

二阶邻居节点信息驱动의节点重要性排序算法

高 颖*, 董洁霜, 潘 杰, 周亦威, 朱美宣

上海理工大学管理学院, 上海

收稿日期: 2025年3月16日; 录用日期: 2025年4月9日; 发布日期: 2025年4月17日

摘 要

在复杂网络中, 挖掘重要节点对精准推荐、交通管控、谣言控制和疾病遏制等应用至关重要。为此, 本文提出一种局部信息驱动的节点重要性排序算法Leaky Noisy Integrate-and-Fire (LNIF)。该算法通过获取节点的二阶邻居信息计算节点重要性, 首先通过信息量衡量节点重要性, 然后叠加符合条件的邻居节点重要性, 进一步区分节点重要程度, 最终获得节点排序。为验证LNIF的有效性, 在11个真实数据集和1个网络模型上进行了实验, 采用平均度、网络效率、最大连通子图系数和SIR传播模型四个指标进行对比分析。实验结果表明, LNIF生成的节点序列在多个关键指标上表现显著优于现有基准方法。LNIF计算出的节点序列使网络平均度、网络效率和最大连通子图系数下降最快, 表明其在优化网络结构和识别关键节点方面效率更高; 在SIR传播模型中, LNIF的节点序列传播能力也优于其他五种算法, 进一步证明了其在遏制和引导信息传播方面的潜力。

关键词

复杂网络, 节点重要性, 鲁棒性, SIR传播模型

Second-Order Neighbor Node-Driven Importance Ranking Algorithm

Ying Gao*, Jieshuang Dong, Jie Pan, Yiwei Zhou, Meixuan Zhu

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 16th, 2025; accepted: Apr. 9th, 2025; published: Apr. 17th, 2025

Abstract

In complex networks, identifying important nodes is crucial for applications like recommendation systems, traffic control, rumor containment, and disease transmission management. To address this,

*通讯作者。

文章引用: 高颖, 董洁霜, 潘杰, 周亦威, 朱美宣. 二阶邻居节点信息驱动的节点重要性排序算法[J]. 建模与仿真, 2025, 14(4): 454-470. DOI: 10.12677/mos.2025.144301

we propose the Leaky Noisy Integration-and-Fire (LNIF) algorithm, which ranks node importance using second-order neighbor information. LNIF first measures node importance by calculating information quantity, then refines the ranking by aggregating the importance of qualified neighbor nodes. Experiments on 11 real datasets and 1 network model, using metrics like average degree, network efficiency, maximum connectivity subgraph coefficient, and the SIR propagation model, demonstrate LNIF's effectiveness. Results show that LNIF outperforms existing methods, with its node sequences causing the fastest declines in average degree, network efficiency, and maximum connected subgraph coefficient, highlighting its efficiency in optimizing network structure and identifying key nodes. Additionally, in the SIR model, LNIF's node sequences exhibit superior propagation capabilities, further proving its potential in controlling and guiding information spread.

Keywords

Complex Network, Node Importance, Robustness, SIR Propagation Model

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在科技驱动的互联网快速发展中, 复杂网络逐渐兴起, 成为社会学、计算机科学、管理学、物理、数学、生物、交通等多个领域的研究热点[1]-[3]。在现实世界中, 许多复杂系统可以抽象为网络结构, 常见的网络结构有: 在社会关系网络中, 有社交网络、邮件网络、科学家合作网络等[4] [5]; 在交通系统中, 有公路、铁路、航空网络等[6] [7]; 在生物领域中的蛋白质网络、神经网络、新陈代谢系统等[8] [9]。在这些多样的网络结构中, 小世界和无标度特性使某些节点具有极大影响力, 其行为对网络功能和结构产生重大影响, 这些节点被称为重要节点[10]-[12]。因此, 准确快速识别复杂网络中的重要节点是当前的热点之一。

挖掘网络中的重要节点对于现实生活具有重要意义。在传染病传播中, 构建传染病网络, 找到关键传染源, 可以大大降低传染病的传播速度[13] [14]; 在设备网络中对关键节点采取保护措施, 可以增强设备网络的整体抗毁性[15]; 在交通网络中, 识别关键路段, 并有效对这些关键路段进行管控, 可以降低高峰期交通网络的堵塞情况[6]; 在谣言传播的网络中[16], 谣言传播经过众多节点, 尤其重要节点能加速信息扩散。快速定位这些节点有助于引导舆论, 发挥其影响力, 有效遏制谣言传播。

研究人员在节点重要性方面做了很多工作, 许多经典的节点重要性算法被提出, 如度中心性、介数中心性、紧密度中心性、特征向量中心性、PageRank 算法与 K-shell 算法等[17]-[19]。近年来, 随着复杂网络理论的发展, 研究者们提出了新的节点重要性算法, 丰富了复杂网络的相关知识。Liu 等人[20]通过结合网络拓扑与信息传播, 利用邻居节点的影响来对网络中的节点重要性等级进行修正。Lü 等人[21]将用于评价学术论文作者影响力的 H 指数引入到复杂网络节点重要性分析领域。王安等人[22]通过 LPA 算法将网络进行社区划分, 然后在社区内部排序处理, 最终得到节点的综合排序结果, 提高了传播性能。Liu 等人[23]将结构空穴和 K-shell 分解算法结合起来区分网络中的重要节点。Yang 等人[24]结合 K-shell 算法与节点的位置提出一种新的重要性算法, 并通过相关实验验证了该算法的合理性和有效性。

在大规模复杂网络中, 全局节点重要性算法因高时间复杂度不实用, 局部算法虽降低复杂度, 但 H-index 和 K-shell 无法区分同值节点。本文提出 LNIF 算法, 结合二阶邻居信息和信息传播因素, 通过计算节点信息量并引入传播阈值修正, 提升准确性。在 11 个真实数据集和 1 个网络模型上的实验显示, LNIF

在平均度、网络效率、最大连通子图系数和 SIR 传播模型中表现最优，能有效识别重要节点并产生最大影响力。

2. 材料和方法

2.1. 相关理论基础

令无向网络 $G=(V,E)$ ， $V=\{v_1, v_2, \dots, v_n\}$ 代表所有节点的集合， $|V|=n$ ； $E=\{e_1, e_2, \dots, e_n\}$ 代表节点间的连接关系， $|E|=m$ ，网络 G 的邻接矩阵用 A 表示，若节点之间存在连边时，则相应的位置上置 1，否则为 0。

1) 度中心性

度中心性是表示与节点直接相连的节点数量，度中心性的值越大，节点就越重要，计算公式如下：

$$DC(v_i) = \sum_{j=1}^N A_{ij} \quad (1)$$

2) 介数中心性

介数中心性衡量的是某个节点在多大程度上介于(between)其它节点对之间，其中 $\sigma(j,k|i)$ 表示节点 j 到节点 k 经过节点 i 的最短路径数， $\sigma(j,k)$ 为节点 j 到节点 k 的最短路径数。

$$BC(v_i) = \sum_{i \neq j \neq k} \frac{\sigma(j,k|i)}{\sigma(j,k)} \quad (2)$$

3) 紧密度中心性

紧密度中心性的思想是：如果一个节点与许多其他节点都很接近，那么节点处于网络中心位置：

$$CCv_i = \frac{1}{\sum_{j \in V} d_{ij}} \quad (3)$$

其中 d_{ij} 为节点 i 到节点 j 的最短距离。

2.2. 基于局部领域信息的中心性算法

2.2.1. 算法描述

通常认为，网络中度数最大的节点最重要，但实际上节点的重要性不仅取决于度数，还与邻居节点的信息贡献有关。如图 1 所示，一个小型网络被分为三块，尽管某节点的度数小于其邻居，但当该节点受攻击时，网络迅速分裂为三个独立部分，破坏性最大。度数仅表示节点直接连接的边数，无法区分相同度数的节点差异，尤其对于“桥节点”，虽然连接度不高，但在网络中地位关键。此外，从信息传播角度看，节点是各块间信息传输的必经之路，因此信息通过该节点能更广泛传播。信息量可用于衡量节点的重要性。

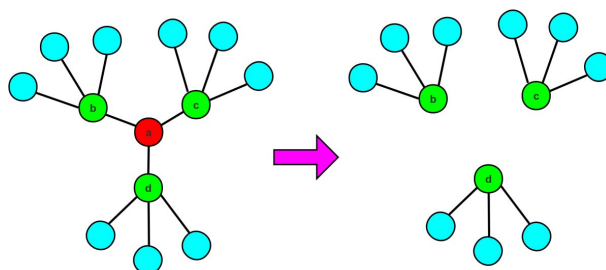


Figure 1. Small networks
图 1. 小型网络

通过衡量信息在网络中的流通来衡量节点的重要性，计算公式如下：

$$\text{LNI}_a = \sum_{i \in \Gamma_a} F_{ai} \quad (4)$$

$$F_{ai} = -[p_{ai} \log p_{ai} + (1 - p_{ai}) \log (1 - p_{ai})] \quad (5)$$

$$p_{ai} = \frac{k_i}{\sum_{j \in \Gamma_a} k_j} \quad (6)$$

其中 k_a 表示节点 a 的度数， Γ_a 表示节点 a 的一阶邻居。

为了更有效筛选出核心节点，比如“桥节点”，将符合条件的节点邻居的重要性叠加到节点本身，来进一步区分节点的重要程度。同时，在信息传播方面，存在一定阈值，例如在疾病传播模型中，当感染阈值大于感染节点与易感节点之间的阈值时，疾病才可以进行传播；当感染阈值小于感染节点与易感节点之间的阈值时，不进行传播。基于上述思想，将符合条件的节点邻居的重要性叠加到节点本身来进一步区分节点的重要程度。计算公式如下：

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$\text{LNIF}(i) = \sum_{j \in \Gamma(i)} f(\text{LNI}(i) - \text{LNI}(j)) \quad (8)$$

若根据节点的度数来确定节点的重要性，则可能会忽略节点的信息，如图 1 中的节点 a 。但是 LNIF 算法综合考虑了节点的二阶内邻居信息以及信息传播的因素来衡量节点的重要性值，若某节点的 LNIF 值越大，则说明该节点在网络中越不能被替代，其重要性就越高，节点就越重要。

算法描述

输入：网络 $G = (V, E)$

输出：节点的 LNIF 重要性值

对于节点 a ：

对于节点 a 的邻居节点 i ：

$$p_{ai} = \frac{k_i}{\sum_{j \in \Gamma_a} k_j}$$

$$F_{ai} = -[p_{ai} \log p_{ai} + (1 - p_{ai}) \log (1 - p_{ai})]$$

结束

$$\text{LNI}_a = \sum_{i \in \Gamma_a} F_{ai}$$

结束

对于网络中的节点：

对于节点 $N(\text{node})$ 的邻居节点 $\text{neighbour}(\text{nbr})$ ：

如果 $f(\text{LNI}(\text{node}) - \text{LNI}(\text{nbr}))$ ：

$$\text{LNIF}(\text{node}) = \text{LNIF}(\text{node}) + f(\text{LNI}(\text{node}) - \text{LNI}(\text{node}))$$

结束

结束

结束

以图 2 中的节点 0 为例来说明算法第一部分如何计算, 首先计算节点 a 与它全部邻居节点的 F 值, 这里计算得到每个节点的结果如下: $F_{01}=0.391$, $F_{02}=0.619$, $F_{03}=0.890$, $F_{04}=0.779$, $F_{05}=0.779$ 。对于节点 0 来说, $LNI(0)=F_{01}+F_{01}+F_{01}+F_{01}+F_{01}=3.46$, 同理可以求出剩下几个节点的 E 值。

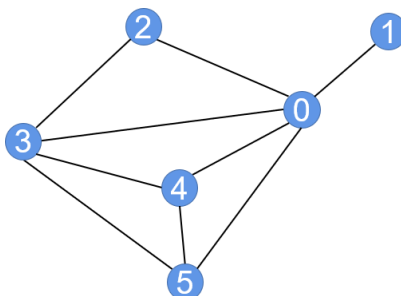


Figure 2. Algorithm illustration diagram
图 2. 算法说明图

2.2.2. 时间复杂度分析

设网络 $G=(V,E)$ 中含有的节点数为 N , 边数为 M , 整个网络的平均度为 $\langle k \rangle$ 。由图 3 可以知道, 首先计算节点的 E 值, 这里面包含了 2 个循环, 一个大循环和一个小循环。其中大循环需要遍历全部节点, 而小循环需要遍历每个节点的邻居节点, 所以这两个循环的时间复杂度为 $O(\langle k \rangle N)$ 。当计算完全部节点的 E 值后, 需要计算各节点的 LIE 值, 此时也包含两个循环, 与上一步一样, 其时间复杂度为 $O(\langle k \rangle N)$, 所以本文的 LIE 算法的总时间复杂度为 $O(2\langle k \rangle N) \sim O(N)$ 。

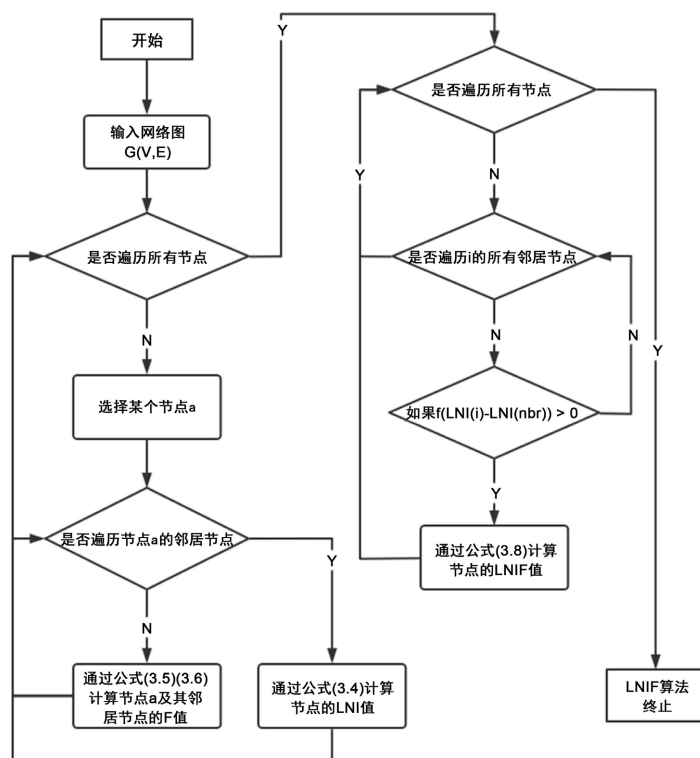


Figure 3. Flow chart of LNIF algorithm
图 3. LNIF 算法流程图

2.3. 评价标准分析

基于复杂网络的鲁棒性对节点重要性的结果进行评价，主要采用网络平均度、网络效率和最大连通子图系数三个指标来量化移除节点及其连边后，网络连通性和网络结构受到的影响，以此来评价节点的结构重要性。此外还有基于网络的传播学原理来衡量节点的重要性算法，如 SI 算法、SIR 算法等，在这些算法中节点的重要性是通过该节点的平均传播范围决定。在本文中采用网络平均度、网络效率、最大连通子图系数和 SIR 传播模型这 4 个指标来衡量 LNIF 算法的优越性。

2.3.1. 最大连通子图系数

在社会关系中，如果一个人没有任何朋友，或者一群人除他们外没有别的朋友，那么在整个网络中，这个(群)人就是孤立存在的，对于整个网络来说，是不连通的。对于复杂网络来说，任何一个节点的行为都有可能影响整体网络的连通性，而最大连通子图系数是用来评估网络连通性的指标。将网络中的所有节点按重要性排序算法进行排序，移除节点后，对网络连通性的影响。最大连通子图系数(G)的计算公式如下：

$$G = \frac{r}{n} \quad (9)$$

其中， r 表示移除节点后的网络中的最大连通子图系数的节点数目， n 表示网络的节点数量。若最大连通子图系数随着节点移除而变小的趋势越明显，说明节点重要性排序算法越准确。

2.3.2. 网络效率

利用网络中节点间信息交换的效率来评价节点重要性排序算法。使用网络效率来量化网络中交换信息的效率，用于评价网络连通性的强弱，当移除某个节点以及所有的连边时，网络某些路径被切断使得一些节点间的最短路径变大，从而增加了整个网络的平均路径长度，影响网络的信息连通性。网络效率(ED)计算公式如下：

$$ED = \frac{1}{n(n-1)} \sum_{i \neq j \in V} \frac{1}{d_{ij}} \quad (10)$$

其中， n 表示网络的节点数量， d_{ij} 表示节点 i 和 j 的最短路径。当节点 i 和节点 j 之间不存在可达路径时， $d_{ij} = +\infty$ ，而 $1/d_{ij} = 0$ ，在无向图中，网络效率的变化范围是 $[0,1]$ 。通过删除一定数量的节点，模拟现实世界中网络遭受攻击时的效果，通过计算网络被攻击前后网络效率下降的比例来量化衡量节点重要性评价指标的准确性。

2.3.3. 平均度

网络中所有节点 i 的度 k_i 的平均值称为网络的平均度，记为 $\langle k \rangle$ ，其计算公式如下：

$$\langle k \rangle = \frac{\sum_{i=1}^N k_i}{N} \quad (11)$$

2.3.4. SIR 传播模型

SIR 传播模型是经典的传染病模型，用于评估节点传播能力，其中 S、I、R 分别表示易感者、感染者和恢复者。由于网络信息传播与传染病传播规律相似，SIR 模型被广泛应用于研究信息传播及其演化规律。它可以模拟计算机病毒在通信网络、危机在经济网络、谣言在社交网络中的传播。因此，节点重要性可通过 SIR 模型在网络上进行传播模拟获得。

2.4. 实验数据集

由于不同的网络具有不一样的结构与特征，为了验证 LNIF 算法的有效性，本文共使用 12 个数据集，

包括真实网络数据集 11 个: Football, Adjnoun, Cond-mat, Karate, Dolphins, Polbooks, Polblogs, Power, Hep-th, As-22july06, Astro-ph 以及 1 个人工模拟网络 NW, 数据集的拓扑统计特征如表 1 所示。其中 N 和 M 分别代表数据集网络总节点数与总连边数, $\langle k \rangle$ 表示平均度数, K_{\max} 表示网络中节点的最大度数, C 表示网络中的平均集聚系数。

- Football 数据集: 2000 年美国秋季常规赛期间学院之间的美式橄榄球比赛网络[25];
- Adjnoun 数据集: Charles Dickens 的小说 David Copperfield 中常用形容词和名词的邻接网络[26];
- Cond-mat 数据集: 该数据集是 1995 年到 1999 年凝聚态物理方面的合作网络数据集[27] [28]。
- Karate 数据集: 该数据集为美国一所大学空手道俱乐部成员之间的社交网络[29];
- Dolphins 数据集: 该数据集是新西兰 62 只海豚之间的关系网络[30] [31];
- Polbooks 数据集: 该数据集是在 2004 年美国总统大选期间出版的书籍网络[32];
- Ploblogs 数据集: 该数据集是 2004 年美国总统大选前 2 个月内网络上自由派与保守派之间的互动情况, 在该数据中, 节点表示博客, 连边表示博客之间存在超链接[33];
- Power 数据集: 该数据集是美国西部各州电网的拓扑结构, 其节点表示变压器、变电站、发电机, 而连边则表示高压输电线[34];
- Hep-th: 高能理论合作, 科学家在高能理论电子印刷档案上的预印的加权网络[27] [35]。在该网络中共有 8361 个节点和 15,751 条连接关系。
- As-22july06 数据集: 该数据集是 Newman 收集的在自治系统级别上的互联网结构的对称快照[36];
- Astro-ph 数据集: 该数据集是 1995 年到 1999 年中科学家在天体物理学方面的合著网络[37], 其中节点表示科学家, 连边表示科学家之间的合作关系;
- NW 网络: NW 网络采用随机增加节点的连边来取代边的随机重连[34] [38]。在本文 NW 网络的参数设置为 $N = 500, k = 3, p = 0.5$ 。由于模型每次成的连边数以及平均度 $\langle k \rangle$ 等不一致, 在本文中, NW 实验结果均取 20 次的均值作为最终结果。

Table 1. Topological characteristics of the experimental data set network
表 1. 实验数据集网络的拓扑特征

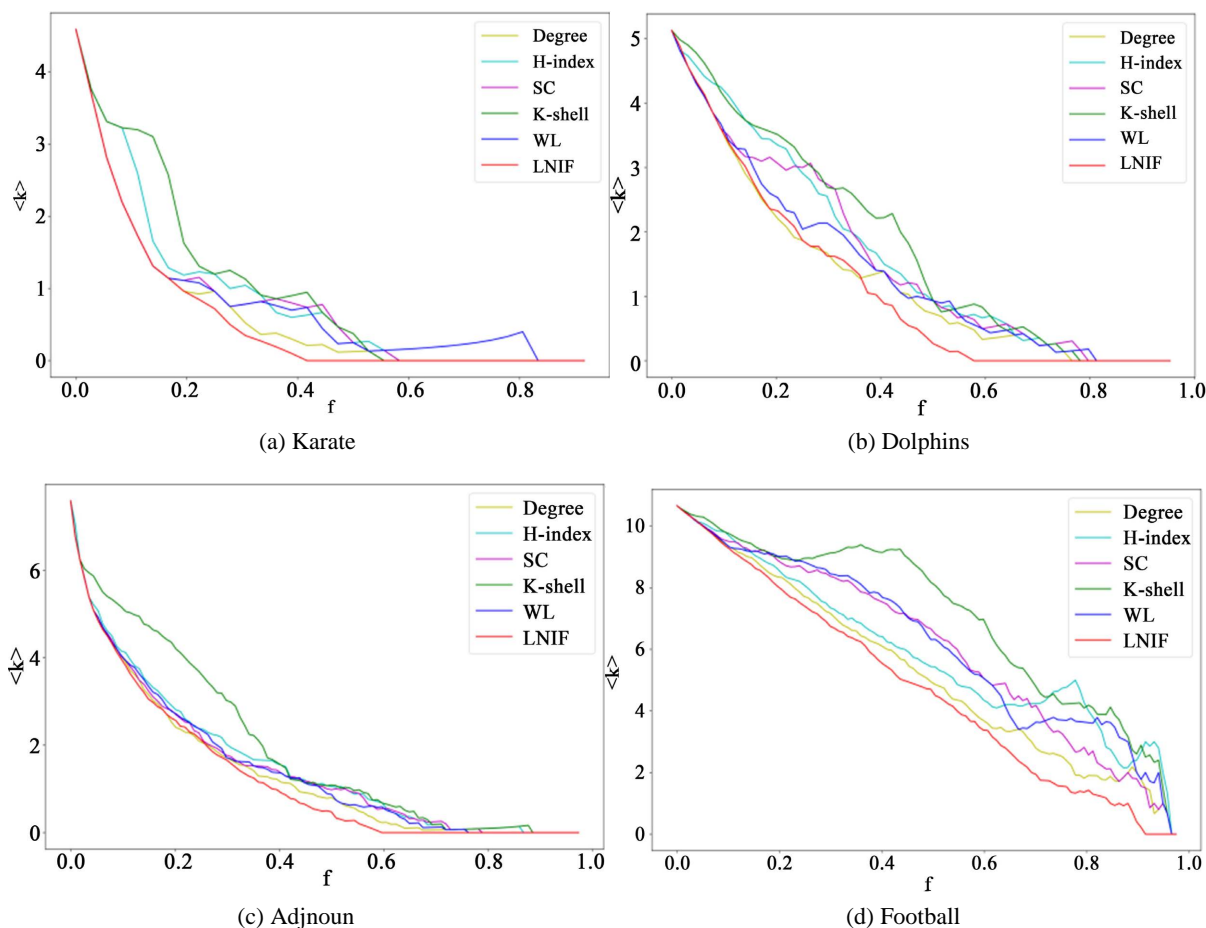
数据集	N	M	$\langle K \rangle$	k_{\max}	C
Adjnoun	112	425	7.59	49	0.17
Football	115	613	10.66	12	0.40
Dolphins	62	159	5.13	12	0.26
Karate	34	78	4.59	17	0.57
Polbooks	105	441	8.40	25	0.49
Polblogs	1490	16,718	22.44	351	0.26
Power	4941	6594	2.67	19	0.08
Hep-th	8361	15,751	3.77	50	0.77
As-22july06	22,963	48,436	4.22	2390	0.23
Cond-mat	16,726	47,594	5.69	107	0.62
Astro-ph	16,706	121,251	14.52	360	0.64
NW	500	750.05	3.0	6.55	0.003

3. 结果

3.1. 平均度结果

基于上述的其中 6 个真实数据集和一个模拟网络, 本文将采用同样是基于局部信息的 Degree、H-index [21]、SC [39]、WL [40]、K-shell [41]与本文的 LNIF 算法进行对比和分析。根据 6 种算法得到的节点排序结果, 以静态攻击的方式将网络中的节点按照排序结果进行移除, 模拟网络遭受攻击时网络平均度 $\langle k \rangle$ 的变化情况, 以此来评价算法的准确性。本文所有实验均采用 Python 3.7。

图 4 展示了不同节点重要性算法下, 移除节点后网络平均度的变化, 横坐标 f 表示移除节点比例, 纵坐标表示平均度变化。实验结果如图 4(a)~(f)所示。在图 4(a)的 karate 网络中, LNIF 算法在移除相同比例节点时下降最快, 前 20% 与 WL、SC、Degree 算法曲线重合, 超过 20% 后 LNIF 曲线位于左下方, 最先使 $\langle k \rangle$ 降至 0。图 4(b)的 Dolphins 网络中, 前 40% 时 LNIF 与 Degree 算法曲线接近, 超过 40% 后 LNIF 曲线下降最快, 仅需移除约 60% 节点即可使 $\langle k \rangle$ 降至 0, 而其他算法需移除约 80%。图 4(c)的 Adjnoun 网络中, 前 30% 时 LNIF 与 Degree 算法位于左下方, 超过 30% 后 LNIF 曲线下降最快, 最先使 $\langle k \rangle$ 降至 0。图 4(d)的 Football 网络中, 前 10% 时除 H-index 和 K-shell 外, 其余算法曲线接近, 超过 10% 后 LNIF 曲线位于最下方。图 4(e)的 Polbooks 网络中, LNIF 曲线与图 4(a)类似, 超过 20% 后表现最优。图 4(f)的 Polblogs 网络中, 除 K-shell 外, 各算法曲线接近, 但 LNIF 仍位于最下方。图 4(g)的 NW 网络模型中, LNIF 曲线下降幅度更大, 更为陡峭。



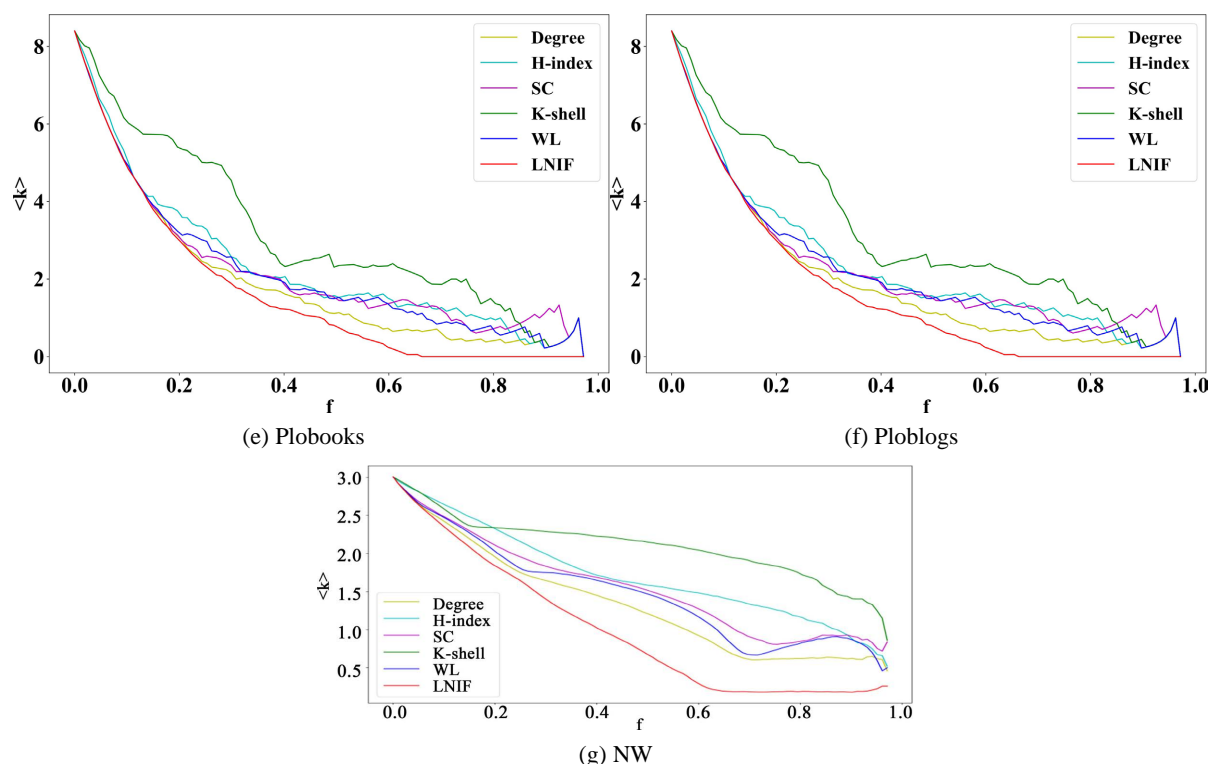


Figure 4. Changes of the average degree $\langle k \rangle$ of the network under different attack methods

图 4. 在不同攻击方法下网络的平均度 $\langle k \rangle$ 的变化

3.2. 网络效率结果

基于上述的其中 6 个真实数据集, 本文将采用同样也是基于局部信息的 Degree、H-index [21]、SC [39]、WL [40]、K-shell [41] 算法与本文的 LNIF 算法进行对比和分析。根据 6 种算法得到的节点排序结果, 以静态攻击的方式将网络中的节点按照排序结果进行移除, 模拟网络遭受攻击时网络效率 ED 的变化情况, 以此来评价算法的准确性。

图 5 展示了不同节点重要性算法下, 移除节点后网络效率的变化, 横坐标 f 表示移除节点比例, 纵坐标表示网络效率。实验仅展示移除 80% 节点的情况, 因网络此时已基本崩溃, 可减少运行时间。在图 5(a) 的 Karate 网络中, LNIF 算法初始略逊于 K-shell, 但之后表现最优。图 5(b) 的 Dolphins 网络中, LNIF 算法在前 20% 与 Degree、WL、SC 算法重合, 20%~40% 时与 Degree 算法表现最佳, 超过 40% 后 LNIF 算法显著优于其他算法。图 5(c) 的 Adjnoun 网络中, LNIF 算法基本最优, 仅小部分略逊于 Degree 算法。图 5(d) 的 Football 网络中, LNIF 算法下降速度远超其他算法。图 5(e) 的 Plobooks 网络中, LNIF 算法在前 20% 与 Degree、WL、SC 算法接近, 随移除比例增加, 优势愈发明显。图 5(f) 的 Ploblogs 网络中, LNIF 算法除初始部分外, 后续表现最佳, 最先使网络效率降至 0。图 5(g) 的 NW 网络模型中, LNIF 算法曲线下降速率更快, 优于其他算法。

3.3. 最大连通子图系数结果

基于上述的其中 6 个真实数据集, 本文将采用同样也是基于局部信息的 Degree、H-index [21]、SC [39]、WL [40]、K-shell [41] 算法与本文的 LNIF 算法进行对比和分析。根据 6 种算法得到的节点排序结果, 以静态攻击的方式将网络中的节点按照排序结果进行移除, 模拟网络遭受攻击时最大连通子图系数 G 的变化情况, 以此来评价算法的准确性。

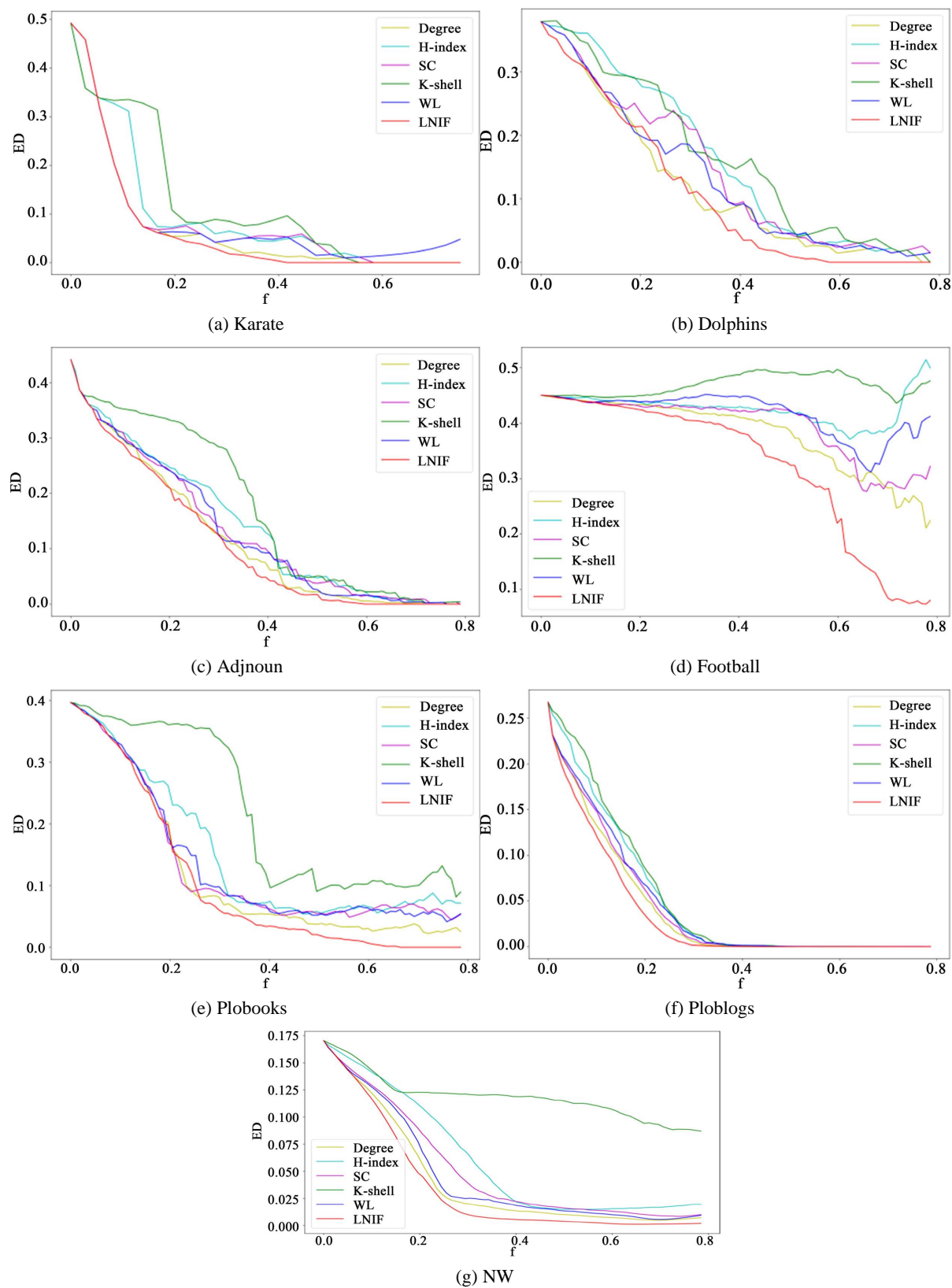
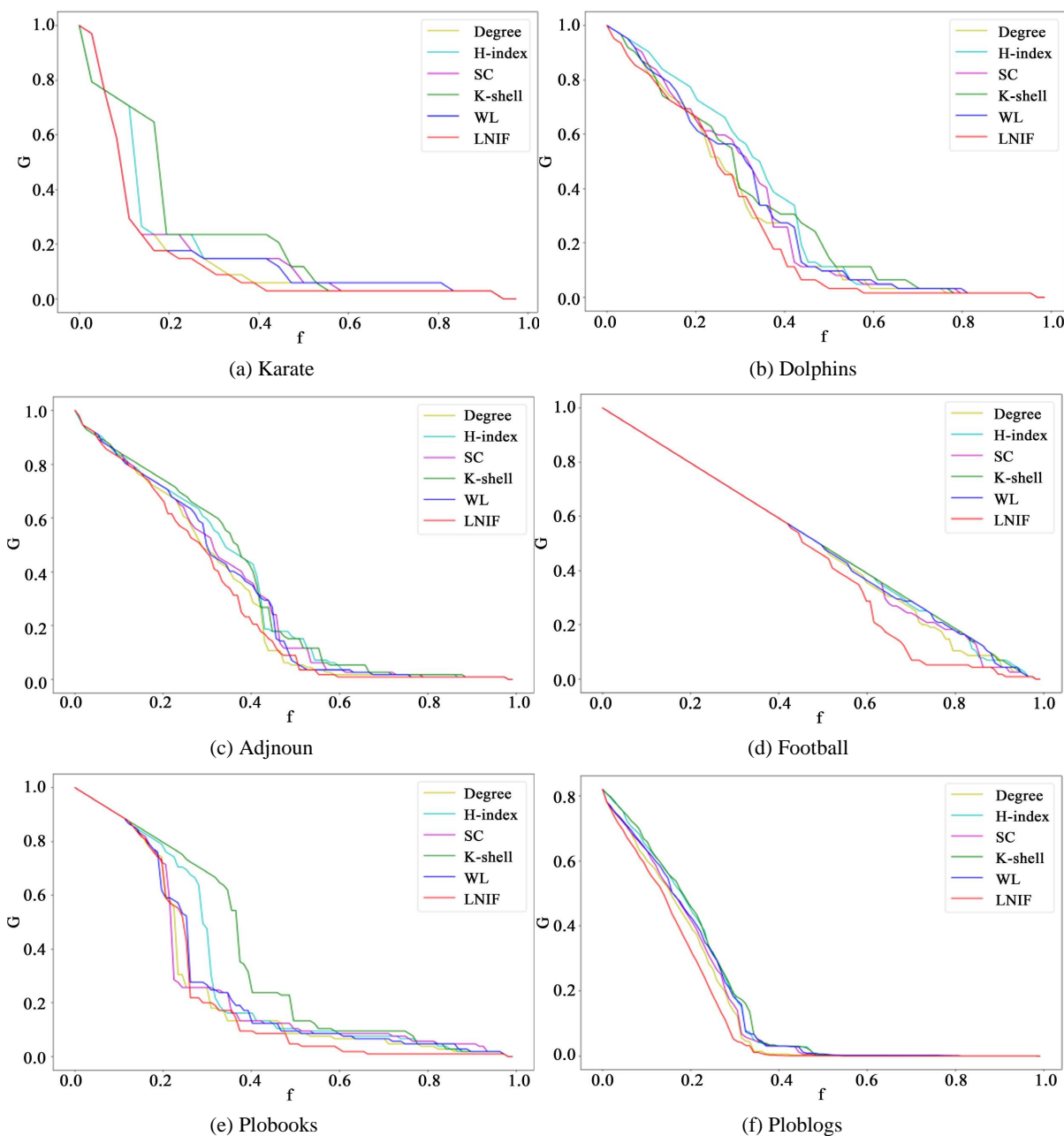


Figure 5. Changes in ED of network efficiency under different attack methods

图 5. 在不同攻击方法下网络的效率 ED 的变化

图 6 展示了不同节点重要性算法下, 移除节点后最大连通子图系数的变化, 横坐标 f 表示移除节点比例, 纵坐标表示最大连通子图系数。实验结果如图 6(a)~(f)所示。在图 6(a)的 Karate 网络中, LNIF 算法初始略差, 但后续表现最优。图 6(b)的 Dolphins 网络中, LNIF 算法在前 30%与 Degree 算法各有优劣, 超过 30%后逐渐领先其他算法。图 6(c)的 Adjnoun 网络中, LNIF 算法曲线基本处于左下方, 仅部分略逊于 Degree 和 WL 算法。图 6(d)的 Football 网络中, 前 40%各算法曲线重合, 超过 40%后 LNIF 算法下降趋势明显, 更能降低网络连通性。图 6(e)的 Plobooks 网络中, LNIF 算法在前 40%略逊于 Degree 和 SC 算法, 与 WL 算法接近, 但优于 H-index 和 K-shell 算法; 超过 40%后逐渐体现优势, 处于其他算法曲线下方。图 6(f)的 Polblogs 网络中, 除重合部分外, LNIF 算法均表现最佳。图 6(g)的 NW 网络模型中, LNIF 算法更能降低网络连通性, 而 K-shell 算法曲线为直线。



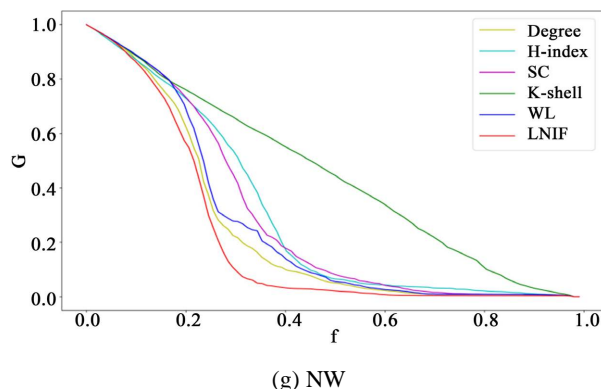


Figure 6. The change of the maximum connectivity subgraph coefficient G under different attack methods
图 6. 在不同攻击方法下网络的最大连通子图系数 G 的变化

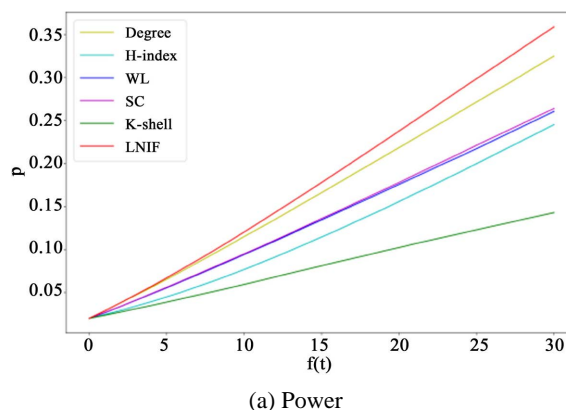
3.4. SIR 实验结果

在 SIR 实验中, 设置的 SIR 的相关参数为感染概率为 0.05, 恢复概率为 0.01, 设置感染轮数为 30, 取不同的初始感染比例 $a = 0.02$ 和 $a = 0.03$ 作为感染源, 所有结果均取 100 次均值。

基于上述的其中 5 个真实数据集, 本文将采用同样是基于局部信息的 Degree、H-index [21]、SC [39]、WL [40]、K-shell [41] 算法与本文的 LNIF 算法进行对比和分析。根据 6 种算法得到的节点排序结果, 选取不同的感染比例来衡量算法的传播能力。对于同一个数据集来说, 不同算法得到的节点排序结果往往不同, 所以选取排序结果的前面部分作为感染源, 以此来衡量算法的优劣。

图 7 表示的是在同一个网络中采用不同的节点重要性算法下, 选取感染比例为 0.02 时, 每次时间下感染比例的变化, 其中横坐标 $f(t)$ 表示感染时间, 纵坐标表示每次时间下的感染比例。实验结果如图 7(a)~(e) 所示。在图 7(a)~(e) 的 Power、Hep-th、Cond-mat、Astro-ph 和 As-22july06 网络中, LNIF 算法的传播性能均优于其余 5 种算法, 尤其是在图 7(a)~(c) 表现得最为明显, 与其他 5 种算法的感染比例相差最大, 而表现最差的算法则是 K-shell 算法。

图 8 表示的是在同一个网络中采用不同的节点重要性算法下, 选取感染比例为 0.03 时, 每次时间下感染比例的变化, 其中横坐标 $f(t)$ 表示感染时间, 纵坐标表示每次时间下的感染比例。实验结果如图 8(a)~(e) 所示。在图 8(a)~(e) 的 Power、Hep-th、Cond-mat、Astro-ph 和 As-22july06 网络中, LNIF 算法的传播性能均优于其余 5 种算法, 在图 8(a) 中的 Power 网络中, LNIF 算法在第 30 次时间的感染比例比 K-shell 算法在第 30 次时间的感染比例的差值要大于 0.2, 这说明 LNIF 算法得到的节点重要性排序结果的传播力更强。LNIF 算法在其他数据集上均有最佳的表现。



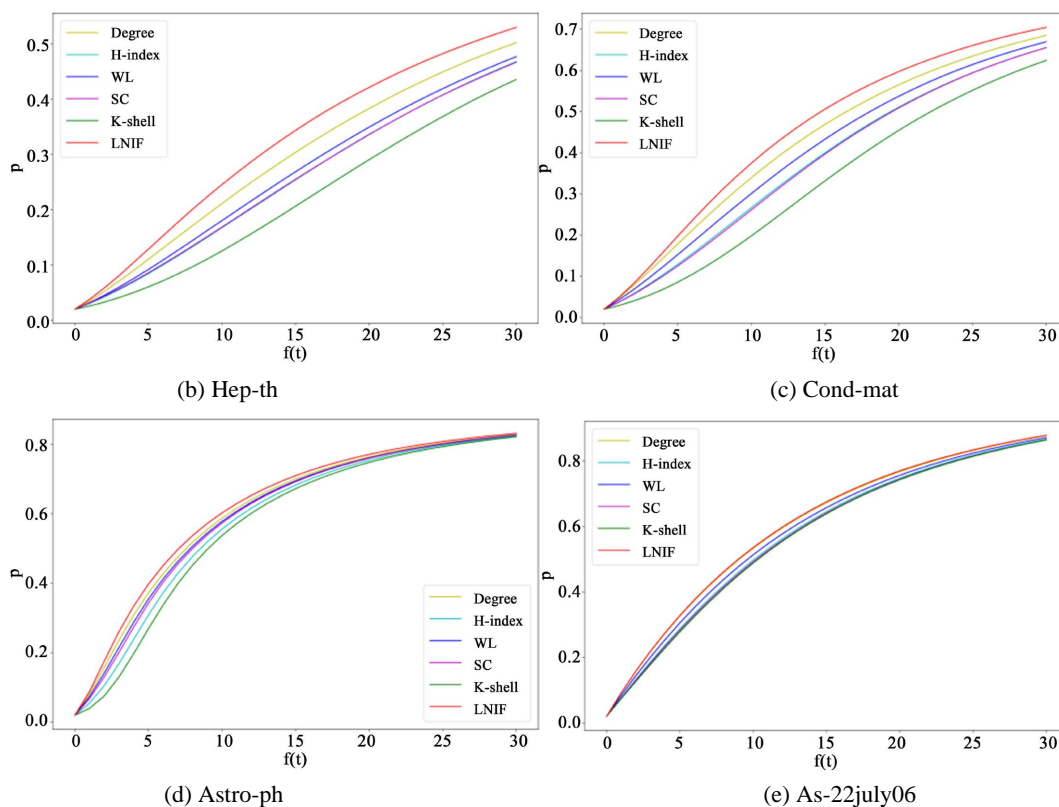
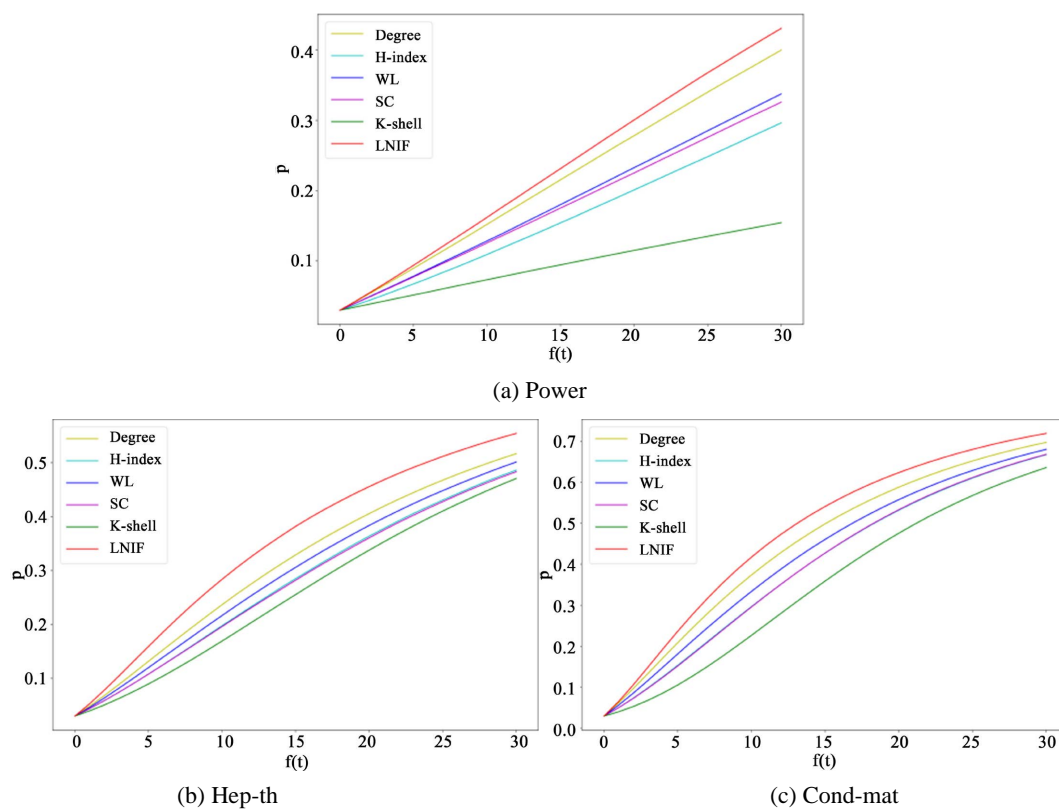


Figure 7. Changes in the proportion of network infection under different node importance algorithms ($a = 0.02$)
图 7. 在不同节点重要性算法下网络的感染比例变化($a = 0.02$)



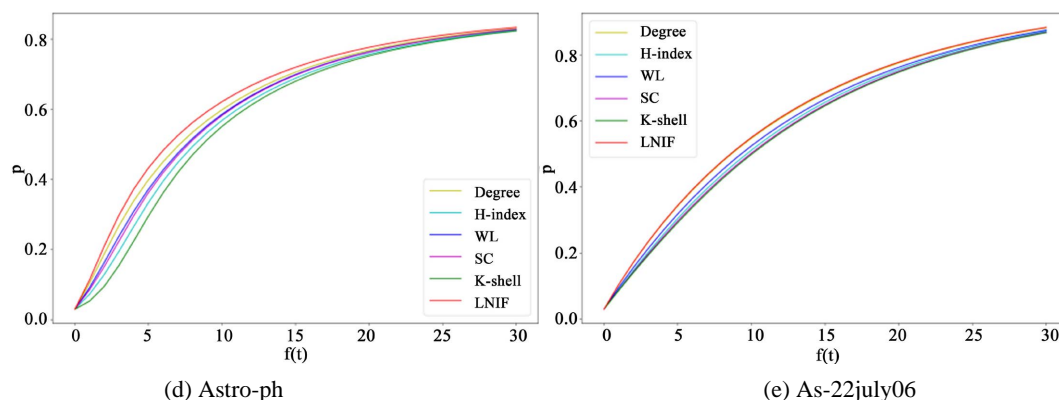


Figure 8. Changes in the proportion of network infection under different node importance algorithms ($a = 0.03$)
图 8. 在不同节点重要性算法下网络的感染比例变化($a = 0.03$)

4. 讨论

节点重要性研究是复杂网络的一个热门方向之一。不同节点重要性算法得到的节点排序往往不同，这主要体现在对节点重要性的定义上的不同[42]-[44]。从网络的拓扑结构来研究节点重要性是常用的一个方法，研究人员在这方面做了很多工作，提出不少经典的节点重要性算法[45]-[47]。从结构上来说分为基于局部信息和基于全局信息来衡量节点的重要性。本文在此总结对比了几种经典的节点重要性算法，详见表 2。为了更直观表示，将设置的平均度 $\langle k \rangle$ 简写成 k 。

Table 2. Comparison of importance algorithms of different nodes
表 2. 不同节点重要性算法对比

Algorithm	Information	Time
LNIF	Local	$O(2kN)$
Degree	Local	$O(N)$
BC	Global	$O(N^3)$
CC	Global	$O(N^3)$
WL	Local	$O(M + kN)$
K-shell	Node Location	$O(M)$

在 3.1 节的平均度实验中，LNIF 算法的曲线大多位于图的左下方，表明其节点序列能更快降低网络平均度，比其他算法更有效地移除高度数节点。在 3.2 节的网络效率实验中，LNIF 算法的曲线同样位于左下方，说明其节点排序能更快降低网络效率，当效率降至 20% 时，网络基本崩溃。实验中，曲线有时会上升，这是由于静态攻击模式下，原本排序靠前的节点(伪核心节点)在网络结构变化后可能不再处于重要位置(如边缘或孤立节点)，移除这些节点可能导致平均度或网络效率不降反升。

在第三节的最大连通子图系数实验中，LNIF 算法在几乎所有数据集中均处于不同算法曲线的下方，表明其移除节点后，网络的最大连通子图下降更快，破坏性更大。此外，部分数据集的曲线在初始阶段就呈现快速下降趋势，而另一些则趋于平缓，这与网络结构有关。例如，图 6(d)中的 Football 网络平均度高，节点连接边数多，移除节点时网络仍保持全连通性，直到移除比例超过一定值后，网络才进入非全连通状态。不同算法因移除顺序不同，最大连通子图的变化速率也不同，而 LNIF 算法的效果最为显著。在 3.4 节的 SIR 实验中，通过设置不同感染比例评估 LNIF 算法的节点排序效果。图 7 和图 8 显示，在相

关数据集上, LNIF 算法的感染曲线高于其他 5 种算法, 表明其节点排序结果的传播能力更强。

在上述实验中, K-shell 算法表现较差, 主要原因是其将节点分层的结果是粗粒度的, 无法区分同数值节点的重要性[48]。H-index 算法也存在类似问题, 部分 H-index 值相同的节点在网络中的重要性可能不同, 导致结果不够精确。相比之下, 本文提出的 LNIF 算法表现优异, 在几乎所有数据集中均表现最佳。综合平均度、网络效率、最大连通子图系数以及 SIR 实验结果, LNIF 算法在多数网络中优于现有的 Degree、H-index、WL、SC 和 K-shell 算法。

5. 结论

在复杂网络中, 核心节点处于网络中的优势地位, 挖掘核心节点具有重要的意义。为了挖掘网络中的重要节点, 本文提出一种考虑了节点的拓扑信息以及信息传播的影响的 LNIF 算法, 与基于全局信息的算法相比, 该算法只需要计算节点的二阶内邻居信息即可计算节点的重要性值, 对于挖掘、寻找大规模网络中的关键节点具有现实意义。

在 12 个数据集上将本文提出的 LNIF 算法与其它 5 种算法(Degree, H-index, WL, SC, K-shell)进行对比, 通过平均度、网络效率、最大连通子图系数和 SIR 来验证 LNIF 算法的优劣。实验结果表明, LNIF 算法要优于同为基于局部信息的 Degree 算法、H-index 算法、WL 算法、LC 算法以及 K-shell 算法。

挖掘重要节点, 是为了后续的研究做准备, 如挖掘团队中的核心人员来管理和维护团队的稳定以及协作任务的精准推荐等。虽然本文提出的 LNIF 算法能够挖掘重要(核心)节点, 但是本文只分析了单层网络的结构, 而在多层网络中的节点相互关联, 一旦节点失效后, 会产生蝴蝶效应从而导致多层网络崩溃, 因此下一步研究方向是如何在多层网络中挖掘重要节点。

参考文献

- [1] Dai, J., Huang, K., Liu, Y., Yang, C. and Wang, Z. (2021) Global Reconstruction of Complex Network Topology via Structured Compressive Sensing. *IEEE Systems Journal*, **15**, 1959-1969. <https://doi.org/10.1109/jsyst.2020.2997713>
- [2] Gebert, L.F.R. and MacRae, I.J. (2018) Regulation of MicroRNA Function in Animals. *Nature Reviews Molecular Cell Biology*, **20**, 21-37. <https://doi.org/10.1038/s41580-018-0045-7>
- [3] Wagg, C., Schlaeppli, K., Banerjee, S., Kuramae, E.E. and van der Heijden, M.G.A. (2019) Fungal-Bacterial Diversity and Microbiome Complexity Predict Ecosystem Functioning. *Nature Communications*, **10**, Article No. 4841. <https://doi.org/10.1038/s41467-019-12798-y>
- [4] Ureña, R., Kou, G., Dong, Y., Chiclana, F. and Herrera-Viedma, E. (2019) A Review on Trust Propagation and Opinion Dynamics in Social Networks and Group Decision Making Frameworks. *Information Sciences*, **478**, 461-475. <https://doi.org/10.1016/j.ins.2018.11.037>
- [5] Liu, J., Meng, H., Wang, W., Xie, Z. and Yu, Q. (2019) Evolution of Cooperation on Independent Networks: The Influence of Asymmetric Information Sharing Updating Mechanism. *Applied Mathematics and Computation*, **340**, 234-241. <https://doi.org/10.1016/j.amc.2018.07.004>
- [6] Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., et al. (2020) T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems*, **21**, 3848-3858. <https://doi.org/10.1109/tits.2019.2935152>
- [7] Wen, X., Tu, C. and Wu, M. (2018) Node Importance Evaluation in Aviation Network Based on “No Return” Node Deletion Method. *Physica A: Statistical Mechanics and its Applications*, **503**, 546-559. <https://doi.org/10.1016/j.physa.2018.02.109>
- [8] Gosak, M., Markovič, R., Dolencsek, J., Slak Rupnik, M., Marhl, M., Stožer, A., et al. (2018) Network Science of Biological Systems at Different Scales: A Review. *Physics of Life Reviews*, **24**, 118-135. <https://doi.org/10.1016/j.plrev.2017.11.003>
- [9] Koutrouli, M., Karatzas, E., Paez-Espino, D. and Pavlopoulos, G.A. (2020) A Guide to Conquer the Biological Network Era Using Graph Theory. *Frontiers in Bioengineering and Biotechnology*, **8**, Article 34. <https://doi.org/10.3389/fbioe.2020.00034>
- [10] Zhang, M., Wang, X., Jin, L., Song, M. and Li, Z. (2022) A New Approach for Evaluating Node Importance in Complex

- Networks via Deep Learning Methods. *Neurocomputing*, **497**, 13-27. <https://doi.org/10.1016/j.neucom.2022.05.010>
- [11] van Elteren, C., Quax, R. and Sloot, P. (2022) Dynamic Importance of Network Nodes Is Poorly Predicted by Static Structural Features. *Physica A: Statistical Mechanics and Its Applications*, **593**, Article 126889. <https://doi.org/10.1016/j.physa.2022.126889>
 - [12] Ruan, Y., Lao, S., Tang, J., Bai, L. and Guo, Y. (2022) Node Importance Ranking Method in Complex Network Based on Gravity Method. *Acta Physica Sinica*, **71**, Article 176401. <https://doi.org/10.7498/aps.71.20220565>
 - [13] Matamalas, J.T., Arenas, A. and Gómez, S. (2018) Effective Approach to Epidemic Containment Using Link Equations in Complex Networks. *Science Advances*, **4**, eaau4212. <https://doi.org/10.1126/sciadv.aau4212>
 - [14] Ibnoulouafi, A., El Haziti, M. and Cherifi, H. (2018) M-Centrality: Identifying Key Nodes Based on Global Position and Local Degree Variation. *Journal of Statistical Mechanics: Theory and Experiment*, **2018**, Article 073407. <https://doi.org/10.1088/1742-5468/aace08>
 - [15] Li, Y., Wang, B., Wang, H., Ma, F., Zhang, J., Ma, H., et al. (2022) Importance Assessment of Communication Equipment in Cyber-Physical Coupled Distribution Networks Based on Dynamic Node Failure Mechanism. *Frontiers in Energy Research*, **10**, Article 911985. <https://doi.org/10.3389/fenrg.2022.911985>
 - [16] Chierichetti, F., Lattanzi, S. and Panconesi, A. (2011) Rumor Spreading in Social Networks. *Theoretical Computer Science*, **412**, 2602-2610. <https://doi.org/10.1016/j.tcs.2010.11.001>
 - [17] Liu, F., Wang, Z. and Deng, Y. (2020) GMM: A Generalized Mechanics Model for Identifying the Importance of Nodes in Complex Networks. *Knowledge-Based Systems*, **193**, Article 105464. <https://doi.org/10.1016/j.knosys.2019.105464>
 - [18] Mo, H. and Deng, Y. (2019) Identifying Node Importance Based on Evidence Theory in Complex Networks. *Physica A: Statistical Mechanics and its Applications*, **529**, Article 121538. <https://doi.org/10.1016/j.physa.2019.121538>
 - [19] Lu, M. (2020) Node Importance Evaluation Based on Neighborhood Structure Hole and Improved Topsis. *Computer Networks*, **178**, Article 107336. <https://doi.org/10.1016/j.comnet.2020.107336>
 - [20] Liu, D., Huang, K., Wu, D. and Zhang, S. (2021) A New Method of Identifying Core Designers and Teams Based on the Importance and Similarity of Networks. *Computational Intelligence and Neuroscience*, **2021**, Article 3711733. <https://doi.org/10.1155/2021/3711733>
 - [21] Lü, L., Zhou, T., Zhang, Q. and Stanley, H.E. (2016) The H-Index of a Network Node and Its Relation to Degree and Coreness. *Nature Communications*, **7**, Article No. 10168. <https://doi.org/10.1038/ncomms10168>
 - [22] 王安, 顾益军. 基于社区划分的节点重要性评估方法[J]. 计算机工程与应用, 2020, 56(8): 42-48.
 - [23] Liu, Y., Wang, J., He, H., Huang, G. and Shi, W. (2021) Identifying Important Nodes Affecting Network Security in Complex Networks. *International Journal of Distributed Sensor Networks*, **17**, Article 155014772199928. <https://doi.org/10.1177/1550147721999285>
 - [24] Yang, X. and Xiao, F. (2021) An Improved Gravity Model to Identify Influential Nodes in Complex Networks Based on K-Shell Method. *Knowledge-Based Systems*, **227**, Article 107198. <https://doi.org/10.1016/j.knosys.2021.107198>
 - [25] Buldú, J.M., Busquets, J., Martínez, J.H., Herrera-Diestra, J.L., Echegoyen, I., Galeano, J., et al. (2018) Using Network Science to Analyse Football Passing Networks: Dynamics, Space, Time, and the Multilayer Nature of the Game. *Frontiers in Psychology*, **9**, Article 1900. <https://doi.org/10.3389/fpsyg.2018.01900>
 - [26] Newman, M.E.J. (2006) Finding Community Structure in Networks Using the Eigenvectors of Matrices. *Physical Review E*, **74**, Article 036104. <https://doi.org/10.1103/physreve.74.036104>
 - [27] Newman, M.E.J. (2001) From the Cover: The Structure of Scientific Collaboration Networks. *Proceedings of the National Academy of Sciences*, **98**, 404-409. <https://doi.org/10.1073/pnas.021544898>
 - [28] 杨雄. 复杂网络影响力节点度量及影响力最大化算法研究[D]: [博士学位论文]. 杭州: 浙江工业大学, 2017.
 - [29] Wei, D., Zhang, X. and Mahadevan, S. (2018) Measuring the Vulnerability of Community Structure in Complex Networks. *Reliability Engineering & System Safety*, **174**, 41-52. <https://doi.org/10.1016/j.res.2018.02.001>
 - [30] Yan, Y., Yu, G., Yan, X. and Xie, H. (2018) Community Cores Expansion for Overlapping Community Detection in Complex Networks. *Modern Physics Letters B*, **32**, Article 1850405. <https://doi.org/10.1142/s0217984918504055>
 - [31] Wei, H., Pan, Z., Hu, G., Zhang, L., Yang, H., Li, X., et al. (2018) Identifying Influential Nodes Based on Network Representation Learning in Complex Networks. *PLOS ONE*, **13**, e0200091. <https://doi.org/10.1371/journal.pone.0200091>
 - [32] Nath, K., Roy, S. and Nandi, S. (2020) Inovin: A Fuzzy-Rough Approach for Detecting Overlapping Communities with Intrinsic Structures in Evolving Networks. *Applied Soft Computing*, **89**, Article 106096. <https://doi.org/10.1016/j.asoc.2020.106096>
 - [33] Adamic, L.A. and Glance, N. (2005) The Political Blogosphere and the 2004 U.S. Election. *Proceedings of the 3rd International Workshop on Link Discovery*, Chicago, 21-25 August 2005, 36-43. <https://doi.org/10.1145/1134271.1134277>

-
- [34] Watts, D.J. and Strogatz, S.H. (1998) Collective Dynamics of ‘Small-World’ Networks. *Nature*, **393**, 440-442. <https://doi.org/10.1038/30918>
 - [35] Si, C., Jiao, L., Wu, J. and Zhao, J. (2017) A Group Evolving-Based Framework with Perturbations for Link Prediction. *Physica A: Statistical Mechanics and Its Applications*, **475**, 117-128. <https://doi.org/10.1016/j.physa.2017.01.087>
 - [36] 刘超. 基于智能计算的复杂网络社区发现算法研究[D]: [硕士学位论文]. 济南: 山东大学, 2020.
 - [37] 赵丽倩. 复杂网络中节点重要性度量与影响力阻塞最大化研究[D]: [硕士学位论文]. 兰州: 兰州大学, 2021.
 - [38] Newman, M.E.J. and Watts, D.J. (1999) Renormalization Group Analysis of the Small-World Network Model. *Physics Letters A*, **263**, 341-346. [https://doi.org/10.1016/s0375-9601\(99\)00757-4](https://doi.org/10.1016/s0375-9601(99)00757-4)
 - [39] Chen, D., Lü, L., Shang, M., Zhang, Y. and Zhou, T. (2012) Identifying Influential Nodes in Complex Networks. *Physica A: Statistical Mechanics and Its Applications*, **391**, 1777-1787. <https://doi.org/10.1016/j.physa.2011.09.017>
 - [40] 王建伟, 荣莉莉, 郭天柱. 一种基于局部特征的网络节点重要性度量方法[J]. 大连理工大学学报, 2010, 50(5): 822-826.
 - [41] Kitsak, M., Gallos, L.K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H.E., *et al.* (2010) Identification of Influential Spreaders in Complex Networks. *Nature Physics*, **6**, 888-893. <https://doi.org/10.1038/nphys1746>
 - [42] Ruan, Y., Tang, J., Wang, H., Guo, J. and Qin, W. (2021) Method for Measuring Node Importance in Complex Networks Based on Local Characteristics. *International Journal of Modern Physics B*, **35**, Article 2150231. <https://doi.org/10.1142/s0217979221502313>
 - [43] Fan, W., He, Y., Han, X. and Feng, Y. (2021) A New Model to Identify Node Importance in Complex Networks Based on DEMATEL Method. *Scientific Reports*, **11**, Article No. 22829. <https://doi.org/10.1038/s41598-021-02306-y>
 - [44] Meng, Y., Qi, Q., Liu, J. and Zhou, W. (2022) Dynamic Evolution Analysis of Complex Topology and Node Importance in Shenzhen Metro Network from 2004 to 2021. *Sustainability*, **14**, Article 7234. <https://doi.org/10.3390/su14127234>
 - [45] Lalou, M., Tahraoui, M.A. and Kheddouci, H. (2018) The Critical Node Detection Problem in Networks: A Survey. *Computer Science Review*, **28**, 92-117. <https://doi.org/10.1016/j.cosrev.2018.02.002>
 - [46] Liu, J., Li, X. and Dong, J. (2020) A Survey on Network Node Ranking Algorithms: Representative Methods, Extensions, and Applications. *Science China Technological Sciences*, **64**, 451-461. <https://doi.org/10.1007/s11431-020-1683-2>
 - [47] Zareie, A., Sheikhamadi, A. and Khamforoosh, K. (2018) Influence Maximization in Social Networks Based on Topsis. *Expert Systems with Applications*, **108**, 96-107. <https://doi.org/10.1016/j.eswa.2018.05.001>
 - [48] Ruan, Y.-R., Lao, S.-Y., Wang, J.-D., *et al.* (2017) Node Importance Measurement Based on Neighborhood Similarity in Complex Network. *Acta Physica Sinica*, **66**, Article 038902. <https://doi.org/10.7498/aps.66.038902>