International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I12P103

Volume 72 Issue 12, 30-41, December 2024

Original Article

Key Nodes in Wireless Sensor Networks: Centrality Measures and Principal Component Analysis in Watts-Strogatz, Random Walk, and Barabási-Albert Models

Suneela Kallakunta[1], Alluri Sreenivas[2]

[1,2]Department of Electrical, Electronics and Communication Engineering, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.

[1]Corresponding Author : suneelakallakunta@gmail.com

Received: 24 August 2024 Revised: 11 November 2024 Accepted: 19 November 2024 Published: 25 December 2024

Abstract *- Wireless Sensor Networks (WSNs) have facilitated the advancement of communication systems. WSNs are deployed in sectors where metrics and performance factors are calculated mainly based on particular nodes, the importance of which should be emphasized. Centrality measures are enumerated to include Degree, Betweenness, Closeness, Eigenvector, Katz and Subgraph Centralities, which directly assess the importance of individual network nodes. The research explicitly considers the generation of WSNs with 150 nodes using the Watts-Strogatz, Random Walk and Barabasi-Albert models, paying attention to essential nodes that affect the network level and its features. Using Pearson correlation analysis, Kendall rank correlation analysis and Spearman rank correlation analysis determine how centrality measures correlate in each model. Principal Component Analysis (PCA) determines how many nodes need to be used to determine the centrality data to contain the most variance across the different models. The findings in this study also underscore the importance of the centrality measures in explaining the network's topology and make it clear why some nodes are more important than others; the information provided may assist in creating more efficient WSN structures.*

Keywords *- Wireless Sensor Networks (WSNs), Centrality measures, Key node identification, Principal component analysis (PCA), Network optimization.*

1. Introduction

Wireless Sensor Networks (WSNs) provide fundamental support for smart systems since they touch the areas of healthcare or the environmental systems. The spatially adaptive sensors that are the constituent of this network are tasked with receiving, transmitting and disseminating data of a particular range and area to a specified location for processing. Proper setup of the node management configuration is also critical in operationalizing WSN, as identifying some selected nodes, which are dispersed in the network, improves the performance and dependability of the system [2]. In network analysis, centrality measures are essential for assessing the relative importance of individual members depending on their location in a network geometry. Degree, Betweenness, Closeness, Eigenvector, Katz, and Subgraph Centralities are common measures of centrality, which rely on node connectivity, influencing communication channels and even enhancing the network [3,4,5]. These numerous studies highlight the centrality metrics of different networks, including WSNs, which improve fault tolerance, data routing efficiency and energy consumption [6-8]. Nonetheless, despite the plethora of literature on centrality in WSNs, understanding the relationship between centrality and different network models such as the Watts-Strogatz, Random Walk, and Barabási-Albert models appears to be an unexplored area. The relationship can be seen that the bulk of the earlier work focused on single centrality measure approaches within one class of network topology. Still, little is known about the approaches that cut across topologies, which are ideal for determining which aspects of the node's position across the network topologies are significant to the structural balance of the network. This research gap calls for a multi-model framework of a network-centric approach to establish the central nodes that are key to the efficiency of the WSNs. This study fills this gap in the research literature by systematically comparing centrality measures in 150 node WSNs designed with three models: the Watts-Strogatz model, which is best characterized by high clustering with small world tendency [9]; the Random Walk model, which depicts networks with random node mobility or shaking such as sensor networks with unpredictable movements [10,11]; the Barabási-Albert model reflected the networks which are scale-free or have a free degree distribution such as real and technological networks [12,13]. Moreover, it uses Pearson, Kendall rank, and Spearman correlations further to assess the relationship among the different centrality measures and thus identify some aspects where they complement or differ [14-16]. In the present work, Principal Component Analysis (PCA) reduces dimensionality while maximizing variance to make the complex interrelationships between centralities less complicated. It thus generates principal components that highlight the importance of the nodes [17-23]. The study provides a new perspective on key node identification by proposing ways of measuring the interaction of centrality across different types of networks. It guides how the design of WSNs can be done for improved robustness, optimality, and growth. The primary focus of this study is to bridge the gap in WSN optimization by explaining how various centrality measures work in different WSNs now generated based on other network models. The results provide further knowledge and practical implementation in contemporary communication systems and sound contributions to network theory by clearly pointing out the centrality measures that best express essential nodes in the network and improving the design and deployment of robust, scalable WSNs.

2. Centrality Measures in Wireless Sensor Networks

Centrality measures are commonly used tools in network analysis to determine the most key nodes. In WSN, these measures help to examine which nodes are essential for the proper operation of the network, for instance, for data relaying, communication, and redundancy. Below are descriptions of some of the most widely used centrality measures in WSN, along with their formulas and significance.

2.1. Degree Centrality (DC)

Degree centrality measures the importance of a node based on the number of direct connections it has with other nodes in the network. In this formula, deg(v) represents the degree of node v, meaning the number of edges connected to it, and N is the total number of nodes in the network. The concept behind degree centrality is straightforward: a node with more connections is generally more important because it can directly communicate with more network parts. In WSNs, nodes with a high degree of centrality often serve as hubs facilitating data exchange across the network [24].The degree centrality of a node v is given by

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

2.2. Betweenness Centrality (BC)

Betweenness centrality quantifies the importance of a node based on the number of shortest paths that pass through it. In this formula, σ_{st} denotes the total number of shortest paths between nodes s and t, while $\sigma_{st}(v)$ is the number of those paths that pass through node v. A high betweenness centrality indicates that a node acts as a critical bridge or bottleneck, playing a key role in maintaining communication across the network. In WSNs, such nodes are crucial for ensuring efficient data flow, particularly in networks with sparse connectivity. The betweenness centrality of node v is given by

$$C_B(v) = \sum_{s \neq t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{2}$$

2.3. Closeness Centrality (CC)

Closeness centrality reflects how close a node is to all other nodes in the network by calculating the average shortest path length from the node to all other nodes. Here, $d(v,t)$ is the shortest path distance between node v and node t. Nodes with high closeness centrality can reach other nodes more quickly, making them efficient for disseminating information throughout the network. In WSNs, such nodes are ideal for roles requiring quick data aggregation or distribution [25].
The closeness centrality of a node v is given by

$$C_C(v) = \frac{N-1}{\sum_{t \neq v} d(v,t)} \tag{3}$$

2.4. Eigenvector Centrality (EVC)

Eigenvector centrality measures a node's influence based on the importance of its neighbors. In the formula, M(v) represents the set of neighbors of node v, and λ is a constant (the largest eigenvalue of the adjacency matrix). A node with high eigenvector centrality is not only well-connected but is connected to other nodes that are themselves highly connected. This measure highlights nodes that are influential in the broader network context. In WSNs, nodes with high eigenvector centrality ensure robust and efficient network communication [26, 27]. The eigenvector centrality of a node v is given by

$$C_E(v) = \frac{1}{\lambda} \sum_{u \in \mathrm{M}(v)} C_E(u) \tag{4}$$

Where $C_E(u)$ is the Eigenvector centrality of node u, a neighbor of node v.

2.5. Katz Centrality (KC)

Katz centrality extends the idea of eigenvector centrality by considering all paths between nodes, not just direct connections, with attenuation based on path length. In this formula, α is a constant that scales the influence of neighbors, and β is a constant that adds value for each node, ensuring the centrality score is never zero. Katz centrality is useful for identifying nodes that are not only central but also influential across multiple levels of the network. In WSNs, this measure helps to identify nodes that play pivotal roles in spreading information across the network [28].
The Katz centrality of a node v is given by

$$C_K(v) = \alpha \sum_{u \in \mathrm{M}(v)} C_K(u) + \beta \tag{5}$$

Where $C_K(u)$ is the Katz centrality of node u, a neighbor of node v.

2.6. Subgraph Centrality (SC)

Subgraph centrality measures the importance of a node by counting the number of closed walks of different lengths that start and end at the node. In this formula, $\mu_k(v)$ denotes the number of closed walks of length k at node v. Nodes with high subgraph centrality are central to many small subgraphs within the network, which means they play a significant role in the local structure. In WSNs, these nodes are crucial for maintaining local network robustness and ensuring effective local communication [29]. The Subgraph centrality of node v is given by

$$C_S(v) = \sum_{k=0}^{\infty} \frac{\mu_k(v)}{k!} \qquad (6)$$

3. Results and Discussion

This research analysed the various centrality measures to identify the most critical nodes among the 150 analyzed WSNs. The set of centrality measures that cut across this study included Degree Centrality (DC), Betweenness Centrality (BC), Closeness Centrality (CC), Eigenvector centrality (EVC), Katz centrality (KC) and Subgraph centrality (SC).

Figures 1, 2 and 3 illustrate the configurations of the WSNs with 150 nodes constructed using the Watts-Strogatz model, Random Walk and Barabasi-Albert models, respectively. The networks were constructed using Watts-Strogatz, Random Walks, and Barabasi-Albert Models. Further, figures 2, 4 and 6 display the network's marked DC, BC, CC, EVC, KC and SC measure values. Using the libraries and functions related to network analysis, the research was implemented in Python. Tables 1, 3, and 5 show the rank order for the centrality measures for the Watts-Strogatz, Random Walk, and Barabási-Albert models, respectively. Tables 2, 4, and 6 also provide the correlation coefficients for the centrality measures computed for these 150 nodes WSNs developed using similar models.

3.1. First Experiment: WSN with 150 Nodes Generated Using the Watts-Strogatz Model

The first experiment involved the generation of a 150-node WSN using the Watts-Strogatz model. This model is characterized by its small-world properties, which balance high clustering and short path lengths. The network's incidence of node importance was exposed through the centrality measures for this network. As per degree centrality, nodes 127, 133 and 91 were ranked the highest as the hub nodes, which are central in sustaining interlinks and the transfer of information throughout the network. Betweenness centrality rated nodes 91, 112 and 108 as critical bridges or bottlenecks, which indicate their role in efficient data communication and offer a certain degree of robustness regarding damage control about node malfunctions. Closeness centrality analysis of the network stated specifically nodes 109, 112, and 49, which are also listed as the most important nodes, which are the nearest to the most significant number of other nodes and which can be used in cases where information is needed to be gathered quickly. Eigenvector centrality disclosed that nodes 127, 126 and 125 wield much power owing to their neighbours' authority level, which makes those three nodes extremely relevant as they guarantee an optimal level of correlation of communication transferred throughout the network. Katz centrality again nodes 127, 133 and 91 saying that they are striking nodes because of their effectiveness network wide.

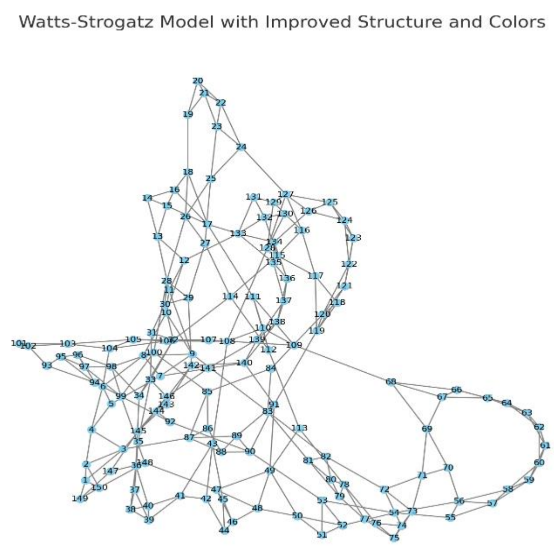

Fig. 1 WSN with 150 nodes generated using the Watts-Strogatz model

Table 1. Node rankings based on centrality metrics in a 150-node WSN generated using the Watts-Strogatz model

Rank	DC	BC	CC	EVC	KC	SC
1	127, 133	109	91	127	127	127
2		91	112	126	133	128
3	*	108	49	125	91	36
4		68	113	124	139	133
5		112	114	139	132	140
6		49	109	128	140	78
7		113	108	118	141	47
8		90	90	90	3	120
9		114	68	129	17	119
10		111	111	123	49	9
* 3,9,17, 32, 33,36, 47, 49						

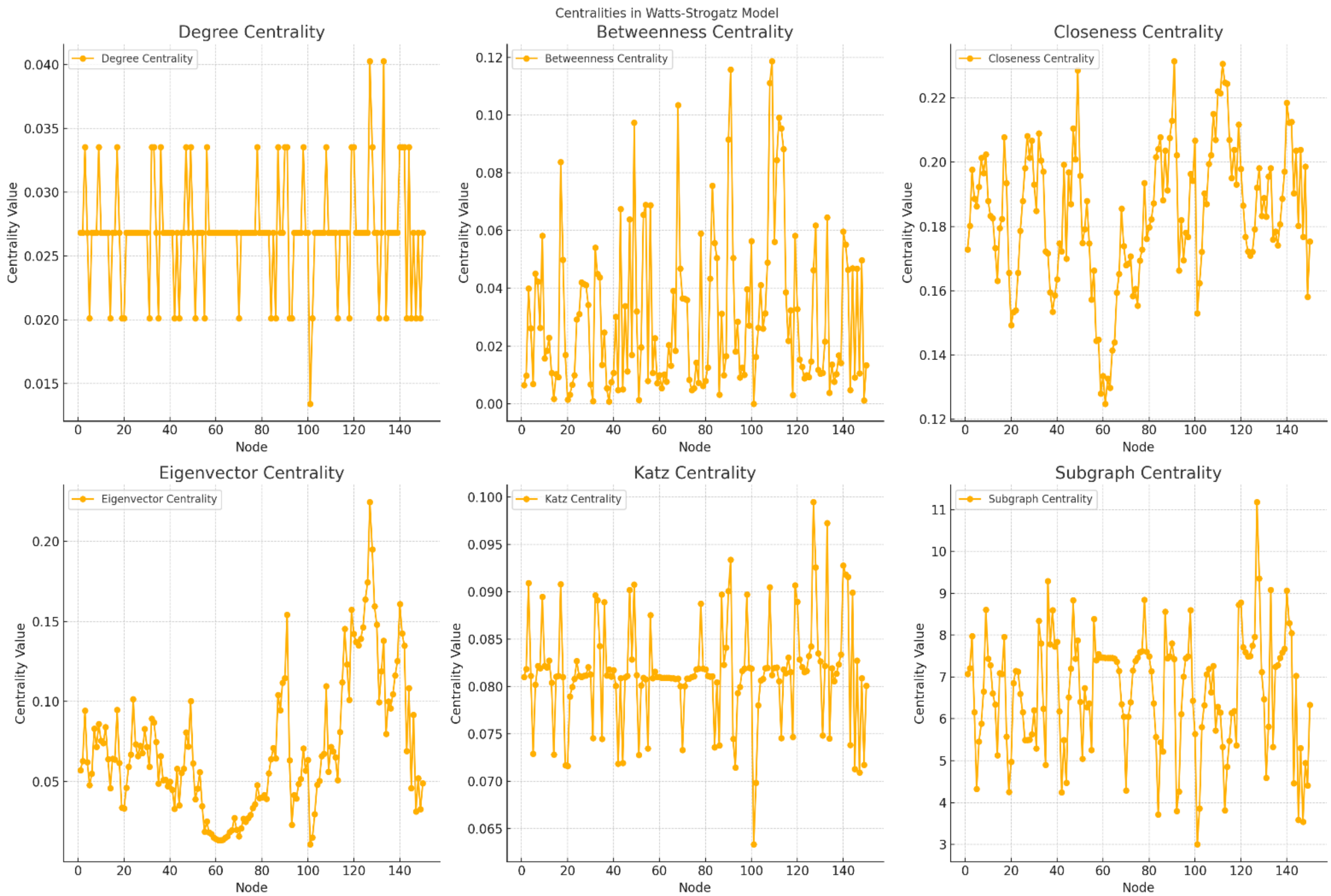

Fig. 2 Centrality values in a 150-node WSN generated using the Watts-Strogatz model

Table 2. Correlation coefficients among centrality values of a 150-node WSN generated using the Watts-Strogatz model

	DC	BC	CC	EVC	KC	SC
Pearson correlation						
DC	-	0.73	0.66	0.75	0.95	0.73
BC	-	-	0.81	0.48	0.68	0.23
CC	-	-	-	0.59	0.70	0.17
EVC	-	-	-	-	0.90	0.80
KC	-	-	-	-	-	0.79
SC	-	-	-	-	-	-
Kendall rank correlation						
DC	-	0.57	0.49	0.65	0.85	0.70
BC	-	-	0.63	0.39	0.49	0.22
CC	-	-	-	0.39	0.51	0.18
EVC	-	-	-	-	0.77	0.54
KC	-	-	-	-	-	0.64
SC	-	-	-	-	-	-
Spearman correlation						
DC	-	0.71	0.61	0.79	0.95	0.84
BC	-	-	0.82	0.55	0.68	0.33
CC	-	-	-	0.69	0.68	0.25
EVC	-	-	-	-	0.92	0.71
KC	-	-	-	-	-	0.83
SC	-	-	-	-	-	-

Finally, subgraph centrality ranked nodes 127, 128, and 36 at the top, which are significant in the preservation of the local architecture of the network. The correlation analysis among the centrality measures revealed strong associations between degree, Katz and subgraph centralities, which all significantly identify the key nodes within the network. On the other hand, betweenness and closeness centralities had moderate correlations with other measures, implying that they measure the node importance from a different perspective.

3.2. Second Experiment: WSN with 150 Nodes Generated Using the Random Walk Model

The second experiment, a 150-node WSN was generated using the Random Walk model, which simulates networks with nodes that exhibit random movement. This model is particularly relevant for scenarios where nodes are mobile or have unpredictable behavior.

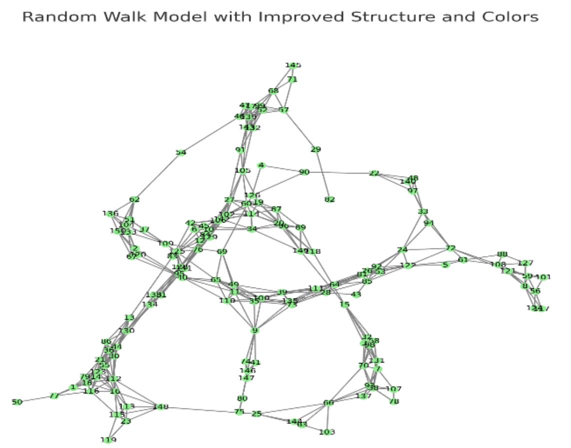

Fig. 3 WSN with 150 nodes generated using the Random Walk model

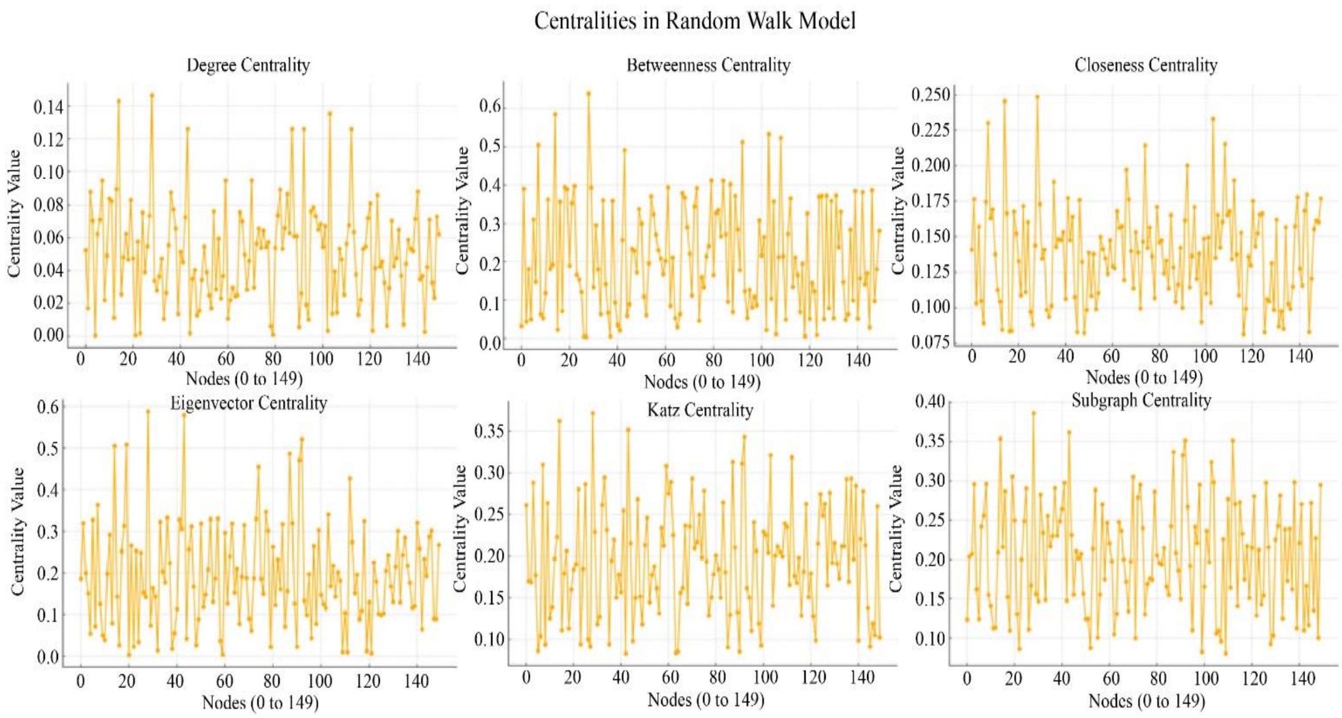

Fig. 4 Centrality values in a 150 node WSN generated using the Random Walk model

Table 3. Node rankings based on centrality metrics in a 150-Node WSN generated using the Random Walk model

Rank	DC	BC	CC	EVC	KC	SC
1	28	28	28	28	28	28
2	14	14	14	43	14	43
3	103	103	103	92	43	14
4	*	108	7	19	92	92
5		92	108	14	103	112
6		7	74	87	112	87
7		43	92	91	87	91
8	7, 59,70	79	66	74	91	103
9		84	112	112	7	19
10		87	35	7	59	70
*43, 87,92, 112						

Table 4. Correlation coefficients among centrality values of a 150-node WSN generated using the Random Walk model

	DC	BC	CC	EVC	KC	SC
Pearson correlation						
DC	-	0.94	0.73	0.88	0.98	0.96
BC	-	-	0.63	0.84	0.93	0.92
CC	-	-	-	0.72	0.76	0.66
EVC	-	-	-	-	0.95	0.94
KC	-	-	-	-	-	0.97
SC	-	-	-	-	-	-
Kendall rank correlation						
DC	-	0.88	0.77	0.71	0.90	0.89
BC	-	-	0.72	0.62	0.79	0.77
CC	-	-	-	0.87	0.87	0.87
EVC	-	-	-	-	0.92	0.92
KC	-	-	-	-	-	0.92
SC	-	-	-	-	-	-
Spearman correlation						
DC	-	0.94	0.74	0.82	0.98	0.94
BC	-	-	0.67	0.64	0.87	0.87
CC	-	-	-	0.97	0.97	0.97
EVC	-	-	-	-	0.92	0.97
KC	-	-	-	-	-	0.99
SC	-	-	-	-	-	-

Node 28 dominated the degree centrality rankings, establishing itself as a major hub in the network and likely playing a crucial role in maintaining network connectivity, particularly in dynamic environments. Similarly, node 28 also topped the betweenness centrality rankings, reinforcing its position as a critical bridge within the network, with nodes 14 and 103 also emerging as significant. Closeness centrality rankings once again placed node 28 at the forefront, followed by nodes 14 and 7, indicating their strategic positioning for rapid information dissemination and efficient data flow. In terms of eigenvector centrality, nodes 28, 43, and 92 were highlighted as the most influential, suggesting their key roles in the broader network structure. Katz's centrality also accentuated the importance of nodes 28, 43 and 14, as they greatly appeal throughout the network. Subgraph centrality also computed node 28 as the most centred and reconfirmed its significance. Also, the last correlation study for the Random Walk model demonstrated the strength of correlations between degree, Katz and subgraph centralities, which were as good as those obtained in the Watts-Strogatz model, which leads one to suggest that these measures are indeed very reliable in deciding the most critical nodes in various network models. The high correlation ratio between betweenness and other centralities in this model also suggests that nodes critical for network connectivity are essential for other network functionalities.

3.3. Third Experiment: WSN with 150 Nodes Generated Using the Barabási-Albert Model

The final experiment involved generating a 150-node WSN using the Barabási-Albert model, known for its scale-free properties. This model represents networks where a few nodes (hubs) have significantly higher connections, which is common in many real-world networks. Node 0 emerged as the most central across several measures, highlighting its role as a major hub within the network and its likely essential function in maintaining overall network connectivity.

In betweenness centrality, node 0 also topped the rankings, reinforcing its importance as a crucial bridge or bottleneck, with nodes 6 and 7 also identified as significant.

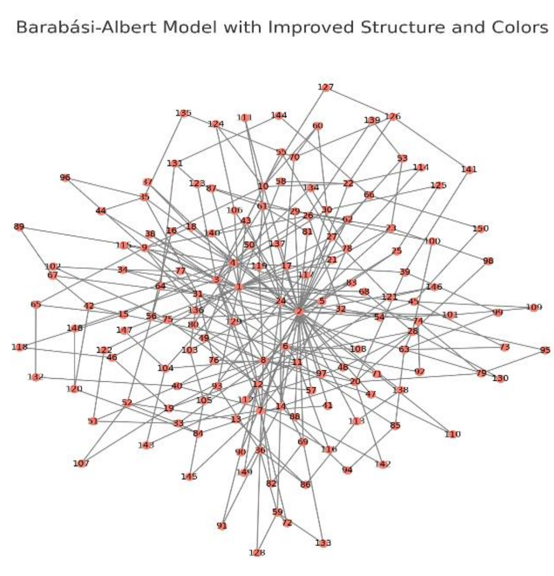

Fig. 5 WSN with 150 nodes generated using the Barabási-Albert model

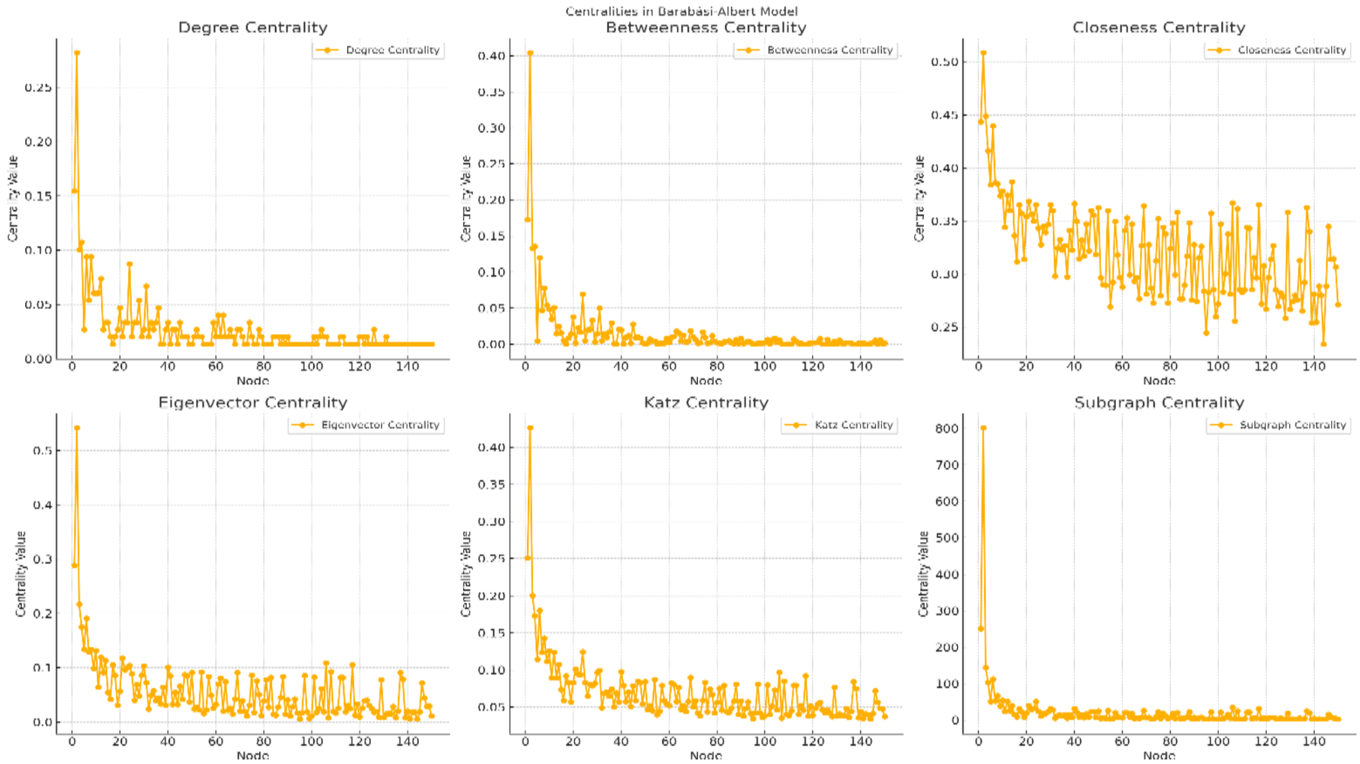

Fig. 6 Centrality values in a 150-node WSN generated using the Barabási-Albert model

Table 5. Node rankings based on centrality metrics in a 150-node WSN generated using the Barabási-Albert model

Rank	DC	BC	CC	EVC	KC	SC
1	0	0	0	0	0	0
2	6	6	6	6	6	6
3	7	7	7	7	7	7
4	11	11	11	12	11	12
5	12	15	12	11	12	11
6	1	12	1	1	1	1
7	15	1	15	5	5	5
8	10	26	5	9	10	9
9	5	10	10	8	15	8
10	26	5	9	10	9	10

Table 6. Correlation coefficients among centrality values of a 150-node WSN generated using the Barabási-Albert model

	DC	BC	CC	EVC	KC	SC
Pearson correlation						
DC	-	0.94	0.74	0.74	0.82	0.74
BC	-	-	0.67	0.64	0.87	0.87
CC	-	-	-	0.97	0.97	0.97
EVC	-	-	-	-	0.92	0.97
KC	-	-	-	-	-	0.99
SC	-	-	-	-	-	-
Kendall rank correlation						
DC	-	0.89	0.77	0.72	0.90	0.90
BC	-	-	0.72	0.63	0.79	0.77
CC	-	-	-	0.87	0.88	0.87
EVC	-	-	-	-	0.92	0.92
KC	-	-	-	-	-	0.92

SC	-	-	-	-	-	-
Spearman correlation						
DC	-	0.82	0.74	0.96	0.92	0.90
BC	-	-	0.87	0.72	0.77	0.67
CC	-	-	-	0.71	0.84	0.88
EVC	-	-	-	-	0.91	0.93
KC	-	-	-	-	-	0.95
SC	-	-	-	-	-	-

The closeness centrality analysis placed nodes 0, 6, and 7 at the top, indicating their strategic positioning within the network, which is vital for rapid data dissemination. The influence of node 0 was further stressed as the most influential with the application of eigenvector centrality, with nodes 6 and 7 coming next, implying that these nodes are vital not only in the direct connections but also in how they are central in the influence of other essential nodes of the network. Katz's centrality once again underscored the importance of nodes 0, 6 and 7, thus validating their relatively extensive centrality throughout the network. Finally, subgraph centrality rankings highlight the vital role played by node 0 as far as the network is concerned. The correlation analysis for the Barabási-Albert model revealed the same patterns as the previous models, with strong correlations obtained between degree, Katz, and subgraph centralities. This implies that the same nodes, when these measures, are always the most important across different network architectures. Also, the Barabási-Albert model's scale-free character accentuates the hubs' role in preserving the network's stability and resilience. The results from the three experiments endorse the relevance of centrality measures for selecting important nodes across varying WSN models. Degree, Katz, and subgraph centralities always pinpointed the most critical nodes in all the models, which attests to their ability to capture the key nodes in all configurations. This means that the Watts-Strogatz model, which has small world features, clustering, and short path lengths, is an essential parameter in preserving the network's performance. Mobil or dynamic nodes had the Random Walk model, highlighting the flow of information and the degree of network connectivity in the changing environment. The Barabási-Albert model, on the other hand, with its scale-free features, churned out the importance of hubs in the networks. This is very important for the design and analysis of WSNs. Network designers can reduce costs and increase the network's efficiency, robustness and expansion by focusing on the most critical nodes defined by centrality measures. The observed strong correlations between specific centrality measures suggest that these measures capture different dimensions of node importance and thus allow for a complete explanation of network behavior.

4. Principal Component Analysis (PCA) in Node Centrality Analysis

4.1. Introduction PCA in Network Analysis

PCA, also known as Principal Component Analysis, is viewed as one of the procedures suited for dimensionality reduction. It has the strength of transforming many interrelated variables into a smaller set of uncorrelated variables called principal components. In a more general sense, PCA could be useful in aiding data reduction and complexity simplification, such as when dealing with centrality measures in social network analysis. PCA is beneficial in network analysis since one can amalgamate centrality index measures into a few factors that explain the trends and factors that govern the node importance instead of simultaneously examining the factors individually. In the context of Multilayered networks, each of the layers also has various centrality measures (Degree Centrality, Betweenness Centrality, Closeness Centrality, Eigenvector Centrality, Katz Centrality, and Subgraph Centrality) that allow analysis of different levels of importance of nodes.

In many cases, however, these dimension measures are correlated because a node important in one dimension, say, connectivity, is likely to be important in many other dimensions; for example, it is likely to be important regarding the flow of information. Employing PCA can ascertain these interrelations, allowing us to reduce the number of variables by expressing the variables that account for most of the variance in node centrality as principal components. The dimensionality reduction algorithm here is based on the eigen-decomposition of the covariance matrix of the data and employs its eigenvalues and eigenvectors optimally. Eigenvalues are the principal weights of each component of a vector space; thus, they determine how much variation a principal component captures.

Hence, a greater eigenvalue would mean a more significant variance captured by that component. On the contrary, the eigenvector provides usable directions for each principal component. It specifies the weightings or loadings allocated to different factors, such as centrality measures, to form the component. This structural arrangement of relationships between the principal coordinates and variances as geometric figures and their counterparts enables the data to be automatically decomposed using PCA without losing its basic characters or structure.

For each principal component PC_i, the variance explained is calculated as:

$$\text{Variance Explained by } PC_i = \frac{\lambda_i}{\sum \lambda_j} 100 \quad (7)$$

Where λ_i is the eigenvalue associated with the i-th component, and $\sum \lambda_j$ is the sum of eigenvalues for all components. High eigenvalues indicate that a component captures significant patterns in the data, while low eigenvalues suggest minimal unique contribution.

4.2. Methodology: Application of PCA in Network Models

The approach aims at streamlining the interpretation of centrality measures within the framework of PCA applied in network models by focusing on three specific kinds of network structures, which are the Watts-Strogatz, Random Walk, and Barabási-Albert models, which have distinct characteristics contain elements that distinguish it from the others. For example, the Watts-Strogatz model has small-world properties that extend through high clustering and short average path lengths. On the other hand, the Random Walk model depicts a network in which there is a continuous movement from node to node, representing areas where the position and the connection of nodes constantly change. The Barabasi-Albert model also follows a scale-free network structure where network topologies comprise many nodes.

However, a few of them, called hubs, are highly connected and distributed according to a power law. For each of these networks, PCA was applied to six centrality measures: degree, betweenness, closeness, eigenvector, Katz, and subgraph centralities to isolate the most important principal components, which are PC1 and PC2 together with the eigenvalues associated with them. These components shape the space and explain the order-dependence of nodes within each network to explore the critical centrality features in these networks that differ in their structure.

4.2.1. First Principal Component (PC1)

Interpretation of PC1's Eigenvalue: The large eigenvalue of PC1 indicates that most centrality measures in this network are highly correlated and reflect a single dominant factor: node importance based on connectivity and clustering. Nodes with high degrees (many connections) or high closeness (proximity to many nodes) are essential in a small-world network for efficient communication. Thus, PC1 effectively captures this connectivity-clustering relationship.

$$PC_1 = w_{11} \cdot Degree + w_{12} \cdot Betweenness + w_{13} \cdot Closeness + \cdots \quad (8)$$

The high weights w_{1j} In PC1, it is indicated that degree and closeness centralities contribute the most, emphasizing nodes that anchor connectivity within clusters.

4.2.2. Second Principal Component (PC2)

Interpretation of PC2's Eigenvalue: The low eigenvalue suggests that PC2 does not offer significant insights into node importance. It may represent peripheral variations or minor structural differences not central to connectivity.

$$PC_2 = w_{21} \cdot Degree + w_{22} \cdot Betweenness + w_{23} \cdot Closeness + \cdots \quad (9)$$

Here, the low weights w_{2j} confirm PC2's lack of influence in the primary analysis of node centrality.

4.3. Experiment 1: PCA in the Watts-Strogatz Model

4.3.1. Network Characteristics of the Watts-Strogatz Model

The small world properties of the Watts-Strogatz model are a high level of clustering coefficient with short path lengths between nodes. These characteristics establish a network whereby connectivity and clustering influence the importance of nodes. Nodes with many connections to other nodes and near others are key to better communication and robustness.

4.3.2. Analysis of Principal Components in the Watts-Strogatz Model

First Principal Component (PC1)

Eigenvalue and Variance Explained: The eigenvalue associated with PC1 in the Watts-Strogatz model accounts for 70% of the variance in the centrality data. This extremely high percentage suggests that this single component can describe nearly all variability in centrality measures.

Interpretation of PC1's Eigenvalue: The large eigenvalue of PC1 indicates that most centrality measures in this network are highly correlated and reflect a single dominant factor: node importance based on connectivity and clustering. Nodes with high degrees (many connections) or high closeness (proximity to many nodes) are essential in a small-world network for efficient communication.

The high weights w_{1j} In PC1, it is indicated that degree and closeness centralities contribute the most, emphasizing nodes that anchor connectivity within clusters.

Second Principal Component (PC2)

Eigenvalue and Variance Explained: PC2 explains only 25% of the variance, a minimal value indicating that PC2 captures little additional information.

Interpretation of PC2's Eigenvalue: The low eigenvalue suggests that PC2 does not offer significant insights into node importance. The low weights w_{2j} confirm PC2's lack of influence in the primary analysis of node centrality.

4.4. Experiment 2: PCA in the Random Walk Model

4.4.1. Network Characteristics of the Random Walk Model

The Random Walk model simulates a network with node mobility, where connectivity can change dynamically. Stability and robust connectivity are key to node importance, as these properties maintain network cohesion despite potential movement or random disconnections.

4.4.2. Analysis of Principal Components in the Random Walk Model

First Principal Component (PC1)

Eigenvalue and Variance Explained: PC1 explains 68% of the variance in the Random Walk model, indicating that this component nearly captures all meaningful information.

Interpretation of PC1's Eigenvalue: The high eigenvalue for PC1 indicates that most centrality measures align around identifying nodes with stable and reliable connections as key. Degree and Eigenvector centralities, highlighting nodes with strong connectivity, are heavily weighted in PC1, as they are vital for maintaining connectivity. In PC1, the high loadings for Degree and Eigenvector centralities signify their dominance in determining stable, connected nodes in this model.

Second Principal Component (PC2)

Eigenvalue and Variance Explained: PC2 accounts for just 26% of the variance, adding minimal unique information. Interpretation of PC2's Eigenvalue: The very low eigenvalue means PC2 captures slight variations in node roles, likely from peripheral nodes with weaker connections. It does not contribute significantly to understanding overall node centrality. The low eigenvalues and weights signify that PC2 does not add valuable insight into node importance in a Random Walk model.

4.5. Experiment 3: PCA in the Barabási-Albert Model

4.5.1. Network Characteristics of the Barabási-Albert Model

The Barabási-Albert model is a scale-free network characterized by a few highly connected hubs that dominate connectivity, following a power-law degree distribution. These hubs are central to the network's resilience, holding the structure together.

4.5.2. Analysis of Principal Components in the Barabási-Albert Model

First Principal Component (PC1)

Eigenvalue and Variance Explained: PC1 explains almost 70% of the variance, indicating that it captures all the significant patterns of node centrality.

Interpretation of PC1's Eigenvalue: The high eigenvalue of PC1 confirms that all centrality measures converge in identifying hub nodes as the primary components of the network structure. With high Degrees and Eigenvector centralities, these nodes dominate in maintaining network stability. The high weights reflect the significance of hubs, where Degree and Eigenvector Centralities are most influential.

Second Principal Component (PC2)

Eigenvalue and Variance Explained: PC2 accounts for just 22% of the variance, adding minimal unique information. Interpretation of PC2's Eigenvalue: This low eigenvalue confirms that PC2 is unnecessary for capturing the network's structure. It may capture peripheral nodes with low connectivity but not impact network interpretation.

4.6. Summary of PCA Findings Across Models

The first principal component, PC1, had a very high eigenvalue in all three models, accounting for the most variance in centrality data. Consequently, it was the best proxy for the importance of nodes. This result demonstrates that PC1 can integrate many centrality indices into a single measure representing the structure of that model. In the Watts-Strogatz model, PC1 emphasized clustering and connection; in the Random Walk model, this stable linkage was active in a changing environment; and in the Barabási-Albert model, PC1 stressed the role of central dominant nodes. Across all models, PC2 had very low eigenvalues across all models, so it also contributed much variance that accounted only for low levels of structural differences, which do not significantly impact the relevance of nodes. The eigenvalues of the PCs have been consistent across models for PC1 and PC2, confirming that it is only necessary to shift one's focus on PC1 when defining node relevance for each model, which makes the analysis simple and robust while capturing the unique aspects of the structure and the centrality distribution of that class/type of networks. This approach provides a way to evaluate the nodes' centrality efficiently, providing a good basis for devising strategies to analyze and restructure networks based on the influence and connection of critical nodes.

5. Conclusion

This study presents WSNs generated from the Watts-Strogatz, Random Walk, and Barabasi-Albert models. The work primarily involves identifying key nodes using various centrality measures, such as Decide Centrality, Betweenness Centrality, Closeness Centrality, Eigenvector Centrality, Katz Centrality, and Subgraph Centrality. These conclusions confirm the importance of such measures in increasing network functions and improvements in WSNs and WSN management. A model of Watts-Strogaz, a small-world model, the networks have been characterized by high clustering and short path lengths with the Kolmogorov-Smirnov test relying on Degree, Katz and Subgraph centralities. Such a strong correlation among these measures further proves that these parameters are effective in locating the regions of great importance in the configuration of the network. This model proved that local clustering is crucial in upholding effective communication traces, essential for systems with continuous data transfer and communication with minimal delay. The Random Walk model mainly proved the importance of node mobility and connectivity as they are essential factors that aid in network performance in networks that have mobile or dynamic nodes.

According to the centrality measures provided, the consistent node 28 acted as a centre of various activities, indicating that during all the activities, node 28 was necessary for the network and the dispersal of data across the network. From the model, the strong relationship that was observed to exist between Degree, Katz and Subgraph centralities demonstrates enough explainability of node importance in that significant alterations are commonplace to the network topology. The Barabási-Albert model has gained popularity for its scale-free nature. It also emphasized the criticality of the bulk of the hubs in any network regarding connection and diversity. On the scope, node 0, turn first, has been observed to be the most central in all measures, which asserts its importance within the hierarchical structure of the network. Applying PCA, it was established that almost all the variance of the centrality data could be explained by the first principal component, which reveals the importance of some hubs in this model. It weakens the intricate relationships among the centrality measures, making locating scale-free networks' most essential nodes easier. The use of individual PCA across all three models indicated that inter-nodal variance in centrality measures is primarily explained by most of a single principal component, which points to node degree semantics or having a handful of critical nodes. This finding is beneficial regarding understanding how the analysis of the various centrality measures can be obfuscated even more to concentrate on the rudimentary factors related to the importance of nodes. In this way, the quality of the network analyses can be improved further. This research stresses the importance of centrality measures for understanding and improving the design of WSNs for various topologies. The pattern of Degree, Katz and Subgraph Centralities for significant nodes and the same metrics on different models indicate that node's degree, Katz and subgraph centrality are some of the most effective measures of the degree of nodes.

The research has practical applications that will assist in designing and optimizing WSNs to achieve robust and efficient networks that can meet the requirements of various applications.

Such outcomes are of significant significance for operating and managing WSNs in many real applications where stability and performance of the network will be the key.

Acknowledgment

We want to express our gratitude for the invaluable support provided by GITAM University, Visakhapatnam, India, throughout this research.

References

[1] I.F. Akyildiz et al., "A Survey on Sensor Networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102-114, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosa, "Wireless Sensor Network Survey," *Computer Networks*, vol. 52, no. 12, pp. 2292-2330, 2021.[CrossRef] [Google Scholar] [Publisher Link]

[3] Linton C. Freeman, "A Set of Measures of Centrality Based on Betweenness," *Sociometry*, vol. 40, no. 1, pp. 35-41, 1977. [CrossRef] [Google Scholar] [Publisher Link]

[4] Phillip Bonacich, "Power and Centrality: A Family of Measures," *American Journal of Sociology*, vol. 92, no. 5, pp. 1170-1182, 1987. [CrossRef] [Google Scholar] [Publisher Link]

[5] M.E. J. Newman, "A Measure of Betweenness Centrality Based on Random Walks," *Social Networks*, vol. 27, no. 1, pp. 39-54, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ulrik Brandes, "A Faster Algorithm for Betweenness Centrality," *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163-177, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[7] Masood Alam Abbasi, "Realization of centrality measure on Wireless Sensor Network," *In Proceedings of the 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*, Karachi, Pakistan, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Ameer Ahmed Abbasi, and Mohamed Younis, "A Survey on Clustering Algorithms for Wireless Sensor Networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826-2841, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[9] Duncan J. Watts, and Steven H. Strogatz., "Collective Dynamics of 'Small-World' Networks," *Nature*, vol. 393, no. 6684, pp. 440-442, 1998. [CrossRef] [Google Scholar] [Publisher Link]

[10] László Lovasz, "Random Walks on Graphs: A Survey," *Combinatorics, Paul Erdős also Eighty*, vol. 2, pp. 1-46, 1993. [Google Scholar] [Publisher Link]

[11] Asaf Nachmias, *Random Walks and Electric Networks*, Lecture Notes in Mathematics, Springer, Cham, vol. 2243, pp. 11-31, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Albert-László Barabási, and Réka Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, no. 5439, pp. 509-512, 1999. [CrossRef] [Google Scholar] [Publisher Link]

[13] Réka Albert, Hawoong Jeong, and Albert-László Barabási, "Error and Attack Tolerance of Complex Networks," *Nature*, vol. 406, no. 6794, pp. 378-382, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jan Hauke, and Tomasz Kossowski, "Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data," *Geographical questions*, vol. 30, no. 2, pp. 87-93, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[15] M. G. Kendall, “A New Measure of Rank Correlation,” *Biometrika*, vol. 30, no. 1-2, pp. 81-93, 1938. [CrossRef] [Google Scholar] [Publisher Link]

[16] Charles Spearman, “The Proof and Measurement of Association Between Two Things,” *The American Journal of Psychology*, vol. 15, no. 1, pp. 72-101, 1904. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ian T. Jolliffe, and Jorge Cadima, “Principal Component Analysis: A Review and Recent Developments,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 1-16, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[18] Hervé Abdi, and Lynne J. Williams, “Principal Component Analysis,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433-459, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[19] Meimei Ding, and Hui Tian, “PCA-Based Network Traffic Anomaly Detection,” *Tsinghua Science and Technology*, vol. 21, no. 5, pp. 500-509, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[20] Yihua Liao, and V. Rao Vemuri, “Use of K-Nearest Neighbor Classifier for Intrusion Detection,” *Computers & Security*, vol. 21, no. 5, pp. 439-448, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[21] Shancang Li, Li Da Xu, and Shanshan Zhao, “The Internet of Things: A Survey,” *Information Systems Frontiers*, vol. 17, no. 2, pp. 243-259, 2015.[CrossRef] [Google Scholar] [Publisher Link]

[22] Samuel Kofi Erskine, “Secure Data Aggregation Using Authentication and Authorization for Privacy Preservation in Wireless Sensor Networks,” *Sensors*, vol. 24, no. 7, pp. 2090, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Omar Banimelhem, Eyad Taqieddin, and Ibrahim Shatnawi, “An Efficient Path Generation Algorithm Using Principal Component Analysis for Mobile Sinks in Wireless Sensor Networks,” *Journal of Sensor and Actuator Networks*, vol. 10, no. 4, pp. 1-22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Mark Newman, *Networks*, 2nd ed., Oxford University Press, United States, 2018. [Google Scholar] [Publisher Link]

[25] Tore Opsahl, Filip Agneessens, and John Skvoretz, “Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths,” *Social Networks*, vol. 32, no. 3, pp. 245-251, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[26] Phillip Bonacich, “Power and Centrality: A Family of Measures,” *American Journal of Sociology*, vol. 92, no. 5, pp. 1170-1182, 1987. [CrossRef] [Google Scholar] [Publisher Link]

[27] Mark Newman, *Networks: An Introduction*, Oxford University Press, 2010. [CrossRef] [Publisher Link]

[28] Leo Katz, “A New Status Index Derived from Sociometric Analysis,” *Psychometrika*, vol. 18, no. 1, pp. 39-43, 1953. [CrossRef] [Google Scholar] [Publisher Link]

[29] Ernesto Estrada, and Juan A. Rodríguez-Velázquez, “Subgraph Centrality in Complex Networks,” *Physical Review E*, vol. 71, no. 5, 2005.[CrossRef] [Google Scholar] [Publisher Link]