IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 March 2024, accepted 29 March 2024, date of publication 2 April 2024, date of current version 9 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3384242

RESEARCH ARTICLE

Convex Isolating Clustering Centrality to Discover the Influential Nodes in Large Scale Networks

BURAN BASHA MOHAMMAD[1], **V SATEESHKRISHNA DHULI**[1], **MURALI KRISHNA ENDURI**[2], **(Member, IEEE), AND LINGA REDDY CENKERAMADDI**[3], **(Senior Member, IEEE)**

[1]Department of Electronics and Communications Engineering, SRM University AP, Amaravathi 522502, India
[2]Algorithms and Complexity Theory Laboratory, Department of Computer Science and Engineering, SRM University AP, Guntur 522502, India
[3]Department of Information and Communication Technology, University of Agder (UiA), 4879 Grimstad, Norway

Corresponding author: Linga Reddy Cenkeramaddi (linga.cenkeramaddi@uia.no)

This work was supported in part by the Indo-Norwegian Collaboration in Autonomous Cyber-Physical Systems (INCAPS) Project through the International Partnerships for Excellent Education, Research and Innovation (INTPART) Program under Grant 287918; and in part by the Low-Altitude Unmanned Aerial Vehicle (UAV) Communication and Tracking (LUCAT) Project through the Information and Communications Technology and Digital Innovation (IKTPLUSS) Program from the Research Council of Norway under Grant 280835.

ABSTRACT Ranking influential nodes within complex networks offers invaluable insights into a wide array of phenomena ranging from disease management to information dissemination and optimal routing in real-time networking applications. Centrality measures, which quantify the importance of nodes based on network properties and relationships of nodes within the network, are instrumental in achieving this task. These measures are typically classified into local and global centralities. Global measures consider the overall structure and connectivity patterns. However, they often suffer from high computational complexity in large-scale networks. On the other hand, local measures focus on the immediate neighborhood of each node, potentially overlooking global information. To address these challenges, we propose a novel metric called Isolating Clustering Centrality (ISCL), which leverages a convex combination approach. By introducing a convex tuning parameter, ISCL enhances the applicability and adaptability of centrality measures across a wide range of real-world network applications. In this study, we assess the efficacy of the proposed measure using real-world network datasets and simulate the spreading process using susceptible-infected-removed (SIR) and independent cascade (IC) models. Our extensive results demonstrate that ISCL significantly improves spreading efficiency compared to conventional and recent centrality measures, while also maintaining better computational efficiency in large-scale complex networks.

INDEX TERMS Isolating centrality, clustering coefficient, convex tuning parameter, influential nodes, complex networks.

I. INTRODUCTION

Complex networks comprised of interconnected nodes, where the relationships between these nodes are not simple and may exhibit diverse structural patterns [1], [2]. Factors such as dynamic environments, evolving relationships, and unexpected events contribute to the complexity and uncertainty inherent in complex networks [1], [3]. These networks can encompass various systems such as social networks, transportation networks, and biological networks [4].

The associate editor coordinating the review of this manuscript and approving it for publication was P. Venkata Krishna.

Exploring these networks allows researchers to predict the behaviors of entities and identify patterns of interaction within the systems. One of the fundamental aspects of analyzing complex networks involves identifying and ranking influential nodes [5], [6]. Influential nodes are those crucial entities within the network that exhibit a significant impact on the overall network structure and dynamics [7], [8]. These nodes act as hubs that facilitate the fast flow of information or resources within the network. Studying influential nodes finds diverse applications in the various domains ranging from healthcare, including epidemic prevention [9] and disease research [10]. It can also be leveraged in commerce, such as e-commerce advertising success [11] and infrastructure resilience. Power grids and internet stability are promising research areas in utilizing these studies [12]. It also plays a pivotal role in understanding social dynamics, often referred to as social physics [13], addressing uncertainty [14], and dissemination efficiency [15]. Moreover, its influence extends to biological research, spanning from identifying drug targets [16] and protein identification [17], as well as impacting marketing strategies, particularly in viral campaigns [18], and mitigating communication breakdowns in both human and technological networks [19]. Centrality measures [20], [21], [22] quantify the prominence of nodes within a network based on various network properties. Conventional centrality metrics capture the various aspects of node importance, such as the number of connections (degree centrality) [23], the extent to which a node lies on the shortest paths between other nodes (betweenness centrality) [24], how quickly a node can reach other nodes in the network (closeness centrality) [25], the influence of a node based not only on its direct connections but also of its neighbors (eigenvector centrality) [26], degree to which nodes in a network tend to cluster together (clustering coefficient centrality) [27], and importance of a node based on the number and quality of its incoming links (PageRank centrality) [28].

Centrality measures developed in the literature are often based on different properties of nodes and their relationships in the network. Based on the different properties of nodes and their relationships within a network, these measures are classified as local and global centrality measures. Local measures [29] focus primarily on the immediate neighborhood of each node and their direct connections, whereas global measures [30] consider the dynamics of the whole network rather than focusing on only local neighborhoods. While both local and global centralities provide valuable insights into the structure and dynamics of networks, they have severe limitations to be addressed [31]. While local centrality measures suffer from accuracy due to their localized perspective, global measures face challenges related to computational efficiency in large-scale networks. Achieving a balance between accuracy and computational feasibility remains a key consideration in designing new centrality approaches. Designing hybrid centrality measures shows promise in overcoming the limitations of local and global approaches. A hybrid approach called local and global centrality is described in the paper [32] to rank the influential nodes in complex networks. By integrating the isolating and clustering coefficient centrality measures, authors proposed an Isolating Clustering Distance Centrality in [33]. In [34], authors propose a generalized measure based on degree, the shortest path between vertices, and global centrality measures and verify their efficiency on various real-world networks. A new indexing method named semi-global triangular centrality is proposed in [35] that maximizes the total collective influence of a spreading method by choosing the best information spreaders in the dense part of a network. A mathematical model is presented in [36] to develop the k-shell hybrid method based on network parameters. The Isolating centrality [37] attempts to identify crucial nodes at the interface between core and peripheral areas by leveraging both local and global network features.

Many centrality measures described in the literature face difficulty in effectively capturing local and global information while maintaining low time complexity. Those measures that prioritize the local information tend to offer good computational efficiency but may sacrifice accuracy. Conversely, global methods can enhance the spreading efficiency but often come with significant time complexity challenges, specifically in large-scale networks. To design an effective centrality measure, it is necessary to consider both local and global structural information with a reasonable time complexity. With this motivation, we develop the convex combination-based centrality measure by integrating the isolating centrality and clustering coefficient. The clustering coefficient focuses on the local density of the network. A high value of the clustering coefficient can accelerate information spread locally (within immediate neighbors), but decrease the information spread globally. Isolating centrality identifies crucial nodes at the interface between core and peripheral areas by leveraging both local and global network features. By integrating these metrics, we are able to leverage both local and global information effectively. Further, a convex tuning parameter (α) is introduced to enhance the applicability and adaptability to a wide range of real-world networks. We extensively conducted the simulation experiments and compared the proposed measure with both conventional and recent measures using SIR and IC models on four real-world datasets. Figure. 1 shows a block diagram that reflects the process of identifying influential nodes with ISCL centrality in complex networks.

II. RELATED WORK

This section discusses the literature that focuses on centrality measures in complex networks. Early studies [38] focused primarily on identifying and removing highly central nodes to break the network into independent components. A study in [39] found that recalculating degree and betweenness centralities is more effective in partitioning complex networks than simply removing nodes based on their initial centrality values. In [32], a centrality method is proposed that considers the local and global topological characteristics of a network. The study in [40] focused on IEEE test systems and modified centrality metrics (degree, proximity, and betweenness) for electrical topology. The connection between cascade failures and centrality in complex networks has been studied in [41]. Motivated by the design of degree centrality, they discovered that nodes with high degree centrality had shallow cascades whereas those with high betweenness centrality had deeper ones. Using real pairwise distances between nodes, research work in [42] studied the distance-based critical node problem. They compared the different centrality metrics with the best solutions using linear integer programming. A similar

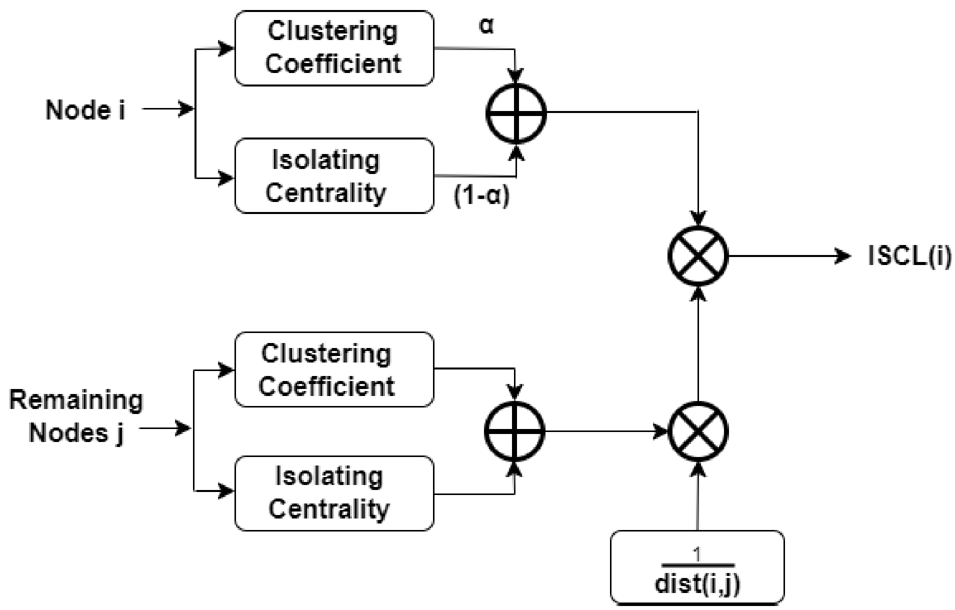

FIGURE 1. A block diagram for identifying influential nodes in complex networks with ISCL centrality.

problem has been studied in [43] with incomplete data containing false information. The isolating centrality [37] is a novel centrality measure that reveals the important nodes for the connectedness of a network. This measure has demonstrated the ability to uncover results that are up to four times more effective than traditional methods. However, it overlooks the nodes positioned within the inner nodes of the network. Since identifying influential nodes is a hard problem, approximate methods have been employed in [44]. To assess the effect of nodes, paper [45] provides a new metric that combines the degree and the average neighbor degree. This approach addresses incompleteness in view of the overall network, often caused by individual node features. In [46], a local degree dimension (LDD) is proposed based on consideration of the influence of neighbor nodes in each layer to determine nodes' importance. The method IDME (Information Diffusion and Matthew Effect aggregation) [47] evaluates node importance by obtaining information from both its own and its multilayer neighbor's information. A K-shell and SH-based centrality (SHKS) [48] considers the node's own and first and second-order neighbor's influences for the node's importance calculation. Global structure model [49], highlights the global influence of the node in the network while also taking self-influence into account, for the identification of influential nodes. A measure known as node propagation entropy [50] combines nodes' clustering coefficients and their first and second-order neighbor's influence from an entropy perspective. Based on structural data from neighbors and nearby neighbors, a metric, nearest neighborhood trust page rank (NTPR) [51] is developed. It includes degree ratio, trust value of neighbors, nearest neighbors, and similarity between nodes. Semi-global triangular centrality [35] indexing technique focuses on the most effective spreaders from the densely populated area of a network. This strategy maximizes the overall collective influence of a spreading process. To identify structural gaps, a semi-local and free parameter centralization metric [27] combines the degree, negative effects, and positive effects of a node's clustering coefficient with that of its second-level neighbors. Deep semi-supervised community detection (DSSC) [12], is a community detection technique that leverages an autoencoder-based deep clustering method to improve it. It is based on deep clustering techniques. In the paper [34], a generalized mixed centrality is designed based on local and global structural information to access nodes' importance. From many of the centrality metrics designed recently, some have gained popularity, including degree, betweenness, closeness, and K-shell centrality, among others. Broadly, these measures can be categorized into three types [52]: local, semi-local, and global. Local measures rely solely on information from immediate neighbors to assess a node's significance, resulting in lower accuracy and high computational efficiency. Conversely, global measures require access to the entire graph's information, leading to higher accuracy but increased computational complexity. Semi-local measures are a new class of measures that have evolved in recent years. These measures offer a high level of accuracy by leveraging more comprehensive information compared to local measures while maintaining nearly linear time complexity. When dealing with large-scale networks, it is essential to devise centrality measures that offer rapid dissemination capabilities with reasonable time complexity. Motivated by these challenges, this work develops a convex combination-based semi-local strategy to discover influential nodes exploiting the intrinsic properties of complex networks. We integrate the clustering coefficient and isolating centrality measures to capture the structural features in complex networks that can be examined at both micro (individual node) and macro (entire graph) levels [53]. Further, we study the influence of convex tuning parameter α on spreading ability. We conduct extensive simulations to verify the efficacy of the proposed measure on real-world networks such as *bio-dmela*, *facebook-combined*, *soc-anybeat*, and *fb-pages-tvshow* using SIR and IC models. The major contributions of our work include:

- We propose an optimal centrality measure, ISCL using the convex combination method.
- Utilizing the SIR model on real-world networks, we demonstrate the superior spreading control capabilities of the proposed measure compared to conventional and recent methods. To further validate the distinctiveness of our approach, we conducted a similarity analysis using Kendall's correlation coefficient with existing measures.
- Finally, to assess the computational efficiency of our proposed measure, we compared its time complexity with conventional and recent methods in the literature.

A. PAPER ORGANISATION

The structure of this article operates as follows: A brief review of centrality measures is defined in section III. Section IV describes the proposed centrality measure and the algorithm. The network datasets and spreading methods used to demonstrate the proposed measure of efficiency are presented in section V. Section VI explained the equipment and software tools used for our simulation. Experimental results are given in section VII. A discussion of the results is mentioned in section VIII. Finally, important observations and findings are presented in the conclusion in section IX.

B. NOTATION

We have given a comprehensive list of symbols and notations in Table 1.

TABLE 1. Notation of symbols.

Symbol	Description
G	Graph
V	Node set
E	Edge set
n	Number of nodes
D_i	Number of immediate neighbors of a node i
$dist(i, j)$	Shortest path distance between nodes i and j
$\sigma_{jk}(i)$	Number of shortest paths between nodes j and k that traverse through node i
σ_{jk}	Number of shortest paths between nodes j and k
$e_{i,j}$	Represents an edge between node i and node j.
$N(i)$	Neighbors of node i
L_i	Number of links exists among the neighbors of node i
$CL(i)$	The clustering coefficient of node i
$IS(i)$	Isolating centrality of node i
A	Adjacency matrix
G_δ	The minimum degree of the network
τ	Kendall's correlation coefficient
n_c	The number of concordant pairs
n_d	The number of discordant pairs
β	Infection probability
γ	Recovery rate

III. REVIEW OF CENTRALITY MEASURES

We discuss the various conventional and recent centrality measures in this section. Primarily, the clustering coefficient (CL), closeness centrality (CC), degree centrality (DC), betweenness centrality (BC), isolating centrality (IS), and local and global centrality (LGC), for an undirected, unweighted network represented as $G(V, E, A)$ with n nodes ($|V|$), m edges ($|E|$), and the adjacency matrix A is used to indicate the network connections. The matrix elements can have one of two values: 1, which indicates that node i and node j are connected, or 0, which indicates that there is no link between them. Degree centrality relies entirely on information from immediate neighbors and is known for its limited accuracy. Degree centrality [54], [55] is described as

$$DC(i) = \frac{\sum_{j \epsilon N(i)} e_{i,j}}{n-1} \tag{1}$$

where j describes the neighbors of node i, $e_{i,j} = 1$ if an edge exist between node i and node j, otherwise 0 and n denotes the number of nodes.

The two most important global centrality metrics are betweenness centrality [56] and closeness centrality [25]. They are not appropriate for massive networks, since they require knowledge of the entire network. According to [57], closeness centrality is based on the notion that nodes that are closely connected in terms of distance to remaining nodes are more relevant than those that are at a larger distance. By taking into account the shortest routes between the node in question and every other node in the network, it determines the significance of each node. It is determined as:

$$CC(i) = \frac{1}{\sum_{j \in G} dist(i, j)} \tag{2}$$

where $dist(i, j)$ denotes the shortest route between nodes i and j.

Betweenness centrality [58], denoted as BC, quantifies a node's ability to regulate information flow along the network's shortest pathways.

$$BC(i) = \sum_{i \neq j \neq k \in G} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{3}$$

where σ_{jk} indicates the shortest route between the nodes j and k. Here, $\sigma_{jk}(i)$ denotes the shortest route between the nodes j and k that passes via node i.

The clustering coefficient centrality [59], quantifies a node's degree of integration inside the network's densely packed clusters. It is calculated as the proportion of edges between a node's nearest neighbors to all possible edges.

$$CL(i) = \frac{2 * L_i}{D_i * (D_i - 1)} \tag{4}$$

where D_i is the degree of a node i and L_i is the total number of real connections that exist among node i's neighbors. To prevent counting edges twice, a factor of 2 is added.

Two mixed centrality measures such as Isolating centrality [37] and local and global centrality [32] in certain cases can provide a comprehensive understanding of network dynamics. Isolating centrality considers a node's degree and the minimum degree of its immediate neighbors, highlighting its relative importance amidst its immediate surroundings.
Integrating a node's degree with its isolation factor defines the isolating centrality, a measure of its ability to disconnect from the entire network. The isolating centrality of a node can be expressed as

$$IS(i) = |N_i \cap G_\delta| \times D_i \tag{5}$$

where N_i denotes the count of the node $i's$ neighbours, while G_δ depicts the minimum degree of the network, and D_i displays the degree of node i. LGC represents the significance of the global structure and optimizes the efficient and effective detection of prominent nodes by studying the node's local and global structure within the network. Integration of local and global structure effects leads to the design of LGC.

It can be expressed as:

$$LGC(i) = \frac{D_i}{n} \times \sum_{i \neq j} \frac{\sqrt{D_j}}{dist(i, j)} \tag{6}$$

where D_i and D_j represents degree of node i and node j respectively. The shortest path distance from node i to node j is expressed as $dist(i, j)$, and n is the total number of nodes. The expression's first part indicates local effects, while the second indicates global effects.

IV. PROPOSED CENTRALITY MEASURE

A novel approach called local and global centrality is described in the paper [32] to highlight important nodes in the network. LGC is derived based on local and global structural information of the network. To enhance accuracy, it calculates a node's local importance using its degree centrality [60]. Degree centrality counts the number of immediate connections to a node but neglects connections among its neighbors.

While the clustering coefficient [61] measure quantifies the interconnectedness of a node's neighbors. As a result, it describes an efficient representation of the network's local structure by quantifying the degree of connectivity among neighbor nodes. Alternatively, degree centrality provides a node's entire connectivity via its immediate connections.
By identifying densely connected regions and accessing the immediate connections of a node, the clustering coefficient has a positive impact on information dissemination and identification of community structure. Whereas the intent of isolating centrality is to identify nodes that significantly affect a network's connectivity. Based on local and global information, isolating centrality calculates node values. For node i the local influence (LI) is defined as:

$$LI(i) = [\alpha CL(i) + (1 - \alpha)\tfrac{IS(i)}{n}] \tag{7}$$

Likewise, for node i the global influence (GI) is defined as:

$$GI(i) = \sum_{i \neq j} \frac{CL(j) + \frac{IS(j)}{n}}{dist(i,j)} \tag{8}$$

More specifically for node i, the isolating-clustering coefficient centrality (ISCL) is formulated as follows:

$$ISCL(i) = [\alpha CL(i) + (1 - \alpha)\tfrac{IS(i)}{n}] \times \sum_{i \neq j} \frac{CL(j) + \frac{IS(j)}{n}}{dist(i,j)} \tag{9}$$

where α is a convex combination parameter, CL(i), CL(j) and IS(i), IS(j) correspond to clustering coefficient and isolating centralities of node i and node j respectively. n being the total number of nodes, $dist(i, j)$ is the shortest path distance between nodes i and j.

Algorithm 1 Algorithm to Determine a Centrality Measure ISCL for a Given Network G

Input: Network $G = (N, L)$
Output: For each node in network G centrality measure ISCL
begin
  $N = Nodelist, L = Edgelist, n = numberofnodes$
  for *all vertices i in N* **do**
    sum=0
    Find the clustering coefficient $CL(i)$ and isolating centrality $IS(i)/n$ of node i
    for *all vertices j in N with* $j \neq i$ **do**
      Find the distance $dist(j, i)$ between (j, i)
      Calculate the clustering coefficient $CL(j)$ and the isolating centrality $IS(j)/n$ of node j
      $sum = sum + \frac{CL(j)+\frac{IS(j)}{n}}{dist(j,i)}$
    $ISCL(i) = [\alpha CL(i) + (1 - \alpha)\frac{IS(i)}{n}]sum$ /*α is a convex combination parameter*/
  return ISCL /* centrality measure for all vertices*/
end

A. TIME COMPLEXITY

In this subsection, the time complexity of isolating the clustering centrality algorithm is determined. By integrating the clustering coefficient, isolating centrality, and the shortest path between the nodes, the isolating clustering centrality (ISCL) time complexity can be estimated. The time complexity of $O(n^2)$, $O(nm)$, and $O(n + m)$ correspond to the node's isolating centrality, clustering coefficient, and the shortest

TABLE 2. Estimation of all centralities' time complexity, where n and m are the numbers of vertices and edges, respectively.

Centrality	Time Complexity
BC	$O(nm)$
CC	$O(nm)$
CL	$O(nm)$
DC	$O(n+m)$
IS	$O(n^2)$
LGC	$O(n^2)$
ISCL	$O(nm)$

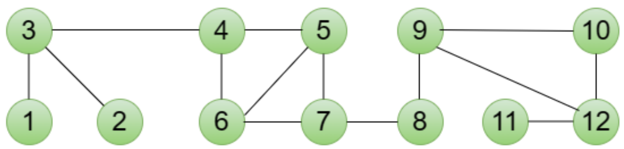

FIGURE 2. A 12-node toy network with 14 edges.

distance between two nodes respectively. Local influence has an effective time complexity of $O(n^2) + O(nm) = O(nm)$. The global influence time complexity of the summing can be expressed as $O(nm) + O(n^2) + O(n + m) = O(nm)$. As a result, we've seen that the time complexity of the Algorithm is $O(nm)$, which is the time required to determine the isolating clustering centrality of a network node. In Table 2, we have shown the time complexity of each centrality. For conventional centralities, the time complexity of ISCL is equivalent to BC, CC, and CL. In the case of the latest measures, the time complexity of ISCL is competitive with that of IS and LGC.

Figure. 2 displays a toy network consisting of 14 links and 12 nodes; we use node 12's effect to demonstrate ISCL.

The convex local influence LI, as well as global influence GI are shown in the Table 3 of the framework for node (12), are computed using ISCL, for $\alpha = 0.6$; $LI(12) = [0.6 * CL(12) + 0.4 * \frac{IS(12)}{12}] = 0.6 * 0.333 + 0.4 * \frac{3}{12} = 0.299$, and, $GI(12, 1) = \frac{CL(1)+\frac{IS(1)}{12}}{dist(12,1)} = \frac{[0+\frac{0}{12}]}{7} = 0$,

the global influence caused by individual nodes is, $GI(12, 1)+GI(12, 2)+GI(12, 3)+GI(12, 4)+GI(12, 5)+ GI(12, 6)+GI(12, 7)+ GI(12, 8) +GI(12, 9)+GI(12, 10) +GI(12, 11) = 1.897$.

Finally, we can determine the effect of node 12 using the algorithm, $ISCL(12) = 0.299 \times 1.897 = 0.56$. Similar calculations have been carried out for the remaining nodes and corresponding centrality values are provided in Table 4.

V. IMPLEMENTATION

In this section, first, we provide an overview of the real-world datasets used in our simulations. The next three key concepts we examine are the results of *SIR*, *IC* models, which simulate the spread of infectious diseases, and Kendall's correlation coefficient, which is used to assess outcomes. At each time step of the *SIR* and *IC* model simulations, we supply a number representing the likelihood of an infected node infecting a susceptible neighbor. It's important to remember that the parameter is essentially stochastic and doesn't have a perfect value, even though some β values in our research, spanning from 0.01 to 0.40, generate better results than others. The recovery rate, which is set to 1, is another parameter in the *SIR* simulation.

A. OVERVIEW OF NETWORK DATASETS

To study the efficiency of the ISCL, we used four popular real-world networks using SIR and IC models. We utilized the *"bio-dmela"*, *"soc-anybeat"*, *"facebook-combined"*, and *fb-pages-tvshow* network datasets to assess the effectiveness of isolating clustering centrality. Bio-dmela belongs to the class of sparse biological networks. Nevertheless, in bio-dmela, proteins serve as nodes, and the edges depict interactions between proteins. Soc-anybeat falls into the category of online social networks. soc-anybeat represents an online community, serving as a public platform where individuals can engage with others, whether they are from their local neighborhood or situated across the globe. The dataset comprises node features (profiles), circles, and ego networks. To ensure privacy, the Facebook data has undergone anonymization, wherein the Facebook-internal IDs for each user have been replaced with new values. Additionally, the feature vectors in this dataset have been presented, but the specific interpretation of these features has been obscured. For instance, if the original dataset included a feature like "political=Democratic Party," the new data now contains "political=anonymized feature 1". Hence, analyzing anonymized data allows us to determine if users are from the same political camp, further, the specific nature of their individual political beliefs remains unrevealed. The fb-pages-tvshow falls within the category of sparse social networks, and it comprises data collected in 2017 from Facebook pages. These datasets specifically depict a network of blue-verified Facebook pages across various categories. In this network, nodes correspond to individual pages, and the edges denote mutual likes between these pages. The dataset comprises node features (profiles), circles, and ego networks. These networks are accessible at this address: [62], [63]. Table 5 provides basic details regarding these data sets.

B. SPREADING MODELS

To examine the fast information spread through the network via the top-ranked nodes, we used the *SIR* and *IC* models.

1) SIR MODEL

Researchers regularly employ the *SIR* model, a widely recognized technique in the area, to investigate how information spreads via complex networks. Susceptibility (S), Infection (I), and Recovery (R) are the three states that the model identifies. The top ten nodes with the greatest centrality ratings [64] are the focus of our experimental setup. With a

TABLE 3. Demonstration of calculating isolating clustering centrality of node 12 (ISCL(12)), dist(12,*i*) and GI(12,*i*) depict shortest distance from node 12 to others and global influence caused by individual nodes with respect node 12.

node i	1	2	3	4	5	6	7	8	9	10	11	12
CL(i)	0	0	0	0.333	0.666	0.666	0.333	0	0.333	1	0	0.333
IS(i)	0	0	6	0	0	0	0	0	0	0	0	3
dist(12,i)	7	7	6	5	4	4	3	2	1	1	1	-
GI(12,i)	0	0	0.083	0.066	0.1666	0.1666	0.083	0	0.333	1	0	-

TABLE 4. The node centrality values of BC (Betweenness Centrality), CC (Closeness Centrality), DC (Degree Centrality), CL (Clustering Coefficient Centrality), LGC (Local and Global Centrality), IS (Isolating Centrality), ISCL (Isolating clustering Centrality) for Figure. 2 is provided in the following table. The red-colored values indicate top-3 influential nodes.

node	1	2	3	4	5	6	7	8	9	10	11	12
DC	1	1	3	3	3	3	3	2	3	2	1	3
BC	0	0	0.345	0.436	0.218	0.218	0.545	0.509	0.436	0	0	0.181
CC	0.229	0.229	0.289	0.343	0.379	0.379	0.392	0.366	0.323	0.261	0.215	0.268
CL	0	0	0	0.333	0.666	0.666	0.333	0	0.333	1	0	0.333
IS	0	0	6	0	0	0	0	0	0	0	0	3
LGC	0.533	0.533	2.083	2.415	2.471	2.471	2.471	1.554	2.269	1.289	0.518	2.015
ISCL	0	0	0.432	0.32	0.557	0.557	0.314	0	0.317	0.604	0	0.546

TABLE 5. Network characteristics of datasets. ACC (Average Clustering Coefficient), APL (Average Path Length), AP(Articulation Points), $|Deg_1|$ (Number of Nodes with degree 1),$|Deg_1|/n$ (Ratio of Number of Nodes having Degree 1 and the Number of Nodes) values of four different networks are displayed.

Network	Nodes	Edges	Max Degree	Avg. Degree	Diameter	ACC	APL	AP	$\|Deg_1\|$	$\|Deg_1\|$/n
fb-pages-tvshow	3892	17262	126	8.87	20	0.373	6.28	552	607	0.155
facebook-comined	4039	88234	1045	43.69	8	0.605	3.69	11	75	0.018
bio-dmela	7393	25569	190	6.917	11	0.011	4.34	1209	2005	0.271
soc-anybeat	12645	49132	4800	7.77	10	0.227	3.17	1257	6260	0.495

TABLE 6. Ranking correlation of the proposed centrality ISCL (Isolating Clustering Coefficient), basic centralities DC (Degree Centrality), BC (Betweenness Centrality), CL (Clustering coefficient), CC (Closeness Centrality), and latest centralities LGC (Local and Global Centrality), IS (Isolating Centrality).

Network	DC	BC	CL	CC	IS	LGC	ISCL
fb-pages-tvshow	0.075231	0.050631	0.111783	0.002139	0.788543	0.003499	0.034552
facebook-combined	0.019184	0.035027	0.022691	0.196366	0.993355	0.008355	0.70275
bio-dmela	0.262978	0.293619	0.540384	0.344921	0.607274	0.393919	0.275555
soc-anybeat	0.275899	0.358969	0.383848	0.029535	0.82156	0.014949	0.119386

transmission probability of β at each time step, an infected node tries to infect its neighbors in the SIR model. The affected node then has the likelihood to recover with a γ probability at each stage. The simulation is over when there are no more infected nodes in the network, which is when this procedure stops. We took into account infection rates ranging from 0.01 to 0.40 for our *SIR* model simulation. For a robust study, we ran this simulation process 100 times.

2) IC MODEL

The Independent Cascade Model (IC) model primarily focuses on information diffusion and influence propagation in social networks [65], [66]. In this model, nodes are typically in either active or inactive states. The propagation process involves the activation of nodes and the information spread through the network. In IC model [67], the activation probabilities of one node in the network do not influence the other nodes. The propagation probabilities (p_{ij}) of edges (E) in this model represent the possibility of impact from node i to node j. It concisely depicts the model using a collection of nodes (V), edges (E), and related propagation probabilities (p), denoted as $G_{ICM} = (V, E, p)$. An idle node j is activated with probability p_{ij} at each step ($t = 1, 2, 3 \ldots$) by its direct neighbor i, which is activated at $t - 1$. Every active node has a one-time influence on its direct neighbors. The eventual activation probability of an inactive node j is unaffected by the order of its many direct neighbors that are activated at $t - 1$. The final activation probability is shown by a measure of influence spread, which continues until no more nodes are accepted. This process follows a cascade pattern, where influence spreads from one node to another node.

C. KENDALL'S CORRELATION COEFFICIENT

In order to evaluate the efficacy of ISCL, we employ Kendall's [68] formula. Let's assume that two node sequences (P, Q) with equal numbers of nodes n are associated, where $P = (p_1, p_2, \cdots, p_n)$ and $Q = (q_1, q_2, \cdots, q_n)$. If both $p_i > p_j$ and $q_i > q_j$ or $p_i < p_j$ and $q_i < q_j$ have the same ordering of their components, then a pair of two node descriptions (p_i, q_i) and (p_j, q_j) for $(i \neq j)$ is said to be concordant. If $p_i > p_j$ and $q_i < q_j$, or $p_i < p_j$ and $q_i > q_j$, they are said to be discordant. If $p_i = p_j$ or $q_i = q_j$, the pair is neither concordant nor discordant. The Kendall's correlation coefficient is given as:

$$(P, Q) = 2 * \frac{(n_c - n_d)}{n(n-1)} \tag{10}$$

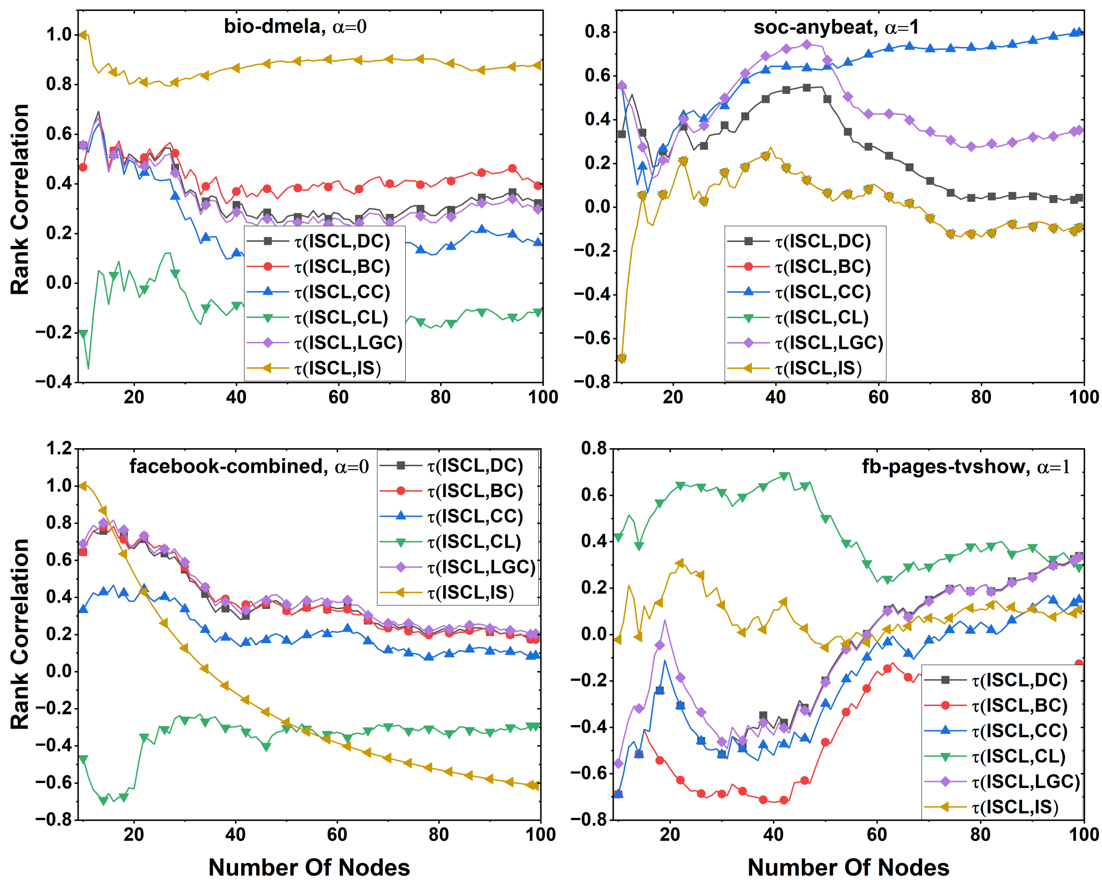

FIGURE 3. Demonstrates the correlation, up to the top- 100 nodes, between ISCL and conventional centralities such as DC, BC, CC, CL, IS, and LGC.

where the terms n_c and n_d, respectively, stand for the number of concordant and discordant pairings, and n is the node count of each sequence.

D. OBTAIN THE CONVEX TUNING PARAMETER α AND TRANSMISSION PROBABILITY β VALUES

We use a convex tuning variable α to examine the effects of the isolating centrality and clustering coefficient for each node. In this work, we investigate the impact of various α values on the total number of infected nodes as a function of the proposed centrality. We evaluate the best possible values for raising the spreading efficiency of the proposed measure by methodically increasing the α value in the range 0.0 to 1.0 in steps of 0.2 throughout our studies. We investigated transmission probability β over the range 0.01 to 0.4 in steps of 0.01. In our experiments, we analyzed the impact of β values on the total number of infected nodes as a function of proposed centrality. We identify the best possible values for elevating the spreading efficiency of the proposed measure. It is found that these values vary for different networks.

VI. EXPERIMENTAL SETUP

By using a 4.10GHz Intel(R) Core(TM) $i5 - 10600K$ CPU and 32GB of RAM as the central processing unit, extensive simulations are accomplished. Python version 3.11.2 is chosen because it provides built-in modules for making graphical models and computing node centralities, making it suitable for accelerating the simulation process. By utilizing NetworkX, a popular Python package for network analysis recognized for its broad set of features and tools developed to produce, analyze, and visualize graphs, we dealt with graphical representations of networks. We employed the OriginPro computer program, which was created by OriginLab Corporation and is a data analysis and graphing application. It is often used for tasks involving data visualization, analysis, and presentation in a range of fields, including scientific and technical research. Our research work is implemented using Python scripting language. The code is available at: https://github.com/BuranMohammad/Code-Link

VII. RESULTS

This section presents the simulation results to compare the isolating clustering centrality measure against the conventional and recent centrality measures. First, we show how the proposed centrality relates to the conventional centrality metrics. To assess the cumulative infected nodes for centrality

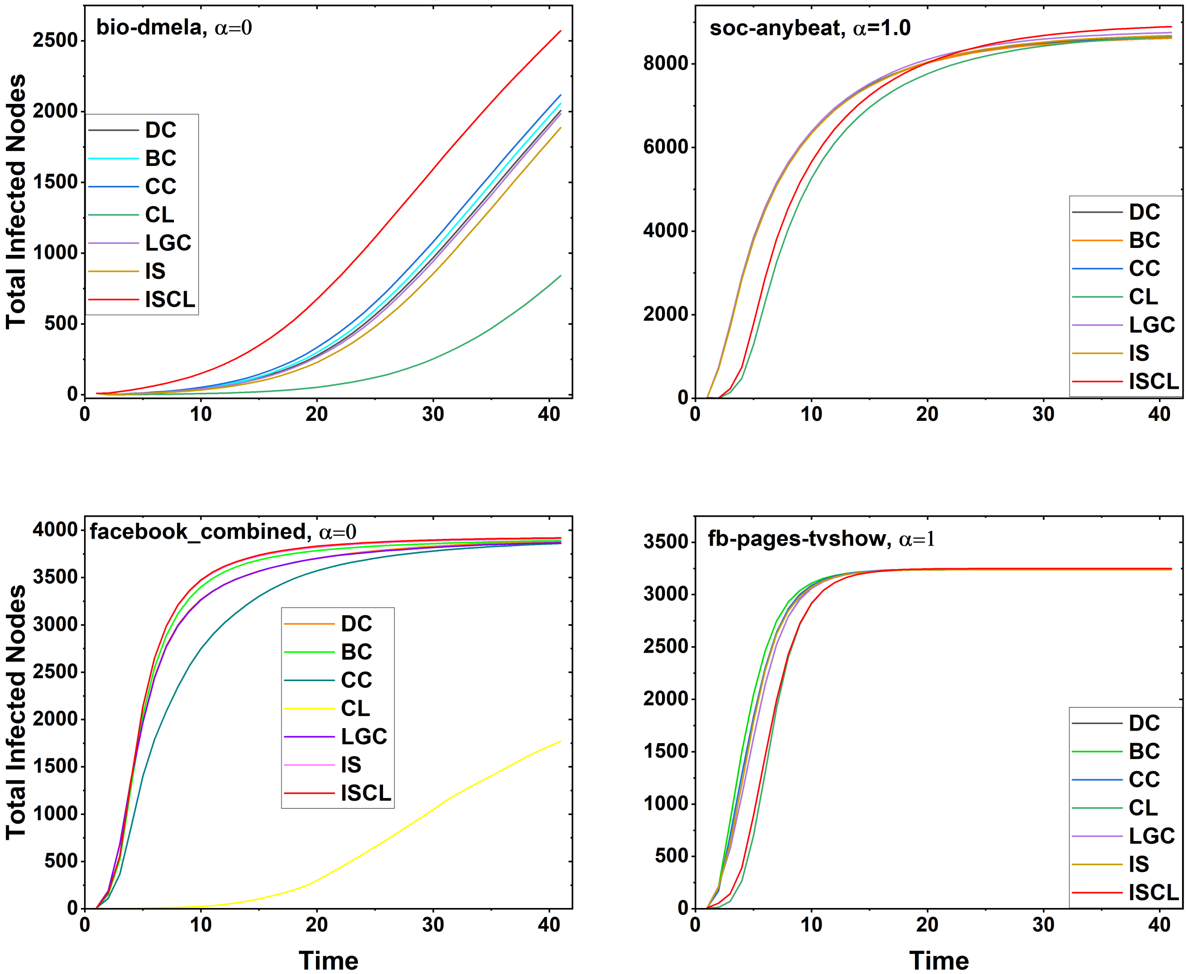

FIGURE 4. SIR model cumulative infected nodes for *bio-dmela*, *soc-anybeat*, *facebook-combined*, and *fb-pages-tvshow* networks (100 simulations in 40 time stamps). The top ten nodes are the best-infected nodes tested by basic centralities (DC,BC,CC,CL), recent centralities (IS, LGC), and the proposed ISCL centrality.

measures like DC, BC, CC, CL, IS, LGC, and ISCL, we use the SIR and IC models.

A. ISCL AND BASIC AS WELL AS RECENT CENTRALITY MEASURES IN CORRELATION

In this subsection, we provide the results of the correlations between ISCL and basic centrality measures in the literature. Kendall's correlation coefficient has been used to compare ISCL with the basic and recent centralities. We provide correlation graphs for the ISCL and other conventional centrality measures in the Figure. 3. However, the ISCL measure does not have a strong correlation with basic centralities in the *bio-dmela* but has a moderate correlation with recent measures isolating centrality. In the case of *soc-anybeat* network, initially, the ISCL is not strongly correlated with conventional centralities, later ISCL is moderately correlated only with closeness centrality. In the case of *facebook-combined* ISCL is slightly correlated with DC, BC, CC and LGC. Coming to *fb-pages-tvshow* network, ISCL is initially moderately correlated with the CL, later has no correlation with it, and ISCL has no correlation with other basic centralities. Table 6 depicts the correlation coefficient values of the ISCL centrality and conventional centrality measures. The proposed centrality is not closely correlated with any of the existing centralities. The remaining subsections explained the cumulative infected nodes and histogram's maximum influence on isolating clustering centrality.

B. ANALYZING SPREADABILITY WITH SIR MODEL

The total number of infected nodes over time is shown in this subsection to demonstrate the impact of information transmission after the first infection brought on by the top-10 seed nodes. The top-10 seed nodes are determined using the proposed isolating clustering centrality (ISCL) along with conventional centrality metrics such as DC, BC, CC, and CL, and the recent measures IS, and LGC. We found that the top-10 seed nodes are initially infected in the SIR model. In the subsequent time step, the surrounding nodes of such seed nodes acquire an infection with a probability of β. We consider that the likelihood of infection, β, is between 0.01 and 0.40. Every infected node ultimately has an opportunity to recover with a specific recovery rate,

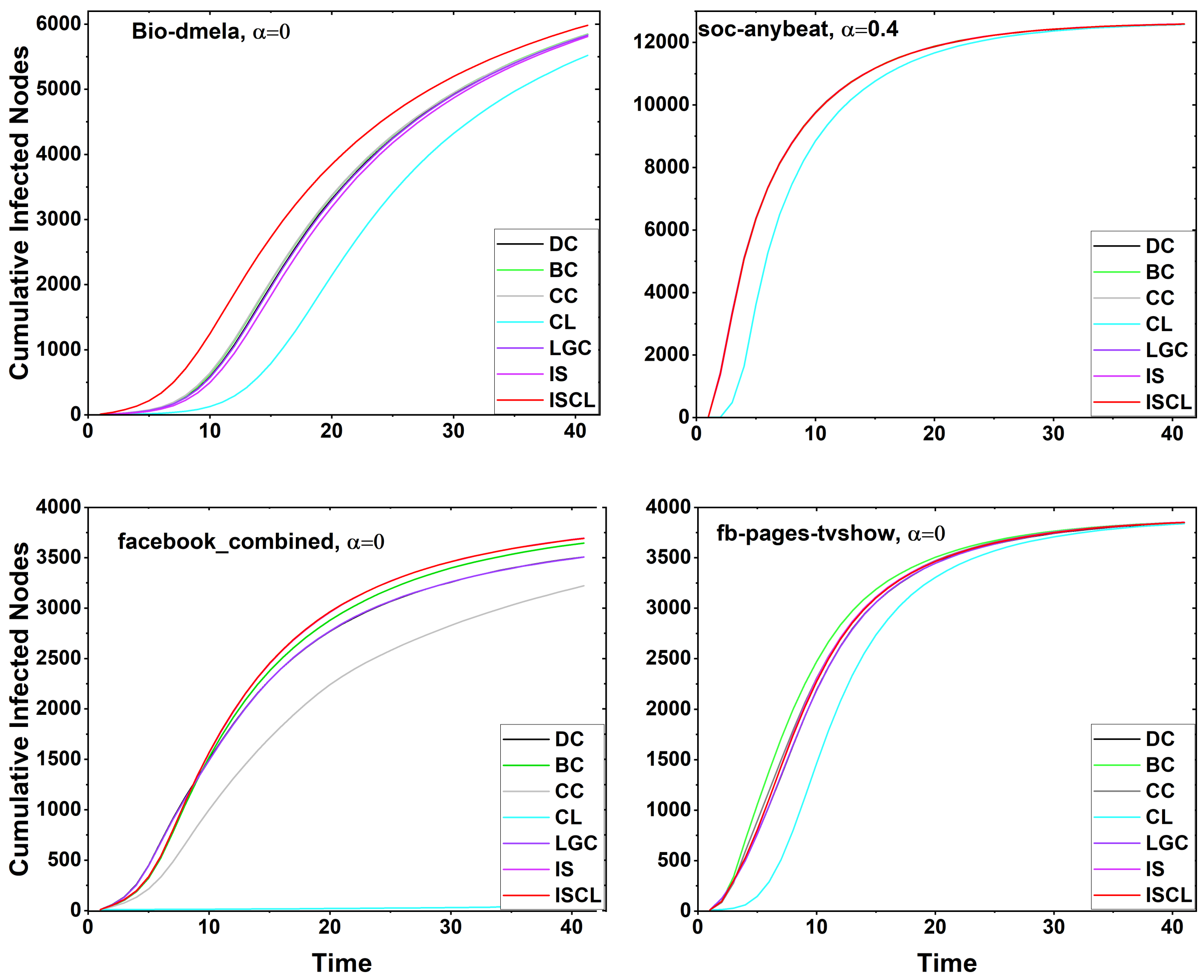

FIGURE 5. IC model cumulative infected nodes for *bio-dmela, soc-anybeat, facebook-combined*, and *fb-pages-tvshow* networks (40 time steps for 100 computer simulations). The ISCL, basic, and recent centralities are investigated by infecting their corresponding top-10 nodes.

or γ, which is taken into consideration. We have done 100 simulations with a 40 time step limit to calculate the cumulative infected nodes.

The SIR results for four real-world networks are shown in Figure. 4. Since the average clustering coefficient of *bio-dmela* is substantially lower and for parameter α is at zero, the proposed measure ISCL beat all conventional and recent centrality metrics from the very first time steps. When the tuning parameter is set to 1 for *soc-anybeat* and *fb-pages-tvshow*, the degree, betweenness, closeness, isolating, and LGC together initially function well up to certain intervals. Our centrality metric, ISCL, then raises to the top.

Further, we've observed that our centrality measure, ISCL, outperforms conventional and recent centralities and the latest in the *facebook-combined* dataset. The aforementioned explanation leads to the conclusion that ISCL performs better than all other conventional centralities and the latest measures. For each of the four datasets as shown in Figure. 6, the histograms of association values for all of the seven centrality metrics (DC, BC, CC, CL, LGC, IS, and ISCL) produced using the SIR model are displayed. For the *bio-dmela* network structure, our proposed metric ISCL can significantly outperform DC, BC, CC, CL, LGC, and IS in terms of the total infected nodes. The ISCL value outperforms other β values in terms of infected nodes, especially when compared to the transmission probability, which is β at 0.01. The proposed measure beat all centralities for β ranging from 0.01 to 0.05 in the Figure. 6.

In comparison to DC, BC, CC, CL, LGC, and IS for the *soc-anybeat* network architecture, our proposed measure, ISCL, could perform better in terms of the number of infected nodes with a convex tuning parameter α of 1. The ISCL value is superior to other β values in terms of infected nodes, especially when considering that the transmission probability of β is 0.06. In all 7 histograms of *soc-anybeat*, the proposed metric ISCL outperforms alternative centralities.

In comparison to DC, BC, CC, CL, LGC, and IS for the *facebook-combined* network structure, our proposed measure, ISCL, can perform well in terms of the number of infected nodes with convex tuning parameter α is 0. When it comes to infected nodes, ISCL is preferred to other β values, especially when it comes to transmission probability, which is

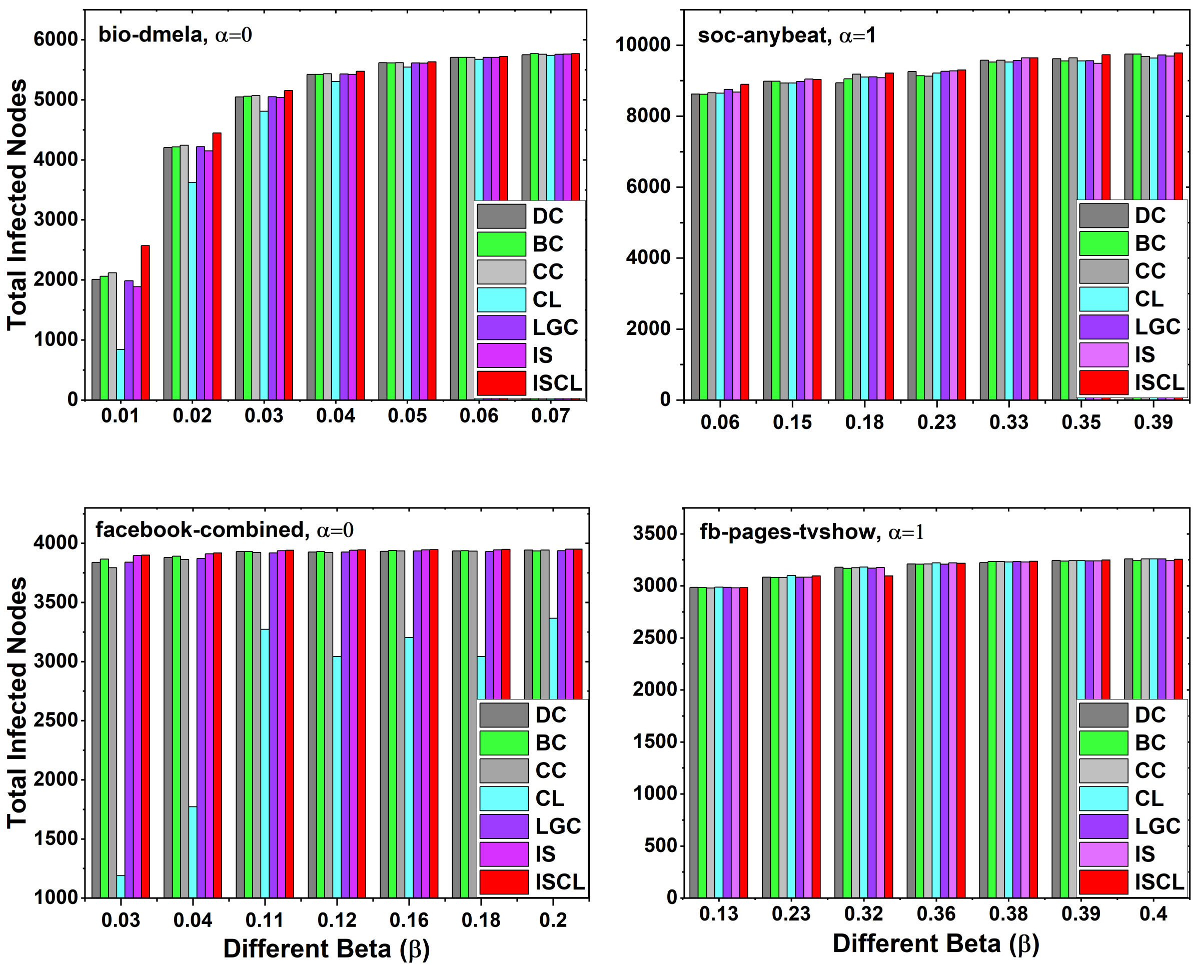

FIGURE 6. For four separate datasets, top-7 histograms of the total infected nodes using the *SIR* model simulated over DC, BC, CC, CL, LGC, IS, and ISCL centrality metrics.

β is 0.11, and ISCL is marginally better than other centralities in the top-7 histograms of the *facebook-combined* network.

With a convex tuning parameter of 1, our proposed measure, ISCL, can outperform DC, BC, CC, LGC, and IS in terms of the number of infected nodes for the *fb-pages-tvshow* network structure. When it comes to the histogram of infected nodes, ISCL is superior to other β values, particularly when the transmission probability is 0.23, and ISCL is on par with different centralities for other β values.

C. ANALYZING SPREADABILITY WITH IC MODEL

This subsection analyzes the impact of information transfer after the initial infection of top-10 seed nodes determined using isolating clustering centrality, basic centrality, and the latest centrality measures. It reveals that these nodes initially have the IC model, and their surrounding nodes acquire an infection with a probability of β. We performed 100 simulations with a 40-step time restriction to calculate the cumulative sum of infected nodes. The ICM results for four real-world networks are shown in Figure 5. In the case of *bio-dmela* network, for α is at zero, the proposed measure outperformed all centrality measures. For *soc-anybeat* with α is at 0.4, the proposed measure is equally competitive with degree, betweenness, closeness, isolating, and local and global centralities, and surpassed clustering coefficient. For *facebook-combined* network with α is at zero, the proposed measure is competitive with the DC and IS centralities and surpasses BC, CC, CL, and LGC centralities. For *fb-pages-tvshow* with α at zero, BC initially surpasses proposed centrality, and finally, proposed centrality reaches the top.

Figure. 7 displays the histograms of the correlation scores for each of the seven centrality measures (DC, BC, CC, CL, LGC, IS, and ISCL) obtained using the IC model for four datasets. For the *bio-dmela* network structure, our proposed metric ISCL can significantly outperform DC, BC, CC, CL, LGC, and IS in terms of the cumulative infected nodes. The ISCL value outperforms other β values in terms of infected nodes, especially when compared to the transmission probability, which is β at 0.03. The proposed measure clearly beat all centralities for β ranging from 0.01 to 0.05 in the Figure. 7

For the *soc-anybeat* network, ISCL can outperform DC, BC, CC, CL, LGC, and IS architecture in terms of the number of infected nodes with a convex tuning parameter

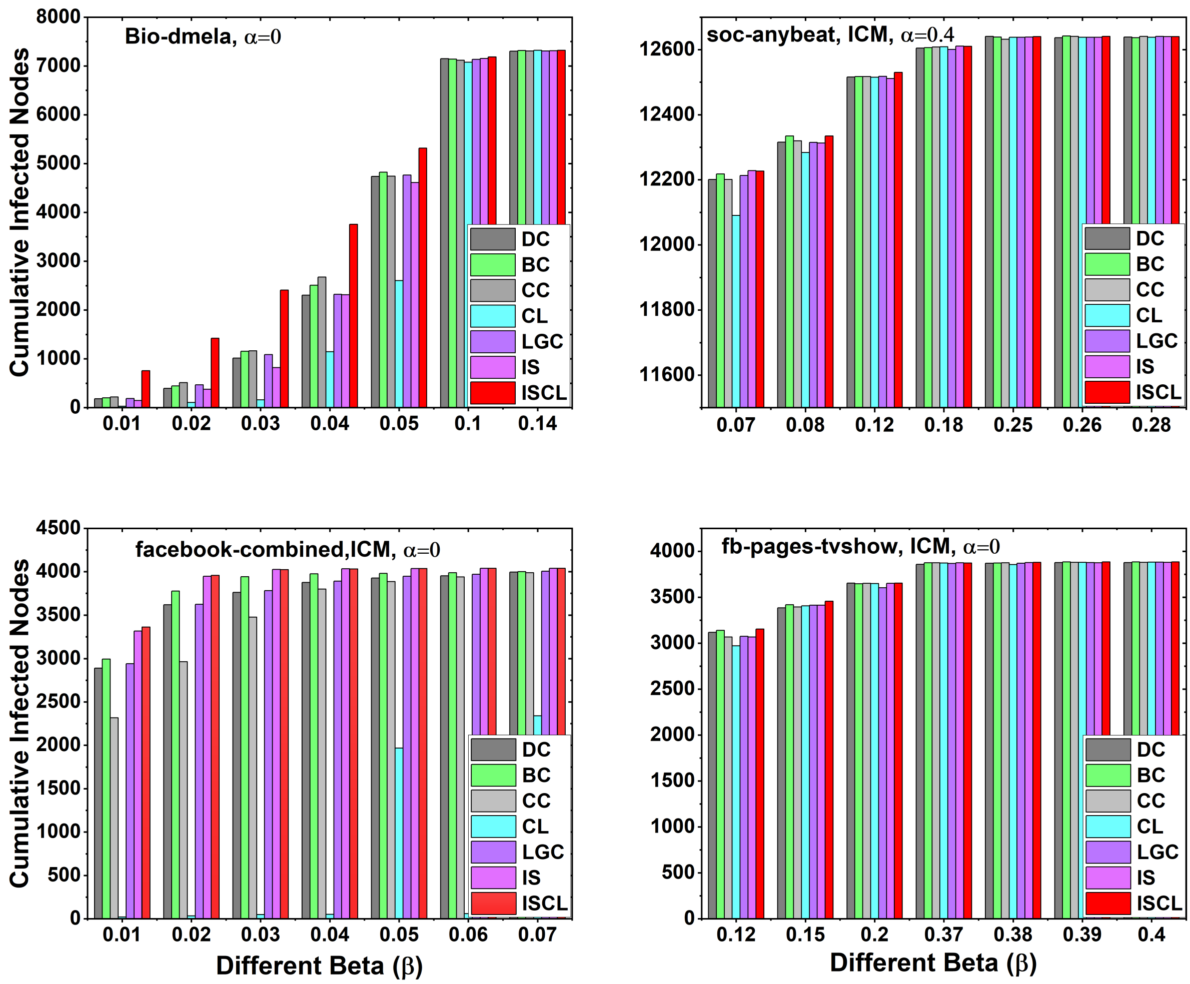

FIGURE 7. For four separate datasets, top-7 histograms of the total infected nodes using the *IC* model-simulated over DC, BC, CC, CL, LGC, IS, and ISCL centrality metrics.

TABLE 7. Results of SIR for ISCL and other centralities for maximum information spread on four real-world networks with average over 100 simulations.

Network	Nodes	DC	BC	CL	CC	IS	LGC	ISCL
fb-pages-tvshow	3892	3244.16	3237.79	3242.05	3241.75	3239.39	3240.45	3249.05
facebook-combined	4039	3878.99	3889.8	1771.9	3861.79	3910.06	3868.92	3919.21
bio-dmela	7393	2007.71	2058.41	841.67	2118.41	1888.02	1986.51	2572.09
soc-anybeat	12645	8621.4	8617.22	8650.54	8658.81	8678.26	8751.37	8894.04

α of 0.4. In terms of infected nodes, ISCL is better than other β values, especially when β is 0.12 for the transmission probability. In comparison to DC, BC, CC, CL, LGC, and IS for the *facebook-combined* network structure, our proposed measure, ISCL, can perform well in terms of the number of infected nodes with convex tuning parameter α is 0. Especially when the transmission probability is 0.01, ISCL is preferred to other β values in terms of infected nodes.

In the case of *fb-pages-tvshow* network, when compared to DC, BC, CC, LGC, and IS structure, our proposed measure, ISCL, can perform well in terms of the number of infected nodes with convex tuning parameter α is 0. In terms of infected nodes, ISCL is better than other β values, especially when the transmission probability is β at 0.15.

For different β values, mean and 95 percent of confidence interval lower and upper values, the standard deviation of the influence spread under the SIR and IC models are presented in Tables 8, 9, 10, and 11. The tables 8, 9, 10, and 11 illustrate how the performance of the proposed method ISCL varies based on the network structure and diffusion models that are utilized. It can also be inferred that the proposed method ISCL functions well under a diffusion model on a particular network, but it might not function as well on other networks. Accordingly, the proposed method ISCL functions well under the SIR model on networks such as soc-anybeat, facebook-combined, fb-pages-tvshow, and bio-dmela. Regarding the IC model, the proposed approach, ISCL, does perform mediocrely on the fb-pages-tvshow, soc-anybeat, facebook-combined, and bio-dmela networks. The

TABLE 8. Standard deviation with 95 percent confidence interval of the proposed centrality ISCL (Isolating Clustering Coefficient) for SIR simulation outcomes. Where β represents the transmission probability, μ denotes the mean, σ represents the standard deviation, and C.I.L and C.I.U represent the lower and upper values of the confidence interval, respectively.

bio-dmela					facebook-combined				
β	μ	**C.I.L**	**C.I.U**	σ	β	μ	**C.I.L**	**C.I.U**	σ
0.01	1951.3	1901.7	2000.8	248.5	0.03	3906.4	3900.2	3912.6	31.0
0.02	4193.2	4169.5	4216.8	118.6	0.04	3923.9	3917.9	3929.9	29.9
0.03	5057.4	5045.5	5069.3	59.5	0.11	3951.9	3947.4	3956.3	22.3
0.04	5437.0	5426.3	5447.6	53.1	0.12	3951.9	3947.9	3955.8	19.8
0.05	5616.2	5606.9	5625.4	46.3	0.16	3957.5	3953.1	3961.8	21.8
0.06	5711.5	5698.8	5724.3	63.8	0.18	3960.2	3956.1	3964.2	20.2
0.07	5776.3	5767.0	5785.5	46.4	0.2	3962.2	3958.1	3966.2	20.2

TABLE 9. Standard deviation with 95 percent confidence interval of the proposed centrality ISCL (Isolating Clustering Coefficient) for SIR simulation outcomes. Where β represents the transmission probability, μ denotes the mean, σ represents the standard deviation, and C.I.L and C.I.U represent the lower and upper values of the confidence interval, respectively.

soc-anybeat					fb-pages-tvshow				
β	μ	**C.I.L**	**C.I.U**	σ	β	μ	**C.I.L**	**C.I.U**	σ
0.06	8613.6	8447.2	8780.0	834.5	0.13	2984.6	2973.4	2995.9	56.5
0.15	9074.8	8920.5	9229.1	773.7	0.23	3102.8	3094.0	3111.6	44.3
0.18	9051.2	8900.0	9202.3	757.9	0.32	3181.3	3171.6	3190.9	48.5
0.23	9336.5	9202.9	9470.1	669.8	0.36	3217.6	3208.6	3226.5	44.9
0.33	9465.7	9343.1	9588.2	614.5	0.38	3246.1	3237.4	3254.8	43.5
0.35	9678.4	9549.8	9807.0	644.8	0.39	3247.8	3239.0	3256.6	44.0
0.39	9853.5	9740.6	9966.3	566.0	0.4	3252.4	3242.9	3261.9	47.5

TABLE 10. Standard deviation with 95 percent confidence interval of the proposed centrality ISCL (Isolating Clustering Coefficient) for ICM simulation outcomes. Where β represents the transmission probability, μ denotes the mean, σ represents the standard deviation, and C.I.L and C.I.U represent the lower and upper values of the confidence interval, respectively.

bio-dmela					facebook-combined				
β	μ	**C.I.L**	**C.I.U**	σ	β	μ	**C.I.L**	**C.I.U**	σ
0.01	2866.1	2846.1	2886.1	100.3	0.01	3695.2	3691.8	3698.6	16.9
0.02	5018.3	5007.9	5028.6	51.9	0.02	3937.7	3935.6	3939.8	10.5
0.03	5988.4	5980.9	5996.0	37.7	0.03	3994.4	3993.1	3995.6	6.0
0.04	6509.9	6504.3	6515.5	28.0	0.04	4015.9	4015.0	4016.7	4.2
0.05	6818.8	6814.0	6823.5	23.8	0.05	4026.0	4025.2	4026.7	3.8
0.1	7316.5	7314.6	7318.3	9.27	0.06	4031.5	4030.9	4032.0	2.6
0.14	7377.4	7376.6	7378.3	4.16	0.07	4034.0	4033.5	4034.4	2.3

mean of spread efficiency and its variation over 95 percent of the confidence interval is displayed in Tables 8, 9, 10, and 11.

D. CONVEX TUNING PARAMETER α AND TRANSMISSION PROBABILITY β ANALYSIS

The effect of the optimal values of α and β on four real-world networks is expressed in this subsection. Using our proposed measure *ISCL*, we conducted experiments by altering the convex tuning parameter α over $[0, 1]$, and transmission probability β over (0.01-0.04). When α is 0 and β is at 0.01, the *bio-dmela* network shows stronger positive outcomes for our proposed measure under the SIR model assessment is shown in Figure. 4. This indicates that when *alpha* is 0, the isolating centrality at the regional level retains more local information than the clustering coefficient measure. Again, using the *facebook-combined* network, isolating centrality has a bigger effect on capturing local network information than the clustering coefficient when α is 0 and β is at 0.11. The metric ISCL for the *soc-anybeat* network produces good results when α is 1.0 and β is 0.06. Stated differently, A clustering coefficient fully captures local information. When α is 1.0 and β is 0.23 for the *fb-pages-tvshow* network, our proposed measure, *ISCL*, has marginally shown favorable results since the clustering coefficient measure contributes more to capturing the local information of the network than does the isolating centrality. In the same way, under the *IC* model evaluation shown in Figure 5. For the *bio-dmela* network, our proposed metric produces superior spreading efficiency when α is 0, and β is at 0.03, indicating that local information can be captured by isolating centrality alone. With α equal to 0 and β at 0.01, the isolating centrality is only responsible for capturing local information about the network for the *facebook-combined*. Similarly for α equal to 0 and β at 0.15, the isolating centrality alone captures local

TABLE 11. Standard deviation with 95 percent confidence interval of the proposed centrality ISCL (Isolating Clustering Coefficient) for ICM simulation outcomes. Where β represents the transmission probability, μ denotes the mean, σ represents the standard deviation, and C.I.L and C.I.U represent the lower and upper values of the confidence interval, respectively.

soc-anybeat					fb-pages-tvshow				
β	μ	C.I.L	C.I.U	σ	β	μ	C.I.L	C.I.U	σ
0.07	12180.4	12175.8	12185.1	23.3	0.12	3795.3	3791.8	3798.8	17.5
0.08	12335.8	12332.0	12339.6	19.2	0.15	3853.4	3850.7	3856.0	13.4
0.12	12586.4	12584.9	12588.0	7.9	0.2	3883.8	3882.6	3885.0	6.1
0.18	12640.9	12640.5	12641.3	2.0	0.37	3892.0	3892.0	3892.0	0.14
0.25	12644.8	12744.7	12644.9	0.4	0.38	3892.0	3892.0	3892.0	0.0
0.26	12644.9	12644.9	12645.0	44.0	0.39	3892.0	3891.9	3892.0	0.32
0.28	12645.0	12645.0	12645.0	47.5	0.4	3892.0	3892.0	3892.0	0.0

information of the textitfb-pages-tvshownetwork. In both cases, our proposed measure, ISCL produced marginally positive results. When α is 0.4 and β is 0.12 for the *soc-anybeat* network, it is observed that isolating centrality has a somewhat greater role in capturing local information than the clustering coefficient measure.

VIII. DISCUSSION

Table 7 depicts the outcomes of centrality measures applied to four real-world networks. To ensure a more accurate analysis, we averaged the SIR results across each network type, having conducted 100 simulations on four distinct networks. Notably, the Facebook-combined and bio-dmela networks pose significant challenges for vital node problems due to their high average degree and low $|Deg1|/n$ values. For the facebook-combined network, the centrality measures when α is set 0 yield high value for ISCL compared to DC, BC, CL, CC, IS, and LGC. Although ISCL appears to have the most favorable results, all remaining centrality measures demonstrate the ability to find solutions of comparable quality. Similarly, the bio-dmela network presents a challenging scenario with an elevated average degree and $|Deg1|/n$ values. Consequently, ISCL centrality exhibits superior results, while other centrality measures compete with each other. For the Facebook-combined, and bio-dmela networks, ISCL centrality exhibiting superior performance is due to the dominance of IS over CL in separating the network into several components. The fb-pages-tvshow and soc-anybeat networks present considerable challenges for vital node problems due to their low average degree and moderate values of $|Deg1|/n$. In the case of the fb-pages-tvshow network, setting α to 1 results in the superior performance of ISCL. While ISCL appears to yield more favorable results, all centrality measures demonstrate the capability to find solutions of comparable quality. Similarly, the soc-anybeat network presents a challenging scenario with a lower average degree and moderate $|Deg1|/n$ values. Once again, ISCL centrality exhibits marginally superior results, while other centrality measures compete with each other. In the case of the soc-anybeat and fb-pages-tvshow networks, the good performance of ISCL centrality can be attributed to the dominance of CL over IS in spreading information across the network.

IX. CONCLUSION

In this work, we have studied the influential nodes in large-scale complex networks using the novel measure ISCL. To design the ISCL, we introduced a convex tuning parameter and integrated the isolating centrality and clustering coefficient to consider the local and global structural information. Extensive simulations employing the SIR and IC models have served to validate the efficacy of centrality measures on four real-world networks such as *bio-dmela*, *soc-anybeat*, *facebook-combined*, and *fb-pages-tvshow*. Our study reveals that, across a range of α values, our proposed measure demonstrates superior performance compared to conventional and recently proposed measures in the literature. This parameter plays a key role in enhancing the applicability to a wide range of real-world networks and enabling their use in large-scale complex network applications. The proposed measure's similarity with conventional and recent measures is quantified using Kendall's correlation coefficient. Interestingly, ISCL and global centralities (BC, CC) exhibit identical time complexity $O(nm)$. ISCL stands out for its considerably improved spreading efficiency while maintaining the same time complexity of BC and CC. Extensive simulations demonstrated that ISCL achieves superior spreading outcomes relative to existing methods while upholding identical time complexity requirements. Extending the proposed method for weighted and directed networks can be an interesting future research direction to analyze the crucial nodes in wireless and social networks.

LIST OF ABBREVIATIONS

The abbreviated terms used in this paper are as follows:

DC	Degree centrality
BC:	Betweenness centrality
CC:	Closeness centrality
CL:	Clustering coefficient centrality
LGC:	Local and global centrality
IS:	Isolating centrality
ISCL:	isolating clustering centrality
SIR:	Susceptible infected recovered
IC:	Independent cascade

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, nos. 4–5, pp. 175–308, 2006.

[2] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, Mar. 2001.

[3] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Mar. 2003.

[4] E. Estrada, *The Structure of Complex Networks: Theory and Applications*. Washington, DC, USA: American Chemical Society, 2012.

[5] X. Zhang, J. Zhu, Q. Wang, and H. Zhao, "Identifying influential nodes in complex networks with community structure," *Knowl.-Based Syst.*, vol. 42, pp. 74–84, Apr. 2013.

[6] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 391, no. 4, pp. 1777–1787, Feb. 2012.

[7] M. Lalou, M. A. Tahraoui, and H. Kheddouci, "The critical node detection problem in networks: A survey," *Comput. Sci. Rev.*, vol. 28, pp. 92–117, Mar. 2018.

[8] S. Pei and H. A. Makse, "Spreading dynamics in complex networks," *J. Stat. Mech., Theory Exp.*, vol. 2013, no. 12, Dec. 2013, Art. no. P12002.

[9] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 3, Feb. 2002, Art. no. 036104.

[10] M. K. Enduri and S. Jolad, "Dynamics of dengue disease with human and vector mobility," *Spatial Spatio-Temporal Epidemiol.*, vol. 25, pp. 57–66, Jun. 2018.

[11] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Trans. Web*, vol. 1, no. 1, p. 5es, 2007.

[12] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the north American power grid," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 2, Feb. 2004, Art. no. 025103.

[13] M. Jusup, P. Holme, K. Kanazawa, M. Takayasu, I. Romić, Z. Wang, S. Gecek, T. Lipić, B. Podobnik, L. Wang, W. Luo, T. Klanjšcek, J. Fan, S. Boccaletti, and M. Perc, "Social physics," *Phys. Rep.*, vol. 948, pp. 1–148, Feb. 2022.

[14] W. Deng and Y. Deng, "Entropic methodology for entanglement measures," *Phys. A, Stat. Mech. Appl.*, vol. 512, pp. 693–697, Dec. 2018.

[15] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, Aug. 2015.

[16] M. Ashtiani, A. Salehzadeh-Yazdi, Z. Razaghi-Moghadam, H. Hennig, O. Wolkenhauer, M. Mirzaie, and M. Jafari, "A systematic survey of centrality measures for protein-protein interaction networks," *BMC Syst. Biol.*, vol. 12, no. 1, pp. 1–17, Dec. 2018.

[17] P. Csermely, T. Korcsmáros, H. J. Kiss, G. London, and R. Nussinov, "Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review," *Pharmacol. Therapeutics*, vol. 138, no. 3, pp. 333–408, 2013.

[18] J. Yang, C. Yao, W. Ma, and G. Chen, "A study of the spreading scheme for viral marketing based on a complex network model," *Phys. A, Stat. Mech. Appl.*, vol. 389, no. 4, pp. 859–870, Feb. 2010.

[19] C. A. Oliveira, P. M. Pardalos, and M. G. Resende, "Optimization problems in multicast tree construction," in *Handbook of Optimization in Telecommunications*. Berlin, Germany: Springer, 2006, pp. 701–731.

[20] N. E. Friedkin, "Theoretical foundations for centrality measures," *Amer. J. Sociology*, vol. 96, no. 6, pp. 1478–1504, May 1991.

[21] A. Landherr, B. Friedl, and J. Heidemann, "A critical review of centrality measures in social networks," *Bus. Inf. Syst. Eng.*, vol. 2, pp. 371–385, Oct. 2010.

[22] P. Hines and S. Blumsack, "A centrality measure for electrical networks," in *Proc. 41st Annu. Hawaii Int. Conf. Syst. Sci. (HICSS)*, Jan. 2008, p. 185.

[23] P. Bródka, K. Skibicki, P. Kazienko, and K. Musial, "A degree centrality in multi-layered social network," in *Proc. Int. Conf. Comput. Aspects Social Netw. (CASoN)*, Oct. 2011, pp. 237–242.

[24] M. Barthelemy, "Betweenness centrality in large complex networks," *Eur. Phys. J. B*, vol. 38, no. 2, pp. 163–168, 2004.

[25] K. Okamoto, W. Chen, and X.-Y. Li, "Ranking of closeness centrality for large-scale social networks," in *Proc. Int. Workshop Frontiers Algorithmics*. Springer, 2008, pp. 186–195.

[26] B. Ruhnau, "Eigenvector-centrality—A node-centrality?" *Social Netw.*, vol. 22, no. 4, pp. 357–365, 2000.

[27] 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 Fractals*, vol. 110, pp. 41–54, May 2018.

[28] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, nos. 1–7, pp. 107–117, Apr. 1998.

[29] N. Gupta, A. Singh, and H. Cherifi, "Centrality measures for networks with community structure," *Phys. A, Stat. Mech. Appl.*, vol. 452, pp. 46–59, Jun. 2016.

[30] C. Lozares, P. López-Roldán, M. Bolibar, and D. Muntanyola, "The structure of global centrality measures," *Int. J. Social Res. Methodol.*, vol. 18, no. 2, pp. 209–226, 2015.

[31] S. Rajeh, M. Savonnet, E. Leclercq, and H. Cherifi, "Comparative evaluation of community-aware centrality measures," *Qual. Quantity*, vol. 57, no. 2, pp. 1273–1302, 2023.

[32] A. Ullah, B. Wang, J. Sheng, J. Long, N. Khan, and Z. Sun, "Identifying vital nodes from local and global perspectives in complex networks," *Expert Syst. Appl.*, vol. 186, Dec. 2021, Art. no. 115778.

[33] M. Chiranjeevi, V. Sateeshkrishna Dhuli, M. Krishna Enduri, and L. Reddy Cenkeramaddi, "ICDC: Ranking influential nodes in complex networks based on isolating and clustering coefficient centrality measures," *IEEE Access*, vol. 11, pp. 126195–126208, 2023.

[34] K. Hajarathaiah, M. K. Enduri, S. Anamalamudi, and A. R. Sangi, "Algorithms for finding influential people with mixed centrality in social networks," *Arabian J. Sci. Eng.*, vol. 48, no. 8, pp. 10417–10428, Aug. 2023.

[35] A. Namtirtha, B. Dutta, and A. Dutta, "Semi-global triangular centrality measure for identifying the influential spreaders from undirected complex networks," *Expert Syst. Appl.*, vol. 206, Nov. 2022, Art. no. 117791.

[36] G. Maji, A. Namtirtha, A. Dutta, and M. Curado Malta, "Influential spreaders identification in complex networks with improved k-shell hybrid method," *Expert Syst. Appl.*, vol. 144, Apr. 2020, Art. no. 113092.

[37] O. Ugurlu, "Comparative analysis of centrality measures for identifying critical nodes in complex networks," *J. Comput. Sci.*, vol. 62, Jul. 2022, Art. no. 101738.

[38] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.

[39] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, "Attack vulnerability of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 5, May 2002, Art. no. 056109.

[40] A. B. M. Nasiruzzaman, H. R. Pota, and M. A. Mahmud, "Application of centrality measures of complex network framework in power grid," in *Proc. 37th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Nov. 2011, pp. 4660–4665.

[41] R. Ghanbari, M. Jalili, and X. Yu, "Correlation of cascade failures and centrality measures in complex networks," *Future Gener. Comput. Syst.*, vol. 83, pp. 390–400, Jun. 2018.

[42] A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao, "Critical nodes for distance-based connectivity and related problems in graphs," *Networks*, vol. 66, no. 3, pp. 170–195, 2015.

[43] C. P. Gillen, A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao, "Robustness of solutions to critical node detection problems with imperfect data: A computational study," *Optim. Methods Softw.*, vol. 32, no. 2, pp. 250–273, Mar. 2017.

[44] H. Habiba and T. Berger-Wolf, "Working for influence: Effect of network density and modularity on diffusion in networks," in *Proc. IEEE 11th Int. Conf. Data Mining Workshops*, Dec. 2011, pp. 933–940.

[45] D. Chen and H. Su, "Identification of influential nodes in complex networks with degree and average neighbor degree," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 13, no. 3, pp. 734–742, Sep. 2023.

[46] S. Zhong, H. Zhang, and Y. Deng, "Identification of influential nodes in complex networks: A local degree dimension approach," *Inf. Sci.*, vol. 610, pp. 994–1009, Sep. 2022.

[47] Z. Sun, Y. Sun, X. Chang, F. Wang, Q. Wang, A. Ullah, and J. Shao, "Finding critical nodes in a complex network from information diffusion and Matthew effect aggregation," *Expert Syst. Appl.*, vol. 233, Dec. 2023, Art. no. 120927.

[48] Z. Zhao, D. Li, Y. Sun, R. Zhang, and J. Liu, "Ranking influential spreaders based on both node k-shell and structural hole," *Knowl.-Based Syst.*, vol. 260, Jan. 2023, Art. no. 110163.

[49] A. Ullah, B. Wang, J. Sheng, J. Long, N. Khan, and Z. Sun, "Identification of nodes influence based on global structure model in complex networks," *Sci. Rep.*, vol. 11, no. 1, p. 6173, Mar. 2021.

[50] Y. Yu, B. Zhou, L. Chen, T. Gao, and J. Liu, "Identifying important nodes in complex networks based on node propagation entropy," *Entropy*, vol. 24, no. 2, p. 275, Feb. 2022.

[51] K. Hajarathaiah, M. K. Enduri, and S. Anamalamudi, "Finding influential nodes in complex networks using nearest neighborhood trust value," in *Proc. 10th Int. Conf. Complex Netw. Their Appl. Complex Netw.* Spain: Springer, 2022, pp. 253–264.

[52] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, "Vital nodes identification in complex networks," *Phys. Rep.*, vol. 650, pp. 1–63, Sep. 2016.

[53] M. Newman, *Networks*. Oxford, U.K.: Oxford Univ. Press, 2018.

[54] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *J. Math. Sociol.*, vol. 2, no. 1, pp. 113–120, 1972.

[55] L. C. Freeman, "Centrality in social networks: Conceptual clarification," in *Social Network: Critical Concepts in Sociology*, vol. 1. London, U.K.: Routledge, pp. 238–263, 2002.

[56] U. Brandes, "A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.

[57] M. S. Mariani, Z.-M. Ren, J. Bascompte, and C. J. Tessone, "Nestedness in complex networks: Observation, emergence, and implications," *Phys. Rep.*, vol. 813, pp. 1–90, Jun. 2019.

[58] M. E. Newman, "A measure of betweenness centrality based on random walks," *Social Netw.*, vol. 27, no. 1, pp. 39–54, 2005.

[59] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[60] J. Zhang and Y. Luo, "Degree centrality, betweenness centrality, and closeness centrality in social network," in *Proc. 2nd Int. Conf. Modelling, Simulation Appl. Math. (MSAM)*. Amsterdam, The Netherlands: Atlantis Press, 2017, pp. 300–303.

[61] J. Wang, H. Mo, F. Wang, and F. Jin, "Exploring the network structure and nodal centrality of China's air transport network: A complex network approach," *J. Transp. Geography*, vol. 19, no. 4, pp. 712–721, 2011.

[62] J. Leskovec and A. Krevl. (Jun. 2014). *SNAP Datasets: Stanford Large Network Dataset Collection*. [Online]. Available: http://snap.stanford.edu/data

[63] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Proc. AAAI*, 2015, pp. 1–2. [Online]. Available: https://networkrepository.com

[64] K. Hajarathaiah, M. K. Enduri, S. Dhuli, S. Anamalamudi, and L. R. Cenkeramaddi, "Generalization of relative change in a centrality measure to identify vital nodes in complex networks," *IEEE Access*, vol. 11, pp. 808–824, 2023.

[65] J. Goldenberg, B. Libai, and E. Müller, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing Lett.*, vol. 12, no. 3, pp. 211–223, 2001.

[66] M. Li, X. Wang, K. Gao, and S. Zhang, "A survey on information diffusion in online social networks: Models and methods," *Information*, vol. 8, no. 4, p. 118, Sep. 2017.

[67] D. Rosa and A. Giua, "On the spread of innovation in social networks," *IFAC Proc. Volumes*, vol. 46, no. 27, pp. 322–327, 2013.

[68] M. G. Kendall, "The treatment of ties in ranking problems," *Biometrika*, vol. 33, no. 3, pp. 239–251, 1945.

BURAN BASHA MOHAMMAD received the B.E. degree in electronics and communication engineering from VTU, Karnataka, India, in 2005, and the M.Tech. degree in electronics and communication engineering from JNTUH, Telangana, India, in 2013. He is currently pursuing the Ph.D. degree in electronics and communication engineering with SRM University AP, Amaravati, India. His research interests include the complex networks and analysis of algorithms, graph machine learning, and wireless sensor networks with the IoT.

V SATEESHKRISHNA DHULI received the B.Tech. degree in electronics and communication engineering from the Maharaj Vijayaram Gajapathi Raj College of Engineering, Vizianagaram, India, in 2008, the M.Tech. degree in electronics and communication engineering from the National Institute of Technology, Warangal, India, in 2011, and the Ph.D. degree in electrical engineering from Indian Institute of Technology Kanpur, Kanpur, in 2019. He is currently an Assistant Professor with SRM University AP, Andhra Pradesh. His research interests include wireless sensor networks, the Internet of Things, and communication systems.

MURALI KRISHNA ENDURI (Member, IEEE) received the M.Sc. degree in mathematics from Acharya Nagarjuna University, Andhra Pradesh, India, in 2008, the M.Tech. degree in systems analysis and computer applications from the National Institute of Technology Karnataka, India, in 2011, and the Ph.D. degree in computer science and engineering from Indian Institute of Technology Gandhinagar, Gujarat, India, in 2018. He is currently an Assistant Professor with the Department of Computer Science and Engineering, SRM University AP, Amaravati, India. He has research experience as a Postdoctoral Researcher with Indian Institute of Technology Madras, Tamil Nadu, India, in 2018. His research interests include algorithms, complexity theory, and complex networks. Within these general areas, he has research experience in designing algorithms for graph isomorphism problems for restricted graph classes, measuring the impact of the diversity of an article, and predicting the spread of disease.

LINGA REDDY CENKERAMADDI (Senior Member, IEEE) received the master's degree in electrical engineering from Indian Institute of Technology Delhi (IIT Delhi), New Delhi, India, in 2004, and the Ph.D. degree in electrical engineering from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2011.

He worked at Texas Instruments on mixed-signal circuit design before joining the Ph.D. Program at NTNU. After finishing the Ph.D. degree, he worked on radiation imaging for an atmosphere–space interaction monitor (ASIM mission to the International Space Station) with the University of Bergen, Bergen, Norway, from 2010 to 2012. He is currently the Leader of the Autonomous and Cyber-Physical Systems (ACPS) Research Group and a Professor with the University of Agder, Grimstad, Norway. He has coauthored over 190 research publications that have been published in prestigious international journals and standard conferences in the research areas Internet of Things (IoT), cyber-physical systems, autonomous systems, robotics and automation involving advanced sensor systems, computer vision, thermal imaging, LiDAR imaging, radar imaging, wireless sensor networks, smart electronic systems, advanced machine learning techniques, connected autonomous systems, including drones/unmanned aerial vehicles (UAVs), unmanned ground vehicles (UGVs), unmanned underwater systems (UUSs), 5G- (and beyond) enabled autonomous vehicles, and socio-technical systems, such as urban transportation systems, smart agriculture, and smart cities. He is also quite active in medical imaging. Several of his master's students won the best master thesis award in information and communication technology (ICT).

Dr. Cenkeramaddi is also a member of ACM and a member of the editorial boards of various international journals and the technical program committees of several IEEE conferences. He is the Principal Investigator and a Co-Principal Investigator of many research grants from the Norwegian Research Council.