Duzce University Journal of Science and Technology
Düzce Üniversitesi Bilim ve Teknoloji Dergisi
14 (1), 2026, 225-239.
https://doi.org/10.29130/dubited.1795730
e-ISSN: 2148-2446
https://dergipark.org.tr/en/pub/dubited

Research Article

Semantic and Structural Perspectives on Node Importance: A Combined Classical–LLM Approach

Levent Sabah[1*], Zeynep Bozdoğan Sabah[1], Resul Kara[2]

[1] Düzce University Department of IT, Düzce, Türkiye. ror.org/04175wc52

[2] Düzce University Department of Computer Engineering, Düzce, Türkiye. ror.org/04175wc52

*Corresponding author:
leventsabah@duzce.edu.tr

Abstract: Research on node importance in complex networks continues to draw on a wide range of approaches, including classical centrality metrics, entropy-based measures, and various machine-learning models. Previous studies often reports that entropy-driven and learning-based techniques yield higher accuracy and faster convergence for tasks such as information spreading or robustness analysis, although these results can vary depending on the dataset and the assumptions of each method. What remains less explored is the extent to which these techniques account for the contextual or semantic characteristics of nodes, which are not directly visible in structural data. In this study, a hybrid framework is introduced to address this gap by combining established centrality metrics with evaluations produced by a large language model (LLM). The approach was examined on several well-known benchmark networks—Zachary's Karate Club, the Krackhardt Kite Graph, and the Jazz Musicians Network—and the resulting rankings showed clearer separation among nodes and, in many cases, more interpretable patterns than the purely structural baselines. To situate these findings within the broader literature, results were compared with three prior studies that employ composite or entropy-based centrality formulations. To explore scalability, three larger collaboration networks from the SNAP repository were also evaluated. The overall evidence suggests that LLM-supported scoring may offer a complementary perspective for complex network analysis, especially when structural information alone provides an incomplete picture.

Keywords: *Complex network, Centrality measure, Large language model, Node importance*

Received: 02.10.2025
Revision: 10.12.2025
Accepted: 15.12.2025
Available Online: 21.01.2026

I. INTRODUCTION

In the study of complex networks, understanding how nodes contribute to the overall structure and dynamics of the system is a central concern. This importance is typically assessed through a range of classical centrality measures—such as degree, closeness, betweenness, and eigenvector centrality—which capture different aspects of a node's position. Although these measures provide clear numerical summaries of structural influence, they often fall short when the goal is to reflect the contextual or role-oriented significance that a node may hold within a network. Large language models (LLMs) offer a complementary perspective. When supplied with structured descriptions of node neighborhoods, an LLM can produce semantic interpretations that go beyond purely topological information. Rather than replacing classical centralities, these model-driven insights can extend them and help distinguish nodes that appear similar under classical measures. In this way, LLM-based assessments enhance the interpretability and separability of nodes across the network.

The concept of centrality was first formalized by Freeman (Freeman, 1978), who introduced degree, closeness, and betweenness centrality. measures such as Katz (Katz, 1953), and later PageRank (Page et al., 1999) and HITS (Kleinberg, 1999), are widely used due to their simplicity and interpretability. More recent studies have examined composite centralities by integrating multiple measures using weighting schemes, entropy methods, or multi-criteria decision-making (e.g., TOPSIS, AHP). Such approaches aim to overcome the biases of single measure. The integration of machine learning into network science is still evolving. LLM, in particular, offers interpretative advantages by generating role-based insights. In recent years, significant progress has been made in the field of node detection and classification using large language models (LLM) and graph neural networks (GNN). In particular, the Zachary's Karate Club dataset is widely preferred in community segregation studies with LLM and GNN-based node detection. However, studies on LLM-based node detection in the Krackhardt Kite Graph and Jazz Musicians Network datasets are still limited (Chaudhary & Singh, 2023; Huang et al., 2022; Li et al., 2024; Lyu et al., 2023; Wang et al., 2025; Zhang & Xiao, 2022). This study addresses this gap by introducing a hybrid framework that combines classical centralities with LLM-based reasoning.

Many methods for detecting important nodes have been studied in the literature. They aims to increase node distinguishability and model information propagation more realistically. (Salavati et al., 2019) proposed the GLR

This is an open access paper distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License.

(Gateway Local Rank) method based on community structures and emphasized the importance of bridges between communities by identifying critical nodes in each community. Their method provides a robust solution especially in cases where information flow is concentrated at community boundaries but not provide a semantic explanation of the functional roles of nodes. (Guo et al., 2020), developed the EnRenew algorithm based on information entropy and showed that nodes with high entropy are more effective in information propagation and so demonstrating the quantitative distinguishing power of entropy-based methods. Entropy-based approaches remain relevant. (Yang et al., 2020) in their study the DCC (Degree–Clustering Coefficient) approach provides a more generic centrality measure by combining the degree, neighborhood structure, and clustering coefficient of nodes with entropy weights. Their study produces more realistic results by accounting for the influence of second-level neighbors. Similarly, (Xu et al., 2020) introduced the concept of 'adjacency information entropy' to measure structured information in networks and demonstrated its success in different network types (directional/weighted). (Şimşek, 2024) formulated the identification of influential nodes as a binary classification problem. In this framework, basic centrality measures were employed as node-level features, while the effectiveness of each node was quantified using the Independent Cascade (IC) propagation model. Nodes corresponding to the top 1% and top 5% influence levels were subsequently labeled as influential and used to train the classifier. The results show that the Decision Tree method performs better than individual centrality measures. Beyond this performance gain, the method contributes by adopting a more realistic propagation model and by simplifying the problem through narrowing the solution space for optimization algorithms.

On the other hand Our LLM-based approach, offers a complementary and progressive perspective to these methods. This can consider not only the topological location of nodes but also their contextual and semantic relationships. In addition the classical centrality and machine learning-based classification approaches in the literature, this proposed method offers a new paradigm for effective individual detection. This study aims to compare the ability of LLMs to distinguish nodes with three reference studies in the literature that derive joint centrality measures from classical measures.

A. Related Studies

Centrality measures are commonly used tools for determining node importance of complex network analysis. However, studies that analyses based on a single centrality measure are often insufficient also different measures can produce same results on the same network. Therefore, in recent years, studies on developing composite methods that combine different centrality measures have increasing. (Keng et al., 2021) in their study analyzing node rankings by combining degree, closeness, and betweenness centralities with different rates. The concept of regular points allows nodes to classifing into the largest possible number of classes and the concept of critical points indicates the coefficient values for node rankings changes. Results on such as Krackhardt Kite Graph and Zachary's Karate Club dataset have reported that convex combinations provide greater distinguishing power than single measures. In addition to (Keng et al., 2021) convex combination approach, (Mohammad et al., 2025) recently proposed a hybrid structure that combines classical centralities (DC, BC, EC) with 'isolating centrality' (ISC) in the network and uses a tuning parameter. This study shows that convex combinations can be used not only for two measures but also for more complex structures that combine local and global information.

(Fei & Deng, 2017) in their study combines information obtained from different centrality measures with relative entropy and TOPSIS methods to evaluate node importance within the framework of multi-criteria decision-making. Experiments on networks such as Zachary's Karate Club, Email, UsAir97, and David Copperfield prove that this method produces more consistent rankings compared to classical centrality measures. Also it was reported that reducing the number of nodes with the same centrality value increases distinctiveness and provides more reliable results, especially in the top-10 node rankings. (Sabah & Şimşek, 2023) in their study, used the Analytical Hierarchy Process (AHP) and entropy-based weighting to combine centrality measures. In their study, degree, closeness, betweenness, eigenvector, PageRank, and HITS measures were considered; coefficients were determined based on entropy values, independent of subjective expert judgments. Analyses of the Krackhardt Kite Graph, Zachary's Karate Club, Jazz Musicians Network and five collaborative networks in the SNAP database, revealed that the proposed method both improves distinguishing power and is significantly more efficient in terms of computational time than competing methods. In particular, increasing the node level from 5-6 to 7 in small networks and analyzing thousands of nodes in large networks in a short time demonstrate the method's scalability. In addition to multi-criteria decision-making methods, physics-based fusion models are also attracting attention in recent literature. (Li et al., 2025) developed a gravity model called DKBC that considers the degree of the nodes, the K-shell value, and the betweenness as 'mass' and takes into account the distance between nodes. This model aims to more accurately predict the spreading capacity by integrating spatial information into the betweenness measurements. A summary of related work is presented in Table 1.

Table 1. Summary of Related Studies.

Study	Method Used	Data Sets	Main Contribution	Results Obtained
(Keng et al., 2021)	Convex combinations of centrality measures	Krackhardt Kite Graph, Zachary's Karate Club	Distinguishing power with regular and critical point concepts	Convex combinations produced more node classes and provided flexibility in rankings.
(Fei & Deng, 2017)	Relative entropy + TOPSIS approach	Karate Club, Email, UsAir97, David Copperfield	Multi-criteria decision-making framework; proximity to the ideal solution with relative entropy	Consistent rankings were achieved, and the number of nodes with the same centrality value was reduced.
(Sabah & Şimşek, 2023)	Composite centrality with AHP + Entropy weighting	Krackhardt Kite Graph, Zachary's Karate Club, Jazz Musicians Network, SNAP collaboration networks	Objective weight determination using entropy, joint with AHP.	It provided high distinguishing power and worked efficiently in large networks.
(Xu et al., 2020)	Adjacency information entropy	Astro, CA, Facebook, Hamster, Email, PGP, Router, Wiki-Vote, P2P, PHD	Two algorithm proposals based on neighborhood degree and choice probability for four different network types (weighted/unweighted, directed/undirected).	It showed a high correlation with degree centrality and performed well in disconnecting the network (independent parts test).
(Mohammad et al., 2025)	Convex Isolating classical Centrality	Ca-netscience, web-polblogs, bio-dmela	Six new measurements (CIBC, CIDC, etc.) combine isolated centrality (ISC) with classical measurements and balance local/global effects using an alpha tuning parameter.	The SIR model achieved higher propagation efficiency compared to classical and current measurements, enabling effective node detection with reasonable time complexity.
(Li et al., 2025)	Degree-k-shell-Betweenness Centrality	Jazz, USAir, NS, Email, PB, Facebook, Power, GrQc, EEC, PG, Router, WV (12 real networks)	A multi-feature fusion combining nodal degree, k-shell position, and interstitial properties based on the Gravity Model.	The highest spreading effect was observed in SIR and IC models; superior accuracy and discrimination (monotonicity) were achieved compared to benchmark algorithms.

II. MATERIAL AND METHOD

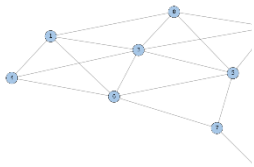
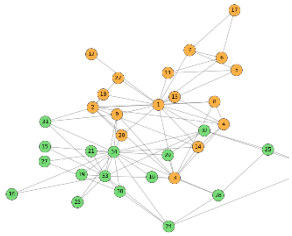
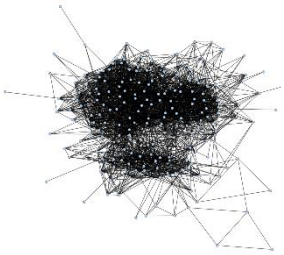
A. Materials

This section provides a detailed overview of the datasets and methods used in our study. First, we introduce three classic datasets in social network analysis: Zachary's Karate Club, Krackhardt Kite Graph, and the Jazz Musicians Network. Their structural properties are summarized. These datasets provide representative network examples of varying scales and characteristics, frequently used for node detection and community segmentation. Additionally, three collaborative networks in the SNAP repository (Leskovec & Krevl, 2014) were used for scalability. Next, we explain the methods used in this study, and discuss the role of LLM-based approaches, which have gained importance in recent years, in node classification and role identification.

A. 1. Datasets

In this study, we chose small and medium-sized complex networks, commonly used in the literature. Classical network datasets (e.g., Zachary's Karate Club, Krackhardt Kite Graph, and the Jazz Musicians Network) are crucial for testing the distinguishability of complex networks. In these networks, automorphism orbits (i.e., structurally identical nodes) are often indistinguishable using classical centrality measures. For example, in datasets like Zachary's Karate Club and the Jazz Musicians Network, strong linear relationships were observed between degree and closeness centrality, leading to inadequate node separation. Especially in small and symmetric structures, classical measures may not be able to meaningfully distinguish between nodes. Details of the datasets used in this study are provided in Table 2.

Table 2. Details and overview of the datasets used in the study.

Dataset	Krackhardt Kite Graph (Krackhardt, 1990)	Zachary's Karate Club (Zachary, 1977)	Jazz Musicians Network (Gleiser & Danon, 2003)
Number of Nodes	10	34	198
Number of Edges	18	78	2742
Features	It is a small-scale network where the differences between centrality measures in complex network analyses are most clearly observed.	It is used as a standard reference set in centrality studies on complex network analyses. The classic social network represents the ties among members of a university club observed in the 1970s.	It offers an example of a larger collaborative network: a medium-sized network representing collaborations among jazz musicians.
Purpose of Usage	Visualize and test how different centrality measures rank nodes differently.	Compare centralization methods and examine community fragmentation.	Measure the performance of centralization combinations in larger networks.
General View			

This study also utilized ca-GrQc, CA-HepTh, and CA-CondMat collaboration networks (Leskovec & Krevl, 2014) from different physics domains on arXiv. In these networks, nodes represent researchers, and edges represent co-authorship relationships. Containing 5,242, 9,877, and 23,133 nodes respectively, these networks are widely used real-world collaboration structures in the literature. Thanks to their different sizes and ensemble characteristics, they provide a suitable basis for evaluating the scalability and performance of our LLM-based node discrimination approach.

B. Methods

In this section firstly we will provide information about the basic centrality measures used in complex network analysis, then on the joint centrality methods used in the three studies compared in our study. Lastly, we give details on the LLM methods used in our study to distinguish nodes in complex networks.

B. 1. Classical Centrality Methods

Classical centrality methods are frequently used to measure the importance of nodes in complex network analyses. Degree centrality determines a node's local influence based on its number of connections, while closeness centrality measures the speed of information dissemination by considering a node's average distance from other nodes. Betweenness centrality identifies which nodes act as intermediaries or bridges along the shortest paths in the network. Eigenvector centrality considers not only the number of connections but also the strength and importance of these connections. Spectral-based methods such as PageRank and HITS (hub and authority) are also important measures that assess node importance from different perspectives. However, because each centrality metric reflects only a specific aspect of the network, when used alone, it can produce different and sometimes contradictory results. For this reason, in recent years, it has become common to evaluate more than one centrality measure together or to develop composite centrality methods derived from them.

B. 2. Joint Centrality Methods

Various approaches have been proposed in the literature for the joint use of centrality measures.

• Convex combination method (Keng et al., 2021): New composite centralities were obtained by weighting two or more measures with different coefficients.

• Relative Entropy + TOPSIS method (Fei & Deng, 2017): Using a multi-criteria decision-making approach, centrality values were re-ranked based on distance from ideal solutions.

• AHP + Entropy weighting (Sabah & Şimşek, 2023): The information contribution of each centrality measure was objectively calculated using entropy values and joint with the Analytical Hierarchy Process to produce composite scores.

These methods were applied to test distinguishability small networks (e.g., separating nodes from 5–6 levels to 7

levels), and to test computational speed and scalability in large networks.

B. 3. Large Language Models (LLMs) and DeepSeek

Large Language Models (LLMs), which have come to the fore in natural language processing in recent years, have the capacity to interpret, summarize, classify, and generate new content from text data using deep learning architectures with billions of parameters. Thanks to their multi-head attention mechanism, these transformer-based models can be adapted not only for language processing but also for the analysis of complex data structures. In the context of network science, LLMs can extract semantic representations of nodes and relationships, enabling the development of new ranking and parsing approaches beyond classical centrality measures. The DeepSeek (DeepSeek-AI et al., 2024) model used in this study is an open-source LLM developed in China with billions of parameters. DeepSeek is optimized to deliver a balanced performance between speed and accuracy, particularly in data-intensive tasks. The model was developed on multilingual training data and can be used in both natural language generation and numerical data interpretation tasks. In our study, DeepSeek was applied for node distinguishing and score generation, resulting in new rankings comparable to classical methods.

B. 3. 1. LLM-Based Approach to Distinguishing Nodes in Complex Networks

In this study, complex network analysis was conducted using Large Language Models (LLMs) with the goal of incorporating node importance assessments that extend beyond purely structural measures. Each node's basic centrality values—such as degree, closeness, betweenness, and eigenvector centrality—were first computed and then provided to the LLM in a structured JSON format. The model was asked to interpret the network and articulate the importance of individual nodes while assigning a numerical importance score between 0 and 1 based on its natural language reasoning capabilities. These LLM-generated scores were subsequently compared to classical centrality-based rankings, and the degree of agreement between the two approaches was evaluated using Kendall's Tau, a non-parametric measure that quantifies rank correlation.

Algorithm 1 presents the pseudocode for distinguishing nodes in a complex network based on centrality metric values using LLM

Algorithm 1
1. Input: centrality_matrix[V][measures], LLM model
2. Output: LLM-based node importance scores

3. For each node v in V:
4. Create prompt P with node id and centrality values
5. Send P to LLM (API call)
6. Parse JSON response: score[v] ∈ [0,1]
7. If JSON parsing fails:
8. Apply safe_json_parse(raw_output)

9. Return llm_scores[V]

B. 3. 2. Methodological Framework for LLM-Based Node Scoring

In this study, the process of distinguishing nodes and generating scores with an LLM-based approach was carried out in two main steps

B. 3. 2. 1. Model Calling Function

The LLM function, designed in Python for interaction with the LLM, directly passes the given prompt text to the LLM model and returns the clean response from the LLM model.

- Model used: *DeepSeek-Chat*
- Diversity parameter: *temperature* = 0.0 (to obtain deterministic output)
- Maximum output length: 2800 tokens

This structure ensured consistent output for the same input data in all experiments.

B. 3. 2. 2. Node-Oriented Prompt Structure

For each network, the prompt provided to the language model was produced automatically from the corresponding node list. The model was asked to return a valid JSON string that assigns a numerical score to every node—ranging from 0.000 to 1.000 with at least three decimal places—and to include a short, informative explanation ("why") for each value. Nothing beyond this structured output was required. Bringing these steps together allowed the LLM-based scoring process to function as a repeatable pipeline: a standardized request, a consistent model response, and a well-formed JSON output. This setup made it possible to interpret the model's semantic judgments separately from the numerical influence values produced by classical centrality measures, giving an additional layer of insight into node roles within

the network. Algorithm 2 outlines this procedure in pseudo-code form and summarizes how the LLM-supported node distinction process operates when joint with classical centrality information.

Algorithm 2
Input: composite_scores[V], llm_scores[V]
Output: Evaluation metrics (Tau correlation, distinguishability)

Step 1: Rank nodes by composite_scores
Step 2: Rank nodes by llm_scores
Step 3: Compute KendallTau(rank_composite, rank_llm)
Step 4: Count unique score levels for both methods
Step 5: Report:
- Tau correlation
- Number of unique levels (distinguishability)

B. 4. Using Joint Centrality vs. LLM

Joint centrality methods, while mathematically repeatable in network analysis, can objectively measure the structural importance of nodes, although their computational cost can be high in large networks. These methods rely solely on the topological properties of the network; that is, they are limited in capturing the contextual, semantic, or contextual roles of nodes. Furthermore, when multiple centrality measures are used together, the selection of appropriate coefficients requires an additional design decision that complicates the applicability of the method.

LLM-based approaches, on the other hand, have the advantage of being able to evaluate nodes in the network not only according to their structural positions but also with contextual and textual descriptions. Thanks to human-like interpretation capabilities, they can define the role of each node with richer expressions. However, these methods have significant computational costs and may not be practical when applied to large networks. Also, the possibility of interpretations changing in different runs on the same network is a significant factor limiting the reproducibility of the method. This approach sometimes deviates from the objective framework provided by mathematical centrality measures, offering a more interpretive evaluation.

B. 5. The Importance of Small-Scale Datasets

Small-scale datasets such as the Krackhardt Kite Graph and Zachary's Karate Club occupy a foundational place in both classical centrality research and LLM-based approaches. Their importance stems various reasons. First, these small networks provide an ideal ground for validating methods and comparing different techniques. Researchers can quickly test the performance of a new method and observe its consistency with existing approaches. The structurally simple and fully traceable nature of these datasets allows for the tracking of relationships between nodes at both visual and analytical levels. This allows for a clear assessment of the validity of the developed methods. Finally, small-scale networks also provide a valuable testing environment for measuring the accuracy of LLM-based annotation generation. Researchers can quickly check the node annotations generated by the LLM and easily identify errors or inconsistencies. Because of these characteristics, small-scale datasets remain indispensable reference points for both classical network science and LLM-supported analyses.

III. RESULTS AND DISCUSSIONS

A. Effective Node Determination with Centrality Measures

Centrality measures such as degree, betweenness, closeness, eigenvector, and PageRank are classic and fast methods for determining the importance of nodes in a network (Ait Rai et al., 2023; D. Chen et al., 2012; Yang et al., 2020). However, these methods typically focus only on local or global information and may not fully reflect interactions across the entire network (D. Chen et al., 2012; Qiu et al., 2021; J. Zhao et al., 2020).

B. Entropy-Based and Hybrid Approaches

Entropy-based methods provide a more dynamic assessment by measuring the uncertainty and potential of nodes in information propagation. Methods such as the EnRenew algorithm have achieved improvements in propagation performance of up to 30% over classic centrality methods by dynamically updating node selection with information entropy (Dong & Zhou, 2021; Guo et al., 2020; Liu et al., 2024; Wang et al., 2022). Additionally, hybrid approaches such as entropy and weighted degree improve accuracy by considering both local and global effects (Dong & Zhou, 2021; Liu et al., 2024; Wang et al., 2022; Yang et al., 2020).

C. Machine Learning and Deep Learning Models

Machine learning methods such as decision trees, regression, graph convolutional networks (GCNs), and Transformer-based models have surpassed classical methods in efficient node detection by combining node features with network structure. These models have demonstrated higher accuracy and propagation success compared to classical centrality measures when tested with SIR simulations (Chen et al., 2024; Wu et al., 2024; Zhao, Jia, Huang, et al., 2020; Zhao, Jia, Zhou, et al., 2020).

D. Benchmarks and Datasets

Entropy-based and machine learning-based methods have particularly demonstrated higher propagation efficiency and better node ranking compared to classical centrality measures (Dong & Zhou, 2021; Guo et al., 2020; G. Zhao, Jia, Huang, et al., 2020; G. Zhao, Jia, Zhou, et al., 2020). Additionally, some hybrid and global-structural models have consistently demonstrated superiority across different network types (e.g., social, biological, and communication) (Mukhtar et al., 2023; Ullah et al., 2021; Yang et al., 2020; J. Zhao et al., 2020).

In the analyses for distinguishing nodes conducted on the Krackhardt Kite Graph dataset by (Fei & Deng, 2017), (Keng et al., 2021), and (Sabah & Şimşek, 2023), as well as in our study, a maximum of 7 levels of distinguishability were achieved. In all studies, it was observed that node pairs (0,1), (2,4), and (5,6) could not be distinguished. These node pairs are structurally equivalent (with the same automorphism orbit). Therefore, any centrality/role metric based solely on structure cannot distinguish them. Based on the results of the analysis performed on DeepSeek using the LLM used in our study, semantic expressions of possible causes were also obtained. LLM cannot break this equivalence; if it does, it is hallucinating.

When the joint centrality formulations used in the three reference studies are compared with the LLM-based DeepSeek evaluation applied here, nodes 5 and 6 consistently appear among the highest-ranked nodes in both approaches. Their positions indicate that the symmetric structure of the Krackhardt Kite Graph is preserved not only by classical joint centrality measures but also by the LLM-driven assessment. Likewise, node 9, which sits at the tail of the graph, is placed at the lowest levels in both analyses. Table 3 provides a side-by-side comparison of the node distinguishability levels obtained from the Krackhardt Kite Graph in prior studies and in the current work.

Table 3. Comparison of node distinguishability levels from related studies and Krackhardt Kite Graph dataset in this study.

	ORIGINAL			**CC-DC**		**BC-DC**				**BC-CC**					**LLM**
Level	**BC**	**CC**	**DC**	**Ct_μ** (Keng et al., 2021)	**Ct_ρ** (Keng et al., 2021)	**FMD** (Fei & Deng, 2017)	(Sabah & Şimşek, 2023)	**Ct_μ** (Keng et al., 2021)	**FMD** (Fei & Deng, 2017)	(Sabah & Şimşek, 2023)	**Ct_μ** (Keng et al., 2021)	**Ct_ρ** (Keng et al., 2021)	**FMD** (Fei & Deng, 2017)	(Sabah & Şimşek, 2023)	**DeepSeek**
1	7	5,6	3	3	3	3,5,6	3	5,6	5,6	7	7	7	5,6,7	7	3
2	5,6	3,7	5,6	5,6	5,6	0,1,7	5,6	3	3,7	5,6	5,6	5,6	3	5,6	5,6
3	8	0,1	0,1	0,1,7	7	2,4	0,1	7	0,1,8	8	3	3	0,1,8	8	0,1
4	3	2,4	2,4,7	2,4	0,1	8	7	0,1	2,4	3	0,1	0,1	2,4	3	2,4
5	0,1	8	8	8	2,4	9	2,4	8	9	0,1	2,4,8	8	9	0,1	7
6	2,4,9	9	9	9	8	-	8	2,4	-	2,4	9	2,4	-	2,4	8
7	-	-	-	-	9	-	9	9	-	9	-	9	-	9	9

The LLM approach, when enriched with additional contextual or semantic information, has the potential to capture details invisible to classical methods. However, in its current form, it falls short of making reliable distinctions in symmetric structures like the Krackhardt Kite Graph.

A detailed illustration of the resulting score and the reason for the node distinguishability levels of the Krackhardt Kite Graph obtained with the LLM-based approach is provided in Table 4. By obtaining information on the number of points assigned to each node in the 0.000-1.000 range and the reason for doing so, more comprehensive conclusions about the nodes in the complex network can be drawn.

Table 4. A detailed illustration of the resulting score and the reason for the node distinguishability levels of the Krackhardt Kite Graph obtained with the LLM-based approach.

Node	Score	Why	Rank
3	0,842	Highest degree, closeness, and eigenvector, most central node of the network	1
5	0,785	Highest degree, closeness, eigenvector, and highest betweenness, bridge role	2
6	0,785	Same characteristics as Node 5: strong bridge and central position	3
0	0,723	High degree, closeness, and eigenvector centrality, strong network of connections, and influential position	4
1	0,723	Same characteristics as Node 0: strong connections and central position	5
2	0,456	Moderate degree of connection, high eigenvector, but low betweenness	6
4	0,456	Same characteristics as Node 2: moderate centrality	7
7	0,412	Moderate degree of connection, high closeness and betweenness, bridge role	8
8	0,234	Low degree, moderate betweenness, limited centrality	9
9	0,085	Lowest centrality values, edge node of the network	10

As shown in Table 5, the methods reported in the literature generally distinguish 26 of the 34 nodes, leaving a small group at equivalent levels. By contrast, the LLM–DeepSeek approach separates all 34 nodes. The highest-ranking nodes (34, 1, 33, 3, and 32) match the literature closely, which supports the reliability of the LLM-based evaluation. A similar level of agreement appears in the middle ranks as well. The lower part of the ranking, however, presents a more nuanced picture: nodes that remain indistinguishable in previous studies are separated by the LLM with very small numerical differences. While this helps reveal fine-grained distinctions, differences on the order of 0.001 also raise the possibility of over-separation or mild hallucinatory effects. Taken together, the findings highlight both the distinguishability strength of the method and the points where its behavior calls for careful interpretation.

Table 5. Comparison of node distinguishability levels obtained from Zachary's Karate Club dataset in this study with related studies.

Level	**BC**	**CC**	**DC**	**Ct_μ** (Keng et al., 2021)	**FMD** (Fei & Deng, 2017)	**BC-CC-DC** (Sabah & Şimşek, 2023)	**BC-CC-DC-EC** (Sabah & Şimşek, 2023)	**LLM - DeepSeek**
1	1	1	34	1	1	1	1	34
2	34	3	1	34	34	34	34	1
3	33	34	33	33	3	33	33	33
4	3	32	3	3	33	3	3	3
5	32	9,14,33	3	32	32	32	32	32
6	9	20	4,32	2	2,9	2	2	2
7	2	2	9,14,24	9	14	9	9	9
8	14	4	6,7,8,28,30,31	14	20	14	14	14
9	20	28,31	5,11,20,25,26,29	4	28	20	20	4
10	6,7	29	10,13, 15,16, 17,18, 19,21, 22,23, 27	20	4,24	6,7	6,7	24
11	28	8	12	28	31	28	28	31
12	24	10	-	24	6,7	24	24	8
13	31	24	-	6,7	30	4	4	28
14	4	6,7,30	-	31	26	31	31	20
15	26	5,11	-	8	25	30	30	30
16	30	18,22, 25,26	-	29	29	26	26	29
17	*25*	*13,15,16, 19,32,33*	-	*30*	*5,8,10,11*	*29*	*29*	6
18	*29*	*12*	-	*26*	*13,15,16,17,1 8,19, 21,22,23,27*	*25*	*8*	7

Table 5. (cont.) Comparison of node distinguishability levels obtained from Zachary's Karate Club dataset in this study with related studies.

19	*10*	*27*	-	*25*	*12*	*8*	*25*	26
20	*5,11*	*17*	-	*10*		*5,11*	*5,11*	10
21	*8,12,13,15,16,17,18,19,21,22,23,27*	-	-	*5,11*	-	*10*	*10*	25
22	-	-	-	*18,22*	-	*18,22,23,21,16,15,13,19,27*	*18,22*	16
23	-	-	-	*13,15,16,19,21,23*	-	*17*	*23,21,16,15,13,19*	23
24	-	-	-	*27*	-	*12*	*27*	21
25	-	-	-	*12*	-	-	*17*	15
26	-	-	-	*17*	-	-	*12*	19
27	-	-	-	-	-	-	-	11
28	-	-	-	-	-	-	-	22
29	-	-	-	-	-	-	-	5
30	-	-	-	-	-	-	-	13
31	-	-	-	-	-	-	-	18
32	-	-	-	-	-	-	-	27
33	-	-	-	-	-	-	-	17
34	-	-	-	-	-	-	-	12

When the results in Table 5 are examined, the top five nodes (34, 1, 33, 3, and 32) form the backbone of the network and have the highest connectivity. Both classical measures and the LLM-based method rank these nodes prominently, which increases confidence in the stability of the upper part of the ranking. The middle-ranked nodes (9, 14, 4, 24, 31, 8, 28, and 20) are also typically distinguishable in the literature. Their positions in the LLM ranking are supported by several centrality measures—such as higher clustering coefficients, favorable PageRank values, or specific eigenvector characteristics—and this part of the network shows strong agreement between the two approaches. The lower-ranked nodes (10, 25, 16, 23, 21, 15, 19, 11, 22, 5, 13, 18, 27, 17, and 12) are characterized by low degree values and generally weak centrality scores. In previous studies, these nodes often remained at the same level, whereas in the LLM-based evaluation they were separated by small differences on the order of 0.001 (Table 6). These distinctions illustrate the method's ability to produce complete separation, yet they also raise the possibility that structurally equivalent nodes may have been "forced apart" by the model.

Table 6. Detailed demonstration of the result score and reason for the node distinguishability levels of Zachary's Karate Club obtained with the LLM-based approach.

Node	Score	Why	Rank	Node	Score	Why	Rank
34	0,914	The most central node in the network with the highest degree, betweenness, closeness, and eigenvector centralities.	1	**7**	0,284	Similar characteristics to node 6, low centrality values	18
1	0,872	One of the most central nodes in the network with high degree, betweenness, closeness, and eigenvector centralities.	2	**26**	0,267	Low degree but relatively high PageRank and eigenvector	19
33	0,801	Very high degree, good betweenness, closeness, and eigenvector centralities.	3	**10**	0,237	Low degree, isolated structure, moderate closeness centrality only	20
3	0,681	High degree and closeness, good betweenness, and eigenvector centralities.	4	**25**	0,218	Low degree, average centrality values, limited influence	21
32	0,674	High degree, good betweenness, closeness, and eigenvector centralities.	5	**16**	0,213	Similar to 15, slightly higher PageRank and eigenvector	22
2	0,634	Good degree and closeness centrality, moderate betweenness and eigenvector.	6	**23**	0,204	Similar to 21, slightly higher PageRank and eigenvector	23

Table 6. (cont.) Detailed demonstration of the result score and reason for the node distinguishability levels of Zachary's Karate Club obtained with the LLM-based approach.

9	0,598	Good degree, high closeness, moderate betweenness, and eigenvector.	7	**21**	0,201	Low degree, average centrality values, limited influence	24
14	0,592	Similar to node 9, good degree and closeness, moderate betweenness.	8	**15**	0,198	Low degree, limited influence, average centrality values	25
4	0,523	Good degree and closeness, low betweenness but high clustering coefficient.	9	**19**	0,196	Similar to 15 and 16, average centrality values	26
24	0,498	Good degree, high PageRank and eigenvector, moderate closeness.	10	**11**	0,195	Similar to node 5, low centrality values	27
31	0,446	Good degree, high closeness and eigenvector, moderate hub/authority.	11	**22**	0,194	Low degree, limited connections, weak centrality measures	28
8	0,425	Perfect clustering, good closeness and eigenvector, zero betweenness.	12	**5**	0,192	Low degree and poor performance on all centrality measures	29
28	0,382	Good degree, high closeness, moderate hub/authority.	13	**13**	0,186	Low degree, limited connections, weak centrality measures	30
20	0,345	Low degree but good Closeness and Eigenvector Centralities	14	**18**	0,184	Low degree, limited connections, weak centrality measures degree, weak centrality measures, limited influence	31
30	0,324	Good degree, moderate performance on all centrality measures	15	**27**	0,179	Low degree, very low closeness and eigenvector centralities	32
29	0,298	Low degree but good closeness centrality	16	**17**	0,143	Lowest degree, isolated node, all centralities are very low	33
6	0,289	Low degree, moderate PageRank, poor other centralities	17	**12**	0,121	The most central node in the network with the highest degree, betweenness, closeness, and eigenvector centralities.	34

Using the Jazz Musicians Network dataset, Sabah and Şimşek reported a maximum distinguishability of 177 for BC centrality and 174 for the CC-DC-EC combination, while other studies reported 191 for Ctμ and 132 for FMD. In our study, 176 Nodes were distinguishable using LLM (Table 7).

Table 7. Comparison of node distinguishability levels obtained from the Jazz Musicians Network dataset in this study and related studies.

Study	**Method**	**Centralities /Approach Used**	**Distinguishability Power**	**Calculation Time**	**Comment**
Fei & Deng (2017)	FMD (Relative entropy + TOPSIS)	DC-CC-BC combinations	**132 level**	Time not reported	Relatively low distinguishing power, uncertain scalability.
Keng et al. (2021)	Ctμ (Convex combinations)	DC-CC-BC combinations	**191 level**	Time not reported (brute-force coefficient selection is costly)	Highest distinguishing power in the literature; however, practical computational time is not given.
Sabah & Şimşek (2023)	Sabah & Şimşek (Entropy + AHP based joint centrality)	DC-CC-BC combinations	**149 level**	0.27–2.51 seconds (small-scale networks), 1.7–68 seconds (SNAP large networks)	Balanced distinguishing power and very fast computation; advantage of practical applicability.
		CC-DC-EC combinations	**174 level**		
This study (2025)	LLM based approach (DeepSeek.)	Direct node-level score estimation	**176 level**	Depends on API calls (minutes–hours)	High distinguishing power; disadvantage of time and cost.

Table 7 shows that the Ctμ method achieved the highest distinguishability level in the literature on the Jazz Musicians Network dataset, but its practical applicability, especially in large networks, is limited due to the lack of reported computational times and the brute-force search for optimal coefficients. The FMD method also lacks time information

and has relatively low distinguishability power. In contrast, the Sabah and Şimşek methods offer full distinguishing power on small-scale networks and a high level of distinguishability on the Jazz Musicians Network dataset. Furthermore, their ability to produce results in seconds makes them applicable to large networks. The LLM-based approach, on the other hand, performed similarly to Sabah and Şimşek in terms of distinguishability power; however, their computational time is in the range of minutes to hours, resulting in a cost disadvantage.

Table 8. Correlation coefficients between LLM-based ranking and classical centrality combinations.

			Kendall's τ-b			Spearman ρ		
Dataset	**Comparison**	**N_nodes**	**Kendall's τ-b**	**95%CI**	**p**	**Spearman ρ**	**95%CI**	**p**
Krackhardt Kite Graph	LLM – BC-CC	10	0.51	[-0.13, 1.00]	0.0466	0.56	[-0.27, 1.00]	0.0897
	LLM – BC-DC		0.51	[-0.13, 1.00]	0.0466	0.56	[-0.27, 1.00]	0.0897
	LLM – CC-DC		0.16	[-0.49, 0.76]	0.601	0.14	[-0.65, 0.90]	0.701
Zachary's Karate Club	LLM – BC-CC	34	0.76	[0.61, 0.89]	2.97e-10	0.89	[0.74, 0.96]	1.9e-12
	LLM – BC-DC		0.77	[0.63, 0.89]	1.67e-10	0.9	[0.77, 0.96]	6.97e-13
	LLM – CC-DC		0.81	[0.69, 0.91]	1.25e-11	0.94	[0.83, 0.98]	3.29e-16
Jazz Musicias Network	LLM – BC-CC	198	0.55	[0.47, 0.62]	1.48e-30	0.72	[0.63, 0.80]	4.37e-33
	LLM – BC-DC		0.61	[0.54, 0.68]	1.45e-37	0.78	[0.70, 0.84]	7.7e-42
	LLM – CC-DC		0.87	[0.84, 0.90]	1.06e-73	0.97	[0.95, 0.98]	8.77e-122
ca-GrQc	LLM – BC-CC	5242	0.75	[0.74, 0.76]	0	0.87	[0.86, 0.88]	0
	LLM – BC-DC		0.6	[0.59, 0.61]	0	0.78	[0.77, 0.80]	0
	LLM – CC-DC		0.77	[0.76, 0.78]	0	0.91	[0.90, 0.91]	0
CA-HepTh	LLM – BC-CC	9877	0.61	[0.60, 0.62]	0	0.76	[0.75, 0.77]	0
	LLM – BC-DC		0.5	[0.49, 0.51]	0	0.67	[0.66, 0.69]	0
	LLM – CC-DC		0.61	[0.60, 0.62]	0	0.76	[0.75, 0.78]	0
CA-CondMat	LLM – BC-CC	23133	0.31	[0.30, 0.32]	0	0.38	[0.37, 0.40]	0
	LLM – BC-DC		0.22	[0.21, 0.23]	0	0.3	[0.28, 0.31]	0
	LLM – CC-DC		0.28	[0.27, 0.29]	0	0.35	[0.34, 0.37]	0

The results in Table 8 suggest that the relationship between LLM-based node rankings and classical centrality measures changes noticeably with network size and structural complexity. In small, well-characterized networks such as Zachary's Karate Club, the LLM aligned closely with classical measures, which implies that the model is able to pick up the essential structural signals present in these graphs. For extremely small networks like the Krackhardt Kite Graph, the broad confidence intervals are largely a consequence of limited sample size rather than a shortcoming of the approach. A different pattern appears in medium-sized networks. In the Jazz Musicians Network, for example, the LLM's rankings were almost indistinguishable from those based on closeness and degree centralities. This suggests that, in settings where the structure is sufficiently rich yet still computationally manageable, the LLM can replicate the differentiations made by classical measures with surprising fidelity. The situation shifts again in large-scale networks. Although correlations remain statistically significant, their magnitudes decline. Here the LLM seems to develop a more composite notion of centrality—one that gradually drifts away from classical formulations. While classical measures tend to compress node scores into narrow ranges in large graphs, making small ranking differences appear disproportionately large, the LLM yields a broader, more internally consistent ranking (with relatively tight confidence intervals) that reflects its own structural intuition.

Table 9. Comparison of node distinction times (in seconds) between related studies and this study.

Dataset	Nodes (n)	Sabah & Şimşek	LLM – DeepSeek
Krackhardt Kite Graph	10	0,27	30,88
Zachary's Karate Club	34	1,00	100,37
Jazz Musicians Network	198	2,51	202,96

The results presented in Table 9 reveal that the Sabah and Şimşek methods and the LLM–DeepSeek approach differ significantly in terms of processing times. For small-scale networks (Krackhardt Kite Graph, 10 nodes; Zachary's Karate Club, 34 nodes), the Sabah and Şimşek methods completed the process in 0.27 and 1.00 seconds, respectively, while the LLM–DeepSeek approach completed the process in 30.88 and 100.37 seconds. Similarly, for the larger Jazz Musicians network (198 nodes), the Sabah and Şimşek method produced results in 2.51 seconds, while this time increased to 202.96 seconds for LLM–DeepSeek. These findings demonstrate that LLM-based approaches are not yet competitive with classical methods in terms of speed. However, the contextual distinguishing power offered by LLMs and the potential to reveal micro-level differences between nodes partially justify this time cost. Therefore, although LLM-based methods need to mature in terms of performance, they provide an alternative and complementary perspective to classical approaches, especially in large and complex networks.

Table 10. Comparison of node distinction times (in seconds) with using batch strategy.

Dataset	Nodes (n)	Mean Time (s)	Std. Dev. (s)	Mean Time (95% CI)
Krackhardt Kite Graph	10	5.69	0.18	[5.68, 5.70]
Zachary's Karate Club	34	19.14	0.56	[19.11, 19.18]
Jazz Musicians Network	198	53.69	1.56	[53.60, 53.80]
CA-GrQc	5242	1297.5	39.67	[1294.98, 1299.91]
CA-HepTh	9877	2264.09	66.37	[2260.07, 2268.32]
CA-CondMat	23133	4995.75	144.61	[4986.82, 5004.84]

In large-scale networks, supplying multiple normalized centrality values for every node to the LLM results in a substantial increase in token usage. Unlike small networks—where the entire graph can be processed in a single request—this becomes impractical as the number of nodes grows and the processing time rises sharply. For that reason, a batch-oriented workflow was adopted to allow the model to evaluate node distinguishability more reliably in datasets with thousands of nodes. In practice, each dataset was partitioned into batches of 15 nodes, and three of these batches were processed in parallel. All experiments ran on Python 3 with a T4 GPU accelerator and a high-memory configuration in Google Colab Pro. This setup improved throughput considerably in large datasets and made the scoring procedure scalable in a way that would not be achievable with single-batch processing.

Table 10 illustrates how effective this batching strategy is. For the smallest networks (10–34 nodes), the dominant cost is simply model initialization; once this overhead is distributed across batches, total processing time drops by more than a factor of five relative to the non-batch scenario. The trend becomes more striking in larger networks. Although the CondMat graph contains over 23,000 nodes—roughly 117 times larger than the Jazz Musicians Network—the total distinction time rises by only about 93 times. This improvement reflects the higher token throughput and the added benefit of parallel decoding in batch mode.

Taken together, the timing results indicate that the LLM-driven approach scales in an almost linear fashion and produces stable, repeatable outputs even in networks with very large node counts. This behavior is especially valuable when compared with computationally intensive measures such as betweenness centrality, which become prohibitively costly at scale.

Effective node identification methods in complex networks offer different advantages depending on the network's structural properties and application area. While classical centrality measures are fast and simple, they may not fully reflect the network's global dynamics (Ait Rai et al., 2023; D. Chen et al., 2012; Yang et al., 2020). Entropy-based and hybrid methods provide higher accuracy by considering dynamics such as information diffusion and uncertainty (Dong & Zhou, 2021; Guo et al., 2020; Liu et al., 2024; Wang et al., 2022). Machine learning and deep learning-based approaches, on the other hand, demonstrate superior performance, particularly in large and complex networks, by integrating node characteristics with network structure (Chen et al., 2024; Wu et al., 2024; Zhao, Jia, Huang, et al., 2020; Zhao, Jia, Zhou, et al., 2020).

The results confirm that classical centralities remain efficient and reliable for structural analysis, but tend to produce ties and limited differentiation in larger or symmetric networks. Hybrid methods improve resolution, while LLM-assisted evaluations add semantic interpretability by contextualizing node roles.

E. *Qualitative Analysis of LLM Explanations and Their Alignment with Centrality Concepts*

The explanations generated by the LLM generally align well with the intuitions behind classical centrality measures. In the Jazz Musicians Network, for example, the model describes central nodes with remarks such as "A highly central node because it has excellent or very high values in all centrality measures" or "It occupies a very central position due to its high degree, closeness, and eigenvector values." These statements echo the basic interpretations of classical centrality: a high degree indicates dense local connectivity, high closeness reflects shorter average path distances, and a strong eigenvector value suggests links to influential or well-connected neighbors. In the CA-GrQc network, phrases like "One of the most central nodes because its degree and eigenvector centralities are excellent (1.000)" or "Highly central because all centrality values are high" similarly capture the structural traits expected of central nodes in large-scale graphs. Taken together, these examples show that the LLM is not merely repeating numerical patterns but forming explanations that reflect an intuitive grasp of how centrality operates within the network. The model's interpretations of low-scoring nodes are equally consistent. In the Jazz Musicians Network, statements such as "Not central because all centrality measures show very low values" or "Not central because degree and betweenness are zero, and others are very low" map neatly onto the structural characteristics of peripheral nodes. Comparable descriptions appear in the CA-GrQc dataset, where the LLM notes that a node is "completely isolated because all centrality values are zero" or "a marginal member of the network because its centrality values are minimal." The same pattern is observed in CA-HepTh and CA-CondMat, where near-zero centrality values are interpreted as "a marginal position in the network because it has no centrality measures." These explanations demonstrate that the model can read and interpret the informational content of centrality values in a manner consistent with classical theory. When both high- and low-scoring examples are viewed together, the LLM's explanations show a coherent interpretive structure. They refer explicitly to each node's centrality profile, remain faithful to the conceptual meanings associated with classical measures, and apply this reasoning consistently across different types of networks. Overall, these findings suggest that LLM-based joint scoring is not only practical in terms of performance but also theoretically grounded and interpretable, reflecting established ideas in centrality research rather than diverging from them.

IV. CONCLUSION AND FUTURE WORK

This study presents a hybrid framework that brings together classical and LLM-based centrality measures, aiming to improve both the quantitative evaluation of node importance and the interpretability of the resulting rankings. In doing so, it highlights the strengths of established approaches—their deterministic behavior, consistency, and sensitivity to structural symmetry—while examining how an LLM-based alternative compares, where it departs from classical expectations, and why it sometimes fails to preserve structural symmetry. The contribution of the study lies not only in proposing an LLM-supported composite scoring method but also in evaluating its constraints and situating it within a balanced comparison against classical joint centrality formulations. Looking ahead, applying this framework to larger real-world networks and investigating domain-specific fine-tuning to enhance LLM accuracy appear to be promising directions for continued work.

DECLARATIONS

Acknowledgements: No acknowledgment statement is included in the manuscript.

Author Contributions: Conceptualization, L.S.; Methodology, L.S. and Z.B.S.; Validation, L.S. and Z.B.S.; Investigation, L.S.; Resources, L.S.; Data Curation, L.S.; Writing—Original Draft, L.S.; Writing—Review & Editing, L.S. and Z.B.S.; Supervision, R.K.

Conflict of Interest: The author/authors declare no conflict of interest.

Supporting Institutions: This research received no external funding.

Ethical Approval: This study does not involve human or animal participants. All procedures followed scientific and ethical principles, and all referenced studies are appropriately cited.

Plagiarism Statement: This article has been evaluated for plagiarism and no instances of plagiarism were detected.

Data Availability: Data sharing is not applicable to this study.

Use of AI Tools: Artificial Intelligence (AI) tools were employed solely for language editing purposes during the preparation of this manuscript. The authors carefully reviewed all AI-assisted edits and take full responsibility for the accuracy and integrity of the final content.

REFERENCES

Ait Rai, K., Machkour, M., & Antari, J. (2023). Influential nodes identification in complex networks: a comprehensive literature review. *Beni-Suef University Journal of Basic and Applied Sciences*, *12*(1), Article 18. https://doi.org/10.1186/s43088-023-00357-w

Chaudhary, L., & Singh, B. (2023). Gumbel-SoftMax based graph convolution network approach for community detection. *International Journal of Information Technology*, *15*(6), 3063–3070. https://doi.org/10.1007/s41870-023-01347-y

Chen, D., Lü, L., Shang, M.-S., Zhang, Y.-C., & Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and Its Applications*, *391*(4), 1777–1787. https://doi.org/10.1016/j.physa.2011.09.017

Chen, L., Xi, Y., Dong, L., Zhao, M., Li, C., Liu, X., & Cui, X. (2024). Identifying influential nodes in complex networks via Transformer. *Information Processing & Management*, *61*(5), Article 103775. https://doi.org/10.1016/j.ipm.2024.103775

DeepSeek-AI, Bi, X., Chen, D., Chen, G., Chen, S., Dai, D., Deng, C., Ding, H., Dong, K., Du, Q., Fu, Z., Gao, H., Gao, K., Gao, W., Ge, R., Guan, K., Guo, D., Guo, J., Guangbo, H., … Zou, Y. (2024). *DeepSeek LLM: Scaling open-source language models with long-termism. arXiv.* https://doi.org/10.48550/arXiv.2401.02954

Dong, S., & Zhou, W. (2021). Improved influential nodes identification in complex networks. *Journal of Intelligent & Fuzzy Systems*, *41*(6), 6263–6271. https://doi.org/10.3233/JIFS-202943

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

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

Gleiser, P. M., & Danon, L. (2003). Community structure in jazz. *Advances in Complex Systems*, *6*(4), 565–573. https://doi.org/10.1142/S0219525903001067

Guo, C., Yang, L., Chen, X., Chen, D., Gao, H., & Ma, J. (2020). Influential nodes identification in complex networks via information entropy. *Entropy*, *22*(2), Article 242. https://doi.org/10.3390/e22020242

Huang, Z., Tang, Y., & Chen, Y. (2022). A graph neural network-based node classification model on class-imbalanced graph data. *Knowledge-Based Systems*, *244*, Article 108538. https://doi.org/10.1016/j.knosys.2022.108538

Katz, L. (1953). A New Status Index Derived from Sociometric Analysis. *Psychometrika*, *18*(1), 39–43. https://doi.org/10.1007/BF02289026

Keng, Y. Y., Kwa, K. H., & McClain, C. (2021). Convex combinations of centrality measures. *The Journal of Mathematical Sociology*, *45*(4), 195–222. https://doi.org/10.1080/0022250X.2020.1765776

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, *46*(5), 604–632. https://doi.org/10.1145/324133.324140

Krackhardt, D. (1990). Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Administrative Science Quarterly*, *35*(2), 342–369. https://doi.org/10.2307/2393394

Li, S., Quan, Y., Luo, X., & Wang, J. (2025). Influential nodes identification for complex networks based on multi-feature fusion. *Scientific Reports*, *15*(1), Article 11440. https://doi.org/10.1038/s41598-025-94193-w

Li, Y., Yang, Y., Zhu, J., Chen, H., & Wang, H. (2024). LLM-empowered few-shot node classification on incomplete graphs with real node degrees. In *Proceedings* of the *33rd ACM International Conference on Information and Knowledge Management* (pp. 1306–1315). https://doi.org/10.1145/3627673.3679861

Liu, W., Lu, P., & Zhang, T. (2024). Identifying influential nodes in complex networks from semi-local and global perspective. *IEEE Transactions on Computational Social Systems*, *11*(2), 2105–2120. https://doi.org/10.1109/TCSS.2023.3295177

Lyu, B., Hamdi, M., Yang, Y., Cao, Y., Yan, Z., Li, K., Wen, S., & Huang, T. (2023). Efficient Spectral Graph Convolutional Network Deployment on Memristive Crossbars. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *7*(2), 415–425. https://doi.org/10.1109/TETCI.2022.3210998

Mohammad, B. B., Dhuli, S., & Enduri, M. K. (2025). Isolating Centrality-Based Generalization of Traditional Centralities to Discover Vital Nodes in Complex Networks. *Arabian Journal for Science and Engineering*, *50*(15), 12003–12025. https://doi.org/10.1007/s13369-024-09628-9

Mukhtar, M. F., Abal Abas, Z., Baharuddin, A. S., Norizan, M. N., Fakhruddin, W. F. W. W., Minato, W., Rasib, A. H. A., Abidin, Z. Z., Rahman, A. F. N. A., & Anuar, S. H. H. (2023). Integrating local and global information to identify influential nodes in complex networks. *Scientific Reports*, *13*(1), Article 11411. https://doi.org/10.1038/s41598-023-37570-7

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web* (Stanford InfoLab Technical Report No. 1999–66). Stanford University. http://ilpubs.stanford.edu:8090/422/

Qiu, L., Zhang, J., & Tian, X. (2021). Ranking influential nodes in complex networks based on local and global structures. *Applied Intelligence*, *51*(7), 4394–4407. https://doi.org/10.1007/s10489-020-02132-1

Sabah, L., & Şimşek, M. (2023). A new fast entropy-based method to generate composite centrality measures in complex networks. *Concurrency and Computation: Practice and Experience*, *35*(10), Article e7657. https://doi.org/10.1002/cpe.7657

Salavati, C., Abdollahpouri, A., & Manbari, Z. (2019). Ranking nodes in complex networks based on local structure and improving closeness centrality. *Neurocomputing*, *336*, 36–45. https://doi.org/10.1016/j.neucom.2018.04.086

Şimşek, A. (2024). Sosyal ağlarda merkezilik ölçütleri kullanılarak makine öğrenmesi ile etkili bireylerin tespiti. *Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *29*(1), 166–172. https://doi.org/10.53433/yyufbed.1348472

Leskovec, J., & Krevl, A. (2014). *SNAP datasets: Stanford large network dataset collection* [Data set]. Stanford University. https://snap.stanford.edu/data/

Ullah, A., Wang, B., Sheng, J., Long, J., Khan, N., & Sun, Z. (2021). Identification of nodes influence based on global structure model in complex networks. *Scientific Reports*, *11*(1), Article 6173. https://doi.org/10.1038/s41598-021-84684-x

Wang, B., Zhang, J., Dai, J., & Sheng, J. (2022). Influential nodes identification using network local structural properties. *Scientific Reports*, *12*(1), Article 1833. https://doi.org/10.1038/s41598-022-05564-6

Wang, S., Huang, J., Chen, Z., Song, Y., Tang, W., Mao, H., Fan, W., Liu, H., Liu, X., Yin, D., & Li, Q. (2025). Graph machine learning in the era of large language models (LLMs). *ACM Transactions on Intelligent Systems and Technology*, *16*(5), 1–40. https://doi.org/10.1145/3732786

Wu, Y., Hu, Y., Yin, S., Cai, B., Tang, X., & Li, X. (2024). A graph convolutional network model based on regular equivalence for identifying influential nodes in complex networks. *Knowledge-Based Systems*, *301*, Article 112235. https://doi.org/10.1016/j.knosys.2024.112235

Xu, X., Zhu, C., Wang, Q., Zhu, X., & Zhou, Y. (2020). Identifying vital nodes in complex networks by adjacency information entropy. *Scientific Reports*, *10*(1), Article 2691. https://doi.org/10.1038/s41598-020-59616-w

Yang, Y., Wang, X., Chen, Y., Hu, M., & Ruan, C. (2020). A novel centrality of influential nodes identification in complex networks. *IEEE Access*, *8*, 58742–58751. https://doi.org/10.1109/ACCESS.2020.2983053

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. Journal of Anthropological Research, 33(4), 452–473.

Zhang, Y., & Xiao, X. (2022). A dynamic community detection method for complex networks based on deep self-coding network. *Computational Intelligence and Neuroscience*, *2022*(1), Article 7084084. https://doi.org/10.1155/2022/7084084

Zhao, G., Jia, P., Huang, C., Zhou, A., & Fang, Y. (2020). A machine learning–based framework for identifying influential nodes in complex networks. *IEEE Access*, *8*, 65462–65471. https://doi.org/10.1109/ACCESS.2020.2984286

Zhao, G., Jia, P., Zhou, A., & Zhang, B. (2020). InfGCN: Identifying influential nodes in complex networks with graph convolutional networks. *Neurocomputing*, *414*, 18–26. https://doi.org/10.1016/j.neucom.2020.07.028

Zhao, J., Wang, Y., & Deng, Y. (2020). Identifying influential nodes in complex networks from global perspective. *Chaos, Solitons & Fractals*, *133*, Article 109637. https://doi.org/10.1016/j.chaos.2020.109637