Expert Systems With Applications 255 (2024) 124590

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

The role of link redundancy and structural heterogeneity in network disintegration

Bitao Dai [a,1], Jianhong Mou [a,1], Suoyi Tan [a], Mengsi Cai [a], Fredrik Liljeros [b,c], Xin Lu [a,*]

[a] College of Systems Engineering, National University of Defense Technology, Changsha 410073, China
[b] Department of Sociology, Stockholm University, Stockholm 10691, Sweden
[c] Department of Global Public Health, Karolinska Institutet, Stockholm 17177, Sweden

ARTICLE INFO

Keywords:
Influence maximization
Complex networks
Network disintegration
Network robustness
Network dismantling

ABSTRACT

While link redundancy has long been acknowledged as a critical factor in network robustness, current approaches frequently neglect the inherent heterogeneity of structure, thereby falling short of optimal network disintegration. This study introduces the novel Neighborhood Dissimilarity Community Heterogeneity (NDCH) framework, which systematically investigates redundancy elimination from a new perspective of neighborhood dissimilarity, while concurrently integrating structural heterogeneity derived from community structure. Extensive experiments on both synthetic and diverse real-world networks reveal that strategies developed under NDCH framework markedly outperform existing state-of-the-art methods in network disintegration, with performance improvements reaching up to 60.151% and 31.000% for Schneider R and the critical removal fraction f_c, respectively. Notably, additional analysis consistently indicates low correlations (Kendall's Tau < 0.6) and distinct value distributions between NDCH-based strategies and the state-of-the-art approaches. In summary, the novel framework underscores the significance of both redundancy and structural heterogeneity when devising network disintegration strategies, offering a substantial leap forward in enhancing network robustness against malicious attacks and various disruptions, especially in infrastructure networks.

1. Introduction

Over the past two decades, the study of network disintegration has garnered increasing attention due to its profound implications across various disciplines, such as curbing the spread of epidemics and rumors, enhancing the robustness of transportation systems, mitigating the impact of cyber-attacks, dismantling terrorist organizations, and managing ecological systems by controlling invasive species and preserving biodiversity (Doyle et al., 2005; Eiselt, 2018; Fan, Zeng, Feng, et al., 2020; Li et al., 2020; Lin et al., 2014; Tan et al., 2016; Wandelt et al., 2022). In recent years, the world has faced a continuous escalation of threats from terrorist organizations, exemplified by the September 11th attacks and the Madrid train bombings, which resulted in severe casualties and property damage. Simultaneously, the global health and security have been threatened by the spread of diseases such as H1N1 and COVID-19. The widespread dissemination of computer viruses such as Nimda and Melissa has compromised cybersecurity on an international level. These events collectively underscore the critical urgency of implementing effective interventions strategies in addressing global challenges (Akhtar et al., 2023; Chen & Carley, 2004; Gao et al., 2014; Nishi et al., 2020; Robins et al., 2023; Wandelt et al., 2022).

Network disintegration, also known as network dismantling, aimed at devising effective attack strategies that facilitate the rapid dismantlement of these hazardous networks. It seeks is to identify a sequence of nodes or links whose removal would significantly disrupt network connectivity, leading to structural destabilization and functional degradation (Albert et al., 2000; Wang et al., 2021; Wu et al., 2011). Previous research has demonstrated that network disintegration is a challenging combinatorial optimization problem classified as NP-hard, meaning it has no exact solution (Braunstein et al., 2016). Consequently, a vast breadth of approximation algorithms has been employed to tackle this issue. Currently, strategies for network disintegration are predominantly categorized into five types: mathematical programming methods, node (edge) centrality (structure-based methods), heuristic

* Corresponding author.
E-mail addresses: daibitao@nudt.edu.cn (B. Dai), moujianhong@nudt.edu.cn (J. Mou), tansuoyi_cn@outlook.com (S. Tan), caimengsi18@163.com (M. Cai), fredrik.liljeros@sociology.su.se (F. Liljeros), xin.lu.lab@outlook.com (X. Lu).
[1] These authors contribute equally.

https://doi.org/10.1016/j.eswa.2024.124590
Received 13 January 2024; Received in revised form 11 June 2024; Accepted 22 June 2024
Available online 24 June 2024
0957-4174/

algorithms, evolutionary algorithms, and machine learning techniques (Artime et al., 2024; Freitas et al., 2022a; Huang et al., 2024; Lou et al., 2023). For instance, Wu et al. have devised strategies for the optimal disintegration of complex networks, including undirected, directed, and multilayer networks by leveraging tabu search techniques (Deng et al., 2019; Deng et al., 2016; Qi et al., 2018). Concurrently, Deng et al. have pursued a distinct avenue, formulating the disintegration task as a constrained optimization problem and employing a genetic algorithm for its resolution (Deng et al., 2018). Wang et al. have proposed algorithms that simultaneously considers both structural topology and information diffusion processes, based on memetic algorithms, and emphasizes the role of edges in the disintegration process (Wang & Liu, 2023; Wang & Tan, 2023a, 2023b). Subsequently, Wang introduced multifactorial evolutionary algorithms for dismantling networks to simultaneously satisfy both targeted attacks and cascading failures in a multiobjective optimization framework (Wang, Ding, et al., 2023; Wang, Jin, et al., 2023; Wang et al., 2021). Helbin et al. and Deng et al. have explored cost-heterogeneous disintegration strategies that reflect real-world scenarios (Deng et al., 2018; Ren et al., 2019). Holme et al. have developed a dynamical framework that recurrently calculates the degree and betweenness centrality to guide the disintegration process (Holme et al., 2002). While these heuristic and evolutionary algorithms can provide relatively satisfactory performance, they encounter scalability issues when applied to large networks.

To surmount this limitation, researchers have turned to structure-based strategies, such as *k*-core decomposition (Kitsak et al., 2010), which are more scalable and effective for large-scale networks. Recently, with the rapid development of deep learning, Fan et al. pioneered a deep reinforcement learning algorithm tailored for network disintegration, while Grassia et al. have devised a graph neural network-based strategy for the same purpose. (Fan, Zeng, Sun, et al., 2020). Both use geometric deep learning to learn an attack strategy on small synthetic networks. While machine learning-based disintegration strategies excel in computational efficiency and generalization capabilities, they present a significant drawback in terms of sacrificing interpretability, which is a critical consideration in practical implementation.

Among the various disintegration strategies proposed, structure-based methods have emerged as particularly effective due to their interpretability and applicability for large-scale networks. These methods rely on the underlying structure of the network to guide the disintegration process, recognizing that the way nodes are connected can significantly influence the strategy's success. Recent studies have particularly highlighted the impact of redundant ties, defined as links that forge multiple alternative paths between nodes, on network robustness (Freitas et al., 2022b; Tan et al., 2019). Redundant ties are a natural outcome of the complex interactions within networks, giving rise to a variety of intricate structures such as community structures, cycles, cliques, and simplices (Shi et al., 2021; Sizemore et al., 2018). These structures significantly influence the network's functionalities, including information dissemination, synchronization, and robustness. Specifically, increasing redundant ties through link reassignment or addition enhances network robustness by providing alternative routes for information flow and other critical functions. However, while increased redundancy can bolster robustness, it can also obstruct the efficacy of disintegration strategies. Redundant ties precipitate the emergence of nodes in close interaction, resulting in tightly coupled local clusters. Within these clusters, nodes often exhibit extensive neighbor-sharing and overlapping spheres of influence, complicating the assessment of their individual significance (Dai et al., 2023; Liu et al., 2015a; Nian et al., 2017). Traditional structure-based methods, such as those that rely on measures like degree, betweenness, and *k*-core, lose their efficacy in network with numerous redundant ties (Bonacich, 1972; Brandes, 2008; Kitsak et al., 2010). For instance, a node with few direct neighbors but highly influential ones may exert a greater impact than a node with abundant immediate neighbors. Conversely, some nodes are tightly interconnected, forming pseudo-cores where they remain within the same core and exhibit significant core value (Liu et al., 2015b, 2015c). These redundant ties create locally dense, core-like structures, potentially distorting the identification of influential nodes when using traditional metrics.

To navigate the complexities introduced by redundant ties in network disintegration, researchers have developed innovative approaches that capture the nuanced effects of redundant ties on node importance. For instance, Chen et al. devised the Cluster Rank, which considers the side effects of clustering coefficients when assessing node importance (Chen et al., 2013). Fan et al. developed the Cycle Ratio, incorporating high-order information to account for the impact of redundant ties (Fan et al., 2021). Dai et al. formulated the Multi-Spanning Tree-based Degree Centrality (MSTDC) measure, which aggregates the degrees of spanning trees constructed with randomly selected root nodes, to mitigate the side effect arising from link redundancy (Dai et al., 2023). Bianconi et al. introduced a simplices-based approach to examine the impact of redundant ties on network connectivity, while Shi et al. proposed the clique vector-space framework (Millán et al., 2020; Shi et al., 2019). These studies have advanced our understanding of the roles of redundant ties in network robustness; however, a common oversight persists in that they generally treat nodes and links with equal weight, neglecting the distinct contributions of different connections. This issue becomes particularly pronounced in scenarios where a bridge node connects disparate communities, such as in linking a neighbor in closely-knit community A with another in distinct community B. In such cases, it is clear that cross-community links, and the nodes they connect, carry more strategic weight than their intra-community counterparts (Bouyer et al., 2023; D'Souza et al., 2023). This issue is further exacerbated by the intricate community structures that abundant redundant ties can foster, amplifying the heterogeneity between nodes and links (D'Souza & Mitzenmacher, 2010; Wandelt et al., 2022; Wandelt et al., 2021). Yet, contemporary approaches rarely address the dual challenges of link redundancy and structural heterogeneity in an integrated manner.

Our research seeks to address this gap by designing a novel framework, the Neighborhood Dissimilarity Community Heterogeneity (NDCH), and developing three approaches within the framework to systematically investigate the integrated effects of redundancy elimination, from the perspective of neighborhood dissimilarity, and structural heterogeneity, quantified from weightings of belonging to different communities, as well as the role of information aggregation methods. Specifically, we recognize that links bridging different communities hold greater importance compared to intra-community connections. Redundant ties in a network contribute to a complex community structure, which can be leveraged to identify and target specific communities for disintegration strategies. Furthermore, this sophisticated community structure amplifies the disparities among nodes and links. So, we quantify the structural heterogeneity through community structure. While our focus here is rooted in community structures, it is worth noting that these structures serve as exemplars, and the framework is adaptable enough to incorporate other structures for heterogeneity assessment.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive introduction to the NDCH framework, including a description of the synthetic and empirical network data employed in our experiments. Section 3 delves into the robustness measures employed, discusses competing methods, and outlines the specific parameter settings used in our experiments. Section 4 presents disintegration results on both synthetic and empirical networks, comparing performance in terms of Schneider R and the critical removal fraction f_c against a range of state-of-the-art approaches. Finally, Section 5 concludes by summarizing the principal contributions of this study and discussing potential directions for future research.

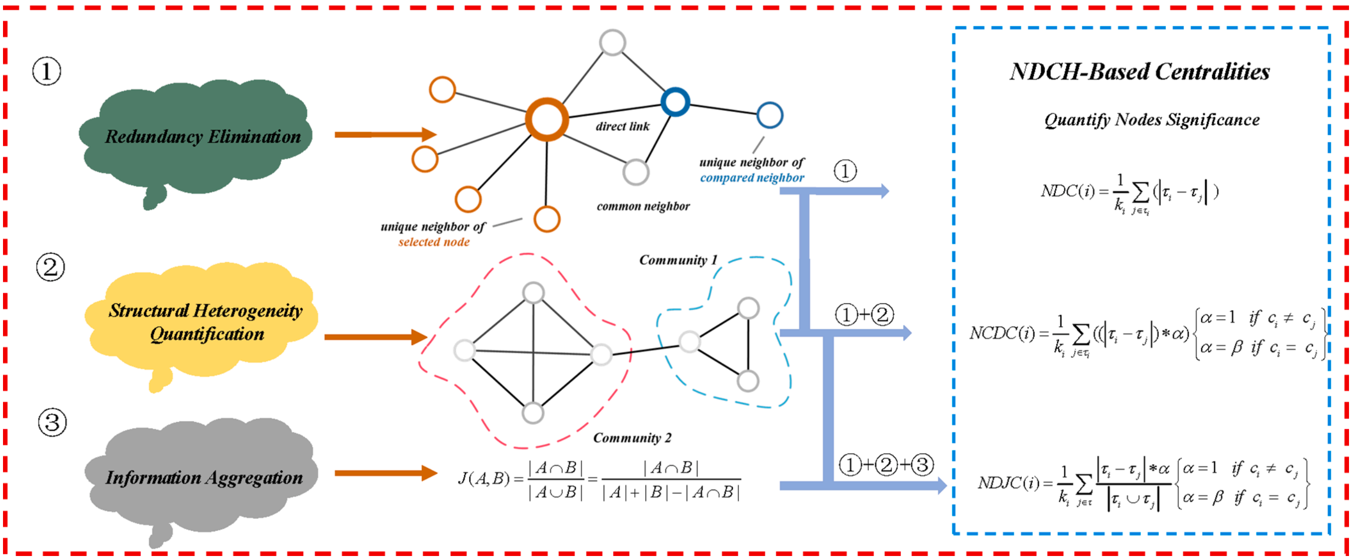

Fig. 1. Neighborhood Dissimilarity Community Heterogeneity Framework.

2. Data and method

In this section, we introduce the Neighborhood Dissimilarity Community Heterogeneity (NDCH) framework that innovatively integrates considerations of redundancy and structural heterogeneity, as well as the information aggregation methods. Furthermore, we provide a detailed exposition of the synthetic and empirical network data sets that we employ in our research. These data sets serve as the testing grounds where we rigorously evaluate the performance and efficacy of the NDCH framework, and they encompass a diverse range of network types and scales to ensure a comprehensive assessment.

2.1. Neighborhood dissimilarity community heterogeneity framework

As depicted in Fig. 1, the Neighborhood Dissimilarity Community Heterogeneity (NDCH) framework is comprised of three core components: redundancy elimination, structural heterogeneity quantification, and information aggregation methods, which are presented below.

Redundancy elimination. To eliminate redundancy and accurately assess node importance, we quantify neighborhood dissimilarity by focusing on the uniqueness of a node's neighbors. Instead of merely considering direct connections, we evaluate the distinct neighbors a node has compared to its immediate surrounding nodes. Specifically, we calculate the number of neighbors that are not shared with the node's immediate neighbors. This approach enables us to quantify a node's unique influence and contributions to connectivity relative to its neighbors, thereby clarifying its individual significance amidst the overlapping influences commonly found in densely interconnected local clusters (communities). For a precise mathematical expression of this method, refer to Equation 1.

Structural heterogeneity quantification. We consider network elements as of different weight and quantify structural heterogeneity from the perspective of community structure. Specifically, we posit that links connecting different communities are of paramount importance compared to those solely within a single community. Redundant ties within a network can foster complex community structure, which can be tapped into to identify and target specific communities for strategic disintegration. Additionally, this intricate community structure amplifies disparities among nodes and links. While our emphasis is primarily on community structures, it is essential to mention that they merely exemplify the capabilities of our framework. Indeed, our framework can readily adapt to assess other structures for heterogeneity. For a more in-depth understanding, refer to Equation 2.

Information aggregation. Information aggregation methods facilities a more coherent and comprehensive way to incorporate redundancy elimination and structural heterogeneity quantification, resulting in a synthesized importance score for node ranking. These techniques aim at capturing the inherent patterns, relationships, and structures within the data, offering a summative view that is crucial for decision-making processes. While various techniques for information aggregation exist including prominent ones like information entropy, we specifically employ the Jaccard similarity as an illustrative example to showcase the concept. It is crucial to acknowledge that the Jaccard similarity may not be the optimal approach for information aggregation. For a thorough understanding, please refer to Equation 2 and Equation 3.

Within the NDCH framework, we have developed three approaches *NDC*, *NCDC*, *NDJC*, could be found in Section 2.2 to conduct a systematic ablation study to investigate the roles of different components in comparison to the state-of-the-arts (SOTA) approaches. *NDC* focuses solely on the redundancy elimination aspect of the NDCH framework, rendering it the most fundamental approach. *NCDC* combines redundancy elimination with structural heterogeneity quantification through community structure. In contrast, *NDJC* explores the impact of the third component of the NDCH framework, information aggregation, employing diverse form of information aggregation. These three metrics are progressively compared. Additionally, we systematically explore the impact of four community detection algorithms on *NCDC* and *NDJC* in Supplementary Note 3. This comprehensive framework is designed to reveal the potential superiority of redundancy elimination, structural heterogeneity, and information aggregation methods in network disintegration.

2.2. NDCH-based centralities

We initiate our exploration by introducing the cornerstone of redundancy elimination. Redundant ties intricately weave multiple paths, contributing to the emergence of densely interconnected local clusters. These clusters consist of nodes that frequently share neighbors, resulting in overlapping influence. This overlap makes evaluating the distinct significance of individual nodes challenging. To address this issue, we present a novel concept—neighborhood dissimilarity. This concept aims to counteract the overarching influence of redundancy arising from neighborhood similarity. It specifically concentrates on quantifying the uniqueness of a node's neighbors compared to its immediate neighbors. Therefore, we present the fundamental approach based on NDCH framework, known as Neighborhood Dissimilarity

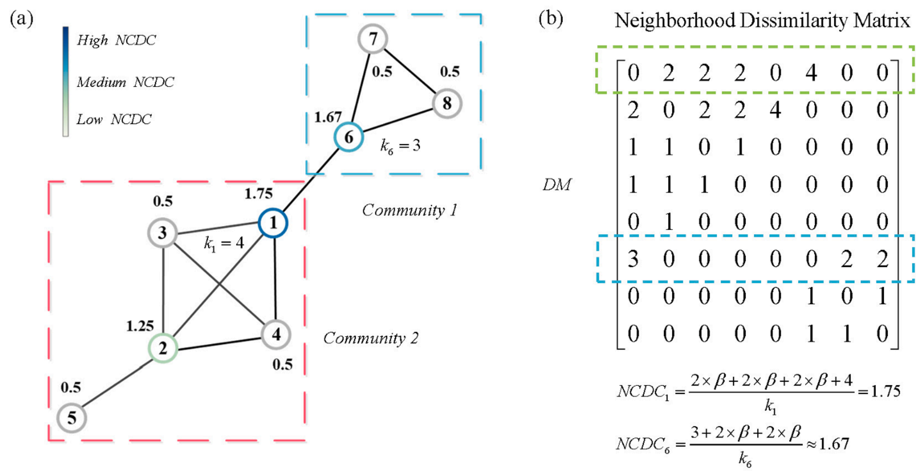

Fig. 2. Neighborhood Community Dissimilarity Centrality in an example network. (a) An example network with *NCDC* of nodes, the dotted lines represent different communities. (b) The neighborhood dissimilarity matrix of the example network in (a) and the calculations of $NCDC_1$ and $NCDC_6$ with $\beta = 0.5$.

Centrality (*NDC*), which is calculated as follows:

$$NDC(i) = \frac{1}{k_i}\sum_{j\in\tau_i}(|\tau_i - \tau_j|) \tag{1}$$

where τ_i and τ_j denotes the set of neighbor for i and j respectively, and $|\tau_i - \tau_j|$ counts the number of nodes in τ_i but not in τ_j. For example in Fig. 2(a), $\tau_1 = \{v_2, v_3, v_4, v_6\}$ and $\tau_2 = \{v_1, v_3, v_4, v_5\}$, so the neighborhood dissimilarity set $|\tau_1 - \tau_2| = |\{v_2, v_6\}| = 2$, where $\{v_2\}$ indicates that there is a direct link between v_1 and v_2. Based on Equation 1, we could construct the neighborhood dissimilarity matrix *DM*, where $DM_{ij} = |\tau_i - \tau_j|$ as shown in Fig. 2(b).

Building upon our initial approach to eliminate the confounding effects of redundant ties, we further incorporate the consideration of another core component of NDCH framework: structural heterogeneity introduced by community structure. Connections spanning different communities play a pivotal role in enhancing network connectivity and facilitating information dissemination. We demonstrate that nodes within the same community tend to have localized impact, whereas nodes connecting different communities enable global information spreading. To address this, we penalize links within the same community and improve *NDC* by designing weighted contributions from each connected node in the neighborhood dissimilarity set. We refer to this modified approach as Neighborhood Community Dissimilarity Centrality (*NCDC*), which can be calculated using the following formula:

$$NCDC(i) = \frac{1}{k_i}\sum_{j\in\tau_i}((|\tau_i - \tau_j|) * \alpha)\left\{\begin{array}{lll}\alpha = 1 & if & c_i \neq c_j \\ \alpha = \beta & if & c_i = c_j\end{array}\right\} \tag{2}$$

in which α accounts for the impact of links belonging to different communities, c_i indicates the community label where node i locates, and $0\leqslant\beta\leqslant1$ is a constant penalty coefficient. With this definition, the contribution of nodal importance to *NCDC* is higher for cross-community links compared to those within the same community. Fig. 2(b) provides a calculation example for v_1 and v_6 with the corresponding value of $\beta = 0.5$ to reflect the structural heterogeneity supplied by the community structure.

In this study, while we employ a specific function i.e., constant penalty coefficient for illustrative purposes, the framework we propose is inherently versatile. Alternative complex mathematical functions, such as exponential functions, can be seamlessly integrated into this framework based on the specific characteristics and needs of the analysis. Furthermore, the community detection algorithm is a variable component that can be adjusted to suit specific circumstances for better applicability in real-world scenarios. This adaptability ensures our approach remains responsive and effective under different network conditions.

Expanding our investigation based on community structure, we explore the final core component of the NDCH framework: information aggregation methods. While a variety of aggregation techniques exists, with notable mentions such as information entropy, our study specifically utilizes the Jaccard similarity to elucidate the principle. However, it is imperative to recognize that, despite its illustrative utility, the Jaccard similarity may not always be the quintessential method for every aggregation scenario. Consequently, we introduce an adjusted method referred to as Neighborhood Community Dissimilarity Jaccard Centrality (*NDJC*) and define it as follows:

$$NDJC(i) = \frac{1}{k_i}\sum_{j\in\tau_i}\frac{|\tau_i - \tau_j| * \alpha}{|\tau_i \cup \tau_j|}\left\{\begin{array}{lll}\alpha = 1 & if & c_i \neq c_j \\ \alpha = \beta & if & c_i = c_j\end{array}\right\} \tag{3}$$

To demonstrate the superiority of our approaches, we employ state-of-the-art techniques outlined in the "Competing Methods" section to calculate node centralities for the network depicted in Fig. 2. Suppose a virus starts from an arbitrary node in Fig. 2(a), it will transmit to its neighbors in the subsequent step. We aim to identify the most critical

Table 1
The values of nodes based on different approaches.

Node	*Cycle Ratio*	*Information Entropy*	*Cluster Rank*	*Leverage Centrality*	*MSTDC*	*NDC*	*NCDC*	*Average infection size* $\bar{I}$
1	3.00	1.39	1.26	0.11	3.00	2.50	1.75	3.57
2	3.00	1.39	1.26	0.22	2.52	2.50	1.25	5.29
3	3.00	1.10	0.30	−0.10	1.22	1.00	0.50	7
4	3.00	1.10	0.30	−0.10	1.22	1.00	0.50	7
5	0.00	0.00	1.00	−0.60	1.00	1.00	0.50	7
6	3.00	1.10	1.39	0.09	2.78	2.33	1.67	4.14
7	3.00	0.69	0.20	−0.10	1.11	1.00	0.50	7
8	3.00	0.69	0.20	−0.10	1.11	1.00	0.50	7

nodes for blocking the transmission. Intuitively, if any node except v_1, v_2, and v_6 is immunized, the virus can still spread globally, irrespective of which node is infected, yielding an average infection size of $\bar{I} = 7$ nodes. However, if v_1 is vaccinated, the virus is contained within a local spread, resulting in an infection size of $\bar{I} = 4$ when initiated from $\{v_2, v_3, v_4, v_5\}$ or $\bar{I} = 3$ when initiated from $\{v_6, v_7, v_8\}$. Hence, immunizing v_1 leads to an expected infection size of approximately $\bar{I} = 3.57$, as calculated by $(4 \times 4 + 3 \times 3)/7$. Similarly, the expected infection sizes are approximately 5.29 and 4.14 when v_2 and v_6 are immunized, respectively. Consequently, v_1, v_6, and v_2 emerge as the most influential nodes in terms of blocking transmission.

Upon comparing the results in Table 1, we observe that only *NCDC* accurately reflects this ranking. Cycle Ratio, Information Entropy, Cluster Rank, and Leverage Centrality perform poorly and even yield counterintuitive outcomes. MSTDC, while identifying the top three nodes effectively, falls short in ranking the remaining nodes. This outcome underscores the critical importance of considering structural heterogeneity in network disintegration. We specifically observed the existence of two distinct communities in the network, with v_1 having connections to the other community, while v_2 primarily connects within its own community. For example, node 6 demonstrates a relatively large dissimilarity in its node set compared to node 2. In subsequent sections, we will rigorously evaluate our NDCH-based approaches by assessing their relative effectiveness across both synthetic and empirical networks.

2.3. Complexity analysis

The pseudocode of *NDJC* calculation is outlined in Algorithm 1, others could be found in Supplementary Note 2. The complexity of *NDJC* is primarily driven by the community detection algorithm and the nested loops which compute the dissimilarity between pairs of neighbors. The complexity of the community detection algorithm varies from practical reality. For instance, Louvain algorithm has a complexity of $O(|E|\log|V|)$. Regarding nested loops, initializing the *NDJC* dictionary takes $O(|V|)$, and for reach node, iterating over its neighbors involves a worst-case complexity of $O(|\Delta|)$, where Δ is the maximum degree of the graph and is dramatically smaller than $|V|$. Therefore, the computational complexity of *NDJC* is $O(|E|\log|V|) + O(|V|\cdot\Delta^2) \approx O(|N|)$, enabling it sufficiently adaptable for large-scale networks. The space complexity is $O(|V| + \Delta) \approx O(|N|)$, accounting for the storage of *NDJC* scores and temporary neighbor sets. This makes the method efficient in terms of memory usage.

Algorithm 1: *NDJC* calculation

Input: Graph $G = (V, E)$, penalty coefficient β, community detection algorithm.
Output: Ranking node list.
Initialize an empty dictionary *dic* with keys as node IDs and values set to zero:
  $dic \leftarrow \{0|i \in V\}$.
Execute specific community detection algorithm to obtain node communities:
  $c \leftarrow \{c_i|i \in V\}$.
For each node $i \in V$ do:
  $\tau_i \leftarrow$ Set of neighbors of node i.
    For each neighbor $j \in \tau_i$ do:
      $\tau_j \leftarrow$ Set of neighbors of node j.
      If $c_i = c_j$ then:
        $\theta = |\tau_i - \tau_j|/|\tau_i \cup \tau_j|^*\beta$.
      Else:
        $\theta = |\tau_i - \tau_j|/|\tau_i \cup \tau_j|^*1 = |\tau_i - \tau_j|/|\tau_i \cup \tau_j|$.
      $dic[i] = dic[i] + \theta$.
    End for.
End for.
For each node $i \in V$ do:
    $dic[i] = dic[i]/k_i$.
End for.
Ranking *dic* by value.
Return index list of *dic*.

Table 2
Basic statistics of empirical networks.

Networks	*N*	*M*	*Diameter*	$\bar{k}$	*C*	$\bar{l}$
Birmingham	14,578	20,913	135	2.87	0.15	41.72
Philadelphia	13,389	21,246	97	3.17	0.02	41.89
Chicago	12,979	20,627	106	3.18	0.04	41.97
Power_Grid	4,941	6,594	46	2.67	0.10	18.99
Parsed_Jazz	198	2,742	6	27.70	0.62	2.34
Parsed_Tap	1,373	6,833	12	9.95	0.53	5.22
Netscience	379	914	17	4.82	0.74	6.04
Openflights	1,485	13,897	9	18.72	0.48	3.15
Facebook	4,039	88,234	8	43.69	0.61	3.69
Circuit	252	399	13	3.17	0.06	5.81
Hamster	2,000	16,098	10	16.10	0.54	3.59
Collins_Yeast	1,004	8,319	15	16.57	0.65	5.55

Note: *N*: network size, *M*: number of links, $\bar{k}$: average degree, *C*: clustering coefficient, and $\bar{l}$: average shortest path length.

2.4. Network data

Synthetic Networks. To examine the effects of redundancy, we employed the Powerlaw-Cluster (PLC) model to generate synthetic networks with power-law degree distribution and adjustable clustering coefficients (Holme & Kim, 2002). The PLC model expands upon the Barabási-Albert (BA) growth model by incorporating an additional step (Albert & Barabási, 2002). After each random link is added, there is a chance of creating a link to one of its neighbors, thereby introducing triangles between nodes and amplifying redundant connections. This modification allows for a higher average clustering coefficient to be achieved compared to the BA model. The model's parameters can be tuned, including the number of nodes *N*, the number of random links added for each new node *m* (half the average degree), and the probability of adding a triangle after adding a random link *p*, which adjusts the average clustering coefficient). We generated 100 networks for each parameter setting (see Table S1 for further details).

Empirical Networks. To capture the intricacies of real-world systems, we additionally conduct experiments on twelve distinct networks spanning various domains including infrastructure networks, collaboration networks, protein–protein interaction (PPI) networks, and social networks. Specifically, the networks are as follows: Birmingham is a transportation network including roads, railways, and public transit facilitating travel within Birmingham, England. Philadelphia is a transportation network that encompasses the metropolitan area of Philadelphia, Pennsylvania, and the broader Delaware Valley region, featuring a complex interplay of freeways, parkways, and other transportation arteries. Chicago is a transportation network detailed representation of the transportation infrastructure in the Chicago region, supporting travel and connectivity throughout the city and its surrounding areas (Wang et al., 2024). Power_Grid is a power grid network located in the Western States of the United States of America, comprising generators, transformers, and substations (Watts & Strogatz, 1998). Parsed_Jazz represents a collaboration network among Jazz musicians and bands from 1912 to 1940 (Gleiser & Danon, 2003). Netscience is a co-authorship network involving scientists working on network science since 2006 (Newman, 2006). Parsed_Tap is a yeast protein–protein binding network generated through tandem affinity purification experiments (Gavin et al., 2006). Openflights is a network depicting regularly scheduled flights between airports across the globe (Openflights: Network data from openflights.org). Facebook captures an ego-social network derived from Facebook users (Leskovec & Mcauley, 2012). Circuit, similar to electronic circuits, comprises cyclic structures, signal propagation pathways, and feedback mechanisms (Milo et al., 2002). In biological systems, it plays a crucial role in processes such as cellular signaling and gene regulation. Hamster network is a friendship and family connections network among users of a website (Kunegis, 2014). It represents the relationships and connections between individuals on the platform. Lastly, Collins_Yeast portrays a network of protein–protein interactions in Saccharomyces cerevisiae (budding yeast), measured through high-throughput affinity purification and mass spectrometry to identify co-complex associations (Collins et al., 2006).

All networks are treated as undirected and unweighted. We extract the giant connected components of each network for detailed analysis. Summary statistics for these components are provided in Table 2.

3. Experimental setting

In order to evaluate the optimal sequence for significantly degrading network connectivity through node removal, we utilize a network disintegration approach. The relative size of the giant connected component (s) serve as a critical measure of network connectivity. As nodes are incrementally removed, the network's connectivity gradually decreases until it reaches a state of complete disconnection, referred to as network collapse. To assess the network's ability against total collapse, we employ the critical removal fraction f_c as an indicator (Albert et al., 2000; Callaway et al., 2000). We record the number of nodes N_r, that are removed up until $s < \frac{\sqrt{N}}{N}$, indicating near-total disconnection and severely limited spreading capability. The critical threshold, f_c, is calculated as:

$$f_c = \frac{N_r}{N} \tag{4}$$

A smaller f_c value signifies a more efficient disintegration strategy.

Considering that the network is severely damaged but not completely collapsed, we employ Schneider R to represent the network's ability to withstand any level of attack intensity and capture its response to the disintegration throughout the entire process (Schneider et al., 2011):

$$R = \frac{1}{N}\sum_{Q=1}^{N} s(Q) \tag{5}$$

in which $s(Q)$ is the remaining GCC size after removing Q nodes. R can be quantified by calculating the area under the robustness curve, where the horizontal axis denotes the fraction of nodes removed, and the vertical axis represents s.

To underscore the superior performance of our approaches, we visualize the differences in performance between our methods and the leading benchmark methods. We quantify this using the promotion ratio φ, defined as the ratio of improvement achieved by our approaches relative to the benchmarks. This comparison is conducted across the evaluation of f_c and R, calculated as follows:

$$\begin{aligned} \varphi^{fc} &= \frac{\left(f_c^{best} benchmarks - f_c^{best} our\ methods\right)}{f_c^{best}\ benchmarks} \times 100\% \\ \varphi^{R} &= \frac{\left(R_c^{best} benchmarks - R_c^{best} our\ methods\ \right)}{R_c^{best} benchmarks} \times 100\% \end{aligned} \tag{6}$$

To systematically evaluate network disintegration, we implement different node removal strategies across varied network scale. For networks with fewer than 1000, we remove nodes incrementally. For larger networks, we removed 1% of the total number of nodes in each step. This ensures a consistent and adaptive strategy that accounts for the varying sizes of the networks, enhancing the overall cohesion and effectiveness of the evaluation process.

3.1. Competing methods

We incorporate a series of the state-of-arts methods on redundancy for comparison. These approaches can be categorized as degree-based centralities (including Leverage Centrality, Cluster Rank, and Information Entropy) and path-based centrality (including Cycle Ratio and MSTDC). It is worth noting that these methods have demonstrated superior performance compared to traditional metrics such as degree, betweenness, and closeness centrality (Chen et al., 2013; Dai et al., 2023; Fan et al., 2021; Joyce et al., 2010; Xu et al., 2020). Therefore, we do not include the traditional metrics in our comparison to avoid repetition for the sake of simplification.

Leverage Centrality considers the extent of connectivity of a node relative to its neighbors. A node with a higher degree among its neighbors is likely to have high leverage centrality (Joyce et al., 2010). The leverage centrality of a node is defined as follows:

$$c_L(i) = \frac{1}{k_i}\sum_{j\in\tau_i}\frac{k_i - k_j}{k_i + k_j} \tag{7}$$

in which, k_i denotes the degree of v_i.

Cluster Rank quantifies the node influence by considering both the clustering coefficient and neighbors' influence (Chen et al., 2013). It reflects the impacts of redundant ties in terms of nodes importance. The cluster rank of a node is defined as follows:

$$s_i = f(c_i)\sum_{j\in\tau_i}(k_j + 1) \tag{8}$$

where c_i is the clustering coefficient of v_i, $f(c_i)$ accounts for the effect of local clustering to the overall cluster rank calculation. Additionally, the term $'+1'$ denotes the contributions of node j itself to the cluster rank. In this study, we adopt $f(c_i) = 10^{-c_i}$ the same as Chen (Chen et al., 2013).

Information Entropy ranks nodes by aggregating their neighbors' information based on entropy theory (Xu et al., 2020). It reflects the average information and excludes redundant information provided by neighbor nodes, which is defined as:

$$E_i = -\sum_{j\in\tau_i}(P_{ij}\log_2 P_{ij}) \tag{9}$$

where $P_{ij} = k_i/A_j$, and $A_i = \sum_{j\in\tau_i} k_j$ i.e., the second-order degree.

Multi-Spanning Tree-based degree centrality (MSTDC) measures node importance by eliminating redundant ties and local coupling (Dai et al., 2023). It aggregates the degrees from spanning trees constructed with a few randomly selected root nodes, which is defined as:

$$MSTDC(i) = \frac{\sum_{r\in\sigma}\sum_{j\in\tau_i} a_{ij}^{(r)}}{T} \tag{10}$$

where r represents the root node of one spanning tree, σ represents the randomly chosen node-set, and T implies the number of spanning trees. In this study, we adopt $T = 30$ the same as Dai (Dai et al., 2023).

Cycle Ratio evaluates the efficacy of redundant ties considering the information of cycle structure (Fan et al., 2021). It defines the cycle number matrix $C = [c_{ij}]_{N\times N}$ to characterize the shortest cycles and is calculated by:

$$r_i = \begin{cases} 0, c_{ii} = 0 \\ \sum_{j, c_{ij}>0}\frac{c_{ij}}{c_{jj}}, c_{ii} > 0. \end{cases} \tag{11}$$

where c_{ii} is the number of cycles in the set of all shortest cycles that contain v_i, and c_{ij} denotes the number of cycles in the set of all shortest cycles that pass through both v_i and v_j.

3.2. Parameters setting

This study involves two adjustable components (the community detection algorithms and the penalty coefficient beta) controlling the aggregation weight on *NCDC* and *NDJC*. First, as the results of community detection algorithms could be various, we adopt the Greedy

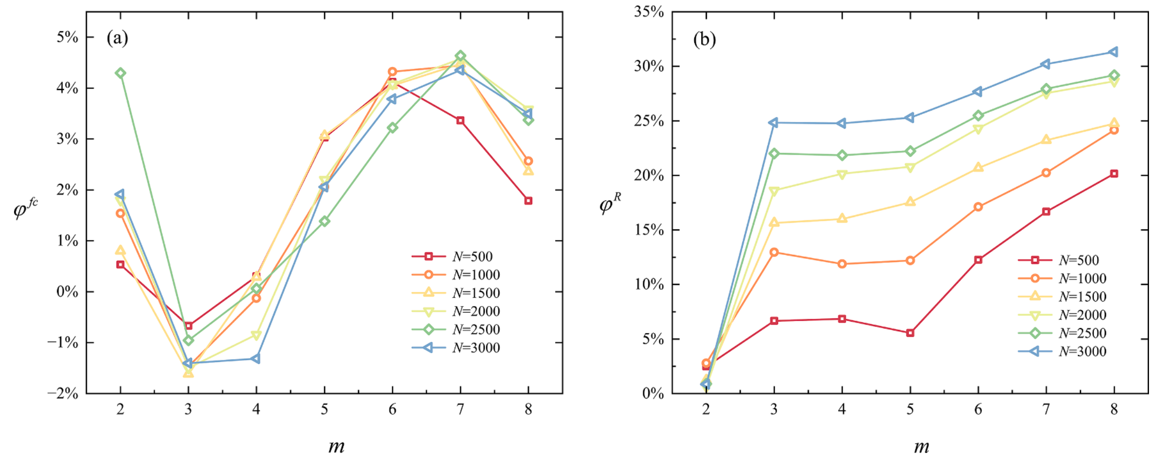

Fig. 3. Comparison of Promotion Ratio Curves in PLC Network. (a) The promotion ratio with respect to f_c. (b) The promotion ratio with respect to R. Each data marker corresponds to an average over 100 independent realizations.

Modularity community detection algorithm in synthetic networks. Because its result remains constant. Additionally, for empirical networks, we compare the efficiency of Greedy Modularity, Infomap, Louvain and Label Propagation community detection algorithms. Second, regarding the penalty coefficient β, we explore values ranging from 0 to 1 in intervals of 0.05. Finally, we conduct comprehensive sensitivity analyses on the combinations of β and the community detection algorithms in the Supplementary Note 3 and report the optimal values on f_c and R in Section 4.

4. Results

In this section, we present a comprehensive evaluation of the NDCH framework. We start with a thorough comparison of our three introduced approaches against five state-of-the-art (SOTA) methods, utilizing both synthetic and empirical networks as testbeds. Subsequently, we delve into the influence of network structure on the performance of our proposed approaches in synthetic networks. To further substantiate the evaluation, we conduct a correlation analysis between our introduced

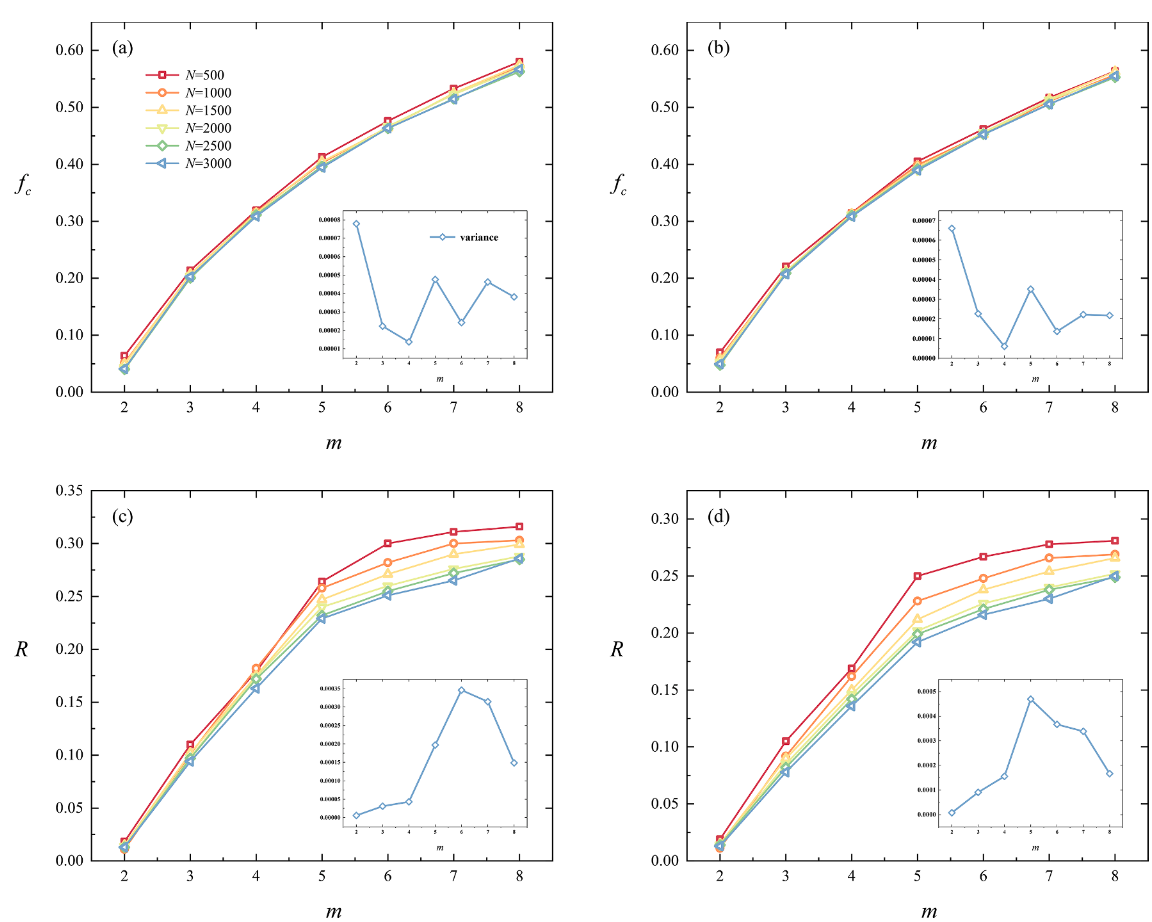

Fig. 4. Performance analysis in PLC Network. (a) The f_c of *NCDC*. (b) The f_c of *NDJC*. (c) The R of *NCDC*. (d) The R of *NDJC*. Each data marker corresponds to an average over 100 independent realizations.

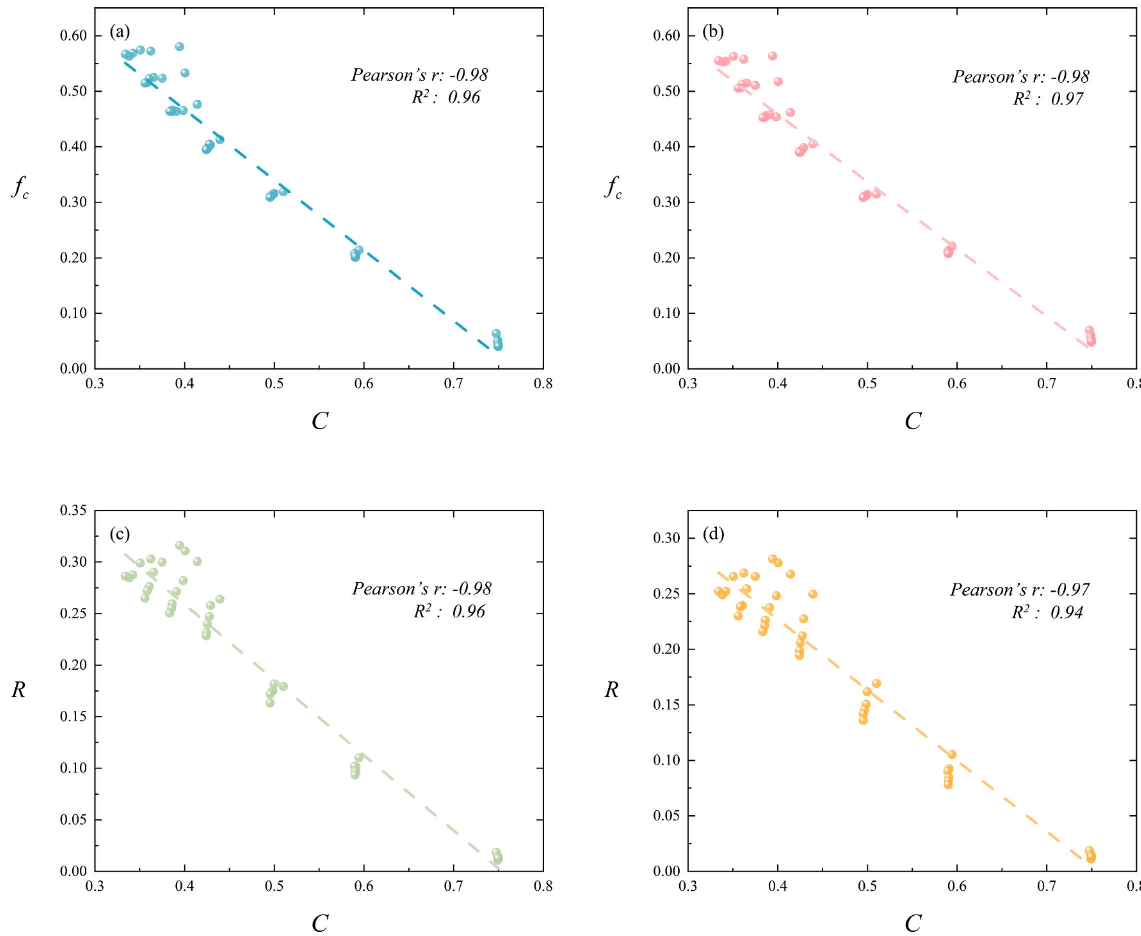

Fig. 5. Correlation analysis in PLC networks. (a) f_c vs. clustering coefficient of *NCDC*. (b) f_c vs. clustering coefficient of *NDJC*. (c) R vs. clustering coefficient of *NCDC*. (d) R vs. clustering coefficient of *NDJC*. Each data marker corresponds to an average over 100 independent realizations.

approaches and existing approaches. This analysis allows us to compare their distribution discrepancies and thereby gain deeper insights into the distinctive strengths of the NDCH framework.

4.1. Synthetic networks

Table S2 and S3 present the performance of various disintegration strategies across synthetic networks. Remarkably, the three introduced approaches under NDCH framework consistently surpass the SOTA methods in terms of f_c, achieving a maximum improvement of 4.64% in the PLC_2000_7_1 (implies the network size of PLC network is 2000, the average degree is 14 and the tunable parameter p is set as 1) network. Among the introduced approaches, *NDJC* generally surpasses *NCDC* across all synthetic networks, with the exception of certain network configurations with $m = 3$ and $m = 4$. In these specific instances, *NCDC* exhibits a slight advantage over *NDJC*, with an average discrepancy of 1.09. Regarding the Schneider R, our introduced approaches significantly outpace the SOTA methods, registering a substantial gain of up to 31.23% in the PLC_3000_8_1 network. Specifically, *NDJC* demonstrates marked superiority over *NCDC*, except when $m = 2$, where their performance is closely matched.

As depicted in Fig. 3, the promotion ratios of f_c and R demonstrate distinct patterns. Apart from a few networks with negligible gaps, our methods consistently outperform the benchmarks in terms of f_c. Initially, φ^{fc} declines from $m = 2$ to $m = 3$, stabilizes with a steady increase, but then decreases again at $m = 8$. Notably, the observed dip at $m = 8$ can be attributed to the accelerated increase of Cluster Rank as m grows. Our experiments, which exclude the Cluster Rank, reveal that when the network size is held constant, φ^{fc} escalates as m increases, commencing from $m = 3$. In terms of the Schneider R, φ^{R} exhibits a significant advantage over the benchmarks. The results underscore robust correlations between this metric and both network size and average degree. Specifically, with a fixed network size, φ^{R} increases as the average degree rises. Conversely, when the average degree remains constant, a larger network size is associated with heightened efficacy, i.e., high φ^{R}.

We further conduct a comprehensive analysis to investigate the effects of network size N and parameter m on the performance of *NCDC* and *NDJC*. In Fig. 4, each main subgraph contains an embedded subgraph. The main subgraphs portray the performance trends of *NCDC* and *NDJC* under varying conditions, while the embedded subgraphs specifically illustrate the variance associated with different network sizes, corresponding to distinct values of average degree. Our observations reveal that, for a given network size, both *NCDC* and *NDJC* exhibit a monotonic increase with the average degree. In contrast, for a fixed average degree, the performances of *NCDC* and *NDJC* are observed to be asymptotically independent of the network size. These results suggest that the efficacy of *NCDC* and *NDJC* is related to the average degree, but remains largely independent of the network size.

Finally, we investigate the influence of the clustering coefficient on the performance of *NCDC* and *NDJC*. As illustrated in Fig. 5, the data exhibit strong linear correlations across various network configurations. Specifically, the minimum absolute value of Pearson's correlation coefficient, exceeds 0.97, and the minimum goodness-of-fit value surpasses 0.94. These robust statistics highlight that our approaches effectively leverage redundant ties, thereby enhancing performance in networks characterized by high clustering coefficients.

Table 3
Comparison of disintegration strategies and promotion ratios of f_c on empirical networks.

Networks	*Cycle Ratio*	*Information Entropy*	*Cluster Rank*	*Leverage Centrality*	*MSTDC*	*NDC*	*NCDC*	*NDJC*	φ^{fc}
Birmingham	0.691	0.741	0.421	0.551	0.290	0.421	0.274	**0.200**	31.000%
Philadelphia	0.590	0.560	0.460	0.360	0.420	0.530	**0.330**	**0.266**	26.306%
Chicago	0.591	0.461	0.571	0.401	0.461	0.481	0.353	**0.318**	20.700%
Power_Grid	0.430	0.147	0.142	0.198	0.233	0.152	**0.139**	0.181	2.113%
Parsed_Jazz	0.692	0.742	0.747	0.702	0.702	0.732	**0.672**	0.720	2.920%
Parsed_Tap	0.591	0.561	0.428	0.581	0.530	0.387	**0.377**	0.469	11.916%
Netscience	0.158	0.182	0.119	0.164	0.216	0.111	**0.100**	0.113	15.556%
Openflights	0.333	0.293	0.253	0.293	0.303	0.242	**0.212**	0.277	16.206%
Facebook	0.741	0.711	0.619	0.700	0.834	0.619	**0.608**	0.629	1.785%
Circuit	0.464	0.234	0.246	0.171	0.298	0.222	0.202	**0.160**	6.433%
Hamster	0.570	0.380	0.340	0.490	0.390	0.310	**0.280**	0.291	17.650%
Collins_Yeast	0.603	0.603	0.340	0.635	0.482	0.329	**0.263**	0.482	22.677%

Note: The approach with the best performance is highlighted in bold, and the second best is underlined. All elements in the table are reserved for three decimal places. Results are averaged over 100 independent realizations.

Table 4
Comparison of disintegration strategies and promotion ratios of R on empirical networks.

Networks	*Cycle Ratio*	*Information Entropy*	*Cluster Rank*	*Leverage Centrality*	*MSTDC*	*NDC*	*NCDC*	*NDJC*	φ^{R}
Birmingham	0.217	0.159	0.160	0.133	0.125	0.133	0.061	**0.054**	56.655%
Philadelphia	0.236	0.189	0.200	0.152	0.172	0.188	0.082	**0.060**	60.151%
Chicago	0.287	0.209	0.209	0.181	0.192	0.207	0.095	**0.083**	54.461%
Power_Grid	0.082	0.066	0.055	0.075	0.058	0.054	**0.034**	0.036	38.386%
Parsed_Jazz	0.422	0.443	0.422	0.416	0.390	0.417	0.378	**0.364**	6.700%
Parsed_Tap	0.280	0.333	0.237	0.265	0.224	0.242	**0.195**	0.196	13.186%
Netscience	0.052	0.057	0.052	0.063	0.066	0.052	**0.046**	0.050	11.335%
Openflights	0.116	0.117	0.104	0.115	0.114	0.104	0.098	**0.095**	8.722%
Facebook	0.275	0.294	0.246	0.277	0.153	0.248	0.111	**0.111**	27.492%
Circuit	0.276	0.145	0.138	0.128	0.161	0.135	0.116	**0.105**	17.969%
Hamster	0.174	0.205	0.174	0.240	0.205	0.179	0.173	**0.164**	5.872%
Collins_Yeast	0.195	0.253	0.152	0.202	0.119	0.155	0.116	**0.113**	5.710%

Note: The approach with the best performance is highlighted in bold, and the second best is underlined. All elements in the table are reserved for three decimal places. Results are averaged over 100 independent realizations.

4.2. Empirical networks

Tables 3 and 4 respectively present the f_c and R of varying disintegration strategies applied to twelve empirical networks. Our proposed methods, *NCDC* and *NDJC*, which integrate redundancy elimination and community heterogeneity, consistently secures superior solutions across all scenarios in terms of f_c, with promotion ratios spanning from 1.785% to 31.000%. Specifically, *NDJC* performs better than *NCDC* in transportation networks, i.e., Birmingham, Philadelphia, Chicago and Circuit networks, indicating that it is extremely suitable for identifying the critical infrastructure of transportation systems, whereas *NCDC* outperforms in other networks. Notably, our basic method *NDC*, which focuses exclusively on redundancy elimination, markedly outperforms the benchmarks in several networks including Parsed_Tap, Netscience, Openflights, Facebook, Hamster, and Collins_Yeast despite not considering structural heterogeneity. This affirms the effectiveness of our novel approach to redundancy elimination. However, the performance of *NDC* is overshadowed by *NCDC* and *NDJC*, highlighting the critical role of structural heterogeneity in connected nodes. It is also worth noting that in benchmark measures, Cluster Rank can thoroughly disintegrate the network faster with better performance.

Turning to the Schneider R, *NCDC* and *NDJC* generally outperforms the benchmarks, with *NDJC* standing out in particular. This result corroborates the importance of different forms of information aggregation through comparison between *NCDC* and *NDJC*. The promotion ratios for R range from 5.710% to 60.151%, suggesting that *NCDC* and *NDJC* is versatile ideally suited when a network is severely damaged, but not completely collapsed. Moreover, *NDC* dramatically outperforms benchmarks in Power_Grid, Netscience, Openflights, and Hamster networks, reinforcing the strength of our redundancy elimination strategy. Furthermore, it is noteworthy that MSTDC consistently demonstrates superior collapse disintegration performance compared to other benchmark measures across the majority of empirical networks.

In conclusion, our comprehensive analysis robustly validates the effectiveness of the NDCH framework. We find that *NDJC* emerges as a more efficient measure for disintegrating complex networks in practical applications, striking a nuanced balance between rapid network collapse and sustained damage resilience. Moreover, *NCDC* proves particularly potent when a complete network collapse is the desired outcome.

To further explore the disintegration process under various strategies, we examine the distribution of robustness curves corresponding to different attack strategies. As depicted in Fig. 6, employing *NCDC* and *NDJC* as the guiding metrics for node removal generally precipitates a more rapid collapse of the empirical networks. Notably, in both the Power_Grid and Netscience networks, our methods enable a substantial portion of the network to be dismantled at a comparatively low cost, specifically when the removed fraction is below 0.05. A further observation in Birmingham,

Philadelphia, Chicago, Power_Grid, Netscience, and Circuit reveals that the maximal horizontal discrepancies between benchmarks and both *NDJC* and *NCDC* exceed 0.5, highlighting the efficiency of our proposed strategies. While the *NDC* strategy also exhibits effective disintegration results, it falls markedly short of the performance achieved by *NDJC* and *NCDC*, confirming the critical role of considering structural heterogeneity in network disintegration. Another intriguing insight is the superior performance of MSTDC relative to other benchmark strategies, especially during the initial stages of node removal in empirical networks. This suggests that MSTDC can expedite network disintegration in the early phase, thereby offering significant resource efficiency in practical applications.

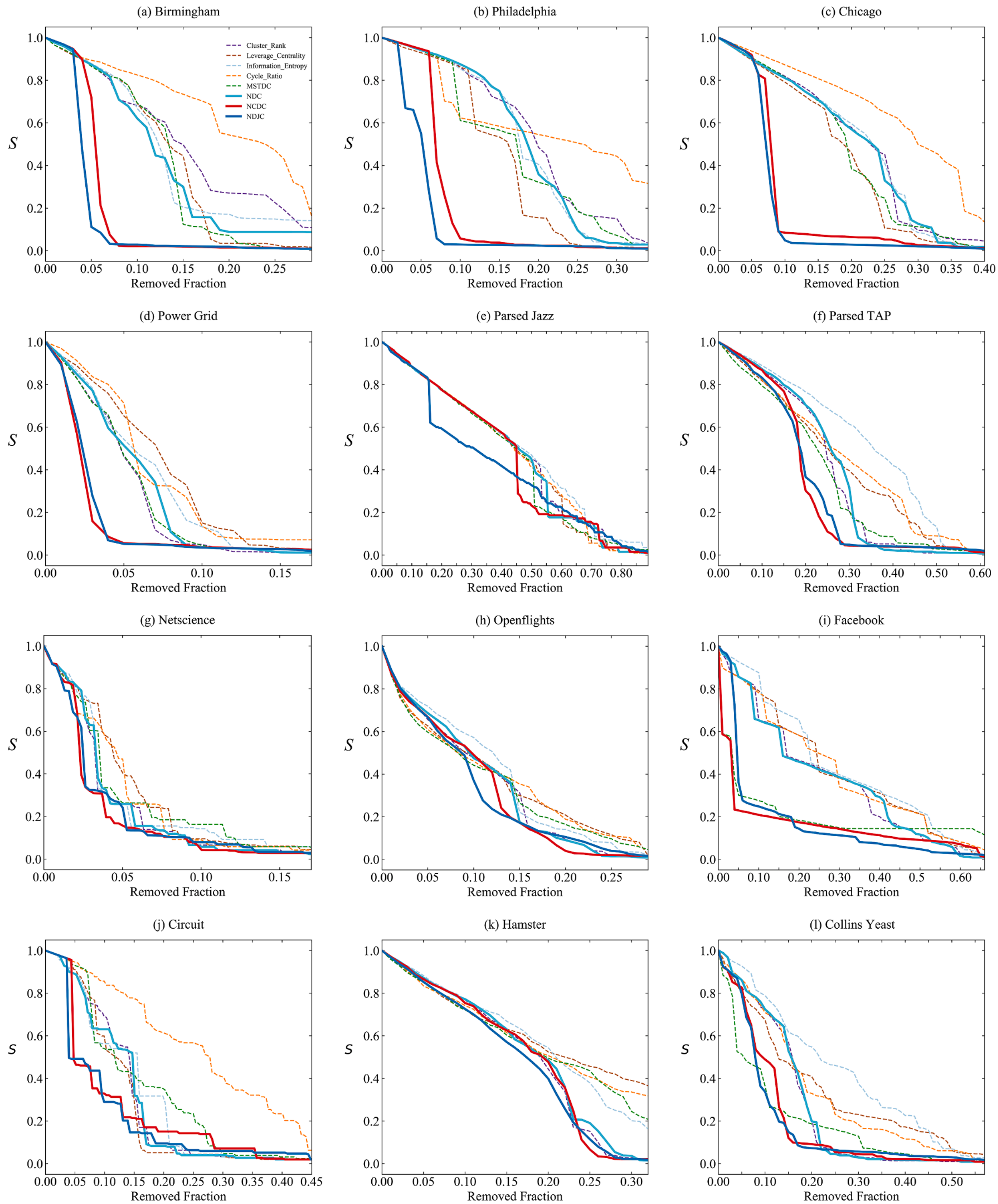

Fig. 6. Robustness curves against various attack strategies observed in empirical networks.

4.3. Correlation analysis

To explore the inherent distinctions between our proposed NDCH methods and the benchmark strategies, we employ Kendall's Tau (τ) as a metric to assess the correlations in the node rankings produced by our introduced approaches and the SOTA approaches (Kendall, 1938). Specifically, we aggregate the ranking results, optimized with respect to R, to characterize the performance of our approaches. As depicted in Fig. 7, the average correlations between our methods and the benchmarks are notably low, falling below 0.6. In contrast, the correlations with *NDC* and *NCDC* are relatively higher, registering $\overline{\tau} = 0.64$ and $\overline{\tau} = 0.67$, respectively. These low correlations with the benchmarks suggest that the node rankings generated by *NDJC* encompass more informative insights compared to those derived from alternative metrics. This validates the effectiveness of our proposed methods and underscores the importance of integrating heterogeneity and information aggregation techniques when designing network disintegration strategies.

Furthermore, our analysis reveals that the minimum value of τ is 0.15, observed when comparing our method with Information Entropy in the Netscience network. In contrast, the maximum τ value of 0.86 is found when comparing our method with Leverage Centrality in the Power_Grid network. Intriguingly, this high τ value of 0.86 appears to contradict the results presented in Section 4.2.

To shed light on the reasons behind the substantial variance in disintegration effects, even with high τ values, we present the distributions of the indices for the Power_Grid network in Fig. 8. Our method

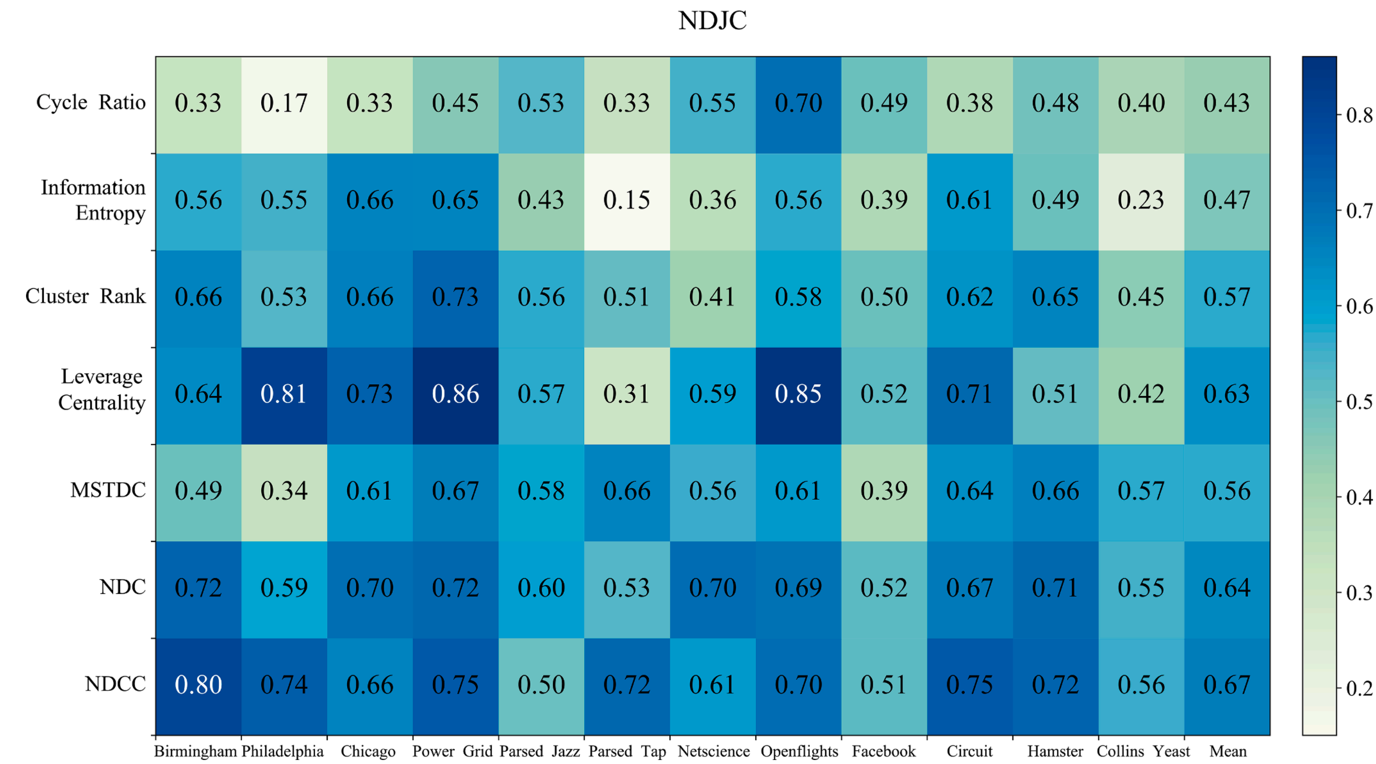

Fig. 7. The correlation matrix between *NDJC* and other approaches on empirical networks.

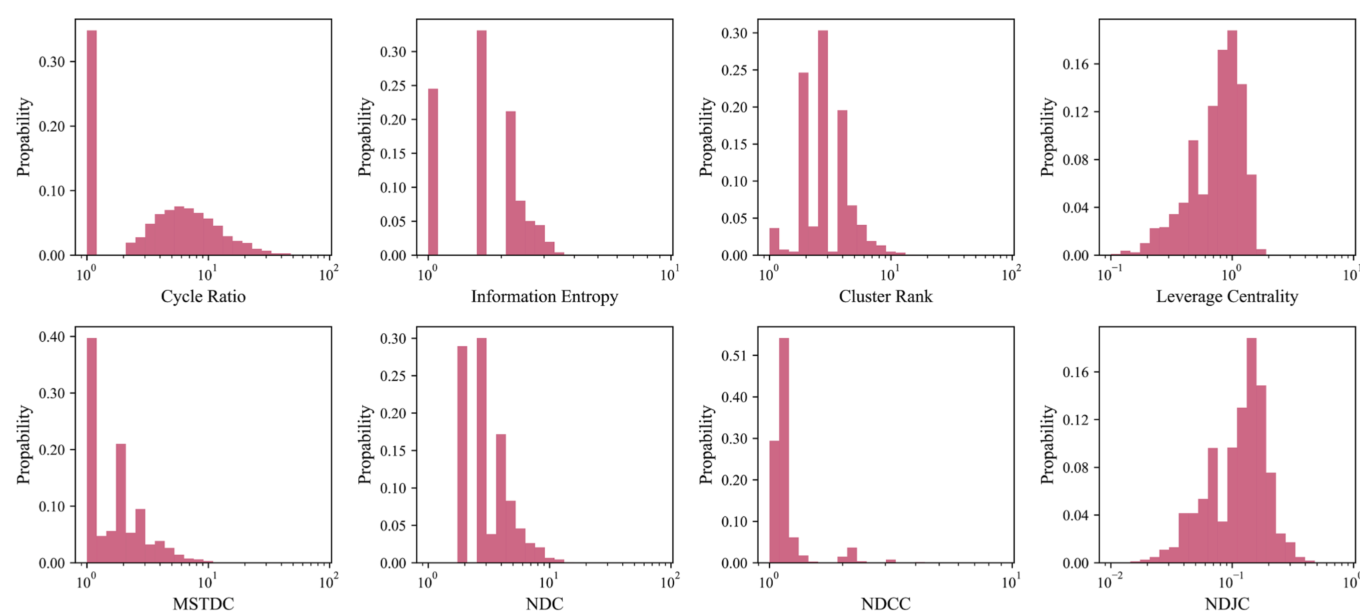

Fig. 8. Log-Binned histograms of different approaches in Power_Grid network. To mitigate the distortions in distribution representation caused by the expansive ranges of centrality indicators, we employ a log-bin method. This approach segments the data into logarithmically scaled bins, producing an arithmetic sequence on the x-axis under a power-law distribution. This technique ensures a uniform distribution of statistical errors across the dataset (Clauset et al., 2009).

demonstrates commendable distinguishability across most fractions, and its distribution exhibits a resemblance to that of Leverage Centrality. Nevertheless, our method more distinctly identifies the influential nodes. This suggests that the high τ values are attributed to the analogous ranking of middle and lower-tier nodes. For further insights, Supplementary Notes 4 and 5 provide the correlation matrix between *NDC* and *NCDC* across the various metrics, as well as the distributions of the approaches for the remaining empirical networks, respectively.

5. Conclusion and discussion

While link redundancy has long been recognized as a critical factor affecting network robustness, it has traditionally constrained the effectiveness of topology-based network disintegration strategies. The present study challenges this paradigm and introduce a novel perspective: the urgency to differentiate between links and nodes when quantifying redundancy, and assigning weights that are sensitive to the underlying structural heterogeneity of the network. We contend that effective network disintegration strategies must transcend the singular focus on redundancy elimination to incorporate a more nuanced understanding of structural heterogeneity.

Our work illuminates the intricate relationships between redundant ties and community structures within networks. Redundant ties contribute to the emergence of community structure within a network and explicit and sophisticated community structures once formed, further exacerbate the heterogeneity among nodes and links. To capture this complicated interplay, we have developed the NDCH framework to quantify structural heterogeneity, leveraging information derived from community structures and employing diverse information aggregation approaches.

Extensive rigorous ablation studies, conducted on both synthetic and real-world network datasets, have demonstrated the superior performance of the NDCH framework. Within the NDCH framework, our introduced strategies consistently outperform state-of-the-art methods in network disintegration, with performance improvements reaching up to 60.151% and 31.000% for Schneider R and the critical removal fraction f_c, respectively. Specifically, our approaches have demonstrated extreme effectiveness in infrastructure networks, compared to the benchmarks, contributing to identifying vulnerable entities and offering insights for designing targeted protection. These findings underscore the substantial potential for performance embedded in the integrated consideration of link redundancy and structural heterogeneity. Empirically, our analyses, which employ Kendall's Tau, reveal consistently low correlations (below 0.6) between our approaches and existing state-of-the-art methods. This outcome is indicative of the distinctive and superior discriminative capability of our strategies.

In summary, our work proposes a groundbreaking and comprehensive framework for network disintegration strategies that emphasizes the critical importance of addressing redundant ties alongside other high-order redundancy structures, such as community structures. This framework not only highlights the potential side effects of redundant ties but also serves as a pragmatic, actionable guide for crafting effective strategies tailored to real-world applications. Our results provide a valuable reference, steering the selection and design of potent strategies for practical network dismantling endeavors, particularly in identifying key components within infrastructure networks.

Looking forward, we believe that incorporating additional fine-grained structural information has the potential to further enhance strategy performance. Notably, given that redundant ties can be quantified by the length and number of redundant paths, potentially linked to the characteristic spectrum of a network, we are eager to integrate this aspect into our research. We anticipate that such integrations will offer valuable insights for the strategic enhancement and calculated dismantling of network robustness, ensuring the security and stability of infrastructure systems.

CRediT authorship contribution statement

Bitao Dai: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Visualization. **Jianhong Mou:** Data curation, Writing – review & editing. **Suoyi Tan:** Writing – review & editing. **Mengsi Cai:** Writing – review & editing. **Fredrik Liljeros:** Writing – review & editing. **Xin Lu:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (72025405, 72088101, 72301285, 72001211), the National Social Science Foundation of China (22ZDA102), and the Natural Science Foundation of Hunan Province (2023JJ40685).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eswa.2024.124590.

References

Akhtar, M. U., Liu, J., Liu, X., Ahmed, S., & Cui, X. (2023). NRAND: An efficient and robust dismantling approach for infectious disease network. *Information Processing & Management, 60*(2), Article 103221. https://doi.org/10.1016/j.ipm.2022.103221
Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47.
Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature, 406*(6794), 378–382.
Artime, O., Grassia, M., De Domenico, M., Gleeson, J. P., Makse, H. A., Mangioni, G., Perc, M., & Radicchi, F. (2024). Robustness and resilience of complex networks. *Nature Reviews Physics*, 1–18.
Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology, 2*(1), 113–120.
Bouyer, A., Beni, H. A., Arasteh, B., Aghaee, Z., & Ghanbarzadeh, R. (2023). FIP: A fast overlapping community-based Influence Maximization Algorithm using probability coefficient of global diffusion in social networks. *Expert Systems with Applications, 213*, Article 118869. https://doi.org/10.1016/j.eswa.2022.118869
Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks, 30*(2), 136–145. https://doi.org/10.1016/j.socnet.2007.11.001
Braunstein, A., Dall'Asta, L., Semerjian, G., & Zdeborová, L. (2016). Network dismantling. *Proceedings of the National Academy of Sciences, 113*(44), 12368-12373.
Callaway, D. S., Newman, M. E., Strogatz, S. H., & Watts, D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters, 85*(25), 5468.
Chen, D.-B., Gao, H., Lü, L., & Zhou, T. (2013). Identifying influential nodes in large-scale directed networks: The role of clustering. *PloS One, 8*(10), e77455.
Chen, L.-C., & Carley, K. M. (2004). The impact of countermeasure propagation on the prevalence of computer viruses. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 34*(2), 823–833. https://doi.org/10.1109/TSMCB.2003.817098
Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review, 51*(4), 661–703. https://doi.org/10.1137/070710111
Collins, S. R., Schuldiner, M., Krogan, N. J., & Weissman, J. S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biology, 7*, 1–14.
D'Souza, R. M., di Bernardo, M., & Liu, Y.-Y. (2023). Controlling complex networks with complex nodes. *Nature Reviews Physics, 5*(4), 250–262.
D'Souza, R. M., & Mitzenmacher, M. (2010). Local cluster aggregation models of explosive percolation. *Physical Review Letters, 104*(19), Article 195702.
Dai, B., Qin, S., Tan, S., Liu, C., Mou, J., Deng, H., Liljeros, F., & Lu, X. (2023). Identifying influential nodes by leveraging redundant ties. *Journal of Computational Science, 69*, Article 102030.
Deng, Y., Wu, J., Qi, M., & Tan, Y. (2019). Optimal disintegration strategy in spatial networks with disintegration circle model. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 29*(6), Article 061102.
Deng, Y., Wu, J., & Tan, Y.-J. (2016). Optimal attack strategy of complex networks based on tabu search. *Physica A: Statistical Mechanics and its Applications, 442*, 74–81.
Deng, Y., Wu, J., Xiao, Y., Zhang, M., Yu, Y., & Zhang, Y. (2018). Optimal disintegration strategy with heterogeneous costs in complex networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50*(8), 2905–2913.
Doyle, J. C., Alderson, D. L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., & Willinger, W. (2005). The "robust yet fragile" nature of the Internet. *Proceedings of the National Academy of Sciences, 102*(41), 14497–14502.
Eiselt, H. (2018). Destabilization of terrorist networks. *Chaos, Solitons & Fractals, 108*, 111–118.
Fan, C., Zeng, L., Feng, Y., Xiu, B., Huang, J., & Liu, Z. (2020). Revisiting the power of reinsertion for optimal targets of network attack. *Journal of Cloud Computing, 9*, 1–13.
Fan, C., Zeng, L., Sun, Y., & Liu, Y.-Y. (2020). Finding key players in complex networks through deep reinforcement learning. *Nature machine intelligence, 2*(6), 317–324.
Fan, T., Lü, L., Shi, D., & Zhou, T. (2021). Characterizing cycle structure in complex networks. *Communications Physics, 4*(1), 272.
Freitas, S., Yang, D., Kumar, S., Tong, H., & Chau, D. H. (2022a). Graph vulnerability and robustness: A survey. *IEEE Transactions on Knowledge and Data Engineering, 35*(6), 5915–5934.
Freitas, S., Yang, D., Kumar, S., Tong, H., & Chau, D. H. (2022b). Graph vulnerability and robustness: A survey. *IEEE Transactions on Knowledge and Data Engineering.*
Gao, J., Liu, Y.-Y., D'souza, R. M., & Barabási, A.-L. (2014). Target control of complex networks. *Nature Communications, 5*(1), 5415.

Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., & Dümpelfeld, B. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature, 440*(7084), 631–636.

Gleiser, P. M., & Danon, L. (2003). Community structure in jazz. *Advances in Complex Systems, 6*(04), 565–573.

Holme, P., & Kim, B. J. (2002). Growing scale-free networks with tunable clustering. *Physical Review E, 65*(2), Article 026107.

Holme, P., Kim, B. J., Yoon, C. N., & Han, S. K. (2002). Attack vulnerability of complex networks. *Physical Review E, 65*(5), Article 056109.

Huang, Y., Wang, H., Ren, X.-L., & Lü, L. (2024). Identifying key players in complex networks via network entanglement. *Communications Physics, 7*(1), 19.

Joyce, K. E., Laurienti, P. J., Burdette, J. H., & Hayasaka, S. (2010). A new measure of centrality for brain networks. *PloS One, 5*(8), Article e12200.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika, 30*(1/2), 81–93.

Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature Physics, 6*(11), 888–893.

Kunegis, J. (2014). Hamsterster full network dataset-KONECT. *konect. uni-koblenz: de/networks/petster-hamster*.

Leskovec, J., & Mcauley, J. (2012). Learning to discover social circles in ego networks. *Advances in neural information processing systems, 25*.

Li, Q., Liu, S.-Y., & Yang, X.-S. (2020). Neighborhood information-based probabilistic algorithm for network disintegration. *Expert Systems with Applications, 139*, Article 112853.

Lin, J.-H., Guo, Q., Dong, W.-Z., Tang, L.-Y., & Liu, J.-G. (2014). Identifying the node spreading influence with largest k-core values. *Physics Letters A, 378*(45), 3279–3284.

Liu, Y., Tang, M., Zhou, T., & Do, Y. (2015a). Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition. *Scientific Reports, 5*(1), 1–8.

Liu, Y., Tang, M., Zhou, T., & Do, Y. (2015b). Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition. *Scientific Reports, 5*(1), 9602.

Liu, Y., Tang, M., Zhou, T., & Do, Y. (2015c). Improving the accuracy of the k-shell method by removing redundant links: From a perspective of spreading dynamics. *Scientific Reports, 5*(1), 13172.

Lou, Y., Wang, L., & Chen, G. (2023). Structural robustness of complex networks: A survey of a posteriori measures [feature]. *IEEE Circuits and Systems Magazine, 23*(1), 12–35.

Millán, A. P., Torres, J. J., & Bianconi, G. (2020). Explosive higher-order Kuramoto dynamics on simplicial complexes. *Physical Review Letters, 124*(21), Article 218301.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science, 298*(5594), 824–827.

Newman, M. E. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E, 74*(3), Article 036104.

Nian, F., Ren, S., & Dang, Z. (2017). The propagation-weighted priority immunization strategy based on propagation tree. *Chaos, Solitons & Fractals, 99*, 72–78.

Nishi, A., Dewey, G., Endo, A., Neman, S., Iwamoto, S. K., Ni, M. Y., Tsugawa, Y., Iosifidis, G., Smith, J. D., & Young, S. D. (2020). Network interventions for managing the COVID-19 pandemic and sustaining economy. *Proceedings of the National Academy of Sciences, 117*(48), 30285–30294.

Openflights: Network data from openflights.org https://openflights.org/data.html.

Qi, M., Deng, Y., Deng, H., & Wu, J. (2018). Optimal disintegration strategy in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 28*(12), Article 121104.

Ren, X.-L., Gleinig, N., Helbing, D., & Antulov-Fantulin, N. (2019). Generalized network dismantling. *Proceedings of the National Academy of Sciences, 116*(14), 6554–6559.

Robins, G., Lusher, D., Broccatelli, C., Bright, D., Gallagher, C., Karkavandi, M. A., Matous, P., Coutinho, J., Wang, P., & Koskinen, J. (2023). Multilevel network interventions: Goals, actions, and outcomes. *Social Networks, 72*, 108–120.

Schneider, C. M., Moreira, A. A., Andrade, J. S., Jr, Havlin, S., & Herrmann, H. J. (2011). Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences, 108*(10), 3838–3841.

Shi, D., Chen, Z., Sun, X., Chen, Q., Ma, C., Lou, Y., & Chen, G. (2021). Computing cliques and cavities in networks. *Communications Physics, 4*(1), 249.

Shi, D., Lü, L., & Chen, G. (2019). Totally homogeneous networks. *National Science Review, 6*(5), 962–969.

Sizemore, A. E., Karuza, E. A., Giusti, C., & Bassett, D. S. (2018). Knowledge gaps in the early growth of semantic feature networks. *Nature Human Behaviour, 2*(9), 682–692.

Tan, S.-Y., Deng, Y., & Wu, J. (2019). Cost-effectiveness analysis of structural robustness in complex networks. 2019 IEEE International Symposium on Circuits and Systems (ISCAS).

Tan, S.-Y., Wu, J., Lü, L., Li, M.-J., & Lu, X. (2016). Efficient network disintegration under incomplete information: The comic effect of link prediction. *Scientific Reports, 6*(1), 1–9.

Wandelt, S., Lin, W., Sun, X., & Zanin, M. (2022). From random failures to targeted attacks in network dismantling. *Reliability Engineering & System Safety, 218*, Article 108146.

Wandelt, S., Shi, X., & Sun, X. (2021). Estimation and improvement of transportation network robustness by exploiting communities. *Reliability Engineering & System Safety, 206*, Article 107307.

Wang, S., Ding, B., & Jin, Y. (2023). A multi-factorial evolutionary algorithm with asynchronous optimization processes for solving the robust influence maximization problem. *IEEE Computational Intelligence Magazine, 18*(3), 41–53.

Wang, S., Jin, Y., & Cai, M. (2023). Enhancing the robustness of networks against multiple damage models using a multifactorial evolutionary algorithm. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Wang, S., Liu, J., & Jin, Y. (2021). A computationally efficient evolutionary algorithm for multiobjective network robustness optimization. *IEEE Transactions on Evolutionary Computation, 25*(3), 419–432.

Wang, S., & Liu, W. (2023). Enhancing the robustness of influential seeds towards structural failures on competitive networks via a Memetic algorithm. *Knowledge-Based Systems, 275*, Article 110677.

Wang, S., & Tan, X. (2023a). Finding robust influential seeds from networked systems against structural failures using a niching memetic algorithm. *Applied Soft Computing, 136*, Article 110134.

Wang, S., & Tan, X. (2023b). A Memetic algorithm for determining robust and influential seeds against structural perturbances in competitive networks. *Information Sciences, 621*, 389–406.

Wang, Z.-G., Deng, Y., Wang, Z., & Wu, J. (2021). Disintegrating spatial networks based on region centrality. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 31*(6), Article 061101.

Wang, Z., Su, Z., Deng, Y., Kurths, J., & Wu, J. (2024). Spatial network disintegration based on kernel density estimation. *Reliability Engineering & System Safety, 110005*.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature, 393*(6684), 440–442.

Wu, J., Barahona, M., Tan, Y.-J., & Deng, H.-Z. (2011). Spectral measure of structural robustness in complex networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 41*(6), 1244–1252.

Xu, X., Zhu, C., Wang, Q., Zhu, X., & Zhou, Y. (2020). Identifying vital nodes in complex networks by adjacency information entropy. *Scientific Reports, 10*(1), 1–12.