

基于非回溯矩阵中心性的超图可靠性研究

彭浩^{1,2}, 钱程¹, 赵丹丹¹, 钟鸣¹, 韩建民¹, 谢紫伊¹, 王伟³

(1. 浙江师范大学计算机与科学技术学院, 浙江 金华 321004;

2. 浙江师范大学浙江省智能教育技术与应用重点实验室, 浙江 金华 321004;

3. 重庆医科大学公共卫生学院, 重庆 400016)

摘要: 近年来, 超图作为网络科学的一个研究热点, 引起了广泛的关注。超图区别于传统图的结构特点在于它的超边可以同时连接多个节点, 从而形成更为复杂和高阶的关系。在这样的网络结构中, 有效地识别重要的节点和超边成为一个关键的挑战。特征向量中心性是一个常见的度量标准, 但当网络中存在着极大度值的枢纽节点时, 使用特征向量中心性度量方法会使结果表现出局域性, 限制了该方法的应用场景。因此, 将超图转化成对应的线图, 在此基础上使用非回溯矩阵中心性这一方法, 该方法在评估超边重要性时表现出更好的均匀性和区分度。此外, 还探讨了特征向量中心性和非回溯矩阵中心性在超图中节点重要性评估上的应用。通过比较这两种方法, 研究发现非回溯矩阵中心性在区分节点重要程度方面具有更明显的优势。研究不仅包括理论分析和模型构建, 还包括对真实世界数据的实证。为了验证所提方法和结论, 选取了6个真实世界超图作为实验对象。通过在这些超图上的应用, 证明了非回溯矩阵中心性在识别重要节点和超边方面的有效性。研究为超图中关键元素的识别提供了一种新的视角和方法, 对于理解和分析实际复杂网络系统, 具有重要的理论和实践意义。

关键词: 超图; 特征向量中心性; 非回溯矩阵中心性; 向量中心性

中图分类号: TP301.6

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024011

Study on the reliability of hypergraphs based on non-backtracking matrix centrality

PENG Hao^{1,2}, QIAN Cheng¹, ZHAO Dandan¹, ZHONG Ming¹, HAN Jianmin¹, XIE Ziyi¹, WANG Wei³

1. School of Computer Science and Technology, Zhejiang Normal University, Jinhua 321004, China

收稿日期: 2023-08-01; **修回日期:** 2023-11-14

通信作者: 王伟, wwzqbx@hotmail.com

基金项目: 国家自然科学基金(62074212, 61902359, 61702148); 信息网络安全公安部重点实验室开放课题(C20607); 重庆医科大学未来青年医学创新计划(W0150)

Foundation Items: The National Natural Science Foundation of China (62074212, 61902359, 61702148), The Open Project Program of the Key Laboratory of Information Network Security, Ministry of Public Security (C20607), The Program for Youth Innovation in Future Medicine, Chongqing Medical University(W0150)

引用格式: 彭浩, 钱程, 赵丹丹, 等. 基于非回溯矩阵中心性的超图可靠性研究[J]. 网络与信息安全学报, 2024, 10(1): 22-32.

Citation Format: PENG H, QIAN C, ZHAO D D, et al. Study on the reliability of hypergraphs based on non-backtracking matrix centrality[J]. Chinese Journal of Network and Information Security, 2024, 10(1): 22-32.

2. Key Laboratory of Intelligent Educational Technology and Application, Zhejiang Normal University, Jinhua 321004, China

3. College of Public Health, Chongqing Medical University, Chongqing 400016, China

Abstract: In recent years, there has been widespread attention on hypergraphs as a research hotspot in network science. The unique structure of hypergraphs, which differs from traditional graphs, is characterized by hyperedges that can connect multiple nodes simultaneously, resulting in more complex and higher-order relationships. Effectively identifying important nodes and hyperedges in such network structures poses a key challenge. Eigenvector centrality, a common metric, has limitations in its application due to its locality when dealing with hub nodes with extremely high degree values in the network. To address this issue, the hypergraphs were transformed into their corresponding line graphs, and non-backtracking matrix centrality was employed as a method to measure the importance of hyperedges. This approach demonstrated better uniformity and differentiation in assessing the importance of hyperedges. Furthermore, the application of both eigenvector centrality and non-backtracking matrix centrality in assessing the importance of nodes in hypergraphs was explored. Comparative analysis revealed that non-backtracking matrix centrality effectively distinguished the importance of nodes. This research encompassed theoretical analysis, model construction, and empirical studies on real-world data. To validate the proposed method and conclusion, six real-world hypergraphs were selected as experimental subjects. The application of these methods to these hypergraphs confirmed the effectiveness of non-backtracking matrix centrality in identifying important nodes and hyperedges. The findings of this research offer a fresh perspective and approach for identifying key elements in hypergraphs, holding significant theoretical and practical implications for understanding and analyzing complex network systems.

Keywords: hypergraph, eigenvector centrality, non-backtracking matrix centrality, vector centrality

0 引言

在现实世界中, 复杂系统由许多相互作用的元素组成^[1]。例如, 电力网络^[2]、社交网络^[3]、交通网络^[4]等。最常用的复杂网络建模方法是将元素视为节点, 元素间的关系视为边。这种只考虑两两节点之间成对交互的网络被称为“简单网络”^[5]。在网络中根据重要程度将节点进行排序一直是网络科学领域中研究的热点^[6-7], 成功地识别和利用这些关键节点^[8], 可以对网络的运行进行有效的干预, 如提高信息传播的效率、增强网络的可靠性或防止有害信息的扩散。

为了识别这些关键节点, 学者提出了多种方法。其中, 度中心性是基本和直观的方法, 通过考虑节点的连接数来判断其重要性^[9-10]。而介数中心性则通过节点在网络中的桥梁作用来度量其重要性^[11-12]。特征向量中心性则通过网络的特征向量来刻画节点的中心性, 并在某些网络中展现出其独特优势^[13-14]。但是, 这些传统中心性指标在处理存在大量回溯路径的网络时会遇到一些挑战。例如, 度中心性简单地衡量一个节点的连接

数。在有许多回溯路径的网络中, 节点的连接数可能会由于回溯路径的存在而变得很高, 但这并不意味着节点在全局上很重要。介数中心性计算依赖于一个节点的最短路径数量。在有大量回溯路径的网络中, 最短路径可能会被重复计算, 因为从一个节点出发可能存在多条相同长度的路径到达同一个目标, 这会夸大某些节点的重要性。特征向量中心性在存在大量回溯路径的网络上计算时, 这些路径上的节点可能会错误地被视为更重要。这样会增加这节点的特征向量分数, 因为它们互相增强了彼此的重要性, 即使这种情况在实际网络中可能并不代表真实的影响力。

以上局限性使非回溯矩阵中心性逐渐受到关注, 它旨在精确表示从任一节点出发, 下一步能够到达的节点集合, 而明确排除了直接返回上一节点的可能性, 从而有效地避免了在路径计算中考虑“回溯”这一行为, 为网络科学研究提供了一个新的视角, 可以更有效地识别网络中的关键节点。

此外, 已有研究证实在无标度网络中, 使用

特征向量中心性度量节点会出现局域性的情况^[15], 即只有少数的枢纽节点中心性高, 而其他大部分节点的中心性极小, 不能有效地量化每个节点的重要性。Martin 等^[16]同样证实在随机网络中, 如果枢纽节点的度值很大, 则通过特征向量中心性计算方法会使网络出现局域性。即中心性大的值集中在网络的少部分节点上, 显著减少了特征向量中心性量化节点重要性的有效性。文献[16]提出了一种新的度量方式: 基于非回溯矩阵的主特征向量中心性, 该非回溯矩阵中心性可以有效地避免中心性出现局域性的情况, 从而在特征向量中心性失效的情况下同样能够给出良好的节点中心性度量结果。

为了更好地描述网络中存在的高阶相互作用, 超图^[17-18]成为一个宝贵的数学工具。相比简单网络只能描述点与点间的成对交互, 超图可以描述多个节点之间的高阶交互, 为揭示复杂网络结构中的高阶关系提供了可能。20 世纪六七十年代, 随着图论和组合数学的迅猛发展, 超图开始被广泛研究并应用于多个领域。这一时期的研究扩展了网络理论的边界, 为解决传统网络中心性度量所面临的局域性问题提供了新的视角和工具。

超图在建模方面有着独特的优势, 超图中的超边可以很自然地表示任意数量节点之间的高阶作用关系, 比传统的简单网络更适用于表示真实世界出现的一些场景^[19], 如在建模科学家合作网络时, 可以将科学家看作节点, 将多位科学家合著完成并发表的论文视为一条超边。通过超边可以清楚地看到多位科学家之间的合作关系, 这在简单网络中是做不到的。近年来, 如何识别超图中的重要节点同样吸引着研究人员, Kovalenko 等^[20]提出了一种新的节点中心性度量方法, 节点中心性不再是一个标量值, 而是一个维度比图中交互的最高阶低一维的向量。它能够揭示同一节点在不同交互阶数下可能扮演的不同角色, 这是不可能通过单个标量测量来检索的信息。Peng 等^[21]根据超图中节点超度的大小, 有选择地蓄意攻击网络中的节点, 得出该方法相比于随机攻击对超图的破坏性更大。同样地, 他们还根据超图中超边基数的大小, 有选择地蓄意攻击网络中的超边

^[22], 得出随着基数大的超边被去除的概率增加, 超图的可靠性变得越来越脆弱这一结论。Tudisco 等^[23]定义和分析了一类谱中心性度量, 用于识别超图中的重要节点和超边。通过利用非线性 Perron–Frobenius 理论的最新发展, 展示了由此产生的约束非线性特征值问题如何具有可以通过非线性幂法迭代有效计算的唯一解。Mishra 等^[24]研究了超边和普通边在驱动超图特征向量局域化中的相互作用, 发现超图特征向量会位于度数远高于或远低于平均度数的节点集上, 即超图中特征向量中心性也会出现局域性。周丽娜等^[25]基于超图中的理论, 提出通过超图的邻接结构熵来识别网络中的关键节点, 该方法通过研究节点及其直接与间接节点间的关系, 利用节点信息熵刻画不同节点在超图中的重要性。

本文将非回溯矩阵中心性的概念扩展到超图上, 希望为识别超图中的关键节点以及超边提供新的策略, 并进一步拓展非回溯矩阵中心性在复杂网络分析中的应用。为此, 本文探究了超图中使用非回溯矩阵中心性能否有效地避免局域性的出现, 并且在超图上将非回溯矩阵中心性这一度量指标与特征向量中心性进行对比实验。为了便于研究超图中的节点和超边的相关性质, 本文引入了线图这一方法^[26], 通过将超图转化为线图, 可以方便地计算出每条超边的中心性。实验结果显示, 与 Mishra 等^[24]得出的结论一致, 超图中使用特征向量中心性同样会出现局域性。而本文提出的非回溯矩阵中心性方法比特征向量中心性在超图中量化节点的重要程度上更加均匀, 即每个节点的中心性的度量结果区分度更高一些。

1 模型

为了探究超图网络中每个节点或每条超边在超图网络中的重要程度。首先定义一个超图 $G=(V, E)$, 其中 V 表示包含 N 个节点的有限集合, 即 $|V|=N$; E 表示超边集, 即 $|E|=H$ 。每条超边 $e \in E$ 包含的节点数目大于或等于 2, 且任意两条超边 $e_i \in E$, $e_j \in E$ 之间可以存在共同节点。节点 $i \in V$ 的超度 d_i 表示节点 i 被超边包含的数量, 用 $\bar{d}$ 表示超图中节点的平均超

度，即

$$\bar{d} = \frac{1}{N} \sum_{i \in V} d_i \quad (1)$$

超边 e_i 的基数大小 m_i 表示超边 e_i 所包含的节点数量 ($m_i \geq 2$)，可以简单地得到如下等式。

$$kN = \sum_{e_i \in E} m_i \quad (2)$$

为了方便研究超边的相关性质，将超图 G 与它对应的线图 $L(G)$ 联系起来。一般来说，线图 $L(G)$ 是一个简单图，超图 G 中的每条超边 e 都映射到 $L(G)$ 中的一个节点，因此 $L(G)$ 包含 H 个节点。 $L(G)$ 中的边表示 G 中超边之间的邻接关系：如果 $e_i, e_j \in E$ 是 G 中任意两条超边，当且仅当 e_i 和 e_j 包含公共节点，即 $e_i \cap e_j \neq \emptyset$ 时， $L(G)$ 中对应的两个节点 e_i 和 e_j 之间存在一条无向边。例如，如图 1(a) 所示，超图 $G = (V, E)$ 由节点集 $V = \{1, 2, 3, 4, 5, 6\}$ 和边集 $E = \{e_1, e_2, e_3, e_4\}$ 构成，图 1(b) 是超图 G 对应的线图 $L(G)$ 。那么如何定义超图中节点的重要程度？

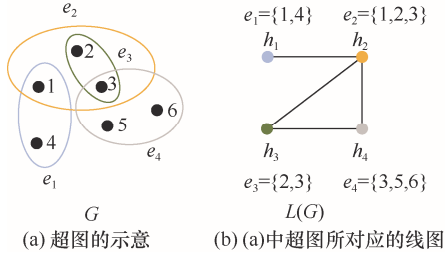


图 1 超图转化成线图的示意
Figure 1 Schematic of converting a hypergraph into a line graph

在超图中，一个节点可以同时存在于多个超边，并且每个超边的大小可以不同，即同一节点在不同阶数的交互下扮演着不同的角色，这导致节点在不同大小的超边中的重要性可能是不同的。为了表示不同大小的超边中节点的重要性，这里借鉴 Kovalenko 等^[20]提出的向量中心性来计算超图中节点的特征向量中心性和非回溯矩阵中心性。

2 方法论

为了得到超图中每条超边的重要性，将该问题转化为衡量线图中每个节点的重要性。在超图对应的线图中，借助邻接矩阵 A 来计算节点的特征向量中心性，即

$$A\mathbf{v} = \lambda\mathbf{v} \quad (3)$$

其中， A_{ij} 是邻接矩阵 A 中的元素，在数学上，其元素 A_{ij} 定义为当且仅当网络中两个节点 i 和 j 之间存在边时， $A_{ij} = 1$ ，否则 $A_{ij} = 0$ 。向量 $\mathbf{v}$ 是邻接矩阵 A 的特征向量，其中 λ 是常数。根据经典的 Perron-Frobenius 定理^[27]，选择与最大特征值 $\lambda_{\max}$ 相对应的特征向量 $\mathbf{v}_{\max}$ ，特征向量 $\mathbf{v}_{\max}$ 的每个分量 v_i （对应于网络中的每个节点 i ）就是节点 i 的特征向量中心性得分。因此，节点的特征向量中心性得分反映了其在网络中的相对重要性，高分的节点被认为在网络中的位置更加中心，对网络的影响力更大。

一般地，线图的非回溯矩阵中心性可以通过非回溯矩阵 B 来计算。需要将一个具有 m 条边的无向网络转换为具有 $2m$ 条边的有向网络，即将无向边 ij 表示为两条有向边 $i \rightarrow j$ 和 $j \rightarrow i$ ，每条有向边 $i \rightarrow j$ 表示为非回溯矩阵 B 的一行或一列，则 B 是一个 $2m \times 2m$ 的非对称矩阵， B 中的元素为

$$B_{i \rightarrow j, k \rightarrow l} = \delta_{jk}(1 - \delta_{il}) \quad (4)$$

其中， δ_{jk} 和 δ_{il} 都是 Kronecker delta 函数，当 $j = k$ 时， $\delta_{jk} = 1$ ，当 $j \neq k$ 时， $\delta_{jk} = 0$ ；同理，当 $i = l$ 时， $\delta_{il} = 1$ ，当 $i \neq l$ 时， $\delta_{il} = 0$ 。因此，当 j 和 k 是同一节点， $i \rightarrow j$ 和 $k \rightarrow l$ 不是同一对节点指向相反方向的有向边时，矩阵元素 $B_{i \rightarrow j, k \rightarrow l} = 1$ ，否则为零。然而，使用 Ihara-Bass 行列式可以更快地计算节点的非回溯矩阵中心性^[28]。定义矩阵

$$C = \begin{bmatrix} A & I - D \\ I & 0 \end{bmatrix} \quad (5)$$

其中， A 是网络的邻接矩阵， I 是单位矩阵， D 是对角矩阵，对角元素为网络中每个节点的度数。非回溯矩阵中心性向量 $\mathbf{v}$ 等于矩阵 C 主特征向量的前 N 个元素。

上述部分中，可以在线图 $L(G)$ 中计算出 H 条超边的特征向量中心性和非回溯矩阵中心性。为了得到超图中节点的特征向量中心性以及非回溯矩阵中心性，使用 $f(e) \in [0, 1]$ 表示每条超边的中心性大小。定义超图 G 中节点 $i \in V$ 的向量中心性为 $\mathbf{v}_i = (v_{i2}, v_{i3}, \dots, v_{iD})$ ^[20]，其中最高阶数 $D = \max\{|e|; e \in E\}$ ， $|e|$ 表示每条超边的大小。对

于每个 $2 \leq m \leq D$ ，分向量

$$\mathbf{v}_{im} = \frac{1}{m} \sum_{i \in e \in E, |e|=m} f(e) \quad (6)$$

表示节点 i 在超图的 m 阶交互中的中心性（超边大小为 m ，是指这个超边连接着 m 个顶点。这可以被理解为表示一个 m 阶的相互交互，因为这个超边表达了这 m 个顶点之间存在某种关系或交互）。

为了更好地表示超图中节点的整体中心性，用线性方法将节点 i 的中心性定义为

$$\mathbf{v}_i = \sum_{2 \leq m \leq D} \frac{\mathbf{v}_{im}}{m} \quad (7)$$

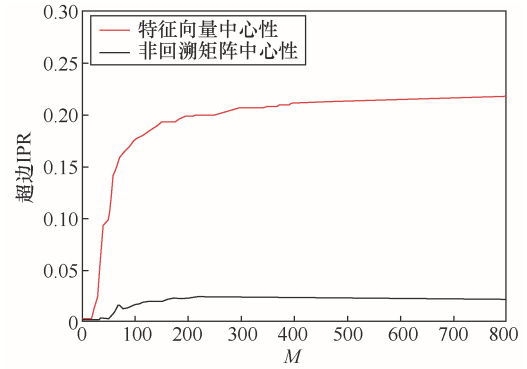
该方法通过计算每个节点在不同大小的超边中的中心角色来评估它的中心性。使用线性组合的方法综合这些不同大小超边的中心性，得到一个表示节点整体中心性的指标。简而言之，该方法通过量化一个节点在所有超边中的相对重要性，来评估其在整个超图网络中的中心性。

3 超图的局域性

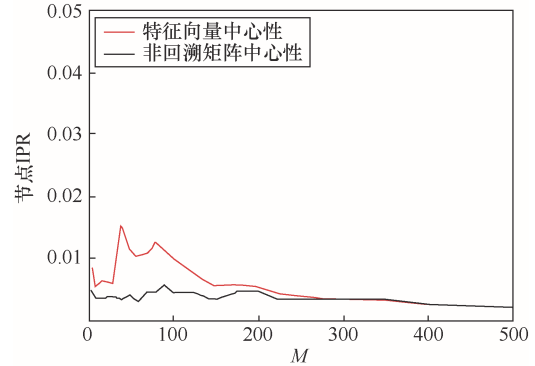
从简单网络扩展到超图，探讨当超边尺寸足够大时，特征向量中心性和非回溯矩阵中心性表现是否不同。假设超图 G 的节点数 $N = 5000$ ，超边数 $H = 5000$ ，节点的平均超度 $\bar{d} = 3$ 。为了满足线图上存在最大度值，将最后一条超边大小设置为 M ($2 \leq M \leq 500$)，其余超边的大小是随机的。随着 M 值的变化，分别计算每个超边 $e \in E$ 的特征向量中心性和非回溯矩阵中心性，记为 $f(e) \in [0, 1]$ 。然后，通过计算逆参与比 (IPR, inverse participation ratio)^[16]来量化超边局域性程度，IPR 值越高，网络局域性程度越大，中心性分布越不均匀。

$$\text{IPR} = \sum_{i=1}^H f(e_i)^4 \quad (8)$$

由图 2(a)可以看出，随着 M 的增大，特征向量中心性的 IPR 值快速增大并趋于稳定，而非回溯矩阵中心性的 IPR 值则缓慢增大并一直处于较低水平。这表明，当超图中存在极大的超边时，超边的特征向量中心性出现局域性。然而，非回溯矩阵中心性可以有效避免这种情况，并且可以更好地衡量超边的重要性。



(a) 超边特征向量中心性和非回溯矩阵中心性的 IPR 值随 M 增大的变化情况



(b) 节点特征向量中心性和非回溯矩阵中心性的 IPR 值随 M 增大的变化情况

图 2 超边和节点特征向量中心性及非回溯矩阵中心性的 IPR 值随 M 增大的变化情况

Figure 2 Changes in the IPR values of hyperedge and node eigenvector centrality, as well as non-backtracking matrix centrality, with increasing M

在超图中，节点似乎比超边受到更多研究。基于超边的中心性，本文进一步探讨 M 的变化对节点中心性的影响，以及节点的特征向量中心性是否会出现局域化，并且通过向量中心性衡量同一节点在不同交互阶数 m 下的重要性。图 2(b)显示了不同 M 下超图中节点特征向量中心性和非回溯矩阵中心性的 IPR 值的变化。可以看出，随着 M 的增加，非回溯矩阵中心性的 IPR 值相对稳定。相比之下，特征向量中心性的 IPR 值先增大后减小。最终两条曲线趋于一致。这说明，在 M 较小的情况下，特征向量中心性可能更加倾向于局部性。但随着 M 的增加，全局性开始起主导作用，从而使得中心性分布变得更加均匀，而节点的非回溯矩阵中心性分布则一直相对均匀。特征向量中心性的 IPR 值先增大后减小是由于超图中总超边基数是固定的，根据式(6)和式(7)，随着 M 的增大，意味着其他 4999 条超边的基数逐渐变小，超图中的结构发生变化，导致某些节点变

得更为重要。但当 M 继续增加, 可能会出现新的高阶相互作用, 使得原先重要的节点不再那么突出。

4 实验与结果

4.1 目标攻击

节点中心性分布是否均匀会极大地影响节点重要性的量化。当 IPR 值较大时, 节点的中心性出现局域性, 大部分权重集中在几个节点上, 影响了其余节点的重要性度量。特别是在目标攻击策略中, 如果无法识别网络中的关键节点, 就会影响目标攻击的效果。

选择几个 M 值分别构造出不同的超图, 并对这些超图进行目标攻击实验。当 $M = 40, 80, 150, 300$ 时, 首先根据两个中心性度量对节点的重要性进行排序, 然后根据排序结果贪婪地选择重要节点进行攻击, 从而探索哪种目标攻击方法可以使网络更快崩溃。本文采用了 3 种攻击方法: 随机攻击、基于特征向量中心性的攻击以及基于非回溯矩阵中心性的攻击。不同 M 值下的 P - R 曲线

如图 3 所示, 其中 P 表示保留的节点比例, R 表示目标攻击 $(1-P)$ 比例的节点后, 超图中极大连通子图所包含的节点与初始节点总数所占比例。

从图 3 可以发现, 当 M 值在一定范围内时, 节点的特征向量中心性出现局域性, 此时非回溯矩阵中心性能够更好地使节点中心性呈现出均匀分布, 从而更好地度量节点的重要性, 即基于非回溯矩阵中心性的攻击方法对网络的破坏效果比基于特征向量中心性的攻击方法好。相比随机攻击以及基于特征向量中心性的攻击, 基于非回溯矩阵中心性的攻击方法只需移除更少节点便可以使得超图完全崩溃, 这反映出面对基于非回溯矩阵中心性的攻击方法, 超图的鲁棒性变得更差。

4.2 随机超图网络

在超图中, 基于具体的实验, 进一步比较分析节点特征向量中心性和非回溯矩阵中心性的差异。例如, 固定 $M = 100$ 生成一个随机超图, 其中节点数 $N = 5\,000$ 、超边数 $H = 5\,000$, 平均超度 $\bar{d}$ 以及平均超边大小 $\bar{m}$ 为 3。图 4 显示了该超

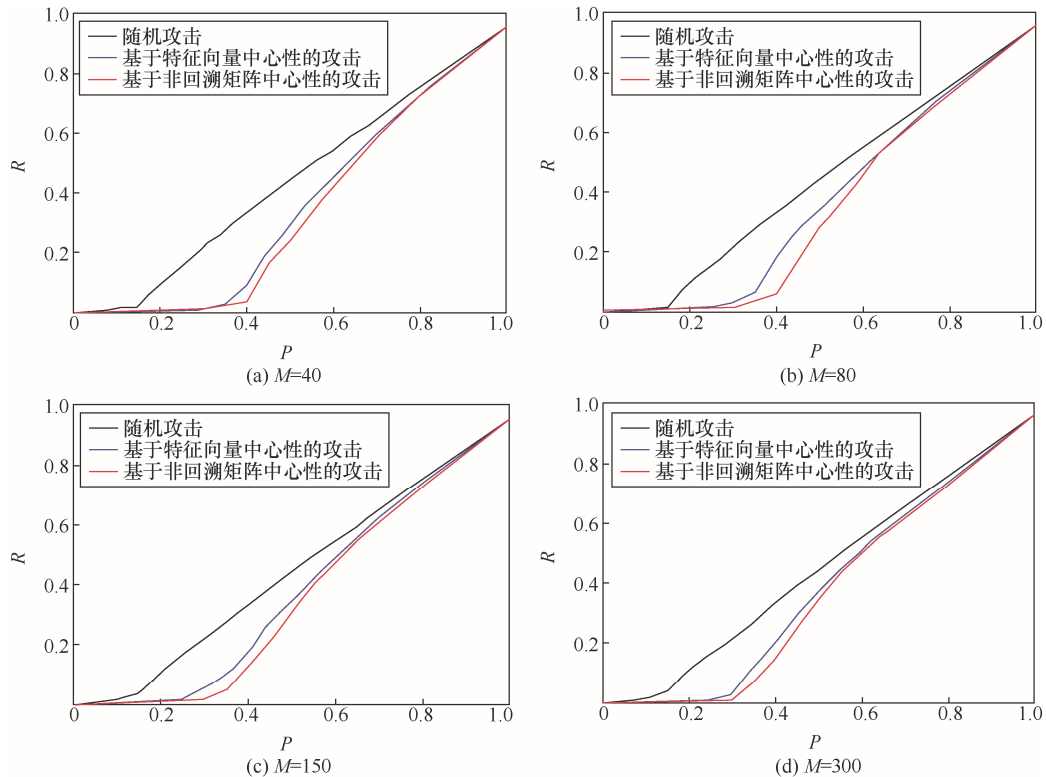


图3 不同 M 值下的 P - R 曲线
Figure 3 The P - R curve of the different M

图的超边大小 m 的分布和节点超度 d 的分布。从图 4 可以看出节点超度分布均匀, 而超边的大小集中分布在 $m=2, 3$, 这是因为节点的平均超度和超边的大小要满足式(2), 而最大超边的大小为 100, 其在整体中占了较大的比重, 从而使得其余超边较小。

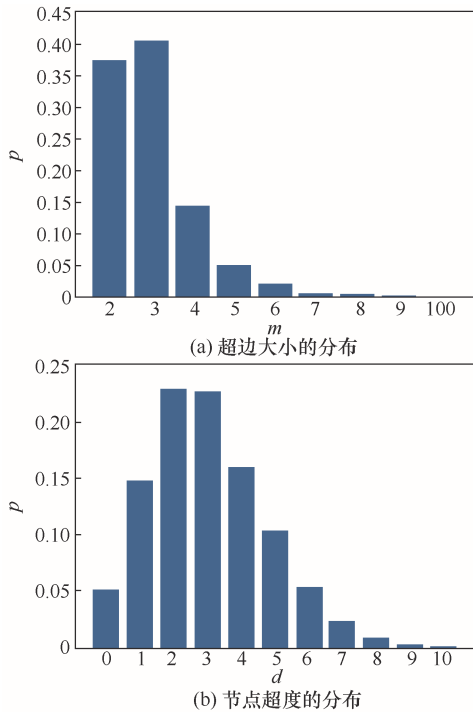


图 4 超边大小 m 的分布和节点超度 d 的分布
Figure 4 Distribution of hyperedge sizes m and node hyperdegrees d

图 5 分别显示了超图中节点特征向量中心性和非回溯矩阵中心性的分布情况。可以看出, 非回溯矩阵中心性分布更加均匀, 而特征向量中心性的权重主要集中在少数节点, 导致低中心度值比例较大, 这进一步证实了先前提到的结论。

在上述内容中, 本文分析了节点特征向量中心性和非回溯矩阵中心性的相关性质和差异。接下来, 本文将深入研究不同交互阶数 m 下, 节点特征向量中心性和非回溯矩阵中心性分量的差异。选择节点的特征向量中心性和非回溯矩阵中心性的前 4 个分量, 其分布情况如图 6 所示, 其中 v_i 表示 v 中的第 i 个分量, $i=2, 3, 4, 5$ 表示相互作用的阶数。可以看出, 各分量中的非回溯矩阵中心性分布比特征向量中心性更均匀, 并且特征向量中心性中低中心性值所占的比例更

大。此外, 还发现各个分量向量的中心性分布不同, 这说明不同交互阶数下节点的影响力也不同。

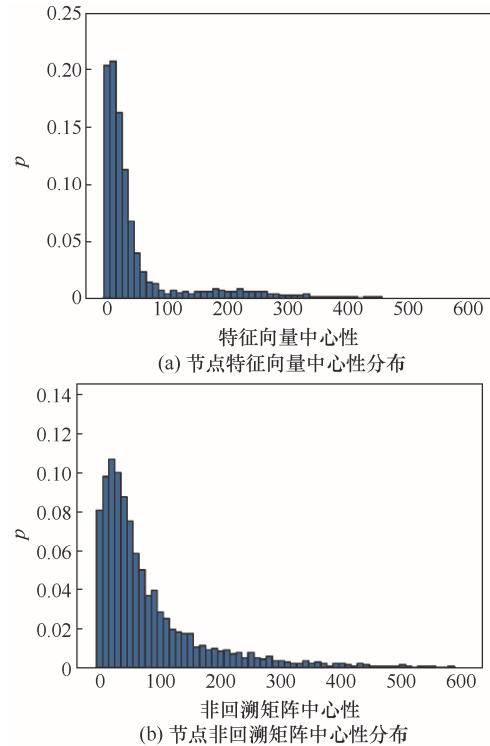


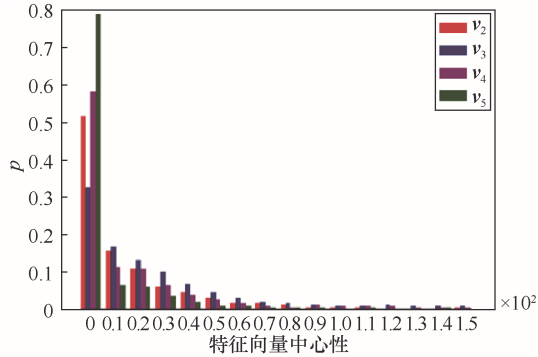
图 5 超图中节点特征向量中心性分布和非回溯矩阵中心性分布
Figure 5 Distribution of eigenvector centrality and non-backtracking matrix centrality in hypergraph nodes

为了探究节点在不同阶数下的平均影响力, 在本文所构造的随机超图中, 超边大小的分布情况如图 4(a)所示, 分别为 $m=2, 3, 4, 5, 6, 7, 8, 9, 100$, 因此这标志着节点的向量中心性有 9 个分量。在每个分量中, 超图中节点在不同阶数下的特征向量中心性和非回溯矩阵中心性的平均值如图 7 所示。

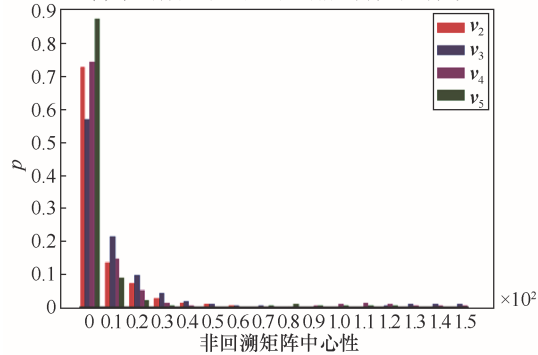
不难发现, 非回溯矩阵中心性的平均值略大于特征向量中心性的平均值, 因为节点的非回溯矩阵中心性的分布更均匀。此外, 还发现节点在各阶相互作用中的平均中心性的分布情况与超图中超边大小的分布基本一致。例如, 超图中 $m=3$ 的超边数量最多, 该类超边的线图 $L(G)$ 计算出的中心性值之和也最大, 节点分量 v_3 的中心性平均值也就最大。

基于特征向量中心性和非回溯矩阵中心性, 对于每个 m 阶交互, 按重要性对节点进行排序, 并分别输出前 100 个重要的节点, 然后比较每对

重要节点之间的重复率。不同阶数下前 100 个重要节点的重复率比较如图 8 所示。可以清楚地看到, 不同阶数交互作用下, 前 100 个重要节点的列表存在显著差异, 这进一步证实了向量中心性在提取节点影响力差异方面的重要性。



(a) 节点特征向量中心性的前4个分向量分布



(b) 节点非回溯矩阵中心性的前4个分向量分布

图6 超图中节点特征向量中心性和非回溯矩阵中心性的前4个分向量分布情况

Figure 6 Distribution of the first four components of eigenvector centrality and non-backtracking matrix centrality in hypergraph nodes

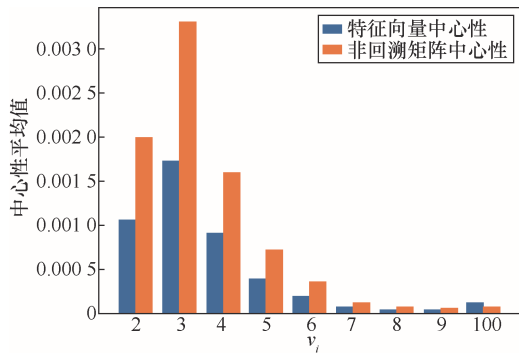


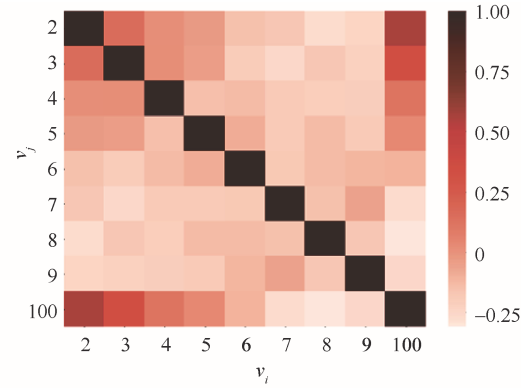
图7 超图中节点在不同阶数下的特征向量中心性和非回溯矩阵中心性的平均值

Figure 7 Average eigenvector centrality and non-backtracking matrix centrality of nodes in hypergraph at different orders

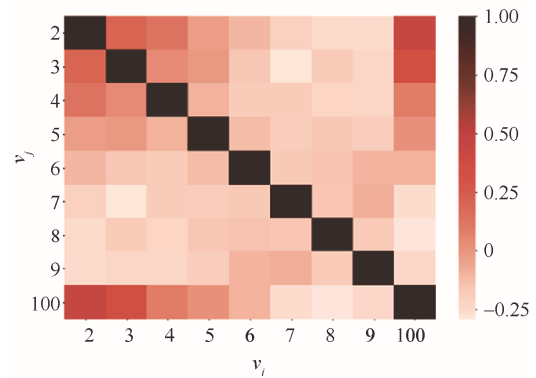
4.3 真实世界超图网络

本文比较了 6 个真实世界超图: 会议、高中、

工作场所^[21]以及 3 个代谢超图网络^[29] (iAB RBC_283、iAF692、iAF1260b) 的特征向量中心性和非回溯矩阵中心性之间的差异。



(a) 向量中心性值基于特征向量中心性



(b) 向量中心性值基于非回溯矩阵中心性

图8 不同阶数下前 100 个重要节点的重复率比较
Figure 8 Comparison of repetition rates of top 100 important nodes at different orders

表 1 显示了 6 个真实世界中的超边和节点的数量以及超边和节点的特征向量中心性和非回溯中心性的 IPR 值。可以发现, 即使在不具有特别大 M 值超边的情况下, 非回溯矩阵中心性的 IPR 值是普遍低于特征向量中心性的, 这说明非回溯矩阵中心性相比于特征向量中心性更不容易出现局域性。

4.4 两种中心性的时间复杂度

要计算特征向量中心性和非回溯矩阵中心性, 关键的步骤是求解它们相关矩阵的主要特征向量。此操作的时间复杂度取决于所用的算法和矩阵的特性。

(1) 特征向量中心性

特征向量中心性基于图的邻接矩阵。如果采用常规的特征分解算法, 那么对于一个 $N \times N$ 的

表 1 6 个真实世界超图的相关数据
Table 1 Relevant data of six real-world hypergraphs

真实世界超图	超边数	节点数	超边特征向量 中心性的 IPR	超边非回溯矩阵 中心性的 IPR	节点特征向量 中心性的 IPR	节点非回溯矩阵 中心性的 IPR
会议	1 436	403	0.003 8	0.003 7	0.137 9	0.137 6
高中	1 453	327	0.004 7	0.004 6	0.070 0	0.069 1
工作场所	662	92	0.005 8	0.005 6	0.146 3	0.142 1
iAB RBC_283	825	342	0.002 25	0.002 24	0.403 3	0.403 2
iAF692	1 203	628	0.001 439	0.001 438	0.435 28	0.435 24
iAF1260b	3 630	1 668	0.000 574	0.000 573	0.494 727	0.494 715

邻接矩阵, 时间复杂度大约是 $O(N^3)$ 。但对于稀疏矩阵或使用迭代算法(如幂方法), 这个时间复杂度可以降低到 $O(N^2)$ 或更低。

(2) 非回溯矩阵中心性

非回溯矩阵的大小是与图的边数有关的。设图有 H 条边, 则非回溯矩阵的大小大约是 $2H \times 2H$ 。同样, 如果采用常规的特征分解算法, 那么时间复杂度是 $O(H^3)$ 。但对于稀疏矩阵或使用迭代方法, 这个时间复杂度可以降低到 $O(H^2)$ 或更低。

4.5 两种中心性的适用条件

(1) 特征向量中心性

特征向量中心性通常适用于度量节点在网络中的影响力, 其假定最重要的节点是那些与其他重要节点有直接联系的节点。在超图网络中, 它可以用来识别不仅在最大超边中, 而且在整个网络结构中有影响力的节点。

(2) 非回溯矩阵中心性

而非回溯矩阵中心性则专门针对网络结构可能导致传统中心性度量失效的复杂网络, 如含有多环结构的网络。它可以减小这些环结构对中心性计算的影响, 因此在分析超图这样的高阶网络结构时尤为适用, 可以更准确地反映节点的中心性。

具体这两种中心性度量的适用条件取决于超图的具体特性, 如节点连接的复杂性、网络的稀疏性以及超边的大小分布等。在实际应用中, 选择合适的中心性度量需要根据网络的具体结构和研究目标来决定。

5 结束语

当超图中存在较大的超边时, 特征向量中心性表现出明显的局域性, 而超边的非回溯矩阵中心性可以有效避免这种局限性。当从超边映射到节点时, 引入了向量中心性, 揭示了同一节点在不同阶数交互作用下的影响。此外, 在不同的交互阶数下, 本文还比较了节点特征向量中心性和非回溯矩阵中心性的平均值, 以及分向量的分布。这些实验结果表明, 非回溯矩阵中心性优于特征向量中心性。这是因为非回溯矩阵中心性能够使节点的重要性程度分布地更加均匀, 从而能够更好地衡量超图中每个节点的重要性。并且, 本文在 6 个真实世界超图上验证了仿真模拟中得到的结论。本文为研究超图上的可靠性提供了思路, 为今后设计更加可靠的超图系统提供了理论依据, 考虑到许多实际网络是动态变化的, 未来的研究可以探索如何在动态超图上应用非回溯矩阵中心性。

参考文献:

- [1] COHEN R, HAVLIN S. Complex networks: structure, robustness and function[M]. Cambridge: Cambridge University Press, 2010.
- [2] TU H, XIA Y, IU H H C, et al. Optimal robustness in power grids from a network science perspective[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2018, 66(1): 126-130.
- [3] LI D, LIU J. Modeling influence diffusion over signed social networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(2): 613-625.
- [4] DING R, UJANG N, HAMID H B, et al. Application of complex networks theory in urban traffic network researches[J]. Networks

- and Spatial Economics, 2019, 19: 1281-1317.
- [5] ALBERT R, BARABÁSI A L. Statistical mechanics of complex networks[J]. Reviews of Modern Physics, 2002, 74(1): 47.
- [6] BOCCALETTI S, LATORA V, MORENO Y, et al. Complex networks: structure and dynamics[J]. Physics Reports, 2006, 424(4-5): 175-308.
- [7] JACKSON M O. Social and economic networks[M]. Princeton: Princeton University Press, 2008.
- [8] 朱会虎, 邱菡, 朱俊虎, 等. 基于传播动力学的域间路由系统关键节点识别方法[J]. 网络与信息安全学报, 2019, 5(5): 9-20.
- ZHU H H, QIU H, ZHU J H, et al. Spreading dynamics based key nodes identification in inter-domain routing system[J]. Chinese Journal of Network and Information Security, 2019, 5(5): 9-20.
- [9] BRÓDKA P, SKIBICKI K, KAZIENKO P, et al. A degree centrality in multi-layered social network[C]//Proceedings of 2011 International Conference on Computational Aspects of Social Networks (CASoN). 2011: 237-242.
- [10] WEI D, LI Y, ZHANG Y, et al. Degree centrality based on the weighted network[C]//Proceedings of 2012 24th Chinese Control and Decision Conference (CCDC). 2012: 3976-3979.
- [11] BARTHELEMY M. Betweenness centrality in large complex networks[J]. The European Physical Journal B, 2004, 38(2): 163-168.
- [12] NEWMAN M E J. A measure of betweenness centrality based on random walks[J]. Social networks, 2005, 27(1): 39-54.
- [13] RUHNAU B. Eigenvector-centrality—a node-centrality[J]. Social Networks, 2000, 22(4): 357-365.
- [14] BONACICH P. Some unique properties of eigenvector centrality[J]. Social Networks, 2007, 29(4): 555-564.
- [15] SCHNEIDER C M, YAZDANI N, ARAÚJO N A M, et al. Towards designing robust coupled networks[J]. Scientific Reports, 2013, 3(1): 1969.
- [16] MARTIN T, ZHANG X, NEWMAN M E J. Localization and centrality in networks[J]. Physical Review E, 2014, 90(5): 052808.
- [17] BRETTO A. Hypergraph theory[M]. Berlin: Springer, 2013.
- [18] BATTISTON F, CENCETTI G, IACOPINI I, et al. Networks beyond pairwise interactions: structure and dynamics[J]. Physics Reports, 2020, 74: 81-92.
- [19] YADATI N, NIMISHAKAVI M, YADAV P, et al. Hypergen: a new method for training graph convolutional networks on hypergraphs[J]. Advances in Neural Information Processing Systems, 2019, (32).
- [20] KOVALENKO K, ROMANCE M, VASILYEVA E, et al. Vector centrality in hypergraphs[J]. Chaos, Solitons & Fractals, 2022, 162: 112397.
- [21] PENG H, QIAN C, ZHAO D, et al. Targeting attack hypergraph networks[J]. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2022, 32(7).
- [22] PENG H, QIAN C, ZHAO D, et al. Disintegrate hypergraph networks by attacking hyperedge[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(7): 4679-4685.
- [23] TUDISCO F, HIGHAM D J. Node and edge nonlinear eigenvector centrality for hypergraphs[J]. Communications Physics, 2021, 4(1): 201.
- [24] MISHRA A, JALAN S, alan S. Eigenvector localization in hypergraphs: pairwise versus higher-order links[J]. Physical Review E, 2023, 107(3): 34311.
- [25] 周丽娜, 常笑, 胡枫. 利用邻接结构熵确定超网络关键节点[J]. 计算机工程与应用, 2022, 58(8): 76-82.
- ZHOU L N, CHANG X, HU F. Using adjacent structure entropy to determine vital nodes of hypernetwork[J]. Computer Engineering and Applications, 2022, 58(8): 76-82.
- [26] BERMOND J C, HEYDEMANN M C, SOTTEAU D. Line graphs of hypergraphs I[J]. Discrete Mathematics, 1977, 18(3): 235-241.
- [27] LANG S. Introduction to linear algebra[M]. Springer Science & Business Media, 2012.
- [28] KRZAKALA F, MOORE C, MOSSEL E, et al. Spectral redemption in clustering sparse networks[J]. Proceedings of the National Academy of Sciences, 2013, 110(52): 20935-20940.
- [29] ZHANG M, CUI Z, JIANG S, et al. Beyond link prediction: predicting hyperlinks in adjacency space[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018.

[作者简介]



彭浩 (1982—), 男, 江苏泰州人, 浙江师范大学教授、博士生导师, 主要研究方向为人工智能安全与网络安全。



钱程 (1997—), 男, 安徽宣城人, 浙江师范大学博士生, 主要研究方向为高阶网络的可靠性。



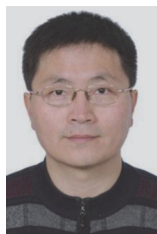
赵丹丹 (1982—), 女, 浙江台州人, 博士, 浙江师范大学副教授, 主要研究方向为复杂网络。



钟鸣（1990— ），男，浙江金华人，博士，浙江师范大学讲师，主要研究方向为人工智能安全与网络安全。



谢紫伊（1998— ），女，浙江台州人，浙江师范大学硕士生，主要研究方向为高阶网络的可靠性。



韩建民（1969— ），男，辽宁大连人，浙江师范大学教授、博士生导师，主要研究方向为数据发布隐私保护。



王伟（1988— ），男，重庆人，重庆医科大学教授、博士生导师，主要研究方向为流行病传播建模与防控、网络医学以及网络科学。