

Six-degree gravity centrality for detecting influential nodes in networks

Jianbo Wang(王建波), Bohang Lin(林渤杭), Zhanwei Du(杜占玮), Ping Li(李平), and Xiao-Ke Xu(许小可)
Citation: Chin. Phys. B, 2025, 34 (**8**): 088902. DOI: 10.1088/1674-1056/adec62
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Node ranking based on graph curvature and PageRank

Hongbo Qu(曲鸿博), Yu-Rong Song(宋玉蓉), Ruqi Li(李汝琦), Min Li(李敏), and Guo-Ping Jiang(蒋国平)
Chin. Phys. B, 2025, 34 (**2**): 028901. DOI: 10.1088/1674-1056/ad9a9b

Identifying influential spreaders in complex networks based on density entropy and community structure

Zhan Su(苏湛), Lei Chen(陈磊), Jun Ai(艾均), Yu-Yu Zheng(郑雨语), and Na Bie(别娜)
Chin. Phys. B, 2024, 33 (**5**): 058901. DOI: 10.1088/1674-1056/ad20d6

Identifying influential spreaders in social networks: A two-stage quantum-behaved particle swarm optimization with Lévy flight

Pengli Lu(卢鹏丽), Jimao Lan(揽继茂), Jianxin Tang(唐建新), Li Zhang(张莉), Shihui Song(宋仕辉), and Hongyu Zhu(朱虹羽)
Chin. Phys. B, 2024, 33 (**1**): 018901. DOI: 10.1088/1674-1056/acd3e0

Robustness of community networks against cascading failures with heterogeneous redistribution strategies

Bo Song(宋波), Hui-Ming Wu(吴惠明), Yu-Rong Song(宋玉蓉), Guo-Ping Jiang(蒋国平),Ling-Ling Xia(夏玲玲), and Xu Wang(王旭)
Chin. Phys. B, 2023, 32 (**9**): 098905. DOI: 10.1088/1674-1056/acd9c3

A novel method for identifying influential nodes in complex networks based on gravity model

Yuan Jiang(蒋沅), Song-Qing Yang(杨松青), Yu-Wei Yan(严玉为),Tian-Chi Tong(童天驰), and Ji-Yang Dai(代冀阳)
Chin. Phys. B, 2022, 31 (**5**): 058903. DOI: 10.1088/1674-1056/ac4226

SPECIAL TOPIC — Computational programs in complex systems

Six-degree gravity centrality for detecting influential nodes in networks

Jianbo Wang(王建波)[1,2,3,†], Bohang Lin(林渤杭)[1], Zhanwei Du(杜占玮)[2], Ping Li(李平)[1], and Xiao-Ke Xu(许小可)[4,‡]

[1] *The School of Computer Science and Software Engineering, Southwest Petroleum University, Chengdu* 610500, *China*
[2] *The School of Public Health, The University of Hong Kong, Hong Kong* 999077, *China*
[3] *The Intelligent Policing and National Security Risk Management Laboratory, Sichuan Police College, Luzhou* 646000, *China*
[4] *The Computational Communication Research Center and The School of Journalism and Communication, Beijing Normal University, Beijing* 100875, *China*

(Received 1 June 2025; revised manuscript received 4 July 2025; accepted manuscript online 7 July 2025)

Identifying critical nodes is a pivotal research topic in network science, yet the efficient and accurate detection of highly influential nodes remains a challenge. Existing centrality measures predominantly rely on local or global topological structures, often overlooking indirect connections and their interaction strengths. This leads to imprecise assessments of node importance, limiting practical applications. To address this, we propose a novel node centrality measure, termed six-degree gravity centrality (SDGC), grounded in the six degrees of separation theory, for the precise identification of influential nodes in networks. Specifically, we introduce a set of node influence parameters — node mass, dynamic interaction distance, and attraction coefficient — to enhance the gravity model. Node mass is calculated by integrating K-shell and closeness centrality measures. The dynamic interaction distance, informed by the six-degrees of separation theory, is determined through path searches within six hops between node pairs. The attraction coefficient is derived from the difference in K-shell values between nodes. By integrating these parameters, we develop an improved gravity model to quantify node influence. Experiments conducted on nine real-world networks demonstrate that SDGC significantly outperforms nine existing classical and state-of-the-art methods in identifying the influential nodes.

Keywords: gravity model, influential nodes, six degrees of separation, semi-local information

PACS: 89.75.Hc

DOI: 10.1088/1674-1056/adec62 **CSTR:** 32038.14.CPB.adec62

1. Introduction

Complex networks serve as powerful tools for modeling real-world systems, such as social, biological, and communication networks.[1–3] The relationships between nodes (representing entities) govern the dynamics of these networks. Consequently, understanding the structure and functionality of networks is crucial for fields such as social media marketing, epidemic control, infrastructure optimization, knowledge dissemination, and financial forecasting.[4] With the rise of big data, the internet, and increasingly complex interactions, network scales have rapidly expanded, becoming more intricate and interconnected. As a result, identifying influential nodes has become a critical challenge,[5] essential for applications such as precision advertising,[6] disease control,[7,8] and information flow optimization.[9] Recent advancements in deep learning, graph theory, and artificial intelligence have highlighted the significance of this task.[10–13] For instance, the algorithm that predicts individual influence in social networks directly impacts viral marketing[14] and political campaigns.[15,16] Similarly, effective interventions in disease or information propagation require accurate identification of key nodes. As networks continue to grow in scale and complexity, the need for precise identification of influential nodes becomes even more urgent.

1.1. Related work

Network analysis often relies on classical centrality measures to evaluate node importance. These measures, based on network structural properties, provide insights into nodes' roles in processes such as information dissemination, network resilience, and other dynamics. Common centrality measures include degree centrality (DC),[17] betweenness centrality (BC),[18] closeness centrality (CC),[19] eigenvector centrality (EC),[20] PageRank,[21] and K-shell centrality (KS).[22] DC measures node importance by counting direct connections, assuming nodes with more connections are more significant. BC assesses how frequently a node lies on the shortest paths between other nodes, with a high value indicating a key role in information flow and influence propagation. CC is the reciprocal of the sum of the shortest path lengths from a node to all other nodes, suggesting that higher CC values enable more efficient information spread. EC extends DC by considering connections to other important nodes. K-shell decomposition

[†]Corresponding author. E-mail: jianbow2021@gmail.com
[‡]Corresponding author. E-mail: xuxiaoke@foxmail.com

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

evaluates node importance based on network structure, particularly in sparse or hierarchical networks, and is often used in resilience analysis. These centrality measures differ in their focus on various network characteristics. For example, CC and BC assess a node's position within the entire network, not just its immediate neighbors, evaluating importance through global topological relationships rather than local neighborhood structures. Ubaida *et al.*[23] proposed a global clustering coefficient-based centrality to analyze profitable product connections in large-scale product networks. Aman *et al.*[24] introduced a method that balances node-specific and global network influences. Yang *et al.*[25] developed a method combining shortest path lengths, path frequencies, and non-shortest paths to identify key nodes. Zhao *et al.*[26] proposed a quantitative model that integrates node importance with global network influence to measure centrality.

In contrast, local centrality methods, such as DC, evaluate node importance based on its neighborhood structure. These methods focus on direct connections or interactions along the shortest paths between nodes. Yang *et al.*[5] emphasized local neighborhood information, making their algorithm computationally efficient for large-scale networks. Zhong *et al.*[27] defined local degree dimensions by analyzing neighbors across different layers around a central node. Xu *et al.*[28] proposed a centrality measure using a local propagation probability model, considering propagation influence scores among neighbors of various orders. Wen *et al.*[29] introduced a method for identifying influencers using local information dimensions, extending the local scale to half the shortest distance from the central node. Zhang *et al.*[30] combined one-step and two-step neighborhood information in a multi-local attribute weighted centrality method using information entropy. Yang *et al.*[31] comprehensively utilized the local attributes, network locations and topological structure information of nodes to evaluate the possibility of each node being selected as a seed.

Recent studies aim to balance global and local information. Ai *et al.*[32] evaluated node importance by integrating self-influence, local structural influence from neighbors and global influence based on topological position and shortest paths. Sheng *et al.*[33] proposed a method that considers global influence through node proximity while accounting for local structural influence from nearest neighbors. Qiu *et al.*[34] improved degree centrality and K-shell decomposition methods to measure both local and global influences, offering a comprehensive evaluation of node importance. Hu *et al.*[35] integrated K-shell hierarchical information with node connectivity and adjacent node values in the GLI algorithm to identify key nodes. Li *et al.*[36] combined local and global methods by analyzing core network nodes using betweenness centrality (BC). Additionally, recent approaches incorporate physics theories to address the balance between global and local information. Su *et al.*[37] divided the network into multiple communities at the global level, and assessed node importance by analyzing the internal and external density entropies of nodes locally. Wang *et al.*[38] introduced a new centrality method based on the centripetal force formula, employing a "seed exclusion algorithm" to avoid the "rich club phenomenon". Yang *et al.*[39] developed an adaptive truncation radius and a relaxed distance formula, incorporating gravitational forces to calculate centrality. Guo *et al.*[40] integrated multiple gravitational models for a comprehensive evaluation of nodes. Zhang *et al.*[41] proposed Laplacian centrality, combining the Laplacian matrix with the gravitational model to more effectively measure node importance. In particular, unlike traditional centrality methods that focus on intrinsic node attributes, gravity-based approaches emphasize the interactions between nodes. By integrating node mass, interaction distance, and attraction coefficients, these methods unify and consolidate multi-dimensional topological information, making them better suited for capturing diffusion dynamics and hierarchical structures in complex networks.

Despite these advancements, challenges remain in practical applications. Traditional centrality methods often focus on either local or global topological features, limiting their accuracy in complex, heterogeneous networks. For example, local methods like DC, PageRank, and EC excel at identifying nodes with tightly connected neighborhoods but fail to account for long-range dependencies. Conversely, global methods like BC and CC assess propagation capabilities comprehensively but are computationally expensive, especially in large networks. Additionally, these methods are sensitive to network noise and outliers, impacting stability and reliability.

Although some advanced methods integrate local and global information, semi-local information remains underexplored. Semi-local information captures the influence that lies between local and global structures. Liu *et al.*[42] proposed a generalized energy model combining information entropy and generalized energy to reflect semi-local influence, using clustering coefficients and inter-node distances to construct a global influence index (GI). Zhang *et al.*[43] introduced a semi-local centrality measure based on relative changes in the average shortest path across the network. Wang *et al.*[44] developed a method based on node dimensions that balances local and global metrics. Farahi *et al.*[45] proposed three methods for identifying key nodes in temporal networks: temporal hypercycle ratio, temporal semi-local integration, and temporal semi-local centrality. Semi-local information plays a critical role in many networks by capturing medium-range dependencies between nodes, which is crucial for assessing node importance comprehensively. However, existing methods have yet to fully incorporate this layer of information. Consequently,

while current methods have improved the accuracy of node importance identification, further exploration is needed to integrate semi-local information and refine the evaluation of node importance.

1.2. Motivation and contributions

Based on the previous analysis, we identify the following key issues: (i) Solely relying on local information may overlook significant interaction relationships between nodes, while global information, though offering a more comprehensive view, suffers from high computational complexity in large-scale networks. (ii) Semi-local information is often neglected, yet it is crucial for assessing node importance. In complex networks, a node's importance depends not only on its immediate neighbors but also on its broader position within the network. (iii) Properly defining the scope of semi-local information is critical, as failure to do so may disrupt the balance between evaluation accuracy and computational complexity, undermining the effectiveness and applicability of the algorithm.[39]

Inspired by Isaac Newton's law of universal gravitation, which states that the gravitational force between two objects is proportional to their masses and inversely proportional to the square of their distance, we propose solutions to these challenges. The main contributions of this paper are as follows.

(i) We propose a new method, the six-degree gravitational centrality (SDGC) metric, to identify key nodes in networks by quantifying node interaction forces based on the "law of universal gravitation". We systematically map the three fundamental components of the "law of universal gravitation" — mass, distance, and gravitational constant — onto corresponding topological features in complex networks, thereby proposing an integrated framework that incorporates local, semi-local, and global information.

(ii) Based on the six-degree separation theory, we limit path searches to six hops, balancing neighborhood propagation range and minimizing computational overhead in large-scale networks. Building on the traditional gravitational model, we introduce cross-layer propagation and replace the fixed gravitational constant with an attraction coefficient, more accurately reflecting node interaction relationships.

(iii) We conduct comprehensive comparative experiments across nine real-world networks, evaluating nine state-of-the-art and classical methods in terms of correlation, propagation capability, monotonicity, and node removal. Our results demonstrate the SDGC method's high accuracy and strong node differentiation capability.

The remainder of this paper is organized as follows: Section 2 details the proposed method; Section 3 covers experimental setup, results, and method comparisons across different networks; and Section 4 concludes the paper and outlines future work.

2. Proposed method

In a given undirected graph $G = (V, E)$, V represents the set of nodes, and E represents the set of edges between nodes. The topological structure of the network is typically described by an adjacency matrix $A = (a_{ij})_{N\times N}$, where N denotes the total number of nodes in the network. The elements a_{ij} in the adjacency matrix describe the connection relationships between nodes: when $a_{ij} = 1$, it indicates that there exists an edge between nodes i and j; when $a_{ij} = 0$, it indicates that there is no connection between nodes i and j.

Existing methods typically consider only local or global information of networks, failing to fully capture the multi-level and dynamic characteristics of node interactions or effectively handle the neighborhood structure in complex networks. To address these limitations, we propose an improved gravity model-based node identification method inspired by the six degrees of separation theory — six-degree gravity centrality (SDGC). SDGC is grounded in the core concept of the gravity model, evaluating node importance through the cumulative effect of interactions between a node and its surrounding nodes. The gravitational force in the model is enhanced by three key factors:

(i) Node "mass": Node mass quantifies the intrinsic importance of a node as a constituent of the network structure. Unlike fixed mass in physical systems, node importance in networks is dynamic. SDGC dynamically assesses a node's mass by integrating its local properties (e.g., neighborhood structure) and global properties (e.g., relative position).

(ii) Dynamic interaction distance between nodes: Dynamic interaction distance measures the ease with which two nodes can interact within the network. The interaction force is inversely proportional to distance. Traditional shortest-path metrics fail to fully reflect the multi-path propagation characteristics in complex networks. To address this, SDGC quantifies interaction distance by searching all possible paths between node pairs and incorporates the six degrees of separation theory to limit the path search scope, balancing the comprehensiveness of network structure information.

(iii) Cross-level attraction coefficient: The attraction coefficient reflects the intensity of interaction between node pairs, with an emphasis on capturing cross-layer interaction phenomena. SDGC improves the gravitational constant in the traditional gravity model by introducing an attraction coefficient based on K-shell value differences between nodes, emphasizing the impact of cross-level interactions on information propagation. The K-shell value reflects a node's relative position in the network hierarchy, and its difference highlights the distinct core layers where nodes reside. This hierarchical difference enables information propagation to transcend layer boundaries, thereby enhancing the strength of connections between nodes.

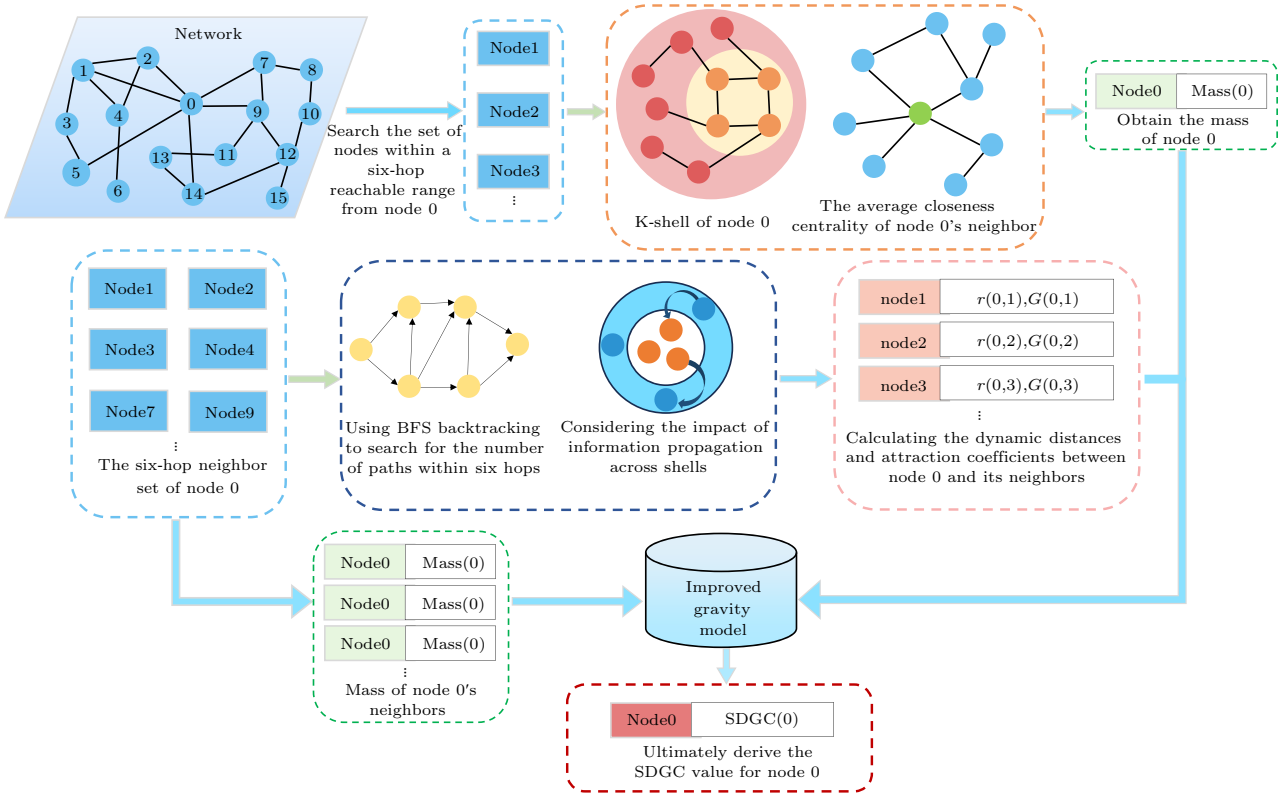


Fig. 1. The framework of the proposed SDGC method.

On the other hand, the six degrees of separation theory posits that any two individuals can be connected through five intermediaries, a concept validated in Milgram's "small-world experiment".[46] This theory reveals that in real-world social networks, information exchange between nodes can be efficiently achieved through a limited number of intermediate paths, with connections beyond this range having significantly reduced influence. Therefore, SDGC defines a node's centrality as the cumulative force of interactions with neighboring nodes (both direct and indirect) within its six-degree neighborhood. The framework of the SDGC method is illustrated in Fig. 1.

2.1. SDGC method

Below, we will provide a detailed description of the SDGC calculation process.

Step 1 Calculation of the node mass.

In gravity model-based centrality methods, node mass is a crucial factor for assessing influence. In the SDGC method, node mass integrates local network properties and the node's global network position. We select closeness centrality (CC) and K-shell value as key factors for the following reasons: CC reflects the likelihood and cost of interactions between a node and others,[47] serving as an important indicator of information dissemination efficiency. In SDGC, we further use the average CC of neighbors within a node's sixth-order neighborhood (i.e., nodes reachable within six hops) to assess the ability of the node's local community to propagate information externally. A higher average CC value indicates lower dissemination costs within the community, facilitating information spread. Thus, the average CC value intuitively reflects the ease of interaction between the local community and external nodes. The pseudocode for the sixth-order neighborhood algorithm is provided in Algorithm 1.

Algorithm 1: Compute six-degree neighborhood.

Input: Graph $G=(V,E)$, target node n, maximum degree $\delta=6$
Output: The set of nodes neighbors$_6$ within six degrees of n
Initialize distances$[n]\leftarrow 0$, $Q\leftarrow[n]$, neighbors$_6\leftarrow\emptyset$
while Q is not empty **do**
    Pop node u from Q
    Add u to neighbors$_6$
    if distances$[u]\geq\delta$ **then**
        break
    end if
    for each neighbor v of u **do**
        if $v\notin$ distances **then**
            distances$[v]\leftarrow$ distances$[u]+1$
            Add v to Q
        end if
    end for
end while
return neighbors$_6$

The K-shell value indicates the node's core layer in the network, serving as a key metric for assessing its global topological status. Nodes in higher core layers are closer to the network center and have greater potential and range for information dissemination. However, K-shell value calculations

often result in duplicate values, leading to insufficient differentiation between nodes, which can bias mass estimation and limit applicability in large-scale networks. To address this, we apply an exponential smoothing process to the K-shell value, reducing the impact of duplicate values while preserving its ability to identify core layers, thereby enhancing the method's performance. This combined approach of using the clustering coefficient and K-shell value balances local and global information while mitigating the adverse effects of K-shell value degeneracy, thereby enhancing its applicability to large-scale networks.

Definition 1 (Node mass). Based on the above analysis, the formula for calculating the mass of node i is as follows:

$$\text{mass}(i) = (\overline{C}_i{}^{\alpha})^{(ks_i)^{\beta}}, \tag{1}$$

where $\overline{C}_i{}^{\alpha}$ represents the average closeness centrality of the neighbor nodes within the six-degree neighborhood of node i, and ks_i denotes the K-shell value of node i. α and β are balancing coefficients used to adjust the influence of the node itself and its neighbors on the centrality calculation.

Step 2 Calculation of dynamic interaction distance between nodes.

In the gravity model, the interaction distance between nodes is crucial for calculating the interaction force, being inversely proportional to the interaction strength. While traditional centrality measures, based on shortest path length, are intuitive, they overlook multi-path interactions and the dynamic nature of information dissemination in complex networks. To address this issue and thereby more accurately model the distance component in the gravitational framework, reflecting the ease of interaction between nodes in complex networks, the SDGC method introduces a dynamic interaction distance based on semi-local information, which better captures the interaction dynamics in complex networks. This measure is derived from the six degrees of separation theory, restricting the path search to the sixth-order neighborhood of each node. This approach retains key topological information while reducing computational complexity. The dynamic interaction distance incorporates both the shortest path length and the number of possible connecting paths. A shorter path indicates lower interaction costs and a higher likelihood of information exchange, thereby reducing the interaction distance. Conversely, a greater number of connecting paths increases the interaction potential, further shortening the dynamic interaction distance. By doing so, the dynamic interaction distance more effectively captures the true nature of node interactions in complex networks, consistent with the principle that interaction strength is inversely proportional to the effective distance measure. The pseudocode for calculating the number of shortest paths between nodes i and j within six hops is provided in Algorithm 2.

Algorithm 2: Compute shortest paths within six degrees.

Input: Graph $G = (V, E)$, source node s, target node t, maximum degree $\delta = 6$
Output: Number of shortest paths within six degrees from s to t
if $s = t$ **then**
    return 1
end if
Initialize distances$[s] \leftarrow 0$, path_counts$[s] \leftarrow 1$
Initialize queue $Q \leftarrow [s]$
while Q is not empty **do**
    Pop u from Q 10:   **if** distances$[u] \geq \delta$ **then**
        break
    end if
    for each neighbor v of u **do**
        if $v \notin$ distances **then**
            distances$[v] \leftarrow$ distances$[u] + 1$
            Add v to Q
        end if
        if distances$[v] =$ distances$[u] + 1$ **then**
            path_counts$[v] \leftarrow$ path_counts$[v] +$ path_counts$[u]$
        end if
    end for
end while
return path_counts$[t]$

Definition 2 (Dynamic interaction distance). Based on the above analysis, the formula for calculating the dynamic interaction distance between nodes i and j is given as follows:

$$r_{ij} = d_{ij}^{(N/P_{ij})^{\gamma}}, \tag{2}$$

where d_{ij} represents the shortest path length between nodes i and j, and P_{ij} denotes the number of connecting shortest paths between nodes i and j within six hops. N is a normalization factor, representing the number of nodes in the network, used to balance the weight of the shortest path count across networks of different scales and to prevent computed values from becoming excessively large or small, thereby mitigating the impact of outliers on subsequent calculations. γ is a balancing parameter that controls the nonlinear influence of the path count on the interaction distance.

Step 3 Calculation of the attraction coefficient between nodes.

In gravity model-based centrality methods, the attraction coefficient between nodes directly determines the strength of interactions, making it a key parameter for modeling information dissemination dynamics. To define the attraction coefficient more effectively, the SDGC method introduces an innovative approach based on K-shell values, which integrates the node hierarchy within the network and their cross-hierarchical propagation characteristics. The K-shell, derived by iteratively removing peripheral nodes, captures the topological hierarchy and the potential for information dissemination. Nodes with low K-shell values are typically located at the periphery of the network, with fewer connections and limited influence, while those with high K-shell values are central, with numerous connections and strong ties to key nodes. Therefore, the dispar-

ity in K-shell values between nodes reflects underlying cross-hierarchical propagation characteristics.[48] Specifically, significant K-shell differences correspond to pronounced hierarchical gaps, driving information flow across hierarchies. This is especially evident in the propagation from low-K-shell to high-K-shell nodes, where information spreads through core nodes to broader regions, enhancing coverage and diffusion efficiency.

Definition 3 (Attraction coefficient). To reflect the positive impact of hierarchical differences on information propagation, we define the attraction coefficient between nodes, with its calculation method presented below:

$$G(i,j) = (|ks_i - ks_j| + 1)^{\sigma}, \tag{3}$$

where ks_i and ks_j represent the K-shell values of nodes i and j, respectively. The parameter σ is used to adjust the influence of the attraction coefficient on the interaction force between nodes. The constant term of $+1$ in the formula ensures that even nodes with identical K-shell values maintain some level of attraction, preventing the calculated value from being zero, which would otherwise render the model incapable of handling interactions for those nodes. To ensure the reasonableness of the attraction coefficient between nodes, we limit the path search range to the sixth-order neighborhood. This restriction enables the differences in K-shell values to more directly reflect the hierarchical disparities between nodes, ensuring that interactions are meaningful. Furthermore, this limitation prevents the information propagation efficiency from approaching zero due to excessively large distances between nodes, thereby enhancing the model's ability to capture the actual characteristics of information propagation in complex networks.

Step 4 Calculation of the interaction force between nodes.

Inspired by the law of universal gravitation, the interaction force between nodes i and j is calculated as follows:

$$F(i,j) = G(i,j)\frac{\text{mass}(i)\cdot\text{mass}(j)}{r(i,j)^2}, \tag{4}$$

where $\text{mass}(i)$ and $\text{mass}(j)$ represent the masses of nodes i and j, respectively, $r(i,j)$ denotes the dynamic interaction distance between nodes i and j, and $G(i,j)$ is the attraction coefficient between nodes i and j.

Step 5: Calculation of the SDGC value of nodes.

In the SDGC method, the centrality of a node is defined as the cumulative interaction force with neighboring nodes within its sixth-order neighborhood. The interaction force between nodes is calculated based on a gravitational model, which not only reflects the direct contribution of local neighbors to centrality but also incorporates global network information through the specification of the neighborhood range.

Definition 4 (SDGC value). Based on the above analysis, the SDGC value of node i is defined as follows:

$$\text{SDGC}(i) = \sum_{i\neq j,\ j\in N_i^{(6)}} F(i,j), \tag{5}$$

where $N_i^{(6)}$ denotes the set of nodes in the sixth-order neighborhood of node i.

To sum up, we propose a novel node centrality measurement method based on a gravitational model, termed SDGC. The SDGC score is the final node centrality value computed under the gravitational framework; it integrates the three previously defined components and encapsulates multiple topological features to comprehensively assess node importance. Unlike traditional centrality metrics, SDGC comprehensively accounts for local network properties (such as closeness centrality and K-shell value) and the dynamic interaction distance between nodes, thereby providing a more holistic assessment of a node's influence. Specifically, by restricting the neighborhood range to six degrees, SDGC effectively captures network topological features while avoiding excessive computational complexity. Furthermore, SDGC mitigates inaccuracies in influence calculations caused by overly distant nodes during information propagation, thus more accurately reflecting a node's importance within the network. The pseudocode of the SDGC method is shown in Algorithm 3.

Algorithm 3: Framework of the proposed SDGC method.

Input: Graph $G = (V,E)$, parameters $\alpha,\beta,\gamma,\sigma$
Output: Gravity centrality $\text{SDGC}(i)$ for each node $i \in V$
Compute closeness centrality and K-shell for all nodes
Initialize centrality$[i] \leftarrow 0$ for each $i \in V$
for each node $i \in V$ **do**
    neighbors_6 $\leftarrow$ Algorithm 1 $(G, i, \delta = 6)$
    Remove i from neighbors
    if neighbors empty **then**
        continue
    end if
    Compute average closeness of neighbors
    Compute mass of node i
    for each $j \in$ neighbors_6, $j \neq i$ **do**
        $P_{ij} \leftarrow$ Algorithm 2 $(G, i, j, \delta = 6)$
        if $P_{ij} > 0$ **then**
            Compute distance $r(i,j)$
            Compute attraction coefficient $G(i,j)$
            Compute gravity force $F(i,j)$
            Update $\text{SDGC}(i)$
        end if
    end for
    Update centrality for node i
end for
return centrality

2.2. Time complexity analysis of the SDGC method

The time complexity of our proposed SDGC method primarily consists of the following components. First, when calculating node mass, neighbors within a six-degree range are

computed using BFS, with a complexity of $O(N+E)$, which is comparable to the complexity of calculating centrality and K-shell, i.e., $O(N\cdot(N+E))$. Second, when computing the number of paths and distances between two nodes, BFS is employed, with a complexity of $O(N+E)$. Finally, when calculating the centrality of the gravity model, each node requires computations with its six-degree neighbors, resulting in an overall complexity of $O(N\cdot(N+E))$. Therefore, the total time complexity of the SDGC method is $O(N\cdot(N+E))$, which simplifies to $O(N^2)$ in sparse graphs. Compared to some existing methods with $O(N^3)$ complexity, the SDGC method offers superior computational efficiency, enabling improved performance for large-scale networks while maintaining high accuracy. Although the complexity of SDGC exceeds that of methods like DC and K-shell, which rely solely on a single topological structure, its node identification performance surpasses these methods. This is because SDGC integrates local, global, and semi-local topological structures, trading off some time complexity to achieve greater accuracy.

3. Experiments

To validate the effectiveness of the SDGC method, we selected nine real-world network datasets and conducted a comparative analysis with nine state-of-the-art and classical centrality identification methods across five experiments.

3.1. Experimental setup

3.1.1. Parameter settings

In this study, the parameter α in Eq. (1) is set to 4.2 to regulate the influence of neighbor nodes' closeness centrality on node mass and to control the interaction weight between local communities and other communities. The parameter β is set to 0.1 to adjust the impact of K-shell values on node mass, emphasizing the importance of the network's core hierarchy in information dissemination and mitigating the issue of insufficient differentiation in K-shell values. The parameter γ in Eq. (2) is set to 0.56 to control the nonlinear effect of path counts on interaction distances. The parameter σ in Eq. (3) is set to 0.32 to modulate the influence of the attraction coefficient on the force between nodes. These parameter values were determined through systematic experimental tuning and validated multiple times to ensure optimal performance across various network datasets.

3.1.2. Datasets

This study selected nine real-world network datasets from diverse sources and with varied characteristics, covering domains such as social, collaboration, transportation, and communication networks. These datasets not only represent a variety of network contexts but also encompass different topological properties, enabling a comprehensive evaluation of the proposed method's applicability and robustness. Specifically, Jazz reflects collaboration relationships among jazz musicians;[49] Haggle describes a human contact network involving individuals carrying wireless devices;[50] USAir represents the transportation network of domestic airlines and airports in the United States;[51] Infectious records visitor interaction data during a gallery exhibition, capturing face-to-face social behaviors;[52] Email documents the email communication network among members of the Universitat Rovira i Virgili in Catalonia;[53] Blogs captures hyperlink relationships among blog webpages during the 2004 US presidential election;[54] Hamsterster describes friendships and family ties within a virtual community;[55] Facebook represents a social network composed of anonymized user circles on the Facebook platform;[56] and GrQc documents collaborations among scientists in the fields of general relativity and quantum cosmology.[57] These datasets range from small-scale networks with tens of nodes and hundreds of edges to large-scale networks with thousands of nodes and tens of thousands of edges, exhibiting varying density and connectivity.

Furthermore, Table 1 summarizes the basic topological properties of each network, including the number of nodes (N), the number of edges (E), the average degree ($\langle k\rangle$), the average shortest path length ($\langle d\rangle$), and the clustering coefficient (C).

Table 1. Statistical characteristics of nine real-world networks.

Network	N	E	$\langle k\rangle$	$\langle d\rangle$	C
Jazz	198	2742	27.6970	2.2350	0.6175
Haggle	274	28244	206.1610	2.4153	0.5664
USAir	332	2126	12.8072	2.7381	0.6252
Infectious	410	17298	84.3805	3.5679	0.4357
Email	1133	5451	9.6222	3.6060	0.1101
Blogs	1224	19025	31.0866	2.7222	0.2260
Hamsterster	2426	16631	13.7106	3.6685	0.2314
Facebook	4039	88234	43.6910	3.6925	0.6170
GrQc	4158	13422	6.4560	6.0494	0.6648

3.1.3. Comparison methods

To verify the advantages of the SDGC method, we compared it with nine classical and state-of-the-art methods, including four classical methods (DC, BC, CC, and K-shell) and five state-of-the-art methods (GC, EGM, LGC, AOGC, and FGM). Detailed descriptions of these methods are provided below.

(i) DC:[17] Degree centrality (DC) measures a node's importance in the network based on its degree (i.e., the number of edges connected to the node). A higher DC value indicates greater node influence.

(ii) BC:[18] Betweenness centrality (BC) quantifies a node's importance as an intermediary in information transmission within the network. Specifically, the BC value reflects the node's role as a "bridge" in the shortest paths between other nodes, indicating its significance in connecting other nodes. A

higher BC value signifies a more prominent role in facilitating information flow.

(iii) CC:[19] Closeness centrality (CC) measures a node's average distance to all other nodes in the network, which is expressed as the reciprocal of the sum of the shortest path distances from node i to all other nodes. A higher CC value indicates stronger connectivity and greater information propagation capability.

(iv) K-shell:[22] K-shell centrality measures a node's proximity to the network's core. Nodes closer to the core have higher K-shell values, implying greater influence within the network.

(v) GC:[58] Gravity centrality (GC), inspired by the law of gravitation, assumes that a node's influence is related to its degree and its distance to other nodes. In this model, a node's degree is treated as its "mass", and the shortest path distance between nodes represents their "distance". A truncation radius, typically set to $0.5 \cdot \langle d \rangle$, where $\langle d \rangle$ is the network's average shortest path length, is used to address noise and computational overhead.

(vi) EGM:[59] Proposed by Li *et al.*, the extended gravity model (EGM) assumes that each node has a distinct influence radius. The influence radius of node i is determined by its path to the farthest node j, calculated under the assumption of equal influence from a virtual node s. The radius accounts for the shortest paths and the equal-influence hypothesis, while node i's value is represented by its degree and the sum of the degrees of its local neighbors. This method integrates the law of gravitation, combining the influence radius and node value to assess node importance.

(vii) LGC:[41] Laplacian gravity centrality (LGC) evaluates a node's influence by considering the degrees of the node and its neighbors, using Laplacian gravity pairs to quantify its propagation capability in the network. By combining Laplacian centrality with a gravity model, LGC provides a more accurate assessment of node importance, particularly in social networks and propagation models.

(viii) AOGC:[39] The adaptive optimal gravity centrality (AOGC), proposed by Yang *et al.*, comprehensively evaluates node interactions by integrating an adaptive truncation radius and all-path distances. This method employs neighborhood structure centrality (NSC) and dynamic quasi-local similarity (DQS) to measure node importance and adaptively determines the truncation radius to balance computational efficiency and accuracy. A node's influence is represented by the sum of interaction forces with neighboring nodes within the truncation radius, reducing computational complexity.

(ix) FGM:[40] The fusion gravity model (FGM), proposed by Guo *et al.*, combines a node's degree and K-shell value to measure its "mass" and introduces the network's normalized maximum eigenvector value as a node weight to better reflect global influence. Node importance is evaluated by integrating degree and K-shell values, considering the shortest path lengths to other nodes, and setting a truncation radius based on the network's average shortest path length to optimize computational efficiency.

3.1.4. Propagation model

Information propagation capability is a key factor in measuring node importance. Due to its simplicity and wide applicability, the susceptible–infected (SI) model is a classic choice for evaluating node propagation capability.[60] Therefore, the experimental section of this paper adopts the SI model to assess the information propagation capability of nodes. In the SI model, nodes in the network are divided into two states: susceptible and infected. Susceptible nodes can be converted into the infected state by infected nodes, and this conversion is irreversible. During the propagation process, infected nodes influence their neighboring susceptible nodes with a propagation probability of β. At each time step t, an infected node has a probability β of converting its susceptible neighbors into infected nodes. The propagation dynamics are characterized by recording the number of infected nodes $F(t)$ at time step t. If the initial infected node has strong propagation capability, the spread will be broader, and the number of infected nodes $F(t)$ will increase accordingly.

3.2. Experimental results

3.2.1. Correlation experiments

This experiment aims to analyze the correlation between the SDGC method and other centrality measurement methods, and to validate the advantages of SDGC in identifying important nodes in complex networks. We conducted correlation analysis on nodes in the Email network to evaluate the relationships between SDGC and six comparative methods (two classical baseline methods and four advanced methods).

We used the SI model to assess node propagation capabilities in the experiment. The infection probability β was set to 0.1, the propagation time was set to $t = 10$, and each network was tested 100 times to obtain average results. The infection capability of each node was quantified by the number of infected neighboring nodes and visualized through color coding, with lighter colors indicating higher infection capability. The results are presented as scatter plots, where each point represents a node, and the coordinates correspond to the SDGC and other centrality scores. Analysis of the scatter plots reveals the correlations between methods. As shown in Fig. 2, most points are concentrated near the main diagonal, indicating that SDGC is strongly correlated with the other methods. The high-centrality nodes identified by SDGC exhibit significant influence. Overall, SDGC demonstrates a strong positive correlation with most centrality methods, confirming

its effectiveness in capturing node influence and propagation capability. Additionally, SDGC outperforms other methods in differentiating nodes, avoiding overly concentrated or redundant centrality scores and more accurately distinguishing nodes with similar propagation capabilities, thus further validating its superiority. It is worth noting that other networks also have similar results, but due to space limitations, they are not presented.

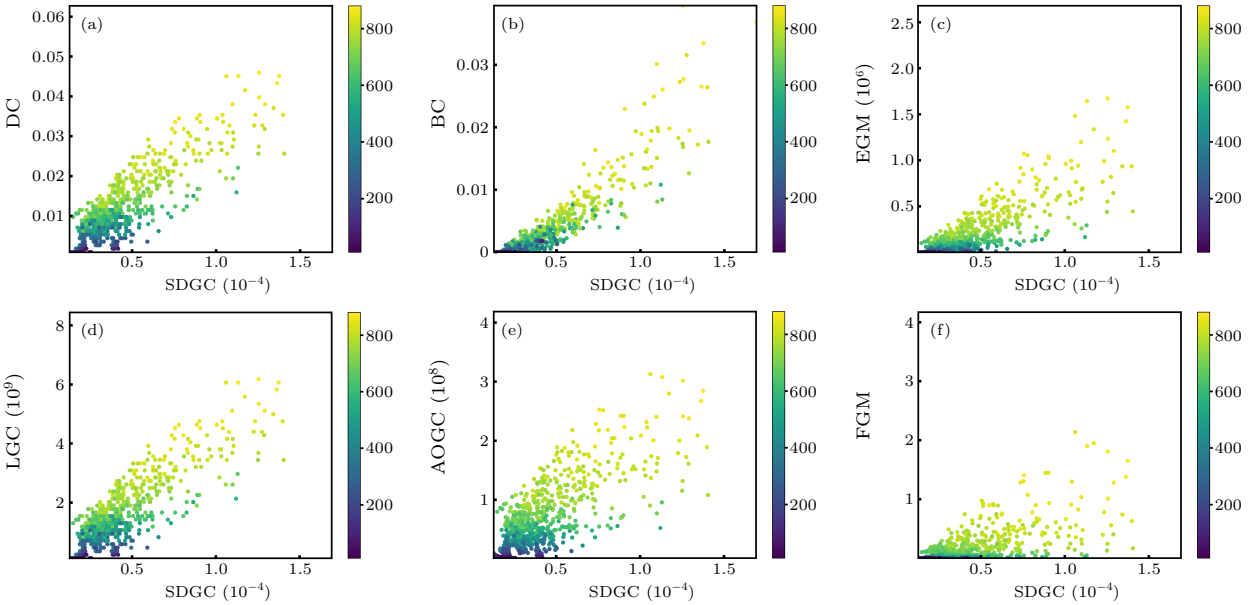


Fig. 2. Correlation analysis between SDGC method and five existing methods in Email network.

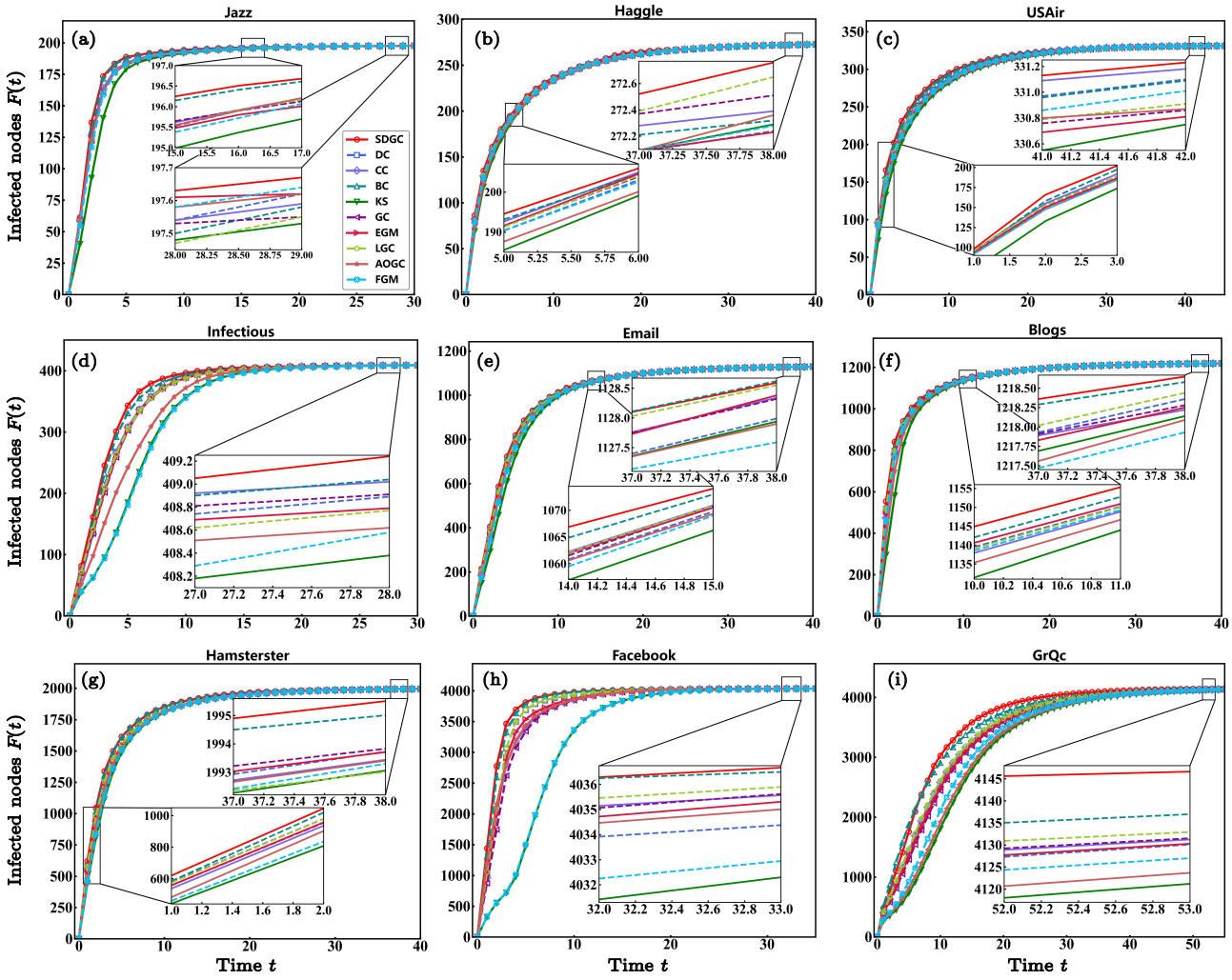


Fig. 3. The infection ability of the top 5% nodes selected by ten methods in nine networks.

3.2.2. Comparison of node infection ability

To evaluate the effectiveness of various methods in identifying key nodes, we employ the SI model to compare the infection capabilities of nodes. The SI model simulates information propagation and reflects the impact of the initially infected node's infection capability on the network.

In this experiment, we select the top 5% ranked nodes from each method as initial infected nodes, with the remaining nodes designated as susceptible. Choosing a proportion rather than a fixed number of initial infected nodes effectively accommodates variations in network scale, preventing experimental results from being skewed by network size. During the experiment, initial infected nodes infect adjacent susceptible nodes with a propagation probability $\beta = 0.1$ over time t. The number of infected nodes, $F(t)$, is used to measure the infection capability of the initial infected nodes. To ensure the stability and objectivity of the results, each experiment is repeated 100 times, and the average is taken as the final result. Node sets with stronger infection capability exhibit faster growth in the number of infected nodes $F(t)$ during propagation, ultimately achieving a larger infection scale. The experimental results are presented in Fig. 3.

Figure 3 compares the infection capability of the top 5% node sets selected by different methods in the Haggle, Blogs, and Hamsterster networks. The results show small differences, but SDGC significantly outperforms others during both the initial rising phase and the final stable phase of the curve, consistently ranking first. In the Email network, SDGC exhibits an advantage in the early propagation stage but slightly underperforms the BC method after stabilization. In the Jazz, USAir, and Infectious networks, SDGC performs comparably to other methods initially, but as propagation progresses, it demonstrates stronger infection effects near the end. In the Facebook network, SDGC maintains superior infection capability throughout the entire propagation process compared to other methods. In the GrQc network, SDGC's infection capability is lower than that of the BC method during the early propagation stage ($0 < t < 7$), but it surpasses most other methods after stabilization. Overall, SDGC exhibits excellent performance in infection capability experiments across networks of varying scales and topological structures, validating its advantages over other methods.

3.2.3. Network efficiency variation

Network efficiency is a core metric for measuring network connectivity, used to evaluate network robustness and

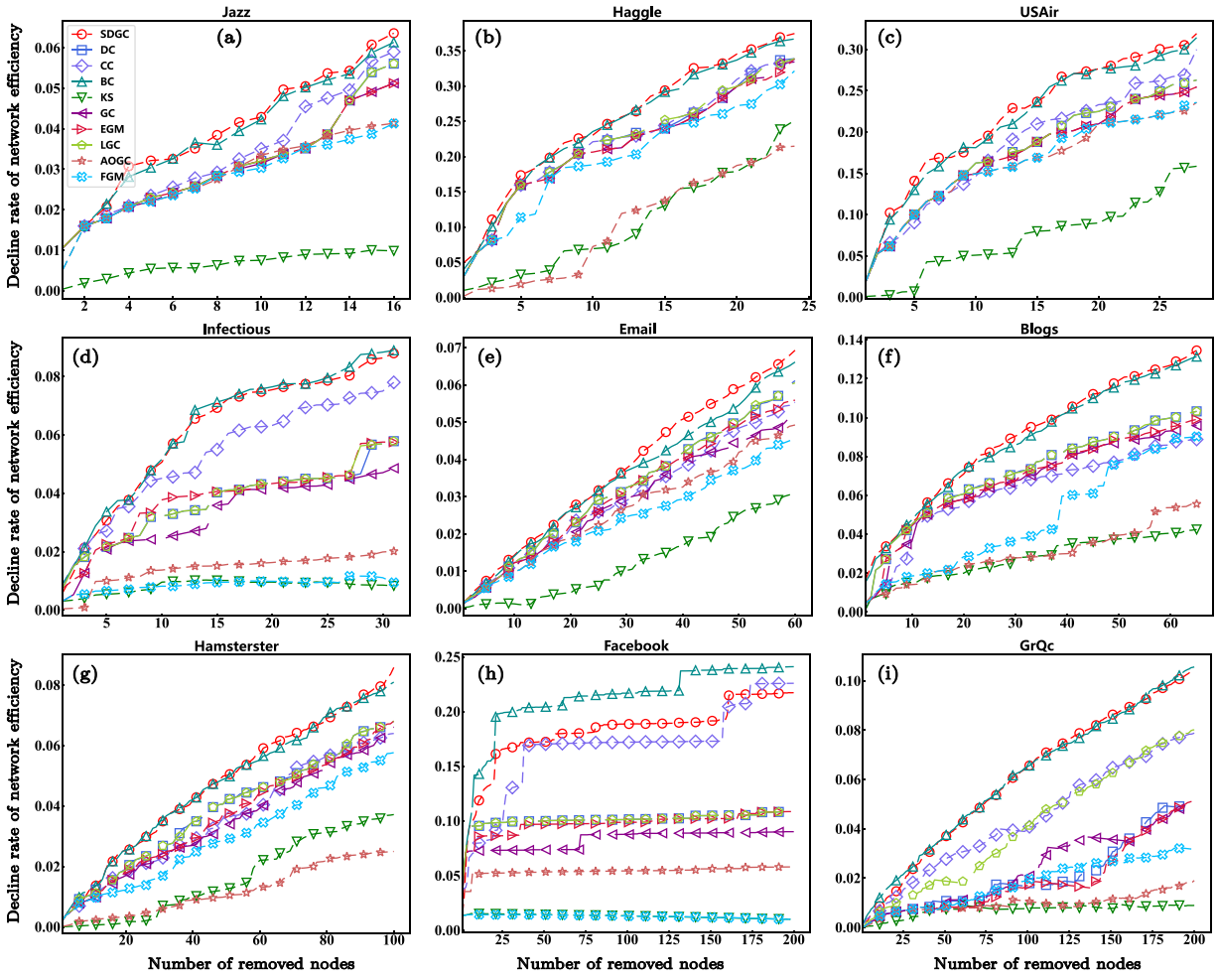


Fig. 4. The network efficiency decline rate of nine networks varies with the number of removed nodes.

stability against attacks.[61] Stronger network connectivity corresponds to higher efficiency values, and the magnitude of efficiency decline upon node removal directly reflects node importance and network vulnerability. The network efficiency $E(G)$ is defined as follows:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}, \tag{6}$$

where N is the total number of nodes in the network, and d_{ij} is the shortest path length between nodes i and j.

This experiment evaluates the changes in network efficiency after removing high-influence nodes selected based on different centrality methods, comparing the accuracy of these methods in ranking node importance. Nodes ranked higher are critical for maintaining network connectivity, and their removal leads to significant declines in network efficiency. Furthermore, to account for differences in network scale and edge density (some networks are small and dense, while others are large but sparse), the number of removed nodes is dynamically adjusted based on the network scale to ensure fairness and clarity in comparisons. The experimental results are presented as curves showing the rate of network efficiency decline with respect to the number of removed nodes, as shown in Fig. 4.

Figure 4 illustrates that in the Infectious, Hamsterster, and GrQc networks, the SDGC method performs comparably to the BC method, both outperforming other methods, with the removal of key nodes identified by SDGC causing particularly significant disruption to network connectivity. In the Email and Blogs networks, SDGC initially performs similarly to the BC method but gradually demonstrates a clear advantage in later stages, surpassing other methods. In the Jazz, Haggle, and USAir networks, SDGC and BC methods exhibit similar performance, consistently ranking among the top two. In the Facebook network, while SDGC's advantage is less pronounced compared to other networks, it still maintains a stable position among the top three. Overall, SDGC outperforms other methods in efficiency experiments across various networks, with the removal of its identified key nodes significantly reducing network connectivity and severely impacting network efficiency, thereby fully demonstrating its superiority in key node identification.

3.2.4. Network robustness

Network robustness is a critical metric for evaluating a network's ability to maintain functionality under attacks or failures. This experiment aims to verify the effectiveness of the SDGC method by analyzing the extent of structural damage caused by the progressive removal of key nodes identified by different methods. We adopt the largest connected component (LCC) as the primary indicator of network robustness. The LCC is the connected subgraph with the most nodes in the network, and its size reflects the network's ability to remain connected and functional after node removal. In the experiment, nodes are removed based on their importance rankings derived from various centrality methods, specifically targeting the top 10% to 30% of nodes, and the changes in LCC are observed. The results show that as key nodes are removed, the LCC gradually decreases, indicating damage to the network structure. A larger reduction in LCC signifies a more significant disruption to connectivity caused by the removed nodes, thereby reflecting the accuracy of the centrality method in identifying important nodes.

The experimental results are shown in Fig. 5. In the Jazz, Infectious, and Email networks, regardless of the method used to remove important nodes, the trend of changes in the largest connected component (LCC) is similar when the removal ratio is the same. This indicates that most nodes in these networks are concentrated in a single connected component, with uniform edge distribution, a lack of local high-density regions, and high average degree, resulting in strong node connectivity. Consequently, these networks maintain basic connectivity even after the removal of some key nodes. However, the nodes identified by the SDGC method cause the most significant damage to network robustness. In the Blogs network, the SDGC method performs particularly well, with the LCC decreasing significantly regardless of the removal ratio. This is attributed to the low clustering coefficient and weak local connectivity of the Blogs network, where key nodes between communities play a critical role. The SDGC method accurately identifies these nodes, leading to rapid network disintegration. In the Haggle and USAir networks, the LCC drops sharply when 20% of nodes are removed, with subsequent changes slowing down, possibly due to the strong scale-free properties of these networks, where many nodes rely on a few highly connected ones. By identifying high-influence nodes, the SDGC method significantly weakens robustness in these networks. In the Hamsterster network, the LCC varies significantly with different node removal methods, reflecting uneven edge distribution. Although its clustering coefficient is low and its local connectivity is moderate, implicit structural features allow it to maintain some connectivity after partial node removal. In the Facebook and GrQc networks, the SDGC method causes substantial LCC reductions across different removal ratios, demonstrating its pronounced impact on the robustness of large-scale networks. The performance of the SDGC method across multiple real-world networks demonstrates its effectiveness in identifying nodes critical to network connectivity, particularly in networks with weak local connectivity or prominent key nodes, where it outperforms other centrality methods.These observations collectively underscore that the variations in LCC decline across different networks are fundamentally rooted in their distinct structural character-

istics. Networks with uniformly distributed edges and high average degrees exhibit strong global connectivity, thereby remaining resilient under node removals. In contrast, networks with scale-free properties or weak local cohesion are more vulnerable, as the removal of a few strategically positioned nodes leads to rapid fragmentation. The consistently strong performance of the SDGC method across networks with such diverse topological profiles further validates its capability to pinpoint structurally critical nodes, especially in challenging settings where traditional centrality measures often fall short.

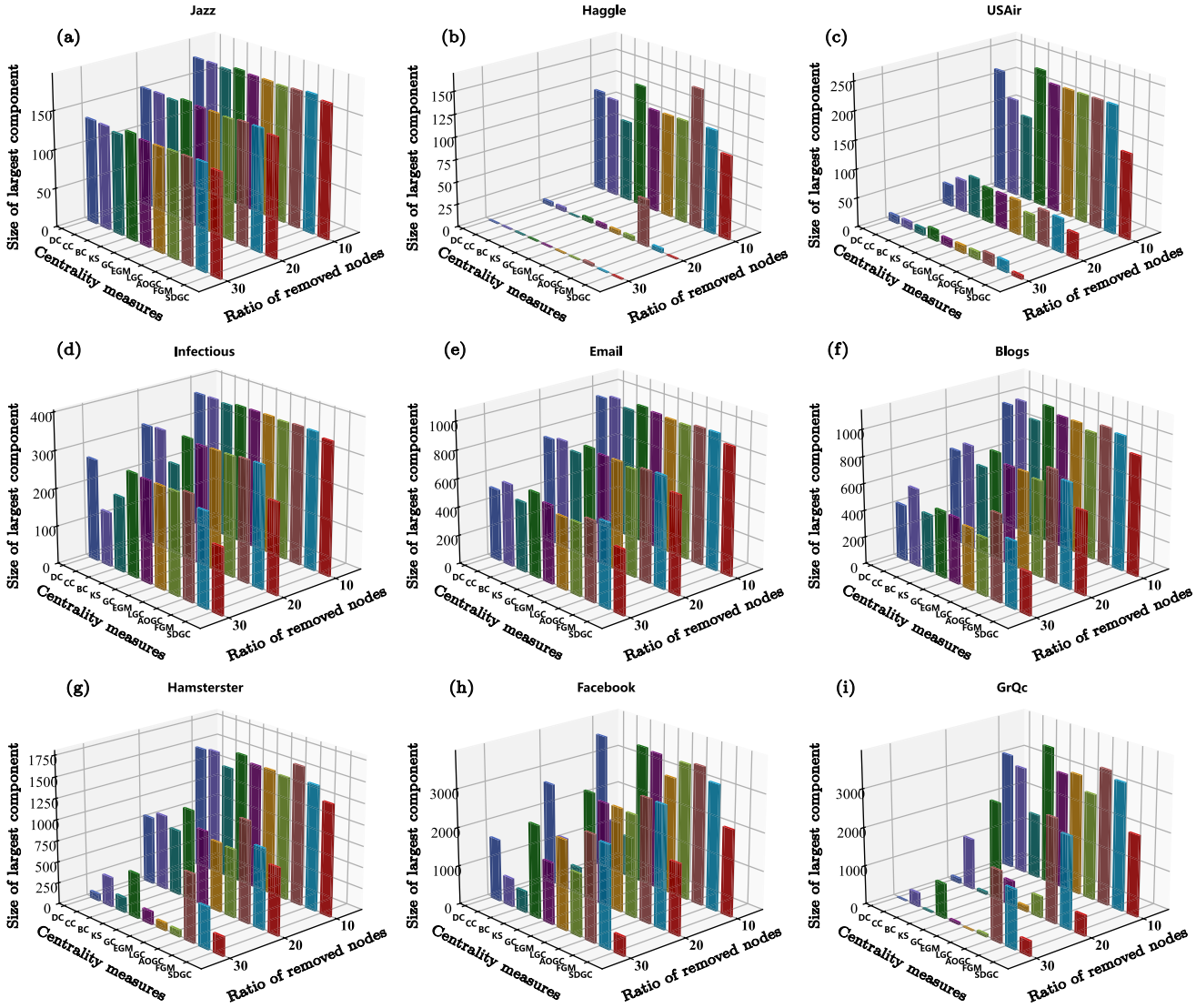


Fig. 5. The size of LCC in nine networks changes with the ratio of removed nodes.

3.2.5. Monotonicity experiment

To evaluate the ability of node centrality measurement methods to distinguish nodes, we adopt monotonicity as the evaluation metric, with its formula given as follows:[62]

$$M(R) = \left(1 - \frac{\sum_{r=1}^{\Phi} \mathrm{num}_r \times (\mathrm{num}_r - 1)}{N \times (N-1)}\right)^2, \tag{7}$$

where $M(R)$ denotes the monotonicity score of the method for node ranking, Φ represents the number of distinct ranks in the node centrality ranking sequence R (i.e., the number of elements in the sequence), num_r indicates the number of nodes assigned to the r-th rank in the ranking, and N is the total number of nodes in the network. The value of $M(R)$ ranges from [0, 1]. When $M(R) = 1$, each rank corresponds to exactly one node, indicating the strongest distinguishing ability of the method; when $M(R) = 0$, all nodes share the same rank, indicating the weakest distinguishing ability of the method.

Table 2 presents the monotonicity results of ten methods across nine real-world networks, with the highest monotonicity value for each network highlighted in bold. Compared to four classical methods, gravity-model-based advanced methods demonstrate superior monotonicity. Notably, our proposed SDGC method achieves monotonicity values of at least 0.98 across all networks, attaining the highest monotonicity in five networks and reaching 0.9999 in the Infectious, Email, and Facebook networks. Additionally, the FGM method has good

node distinguishing ability, achieving optimal results in six networks, though it is slightly outperformed by SDGC in the Infectious and Hamsterster networks. Overall, SDGC exhibits exceptional and stable performance across diverse real-world networks, highlighting its efficacy in providing precise node rankings for network analysis.

Table 2. Monotonicity of ten methods in nine real-world networks.

Network	DC	CC	BC	KS	GC	EGM	LGC	AOGC	FGM	SDGC
Jazz	0.9659	0.9878	0.9885	0.7944	0.9991	0.9995	0.9995	0.9995	**0.9997**	**0.9997**
Haggle	0.7255	0.9757	0.1225	0.6861	0.9836	0.9857	0.9887	0.9908	**0.9998**	0.9869
USAir	0.8586	0.9892	0.6969	0.8114	0.9933	0.9953	0.9933	0.9900	**0.9997**	0.9963
Infectious	0.9278	0.9967	0.9889	0.8182	0.9990	0.9969	0.9969	**0.9999**	0.9989	**0.9999**
Email	0.8874	0.9988	0.9400	0.8088	0.9977	0.9998	**0.9999**	0.9984	**0.9999**	**0.9999**
Blogs	0.9324	0.9980	0.9480	0.9059	0.9991	0.9993	0.9963	**0.9999**	**0.9999**	0.9993
Hamsterster	0.8979	0.9851	0.7123	0.8714	0.9843	0.9512	0.9853	**0.9857**	0.9825	**0.9857**
Facebook	0.9739	0.9967	0.9855	0.9419	0.9998	0.9998	**0.9999**	**0.9999**	**0.9999**	**0.9999**
GrQc	0.7916	0.9992	0.4850	0.6925	0.9994	**0.9996**	0.9994	**0.9996**	0.9994	0.9994

3.2.6. Ablation study

To rigorously examine the contributions of key components within the SDGC framework, we conducted a series of ablation studies in which each of the three fundamental elements — node mass, dynamic interaction distance, and the attraction coefficient — was independently removed. By comparing the performance of these ablated variants with that of the complete model, we quantitatively assessed the individual influence of each parameter on the overall effectiveness of SDGC. Specifically, we formulated the following three ablation configurations: (1) SDGC-wo-Mass (SDGC without node mass): This variant adopts the classical K-shell index as a proxy for node mass. (2) SDGC-wo-Distance (SDGC without dynamic interaction distance): Here, the dynamic interaction distance is replaced with the traditional shortest-path metric. (3) SDGC-wo-Attraction (SDGC without attraction

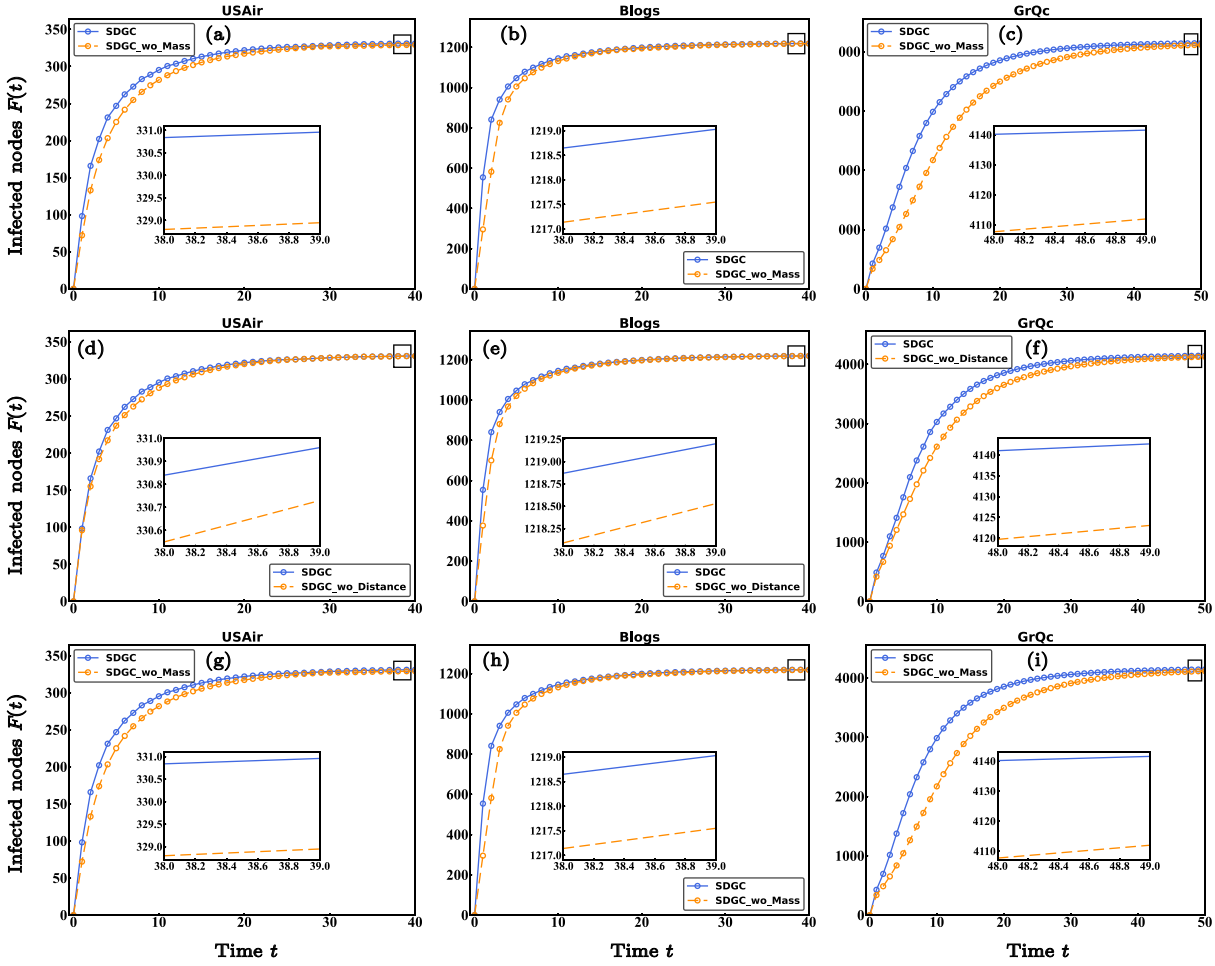


Fig. 6. The infection ability of the top 5% nodes selected by SDGC and three ablation variants in three networks.

coefficient): In this setting, the attraction coefficient is eliminated by fixing the gravitational constant to a unit value (set to 1 here). We conducted comparative experiments on three representative network datasets of varying scales: small-scale (USAir), medium-scale (Blogs), and large-scale (GrQc), using the SI model to evaluate the infection capacity of selected node sets. The experimental setup was fully consistent with that described in Subsection 3.2.2. The results are presented in Fig. 6.

First, in the SDGC-wo-Mass variant, node mass is solely determined by the traditional K-shell value. This limitation is particularly evident in the USAir network, which exhibits dense local connectivity and prominent hub nodes. Relying exclusively on the K-shell value results in a decline in overall propagation efficiency, indicating that our proposed definition of node mass more accurately identifies functional key nodes. Second, SDGC-wo-Distance adopts the conventional shortest path length to measure interaction distance between nodes, thereby disregarding the significant role of path multiplicity in the spreading process. This issue is especially pronounced in large-scale sparse networks such as GrQc. The observed performance degradation of this variant on the GrQc network further validates the critical importance of dynamic interaction distance in identifying influential nodes. Finally, SDGC-wo-Attraction treats the attraction coefficient in the gravitational model as a fixed constant, failing to capture variations in spreading intensity caused by structural hierarchical differences among nodes. In networks with pronounced hierarchical structures, such as Blogs, this variant is unable to accurately identify key nodes, resulting in a reduced spreading capacity of the selected node sets. In summary, these three core mechanisms exhibit distinct yet complementary functions across networks with diverse structural characteristics. The node mass definition enhances the identification of nodes with high spreading potential; dynamic interaction distance improves the fidelity of path modeling; and the attraction coefficient reinforces the representation of hierarchical information diffusion. Collectively, these design elements form the theoretical and empirical basis for the superior performance of the SDGC method in complex networks.

4. Conclusion

Identifying influential nodes in networks is crucial for both theoretical insights and practical applications. This study introduces the six-degree gravity centrality (SDGC) method, inspired by Newton's law of universal gravitation, to quantify internode interactions. SDGC integrates global, local, and semi-local node information, defining influence through a node's "mass" and "distance". Node mass encapsulates local neighborhood structures and global significance, reflecting a node's role within its community and the broader network. Unlike conventional shortest-path distance metrics, SDGC leverages the six-degree separation theory to capture latent connections, analyzing path distributions within six-degree neighborhoods for a more accurate assessment of interaction forces. Additionally, SDGC incorporates a gravitational constant adjusted by differences in node K-shell values, accounting for interaction dynamics across network hierarchies.

Compared to traditional centrality measures, which often overemphasize local or global information, SDGC offers a balanced approach, enabling precise quantification of influence. Comparative experiments on nine real-world network datasets, benchmarked against nine classical and state-of-the-art methods, demonstrate SDGC's superior performance in identifying highly influential nodes, particularly in social networks, where it shows strong discriminative power. The results also confirm SDGC's robustness and applicability across diverse network conditions, including large-scale complex networks.

Future work will focus on optimizing SDGC's computational efficiency for large-scale networks while maintaining accuracy. Additionally, integrating SDGC with existing centrality methods to refine node selection strategies will be explored. Extending SDGC's applicability to multilayer networks will also be investigated to broaden its scope.

Program availability

The source code is openly available in the Science Data Bank at https://doi.org/10.57760/sciencedb.27503 and on Github at https://github.com/nok42/SDGC-.git.

Acknowledgments

Project supported by the National Natural Science Foundation of China (Grant No. 62173065), the Natural Science Foundation of Beijing (Grant No. 4242040), the Intelligent Policing and National Security Risk Management Laboratory Open Topics for the year 2025 (Grant No. ZHKFYB2503), and the Intelligent Policing and National Security Risk Management Laboratory Open Topics for the year 2024 (Grant No. ZHKFZD2401). We thank the High Performance Computing Center of Southwest Petroleum University for performing the numerical calculations in this paper on their supercomputing system.

References

[1] De Souza R C, Figueiredo D R, Rocha A A de A and Ziviani A 2020 *Inf. Sci.* **514** 369
[2] Vermeulen R, Schymanski E L, Barabási A L and Miller G W 2020 *Science* **367** 392
[3] Pelusi D 2019 *IEEE Access* **7** 112171
[4] Havlin S, Kenett D Y, Ben-Jacob E, Bunde A, Cohen R, Hermann H, Kantelhardt J W, Kertész J, Kirkpatrick S, Kurths J, *et al.* 2012 *Eur. Phys. J. Spec. Top.* **214** 273
[5] Yang H H and An S 2020 *Symmetry* **12** 123
[6] Zhang Y, Wang Z Q and Xia C L 2010 *Proc. IEEE Int. Conf. Adv. Inf. Netw. Appl. Workshops* 644

[7] De Arruda G F, Barbieri A L, Rodriguez P M, Rodrigues F A, Moreno Y and Costa L F 2014 *Phys. Rev. E* **90** 032812
[8] Kumar S, Jha S and Rai S K 2020 *Indian J. Public Health* **64** 139
[9] Subbian K, Aggarwal C and Srivastava J 2016 *ACM Trans. Knowl. Discov. Data* **10** 1
[10] Rashid Y and Bhat J I 2024 *Multimed. Syst.* **30** 57
[11] Rezaei A A, Munoz J, Jalili M and Khayyam H 2023 *Expert Syst. Appl.* **214** 119086
[12] Kilkenny M and Nalbarte L 2020 *Reg. Res. Inst., West Virginia Univ.*
[13] Yuan B, Song T and Yao J 2024 *Proc. Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)* 11
[14] Aggarwal K and Arora A 2023 *Soc. Netw. Anal. Min.* **13** 146
[15] Liu P, Li L, Fang S and Yao Y 2021 *Chaos Solitons Fractals* **152** 111309
[16] Casero-Ripollès A 2021 *Sustainability* **13** 2851
[17] Bonacich P 1972 *J. Math. Sociol.* **2** 113
[18] Freeman L C 1977 *Sociometry* 35–41
[19] Freeman L C 2002 *Soc. Netw.: Crit. Concepts Sociol.* **1** 238
[20] Bonacich P and Lloyd P 2001 *Soc. Netw.* **23** 191
[21] Brin S and Page L 1998 *Comput. Netw. ISDN Syst.* **30** 107
[22] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 *Nat. Phys.* **6** 888
[23] Fatima U, Hina S and Wasif M 2023 *J. Comput. Sci.* **70** 102008
[24] Ullah A, Wang B, Sheng J, Long J, Khan N and Sun Z 2021 *Sci. Rep.* **11** 6173
[25] Yang Y, Wang X, Chen Y and Hu M 2020 *IEEE Access* **8** 32904
[26] Zhao J, Wang Y and Deng Y 2020 *Chaos Solitons Fractals* **133** 109637
[27] Zhong S, Zhang H and Deng Y 2022 *Inf. Sci.* **610** 994
[28] Xu G and Meng L 2023 *Chaos Solitons Fractals* **168** 113155
[29] Wen T and Deng Y 2020 *Inf. Sci.* **512** 549
[30] Zhang J, Zhang Q, Wu L and Zhang J 2022 *Entropy* **24** 293
[31] Yang P, Zhao L, Dong C, Xu G and Zhou L 2023 *Chin. Phys. B* **32** 058901
[32] Ai D, Liu X, Kang W, Li L, Lü S and Liu Y 2023 *Chin. Phys. B* **32** 118902
[33] Sheng J, Dai J, Wang B, Duan G, Long J, Zhang J, Guan K, Hu S, Chen L and Guan W 2020 *Phys. A* **541** 123262
[34] Qiu L, Zhang J and Tian X 2021 *Appl. Intell.* **51** 4394
[35] Hu H, Sun Z, Wang F, Zhang L and Wang G 2022 *Sci. Rep.* **12** 22506
[36] Mijia L, Hongquan W, Yingle L and Shuxin L 2021 *J. Phys. Conf. Ser.* **1738** 012026
[37] Su Z, Chen L, Ai J, Zheng Y and Bie N 2024 *Chin. Phys. B* **33** 058901
[38] Wang Y, Li H, Zhang L, Zhao L and Li W 2022 *Chaos Solitons Fractals* **162** 112513
[39] Yang P, Meng F, Zhao L and Zhou L 2023 *Chaos Solitons Fractals* **166** 112974
[40] Guo H, Wang S, Yan X and Zhang K 2024 *Chaos Solitons Fractals* **183** 114924
[41] Zhang Q, Shuai B and Lü M 2022 *Inf. Sci.* **618** 98
[42] Liu W, Lu P and Zhang T 2023 *IEEE Trans. Comput. Soc. Syst.* **11** 2105
[43] Zhang K, Zhou Y, Long H, Wang C, Hong H and Armaghan SM 2023 *J. King Saud Univ. - Comput. Inf. Sci.* **35** 101798
[44] Wang X, Slamu W, Guo W, Wang S and Ren Y 2022 *Chaos Solitons Fractals* **158** 112037
[45] Farahi Z, Kamandi A, Abedian R and Rocha L E C 2024 arXiv:2409.15142
[46] Milgram S and others 1967 *Psychol. Today* **2** 60
[47] Zhang J and Luo Y 2017 *Proc. 2nd Int. Conf. Model. Simul. Appl. Math. (MSAM2017)* 300
[48] Liu P, Xu Y and Deng C 2024 *J. Netw. Intell.* **9** February
[49] Newman M E J 2006 *Phys. Rev. E* **74** 036104
[50] Su J, Scott J, Hui P, Crowcroft J, De Lara E, Diot C, Goel A, Lim M H and Upton E 2007 *Proc. Int. Conf. Ubiquitous Comput.* 391
[51] Colizza V, Pastor-Satorras R and Vespignani A 2007 *Nat. Phys.* **3** 276
[52] Isella L, Stehlé J, Barrat A, Cattuto C, Pinton J F and Van den Broeck W 2011 *J. Theor. Biol.* **271** 166
[53] Guimera R, Danon L, Diaz-Guilera A, Giralt F and Arenas A 2003 *Phys. Rev. E* **68** 065103
[54] Adamic L A and Glance N 2005 *Proc. 3rd Int. Workshop Link Discov.* 36
[55] Kunegis J 2013 *Proc. 22nd Int. Conf. World Wide Web* 1343
[56] Mcauley J and Leskovec J 2014 *ACM Trans. Knowl. Discov. Data* **8** 1
[57] Leskovec J, Kleinberg J and Faloutsos C 2007 *ACM Trans. Knowl. Discov. Data* **1** 2
[58] Li Z, Ren T, Ma X, Liu S, Zhang Y and Zhou T 2019 *Sci. Rep.* **9** 8387
[59] Li S and Xiao F 2021 *Inf. Sci.* **578** 725
[60] Yang M, Chen G and Fu X 2011 *Phys. A Stat. Mech. Appl.* **390** 2408
[61] VragoviI, Louis E and Díaz-Guilera A 2005 *Phys. Rev. E* **71** 036122
[62] Lü L, Chen D, Ren X L, Zhang Q M, Zhang Y C and Zhou T 2016 *Phys. Rep.* **650** 1