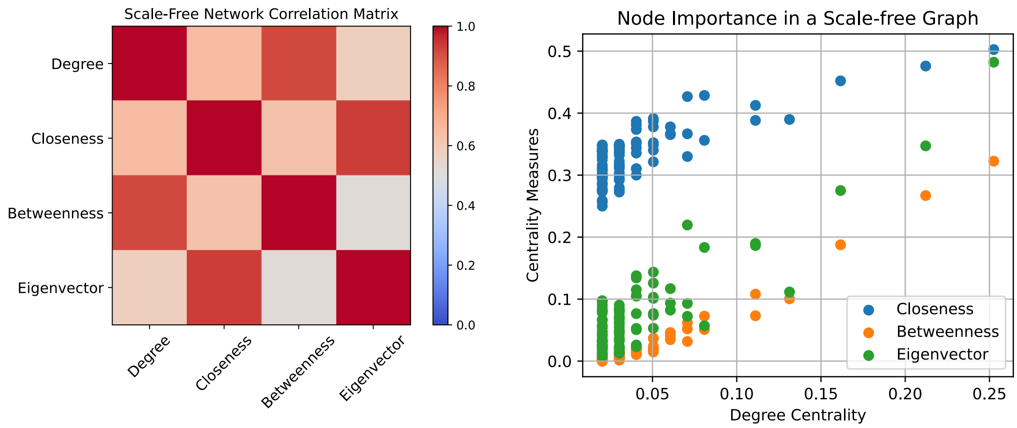
(a) Correlation matrix of centrality measures

(b) Scatter plot of centrality measures

Figure 4.1: Scatter Plot and Correlation Matrix of an Erdős–Rényi Random Network with $n = 100$ nodes.

To graphically represent the correlation between degree centrality and other centrality measures, a scatter plot, Fig. 4.1b, was created. The scatter plot indicates how the various metrics correlate in a random network by showing how the degree centrality values of nodes align with betweenness, eigenvector, and closeness centrality values. The positive correlation is evident across all three pairs of centrality measures. High-degree nodes typically have high values for all three centrality measures, demonstrating their importance in bridging multiple paths and efficiently transmitting information across a random network. This analysis underscores the significance of well-connected nodes ensuring network efficiency.

In our analysis of a scale-free network generated using the Barabási-Albert model [5], we calculated the correlations between the same centrality measures to gain insights into the relationships between nodes, see Fig. 4.2a. The strong positive correlation between degree centrality and betweenness centrality (0.917) suggests that high-degree nodes in a scale-free network play crucial roles in maintaining efficient communication pathways and facilitating the flow of information. Similarly, the positive correlation between degree centrality and eigenvector centrality (0.579) indicates that well-connected nodes in a scale-free network may hold influential positions within the network. Another strong positive correlation observed was between closeness centrality and eigenvector centrality (0.935), implying that nodes with high eigenvector centrality are well-connected and easily reachable from other nodes in the network. Additionally, closeness centrality was found to be positively correlated with betweenness centrality (0.631), indicating its role in bridging different parts of the network.



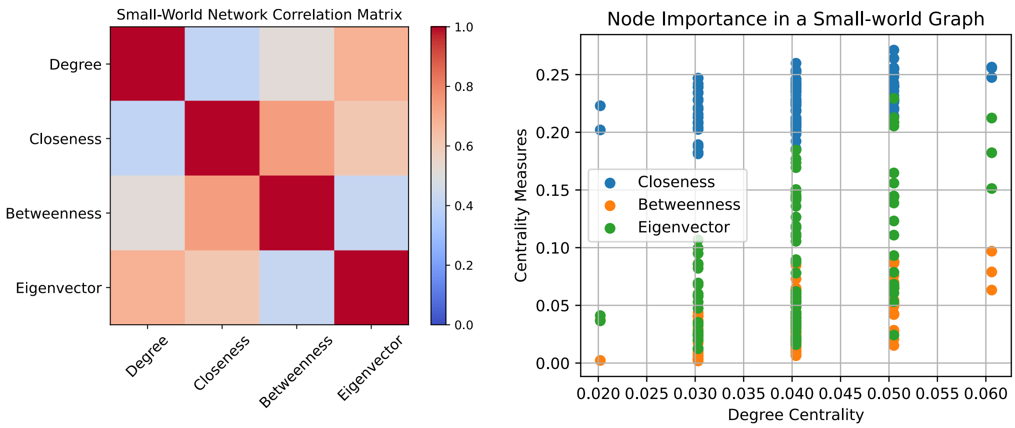
(a) Correlation matrix of centrality measures

(b) Scatter plot of centrality measures

Figure 4.2: Scatter Plot and Correlation Matrix of a Barabási-Albert Scale-Free Network with $n = 100$ nodes.

To visualize these findings, we employed a similar scatter plot that visually represents the correlations between degree centrality and other centrality measures, see Fig. 4.2b. This scatter plot offers a clear picture of how degree centrality is positively correlated with closeness centrality, betweenness centrality, and eigenvector centrality within a scale-free network. This analysis reaffirms the central role of highly connected nodes in ensuring network efficiency and connectivity.

For a small-world network, created using the Watts-Strogatz model [4], we calculated the correlations among these centrality measures, see Fig. 4.3a. A moderately strong positive correlation was observed between degree centrality and betweenness centrality (0.518), suggesting that highly connected nodes in a small-world network may play a crucial role in connecting different parts of the network. Additionally, the correlation between degree centrality and eigenvector centrality was strong (0.686), indicating that nodes with higher degrees often hold influential positions with a wide reach in the network. Closeness centrality exhibited the highest positive correlations with other measures in a small-world network, particularly with betweenness centrality (0.738) and eigenvector centrality (0.613), emphasizing its central role in bridging different parts of the network. While the correlations were not as strong as the other networks, they still indicated significant relationships between centrality measures in a small-world network.



(a) Correlation matrix of centrality measures

(b) Scatter plot of centrality measures

Figure 4.3: Scatter Plot and Correlation Matrix of a Watts-Strogatz Small-World Network with $n = 100$ nodes.

We created a scatter plot, shown in Fig. 4.3b, to visually represent the connections between degree centrality and the other centrality measures in order to better understand and interpret these findings. This scatter plot provides a clear, graphical representation of how degree centrality is related to closeness centrality, betweenness centrality, and eigenvector centrality within a small-world network.

In this part, the analysis involves identifying nodes of critical importance by evaluating various centrality measures, such as degree, closeness, betweenness, and eigenvector centrality. These measures provide insights into the roles played by different nodes in various aspects of network dynamics. In the subsequent section, we focus on the impact of removing these high-degree and influential nodes on the network's structure and functionality.

4.2 The Impact of Node and Edge Removal Strategies on Centrality Measures of Complex Networks

This section provides the results of our second research question on the impact of node and edge removal strategies on the same three types of complex networks. This section is divided into four subsections, three representing different node removal strategies and one representing random edge removal strategy. Each removal strategy is applied to a random network, a scale-free network, and a small-world network, all generated using the same parameters as described in the methodology section. The analysis includes box plots that show the absolute differences in degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. These differences are shown for both before and after applying the removal strategy, with removal percentages ranging from 5% to 25%.

4.2.1 High-degree node removal strategy

The analysis of an Erdős-Rényi (ER) random network after high-degree node removal, see Figure 4.4, indicates that the differences in degree centrality, closeness centrality and eigenvector centrality are higher in comparison to betweenness centrality and the differences increase as the percentage of node removal increases from 5% to 25%. The variation in degree centrality differences indicates the influence of these high-degree nodes in sustaining the network's overall connectivity. Simultaneously, closeness centrality experiences considerable shifts due to high-degree node removal. These differences highlight the network's sensitivity to alterations in the distribution of node importance and raise questions about the stability of random networks when subjected to the removal of highly connected nodes. These findings may contribute to our understanding of network resilience and vulnerability and provide insights into optimizing the robustness of diverse systems.

In the analysis of a scale-free network after high-degree node removal, a different response was observed, see Figure 4.5. Specifically, the decrease in betweenness centrality difference within a scale-free network as node removal percentages increase is indicative of the network's vulnerability to disconnection. In scale-free networks, high-degree nodes often serve as connectors between clusters, facilitating communication and information flow. Their removal disrupts the network's structure, reducing the number of crucial hubs and leading to increased path lengths and decreased network efficiency. This fragmentation of the network is reflected in the decreasing differences in betweenness centrality. The phenomenon is closely tied to the network's highest cluster and the decline in betweenness centrality differences implies the pivotal function of high-degree nodes in maintaining overall network connectivity and shows the consequences of their removal in scale-free

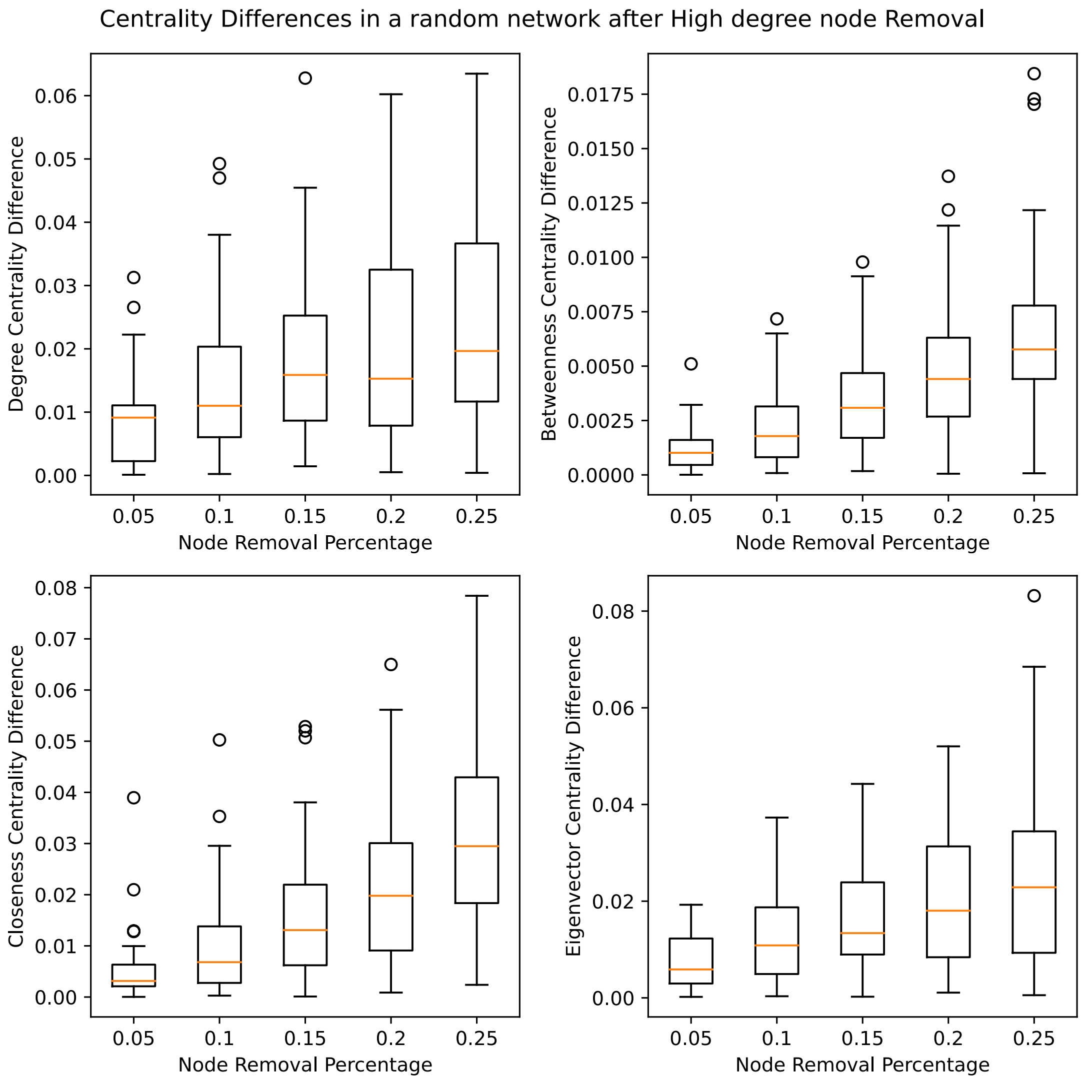


Figure 4.4: Impact of High-degree Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes.

networks, resulting in network fragmentation and reduced betweenness centrality.

Furthermore, As node removal percentages vary, the degree centrality difference increases. This behavior suggests that scale-free networks exhibit an intricate response to node removal, with the degree centrality measure experiencing fluctuations in importance. The differences in closeness centrality are much higher in scale-free networks. These differences substantially increase up to 15 percent node removal and then stabilize, remaining relatively constant as removal percentages continue to rise to 25 percent. The eigenvec-

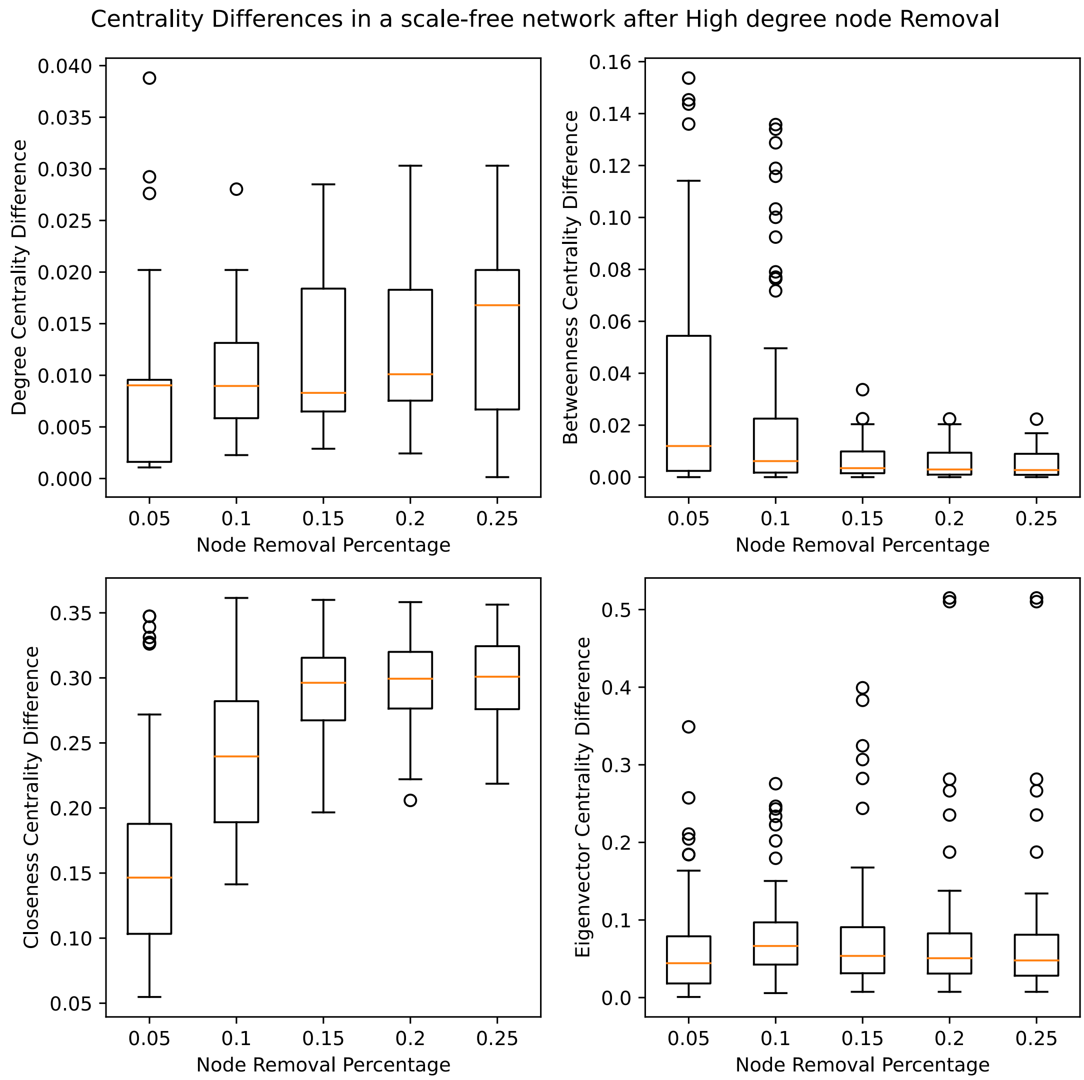


Figure 4.5: Impact of High-degree Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes.

tor centrality difference, while it may show some increases and decreases, it generally remains relatively stable throughout the removal process.

The box plots in Figure 4.6 present the differences in centrality measures before and after high-degree node removal in a small-world network. When examining degree centrality, we notice a consistent increase in the difference as the percentage of high-degree node removal increases. Betweenness centrality difference in a small-world network remains relatively unchanged with increasing node removal percentages, indicating the network's resilience to high-degree node removal. This resilience can be attributed to the network's structure and high clustering, which provide alternative pathways for information transmission even when high-degree nodes are removed.

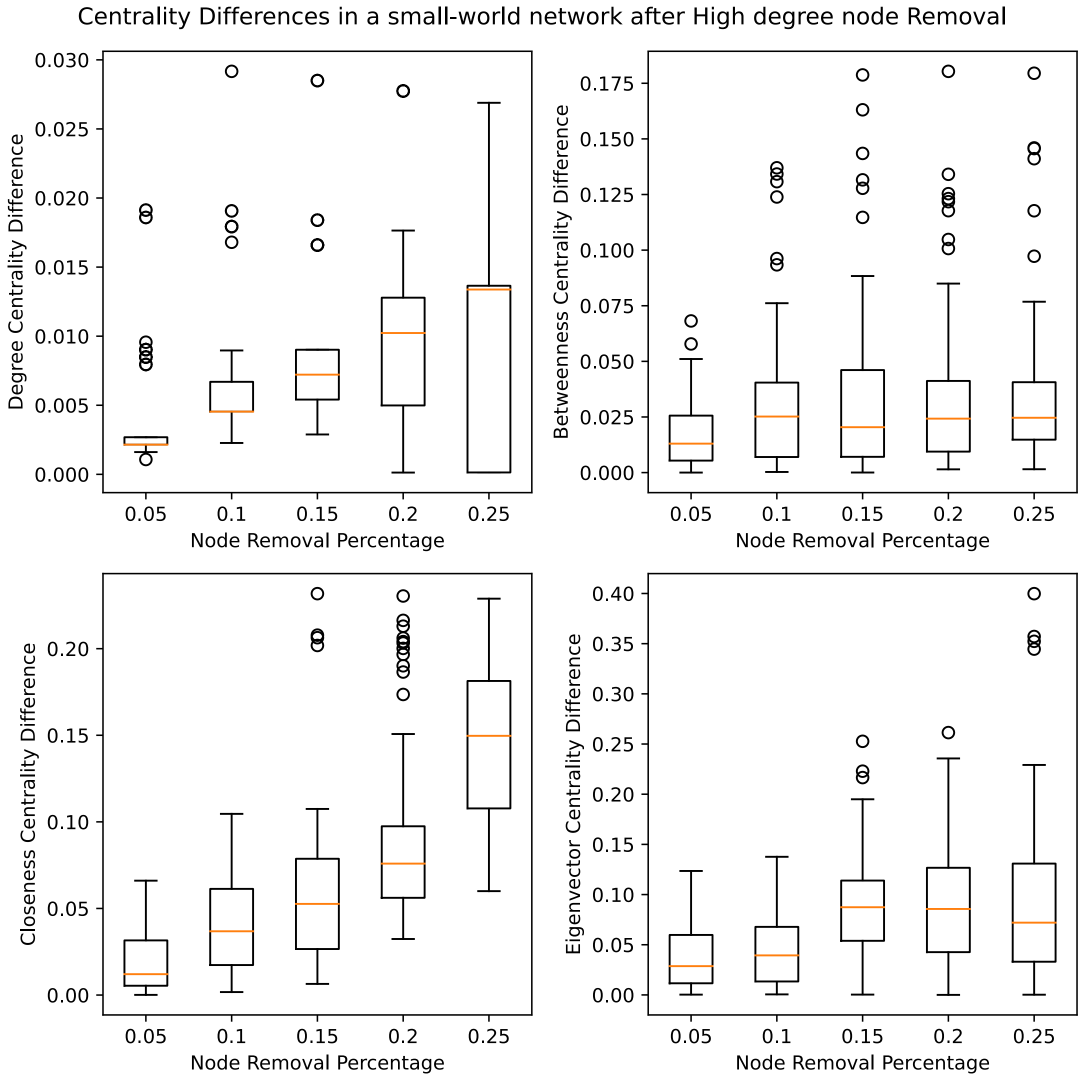


Figure 4.6: Impact of High-degree Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes.

When looking at closeness centrality, we observed that the differences are somewhat higher than those seen for degree and betweenness centrality. Closeness centrality differences increase gradually up to 20 percent node removal, and then a sudden and more pronounced rise in centrality difference is observed. This suggests the network's sensitivity to disruptions but also its ability to maintain closeness efficiency up to a certain threshold. Beyond this threshold, the sudden increase in differences indicates a more profound impact on the network's ability to transmit information efficiently. Eigenvector centrality differences in small-world networks also exhibit higher values. The average eigenvector centrality difference may vary at different removal percentages, but overall, they remain relatively stable during the removal process.

4.2.2 Low-degree node removal strategy

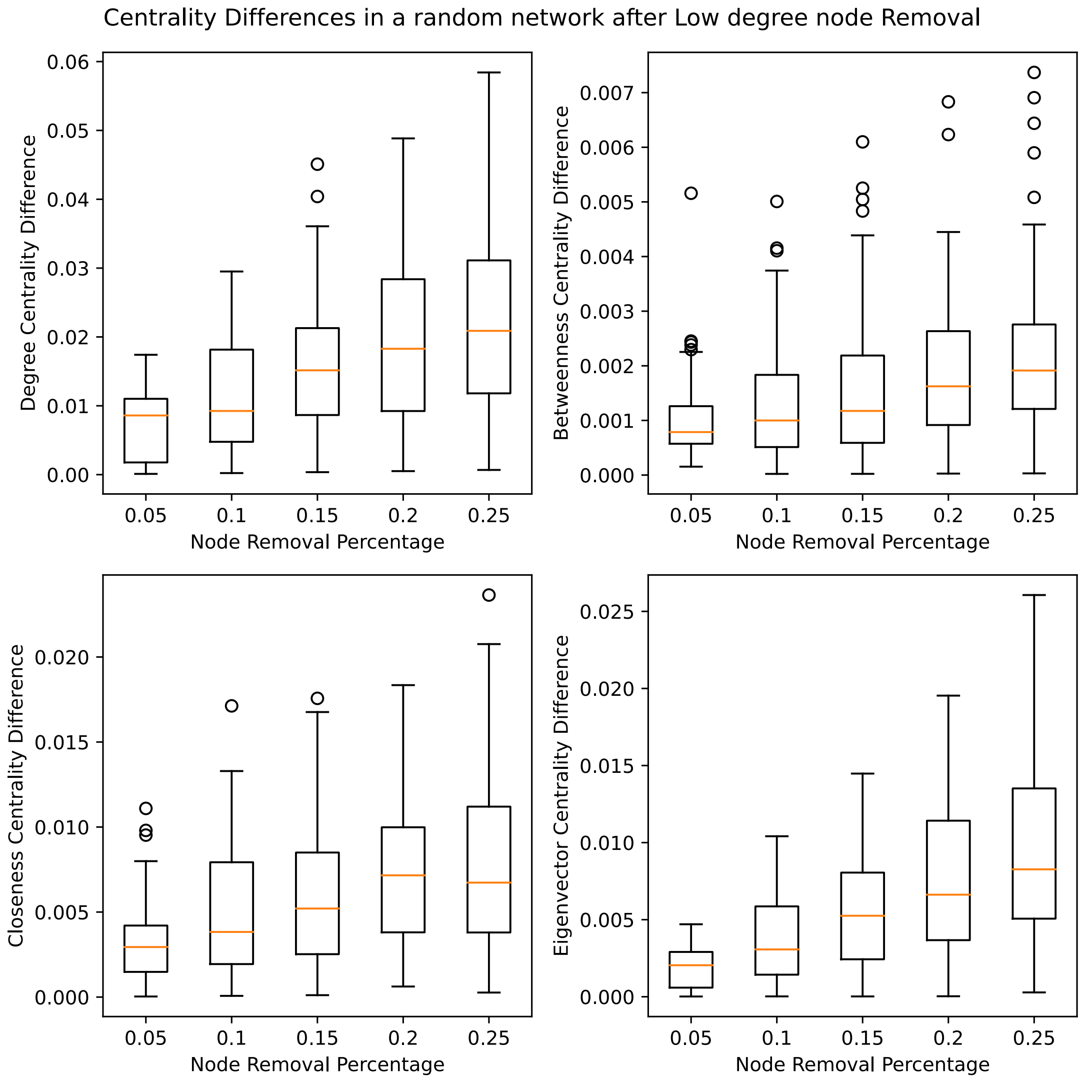


Figure 4.7: Impact of low-degree Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes.

The analysis of low-degree node removal in a random network, as shown in Figure 4.7, shows that as the percentage of low-degree nodes removed increases, the impact on degree centrality difference becomes more pronounced. This suggests that even nodes with a small number of connections play a crucial role in shaping the network's structure. Their removal leads to noticeable changes in the network's layout, demonstrating the importance of low-degree nodes in random networks.

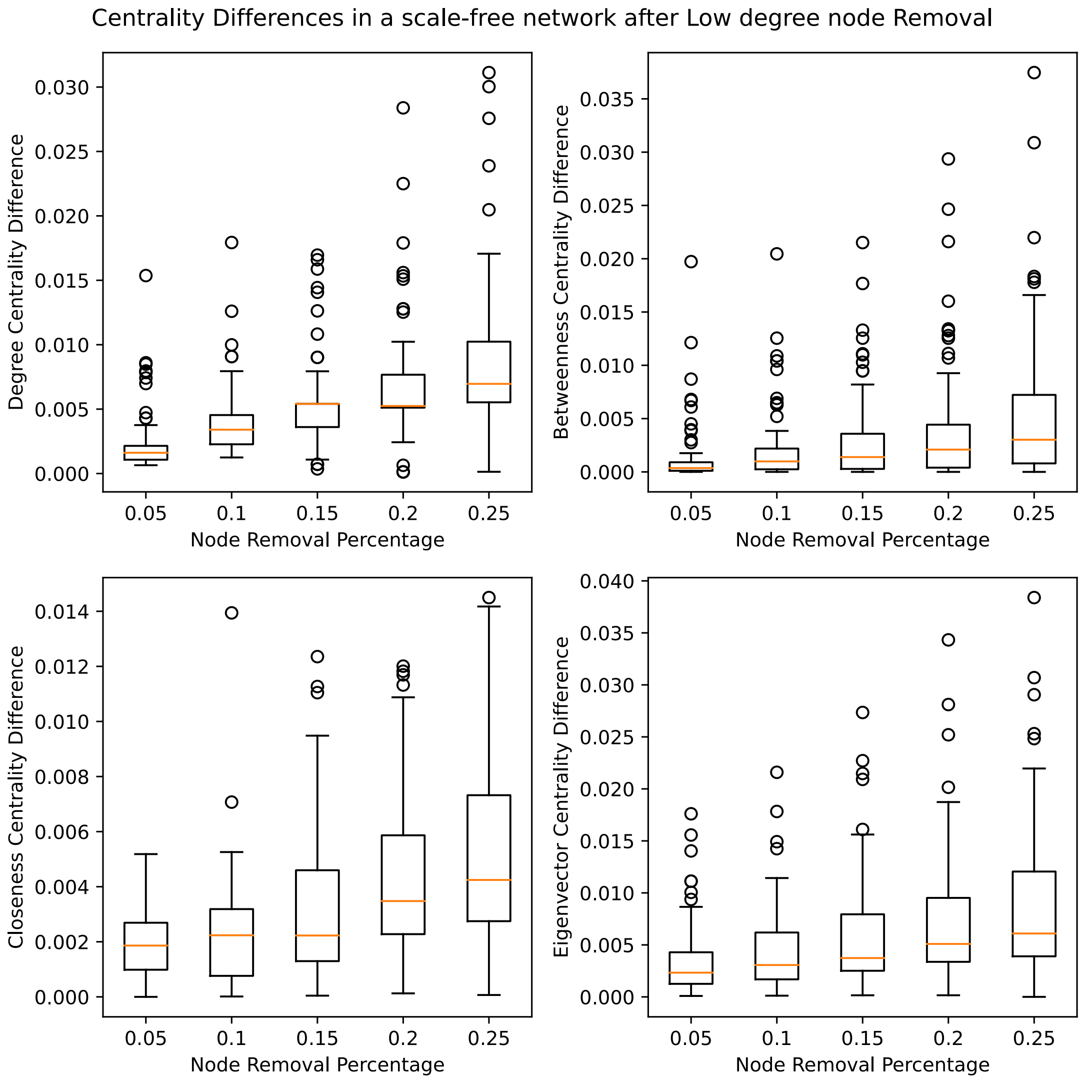


Figure 4.8: Impact of low-degree Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes.

The changes in other measures like betweenness, closeness, and eigenvector centrality were not as higher. These changes indicate, how a random network can withstand the removal of a small number of low-degree nodes without losing their structure. However, they can be made more robust to low-degree node removal by increasing the number of connections between nodes.

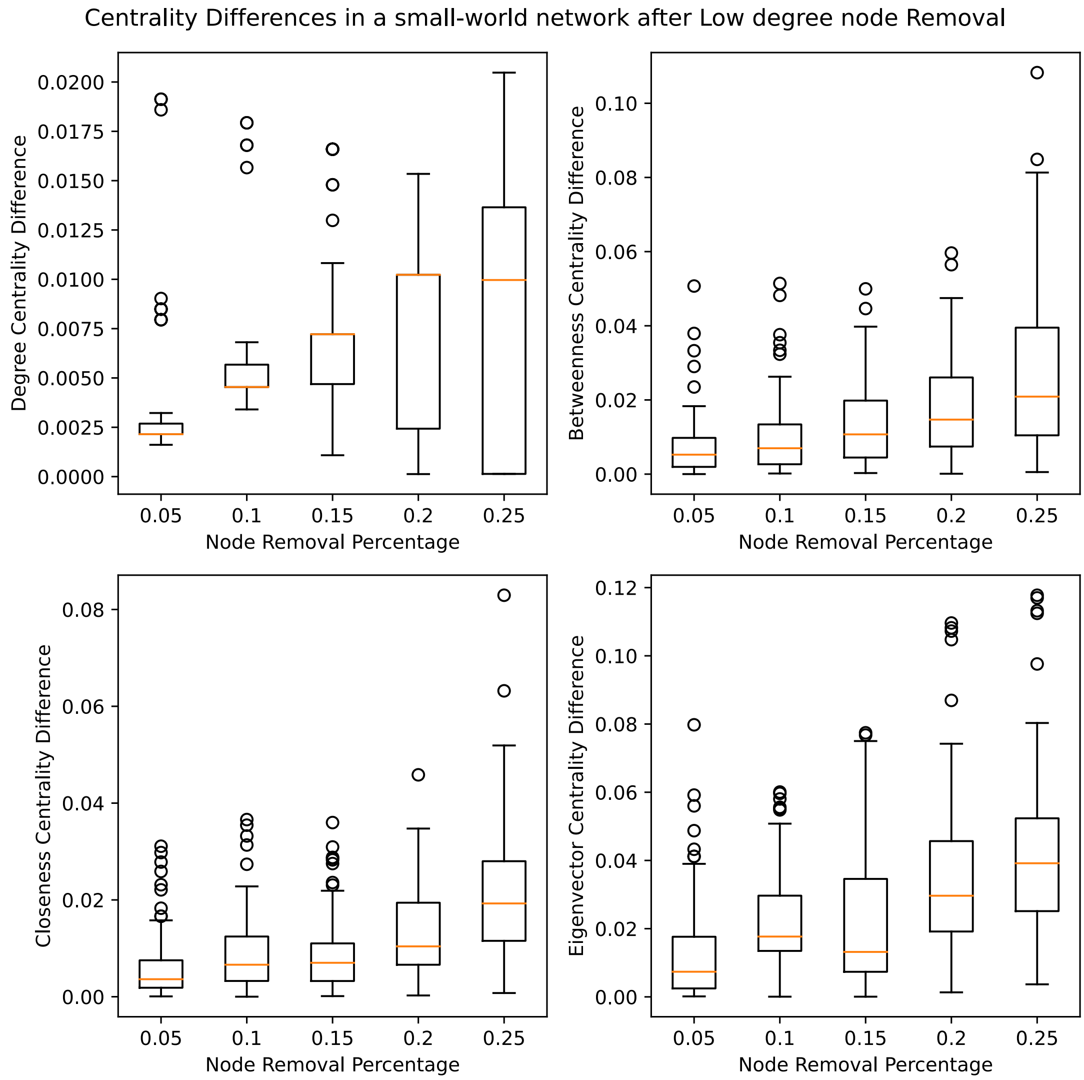


Figure 4.9: Impact of low-degree Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes.

In a scale-free network, as more low-degree nodes were removed from a network as presented in Figure 4.8, an analysis of the removal of these nodes showed a consistent increase in all centrality differences. Compared to a random network, betweenness centrality differences were higher. This is simply because too much removal of low-degree nodes can disrupt the topology of a scale-free network, making it more difficult for information to flow between nodes. When compared to the results of the targeted high-degree node removal strategy, centrality differences were smaller. In scale-free networks, the presence of high-degree nodes (hubs) can play a crucial part in the network structure, which accounts for their resilience. These hubs maintain connectivity even when many low-degree nodes are removed. For the network to continue to have an efficient information flow, this robustness is necessary.

In the analysis of low-degree node removal within a small-world network, from the box plots in Figure 4.9, we can see that as the removal percentage of low-degree nodes increases, the degree centrality difference between nodes increases, but the difference stayed minor. The betweenness centrality and closeness centrality differences moderately increased and the eigenvector centrality was more affected as more low-degree nodes were removed from a small-world network.

This analysis highlights the importance of nodes with relatively few connections in the network's structure. Despite a small-world network's shorter average path length, which facilitates rapid information flow, these nodes are also more vulnerable to disruptions. Low-degree nodes play a crucial role in maintaining a small-world network's connectivity and their removal can lead to network fragmentation.

4.2.3 Random node removal strategy

Random node removal is the third approach we employed to assess the impact of node removal on centrality measures in these networks. In the analysis of random node removal within a random network, see Figure 4.10, we observed that as the percentage of nodes removed increased, the degree centrality differences increased. In contrast, betweenness centrality differences were relatively minor, which indicates that random node removal is more likely to disrupt the network's connectivity than its communication paths. Closeness and eigenvector centrality differences moderately increased but not as higher as the degree centrality changes. This indicates that random node removal can disrupt the network's topology and make it more difficult for information to flow through the network, which can lead to decreases in closeness and eigenvector centrality.

Random networks are relatively resilient to the random node removal strategy. This may be because random networks typically have a high degree of redundancy, the presence of multiple alternative paths connecting any two nodes in the network. This redundancy makes it difficult for random node removal to completely disrupt the network's connectivity. However, random node removal can still have a significant impact on the resilience of random networks. Random node removal can reduce the number of connected components in the network, making it more vulnerable to further disruptions.

Our examination of random node removal within a scale-free network, see Figure 4.11,

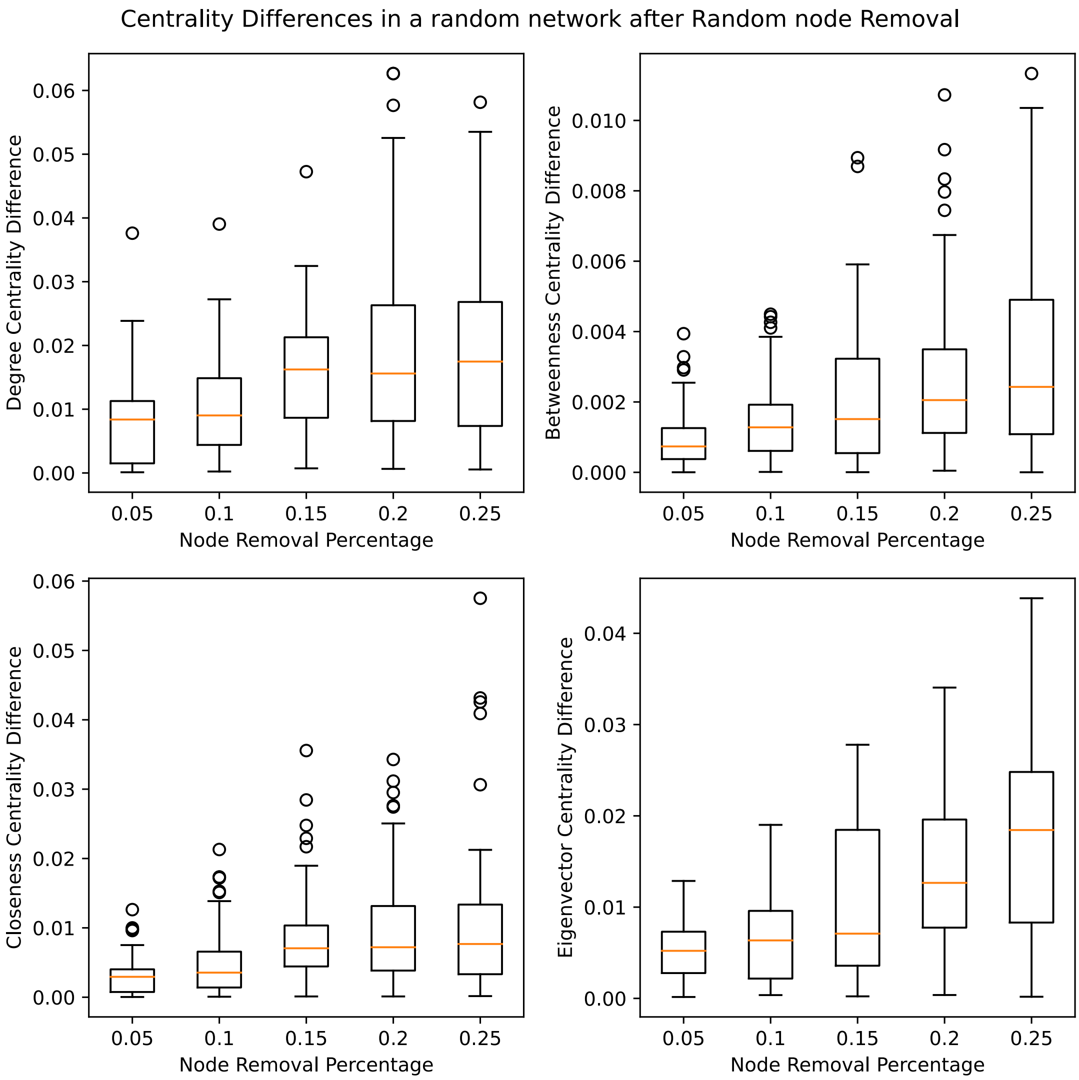


Figure 4.10: Impact of Random Node Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes.

revealed that as the proportion of nodes removed gradually increased, the differences in degree centrality remained relatively minor. The observation that the removal of random nodes in the early stages had a relatively minor impact on the network's structure can be attributed to the presence of hubs. In contrast to degree centrality difference, other centralities like betweenness, closeness, and eigenvector centrality exhibited something different. These centralities initially experienced a rise in differences up to the removal of 15% of nodes and after that the differences began to decrease. This indicates that beyond a certain threshold, the removal of random nodes started to disrupt the network's

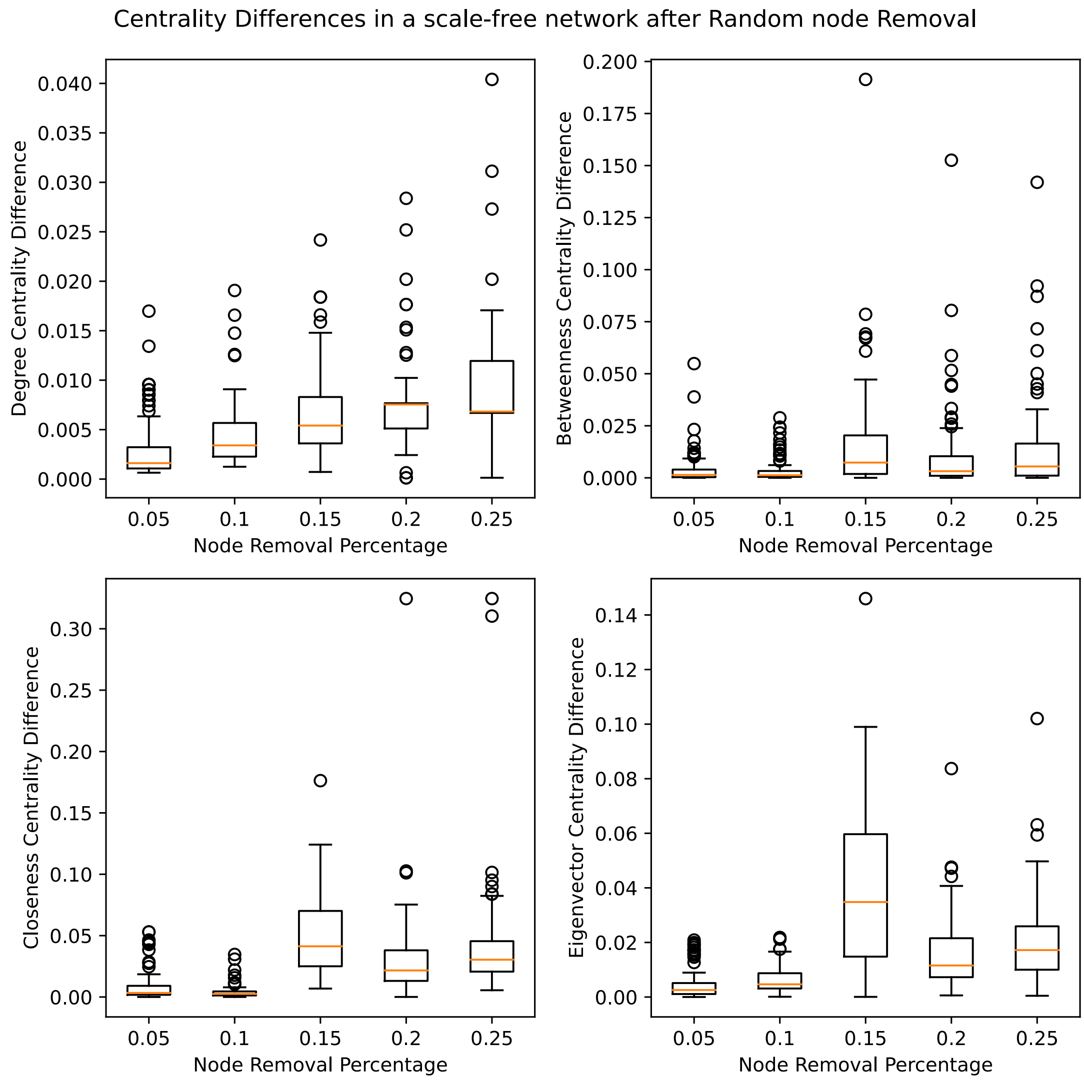


Figure 4.11: Impact of Random Node Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes.

connectivity, making it less efficient in facilitating short paths between nodes and information flow. This decrease suggests that the network might be fragmented, and only the centralities of the largest component were considered.

The analysis of random node removal within a small-world network was presented in Figure 4.12. As we progressively increased the proportion of nodes removed, we observed an increase in the degree centrality differences. While these differences showed a minor increase with the removal percentages, they remained relatively modest throughout the

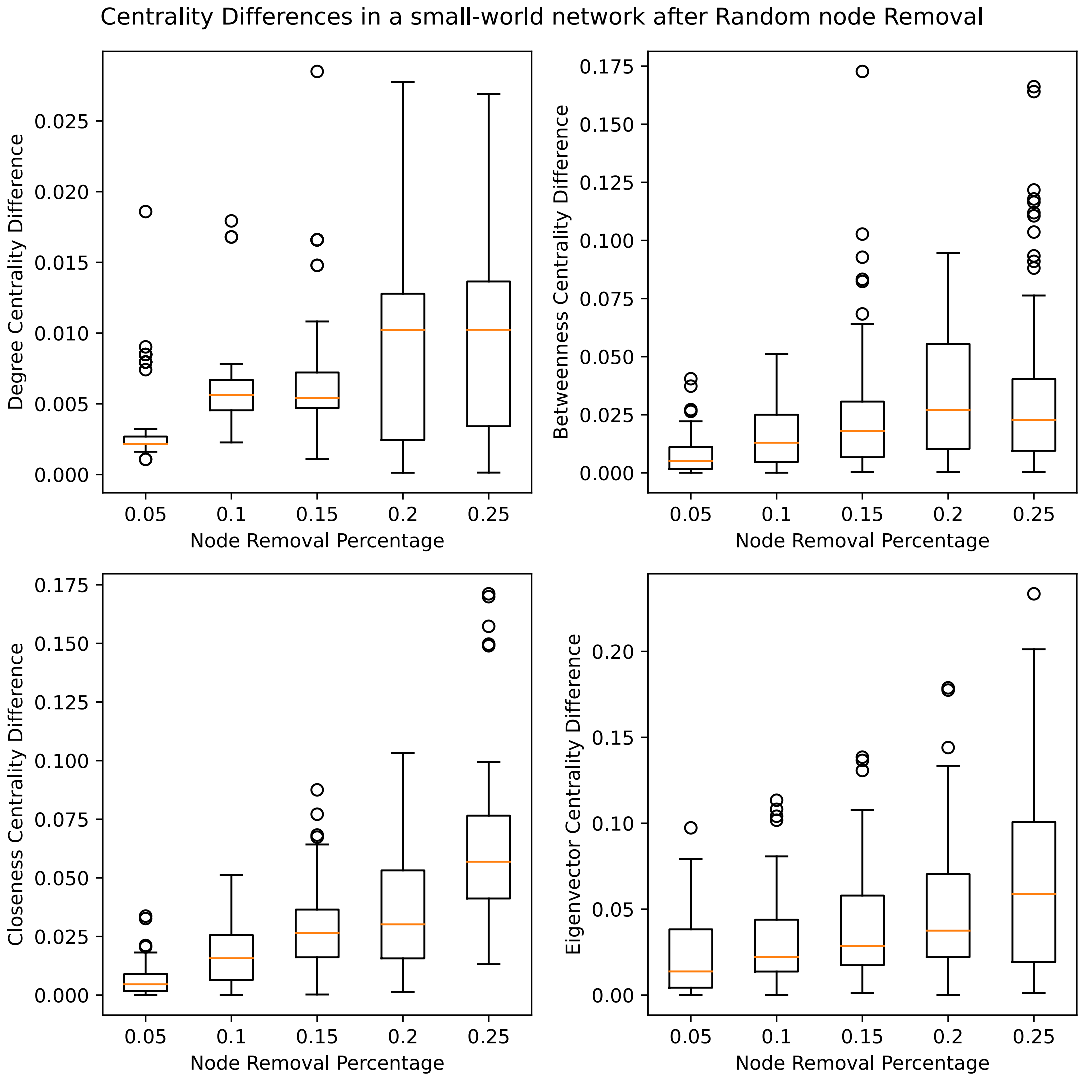


Figure 4.12: Impact of Random Node Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes.

process. This phenomenon suggests that, like scale-free networks, the removal of random nodes in small-world networks had a relatively limited impact on the network's structure. On the other hand, when we examined other centralities such as betweenness, closeness, and eigenvector centrality, we noticed that as the removal percentage increased in a small-world network, the differences in these centralities also increased. Notably, the differences were higher as we continued to remove nodes from the network. This indicates that random node removal had a profound and escalating impact on these centralities within a small-world network. The network's connectivity might be disrupted, leading to substantial differences in centralities. These findings shows the sensitivity of small-world networks to random node removal and emphasize the substantial changes in centrality measures as more nodes were removed.

4.2.4 Random edge removal strategy

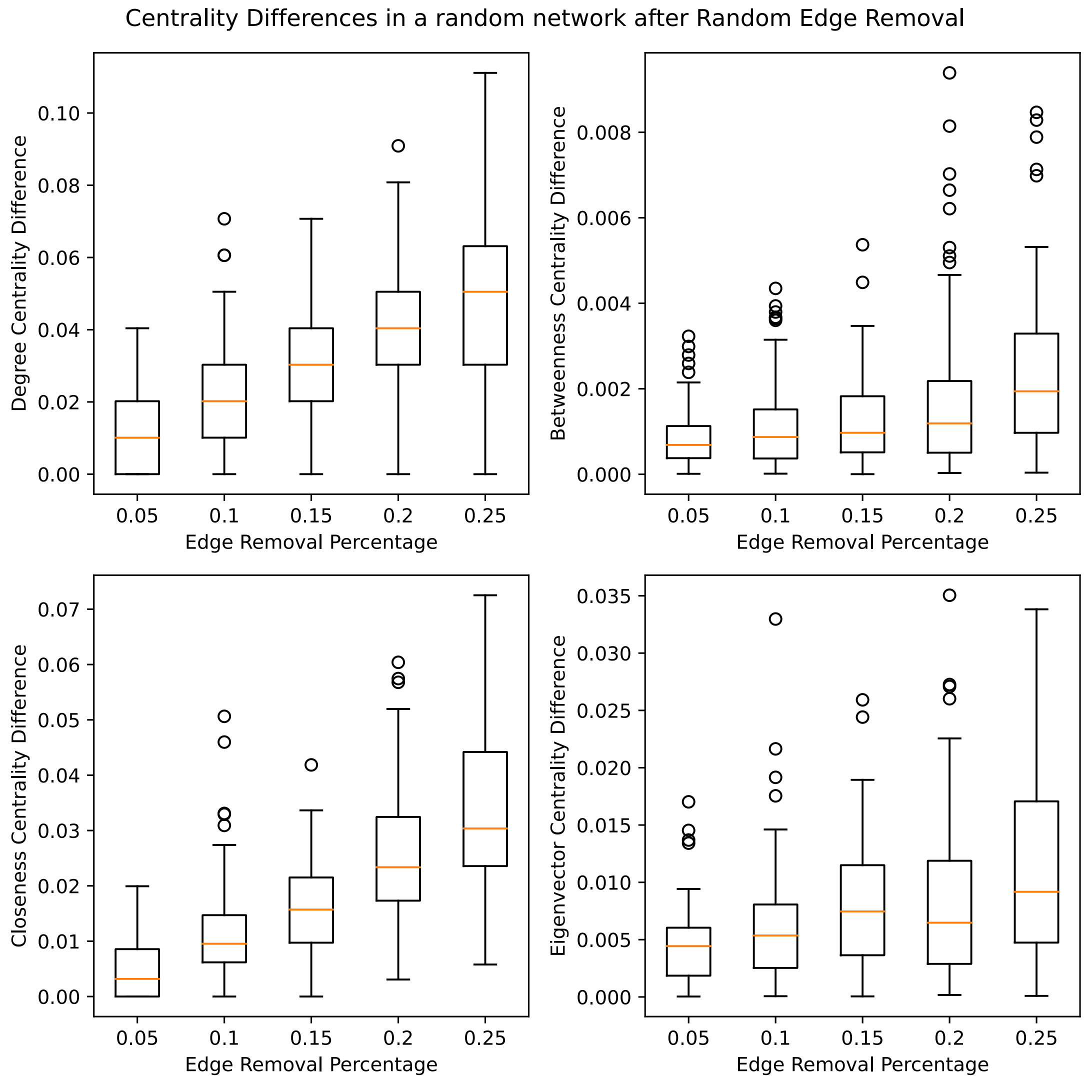


Figure 4.13: Impact of Random Edge Removal on Centrality measures of an Erdős–Rényi Random Network with $n = 100$ nodes.

In this analysis, we examined the effects of random edge removal on these four centrality measures in complex networks. In a random network, see Figure 4.13, we observed

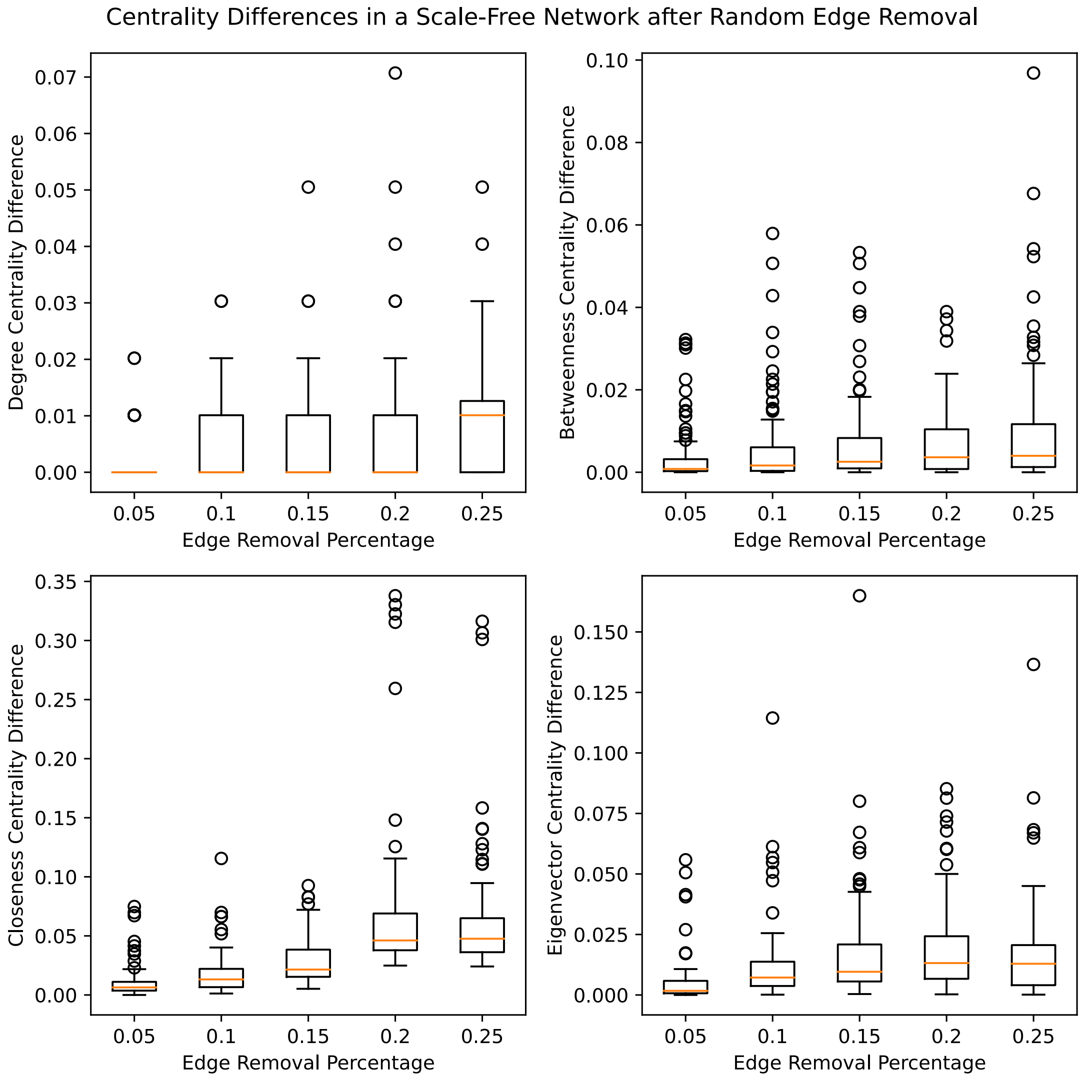


Figure 4.14: Impact of Random Edge Removal on Centrality measures of a Barabási-Albert Scale-free Network with $n = 100$ nodes.

a noticeable difference in degree and closeness centrality as the percentage of random edge removal increased. This difference became more pronounced with higher removal percentages, indicating that network connectivity was affected. On the other hand, the difference in betweenness centrality and eigenvector centrality were relatively small, showcasing the network's resilience in terms of information flow. This indicates that, despite the disruptions caused by edge removal, the network maintains its capacity for efficient information transmission.

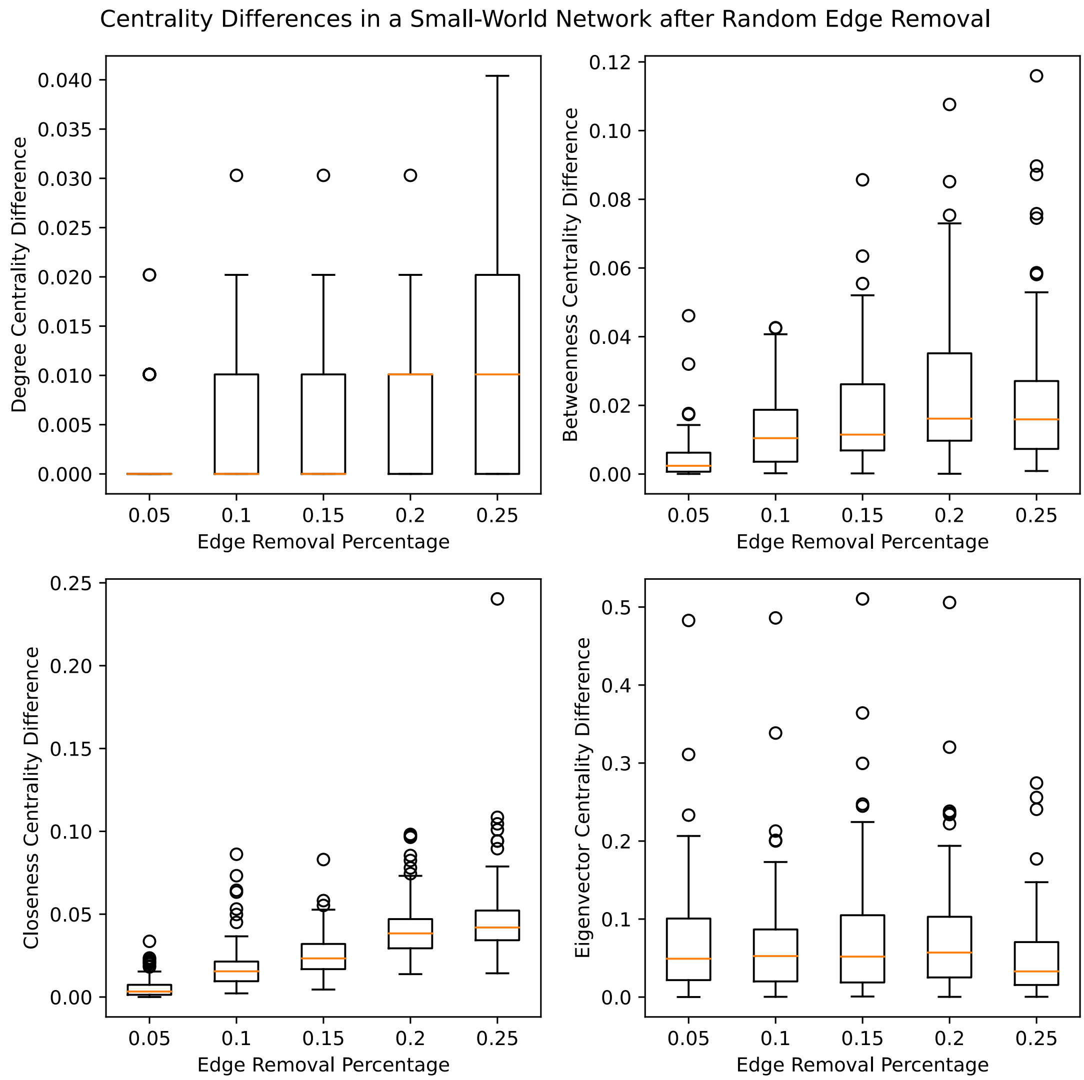


Figure 4.15: Impact of Random Edge Removal on Centrality measures of a Watts-Strogatz Small-world Network with $n = 100$ nodes.

The impact of random edge removal in a scale-free network is presented in Figure 4.14. Degree centrality differences remained relatively small even at higher removal percentages, indicating the network's resilience in terms of connectivity. The differences in betweenness and eigenvector centrality were higher compared to degree centrality differences but still relatively minor, suggesting a slight susceptibility of high-degree nodes to random edge removal and a potential disruption of critical information pathways. Closeness centrality differences were more pronounced, indicating a reduction in the network's clustering. The scale-free nature of this network, characterized by a few highly connected hubs, contributed to its robustness against random edge removal. This resilience against edge removal demonstrates the network's ability to maintain its structure and information flow even when subjected to significant disruptions.

The impact of random edge removal on a small-world network was presented in Figure 4.15. The difference in degree centrality stayed minor, showing that nodes stayed connected even when higher percentages of edges were randomly removed. However, the increasing difference in betweenness centrality with a higher edge removal percentage indicates a more substantial impact on the ability of the network's information flow. Similarly, the higher closeness centrality difference at higher removal percentages suggests a reduction in network clustering. In contrast, eigenvector centrality differences were the highest among all centrality measures, indicating a more significant impact on the influence of nodes within the network. While a small-world network's structure provides efficient information flow, its higher degree of interconnectedness makes it more vulnerable to disruptions caused by random edge removal.

The above observations gives the importance of considering network structure and the centrality measure when assessing the impact of different removal strategies.

4.3 Upper and Lower Bounds of Degree Centrality after removal

In the analysis of upper and lower bounds of degree centrality across random, scale-free, and small-world networks, we initiated by generating three type of networks, each composed of 100 nodes using the Erdos-Renyi model, Barabasi-Albert model and Watts-Strogatz model respectively. This analysis of upper and lower bounds of degree centrality in these networks subjected to the random removal of 10% of its nodes. While we used random node removal strategy to assess the accuracy of the bounds, the bounds can be effectively applied to other node removal strategies, including high-degree, low-degree and random strategies. This analysis involves computing degree centrality values both before and after the node removal and determining the upper and lower bounds for each node's degree centrality using the above derived Equations 3.3 and **??** respectively.

In the analysis of a random network, see Figure 4.16, it was evident that after removing 10 percent of the nodes at random, the degree centrality values for most nodes in the network experienced a notable decrease. The upper and lower bounds offered critical insights into the resilience of individual nodes. Nodes which are not connected to any removed nodes approached the upper limit of their degree centrality, highlighting their independence from the influence of the removed nodes. Conversely, nodes with stronger connections to the removed nodes displayed near the lower bounds, indicating a higher loss in degree centrality due to their associations with the removed nodes.

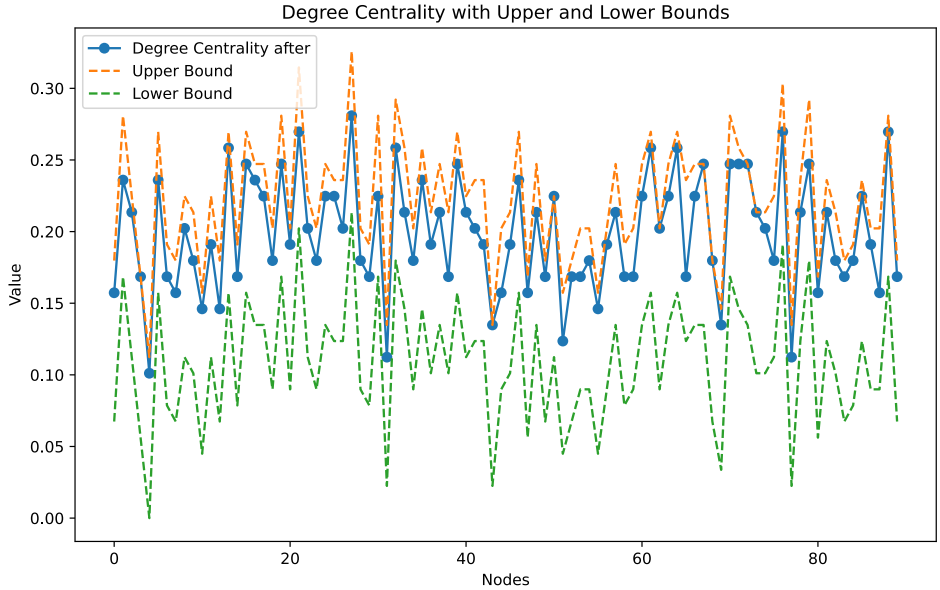


Figure 4.16: Upper and Lower Bounds of Degree Centrality in an Erdős–Rényi Random Network.

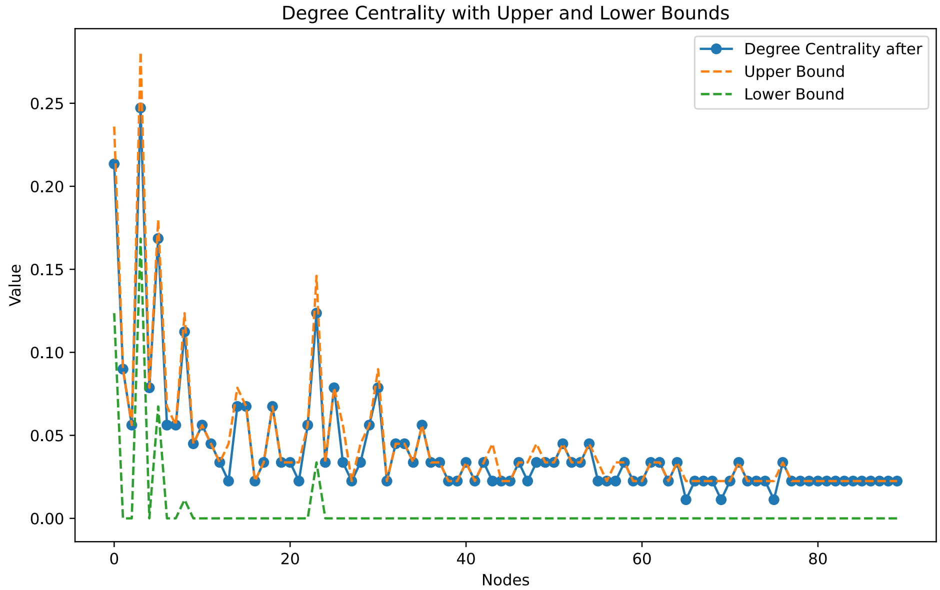


Figure 4.17: Upper and Lower Bounds of Degree Centrality in a Barabási-Albert Scale-Free Network.

In the context of a scale-free network, the analysis after removing 10 percent of nodes randomly revealed that the nodes remained influential and relatively unscathed. The upper and lower bounds in Figure 4.17 further emphasized this phenomenon. Mostly Nodes in a scale-free network approached the upper bound of their degree centrality, demonstrating their resilience. This observation is consistent with the properties of scale-free networks, which are characterized by a large proportion of nodes with lesser connections and a small number of hub nodes that are highly connected. Consequently, the degree centrality of nodes, especially those with limited connections, tended to lean towards the upper bound after random node removal.

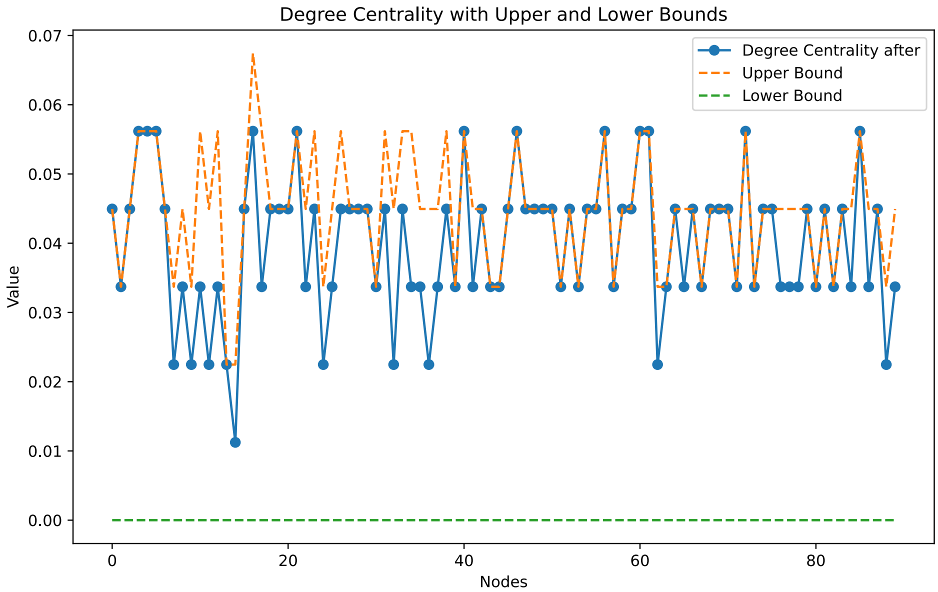


Figure 4.18: Upper and Lower Bounds of Degree Centrality in a Watts-Strogatz Small-World Network.

When assessing a small-world network after random node removal, see Figure 4.18, degree centrality values of some nodes fell within the range defined by the upper and lower bounds. This implies that those nodes were connected to the removed nodes. Node values that were closer to the lower bound were associated with more removed nodes, whereas those closer to the upper bound were associated with fewer removed nodes. This detailed perspective provides insights into the responses of nodes within a small-world network to random node removal. In small-world networks, the lower bound of degree centrality for all nodes reaches zero. This indicates that the average degree of a small-world network is smaller than the number of removed nodes.

These analyses presents the different response of random, scale-free, and small-world networks to random node removals and provide insights on the role of node connectivity in network resilience.

4.4 Structural changes in Complex Networks

In this study, we examined the structural changes within random networks, scale-free networks, and small-world networks, employing the same approach used earlier in our analysis. This approach involved different node and edge removal processes to observe the network's evolution. For each network type, we applied four distinct scenarios: random node removal, high-degree node removal, low-degree node removal, and random edge removal. Our primary goal remained consistent: to reveal the impact of these perturbations on the networks' topology and connectivity.

4.4.1 Change in Network Connectivity

In the analysis of random networks, we employed a consistent approach to assess their structural robustness under four distinct perturbation scenarios: random node removal, high-degree node removal, low-degree node removal, and random edge removal. Here in each iteration 1% of nodes in node removal approaches and 1% of edges in edge removal approach are removed.

The removal of random nodes from a random network exhibited remarkable resilience, see Figure 4.19. Despite the stochastic nature of this process, the network displayed significant robustness. Even when a substantial fraction of nodes was removed, the structural integrity of the network was maintained, with no disconnection observed.

The same resilience was observed for the removal of high-degree and low-degree nodes, see Figure 4.19. The size of the largest component remained unchanged for the removal of high-degree and low-degree nodes (the respective green and yellow color line are covered by the red line in the Figure 4.19). The network stayed robust and connected even though the critical high-degree nodes were removed. This observation of resilience can be attributed to the random and homogeneous nature of a network.

The most striking structural robustness occurred while randomly removing edges from the network. The network's structural integrity remained totally unaffected. The network showed the robustness up to 81 iterations (81% of edge removal from the original network) with the largest component of 100 nodes (original generated network with no disconnected nodes).

The network demonstrated both resilience and robustness in response to each of the four

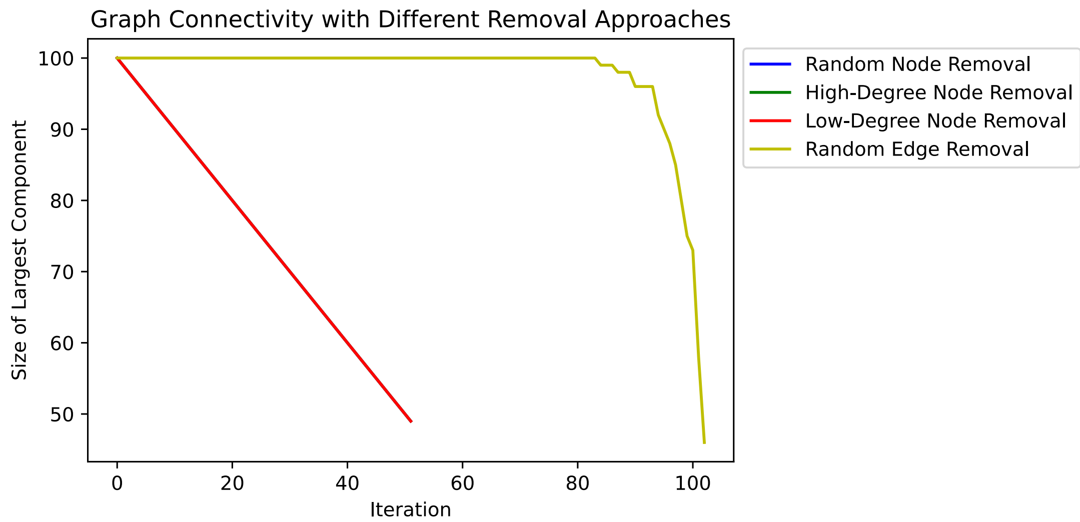


Figure 4.19: Impact of Different Removal Strategies on Network Connectivity in an Erdős–Rényi Random Network.

distinct types of perturbations, highlighting its ability to adapt and maintain structural integrity under a variety of challenges. The random and uniform nature of these networks contributes to their resilience under different random node and edge removal strategies. Understanding these dynamics is essential for applications in diverse fields, from information dissemination to transportation systems, where random network structures often play a vital role.

The removal of random nodes from a scale-free network demonstrated remarkable resilience, as shown in Figure 4.20. The network remained connected up to 21 iterations of random node removal, highlighting its robustness against the stochastic nature of this process. However, at the 35th iteration, the network faced again the disconnection, emphasizing its vulnerability to extensive random node removal.

Similarly, for the removal of high-degree nodes in a scale-free network, the network's vulnerability was evident. After just 4 iterations, the network got disconnected, showcasing its susceptibility to the targeted removal of high-degree nodes. This disconnection resulted into 30 percent of the network as it's largest connected component after only 10 iterations.

Conversely, a scale-free network exhibited robustness in response to low-degree node removal. The network remained connected and adaptive, demonstrating its capacity to withstand the perturbation without significant structural impact.

When subject to random edge removal, a scale-free network displayed impressive re-

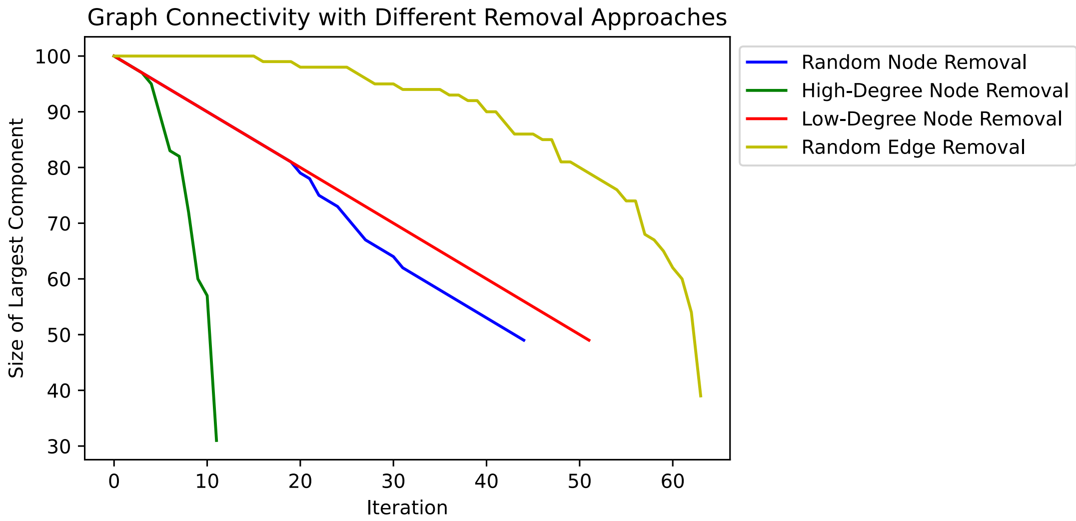


Figure 4.20: Impact of Different Removal Strategies on Network Connectivity in a Barabási-Albert Scale-Free Network.

silience. It maintained connectivity up to 38 iterations of edge removal, highlighting its ability to adapt and stay intact under this perturbation.

These findings emphasize the behavior of scale-free networks under different perturbations and emphasize the importance of understanding their vulnerabilities and strengths in real-world applications.

In the case of random node removal within a small-world network, see Figure 4.21, the network displayed resilience up to 25 iterations before facing disconnection.

When subjected to high-degree node removal, a small-world network demonstrated vulnerability, with disconnection occurring after 16 iterations. However, the network's susceptibility became apparent as early as 25 iterations, with almost a 50 percent reduction in its original size, highlighting the consequences of high-degree node removal. Conversely, low-degree node removal posed no threat to a small-world network's robustness. It remained connected throughout the removal process, even as 50 percent of the random nodes were removed, demonstrating its remarkable structural integrity.

In the analysis of random edge removal in a small-world network, the network showed high resilient upto around 40% of random edge removal but around 45th iteration, the network faced a significant drop, going from a 90 percent largest connected component to a 50 percent largest connected component. This dynamic response suggests that the network initially demonstrated the capability to adjust and maintain its connectivity in response to these random edge removal. However, as the edge removal increased, the

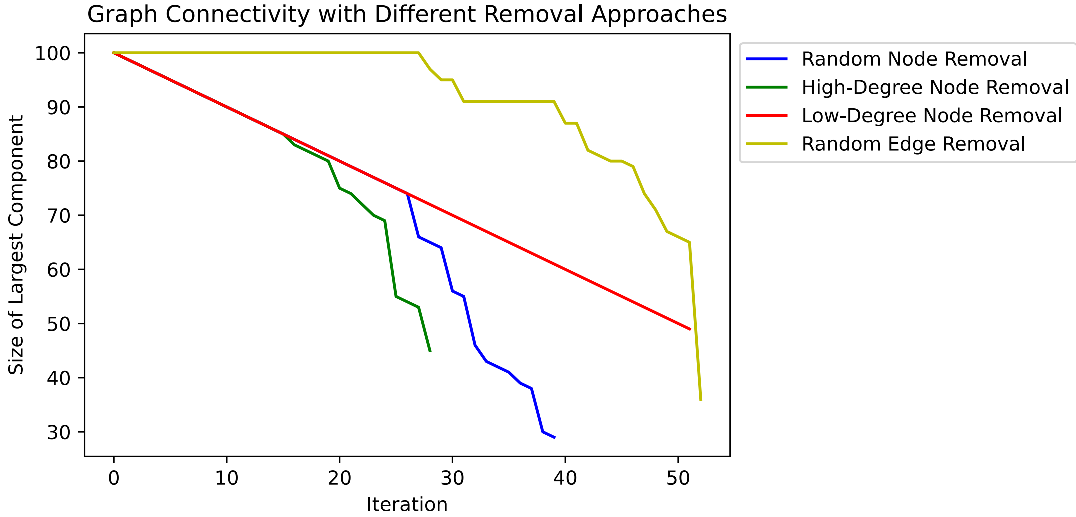


Figure 4.21: Impact of Different Removal Strategies on Network Connectivity in a Watts-Strogatz Small-World Network.

network eventually reached a point where it could no longer adapt effectively and started to break down and lose its resilience.

These observations emphasize the behavior of small-world networks under various perturbation scenarios and underscore the importance of understanding their vulnerabilities and strengths in practical applications.

4.4.2 Change in Network Structure

In this subsection, we explore the impact of the node and edge removal strategies on the structure of the complex networks. We conducted a series of experiments to assess how the network's degree distribution evolves as we apply various removal strategies. These experiments were conducted on the same three types of random, scale-free, and small-world networks with the same parameters as discussed before in the methodology. Also, the same removal strategies included high-degree node removal, low-degree node removal, random node removal, and random edge removal and the removal percentage from 5% to 25%. By conducting these experiments on different network types, we aimed to gain insights into how each removal strategy influenced the degree frequency distribution of the network and, consequently, the overall network structure. From plot figures, the degree distribution plot where each plot illustrates a distinct removal strategy, and each color line represents the degree distribution in the corresponding network before and after a range of removal percentages. We used Non-parametric Kolmogorov-Smirnov (K-

S) and Cramér-von Mises (CvM) tests with a significance level of 0.05 to statistically evaluate these changes.

Impact of removal strategies on Random Network

The frequency distribution plots of a random network before and after high-degree nodes removed are displayed in Figure 4.22. The results of the high-degree node removal strategy are summarized in Table 4.1. The CvM and K-S statistics show a significant increase in the proportion of node removal. The CvM statistics consistently deviate from the original degree distribution, rising from 0.8888 at 5% removal to 11.0364 at 25% removal. The CvM statistics' p-values correspondingly decreased, signifying a significant deviation from the original distribution. Similar trends can be seen in the K-S statistics, which rise from 0.1984 at 5% removal to 0.79 at 25% reduction. The corresponding K-S test p-values similarly decreased to zero, confirming the significant effect of removing high-degree nodes on the degree distribution of the network. Based on these findings, the null hypothesis which states that the degree distribution of a random network is the same before and after a removal of high-degree nodes is rejected.

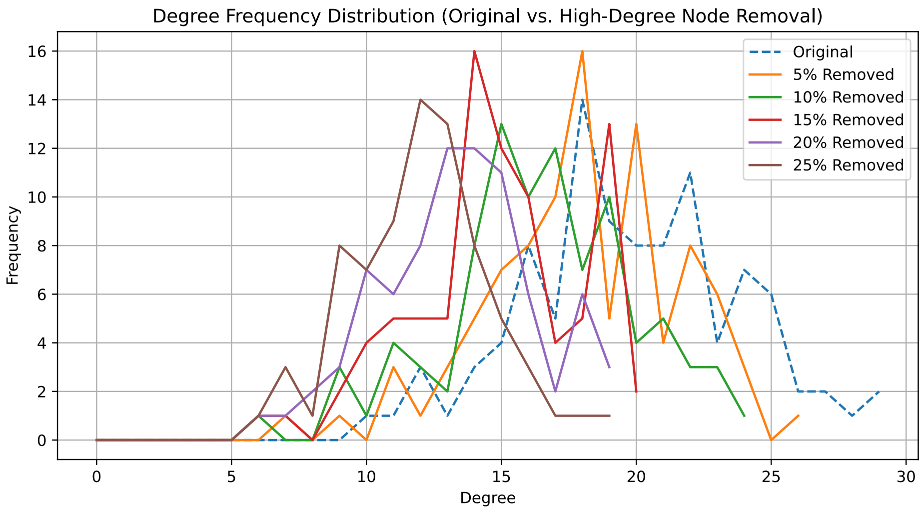


Figure 4.22: Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of High-Degree Node Removal.

The frequency distribution plots of a random network before and after the removal of low-degree nodes are shown in Figure 4.23. Table 4.2 provides an overview of the low-degree node removal strategy's results. As the percentage of nodes removed increases from 5% to 25%, there is a small but noticeable change in the degree distribution plot.

Table 4.1: Erdős–Rényi Random Network: Statistical Test Results after High-Degree Node Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	0.8888	0.0045	0.1984	3.60×10^{-2}
10%	3.0736	5.02×10^{-8}	0.3733	2.11×10^{-6}
15%	6.2458	1.97×10^{-11}	0.5076	2.46×10^{-11}
20%	8.9809	2.73×10^{-10}	0.6575	1.09×10^{-18}
25%	11.0364	2.53×10^{-10}	0.7900	4.36×10^{-27}

This change is reflected in both the CvM and K-S statistics, which increase with increasing node removal. The decreasing p-values associated with the CvM and K-S statistics demonstrate that the observed changes in the degree distribution are statistically significant. As expected, the CvM test is more sensitive to the long tail of the plot than the K-S test. Given the importance of the long tail in network analysis, we can reject the hypothesis that the degree distribution remains unchanged even after removing 15% of low-degree nodes.

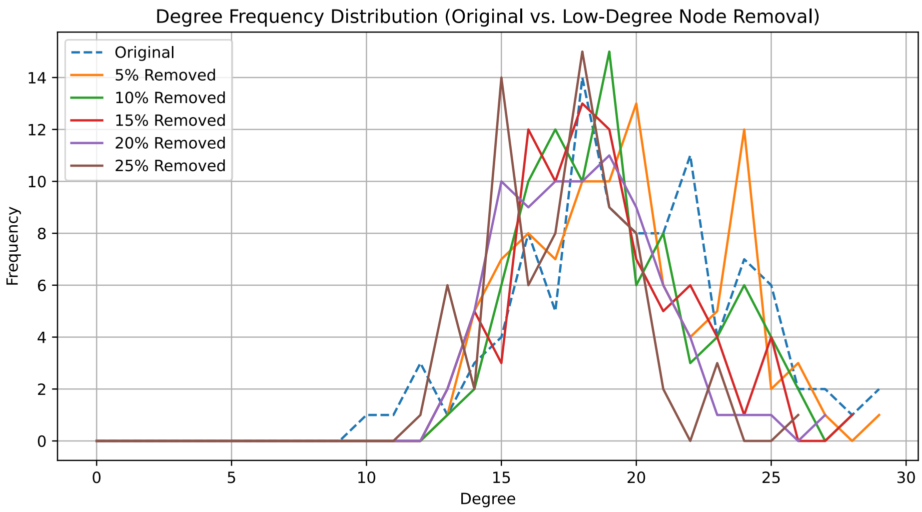


Figure 4.23: Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Low-Degree Node Removal.

The results of the random node removal method are displayed in the frequency distribution plot (see Figure 4.24 and Table 4.3). At 5% removal, the CvM statistics show a slight increase to 0.1355, with a p-value of 0.4404, indicating no significant departure from the original distribution. However, as the percentage of nodes removed increases, the CvM statistics rise notably, reaching 5.5217 at 25% removal, accompanied by a very low p-value. The K-S statistics also demonstrate a progressive increase, from 0.0826

Table 4.2: Erdős–Rényi Random Network: Statistical Test Results after Low-Degree Node Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	0.0599	0.8219	0.0721	0.9394
10%	0.2897	0.1451	0.1322	0.3430
15%	0.6399	0.0178	0.1829	0.0791
20%	1.2772	0.0006	0.2550	0.0051
25%	2.2123	4.16×10^{-6}	0.35	3.79×10^{-5}

at 5% removal to 0.54 at 25% removal, with corresponding p-values diminishing significantly, reinforcing the significant deviation from the original distribution. By combining the results, we fail to reject the null hypothesis for the removal of 5% of random nodes, but we reject it for the remaining removal percentages.

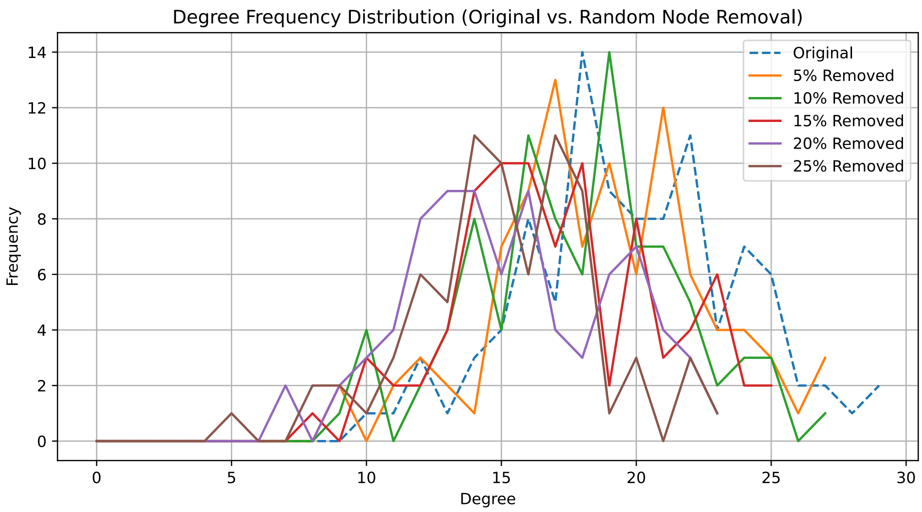


Figure 4.24: Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Random Node Removal.

Table 4.3: Erdős–Rényi Random Network: Statistical Test Results after Random Node Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	0.1355	0.4404	0.0826	0.8563
10%	0.9319	0.0035	0.2	0.0383
15%	2.8869	1.30×10^{-7}	0.3518	1.42×10^{-5}
20%	3.4546	7.19×10^{-9}	0.4025	6.09×10^{-7}
25%	5.5217	3.51×10^{-10}	0.54	5.85×10^{-12}

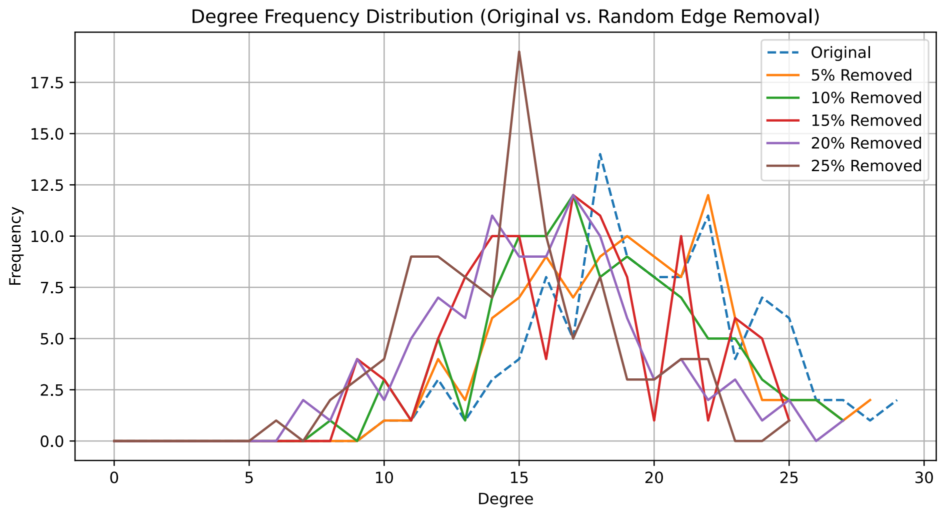


Figure 4.25: Degree Frequency Distribution in Erdős–Rényi Random Network at different percentages of Random Edge Removal.

Table 4.4: Erdős–Rényi Random Network: Statistical Test Results after Random Edge Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.2962	0.1390	0.12	0.4695
10%	1.1043	0.0014	0.23	0.0099
15%	2.5327	8.03×10^{-7}	0.33	3.21×10^{-5}
20%	4.1507	2.12×10^{-10}	0.43	1.12×10^{-8}
25%	6.0382	9.28×10^{-12}	0.5	1.00×10^{-11}

The results of the random edge removal strategy, and frequency distribution plot are shown in Figure 4.25 and Table 4.4 respectively. The CvM statistics increase from 0.2962 at 5% removal with a p-value of 0.1390 to 6.0382 at 25% removal with an extremely low p-value, indicating the statistical significance of structure change. The K-S statistics also show a progressive rise, from 0.12 at 5% removal with a p-value of 0.4695 to 0.5 at 25% removal with a very low p-value. The results of this study indicate how removing more than 5% of random edges significantly affects the degree distribution, causing a noticeable and statistically significant deviation from the original distribution. Consequently, for the remaining removal percentages, we reject the null hypothesis.

Impact of removal strategies on Scale-free Network

The results of the high-degree node removal strategy in a scale-free network are the frequency distribution plots in Figure 4.26 and associated statistical test results in Table 4.5. The CvM statistics exhibit a significant increase, from 2.0114 at 5% removal to 10.1317 at 25% removal, with very low p-values consistently indicating statistical significance. Similar gradual changes are also shown by the K-S statistics, which indicate a significant shift in the degree distribution. These findings show the substantial impact of even 5% of high-degree node removal on altering the network's degree distribution in a scale-free network. Hence, we reject the null hypothesis for each removal percentage.

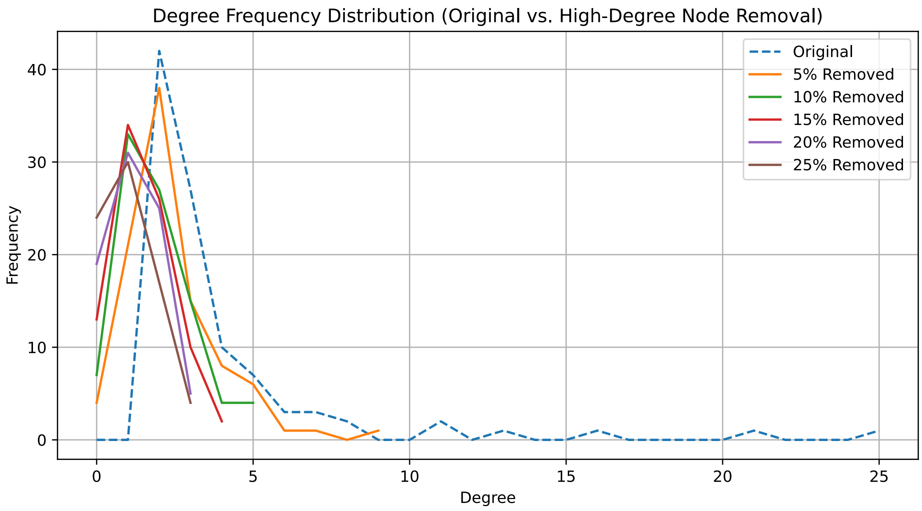


Figure 4.26: Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of High-Degree Node Removal.

Table 4.5: Barabási-Albert Scale-Free Network: Statistical Test Results after High-Degree Node Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	2.0114	1.18×10^{-5}	0.2632	1.78×10^{-3}
10%	4.5115	3.88×10^{-11}	0.4444	5.88×10^{-9}
15%	7.2186	3.69×10^{-10}	0.5529	1.79×10^{-13}
20%	9.0646	3.25×10^{-10}	0.6250	8.85×10^{-17}
25%	10.1317	2.38×10^{-9}	0.7200	5.89×10^{-22}

The results from the low-degree node removal strategy in a scale-free network, the frequency distribution plot shown in Figure 4.27, indicate a relatively mild impact on the degree distribution, the associated statistical results are presented in Table 4.6. The CvM statistics display a gradual increase from 0.0241 at 5% removal to 0.2861 at 25%

removal, with p-values consistently higher than the significance value 0.05, suggesting no statistically significant deviation from the original distribution. Alongside, the K-S test demonstrates similar results, with the p-value higher than the significance value of 0.05 for all removal percentages. This suggests that removing low-degree nodes has a limited effect on the degree distribution in a scale-free network. Consequently, we fail to reject the null hypothesis.

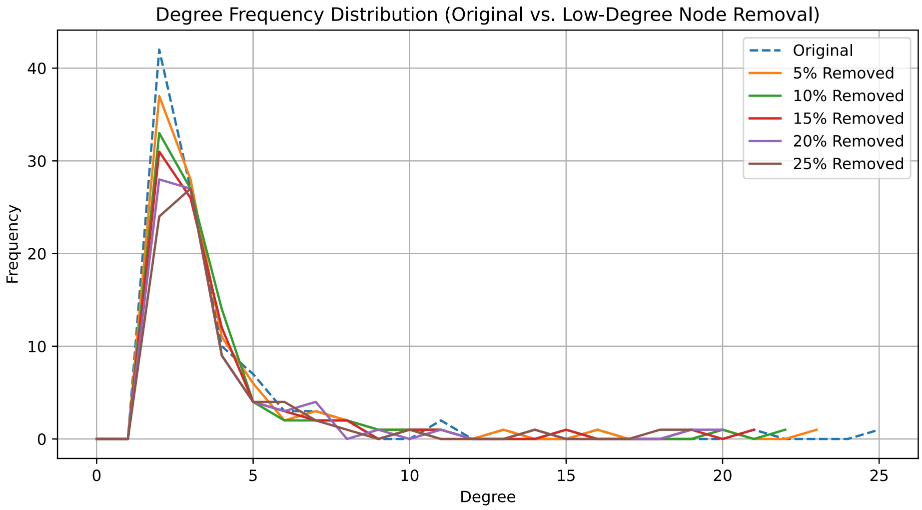


Figure 4.27: Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Low-Degree Node Removal.

Table 4.6: Barabási-Albert Scale-Free Network: Statistical Test Results after Low-Degree Node Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.0241	0.9934	0.0305	0.9999
10%	0.0840	0.6754	0.0533	0.9978
15%	0.1083	0.5505	0.0553	0.9969
20%	0.1637	0.3528	0.0700	0.9707
25%	0.2861	0.1485	0.1000	0.7523

The results from a scale-free network, as the frequency distribution plot shown in Figure 4.28 and the statistical results shown in Table 4.7, indicate that random node removal has a minimal impact on the degree distribution for removal percentages up to 15%. The CvM statistics show deviations, with p-values remaining lesser than the significance value of 0.05. However, at 20% and 25% removal, a more noticeable shift is observed, with CvM statistics reaching 0.5121 and 1.3568, respectively, and p-values dropping to 0.0371 and 0.0004, indicating a statistically significant deviation from the original distribution. Additionally, the K-S test also showed similar results up to 20% of random node removal with K-S statistics 0.15 and p-value 0.2456. This suggests that the effect of random node removal on the degree distribution becomes more pronounced as the percentage of removed nodes increases, leading to the rejection of the null hypothesis for 25% of random node removal.

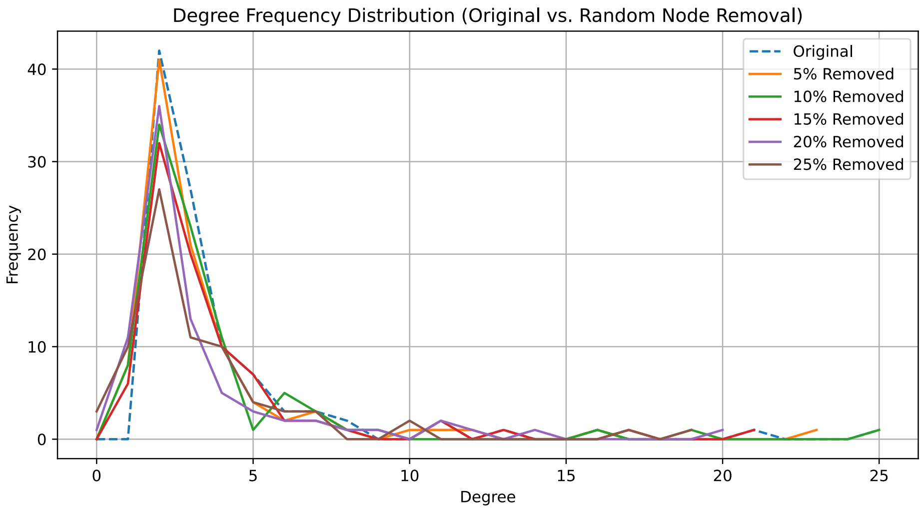


Figure 4.28: Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Random Node Removal.

Table 4.7: Barabási-Albert Scale-Free Network: Statistical Test Results after Random Node Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	0.0363	0.9564	0.0326	0.9999
10%	0.1855	0.2993	0.0889	0.8092
15%	0.2839	0.1508	0.0976	0.7269
20%	0.5121	0.0371	0.1500	0.2456
25%	1.3568	0.0004	0.2133	0.0354

The results of a scale-free network after random edge removal are illustrated in Figure 4.29 and associated statistical results in Table 4.8. The CvM statistics progressively increase from 0.0380 at 5% removal to 1.3194 at 25% removal, with p-values decreasing significantly from 15% removal. Simultaneously, the K-S test provided similar results with a p-value below the significant value of 0.05 from 20% removal, further supporting the significant deviation from the original distribution. Here, we reject the null hypothesis at 20% and 25% removal of random edges.

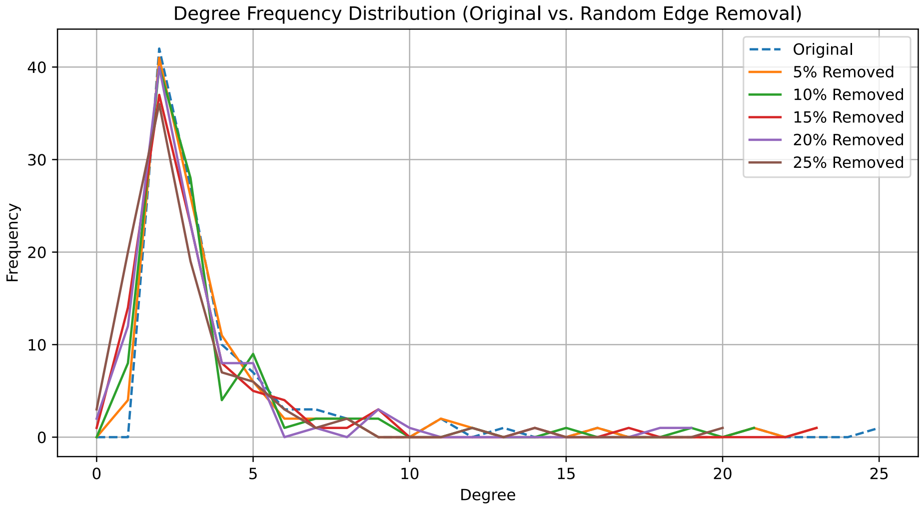


Figure 4.29: Degree Frequency Distribution in Barabási-Albert Scale-Free Network at different percentages of Random Edge Removal.

Table 4.8: Barabási-Albert Scale-Free Network: Statistical Test Results after Random Edge Removal.

Percentage	CvM Statistic	P-Value (CvM)	K-S Statistic	P-Value (K-S)
5%	0.0380	0.9488	0.0500	0.9997
10%	0.1921	0.2849	0.0800	0.9084
15%	0.4865	0.0431	0.1500	0.2112
20%	0.7992	0.0073	0.2100	0.0241
25%	1.3194	4.42×10^{-4}	0.2100	0.0241

Impact of removal strategies on Small-world Network

The frequency distribution plots in Figure 4.30 shows the impact of high-degree node removal on the degree distribution of a small-world network. The associated statistical results are presented in Table 4.9. The CvM test is more sensitive to these changes, at 5% removal, the CvM statistic is 0.5494 with a p-value of 0.0299, indicating a significant deviation from the original degree distribution. The K-S test, on the other hand, showed lower statistics and a p-value higher than the significance level of 0.05 for both 5% and 10% removal. However, after 10% removal, the K-S test also indicated a significant deviation with a p-value lower than the significance level. This suggests that the CvM test is more sensitive to changes in the degree distribution of high-degree nodes so, we reject the null hypothesis at higher removal percentages including at 10% removal of high-degree nodes.

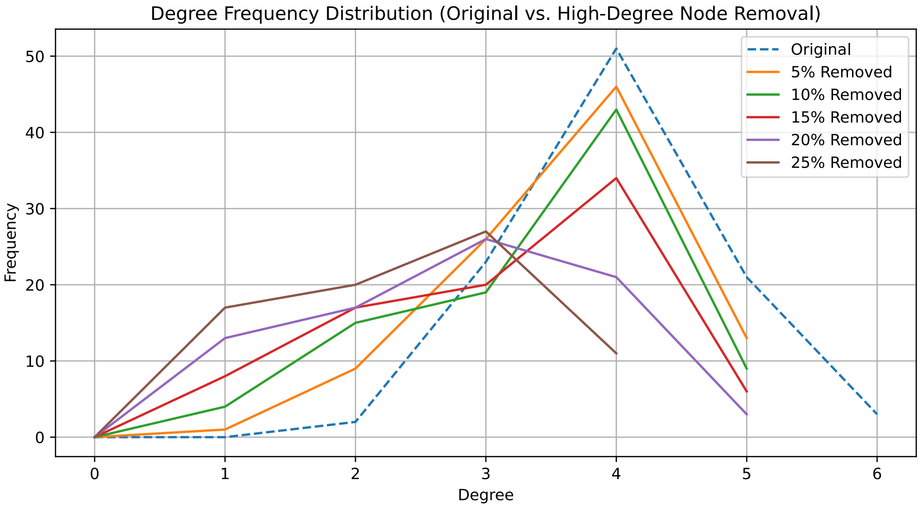


Figure 4.30: Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of High-Degree Node Removal.

Table 4.9: Watts-Strogatz Small-World Network: Statistical Test Results after High-Degree Node Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.5494	0.0299	0.1289	0.3528
10%	1.1332	1.19×10^{-3}	0.1911	0.0538
15%	2.3875	1.69×10^{-6}	0.2794	1.14×10^{-3}
20%	4.9056	3.68×10^{-11}	0.4500	1.29×10^{-8}
25%	8.0887	3.35×10^{-11}	0.6033	5.00×10^{-15}

Figure 4.31 visually represents the impact of the Low-Degree Node Removal strategy on a small-world network's degree distribution, with detailed statistical results presented in Table 4.10. Initially, at a 5% removal rate, the CvM statistics of 0.0314 and a p-value of 0.9754 indicated no noteworthy deviation. However, as the removal percentage increased, statistical significance became apparent, with a CvM statistic of 0.9680 and a significantly low p-value of 0.0029 at 20% removal, further decreasing at higher percentages. The corresponding K-S test at 20% removal, featuring K-S statistics of 0.1750 and a p-value of 0.1172, suggests that low-degree node removal did impact the degree distribution, although the effect was not significant. A statistically significant shift in distribution was detected at 25% removal, with K-S statistics of 0.2700 and a p-value of 0.0032. These findings demonstrate that the removal of low-degree nodes has a gradual impact on the degree distribution of a small-world network. Here, we reject the null hypothesis only at 25% removal of low-degree nodes.

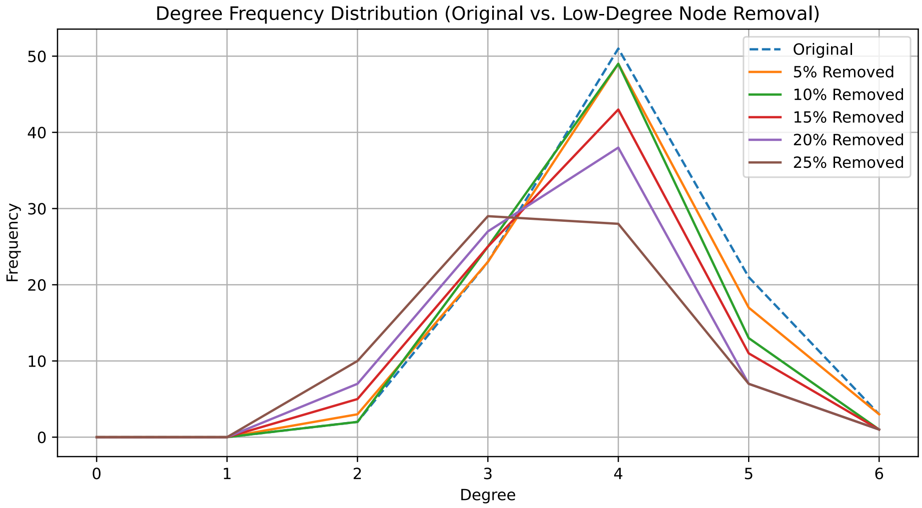


Figure 4.31: Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Low-Degree Node Removal.

Table 4.10: Watts-Strogatz Small-World Network: Statistical Test Results after Low-Degree Node Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.0314	0.9754	0.0295	0.9999
10%	0.1702	0.3356	0.0844	0.8536
15%	0.4109	0.0677	0.1029	0.6668
20%	0.9680	2.90×10^{-3}	0.1750	0.1172
25%	1.7561	4.44×10^{-5}	0.2700	3.16×10^{-3}

The effect of the random node removal strategy on the degree distribution of a small-world network is presented in Figure 4.32 and the corresponding statistical results are presented in Table 4.11. At 5% removal, the CvM statistics of 0.2938 and a p-value of 0.1412 indicated no significant deviation from the original distribution. However, as removal percentages increased, we can see the statistical significance. Notably, at 15% removal, a CvM statistic of 2.3229 and a very low p-value highlighted a substantial deviation from the original distribution. This was supported by the corresponding K-S statistic, which also increased to 0.3265 at 15% removal, along with a p-value lower than the significance level of 0.05. The higher sensitivity of the CvM test to changes in the distribution of high-degree nodes suggests that the observed deviation is likely due to the removal of these nodes. So, We reject the null hypothesis for the higher percentage including 15% removal of random nodes.

In Figure 4.33, we present the effects of random edge removal strategy on a small-world

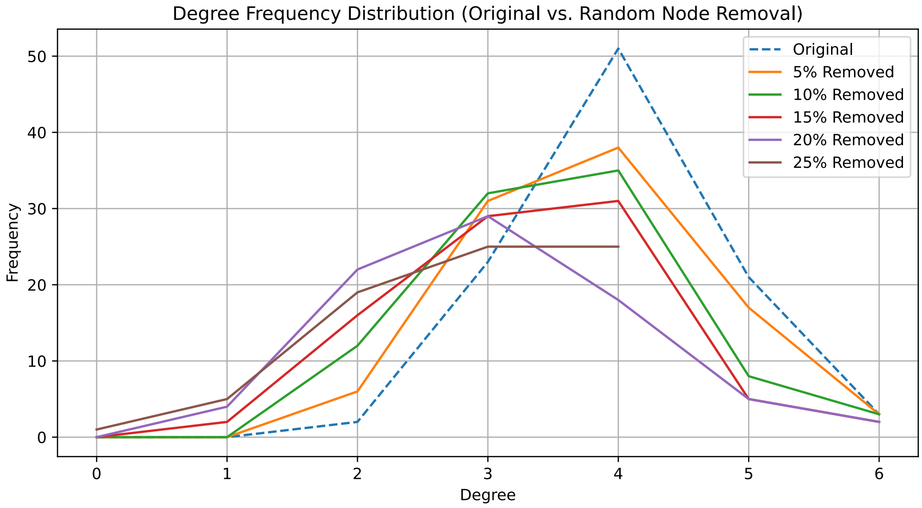


Figure 4.32: Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Random Node Removal.

Table 4.11: Watts-Strogatz Small-World Network: Statistical Test Results after Random Node Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.2938	0.1412	0.0926	0.7507
10%	0.6479	0.0170	0.1722	0.1050
15%	2.3229	2.36×10^{-6}	0.3265	7.45×10^{-5}
20%	3.3379	1.30×10^{-8}	0.3750	4.56×10^{-6}
25%	4.8054	2.83×10^{-11}	0.4567	1.47×10^{-8}

network's degree distribution and the associated statistical results are presented in Table 4.12. At 5% removal, CvM statistic 0.2243 and K-S statistic 0.09 indicated no significant change, aligning with p-values of 0.2260 and 0.8154, respectively. However, at 10% removal and higher removal percentages, both tests illustrated progressively significant deviations from the original distribution. By 25% removal, the CvM p-value and K-S p-value became very low, indicating the strategy's significant impact on the network's structural integrity. From these results, we reject the hypothesis for higher removal percentages including 10% removal of random edges.

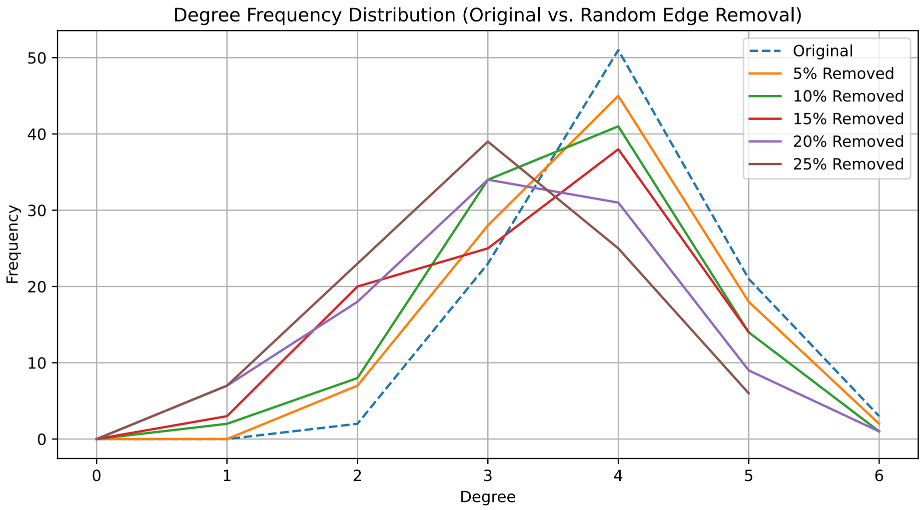


Figure 4.33: Degree Frequency Distribution in Watts-Strogatz Small-World Network at different percentages of Random Edge Removal.

Table 4.12: Watts-Strogatz Small-World Network: Statistical Test Results after Random Edge Removal.

Percentage	**CvM Statistic**	**P-Value (CvM)**	**K-S Statistic**	**P-Value (K-S)**
5%	0.2243	0.2260	0.0900	0.8154
10%	1.0744	1.64×10^{-3}	0.2300	0.0099
15%	1.9970	1.27×10^{-5}	0.2900	4.12×10^{-4}
20%	3.4719	6.63×10^{-9}	0.3900	3.57×10^{-7}
25%	4.1784	1.84×10^{-10}	0.4200	2.75×10^{-8}

5 Conclusion and Discussion

In this paper, we studied the concept of node importance in complex networks and employed different centrality measures to quantify the influence of individual nodes in a network. Furthermore, we studied the impact of different removal strategies on centrality measures and predicted the bounds of degree centrality values after node removal to show the fluctuations in node influence within the network. At last, the study was focused on the importance of considering network connectivity and structural changes resulting from node and edge removal, uncovering critical thresholds. Through our findings, we aim to clarify the complex relationships between node importance, network structure, and network resilience.

Random networks showed a strong correlation between different centrality measures, indicating that all nodes play a relatively important role in information flow. However, while the network was resilient to the removal of low-degree and random nodes, targeting high-degree nodes or removing random edges had a significant impact on centrality measures. This suggests that highly connected nodes and overall network structure play a critical role in maintaining network resilience. The decrease in degree centrality values after random node removal highlights the impact of connectivity on node resilience. The calculated upper and lower bounds of degree centrality further emphasize this point, as unaffected nodes tended towards the upper limit while those connected to removed nodes approached the lower bounds. These findings provide insights into the relationship between node importance and network resilience in random networks. Furthermore, random networks lose their randomness at a removal threshold between 5% and 10%, indicating a structural transition. High-degree node removal causes structural changes at lower than 5% removal.

Scale-free networks rely on a small number of highly connected nodes called hubs to maintain efficient communication pathways. These hubs are crucial for maintaining the network's overall structure. While the network shows high resilience to random node removal, it becomes vulnerable when high-degree nodes, especially hubs, are targeted. The removal of hubs significantly impacts centrality measures, indicating their importance in preserving network integrity. The observation that nodes tend to approach the upper bounds after random node removal suggests that the network structure itself contributes to the resilience of scale-free networks [8]. Additionally, scale-free networks become fragmented at 4% high-degree node removal but remain connected under random node and edge removal strategies at 20%. Statistical tests indicate structural changes at 5% high-degree node removal and structural preservation against random node and edge removal up to as high as 15%.

Small-world networks balance redundancy and efficiency, resulting in a moderate positive correlation between centrality measures. The removal of high-degree nodes, often acting as bridges between different communities, can significantly disrupt network connectivity and structure. However, degree centrality values in small-world networks, after random node removal, consistently fall within the upper and lower bounds. The observation that nodes in small-world networks have varying degrees of connectivity to the removed nodes indicates the adaptability of these networks. This adaptability allows these networks to maintain their overall structure even when some nodes are removed. Furthermore, small-world networks maintained connectivity up to 15% high-degree node removal and approximately 25% random node and edge removal. Statistical analysis reveals structural preservation against high-degree node removal up to 5%, followed by structural loss at higher removal percentages. Similar resilience is observed against other removal strategies, maintaining structural integrity up to 5% removal and experiencing structural changes at higher percentages.

In summary, random networks showed high resilience to connectivity disruptions but lost their structural properties at a relatively low removal threshold. Scale-free networks exhibited vulnerability to high-degree node removal but showed resilience to other removal strategies to a moderate threshold. Small-world networks maintained connectivity up to substantial removal percentages but lost the structural integrity at a lower removal percentage, indicating a similar level of resilience as random networks. These findings highlight the complex relationship between node importance, network structure, and network resilience in complex networks.

Through our study, we have gained a more detailed understanding of the relationships between node importance, network resilience, and connectivity patterns in different network types. While our study contributes valuable insights into node importance, network resilience, and connectivity patterns, it is crucial to acknowledge certain limitations. Firstly, the focus on a relatively small network with 100 nodes may restrict the application of these findings to larger and more complex real-world networks. Secondly, the use of static network analysis may overlook the dynamic aspects of real-world networks, which are constantly evolving and adapting to changing conditions. To gain a more comprehensive understanding of network dynamics, future research may focus on temporal network analysis, which examines how networks change over time. Additionally, while the study examines various network types, the findings may not be universally applicable to all network structures. Future research could address these limitations by studying larger and more intricate networks, combining the dynamic elements. This would ultimately strengthen the validity and applicability of the study's conclusions.

Bibliography

[1] Stefano Boccaletti and Vito Latora. Complex networks: Structure and dynamics. *Chaos*, 16(4):045102, 2006.

[2] Paul Erdős and Alfréd Rényi. On random graphs i. *Publicationes Mathematicae*, 6:290–297, 1959.

[3] Edgar N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[4] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[5] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[6] Stephen P. Borgatti. Centrality and network flow. *Social Networks*, 27:55–71, 2005.

[7] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.

[8] R. Albert, H. Jeong, and A.L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 2000.

[9] Anna D. Broido and Aaron Clauset. Scale-free networks are rare. *Nature communications*, 10(1):1–12, 2019.

[10] Taylor B. Arnold and John W. Emerson. Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal*, 3(2):34–39, 2011.

[11] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[12] Béla Bollobás. *Modern Graph Theory*. Graduate Texts in Mathematics. Springer, New York, 1998.

[13] M.E.J. Newman. *The Structure and Dynamics of Networks.* Princeton University Press, 2003.

[14] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Physical Review Letters*, 85(21):4633, 2000.

[15] Matus Medo, Sergey N. Dorogovtsev, and Alexander V. Goltsev. Temporal networks. *Physical Review Letters*, 107(23):238701, 2011.

[16] M. E. J. Newman. *Power laws, Pareto distributions and Zipf's law.* Springer, 2005.

[17] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–233, 2003.

[18] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–77, 2002.

[19] L. C. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3):215–239, 1978.

[20] M. E. Shaw. Group structure and the behavior of individuals in small groups. *Journal of Psychology*, 38:139–149, 1954.

[21] W. L. Garrison. Connectivity of the interstate highway system. *Papers and Proceedings of the Regional Science Association*, 6:121–137, 1960.

[22] K. D. Mackenzie. Structural centrality in communications networks. *Psychometrika*, 31:17–25, 1966.

[23] F. R. Pitts. A graph theoretic approach to historical geography. *The Professional Geographer*, 17:15–20, 1965.

[24] D. L. Rogers. Sociometric analysis of interorganizational relations: application of theory and measurement. *Rural Sociology*, 39:487–503, 1974.

[25] Y. Kajitani and T. Maruyama. Functional expression of centrality in a graph - an application to the assessment of communication networks. *Electronics and Communication in Japan*, 59(24):9–17, 1976.

[26] J Nieminen. On the centrality in a directed graph. *Social Science Research*, 2:371–378, 1973.

[27] J Nieminen. On centrality in a graph. *Scandinavian Journal of Psychology*, 15:322–336, 1974.

[28] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.

[29] Stefan Wuchty. Scale-free behavior in protein domain networks. *Molecular Biology and Evolution*, 18(9):1694–1702, 2001.

[30] Linton C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[31] A. Bavelas. A mathematical model for group structure. *Applied Anthropology*, 7:16–30, 1948.

[32] Bernard S. Cohn and McKim Marriott. Networks and centres of integration in indian civilization. *Journal of Social Research*, 1:1–9, 1958.

[33] J. M. Anthonisse. The rush in a graph. Amsterdam: Mathematisch Centrum (mimeo), 1971.

[34] Mark EJ Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132, 2001.

[35] Peter V Marsden. Egocentric and sociocentric measures of network centrality. *Social Networks*, 24(4):407–422, 2002.

[36] B. Berche, C. von Ferber, T. Holovatch, and Yu. Holovatch. Resilience of public transport networks against attacks. *The European Physical Journal B*, 71(1):125–137, August 2009.

[37] Olaf Sporns, Christopher J Honey, and Rolf Kötter. Identification and classification of hubs in brain networks. *PloS one*, 2(10):e1049, 2007.

[38] A. Bavelas. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, 22:271–282, 1950.

[39] M. A. Beauchamp. An improved index of centrality. *Behavioral Science*, 10:161–163, 1965.

[40] R. L. Moxley and N. F. Moxley. Determining point-centrality in uncontrived social networks. *Sociometry*, 37:122–130, 1974.

[41] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31:581–603, 1966.

[42] Dirk Koschützki and Falk Schreiber. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regul Syst Bio*, 2:193–201, 2008.

[43] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of mathematical sociology*, 2(1):113–120, 1972.

[44] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953.

[45] Clarence H. Hubbell. An input–output approach to clique identification. *Sociometry*, 28:377–399, 1965.

[46] M. Taylor. Influence structures. *Sociometry*, 32:490–502, 1969.

[47] James S. Coleman, Elihu Katz, and Herbert Menzel. *Medical Innovation: A Diffusion Study*. Bobbs-Merrill, Indianapolis, 1966.

[48] Noah E. Friedkin. Theoretical foundations for centrality measures. *American Journal of Sociology*, 96:1478–1504, 1991.

[49] Akrati Saxena and Sudarshan Iyengar. Centrality measures in complex networks: A survey. *CoRR*, abs/2011.07190, 2020.

[50] Thomas W Valente. Social network thresholds in the diffusion of innovations. *Social Networks*, 18(1):69–89, 1996.

[51] Sebastian Wandelt, Xiaoran Sun, Zanin Massimiliano Feng, Daozhong, and Shlomo Havlin. A comparative analysis of approaches to network-dismantling. *Scientific Reports*, 8:13513, 2018.

[52] Erik Jenelius and Lars-Göran Mattsson. *Resilience of Transport Systems*. 05 2020.

[53] P. Holme, B.J. Kim, C.N. Yoon, and S.K. Han. Attack vulnerability of complex networks. *Physical Review E*, 2002.

[54] J. Smith and L. Johnson. Edge removal and network reconstruction in biological networks. *PLoS ONE*, 10(7):e0131180, 2015.

[55] Michele Bellingeri, Daniele Bevacqua, Francesco Scotognella, and et al. A compara-

tive analysis of link removal strategies in real complex weighted networks. *Scientific Reports*, 10(1):3911, 2020.

[56] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical Review Letters*, 87(19):198701, 2001.

[57] Yang Yang, Takashi Nishikawa, and Adilson E Motter. Small vulnerable sets determine large network cascades in power grids. *Science*, 358, 2017.

[58] Jose L Caldu-Primo, Elena R Alvarez-Buylla, and Jose Davila-Velderrain. Structural robustness of mammalian transcription factor networks reveals plasticity across development. *Scientific Reports*, 8(1):1–15, 2018.

[59] Lazaros K Gallos, Reuven Cohen, Panos Argyrakis, Armin Bunde, and Shlomo Havlin. Stability and topology of scale-free networks under attack and defense strategies. *Physical Review Letters*, 94(18):188701, 2005.

[60] Massimiliano Zanin and Fabrizio Lillo. Modelling the air transport with complex networks: A short review. *The European Physical Journal Special Topics*, 215:5–21, 2013.

[61] Christoforos N Schneider and Iordanis Koutsopoulos. Mitigation of malicious attacks on networks. *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, 2011.

[62] Fabiola Baltar and Ignasi Brunet. Unintended consequences and new challenges of online social systems: Methods, challenges, and applications. *Journal of Information Technology & Politics*, 9(1):1–20, 2012.

[63] Michele Bellingeri, Davide Cassi, and Simone Vincenzi. Efficiency of attack strategies on complex model and real-world networks. *Physica A: Statistical Mechanics and its Applications*, 414:174–180, November 2014.

[64] Duncan S Callaway, M E J Newman, Steven H Strogatz, and Duncan J Watts. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25):5468, 2000.

[65] Xiangtao Chen and Juan Li. Community detection in complex networks using edge-deleting with restrictions. *Physica A: Statistical Mechanics and its Applications*, 519:181–194, 2019.

[66] Sadegh Aliakbary, Jafar Habibi, and Ali Movaghar. Quantification and comparison

of degree distributions in complex networks. In *7th International Symposium on Telecommunications (IST)*, Tehran, Iran, September 2014.

[67] M. Faloutsos, P. Faloutsos, and C. Faloutsos. The structure of the internet graph. *IEEE/ACM Transactions on Networking*, 7(3):397–412, 1999.

[68] Table of critical values for the two-sample test. `https://web.archive.org/web/20130613002106/http://www.soest.hawaii.edu/wessel/courses/gg313/Critical_KS.pdf`. Memento from June 13, 2013 in the Internet Archive.

[69] I. M. Chakravarti, R. G. Laha, and J. Roy. *Engineering Statistics Handbook*. John Wiley and Sons, New York, NY, 1967.

[70] Gueorgi Kossinets and Duncan J Watts. Empirical analysis of an evolving social network. *Science*, 311:88–90, 2006.

[71] Harald Cramér. On the composition of elementary errors. *Skandinavisk Aktuarietidskrift*, 11:141–180, 1928.

[72] Richard von Mises. *Wahrscheinlichkeit, Statistik und Wahrheit*. Julius Springer, Vienna, Austria, 1928.

[73] Theodore W Anderson. On the distribution of the two-sample cramer-von-mises criterion. *The Annals of Mathematical Statistics*, pages 1148–1159, 1962.

[74] M. A. Stephens. Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347):730–737, 1974.

[75] Aric A Hagberg, Daniel A Schult, and Pieter J Swart. Exploring network structure, dynamics, and function using networkx. *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, 11:11–15, 2008.

[76] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

[77] Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The dynamics of viral marketing. In *Proceedings of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 568–576, 2007.

[78] Albert-László Barabási, Hawoong Jeong, Zoltán Néda, Erzsébet Ravasz, András Schubert, and Tamás Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3-4):590–614, 2002.

[79] Mark EJ Newman. *Networks*. Oxford University Press, 2018.

[80] Ernesto Estrada. Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103, 2005.

[81] Significance level. 0.05, n.d. Typical significance level used in hypothesis testing.