Comparative Evaluation of Centrality Measures for Detecting Significant Nodes in Social Networks

Hardeep Singh*, Supreet Kaur, Karman Singh Sethi, Jai Sharma
Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

***Abstract*—Social networks are certainly a crucial platform for bringing people together globally. Detecting the significant nodes inside the social network remains an open problem because of the broad variety of network sizes. To solve this problem, different centrality measures have been introduced. Detecting the significant nodes is essential for speeding up or slowing down the spread of information, managing diseases and rumors, and more. This paper presents a comparative evaluation of 12 centrality measures to determine the most effective measure on the basis of accuracy, differentiation capability, and runtime. To validate performance, a series of experiments is conducted on four social networks using the validation metrics such as monotonicity, the SIR model, and Kendall tau. The experimental outcomes indicate that the gravity-based measures have superior accuracy and differentiation capability as compared to other measures. Finally, this paper outlines future research directions for enhancements based on centrality measures.**

***Keywords*—Social networks; centrality measures; significant nodes; validation metrics; extended gravity**

I. Introduction

Social networks depict the structure of actual social systems in which vertices denote individuals and edges denote their connections. Globally, billions of people own smartphones, and they are adept at using social media applications, including Facebook, Threads, Bluesky, and others, to communicate and disseminate information. Consequently, social networks offer an environment for promoting items, disseminating information, and other purposes. Table I lists a few instances of social networks. Many tools for exploring social networks have been provided by technology in recent years, like NetworkX, Pajek, VOSviewer, etc.

Detecting the significant nodes is crucial for carrying out the aforementioned tasks, like promoting products, spreading information, etc., as a node with a better spreading capacity is regarded as significant in social networks [1]. Significant nodes have a larger effect on the topology of the network in contrast to the remaining nodes [2]. Hence, detecting the significant nodes aids in a thorough analysis of the whole network. For instance, in the management of disease spread, we can control the spread of disease by quickly finding the most critical node of disease transmission [3]. The significant node inside a network is designated by its centrality [4]. As a result, determining the measure of centrality is crucial. Different centralities have been stated due to the enormous growth of social networks, like closeness [5], degree [6], betweenness [7], k-shell decomposition [8], gravity [9], etc. These measures can be classified as local, semi-local, and global, based on node connectivity [10], as displayed in Fig. 1.

Local knowledge regarding the node is utilized to develop local centralities. For instance, degree [6], ProfitLeader [11], etc. Recent years have seen the development of a novel class of centralities known as semi-local, which can retain a higher accuracy than local centralities. For instance, h-index [12], semi-local centrality [13], local gravity model [14], etc. Centralities that necessitate the whole framework of a network are known as global. For instance, eigenvector [6], PageRank [15], global and local structure [16], electric potential [17], etc. This paper evaluates 12 centralities and highlights the most effective measure. The findings suggest that the gravity-based measures, that is, gravity centrality and extended gravity centrality, have superior accuracy and differentiation capability as compared to other measures that include eigenvector, degree, closeness, betweenness, PageRank, h-index, k-shell, ProfitLeader, global and local structure, and electric potential.

This paper's remaining sections are laid out this way: Section II discusses the related studies, Section III describes several centrality measures, and Section IV presents the experimental methodology. Lastly, Sections V and VI conclude the paper and highlight future directions.

TABLE I. FEW INSTANCES OF SOCIAL NETWORKS

Social Networks	
Examples	***Applications***
Friendship	WhatsApp, Facebook, Twitter, etc.
Interaction	Emails, Snapchat, Phone Calls, etc.
Co-authorship	DBLP, ScienceDirect, Wikibooks, etc.
Follower	LinkedIn, Twitter, etc.

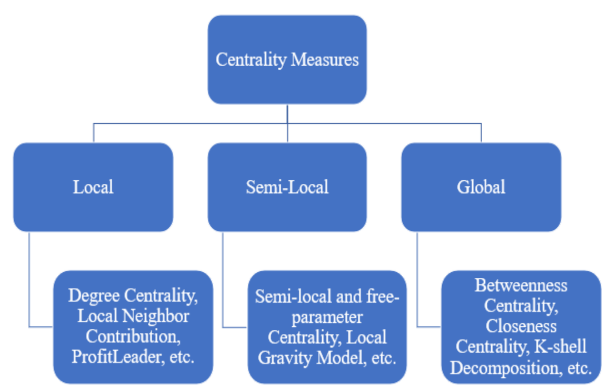

Fig. 1. Classification of centrality measures.

*Corresponding author.

II. Related Studies

In recent years, discovering the highly significant nodes in social networks has emerged as an important research challenge. Many centrality measures have been designed for detecting significant nodes, and each measure has advantages and disadvantages. Bavelas originated the concept of a centrality measure for connected networks [18]. Freeman built three well-known measures of centrality: closeness, betweenness, and degree [7]. Closeness is the global measure that considers both the direct and indirect connections. Betweenness is also the global measure that is used to identify the bridge nodes. Both these global measures have a high computational cost. The local measure, referred to as degree, considers only direct connections. A node becomes significant according to Bonacich's eigenvector measure if it has significant neighbors [6]. Search engines widely and efficiently use the global measure known as PageRank [15], which Brin & Page invented, to rank web pages. A web page attains high quality when many high-quality web pages link to it.

A semi-local measure called the h-index was created [12]; according to this measure, a node is considered more significant when it is linked to many large-degree neighboring nodes. In order to investigate human brain networks, a unique measure known as leverage centrality was invented [19]. Kitsak et al. designed the global measure called the k-shell that has low computational cost [8]. Bae & Kim invented the coreness measure, which calculates a node's influence by considering its neighbors' k-shell values [20]. Ma et al. built gravity-based measures that depend on Newton's gravitational formula [9]. A k-shell hybrid technique was created to detect the significant nodes from both the outer shells and the core [21]. Yu et al. introduced ProfitLeader, a measure that evaluates the importance of a node through its profit capacity [11]. A node's importance increases when it offers greater profit for others.

A method based on local neighbor contribution was suggested, which can be utilized with networks of different sizes and has a relatively low computational cost [22]. Sheng et al. suggested a measure called global and local structure that takes into account both local influence and global influence [16]. Through the use of the effective distance, a gravity model was developed that integrates static as well as dynamic information [23]. To detect the significant nodes, extended hybrid characteristic centrality was designed by using an extended version of the degree and E-shell technique [24]. A gravity-based model was designed that incorporates degree, k-shell, and eigenvector centrality measures [25]. An electric potential-based measure was developed that includes both local and global features of networks [17]. A method named OVED-Rank was suggested that uses orbital velocity formula and effective distance to detect significant nodes in social networks [26]. The previous studies on different centrality measures are presented in Table II.

TABLE II. PREVIOUS STUDIES ON DIFFERENT CENTRALITY MEASURES

Centrality Measure	Category	Strengths	Weaknesses/ Future Work
Degree [6]	Local	Simple with low computational cost	Low accuracy
Eigenvector [6]	Global	Concentrates on both the quantity and quality of neighboring nodes	Applicable solely to connected networks
H-index [12]	Semi-local	Better node ranking than degree	Poor differentiation capability
Gravity and its enhanced variant [9]	Global	Superior accuracy	High computational cost
ProfitLeader [11]	Local	Low computational cost	Not very effective for small networks
Local gravity model [14]	Semi-local	Introduces the truncation radius R to minimize time complexity	To compute the truncation radius R as an auxiliary parameter
Global and local structure [16]	Global	Superior accuracy	High computational cost
Local and global influence [27]	Global	Superior accuracy and differentiation capability	Less efficiency
Effective distance gravity model [23]	Global	More comprehensive	High computational cost
Extended hybrid characteristic [24]	Global	Superior accuracy and differentiation capability	Restricted to undirected unweighted networks
OVED-Rank [26]	Global	Superior accuracy	Restricted to undirected unweighted networks

III. Preliminary Concepts

This section provides 12 currently and previously created centralities. A social network is represented as a graph $G = (V, E)$, where V and E designate the node set and edge set, respectively. Various measures are given as follows:

A. Degree

It represents a simple approach for calculating a node's importance by merely taking into account its closest neighbors [6]. The degree centrality C_D [7] of a node w can be computed as:

$$C_D(w) = \frac{\sum_{x=1}^{N} a_{wx}}{N-1} \quad (1)$$

where, $A = \{ a_{wx}\}$ designates the adjacency matrix, $a_{wx} = 1$ if w and x are connected, and 0 if not; N designates the node count.

B. Betweenness

The count of shortest paths that travel through a node w determines its betweenness centrality C_B [7], which is computed as:

$$C_B(w) = \frac{2}{(N-1)\,(N-2)} \sum_{i\neq w, i\neq j, w\neq j} \frac{g_{ij}(w)}{g_{ij}} \quad (2)$$

where, $g_{ij}(w)$ designates the number of shortest paths joining nodes i and j that travel through a node w.

C. Closeness

It calculates a node's closeness to every other node in the network [5]. The closeness centrality C_c [7] can be computed as:

$$C_c(w) = \frac{N-1}{\sum_{x \neq w} d(w,x)} \quad (3)$$

where, $d(w,x)$ designates the length of the shortest path between nodes w and x.

D. Eigenvector

It measures a node's importance by taking into account both the number of its adjacent nodes and their importance [6]. The eigenvector centrality C_E can be computed as:

$$Az = \lambda z, C_E(w) = z_w = \frac{1}{\lambda} \sum_{x=1}^{N} a_{wx} z_x \quad (4)$$

where, z designates the eigenvector and λ designates the constant.

E. PageRank

It is a variant of the eigenvector measure [15]. It demonstrates that a web page's importance can be measured based on both the count of linking pages and the quality of those pages. The PageRank $C_{PR}^{(k)}$ of a node w at the k-th iteration is computed as:

$$C_{PR}^{(k)}(w) = \sum_{x=1}^{N} a_{xw} \frac{C_{PR}^{(k-1)}(x)}{k_x^{out}} \quad (5)$$

where, k_x^{out} designates the count of edges from node x to w.

F. K-Shell Decomposition

It calculates the importance of a node according to its topological position [8]. It allocates each node a k-shell index through an iterative process. The most significant node holds the maximum k-shell value and is located close to the network's core. Core nodes are more significant than non-core nodes. Within the same core level, it cannot distinguish between influential nodes.

G. H-Index

The H-index (Hirsch index) was first employed to evaluate the scholarly contributions of scholars or journals according to their scholarly works and citation counts [28]. The H-index C_{HI} of a node w is expressed as the highest value h such that w has at least h neighbors for each with a degree no less than h [12]. A node's higher H-index value specifies its greater influence.

H. Gravity

It is a hybrid measure, which makes use of the path information as well as neighborhood information [9]. The gravity centrality C_G can be computed as:

$$C_G(w) = \sum_{x \epsilon \psi_w} \frac{KS(w)KS(x)}{d^2(w,x)} \quad (6)$$

where, $KS(w)$ and $KS(x)$ designate the k-shell value of nodes w and x, respectively, and ψ_w designates the set of third-order neighbors.

I. Extended Gravity

This enhanced variant of gravity measure computes a node's importance by summing the influence $C_G(x)$ of all nearest neighbors [9]. The extended gravity centrality C_{EG} can be computed as:

$$C_{EG}(w) = \sum_{x \epsilon \Lambda_w} C_G(x) \quad (7)$$

where, Λ_w designates the set of adjacent nodes of a node w.

J. ProfitLeader

ProfitLeader uses each node's profit capacity to evaluate its influence [11]. The profit capacity C_{PL} can be computed as:

$$C_{PL}(w) = \sum_{x \epsilon \Gamma_w} AR_{(w \to x)} \cdot SP_{(w \to x)} \quad (8)$$

where, Γ_w designates the neighbors of node w, $AR_{(w \to x)}$ designates the available resource of a node w for another node x, and $SP_{(w \to x)}$ designates the sharing probability (or similarity) between two nodes w and x.

K. Global and Local Structure

It computes the node's influence with two key factors: local influence and global influence [16]. The global and local structure C_{GLS} is computed as:

$$C_{GLS}(w) = I_G(w) \times I_L(w) \quad (9)$$

The global influence $I_G(w)$ can be computed as:

$$I_G(w) = D(w) \sum_{x \epsilon Nei_w} pow(A, \#Com(w,x)) \quad (10)$$

where, A is a constant, $\#Com(w,x)$ designates the count of mutual neighbors of nodes w and x, and $D(w)$ designates the degree and is computed as:

$$D(w) = \sum_{x=1}^{N} a_{wx} \quad (11)$$

The other factor is local influence. The $I_L(w)$ is computed as:

$$I_L(w) = \sum_{x \epsilon Nei_w} C_D(x) p(x) \quad (12)$$

where, Nei_w designates all nearest neighbors of a node w, and every neighbor x has a probability p of contributing to node w.

L. Electric Potential Centrality

It relies on the principle of electric potential, which represents the network's structural features on both a local and global level [17]. The electric potential C_{EP} of a node w is computed as:

$$C_{EP}(w) = LOC(w) \times GOC(w) \quad (13)$$

where, $LOC(w)$ and $GOC(w)$ designate the locality and globality of a node w, respectively.

The $LOC(w)$ is computed as:

$$LOC(w) = D(w) \times e^{\frac{D(w) + \varphi(w)}{N}} \quad (14)$$

where, e designates the base of the natural logarithm, and $\varphi\ (w)$ is computed as:

$$\varphi\ (w) = \frac{2\times\alpha}{(|D(w)-1|+1)} \quad (15)$$

where, α designates a tunable coefficient.

The $GOC(w)$ is computed as:

$$GOC(w) = \frac{D(x)\times\beta}{d\,(w,x)} \quad (16)$$

where, β designates a tunable coefficient.

So, the $C_{EP}(w)$ is computed as:

$$C_{EP}(w) = \sum_{x\epsilon\gamma_w}(D(w)\times\ e^{\frac{D(w)+\varphi\ (w)}{N}})\times\frac{D(x)\times\beta}{d\,(w,x)} \quad (17)$$

IV. Experimental Methodology

This section assesses the 12 centrality measures utilizing the three parameters mentioned below:

a) Accuracy: It reflects how effectively the measure's ranking order correlates with the true ranking order [27].

b) Differentiation Capability: It depicts the effectiveness of the measure to discriminate the nodes' importance [27].

c) Efficiency: It indicates the runtime required by the centrality for evaluating the importance of the nodes [27].

A. Validation Metrics

1) SIR Model: It is applied to determine the propagating influence of ranked nodes [29], [30] and has three states:

- Susceptible *(S)*: People with good health who are prone to contamination from other people are included in this state.
- Infected *(I)*: People with the illness may be likely to infect other people.
- Recovered *(R)*: After recovering, people with the illness are unable to infect other people.

All nodes are initially within the *S* state, with just node *y* within the *I* state. Every time step, the *I* nodes aim to infect those around them within the *S* state with an infection probability β. Nodes within the *S* state move to the *I* state upon infection. The *I* nodes then move to the *R* state with a recovery probability μ. This cycle repeats until the network contains no *I* nodes. Once all *I* nodes have shifted to the *R* state, we count the total *R* nodes to quantify the influence of node *y*, as displayed in Fig. 2. The threshold value of β can be computed as:

$$\beta_{th} = \frac{<K>}{<K^2>-<K>} \quad (18)$$

where, $< K >$ designates the average degree and $< K^2 >$ designates the second-order average degree [31].

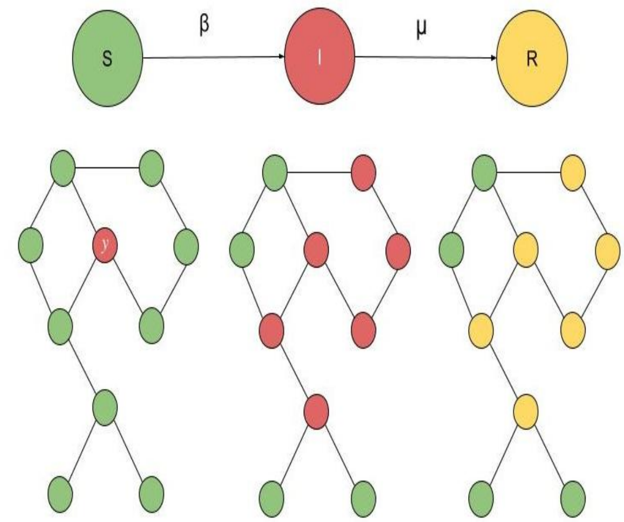

Fig. 2. The SIR model.

2) Kendall Tau τ: Kendall τ [32], [33] is applied to quantify the correlation of the ranking list of the measure with that of the SIR model. Suppose two ranking lists, Q and R, are considered for correlation and have similar nodes $N: Q = (q_1, q_2, \ldots, q_N)$ and $R = (r_1, r_2, \ldots, r_N)$.

Kendall τ between Q and R is computed as:

$$\tau(Q,R) = \frac{N_c - N_d}{0.5N(N-1)} \quad (19)$$

where, N_c and N_d designate the count of concordant and discordant pairs, respectively. A ranked list produced by centrality is more accurate when the τ value is higher [3].

B. Data Description

The topological properties of social networks are mentioned in Table III. The graphical visualization of four networks using the NetworkX Python Package is displayed in Fig. 3. These networks are openly accessible and available for download from the site https://networkrepository.com/

1) LastFM: In March 2020, the public API was used to gather this LastFM user social network, with the nodes signifying LastFM members from Asian nations and the edges signifying reciprocal follower ties.

2) Power: This is a representation of the structure of the US power grid in the Western States, where nodes signify power stations and edges signify lines between them.

3) Webspam: The Purdue University graph database created this dataset. Nodes are signified by web pages, and edges by hyperlinks.

4) Dolphins: It describes the network of bottlenose dolphins. Every edge signifies a dolphin-to-dolphin link, and every node signifies a bottlenose dolphin.

TABLE III. TOPOLOGICAL PROPERTIES OF FOUR SOCIAL NETWORKS

Network	N (Node count)	M (Edge count)	K_{max} (Maximum degree)	$<K>$ (Average degree)	$<CC>$ (Average clustering coefficient)	β_{th} (Infection threshold)
LastFM	7624	27806	216	7.2943	0.2194	0.0409
Power	4941	6594	19	2.6691	0.0801	0.3483
Webspam	4767	37375	477	15.6807	0.2860	0.0140
Dolphins	62	159	12	5.1290	0.2590	0.1723

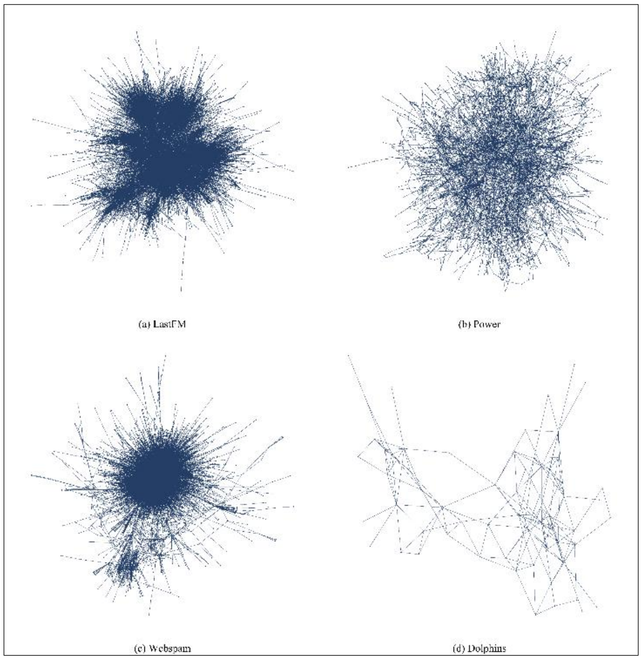

Fig. 3. Graphical visualization of four networks using the NetworkX Python Package.

C. Results

We have employed the validation metrics to compare the effectiveness of 12 measures in detecting the significant nodes. Firstly, we employ a sample network, consisting of 15 nodes and 18 edges, as displayed in Fig. 4. By utilizing this sample network, we first obtain the ranking lists based on 12 measures and the SIR model, as depicted in Table IV. We then compare the ranking list of each measure with that of the SIR model to identify the common nodes between the two lists. With this comparison, we observe that the number of common nodes for C_G, C_{EP}, C_{EG}, C_{PL}, C_C, C_{GLS}, C_{HI}, C_B, C_E, C_{PR}, C_{KS}, and C_D with the SIR model is 10, 7, 6, 6, 5, 5, 4, 3, 3, 2, 2, and 1, respectively. According to the above analysis, it can be seen that the C_G measure has the most common nodes, which means the ranking list attained from the C_G measure and the SIR model match well. Therefore, the C_G measure detects the significant nodes better than the other measures. In addition, we compare these measures on the basis of Kendall τ, as depicted in Fig. 5, and find that the C_G measure performs superior to the other measures.

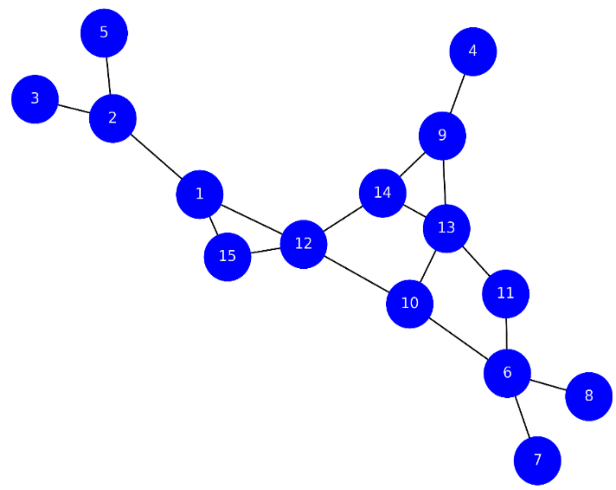

Fig. 4. A sample network with 15 nodes and 18 edges.

TABLE IV. SAMPLE NETWORK RANKING LISTS

Rank	C_D	C_B	C_C	C_E	C_{PR}	C_{HI}	C_{KS}	C_G	C_{EG}	C_{PL}	C_{GLS}	C_{EP}	SIR
1	6	12	12	13	6	13	1	12	13	12	13	12	12
2	12	1	10	12	12	14	6	13	12	13	12	13	13
3	13	10	14	14	13	10	14	10	10	6	10	6	10
4	10	6	13	10	2	12	12	14	14	14	14	10	14
5	9	2	1	9	1	1	11	6	9	1	6	14	6
6	2	13	6	6	9	6	10	1	1	10	1	1	1
7	1	14	15	1	10	9	9	9	6	9	9	9	9
8	14	9	9	11	14	15	15	11	11	2	11	2	15
9	11	11	11	15	11	11	13	15	15	11	2	11	11
10	15	4	2	2	15	7	7	2	2	15	15	15	2
11	3	3	7	4	3	3	3	7	7	3	7	7	8
12	7	7	8	7	5	2	2	8	8	4	8	8	4
13	5	5	4	8	7	5	5	4	4	5	4	4	7
14	4	8	5	5	8	4	4	5	5	7	3	5	5
15	8	15	3	3	4	8	8	3	3	8	5	3	3

The research work involves four experiments based on the three parameters outlined below:

1) Accuracy:

a) Experiment: Comparison of centrality measures using Kendall's τ values.

Kendall's τ serves as a validation metric to quantify the accuracy of centralities. The accuracy of the 12 measures, evaluated by Kendall's τ, is shown in Table V. Table V highlights that C_{EG} and C_G, the gravity-based measures, are highly effective and deliver the best outcomes among all 12 measures. The extended gravity measure performs the best, while the gravity measure ranks second. Among all measures, the C_B and C_{PR} measures perform the worst, with minimal correlation in most networks. Table V emphasizes the three best-performing measures in bold.

TABLE V. VALUES OF KENDALL TAU FOR 12 MEASURES

Network	C_D	C_B	C_C	C_E	C_{PR}	C_{HI}	C_{KS}	C_G	C_{EG}	C_{PL}	C_{GLS}	C_{EP}
LastFM	0.5829	0.4112	0.6821	0.6761	0.4222	0.6016	0.6101	**0.8081**	**0.8468**	0.5779	**0.7456**	0.6617
Power	0.4401	0.3275	0.3850	**0.6306**	0.2441	0.4188	0.3451	**0.6843**	**0.7906**	0.4327	0.6065	0.5498
Webspam	0.7124	0.5132	0.8093	**0.8456**	0.5274	0.7267	0.7315	0.8252	**0.8592**	0.7011	**0.8329**	0.7742
Dolphins	0.7472	0.5410	0.6584	0.6986	0.6647	0.7451	0.5219	**0.8488**	**0.8974**	0.7409	0.8202	**0.8308**

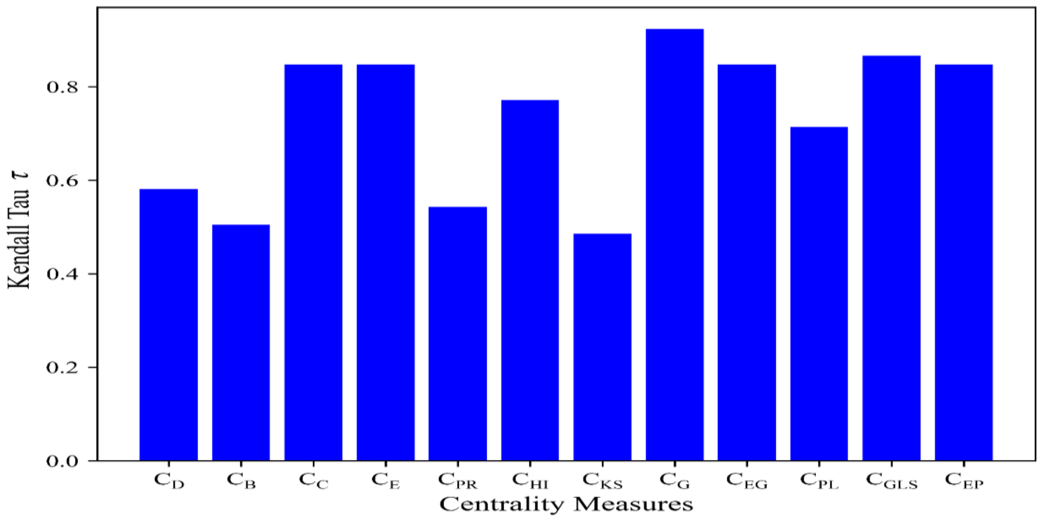

Fig. 5. Kendall τ outcomes of the sample network.

b) Experiment: Comparison of the centrality measures based on propagation influence.

To assess the effectiveness of 12 centralities, we compute the propagating ability of ranked nodes by employing the SIR model. For differentiating the significant nodes, each node's influence is first computed using the 12 centrality measures and then ranked in descending order. Second, every ranked node is utilized as an initial infected node (seed node) to infect susceptible nodes with an infection probability β. Finally, the count of nodes infected by each seed node is computed. The transmission rate of ranked nodes is plotted across the networks, as depicted in Fig. 6. The ranked index signifies the nodes' ranking order in descending order according to their influence, whereas *F(t)* signifies the total of infected and recovered nodes at time *t*. The SIR model states that highly significant nodes have the capacity to infect a greater number of other nodes. For a measure to be efficient, its infection curve should show a consistent reduction in the number of infected nodes when the node influence decreases from left to right. As shown in Fig. 6, the C_{EG} measure exhibits a remarkable infection capability compared to the other measures across all networks, with its curves appearing smooth and stable with minimal fluctuations. The C_G measure also performs well in all networks, but not as well as the C_{EG} measure. In contrast, the infection curves of the C_B and C_{PR} measures display numerous peaks and troughs in all networks. Thus, both C_B and C_{PR} measures show the worst performance.

2) Differentiation capability:

a) Experiment: Comparison of centrality measures based on monotonicity M values.

The monotonicity M [20] is applied to quantify the differentiation ability of centrality measures and is computed as:

$$M(Y) = \left[1 - \frac{\sum_{r \epsilon Y} N_r(N_r - 1)}{N(N-1)}\right]^2 \quad (20)$$

where, Y designates the ranking pattern, and N_r designates the node count with the identical rank r.

The monotonicity M values fall between 0 and 1. The higher the monotonicity value, the better the differentiation capability. In Table VI, it can be clearly shown that the C_{EP}, C_E, C_{PR}, and C_{EG} measures have the best differentiation capability as compared to other measures. In every network, the C_{KS} and C_{HI} measures exhibit the poorest performance. Table VI emphasizes the best-performing measures in bold.

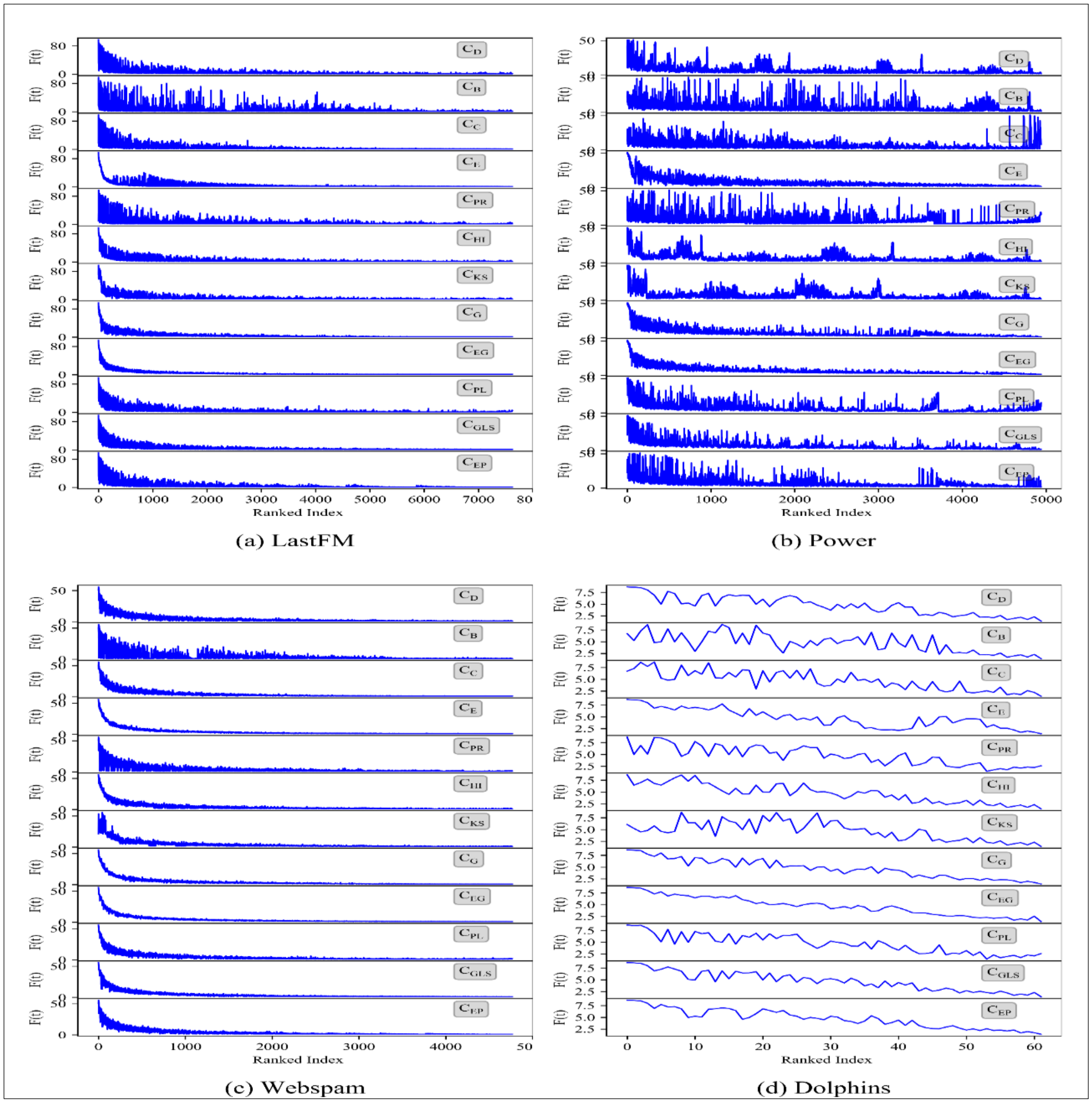

Fig. 6. Influence of propagation on the rankings using 12 measures.

TABLE VI. MONOTONICITY VALUES FOR 12 MEASURES

Network	C_D	C_B	C_C	C_E	C_{PR}	C_{HI}	C_{KS}	C_G	C_{EG}	C_{PL}	C_{GLS}	C_{EP}
LastFM	0.7978	0.8347	0.9997	**0.9999**	**0.9999**	0.7570	0.7348	0.9998	**0.9999**	0.9959	0.9975	**0.9999**
Power	0.5927	0.8319	0.9998	**0.9999**	**0.9999**	0.3930	0.2460	0.9979	**0.9997**	0.9624	0.9754	**0.9999**
Webspam	0.8640	0.8471	0.9996	**0.9998**	**0.9998**	0.8455	0.8374	0.9998	**0.9998**	0.9981	0.9988	**0.9998**
Dolphins	0.8312	0.9623	0.9737	**0.9979**	**0.9979**	0.6841	0.3769	0.9979	**0.9979**	0.9905	0.9958	**0.9979**

3) Efficiency

a) Experiment: Comparison of centrality measures using runtime.

In this experiment, we have analyzed the runtime of 12 measures in four networks to measure the efficiency of each centrality measure. As shown in Table VII, the run time of the C_D measure is significantly lower than that of the other measures across all networks. The C_{KS} and C_{HI} measures also have a low runtime after the C_D measure in all networks. Among all the measures, the runtime of the C_B measure is the highest.

The runtime trends in Fig. 7 indicate that the C_B measure rises sharply with network size *N*, unlike the other measures. Table VII emphasizes the three best-performing measures in bold.

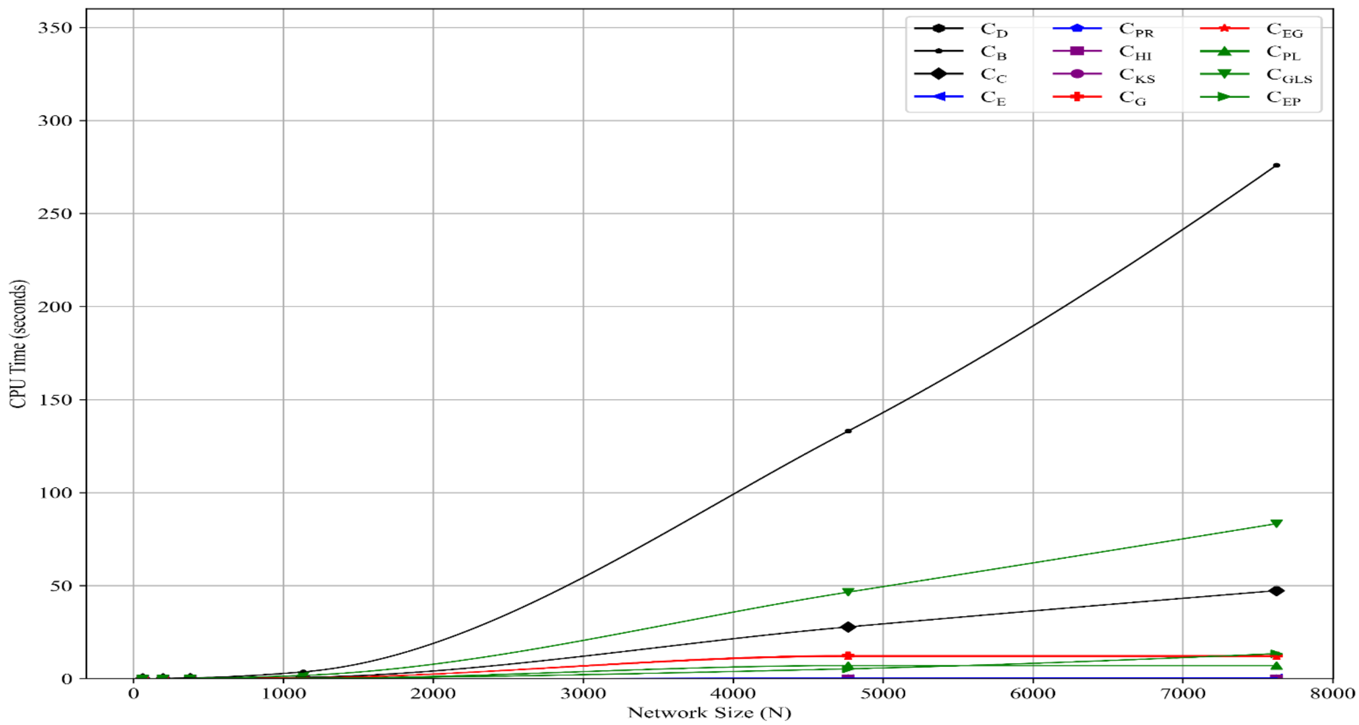

Fig. 7. The runtime (in seconds) of the 12 measures as a function of the Network size N.

TABLE VII. RUNTIME (IN SECONDS) OF 12 MEASURES

Network	C_D	C_B	C_C	C_E	C_{PR}	C_{HI}	C_{KS}	C_G	C_{EG}	C_{PL}	C_{GLS}	C_{EP}
LastFM	**0.0036**	350.6235	57.4717	0.6420	0.0914	**0.0726**	**0.0481**	4.7394	4.8089	1.5479	93.8552	21.0344
Power	**0.0035**	91.9589	11.3889	0.1699	0.0189	**0.0107**	**0.0108**	0.1051	0.1125	0.0870	24.3106	5.5635
Webspam	**0.0035**	186.5113	38.4280	0.2338	0.2421	**0.0457**	**0.0402**	14.7700	15.1397	8.9812	51.2695	6.6649
Dolphins	**0.00004**	0.0055	0.0014	0.0040	0.0018	**0.0002**	**0.0002**	0.0017	0.0017	0.0027	0.0093	0.0006

D. Discussion

Table VIII presents the findings of our research work, which indicates that the C_{EG} and C_G measures perform better than the other measures in every experiment except the run-time. Gravity-based measures have higher accuracy as compared to other measures because they are hybrid measures that provide both neighborhood and path information [34]. In most of the experiments, the C_{EG} measure attains the first position, and the C_G measure is at second because the C_{EG} measure includes a broader set of neighboring nodes as compared to the C_G measure. Table VIII depicts the three best-performing measures that outperform other measures.

TABLE VIII. COMPREHENSIVE COMPARATIVE EVALUATION TABLE

Network	Kendall's tau (τ)	F(t) vs Ranked Index	Monotonicity (M)	Runtime
LastFM	C_{EG}, C_G, C_{GLS}	C_{EG}, C_G, C_{KS}	(C_{EP}, $C_{PR,}$ C_{EG} C_E) *, C_G, C_C	C_D, C_{KS}, C_{HI}
Power	C_{EG}, C_G, C_E	C_{EG}, C_G, C_{GLS}	(C_{EP}, $C_{E,}$ C_{PR}) *, C_C, C_{EG}	C_D, C_{HI}, C_{KS}
Webspam	C_{EG}, C_E, C_{GLS}	C_{EG}, C_E, C_G	(C_{EP}, C_E) *, (C_{EG}, C_{PR}) *, C_G	C_D, C_{KS}, C_{HI}
Dolphins	C_{EG}, C_G, C_{EP}	C_{EG}, C_{EP}, C_G	(C_E, C_{PR}, C_G, C_{EG}, C_{EP}) *, C_{GLS}, C_{PL}	C_D, C_{HI}, C_{KS}

*Point out those measures that have gained the identical position.

V. CONCLUSION

This research has evaluated several centralities to pinpoint the significant nodes across diverse social networks. We measured the accuracy of several centralities by using the Kendall tau and SIR model among different social networks. In addition, we applied monotonicity to determine the differentiation ability of several centrality measures. It has been shown through empirical comparisons that the accuracy of both gravity-based measures, extended gravity centrality and gravity centrality, is superior to that of other centrality measures. Extended gravity measure emerges as the best performer, with the gravity measure ranking second. Furthermore, these gravity-based measures exhibit superior performance in terms of differentiation capability. However, both still require refinement to achieve higher efficiency in terms of runtime.

VI. Future Work

To increase the efficiency of the gravity-based measures, we will refine these measures in the future. This research work is limited to undirected unweighted networks; it can be extended to directed weighted networks to get deeper insights into the effectiveness of several measures.

Furthermore, the scalability and robustness of these measures can be assessed in dense networks. In order to improve the comprehensiveness of our analysis, we will also intend to incorporate new measures that have been built throughout this research.

References

[1] A. Zareie and A. Sheikhahmadi, "A hierarchical approach for influential node ranking in complex social networks," Expert Syst. Appl., vol. 93, pp. 200–211, 2018, doi: 10.1016/j.eswa.2017.10.018.

[2] A. Ullah, B. Wang, J. F. Sheng, and N. Khan, "Escape velocity centrality: escape influence-based key nodes identification in complex networks," Appl. Intell., vol. 52, no. 14, pp. 16586–16604, 2022, doi: 10.1007/s10489-022-03262-4.

[3] A. Ullah, B. Wang, J. Sheng, J. Long, and N. Khan, "Identification of Influential Nodes via Effective Distance-based Centrality Mechanism in Complex Networks," Complexity, vol. 2021, 2021, doi: 10.1155/2021/8403738.

[4] K. Das, S. Samanta, and M. Pal, "Study on centrality measures in social networks: a survey," Soc. Netw. Anal. Min., vol. 8, no. 1, 2018, doi: 10.1007/s13278-018-0493-2.

[5] S. P. Borgatti, "Centrality and network flow," Soc. Netw., vol. 27, no. 1, pp. 55–71, 2005, doi: 10.1016/j.socnet.2004.11.008.

[6] R. Grassi, S. Stefani, and A. Torriero, "Some new results on the eigenvector centrality," J. Math. Sociol., vol. 31, no. 3, pp. 237–248, 2007, doi: 10.1080/00222500701373251.

[7] L. C. Freeman, "Centrality in social networks conceptual clarification," Soc. Networks, vol. 1, no. 3, pp. 215–239, 1978.

[8] M. Kitsak et al., "Identification of influential spreaders in complex networks," Nat. Phys., vol. 6, no. 11, pp. 888–893, 2010.

[9] L. L. Ma, C. Ma, H. F. Zhang, and B. H. Wang, "Identifying influential spreaders in complex networks based on gravity formula," Phys. A Stat. Mech. its Appl., vol. 451, pp. 205–212, Jun. 2016, doi: 10.1016/j.physa.2015.12.162.

[10] K. Berahmand, A. Bouyer, and N. Samadi, "A new centrality measure based on the negative and positive effects of clustering coefficient for identifying influential spreaders in complex networks," Chaos, Solitons and Fractals, vol. 110, pp. 41–54, 2018, doi: 10.1016/j.chaos.2018.03.014.

[11] Z. Yu, J. Shao, Q. Yang, and Z. Sun, "ProfitLeader: Identifying leaders in networks with profit capacity," World Wide Web, vol. 22, pp. 533–553, 2019, doi: 10.1007/s11280-018-0537-6.

[12] A. Korn, A. Schubert, and A. Telcs, "Lobby index in networks," Physica A Stat. Mech. Its Appl., vol. 388, no. 11, pp. 2221–2226, 2009, doi: 10.1016/j.physa.2009.02.002.

[13] D. Chen, L. Lü, M. S. Shang, Y. C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," Phys. A Stat. Mech. Its Appl., vol. 391, no. 4, pp. 1777–1787, 2012, doi: 10.1016/j.physa.2011.09.017.

[14] Z. Li, T. Ren, X. Ma, S. Liu, Y. Zhang, and T. Zhou, "Identifying influential spreaders by gravity model," Sci. Rep., vol. 9, 2019, doi: 10.1038/s41598-019-44930-9.

[15] M. Bianchini, M. Gori, and F. Scarselli, "Inside PageRank," ACM Trans. Internet Technol., vol. 5, no. 1, pp. 92–128, 2005, doi: 10.1145/1052934.1052938.

[16] J. Sheng et al., "Identifying influential nodes in complex networks based on global and local structure," Phys. A Stat. Mech. its Appl., vol. 541, Mar. 2020, doi: 10.1016/j.physa.2019.123262.

[17] A. Ullah, S. U. Din, N. Khan, C. B. Mawuli, and J. Shao, "Towards investigating influencers in complex social networks using electric potential concept from a centrality perspective," Inf. Fusion, vol. 109, p. 102439, Sep. 2024, doi: 10.1016/j.inffus.2024.102439.

[18] A. Bavelas, "Communication patterns in task‐oriented groups," J. Acoust. Soc. Am., vol. 22, no. 6, pp. 725‐730, 1950.

[19] K. E. Joyce, P. J. Laurienti, J. H. Burdette, and S. Hayasaka, "A new measure of centrality for brain networks," PLoS One, vol. 5, no. 8, p. e12200, 2010.

[20] J. Bae and S. Kim, "Identifying and ranking influential spreaders in complex networks by neighborhood coreness," Phys. A Stat. Mech. its Appl., vol. 395, pp. 549–559, Feb. 2014, doi: 10.1016/j.physa.2013.10.047.

[21] A. Namtirtha, A. Dutta, and B. Dutta, "Identifying influential spreaders in complex networks based on kshell hybrid method," Phys. A Stat. Mech. its Appl., vol. 499, pp. 310–324, Jun. 2018, doi: 10.1016/j.physa.2018.02.016.

[22] J. Dai et al., "Identifying Influential Nodes in Complex Networks Based on Local Neighbor Contribution," IEEE Access, vol. 7, pp. 131719–131731, 2019, doi: 10.1109/ACCESS.2019.2939804.

[23] Q. Shang, Y. Deng, and K. H. Cheong, "Identifying influential nodes in complex networks: Effective distance gravity model," Inf. Sci. (Ny)., vol. 577, pp. 162–179, Oct. 2021, doi: 10.1016/j.ins.2021.01.053.

[24] J. Liu and J. Zheng, "Identifying important nodes in complex networks based on extended degree and E-shell hierarchy decomposition," Sci. Rep., vol. 13, no. 1, pp. 1–10, 2023, doi: 10.1038/s41598-023-30308-5.

[25] H. Singh and H. Singh, "An enhanced gravity model for determining crucial nodes in social networks based on degree k-shell eigenvector index," IEEE Access, vol. 12, pp. 23163–23180, 2024, doi: 10.1109/ACCESS.2024.3363635.

[26] A. Ullah, Y. Meng, and J. F. F. Mendes, "OVED-Rank: A ranking scheme to evaluate complex network spreaders' influence through the concept of effective distance and orbital velocity," Inf. Process. Manag., vol. 62, no. 5, p. 104201, Sep. 2025, doi: 10.1016/j.ipm.2025.104201.

[27] L. Qiu, J. Zhang, and X. Tian, "Ranking influential nodes in complex networks based on local and global structures," Appl. Intell., vol. 51, pp. 4394–4407, 2021, doi: 10.1007/s10489-020-02132-1.

[28] J. E. Hirsch, "An index to quantify an individual's scientific research output," Proc. Natl. Acad. Sci., vol. 102, no. 46, pp. 16569–16572, 2005.

[29] R. Xu and Y. Du, "A delayed SIR epidemic model with saturation incidence and a constant infectious period," J. Appl. Math. Comput., vol. 35, pp. 229–250, 2011, doi: 10.1007/s12190-009-0353-3.

[30] Y. Gang, Z. Tao, W. Jie, F. Zhong-Qian, and W. Bing-Hong, "Epidemic spread in weighted scale-free networks," Chinese Phys. Lett., vol. 22, no. 2, p. 510, 2005.

[31] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, "Epidemic outbreaks in complex heterogeneous networks," Eur. Phys. J. B-Condensed Matter Complex Syst., vol. 26, no. 4, pp. 521–529, 2002.

[32] D. G. Bonett and T. A. Wright, "Sample size requirements for estimating pearson, kendall and spearman correlations," Psychometrika, vol. 65, pp. 23–28, 2000, doi: 10.1007/BF02294183.

[33] D. Christensen, "Fast algorithms for the calculation of Kendall's τ," Comput. Stat., vol. 20, pp. 51–62, 2005, doi: 10.1007/BF02736122.

[34] H. Singh and H. Singh, "Analysis of different measures of centrality to identify vital nodes in social networks," in Advanced Network Technologies and Intelligent Computing, vol. 2090, A. Verma, P. Verma, K. K. Pattanaik, S. K. Dhurandher, and I. Woungang, Eds., Commun. Comput. Inf. Sci., 2024, pp. 101–115, doi: 10.1007/978-3-031-64076-6_8.