

DOI:10.11918/202306047

基于集体影响力的重叠社区检测算法

刘彬颖, 刘三阳, 白艺光

(西安电子科技大学 数学与统计学院, 西安 710126)

摘要: 随着大数据时代的到来,网络结构日益复杂,探索复杂网络的社区结构对理解其功能和组织机制具有重要意义。关于社区检测已有大量研究,其中标签传播算法(LPA)具有接近线性的时间复杂度,适用于大规模复杂网络,但其随机性较强,准确性不高。提出基于集体影响力的标签传播算法(CILPA),用于发现重叠社区结构。CILPA引入集体影响力这一全局指标,融合节点自身信息和网络全局信息重新定义节点重要性,并依此固定节点更新顺序以提升算法稳定性;在标签传播过程中,设计一种标签选择策略,并设置自适应过滤因子,防止错误标签的干扰,从而提高算法的准确性和鲁棒性。最后在不同规模、复杂度及重叠率的人工网络和真实网络上进行实验,结果表明,CILPA的模块度和标准化互信息均优于COPRA、SLPA等主流算法,且标准差更小。说明本方法在重叠社区检测中兼具有效性和稳定性,为大规模复杂网络的重叠社区分析提供了可靠方法。

关键词: 社区检测;重叠社区;集体影响力;标签传播;节点重要性

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 0367-6234(2026)03-0197-08

Overlapping community detection algorithm based on collective influence

LIU Binying, LIU Sanyang, BAI Yiguang

(School of Mathematics and Statistics, Xidian University, Xi'an 710126, China)

Abstract: With the advent of the era of big data, network structures are becoming more and more complex, and exploring the community structure of complex networks holds great significance in understanding their function and organization mechanisms. Many studies have been conducted for community detection, among which the label propagation algorithm (LPA) has a near linear time complexity and is applicable for large-scale complex networks. However, it has excessive randomness and relatively low accuracy. This paper proposed a collective influence-based label propagation algorithm (CILPA) for discovering overlapping community structures. CILPA introduced collective influence as a global indicator, redefined the node importance by integrating the node's own information and global network information, and fixed node update order according to node importance to improve the algorithm's stability. In the label propagation process, a label selection strategy was designed, and the adaptive filtering factor were set to prevent the interference of wrong labels, thereby improving the accuracy and robustness of the algorithm. Finally, experiments were conducted on artificial and real networks with different scales, complexities, and overlap rates. The results show that the modularity and normalized mutual information of CILPA are superior to those of mainstream algorithms such as COPRA and SLPA, with a smaller standard deviation. This indicates that the proposed method possesses both effectiveness and stability in overlapping community detection, providing a reliable method for the analysis of overlapping communities in large-scale complex networks.

Keywords: community detection; overlapping community; collective influence; label propagation; node importance

在现实世界的诸多领域,都可以通过复杂网络进行表征^[1],如社交网络^[2]、生物网络^[3]、科学协作网络^[4]等。要分析这些网络,就需要了解其拓扑特征,如小世界特性^[5]、无标度特性^[6]等。社区结构作为复杂网络的重要属性^[7],近年来已成为复杂网

络领域的研究热点,有助于深入挖掘网络蕴含的潜在信息。早期社区检测研究多聚焦于非重叠社区划分^[8-9],但在现实世界中,大部分网络都是重叠网络,即一个节点可能同时属于多个社区,例如,推荐系统(如购物软件的智能推荐功能)^[10]、科学协作网

收稿日期:2023-06-11;录用日期:2023-07-11;网络首发日期:2024-06-25

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20240624.1725.024>

作者简介:刘彬颖(1999—),女,硕士研究生;刘三阳(1959—),男,教授,博士生导师;白艺光(1994—),男,讲师

通信作者:刘彬颖, binyingliu6@163.com

络(如学者跨学科研究)、信息网络(如学术论文引文网络)^[11]、电网(如家庭与企业配电网)^[12]等。因此,重叠社区检测更具现实意义。

重叠社区检测的算法大致可以分为三类:第一类是基于派系过滤的方法^[13],这类方法虽然能识别重叠社区,但只适用于子图完全连通的网络;第二类是基于线图和边分割的方法^[14],核心思想是利用边而非节点进行分区,但其检测质量不如节点检测方法;第三类是基于模糊聚类的方法^[15],需要提前定义社区个数和模糊聚类的相关参数,这些参数都很难确定最优值。

上述算法在时间和空间上消耗较大,不适用于大规模社交网络结构检测。在大数据时代,随着网络规模持续扩大,网络结构愈发复杂,亟需低时间复杂度的高效算法。标签传播算法(LPA)^[16]是由 Raghavan 等学者提出的一种半监督算法,具有接近线性的时间复杂度,适用于大规模网络的高效社区检测。该算法受流行病传播机制的启发,通过节点标签的迭代传播实现社区检测,直到收敛(具有相同标签的节点构成社区)。COPRA 是首个用于重叠社区检测的 LPA 改进算法^[17],其规定每个节点可以包含多个社区标签,并为每个标签定义隶属度,在传播过程中根据邻居节点信息更新标签及隶属度,但该算法认为所有节点的重要性是相同的,一定程度上降低了检测精度,易产生小社区。SLPA 引入了“speaker-listener”规则^[18],给每个节点分配内存用来储存历史标签,最后根据内存里的历史标签确定最终的社区标签,并通过参数 r 删除低隶属度标签,但 r 值的确定具有较强主观性。MGA 算法通过分析标签传播的 Newman 模块化增益函数^[19],提出加速模块化增益检测方法,但受模块化分辨率限制,算法存在局限性。Zhao 等^[20]提出基于图压缩和标签传播的大规模社区检测方法,通过压缩低度数节点并搜索社区中心,根据相似性等信息检测社区结构,但社区中心的搜索结果对最终检测效果影响显著。Qiong 等^[21]提出基于边界节点的标签传播算法 LBN,通过核心节点给边界节点赋予权重并进行社区划分,但边界节点的计算增加了时间复杂度。LPANNI 定义节点相似度和邻居节点影响力^[22-23],提出基于邻居节点影响力和历史标签偏好的更新策略,但在小规模网络中效果欠佳。NI-LPA 基于节点自身属性定义节点重要性^[24],依据重要性进行标签传播,虽提高了准确性,但只考虑了节点的局部信息,没有充分考虑整个网络的拓扑信息。

针对上述问题,本文提出改进的标签传播算法

CILPA。引入集体影响力指标衡量节点的重要性,综合考虑节点自身信息和全局信息定义节点重要性,提升重叠社区检测准确性;设计新型节点标签更新策略,在保证准确性的同时,维持了低时间复杂度;定义自适应过滤因子,删除不重要标签。通过在真实网络和 LFR 基准网络上进行实验,验证了此算法兼具检测重叠社区的有效性和稳定性。

1 基于集体影响力的标签传播算法

1.1 基本定义

给定包含 n 个节点的无向图 $G=(V,E)$,其中 V 为节点集, $V=\{v_1,v_2,\dots,v_n\}$, E 为边集, $E=\{(v_i,v_j)|1\leq i\neq j\leq n\}$ 。 A 是图 V 的邻接矩阵。重叠社区检测旨在将 V 中所有的节点分到 m 个社区中,得到社区划分 $P=\{C_1,C_2,\dots,C_m\}$,每个社区之间可以互相连接,同一社区内节点连接紧密,不同社区间节点连接稀疏,且每个节点可以同时属于多个社区,即 $\sum_{1\leq i\neq j\leq c} C_i\cap C_j\neq 0$ 。

1.2 核心指标定义

1) 集体影响力

集体影响力 (collective influence, CI) 是量化节点影响力的全局指标^[25]。节点 i 的集体影响力记为 $CI_l(i)$,定义为节点 i 与其 l 阶邻居的节点度之和的乘积:

$$CI_l(i) = (k_i - 1) \sum_{d(i,j)=l} (k_j - 1) \quad (1)$$

式中: k_i 是节点 i 的度; $d(i,j)$ 是节点 i 和 j 的最短距离。CI 通过 l 阶邻居描述节点的重要性。等式右边的和包含了围绕中心节点 i 的球表面上的节点的贡献度,每个节点的权重都是 $k-1$ 。这表明,即使节点自身度不高,若其周围存在大量连接,仍可能具有较高影响力,如图 1 所示。

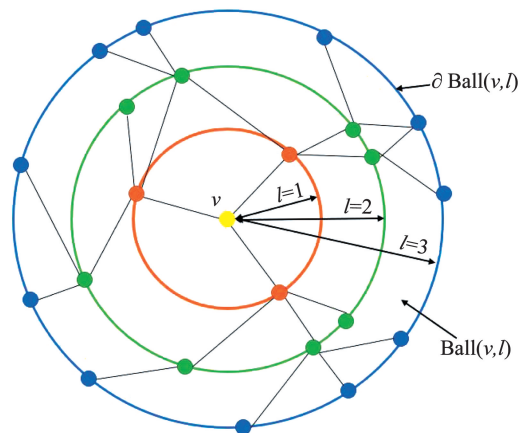


图 1 集体影响力示意图

Fig. 1 Collective influence

2) 节点重要性

节点重要性用于衡量标签传播过程中衡量节点所携带信息的传播优先级。一个节点的重要性越高,该节点发送的标签在传播阶段的优先级就越高。节点重要性也可以反映节点成为社区中心的概率。仅通过节点度来度量节点重要性存在局限性,因为具有相同度的节点在复杂网络中可能不会发挥相同的作用。现有研究^[23-24]考虑节点邻域信息来量化节点重要性,但仍局限于局部视野。CI作为全局指标,能更全面反映拓扑特征,量化节点重要性。由于每个节点的CI会随参数*l*的增加而增加,需进行归一化处理。本文将节点度和归一化的集体影响力结合,重新定义节点重要性:

$$NI(v) = k_v \times \frac{CI_l(v)}{\langle CI \rangle} \quad (2)$$

式中 $\langle CI \rangle$ 是所有节点集体影响力的均值。

1.3 CILPA 算法

1) 初始化

针对重叠社区检测需求,每个节点对应一个标签集。标签集里的每个元素为一个标签对,包含标签和对应的隶属度,定义如下:

$$L_v = \{(c_1, b_1), (c_2, b_2), \dots, (c_L, b_L)\} \quad (3)$$

式中:*c*是节点*v*的一个社区标签;*b*是标签的隶属度,既表示节点属于该社区的概率,也代表该标签的传播优先级。

在初始化时,将每个节点视为一个独立社区,为每个节点的标签赋予一个隶属度,初始化时所有节点的隶属度设置为节点重要性的值,即 $L_v = \{(v, NI_v)\}$ 。

同时,为了缓解传播过程中随机选择更新节点得出结果的不稳定性,将所有节点按重要性进行降序排序,并更新标签,得到序列 N_{sort} 。

2) 传播策略

为了降低算法的复杂度,同时减少不重要标签对结果的干扰,借鉴DLPA^[30]的做法,在传播过程中,仅传播每个节点隶属度最大的标签,定义主导标签为

$$\text{dom}_v = \arg \max_c b(c, v) \quad (4)$$

采取异步更新模式避免震荡问题。在第*t*次迭代中,节点标签通过邻居节点第*t*-1次迭代的主导标签(未更新邻居)和第*t*迭代的主导标签(已更新

邻居)共同更新,公式如下:

$$b_t(c, v) = \sum_{u \in N_1(v)} b_{t-1}(c, u) + \sum_{u \in N_2(v)} b_t(c, u) \quad (5)$$

式中, $N_1(v)$ 是*v*的邻居中未更新的节点集, $N_2(v)$ 是*v*的邻居中已更新的节点集。

在接受邻居所有信息后,得到一个新的标签集,此时标签集里可能会有干扰信息,需要对这个标签集进行过滤。由于固定参数难以确定,选择自适应的过滤因子 $\rho = \frac{1}{|L_v| + 1}$,删除隶属度小于 ρ 的标签,并对标签集进行归一化处理,得到更新后的标签集,并从中提取隶属度最大的标签作为主导标签,用于下次更新。

按 N_{sort} 中的节点顺序进行标签更新与传播,具体更新策略如下:

- a. 对于 N_{sort} 中的每个节点*v*,从其所有邻居接收主导标签信息,并按公式(5)更新*v*的标签集 L_v 。
- b. 将更新后的标签集 L_v 进行过滤,删除隶属度小于 ρ 的标签,得到新的标签集 L_u 。
- c. 对 L_u 进行归一化处理,找到隶属度最大的标签,更新*v*的主导标签 dom_v 。
- d. 如果 dom_v 不变,则终止迭代;否则,重复步骤a~c。

标签传播完成后,网络中每个节点均保留标签集信息,最后再将含有相同标签的节点划分到同一个社区,得到社区划分*P*。

标签传播过程如图2所示,黑色节点为待更新节点,接收来自邻居节点*a*、*b*、*c*的主导标签信息(图中加粗部分)得到一个标签集,然后将其过滤后进行归一化处理,计算 $\rho = 0.25$,删除不满足条件的标签*b*,得到最终标签集,其中*c*为主导标签,用于下一次传播。

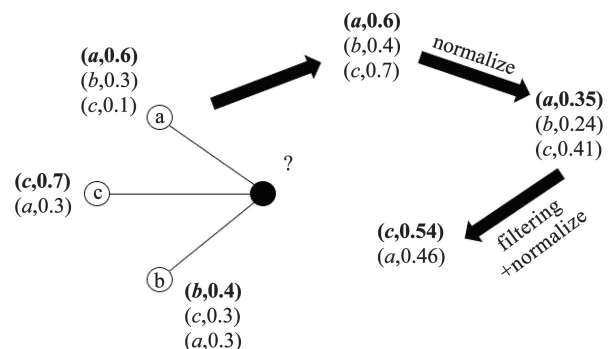


图2 标签传播过程

Fig. 2 Label propagation process

3) 算法描述

输入: 无向图 G , 最大迭代次数 t

输出: 重叠社区划分结果 P

1. $P = \emptyset$
2. for each node v in V do
3. 计算节点重要性值 $NI(v)$
4. 初始化节点标签: $L_{node} = \{ (node, NI) \}$
5. end for
6. 将所有节点按节点重要性值 $NI(v)$ 进行降序排序, 得到节点序列 N_{sort}
7. for u in N_{sort} do
8. for v in u 的邻居节点集 do
9. if u 的标签集包含 v 的标签:
10. $b_u = b + b_u$
11. else
12. 将 v 的标签加入到 u 的标签集中
13. end if
14. end for
15. for b in L_v do // 过滤
16. if $b < \rho$ then
17. delete b from L_v
18. end if
19. end for
20. normalize L_u // 对标签集进行归一化
21. 找到 L_u 中 b 最大的标签作为节点的主导标签, 用于传播
22. if u 的标签不再变化
23. 结束此次循环
24. else
25. $t = t + 1$
26. end if
27. end for
28. 将具有相同标签的节点划分到同一个社区, 得到划分结果 P

2 实验与分析

2.1 评价指标

1) 标准化互信息 (NMI)

Danon 等^[26] 提出标准化互信息 (normalized mutual information, NMI) 用于评价算法得到的社区与真实网络社区分布的相似程度。然而 NMI 在部分场景下可能会高估两个社区划分的相似性, 例如: 当一个多社区划分与一个单社区划分存在完全重合的社区时, NMI 值可能偏高。McDaid 等^[27] 提出了一种改进的标准化互信息 (NMI_{max}), 可有效解决该问题, 定义如下:

$$NMI_{max}(X, Y) = 1/2 \times$$

$$\frac{H(X) + H(Y) - H(X|Y) - H(Y|X)}{\max(H(X), H(Y))} \quad (6)$$

式中: X 为算法得到的社区划分; Y 为真实网络社区划分; $H(X)$ 和 $H(Y)$ 分别为随机变量 X 、 Y 的信息熵; $H(X|Y)$ 为条件熵。当算法得到的社区结果与真实划分完全一致时, 则 $NMI_{max} = 1$; 当两个社区完全独立时, $NMI_{max} = 0$ 。

2) 模块度

在真实网络中, 正确的社区划分通常是未知的, 为了评价社区划分的优劣, Newman 等提出了模块度的概念。优质社区划分应满足社区内部节点连接稠密、社区间连接稀疏, 此时模块度值更大。这是针对非重叠社区的模块度定义。Nicosia 等^[28] 将该模块度扩展到了重叠社区检测, 定义如下:

$$Q = \frac{1}{2m} \sum_{q=1}^l \sum_{i,j \in C_q} \frac{1}{O_i O_j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \quad (7)$$

式中: m 是网络中边的个数; O_i 是节点 i 属于的社区个数; A 是网络的邻接矩阵。

2.2 对比算法和实验数据

1) 对比算法

本文选择了 5 种主流重叠社区检测算法作为对比算法, 分别为 COPRA、SLPA、CPM、LPANNI、ENDNTM, 各比较算法的参数均参考原始文献设置。

2) 真实数据

选择 5 个不同规模的真实数据集, 详细参数见表 1。

表 1 真实网络参数表

Tab. 1 Parameters of real network

网络名	节点数	边数	平均度
Dolphins	62	159	5.1
Football	115	613	10.6
Pol	105	441	8.4
Email	1 133	5 451	9.6
Power	4 941	6 594	2.7

3) 人工合成数据

由于真实数据有限且多数真实网络的社区划分未知, 本文采用 LFR 模型生成基准网络^[29]。该模型可以通过调节不同参数生成不同特性的人工网络。主要参数包括: N (节点数, 反映网络规模); k (平均度, 通常设为 10, 以接近真实社会网络); μ (社团内部连接强度系数, 值越大代表连接越弱); C_{min}

和 C_{max} (社团规模上下限); on (网络重叠节点占比); om (重叠节点最多属于几个社区), on 和 om 共同决定复杂网络中社区的重叠度。

为了验证算法在不同类型网络中的有效性,用 LFR 生成 4 种不同性质的人工网络,参数见表 2。通过改变 N 、 μ 、 om 和 on , 验证算法在不同规模、复杂度、社区重叠度网络中的有效性。

表 2 LFR 网络参数表

Tab. 2 Parameters of LFR network

参数	LFR1	LFR2	LFR3	LFR4
N	1000, 5 000	1 000, 5 000	1 000, 5 000	5 000, 10 000
k	10	10	10	10
k_{max}	50	50	50	50
μ	0.1	0.1 ~ 0.3	0.1	0.1
C_{min}	10	10	10	10
C_{max}	50	50	50	50
on	10%	10%	5% ~ 35%	10%
om	2, 4, 8	2, 8	3	2 ~ 8

2.3 结果分析

1) 参数 l 的确定

在 CILPA 中引入集体影响力,由公式(1)可知,CI 表达式中的参数 l 需合理确定。 l 越大,CI 包含的拓扑信息越丰富,有助于提高算法准确性,但同时也会增加时间复杂度。为在准确性和时间复杂度之间取得平衡,设置实验检验不同 l 值对社区检测的影响,并确定合适的 l 值。实验在 LFR1 网络上进行,结果如图 3 所示。

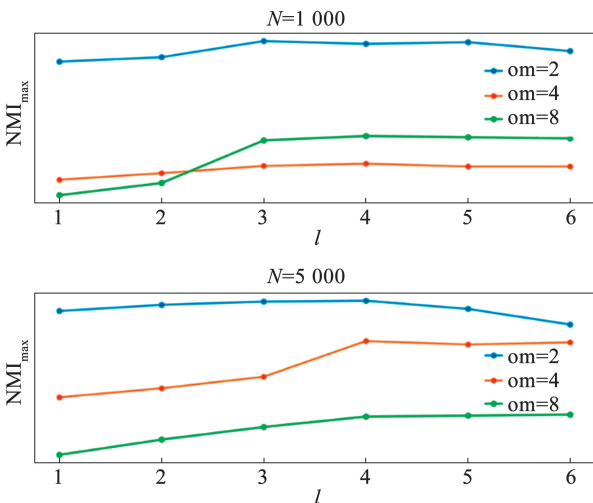


图 3 CILPA 在不同 l 下的结果

Fig. 3 Results of CILPA under different l values

由图 3 可以看出,当 $N=1 000$ 时,在不同 om 值下, NMI_{max} 在 $l=3$ 或 $l=4$ 的时候达到顶峰, $l > 4$ 后趋于平稳或者小幅度下降;当 $N=5 000$ 时,不同 om 值下, NMI_{max} 在 $l=4$ 处达到顶峰。综合不同规模和复杂度网络来看,更丰富的拓扑信息确实能够提高算法的准确性,但过大的 l 也可能降低精度并增加了时间复杂度。

根据实验结果,当 $l=4$ 时,算法效果最好,既能准确检测重叠社区,又能维持较低的时间复杂度。

2) 真实网络实验结果

在 5 个不同规模的真实网络中,从准确性(模块度均值 Q_{avg})和稳定性(模块度标准差 Q_{std})两方面对比五种算法。图 4 为不同网络中模块度均值的对比结果。可以看到,CILPA 的模块度值一直较高。

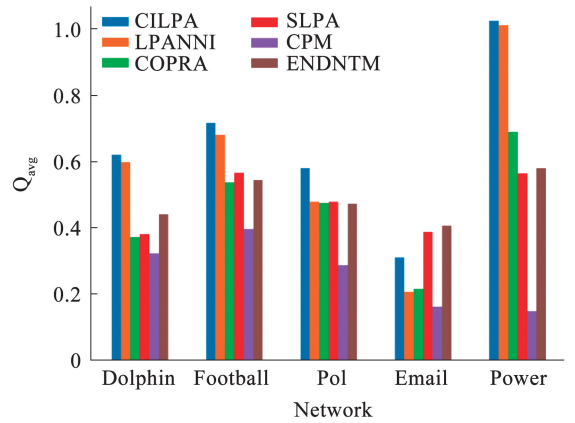


图 4 真实网络中模块度均值的比较结果

Fig. 4 Comparison result of average modularity in real networks

表 3 列出各算法的模块度均值和标准差。可以看出,CILPA 在多数网络中的精度和稳定性方面都表现较好;ENDNTM 在 Email 网络中表现更好;LPANNI 和 SLPA 整体表现处于上游水平,但 CILPA 的 Q_{avg} 还是更高。CPM 的稳定性也表现很好,但精度方面是 5 种算法中最差的。综合来看,CILPA 在真实网络中发现重叠社区的能力更好,并兼具准确性和稳定性。

表 3 真实网络中的结果

Tab. 3 Result on real network

	index	Dolphins	Football	Pol	Email	Power
CILPA	Q_{avg}	0.62	0.72	0.58	0.31	0.93
	Q_{std}	<0.001	<0.001	0.013	<0.001	<0.001
LPANNI	Q_{avg}	0.60	0.68	0.49	0.21	0.91
	Q_{std}	<0.001	<0.001	0.022	0.017	<0.001

表 3(续)

	index	Dolphins	Football	Pol	Email	Power
SLPA	Q_{avg}	0.36	0.56	0.48	0.39	0.56
	Q_{std}	0.048	0.025	0.025	0.057	0.007
COPRA	Q_{avg}	0.37	0.53	0.48	0.16	0.69
	Q_{std}	0.021	0.064	0.049	0.179	0.004
CPM	Q_{avg}	0.32	0.39	0.29	0.16	0.15
	Q_{std}	<0.001	<0.001	0.027	<0.001	<0.001
ENDNTM	Q_{avg}	0.44	0.54	0.47	0.40	0.58
	Q_{std}	<0.001	0.011	0.016	<0.001	<0.001

3) LFR 实验结果分析

首先通过改变 μ 值调整网络复杂度。 μ 是混合参数,表示社区内节点与其他社区的连接强度, μ 值越大,代表网络越复杂,社区的边界越模糊。图 5 结果显示,随着 μ 值的增加,每个算法的 NMI_{max} 值均有下降。其中,CILPA 在 $\mu \leq 0.1$ 时, NMI_{max} 接近于 1,且始终优于其他算法;COPRA 在 $\mu > 0.1$ 后性能迅速下降。这表明在较为复杂的网络中,CILPA 能够更有效地检测重叠社区。

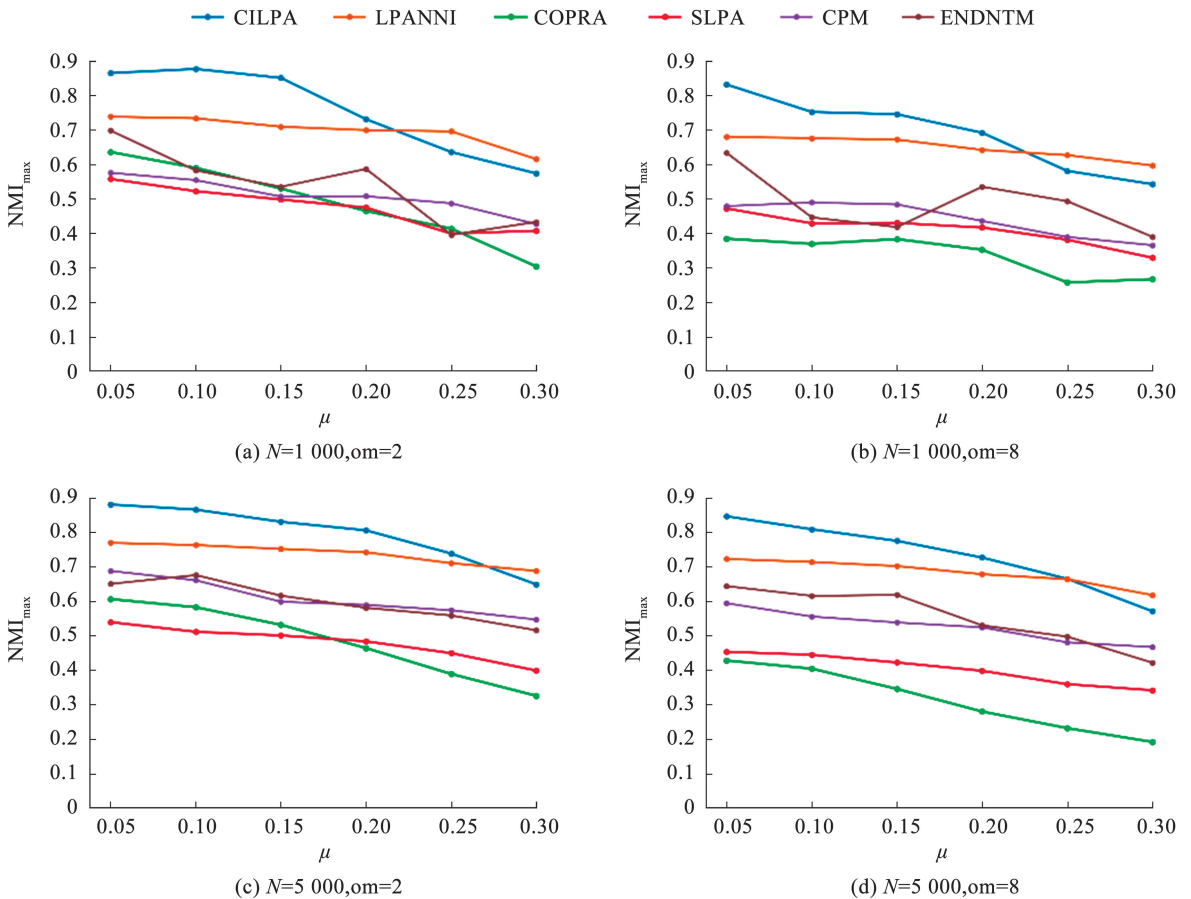


图 5 不同复杂度网络下各算法比较结果

Fig. 5 Comparison results of algorithms on networks with different complexities

其次,通过改变 on (重叠节点占比)和 om (重叠节点隶属社区数)来调整网络的重叠度。随着 om 和 on 的增加,网络中的重叠社区结构变得越来越复杂。图 6 显示,所有算法的 NMI_{max} 都随着 om 和 on 的增加而有所下降。在 om 从 2 增加的 8 的过程中,CILPA 下降幅度相对较小。在 $N = 10\ 000$ 时,COPRA 和 SLPA 在 $om = 8$ 时 NMI_{max} 几乎降为了 0,而 LPANNI、CPM 和 CILPA 虽然也有

所下降,但仍在可接受范围内,其中 CILPA 表现最优。而随着 on 的增加,所有算法识别重叠社区的能力都有所下降。综合来看,LPANNI、CPM 和 CILPA 的效果是比较好的,其中,CILPA 的效果更好。

同时,综合图 5、图 6 的结果来看,在规模更大、结构更复杂的网络中,对比其他算法,CILPA 仍能有效准确识别重叠社区。

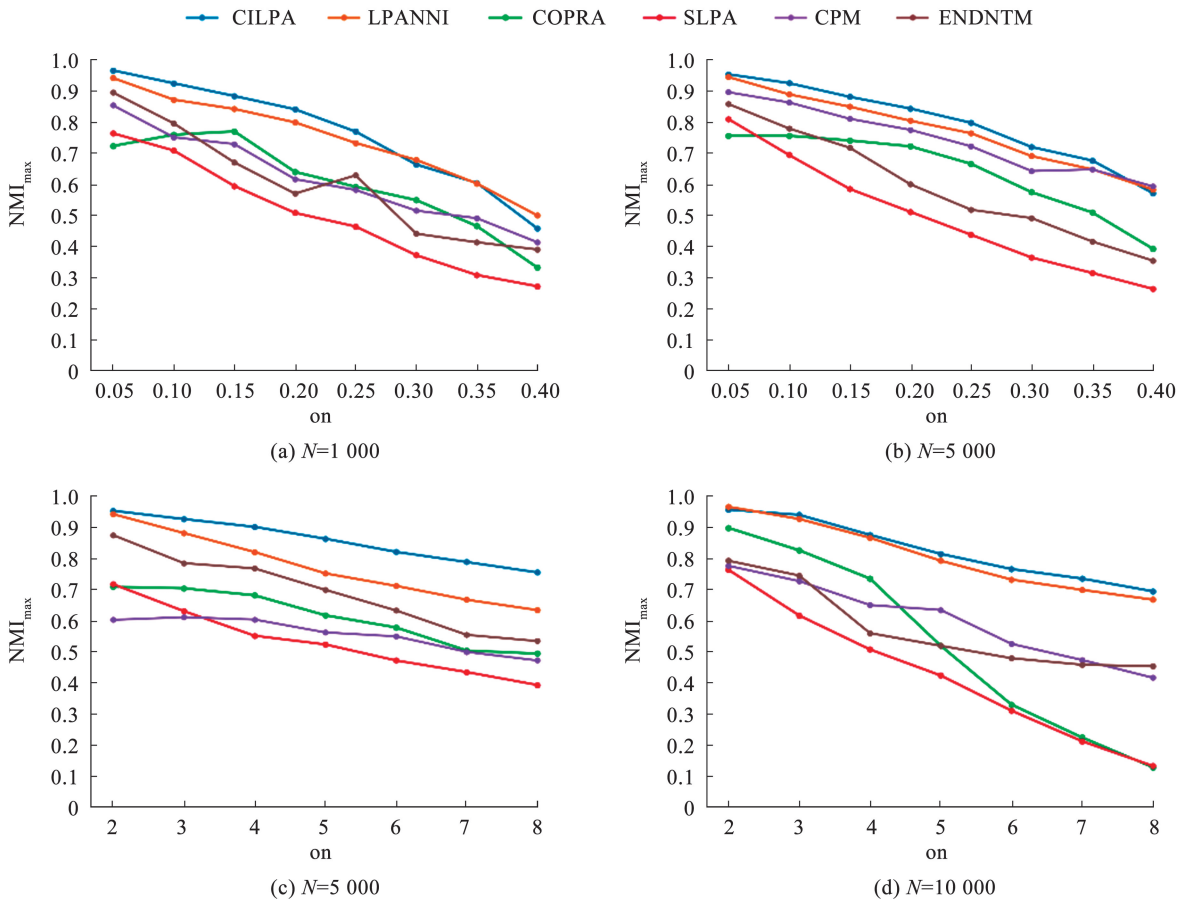


图 6 不同重叠度下各算法的比较结果

Fig. 6 Comparison results of algorithms under different overlapping degrees

3 结 语

本文提出了一种基于集体影响力的重叠社区检测算法 CILPA,该算法保持了 LPA 的低时间复杂度,适用于大规模复杂网络的重叠社区检测。CILPA 通过集体影响力衡量节点重要性,综合考虑节点自身信息和全局信息,并以此来改进标签传播过程,显著提高了重叠社区检测的准确性。此外,算法引入自适应过滤因子以删除不重要标签,大大提高了检测精度。通过调节 CI 中的参数 l ,可在准确性与时间复杂度之间取得良好平衡,实验表明 $l = 4$ 适用于多数网络。

在人工和真实网络上的实验表明,CILPA 能够识别出最接近真实划分的社区结构,在精度和稳定性方面均具优势。然而,本文提出的 CILPA 目前只适用于静态无权网络。未来研究将把该算法进一步推广至加权网络以及动态网络的重叠社区检测场景。

参考文献

[1]FURHT B. Handbook of social network technologies and applications

[M]. New York, NY: Springer, 2010. DOI:10.1007/978-1-4419-7142-5

[2]CHEN Y L, CHUANG C H, CHIU Y T. Community detection based on social interactions in a social network [J]. Journal of the Association for Information Science and Technology, 2014, 65(3): 539. DOI:10.1002/asi.22986

[3]PIZZUTI C, ROMBO S E. Algorithms and tools for protein-protein interaction networks clustering, with a special focus on population-based stochastic methods[J]. Bioinformatics, 2014, 30(10):1343

[4]NEWMAN M E J. From the cover: The structure of scientific collaboration networks[J]. Proceedings of the National Academy of Sciences, 2001, 98(2):404. DOI:10.1073/pnas.021544898

[5]WATIS D J, STROGATZ S H. Collective dynamics of ‘small-world’ networks[J]. Nature, 1998, 393(6684):440

[6]BARABÁSI A, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439):509

[7]ZHANG W, SHANG R, JIAO L. Complex network graph embedding method based on shortest path and MOEA/D for community detection [J]. Applied Soft Computing, 2020, 97:106764. DOI:10.1016/j.asoc.2020.106764

[8]FANG W, WANG X, LIU L, et al. Community detection through vector-label propagation algorithms[J/OL]. arXiv, 2020. DOI:10.48550/arXiv.2011.08342

[9]YUE Y, WANG G, HU J, et al. An improved label propagation algorithm based on community core node and label importance for community detection in sparse network [J]. Applied Intelligence,

- 2023, 53(1):1. DOI:10.1007/s10489-022-04397-0
- [10] CAI B, WANG Y, ZENG L, et al. Edge classification based on convolutional neural networks for community detection in complex network[J]. *Physica A: Statistical Mechanics and its Applications*, 2020, 556:124826
- [11] RUEDA D F, CALLE E, MARZO J L. Robustness comparison of 15 real telecommunication networks: Structural and centrality measurements[J]. *Journal of Network and Systems Management*, 2017, 25(2):269. DOI:10.1007/s10922-016-9391-y
- [12] PAGANI G A, AIELLO M. The power grid as a complex network: A survey[J]. *Physica A: Statistical Mechanics & Its Applications*, 2013, 392(11):2688. DOI:10.1016/j.physa.2013.01.023
- [13] PALLA G, DERÉNYI I, FARKAS I, et al. Uncovering the overlapping community structure of complex networks in nature and society[J]. *Nature*, 2005, 435(7043):814. DOI:10.1038/nature03607
- [14] AHN Y Y, BAGROW J P, LEHMANN S. Link communities reveal multiscale complexity in networks[J]. *Nature*, 2010, 466(7307):761
- [15] BISWAS A, BISWAS B. FuzAg: Fuzzy agglomerative community detection by exploring the notion of self-membership[J]. *IEEE Transactions on Fuzzy Systems*, 2018, 26(5):2568. DOI:10.1109/TFUZZ.2020.2980502
- [16] RAGHAVAN U N, ALBERT R, KUMARA S. Near linear time algorithm to detect community structures in large-scale networks[J]. *Physical Review E*, 2007, 76(3):036106
- [17] GREGORY S. Finding overlapping communities in networks by label propagation[J]. *New Journal of Physics*, 2010, 12(10):103018. DOI:10.1088/1367-2630/12/10/103018
- [18] XIE J, SZYMANSKI K B, LIU X. SLPA: Uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process[J]. *CoRR*, 2011, abs/1109.5720
- [19] YAZDANPARAST S, JAMALABDOLAH M, HAVENS T. Linear time community detection by a novel modularity gain acceleration in label propagation[J]. *IEEE Transactions on Big Data*, 2020. DOI:10.1109/TBDATA.2020.2995621
- [20] ZHAO J. A community detection algorithm based on graph compression for large-scale social networks[J]. *Information Sciences*, 2021, 551:1
- [21] GUI Q, DENG R, XUE P, et al. A community discovery algorithm based on boundary nodes and label propagation[J]. *Pattern Recognition Letters*, 2018, 109:76
- [22] RAJESH J, SHEELA R. Detecting overlapping communities using ensemble-based distributed neighbourhood threshold method in social networks[J]. *Intelligent Decision Technologies*, 2021, 15(2):217
- [23] LU M, ZHANG Z, QU Z, et al. LPANNI: Overlapping community detection using label propagation in large-scale complex networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019, 31(9):1736
- [24] KOUNI E B I, KAROU I W, ROMDHANE B L. Node importance based label propagation algorithm for overlapping community detection in networks[J]. *Expert Systems with Applications*, 2020, 160:113668
- [25] MORO A H, MAKSE H A. Influence maximization in complex networks through optimal percolation[J]. *Nature*, 2015, 524(7563):65
- [26] LANCICHINETTI A, FORTUNATO S, KERTÉSZ J. Detecting the overlapping and hierarchical community structure in complex networks[J]. *New Journal of Physics*, 2009, 11(3):033015
- [27] MCDAID F A, GREENE D, HURLEY J N. Normalized mutual information to evaluate overlapping community finding algorithms[J]. *CoRR*, 2011, abs/1110.2515
- [28] NICOSIA V, MANGIONI G, CARCHIOLO V, et al. Extending the definition of modularity to directed graphs with overlapping communities[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2009, 2009(3):P03024
- [29] LANCICHINETTI A, FORTUNATO S, RADICCHI F. Benchmark graphs for testing community detection algorithms[J]. *Physical Review E*, 2008, 78(4):046110
- [30] SUN H L, HUANG J B, TAN Y Q, et al. Detecting overlapping communities in networks via dominant label propagation[J]. *Chinese Physics B*, 2015, 24(1):018902

(编辑 丁晓清)