MJFAS
Malaysian Journal of Fundamental and Applied Sciences

Moiz *et al.* | Malaysian Journal of Fundamental and Applied Sciences, Vol. 21 (2025) 2576–2587

RESEARCH ARTICLE

A Novel Weighted TOPSIS-based Method for Node Centrality Analysis in Complex Networks

Abdul Moiz[a,*], Ubaida Fatima[b], Muhammad Zeeshan Ul Haque[a]

[a]Department of Biomedical Engineering, Salim Habib University, Karachi 74900, Pakistan; [b]Department of Mathematics, NED University of Engineering & Technology, Karachi 75270, Pakistan

Abstract Complex networks are prevalent across domains such as biology, social science, and engineering, where identifying critical nodes is essential for applications such as disease control, traffic optimization, and discovering pivotal proteins. Various centrality metrics exist to assess node importance, each with unique strengths and limitations in capturing significance. Recently, the Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) method is used to pinpoint key nodes by considering centrality measures as multiple attributes of the network. This paper introduces a Weighted Modified-TOPSIS (WM-TOPSIS) method, enhancing TOPSIS by integrating multiple topological centrality measures and proposing a novel benchmark-based, data-driven algorithm to dynamically assign weights to attributes. The effectiveness of WM-TOPSIS was evaluated on Protein-Protein Interaction (PPI) networks of Saccharomyces cerevisiae (Yeast) and Escherichia coli (E. coli). On Yeast, WM-TOPSIS achieved a sensitivity (S_N) of 0.5150, specificity (S_P) of 0.7455, and accuracy of 0.6927. On E. coli, it recorded an S_N of 0.8031, S_P of 0.4355, and accuracy of 0.4697. These results highlight WM-TOPSIS's potential to improve the detection of critical nodes, offering a robust framework for biological network analysis and broader applications.

Keywords: Complex network, Node centrality, Critical nodes, Benchmark-based weighting, Weighted Modified-TOPSIS (WM-TOPSIS).

Introduction

The study of complex networks is gaining significance within network science due to their ability to effectively model intricate systems that exist in real life. In recent years one vital area has been stressed in the literature; this is finding critical nodes embedded in complex networks. Within the study of complex networks, node centrality analysis plays a crucial role in identifying the most influential nodes. Understanding which nodes are central is vital for various applications, such as cybersecurity, wireless sensor networks, cryptocurrency, water distribution systems, transportation, social networks, and biological systems. In cybersecurity, Hamouda *et al*. [1] leverage node centrality to predict cyberattack vulnerabilities, optimizing resource allocation to enhance network resilience through nonlinear models. In wireless sensor networks, centrality measures optimize energy efficient routing by identifying key communication hubs. Kallakunta and Sreenivas [2] apply diverse centrality metrics to select critical nodes, while Jain *et al*. [3] propose a novel cluster centrality measure to enhance scalability and algorithm efficiency. Rehman *et al*. [4] proposed the Digital Currency Network Centrality Measure (DCNC) for cryptocurrency transaction networks, correlating it with standard centrality metrics to identify key nodes, detect irregularities, and deepen insights into decentralized digital currency systems. For water distribution systems, Gopalsamy *et al*. [5] introduce Distance Laplacian Energy Centrality (DLC), integrating topological and hydraulic properties to pinpoint critical nodes, outperforming traditional measures in reliability.

In biological systems, the analysis of protein-protein interaction (PPI) networks has gained significant attention, as these networks are often depicted as graphs, where nodes symbolize proteins and edges

***For correspondence:**
abdul.moiz@shu.edu.pk

Received: 12 May 2025
Accepted: 13 August 2025

signify interactions between them. Within graph theory, a significant area of research focuses on identifying critical nodes (proteins) [6], [7], [8], [9], [10]. Identifying critical proteins in PPI networks is crucial for understanding disease mechanisms. Without them, disorders can spread quickly. Although, pinpointing critical proteins in PPI networks poses a challenge due to the data being noisy and incomplete [11]. To this end, various computational approaches have been proposed for solving this problem. According to their essential differences, these approaches can be categorized into topology-based, biological information-based, and hybrid-based (combination of both) techniques.

Topology-based methods depend on the structural properties of PPI networks, encompassing degree, closeness, betweenness, eigenvector, subgraph, and network centrality, for assessing the significance of proteins. These approaches operate under the assumption that proteins exhibiting higher scores are more prone to being essential. Nevertheless, these methods come with certain limitations, including oversight of the functional roles of the proteins, and an inability to identify low-connectivity essential proteins. Some examples of topology-based methods are DC [12], SC [13], BC [14], IC [15], EC [16], CC [17], LAC [18], and GCCDC [19].

Biological data, such as gene expression, subcellular localization, gene ontology, protein complexes, and functional modules, is utilized in biological information-based approaches to increase critical protein identifications. Unlike topology-based methods, these methods can reduce the impact of false-positive interactions and improve the biological relevance of predictions. However, these methods are often more complex and computationally expensive to run, and their effectiveness depends on the availability and quality of the biological data. Some examples of methodologies that utilize biological information include PEMC [20], SPP [21], RSG [22], and GOS [23] to improve the accuracy of identifying critical proteins.

A combination of both topology and biological information-based (hybrid) methods assume that proteins with high centrality scores and high functional similarity or specificity are considered significant. Hybrid methods combine the strengths of both methods, capturing structural and functional aspects of protein importance. However, these methods are more complex and computationally intensive. Some examples of hybrid methods include LIDC [24], PeC [25], CoEWC [26], and ION [27].

Current studies have explored ***Multi-Criteria Decision Making (MCDM)*** methods to integrate different centrality metrics for pinpointing critical nodes in intricate network structures. Within the MCDM domain, the ***Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)***, proposed by Hwang and Yoon (1981), is highly esteemed and widely adopted due to its simplicity and the fundamental concept that the optimal solution is the one that closest to the positive ideal solution and furthest from the negative ideal solution [28].

Several studies have explored the application of TOPSIS in this domain. Du *et al*. [29] were among the first to apply TOPSIS to identify influential nodes, recognizing its ability to integrate different centrality measures. However, the traditional TOPSIS method assumes equal weights for all attributes, which is often unrealistic. Consequently, several studies have focused on developing methods to assign weights to attributes. For instance, J. Hu *et al*. [30] proposed a W-TOPSIS that assigned weights based on node spreading ability simulated via the SIR model, outperforming classical TOPSIS in accuracy. Fei *et al*. [31] introduced a relative entropy-based method combined with TOPSIS, leveraging multiple centrality measures to rank nodes across four real networks. Their approach improved ranking accuracy but focused narrowly on entropy, potentially overlooking other structural properties. Y. Yang *et al*. [32] developed an MCDM-based method integrating degree, closeness, and betweenness centrality, with entropy weighting to reduce subjective biases. Tested on four real networks, their approach offered comprehensive node rankings but was constrained by its focus on static networks. The application of TOPSIS has also been specifically explored in the context of social networks. Ishfaq *et al*. [33] employed an entropy weighting technique in conjunction with TOPSIS to estimate node importance in social networks.

In addition to entropy, various weighted approaches have been employed to address node centrality issues. P. Yang *et al*. [34] proposed a dynamic weighted TOPSIS method, assigning weights using grey relational analysis and the Susceptible-Infected-Recovered (SIR) model. Applied to three real networks, their method outperformed single-indicator and weighted TOPSIS approache that discussed in [30]. However, the reliance on SIR simulations increases computational complexity, limiting scalability. P. Yang *et al*. [35], also in 2019, introduced a multi-attribute ranking method using grey relational analysis (GRA) and a new local centrality measure incorporating multi-layer neighbors and clustering coefficients. Applied to nine networks, their modified dynamic weighted TOPSIS method improved accuracy, yet GRA's sensitivity to reference sequences limits robustness. Y. Yang *et al*. [36] proposed an improved TOPSIS method integrating degree, closeness, and betweenness centrality, with objective weight

assignment. Validated through Susceptible-Infected (SI) model experiments. However, its reliance on a limited set of centrality measures reduces comprehensiveness

This research introduces a **Weighted Modified-TOPSIS (WM-TOPSIS)** approach to evaluate the importance of nodes in complex network structures. WM-TOPSIS enhances traditional TOPSIS by incorporating a benchmark based, data-driven weighting approach that adapts to the topological structure of the network.

This research contributes the following advancements through the Weighted Modified-TOPSIS (WM-TOPSIS) approach:

- A novel dynamic weighting mechanism driven by ground truth influential nodes, overcoming equal and subjective weight limitations.
- Integration of eight topological metrics, including the Global Clustering Coefficient-Dependent Degree Centrality, enhancing multi-attribute analysis by capturing clustering effects in PPI networks.
- Validation through experiments on real-world biological networks, confirming the efficacy of the introduced method.

Materials and Methods

Centrality Metrics as Evaluation Criteria

To capture the multidimensional nature of node influence, eight topological centrality measures are used as evaluation criteria. These measures are carefully selected to capture a diverse range of perspectives on node importance, considering both local and global structural properties. Table 1 summarizes these metrics. These measures serve as input attributes for the decision matrix in the TOPSIS framework.

Table 1. Summary of topological centrality metrics

Centrality metrics	Definition	Ref.
Degree centrality	This metric helps to identify the critical node based on its number of interactions or connections.	[12]
Betweenness centrality	This measure of a node shows how often it lies on the fastest (shortest) routes between other nodes in the network.	[14]
Closeness centrality	CC is a metric that determines the extent to which a node is to all other nodes. In PPI networks, CC can be useful in identifying proteins that are well-connected to other proteins.	[17]
Eigenvector centrality	Eigenvector centrality is a metric that evaluates how important a protein is in terms of the number and quality of its interactions with other proteins. A protein that interacts with many other proteins, especially those that are also highly connected, is likely to have a high EC value.	[16]
Harmonic centrality	Harmonic centrality is a type of closeness centrality. Within a graph, the CC of a vertex (let's call it 'v') is determined by taking the reciprocal of the mean distance of the shortest paths from 'v' to all other nodes. On the other hand, harmonic centrality takes a different approach. Instead of taking the reciprocal of the average distance, it sums up the reciprocals of the distances from 'v' to each vertex in the graph.	[37]
PageRank centrality	PageRank centrality is a modification of the EC metric. It determines the relevance of a node by evaluating both the number and the strength of its connections.	[38]
Subgraph centrality	The subgraph centrality, denoted as SC, quantifies the total count of closed walks that involve protein or node "i". It assigns greater importance to closed walks that are shorter in length.	[13]
Global clustering coefficient-dependent degree centrality	GCCDC proposed by U. Fatima *et al*., is a novel centrality metric that is developed using the global clustering coefficient to provide a comprehensive view of the clustering in the network.	[19]

Classical TOPSIS Framework

TOPSIS ranks nodes by comparing their Euclidean distances from the ideal and non-ideal solutions. The standard procedure includes [28]:

1. Construct a decision matrix that shows the values of each option for each criterion.

$$\mathrm{DM} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m1} & \cdots & y_{mn} \end{bmatrix} \quad (1)$$

2. Normalize the decision matrix so that the values are comparable and have the same scale.

$$n_{ij} = \frac{y_{ij}}{\sqrt{\sum_{j=1}^{m} {y_{ij}}^2}}, i = 1,\ldots,m;\ j = 1,\ldots,n. \quad (2)$$

where:

n_{ij} is the normalized value.
y_{ij} is the original value.

3. Compute the weights of each criterion as follows;

$$w_k = \frac{1}{n};\ n\ is\ the\ total\ number\ of\ the\ criteria. \quad (3)$$

where:

w_k is the weight of criterion k. The weight remains consistent for each 'k' criterion

4. Multiply the normalization matrix by the weights w_k to get a weighted decision matrix (WM).

$$\mathrm{WM} = \begin{bmatrix} y_{11}w_1 & y_{12}w_2 & \cdots & y_{1n}w_k \\ y_{21}w_1 & y_{22}w_2 & \cdots & y_{2n}w_k \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1}w_1 & y_{m1}w_2 & \cdots & y_{mn}w_k \end{bmatrix} \quad (4)$$

5. Determine the ${P_i}^+$ and ${P_i}^-$ ideal solutions by identifying the highest and lowest values for each criterion, accordingly.

$${P_i}^+ = \left\{ {}^{max}_{i}(wm)_{ij} \mid j \in C \right\} \quad (5)$$

$${P_i}^- = \left\{ {}^{min}_{i}(wm)_{ij} \mid j \in C \right\} \quad (6)$$

Where the set of different criteria is represented by C.

6. Calculate distances from ideal solutions ${D_i}^+$ and ${D_i}^-$. The distances of each option from ${P_i}^+$ and ${P_i}^-$ are calculated using Euclidean distance. The formulas are:

$$D_i^{+} = \sqrt{\sum_{j=1}^{m} (wm_{ij} - P_j^{+})^2} \quad (7)$$

$$D_i^{-} = \sqrt{\sum_{j=1}^{m} (wm_{ij} - P_j^{-})^2} \quad (8)$$

7. Calculate each option's relative closeness (RC) to the ideal solution by dividing D_i^{-} to $(D_i^{+} + D_i^{-})$.

$$RC_i = \frac{D_i^{-}}{D_i^{+} + D_i^{-}} \quad (9)$$

Benchmark based Weighting (Proposed Modification)

Instead of assigning equal or subjective weights, our method computes data-driven dynamic weights based on how well each centrality measure captures known influential nodes. The computation follows these steps:

1. Nodes are ordered from highest to lowest according to their centrality measures.
2. A threshold t is set for each centrality metric to identify top-ranked nodes (e.g., top 40% nodes, adjustable based on network size).
3. The top t nodes for each metric are compared with a predefined benchmark (ground truth) set of influential nodes.
4. Each centrality metric is assigned a grade g (ranging from 1 to n) based on its overlap with ground truth, where (g = 1) indicates the highest alignment.
5. The final weight w'_k for each centrality metric is computed using:

$$w'_k = \frac{\frac{1}{g_i}}{\sum_i^n \frac{1}{g_i}} \quad (10)$$

Where:
w'_k is the weight of criterion k (centrality).
n is the total number of the criteria.

This benchmark-based weighting adapts to the network's topology, overcoming the equal and subjective weight limitation, while remaining computationally efficient compared to simulation-based methods.

Example Explanation

To illustrate the benchmark-based weighting, consider a hypothetical simple Protein-Protein Interaction (PPI) network with six nodes (P1–P6), as shown in Figure 1, where P1, P2, P3, and P4 are critical benchmark nodes (ground truth proteins).

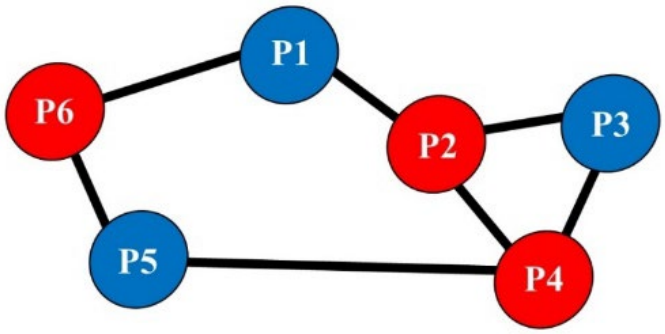

Figure 1. A small protein-protein interaction (PPI) network with six proteins

Each node represents a protein, with undirected, unweighted edges indicating interactions. The eight-centrality metrics: Degree (DC), Harmonic (HC), Closeness (CC), PageRank (PR), Betweenness (BC), Subgraph (SC), Eigenvector (EC), and Global Clustering Coefficient-Dependent Degree Centrality (GCCDC or GD) are computed as shown in Table 2.

Table 2. Sorted nodes of the network based on descending order of topological centrality measures

Nodes	DC	Nodes	BC	Nodes	CC	Nodes	EC
P2	6	P2	2.5	P4	0.166	P2	0.523
P4	6	P4	2.5	P2	0.143	P4	0.523
P1	4	P1	1.5	P1	0.125	P3	0.420
P3	4	P5	1.5	P5	0.125	P1	0.317
P5	4	P6	1	P3	0.111	P5	0.317
P6	4	P3	0	P6	0.111	P6	0.273
Continue…							
Nodes	HM	Nodes	PR	Nodes	SC	Nodes	GD
P2	4	P2	0.2065	P2	36.55397	P2	4.2
P4	4	P4	0.2065	P4	36.55397	P4	4.2
P1	3.5	P6	0.1502	P3	25.63724	P1	2.8
P5	3.5	P1	0.1474	P1	16.21632	P3	2.8
P3	3.333	P5	0.1474	P5	16.21632	P5	2.8
P6	3.333	P3	0.142	P6	13.45538	P6	2.8

Nodes are ranked by each metric, we then select a threshold level, t = 4, for each metric. These are compared to the benchmark set, graded based on overlap (1 for highest, up to 8 for lowest), and weights are calculated each metric using Eq. (10). Table 3 summarizes the weighting process for the example network.

Table 3. Benchmark based weight computation for eight centrality metrics

Metric	Top-4 Nodes	Overlap with Benchmark Nodes (P1, P2, P3, P4)	Grade (g)	Weight (w'_k)
DC	P2, P4, P1, P3	P1, P2, P3, P4 (4/4)	1	0.1714
BC	P2, P4, P1, P5	P2, P4 (2/4)	3	0.0571
CC	P4, P2, P1, P5	P1, P2, P4 (3/4)	2	0.0857
EC	P2, P4, P3, P1	P1, P2, P3, P4 (4/4)	1	0.1714
HM	P2, P4, P1, P5	P1, P2, P4 (3/4)	2	0.0857
PR	P2, P4, P6, P1	P1, P2, P4 (3/4)	2	0.0857
SC	P2, P4, P3, P1	P1, P2, P3, P4 (4/4)	1	0.1714
GD	P2, P4, P1, P3	P1, P2, P3, P4 (4/4)	1	0.1714

These weights, reflecting the network's topology, are applied to the decision matrix for node ranking. The proposed method's flow chart is depicted in Figure 2.

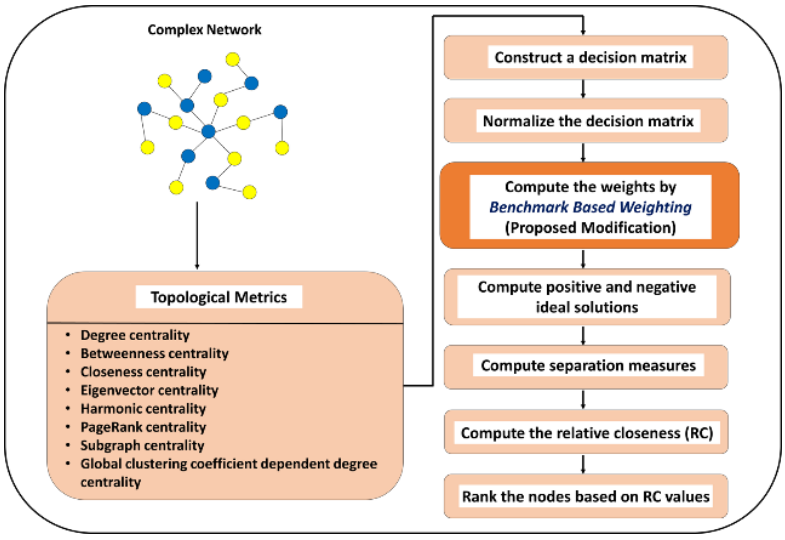

Figure 2. Flowchart of a novel weighted TOPSIS-based method for node centrality analysis in complex networks

Results and Discussion

Experimental Datasets

To assess the performance of the proposed method, we applied it to ***Saccharomyces cerevisiae (Yeast)*** and ***Escherichia coli (E. coli)*** networks to identify critical proteins. The datasets were retrieved from ***DIP*** [39]. In the Saccharomyces cerevisiae network, we considered 5093 distinct proteins and 24743 interactions, while in the E. coli network, we considered 2727 proteins and 11803 interactions. These numbers were obtained after filtering out self-interactions and repeated interactions during the experiment. Table 4 provides a comprehensive overview of the datasets. For benchmarking, we have used a dataset of ground truth critical proteins from [40]. The dataset for Saccharomyces cerevisiae contains 1167 critical proteins, while the E. coli dataset contains 254.

Table 4. Key traits of the saccharomyces cerevisiae and escherichia coli network

Organism	Nodes	Edges	Network diameter	Density	Modularity
Saccharomyces cerevisiae	5093	24743	10	0.002	0.497
Escherichia coli	2727	11803	12	0.003	0.473

Comparison between Proposed Approach and Other Metrics

In this experiment, we evaluated our proposed approach with the classical original TOPSIS (equal weights), and eight topological centrality metrics (DC, HM, PR, BC, SC, CC, EC, and GD). These methods were applied to Saccharomyces cerevisiae and E. coli networks, which are real-world biological networks. First, we calculated the score of each measure and sorted them in descending order. We then compared the top 100, 500, 600, 800, 1200, and 1600 proteins to ground truth benchmark proteins to determine what proportion of them were actually influential proteins. The comparison results are displayed in Table 5. Variations in the fundamental ideas and primary areas of interest across different methods could yield diverse outcomes.

The results show that the proposed WM-TOPSIS method performed well in identifying influential or critical proteins in both Saccharomyces cerevisiae and E. coli networks. In Saccharomyces cerevisiae, the proposed method (WM-TOPSIS) achieved the highest number of true influential proteins in three levels (600, 1200, 1600) among all the measures discussed above. In the top 100 level in Saccharomyces cerevisiae, it was tied with DC and GD. In the top 500 and 800 levels, DC and GD were

slightly better than the proposed approach. In E. coli, the proposed method achieved the highest number of true influential proteins in all levels, except for the top 100 and 500 levels. The results also shows that some topological centrality measures performed well in identifying critical proteins, but they were not consistent in different networks. For instance, DC and GD were the second-best measures in the Saccharomyces cerevisiae network for identifying critical proteins, but they were not as good in the Escherichia coli network. CC was the second-best measure in the E. coli network, but it was the worst measure among all topological measures in the Saccharomyces cerevisiae network. The reason is that different networks have different topologies, so a single topological measure could not capture the complex and multifaceted nature of the network influence.

Table 5. Number of critical proteins identified by WM-TOPSIS, TOPSIS, DC, BC, CC, EC, HM, PR, SC, and GD in Saccharomyces cerevisiae and E-Coli networks

Saccharomyces cerevisiae						
Measures	Top 100	Top 500	Top 600	Top 800	Top 1200	Top 1600
DC	46	201	251	330	481	601
BC	44	177	220	283	417	521
CC	29	175	217	284	412	533
EC	37	192	221	293	437	566
HM	41	188	228	294	423	543
PR	47	196	242	332	467	579
SC	37	192	221	293	437	566
GD	46	201	251	330	481	598
TOPSIS	43	197	239	312	461	587
WM-TOPSIS	**46**	**200**	**252**	**327**	**482**	**601**
Escherichia coli						
Measures	Top 100	Top 500	Top 600	Top 800	Top 1200	Top 1600
DC	27	106	121	141	170	198
BC	31	97	109	146	164	195
CC	27	105	123	120	175	203
EC	22	105	118	135	169	201
HM	28	105	121	139	175	203
PR	27	110	118	130	170	186
SC	22	105	118	135	169	201
GD	27	106	121	141	170	196
TOPSIS	25	102	114	136	158	192
WM-TOPSIS	**24**	**109**	**122**	**141**	**174**	**204**

The WM-TOPSIS also outperformed classical TOPSIS in Saccharomyces cerevisiae and E. coli networks, except for the top 100 in E. coli. The results confirmed that the WM-TOPSIS, which combined multiple centrality measures with a novel weighting mechanism, was more effective and reliable in identifying critical proteins in biological networks.

Validation by Statistical Parameters

To additionally assess the comprehensive effectiveness of the WM-TOPSIS method, various statistical metrics, such as sensitivity (S_N), specificity (S_P), negative predictive value (NPV), positive predictive value (PPV), F-measure, and Accuracy, are utilized. These values are typically calculated using a confusion matrix, as illustrated in Figure 3. The top 1600 proteins from all methods' rankings were designated as **"critical"** in both Yeast and E. coli. The rest were labelled **"noncritical"**.

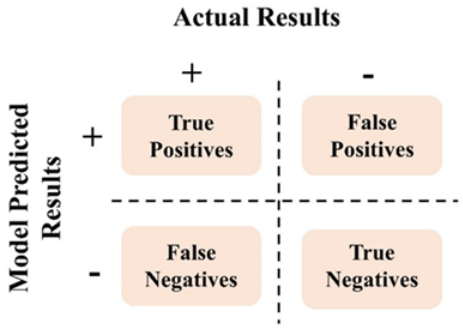

Figure 3. Confusion matrix Illustrating Actual vs. Predicted classifications

The results in Table 6 strongly support the performance of our proposed WM-TOPSIS method. For Saccharomyces cerevisiae data, WM-TOPSIS achieves higher values across all evaluation criteria: sensitivity (S_N) of 0.5150, specificity (S_P) of 0.7455, positive predictive value (PPV) of 0.3756, negative predictive value (NPV) of 0.8380, F-measure of 0.4344, and accuracy of 0.6927. Notably, proposed WM-TOPSIS outperforms all other prediction methods on every criterion except for DC, where it ties. Similarly, on Escherichia coli data, the values of S_N, S_P, PPV, NPV, F- Measure and Accuracy of WM-TOPSIS are 0.8031, 0.4355, 0.1275, 0.9556, 0.2201 and 0.4697 respectively, which also outperforms all other methods listed in Table 6.

Table 6. Comparative performance evaluation of WM-TOPSIS, TOPSIS, and centrality metrics using statistical parameters

Saccharomyces cerevisiae						
Measures	S_N	S_P	PPV	NPV	F-Measure	Accuracy
DC	0.5150	0.7455	0.3756	0.8380	0.4344	0.6927
BC	0.4464	0.7252	0.3256	0.8151	0.3766	0.6613
CC	0.4567	0.7282	0.3331	0.8185	0.3853	0.666
EC	0.4850	0.7366	0.3538	0.8279	0.4091	0.6790
HM	0.4653	0.7308	0.3394	0.8214	0.3925	0.6699
PR	0.4961	0.7399	0.3619	0.8317	0.4185	0.6841
SC	0.4850	0.7366	0.3538	0.8279	0.4091	0.6790
GD	0.5124	0.7448	0.3738	0.8371	0.4322	0.6915
TOPSIS	0.5030	0.7420	0.3669	0.8340	0.4243	0.6872
WM-TOPSIS	**0.5150**	**0.7455**	**0.3756**	**0.8380**	**0.4344**	**0.6927**
Escherichia coli						
Measures	S_N	S_P	PPV	NPV	F-Measure	Accuracy
DC	0.7795	0.4331	0.1238	0.9503	0.2136	0.4653
BC	0.7677	0.4319	0.1219	0.9476	0.2104	0.4631
CC	0.7992	0.4351	0.1269	0.9547	0.2190	0.4690
EC	0.7913	0.4343	0.1256	0.9530	0.2168	0.4675
HM	0.7992	0.4351	0.1269	0.9547	0.2190	0.4690
PR	0.7323	0.4282	0.1163	0.9397	0.2006	0.4565
SC	0.7913	0.4343	0.1256	0.9530	0.2168	0.4675
GD	0.7717	0.4323	0.1225	0.9485	0.2114	0.4639
TOPSIS	0.7559	0.4307	0.1200	0.9450	0.2071	0.4609
WM-TOPSIS	**0.8031**	**0.4355**	**0.1275**	**0.9556**	**0.2201**	**0.4697**

Furthermore, the accuracy of the WM-TOPSIS surpasses that of classical (original) TOPSIS in both organisms, as illustrated in Figure 4 (a and b). Based on these comprehensive analyses across six statistical measures, we conclude that proposed WM-TOPSIS emerges as a highly promising method for identifying critical proteins, demonstrating consistent and superior performance across diverse organisms.

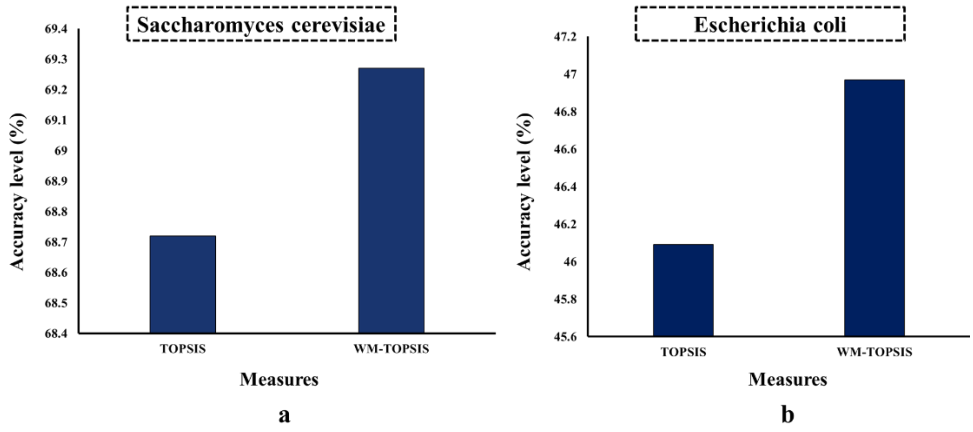

Figure 4. Performance comparison of proposed WM-TOPSIS in (a) yeast and (b) E. coli network

Analysis of Node Frequency with Identical Centrality Values Across Various Methods

In this experiment, the focus was on comparing the frequency of nodes with identical centrality values across different metrics. In a real-world network, nodes that share the same centrality value are essentially ranked equivalently, leading to challenges in distinguishing their relative importance. The ability to effectively distinguish the significance of each node is a key characteristic of an efficient method [41]. Thus, it is necessary to compare the frequency of nodes that have the same centrality value obtained by different methods to test the effectiveness of proposed approach.

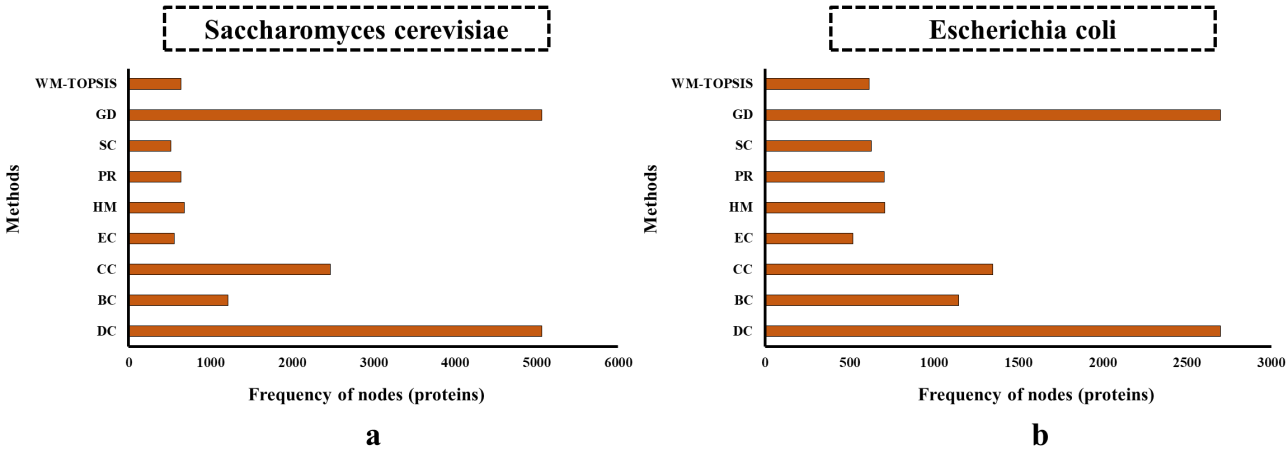

Figure 5. Frequency distribution of nodes with identical centrality scores across different methods in (a) yeast and (b) E. coli network

In the networks of Saccharomyces cerevisiae and Escherichia coli, the centrality value of most proteins obtained by proposed WM-TOPSIS are unique and only a few proteins have same centrality value. However, as depicted in Figure 5 (a and b), the outcomes of DC, BC, CC and GD reveal that the majority of nodes possess identical centrality values. This suggests that DC, BC, CC, and GD are unable to effectively differentiate the significance of each protein within the network. In conclusion, the results

demonstrate that the proposed WM-TOPSIS effectively discerns the importance of each protein, regardless of the network it operates on, thereby proving its efficacy.

Conclusions

This study introduces Weighted Modified-TOPSIS (WM-TOPSIS) for identifying vital proteins in Protein-Protein Interaction (PPI) networks by integrating eight topological centrality measures: DC, HC, PR, SC, BC, CC, EC, and GCCDC. Recognizing that varied network topologies challenge single-metric approaches, WM-TOPSIS employs a benchmark-based, data-driven weighting scheme to prioritize each measure, comprehensively assessing node importance. The accuracy of the WM-TOPSIS surpasses that of classical (original) TOPSIS in both organisms (Yeast and E. coli). Future research will explore WM-TOPSIS's applicability to larger dynamic networks by incorporating temporal data to capture evolving interactions. Additionally, integrating biological data could enhance precision, potentially extending WM-TOPSIS's utility to more complex network analyses.

Conflicts of Interest

The author(s) confirm(s) that there are no conflicts of interest associated with the publication of this article.

Acknowledgment

The authors acknowledge Salim Habib University and NED University of Engineering & Technology for providing computational resources and research facilities that supported the analysis of PPI networks in this study.

References

[1] Hamouda, E., ElHafsi, M., & Son, J. (2024). Securing network resilience: Leveraging node centrality for cyberattack mitigation and robustness enhancement. *Information Systems Frontiers.* Advance online publication. https://doi.org/10.1007/s10796-024-10477-y.

[2] Kallakunta, S., & Sreenivas, A. (2024). Optimizing wireless sensor networks by identifying key nodes using centrality measures. *Momona Ethiopian Journal of Science, 16*(2), 289–295. https://doi.org/10.4314/mejs.v16i2.7.

[3] Jain, A., & Reddy, B. V. R. (2013). Node centrality in wireless sensor networks: Importance, applications and advances. In *Proceedings of the 2013 3rd IEEE International Advance Computing Conference (IACC 2013)* (pp. 127–131). IEEE. https://doi.org/10.1109/IADCC.2013.6514207.

[4] Ur Rehman, S., & Ubaida Fatima, P. (2025). Digital currency network centrality measure (DCNC): A new centrality measure for cryptocurrency datasets. [Preprint].

[5] Gopalsamy, T., Thankappan, V., & Chandramohan, S. (2025). Investigating critical node identification in water networks through distance Laplacian energy centrality. *Environmental Science and Pollution Research.* Advance online publication. https://doi.org/10.1007/s11356-025-36118-8.

[6] Rout, T., Mohapatra, A., Kar, M., & Muduly, D. K. (2025). Centrality-based approach for identifying essential cancer proteins in PPI networks. *SN Computer Science, 6*(1). https://doi.org/10.1007/s42979-024-03480-2.

[7] Moiz, A., Fatima, U., & Zeeshan Ul Haque, M. (2024). A new framework for pinpointing crucial proteins in protein-protein interaction networks. *IEEE Access, 12,* 108425–108444. https://doi.org/10.1109/ACCESS.2024.3437215.

[8] Lawson, S., Donovan, D., & Lefevre, J. (2024). An application of node and edge nonlinear hypergraph centrality to a protein complex hypernetwork. *PLoS One, 19*(10), e0311433. https://doi.org/10.1371/journal.pone.0311433.

[9] Moiz, A., & Fatima, U. (2024). Key element identification in large biological datasets: An MCDM comparative study. In *2024 International Visualization, Informatics and Technology Conference (IVIT 2024)* (pp. 153–158). IEEE. https://doi.org/10.1109/IVIT62102.2024.10692773.

[10] Xu, W., Dong, Y., Guan, J., & Zhou, S. (2022). Identifying essential proteins from protein–protein interaction networks based on influence maximization. *BMC Bioinformatics, 23,* 466. https://doi.org/10.1186/s12859-022-04874-w.

[11] Sprinzak, E., Sattath, S., & Margalit, H. (2003). How reliable are experimental protein-protein interaction data? *Journal of Molecular Biology, 327*(5), 919–923. https://doi.org/10.1016/S0022-2836(03)00239-0.

[12] Wuchty, S., & Stadler, P. F. (2003). Centers of complex networks. *Journal of Theoretical Biology, 223*(1), 45–53. https://doi.org/10.1016/S0022-5193(03)00071-7.

[13] Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review E, 71*(5), 056103. https://doi.org/10.1103/PhysRevE.71.056103.

[14] Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks, 27*(1), 39–54. https://doi.org/10.1016/j.socnet.2004.11.009.

[15] Stephenson, K., & Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks, 11*(1), 1–37. https://doi.org/10.1016/0378-8733(89)90016-6.

[16] Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks, 29*(4), 555–564. https://doi.org/10.1016/j.socnet.2007.04.002.

[17] Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7.

[18] Li, M., Wang, J., Chen, X., Wang, H., & Pan, Y. (2011). A local average connectivity-based method for identifying essential proteins from the network level. *Computational Biology and Chemistry, 35*(3), 143–150. https://doi.org/10.1016/j.compbiolchem.2011.04.002.

[19] Fatima, U., Hina, S., & Wasif, M. (2023). A novel global clustering coefficient-dependent degree centrality (GCCDC) metric for large network analysis using real-world datasets. *Journal of Computational Science, 70,* 102008. https://doi.org/10.1016/j.jocs.2023.102008.

[20] Zhao, B., Wang, J., Li, X., & Wu, F. X. (2016). Essential protein discovery based on a combination of modularity and conservatism. *Methods, 110,* 54–63. https://doi.org/10.1016/j.ymeth.2016.07.005.

[21] Li, M., Li, W., Wu, F. X., Pan, Y., & Wang, J. (2018). Identifying essential proteins based on sub-network partition and prioritization by integrating subcellular localization information. *Journal of Theoretical Biology, 447,* 65–73. https://doi.org/10.1016/j.jtbi.2018.03.029.

[22] Lei, X., Zhao, J., Fujita, H., & Zhang, A. (2018). Predicting essential proteins based on RNA-Seq, subcellular localization and GO annotation datasets. *Knowledge-Based Systems, 151,* 136–148. https://doi.org/10.1016/j.knosys.2018.03.027.

[23] Li, M., Niu, Z., Chen, X., Zhong, P., Wu, F., & Pan, Y. (2016). A reliable neighbor-based method for identifying essential proteins by integrating gene expressions, orthology, and subcellular localization information. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 13*(5), 845–855. https://doi.org/10.1109/TCBB.2016.2530067.

[24] Luo, J., & Qi, Y. (2015). Identification of essential proteins based on a new combination of local interaction density and protein complexes. *PLoS One, 10*(6), e0131418. https://doi.org/10.1371/journal.pone.0131418.

[25] Li, M., Zhang, H., Wang, J. X., & Pan, Y. (2012). A new essential protein discovery method based on the integration of protein-protein interaction and gene expression data. *BMC Systems Biology, 6,* 15. https://doi.org/10.1186/1752-0509-6-15.

[26] Zhang, X., Xu, J., & Xiao, W. X. (2013). A new method for the discovery of essential proteins. *PLoS One, 8*(3), e58763. https://doi.org/10.1371/journal.pone.0058763.

[27] Peng, W., Wang, J., Wang, W., Liu, Q., Wu, F. X., & Pan, Y. (2012). Iteration method for predicting essential proteins based on orthology and protein-protein interaction networks. *BMC Systems Biology, 6,* 87. https://doi.org/10.1186/1752-0509-6-87.

[28] Tzeng, G. H., & Huang, J. J. (2011). *Multiple attribute decision making: Methods and applications.* CRC Press.

[29] Du, Y., Gao, C., Hu, Y., Mahadevan, S., & Deng, Y. (2014). A new method of identifying influential nodes in complex networks based on TOPSIS. *Physica A: Statistical Mechanics and Its Applications, 399,* 57–69. https://doi.org/10.1016/j.physa.2013.12.031

[30] Hu, J., Du, Y., Mo, H., Wei, D., & Deng, Y. (2016). A modified weighted TOPSIS to identify influential nodes in complex networks. *Physica A: Statistical Mechanics and Its Applications, 444,* 73–85. https://doi.org/10.1016/j.physa.2015.09.028.

[31] Fei, L., & Deng, Y. (2017). A new method to identify influential nodes based on relative entropy. *Chaos, Solitons & Fractals, 104,* 257–267. https://doi.org/10.1016/j.chaos.2017.08.010.

[32] Yang, Y., Yu, L., Zhou, Z., Chen, Y., & Kou, T. (2019). Node importance ranking in complex networks based on multicriteria decision making. *Mathematical Problems in Engineering, 2019,* 9728742. https://doi.org/10.1155/2019/9728742.

[33] Ishfaq, U., Khan, H. U., & Iqbal, S. (2022). Identifying the influential nodes in complex social networks using centrality-based approach. *Journal of King Saud University - Computer and Information Sciences, 34*(10), 9376–9392. https://doi.org/10.1016/j.jksuci.2022.09.016.

[34] Yang, P., Liu, X., & Xu, G. (2018). A dynamic weighted TOPSIS method for identifying influential nodes in complex networks. *Modern Physics Letters B, 32*(19), 1850216. https://doi.org/10.1142/S0217984918502160.

[35] Yang, P., Xu, G., & Chen, H. (2018). Multi-attribute ranking method for identifying key nodes in complex networks based on GRA. *International Journal of Modern Physics B, 32*(32), 1850363. https://doi.org/10.1142/S0217979218503630.

[36] Yang, Y., Yu, L., Wang, X., Chen, S., Chen, Y., & Zhou, Y. (2020). A novel method to identify influential nodes in complex networks. *International Journal of Modern Physics C, 31*(2), 2050029. https://doi.org/10.1142/S0129183120500229.

[37] Mari, J., Ortega, E., Eballe, R. G., & High, P. S. (2022). Harmonic centrality in some graph families. *Mathematics, 10*(21), 4128. https://doi.org/10.3390/math10214128.

[38] Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. In *Proceedings of the 7th International World Wide Web Conference (WWW 1999).*

[39] Database of Interacting Proteins (DIP). (2023, July 2). DIP::Stat: Database. Retrieved from https://dip.doe-mbi.ucla.edu/dip/Stat.cgi?SM=0.

[40] Zhong, J., Wang, J., Zhang, Y., Pan, Y., Li, M., & Wu, F. X. (2021). A novel essential protein identification method based on PPI networks and gene expression data. *BMC Bioinformatics, 22,* 580. https://doi.org/10.1186/s12859-021-04175-8.

[41] Zhong, S., Zhang, H., & Deng, Y. (2022). Identification of influential nodes in complex networks: A local degree dimension approach. *Information Sciences, 610,* 994–1009. https://doi.org/10.1016/j.ins.2022.07.172.