

DOI: 10.13718/j.cnki.xdzk.2024.02.015

魏会廷, 陈永光. 面向社交网络重要信息传播的重叠节点挖掘模型研究 [J]. 西南大学学报(自然科学版), 2024, 46(2): 150-158.

面向社交网络重要信息传播的 重叠节点挖掘模型研究

魏会廷¹, 陈永光²

1. 许昌学院 实验室与设备管理中心, 河南 许昌 461000; 2. 周口师范学院 教育科学学院, 河南 周口 466001

摘要: 针对动态社交网络中的社区检测问题, 提出一种面向社交网络重要信息传播的重叠节点挖掘模型 (SNONMM), 结合标签传播算法 (LPA) 和扩散激活原理, 实现对动态社交网络中重叠社区的高效检测. 该模型的新节点在社交网络中向其他节点传播其标签的机会大于旧节点, 从而使新节点更容易被发现并纳入相应的社区. 同时, 引入激活值来表示每个标签的传播强度, 有助于更准确地捕捉社区结构的变化. 为了验证该方法的有效性, 通过两个真实数据集和一个人工合成网络对其性能进行评估. 实验结果表明, 该方法在检测社区准确性方面优于其他可用方法.

关键词: 动态社交网络; 社区检测; 标签传播算法; 扩散激活

中图分类号: TP391 **文献标志码:** A

文章编号: 1673-9868(2024)02-0150-09

开放科学(资源服务)标识码(OSID):



Research on Overlapping Node Mining Model for Important Information Dissemination in Social Networks

WEI Huiting¹, CHEN Yongguang²

1. Laboratory and Equipment Management Center, Xuchang University, Xuchang Henan 461000, China;

2. School of Education Science, Zhoukou Normal University, Zhoukou Henan 466001, China

Abstract: This paper addresses the problem of community detection in dynamic social networks and proposes a Social Network Overlapping Node Mining Model (SNONMM), and aims at the efficient detection of overlapping communities in dynamic social networks. The model combines the Label Propagation Algorithm (LPA) and the Spreading Activation principle to achieve efficient detection of overlapping communities in dynamic social networks. In this approach, new nodes have a greater chance of spreading their labels to other nodes in the social network than that of old nodes, making new nodes more easily discovered and incorporated into their respective communities. At the same time, activation values are introduced to

收稿日期: 2023-05-10

基金项目: 国家自然科学基金项目(120001471); 河南省 2023 年哲学社会科学规划项目(2023BJY041).

作者简介: 魏会廷, 硕士, 副教授, 主要从事信息技术应用研究.

represent the propagation strength of each label, which helps to more accurately capture changes in community structure. To validate the effectiveness of the proposed method, its performance was evaluated using two real-world datasets and a synthetic network. Experimental results demonstrate that the proposed method outperforms other available methods in terms of community detection accuracy.

Key words: dynamic social networks; community detection; label propagation algorithm; diffusion activation

随着互联网和智能手机的普及, 社交网络已经成为现代社会中人们日常生活和工作的重要组成部分. 而社交网络中信息传播迅速, 影响力巨大, 关键节点在信息传播过程中发挥着重要作用. 对于社交网络中重要信息传播的研究, 挖掘重叠节点具有重要的理论和实际价值. 特别是在信息传播和舆情监控等领域, 重叠节点的挖掘可以帮助人们更好地理解信息是如何在社交网络中扩散, 以及如何更有效地监控和管理这些传播过程.

社交网络信息传播是指信息在社交网络中的扩散过程^[1]. 随着社交网络的普及和快速发展, 信息传播研究逐渐成为学术界和产业界关注的焦点. 在这一领域的研究中, 基本理论和传播模型是两个主要方向^[2]. 基本理论主要关注信息传播规律、影响因素和机制^[3], 如网络结构、节点属性、信息内容等对信息传播过程的影响^[4]. 网络结构分析主要关注社交网络中信息传播的特点, 如小世界现象和幂律分布等^[5]; 节点属性研究则关注个体在信息传播中的角色, 包括意见领袖、桥接节点等; 信息内容研究主要关注信息本身的特性, 如话题、情感和传播力等, 以及这些特性如何影响信息在社交网络中的传播^[6].

传播模型则主要用于描述和预测信息传播过程^[7]. 国内外的学者提出了很多模型, 而其中最著名的是独立级联模型(ICM)^[8]和线性阈值模型(LTM)^[9]. 在独立级联模型中每个节点有一定的概率被其邻居影响, 从而采纳某种行为或观点. 这个概率通常根据节点间的关系强度来确定^[10]. 线性阈值模型则认为节点在邻居中所占的影响力达到一定阈值时, 才会采纳新的行为或观点^[11]. 此外, 还有其他的传播模型, 例如信息扩散模型(IDM), 为国内外学者提供了多种信息传播过程的建模方法^[12]. 尽管学者们在社交网络信息传播研究中取得了显著成果, 但这一领域的研究仍然面临诸多挑战. 社交网络中信息传播往往受多方面因素的影响, 如时效性、跨平台传播和社交媒体上的算法推荐等. 这些因素使信息传播的规律和模式变得更加复杂, 给信息传播研究和预测带来了困难.

由于重叠节点在社交网络中起到关键作用, 尤其在信息传播过程中具有特殊意义^[13], 在过去的研究中尽管学者们提出了许多方法来挖掘这些重叠节点(包括基于社区划分的方法^[14]、基于网络结构的方法^[15]、基于随机游走的方法^[16]及基于机器学习的方法)^[17], 但是由于节点的动态性、社区结构的复杂性及信息传播的多样性, 仍然对社交网络中挖掘重叠节点造成了困难和挑战. 在动态社交网络中, 节点和边的变化频繁, 传统的社区检测方法往往难以应对这种变化^[18]. 此外, 社交网络中的社区结构通常具有重叠性, 即一个节点可能属于多个社区, 这种状态给社区检测带来了额外的复杂性^[19].

为了解决目前研究的不足, 本文提出一种面向社交网络重要信息传播的重叠节点挖掘模型(SNONMM), 该模型综合了标签传播算法(LPA)和扩散激活原理, 旨在克服现有方法的局限性. 标签传播算法通过在网络中传播节点标签来发现社区结构; 而扩散激活原理则模拟了信息在网络中的传播过程. 两者相结合使 SONMM 模型能够有效地挖掘社交网络中的重叠节点.

1 社交网络重要信息传播的重叠节点挖掘模型(SNONMM)

1.1 动态网络和扩散激活过程

动态网络通常由一个网络 $G(V, E)$ 表示, 其中 V 是顶点的集合, 而 E 是边的集合. 本文定义动态网络 SG 为一组顺序网络 $SG = \{G_1, G_2, \dots, G_t\}$, 其中每个网络对应一次迭代. 例如, G_1 对应第一次迭代, t 表示从 1 开始的迭代编号. 由于本文方法和其他可用方法都是为动态重叠社区检测提供的, 因此本文模型考虑了这类社区. 重叠动态社区检测是指在 SG 中找到密集连接的重叠子网络任务.

扩散激活过程是一种搜索方法, 通过在网络中随机选择一个或多个节点作为源节点开始. 每个节点都对应一个初始化激活值, 源节点的激活值为 1, 其他节点的激活值为 0. 每个节点的激活值表示该节点传播信息所需的激活量. 每个源节点通过广度优先搜索方法将其激活值发送给其他节点. 如式(1)所示, 接收节点根据以下 3 个参数更新它们的激活值: ① 源节点的激活值; ② 源节点与目标节点之间的链接权重; ③ 衰减因子.

$$A[j] = A[j] + \sum_{i \in \text{neighbours}(j)} (A[i] * W[i, j] * D) \quad (1)$$

衰减因子是表示网络边缘扩散强度减小的参数. 当从多条路径到达一个节点时, 该过程停止. 每个源节点创建一个由其可访问节点组成的树. 总的来讲, 节点 j 的激活值是基于节点 i 的激活值计算获得, 其中 i 是源节点, j 是目标节点. $A[i]$ 和 $A[j]$ 分别是节点 i 和 j 的激活值. D 是衰减因子, $W[i, j]$ 表示节点 i 和 j 之间的链接权重.

1.2 SNONMM 原理描述

SNONMM 主要由 4 个部分组成. 第 1 部分是初始化, 其中每个节点都被标记为一对标签名称, 激活值. 本文首先将标签名称设为节点编号, 激活值设置为 1. 标签名称用于区分节点, 而激活值表示节点的状态或活跃程度. 第 2 部分是网络权重分配, 应用了名为社交行为信息扩散模型(SBIDM)的权重算法. 这种算法为每个节点分配一个分数, 并从得分最高的邻居即最佳邻居处获取基于信息. SBIDM 算法的基本原理是通过节点社交行为来分配权重, 从而使信息传播更加高效. 在确定最佳邻居后, 通过增加节点与其最佳邻居之间的边来提升该节点的状态值:

$$\text{stateValue}_v^{t+1} = \left\{ \begin{array}{l} \text{stateValue}_v^t + \text{neighbour}f_v^{t+1} \\ \text{stateValue}_v^t + \text{randNum} \\ \text{stateValue}_v^t + \text{neighbour}f_v^{t+1} + \text{randNum} \end{array} \right\} \quad (2)$$

式(2)中:

$$\text{neighbour}f_v^{t+1} = \text{neighbour}f_v^t + a * (\text{bestneighbour}f_v^t - \text{stateValue}_v^t) \quad (3)$$

这种权重算法旨在实现目标边的权重越大, 则信息流量越大. 本文在具有 7 个顶点的网络中进行了应用, 其中图 1 和图 2 显示了 SNONMM 在具有 7 个顶点的网络中第 1 部分和第 2 部分的结果.

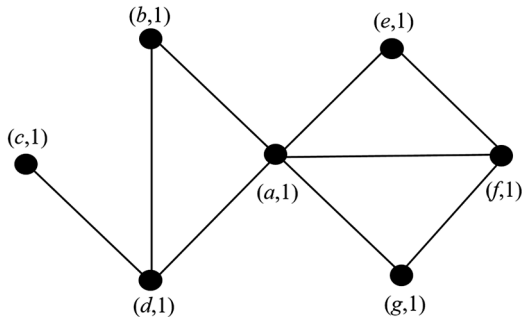


图 1 SNONMM 第 1 部分结果

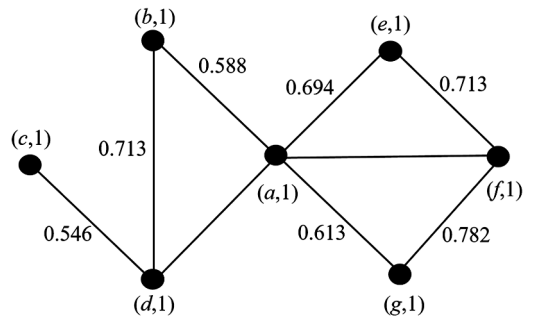


图 2 SNONMM 第 2 部分结果

第 3 部分为所有节点考虑预定义权重的自循环(图 3), 应用了扩散激活过程, 其中所有新节点而非少数节点被选择为源节点. 每个源节点首先将其标签激活值发送给自身, 然后将其发送给通过广度优先搜索访问的其他节点, 直到达到目标状态(激活值低于阈值). 每个目标节点根据 3 个参数更新接收标签的激活值. 第 1 个参数是源节点的激活值, 第 2 个参数是源节点和目标节点之间的边的权重, 第 3 个参数是衰减因子. 当过程完成时, 激活值较低的标签将从节点中移除.

第 4 部分首先将更改的节点按随机顺序排列, 然后从该顺序中选择一个节点作为接收节点, 提取其邻居并将其视为发送节点, 每个发送节点选择激活值大于其他标签的一个标签. 发送节点将其选定的标签发送给接收节点, 接收节点根据发送节点和接收节点之间的边的权重计算发送标签的新激活值. 接收节点对相同标签的激活值求和, 然后接受激活值之和大于其他标签的标签. 如果接收标签的激活值大于 1, 算法

将其设置为 1, 作为激活值范围的结果. 第 4 部分重复进行 5~20 次, 删除激活值较低的标签. 具有近似激活值标签的节点形成社区. 图 4 显示了 SNONMM 在网络中第 4 部分的结果, 表明节点 (a, e, f, g) 和 (a, b, c, d) 在同一个社区中.

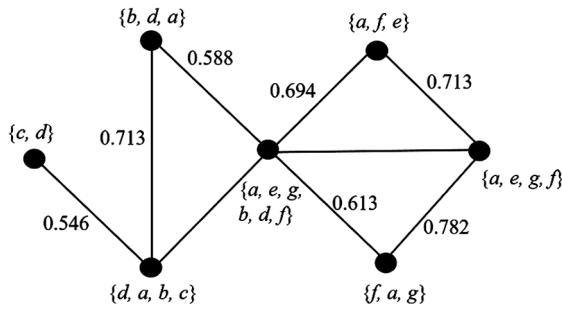


图 3 SNONMM 第 3 部分结果

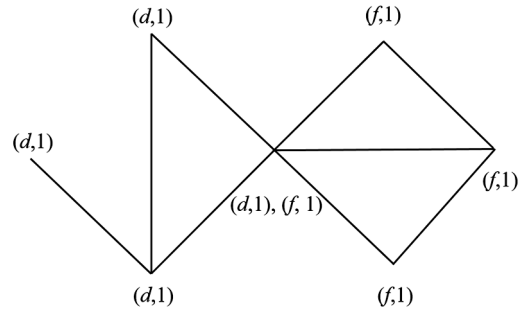


图 4 SNONMM 第 4 部分结果

1.3 模型的算法实现

算法首先为所有新节点分配最大激活值, 然后开始边权重处理过程. 算法 1 首先将所有边的权重初始化为 1, 然后随机选择一个节点并计算其邻居的分数. 在邻居中, 分数大于其他节点的节点被选为最佳邻居. 本文将节点与其最佳邻居之间的权重加上该节点的状态值. 函数 $\text{Normalize}(G(V, E))$ 调整边的权重, 使每条边的权重不超过 1.

算法 1 SBIDMEdgeWeighting(Gts)

Input: $G=(V, E)$

Output: WG: weighted graph

Method:

1. for each e_m in E
2. set $w(e_m)=1$
3. end for each
4. for $i:=1$ to n do
5. Choose node v randomly
6. Calculate score for neighbors of v
7. best_neighbor=select_best_neighbor(neighbor_scores)
8. if best_neighbor exists then
9. Update edge weight between v and best_neighbor based on Eq. (3)
10. Update activation values of v and best_neighbor based on Eq. (2)
11. end if
12. end for
13. for each node v in V do
14. for each neighbor u of v do
15. Normalize edge weight between v and u
16. end for each
17. end for
18. $WG=G(V, E)$

算法 2 中函数 $\text{Spreading Activation}(G(t))$ 计算节点的激活值. 首先为每个节点创建一个名为 UnFiredNodeSet 的空集, 然后根据广度优先搜索进行更新. 从这个集合中提取节点, 并计算它们的激活值, 直到激活值低于阈值或者 UnFiredNodeSet 为空.

算法 2 Spreading Activation ($G(t)$, new nodes)Input: $G(V, E)$ Output: $G(V+, E)$

Method:

1. for each node v in V do
2. Create empty set UnFiredNodeSet
3. Add v to UnFiredNodeSet
4. Neighbors= v . ObtainNbs()
5. Add Neighbors to UnFiredNodeSet
6. while UnFiredNodeSet is not empty do
7. Extract nv from UnFiredNodeSet
8. Compute nv . label. activation_value
9. If nv . label. activation_value < threshold
10. continue
11. end if
12. Neighbors= nv . ObtainNbs()
13. Add Neighbors to UnFiredNodeSet
14. end while
15. end for

算法 3 开始标签传播. 首先通过 ShuffleOrder 函数将更改的节点按随机顺序排列, 并从这个顺序中选择一个节点作为接收节点, 通过函数 ObtainNbsO 提取接收节点的邻居作为发送节点. GetLabelsO 函数选择每个发送节点中激活值最大的一个标签, 创建发送标签的候选标签集. 然后, 接收节点接受候选标签集中激活值之和最大的一个标签. 最后, 删除激活值较低的标签, 并根据标签激活值的相似性提取社区.

算法 3 SNONMM 算法Input: $SG = \{L_1 = \langle V_1, E_1 \rangle, L_2 = \langle V_2, E_2 \rangle, \dots, L_n = \langle V_n, E_n \rangle\}$, T Output: set of communities of SG

Method:

1. Initialize communities
2. for $ts=1; T$ do // ts stands for timestamp
3. for each node v in SG do
4. Node(v). label= v
5. Node(v). label. activation_value=1
6. end for
7. Initialize $G(ts)$
8. end for
9. Initialize ΔV
10. $\Delta V = \{v | v \in V(ts)\}$. 3
11. for $ts=1; T$ do
- // Part 2: Edge Weighting
12. WERW-KpathEdgeWeighting ($G(ts)$) or SBIDMEdgeWeighting ($G(ts)$)
- // end Part 2
- // Part 3: Spreading Activation
13. Spreading Activation ($G(ts)$)

```

// end Part 3
// Part 4: label propagation
14. ChangedNodes = {u, v | (u, v) ∈ E(ts)}
15. V_old = {v | v ∈ G(ts) ∩ v ∉ G(ts+1)}
16. ChangedNodes = ChangedNodes - V_old
17. for it=1: IT do // it means iteration
18. ChangedNodes.ShuffleOrder()
19. for i=1: ChangedNodes.count do
20. Receiver = ChangedNodes(i)
21. Senders = ChangedNodes(i).ObtainNbs()
22. CandidateLabels = Senders.GetLabels()
23. Receiver.update(CandidateLabels.TheMostActivationValueLabel)
24. end for
25. end for
26. Remove Nodes(i) labels with activation value < r
27. Update communities
28. Update SG
// end Part 4
29. end for

```

1.4 复杂度分析

本文方法分为 4 个部分. 第 1 部分中初始化标签需要 $O(n)$, 其中 n 是节点数量. 第 2 部分中应用了权重算法. 在最坏的情况下这一部分需要 $O(m)$, 其中 m 是边的数量. 第 3 部分中由于本文方法对网络每个节点使用广度优先搜索, 因此总共需要 $O(n \times m)$. 第 4 部分包括发送和接收标签, 需要 $O(n)$. 从上述事实可以得出, 本文方法在每次迭代中需要 $O(n \times m)$. 由于迭代次数是一个常数 T , 所以所有社区检测的总复杂度为 $O(n \times m)$.

2 实验与评估

本文选择 2 个引文网络和 1 个合成网络对 SNONMM 进行评估. 在引文网络中只能添加节点, 因此与前一次迭代相比节点数量有所增加. 在本文中, 有向边被视为无向边.

2.1 数据集介绍

Arxiv HEP-PH 是一个引文网络, 涵盖了 1993 年 1 月至 2003 年 4 月期间发表的论文, 包含 34, 546 个节点和 421, 578 条边, 其中节点是论文, 边是论文之间的引用. 因此, 从节点 i 到节点 j 的边被视为论文 i 和 j 之间的引用. 该数据集将 1 个月的论文和交流视为 1 次迭代.

Arxiv HEP-TH 是一个覆盖高能物理理论出版物的引文网络, 论文被视为节点, 论文之间的引用被视为边, 该网络包含 27, 770 个节点和 352, 807 条边. 数据集涵盖了 1993 年 1 月至 2003 年 4 月期间的论文, 且将两个月的论文和交流视为 1 次迭代. 引文数据集的基本信息如表 1 所示.

表 1 引文数据集基本信息

数据集	Arxiv HEP-PH	Arxiv HEP-TH
节点数	34 546	27 770
边数	421 578	352 807
迭代次数	124	62
类型	有向、时序、无权重	有向、时序、无权重

除 2 个真实引文网络外, 本文还将 SNONMM 与其他可用方法的性能在一个人工合成网络中进行评估, 生成这些网络的核心参数如表 2 所示. 其中, S 表示迭代次数, n 表示节点数量, $avgd$ 和 $maxd$ 分别表示节点的平均度和最大度, 符号 $minc$ 和 $maxc$ 分别表示社区的最大和最小规模, 符号 O_n 和 O_{nc} 分别表示重叠节点的数量和每个重叠节点所属的社区数量, μ 是设定为 0.3 的混合参数.

表 2 人工合成网络的参数设置

S	n	$avgd$	$maxd$	$minc$	$maxc$	O_n	O_{nc}	μ
10	7 000	14	40	25	45	200	4	0.3

2.2 评价指标

本文中由于真实网络缺乏已知的基准社区, 因此采用模块度衡量和标准互信息方法来评估检测模型的准确性.

2.2.1 模块度衡量

尽管原始模块度主要处理非重叠社区, 但在本文中社区成员可以相互重叠. 因此, 如式(4)所示, 本文采用了一种拓展模块度方法, 以便衡量具有重叠结构的社区.

$$Q_{ov}^c = \frac{1}{2m} \sum_{c \in C} \sum_{i, j \in V} [A_{i, j} - \frac{k_i k_j}{2m}] \beta_{ic} \beta_{jc} \quad (4)$$

式(4)中:

$$\beta_{ic} = \frac{k_{ic}}{\sum_{c'} k_{ic'}} \quad (5)$$

式(4)、式(5)中, 符号 k_i 和 k_j 分别表示顶点 i 和 j 的度. 符号 m 表示边的数量, $A_{i, j}$ 是邻接矩阵. 符号 β_{ic} 和 β_{jc} 分别表示节点 i 和 j 属于社区 c 的强度. 符号 k_{ic} 和 k_{jc} 分别表示节点 i 和 j 到社区 c 的链接总权重.

2.2.2 标准互信息(NMI)

由于在人工合成网络中已知社区结构, 因此使用标准互信息(Normalized Mutual Information, NMI)^[20]作为评估社区检测的指标. NMI 值的范围在 0~1 之间, 当算法划分的社区结构更准确时, NMI 值较大; 若不太准确, 则 NMI 值较小.

2.3 实验结果

本文选择了动态的 Speaker-Listener 标签传播算法(SLPAD)^[21]和基于动态网络标签传播的进化聚类算法(DLPAE)^[22]来与 SNONMM 在真实网络和人工合成网络中的结果进行对比, 分别显示在图 5 和图 6 中. 选择 SLPAD 和 DLPAE 方法是因为它们都能在动态社交网络中检测到重叠社区.

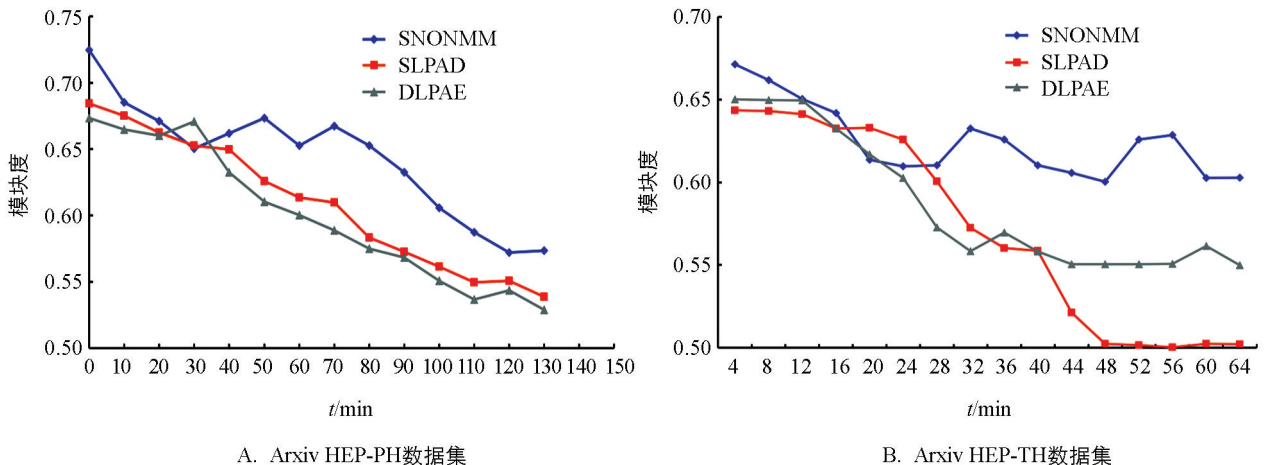


图 5 不同算法在 Arxiv HEP-PH 和 Arxiv HEP-TH 网络上获得的模块度值

图 5A 绘制了 Arxiv HEP-PH 数据集上本文方法与其他可用方法在模块度方面的比较结果. 从图 5A 的数据中可以明显看出所有方法(包括本文的 SNONMM 和其他可用方法)呈现出递减趋势. 这是因为所有方法在第 1 次迭代中考虑了整个网络, 而在迭代 2 及以后的迭代中与前一次迭代相比, 考虑了发生变化的边和节点. 由于迭代 2 及以后的迭代中考虑了变化的边和节点, 因此模块度的减少不可避免. 此外, SNONMM 在早期迭代中的模块度减少量大于 SLPAD 和 DLPAE. 然而, 在中间和最后的迭代中, 本文方法模块度的减少量小于其他可用方法. 从以上事实可以得出, SNONMM

在节点和边变化方面的性能优于其他可用方法. 图 5B 绘制了 Arxiv HEP-TH 数据集上 SNONMM 方法与 SLPAD^[20] 和 DLPAE^[21] 方法在模块度方面的比较. Arxiv HEP-TH 数据集中迭代数量是 Arxiv HEP-PH 数据集的一半, 而数据集中的节点数量大致相同, 由于在此数据集中节点仅被添加而没有被删除, 新节点的数量远低于 Arxiv HEP-PH 数据集. 在早期迭代中 SLPAD 和 DLPAE 方法表现出增长趋势, 而在后期迭代中它们的准确性显著降低. 除了早期迭代外, SNONMM 在大部分迭代中都实现了更高的模块度. 图 6 显示了 3 种方法在人工合成网络中得出的结果, 可以看出 SNONMM 在人工合成网络中的 NMI 值同样大于 SLPAD^[20] 和 DLPAE^[21] 方法, 即 SNONMM 在社区划分时得到的结果相较于其他两种方法具有更高的准确度.

基于上述结果, SNONMM 方法在上述 3 个数据集中取得了比其他两种方法更好的性能.

3 结论和展望

本文针对社交网络中重要信息传播的重叠节点挖掘问题进行了深入研究, 并提出一种面向社交网络重要信息传播的重叠节点挖掘模型(SNONMM). 该模型结合标签传播算法(LPA)和扩散激活原理, 能够在动态社交网络中提高社区检测的准确性, 使新节点在社交网络中向其他节点传播标签的机会大于旧节点, 且每个标签都有一个表示其传播强度的激活值. 通过对 2 个真实网络和 1 个人工合成网络的实验评估, 本文发现 SNONMM 方法在检测社区准确性方面优于其他可用方法, 表明 SNONMM 在解决动态社交网络中重叠节点挖掘问题方面具有潜在优势. 然而, 动态网络中的社区检测仍然具有挑战性, 尤其是在动态社区发现方面, 尽管本文 SNONMM 在一定程度上解决了这个问题, 但未来的研究仍需进一步探讨和优化这些方法, 以提高其在更广泛应用场景中的性能. 在未来的研究中, 可以考虑对网络结构和社交媒体上的内容进行融合, 以提高社区检测的准确性和效率, 并结合多层次和多维度数据, 更好地捕捉和描述动态社交网络中的社区结构变化, 同时可以开发更多针对动态网络的社区检测算法, 并对它们进行跨学科综合研究, 以提高算法的通用性和适应性.

参考文献:

- [1] RAZAQUE A, RIZVI S, KHAN M J, et al. State-of-Art Review of Information Diffusion Models and Their Impact on Social Network Vulnerabilities [J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(1): 1275-1294.
- [2] LI Z L, FANG X A, SHENG O R L. A Survey of Link Recommendation for Social Networks [J]. ACM Transactions on Management Information Systems, 2018, 9(1): 1-26.
- [3] WANG G H, WANG Y F, LI J M, et al. A Multidimensional Network Link Prediction Algorithm and Its Application

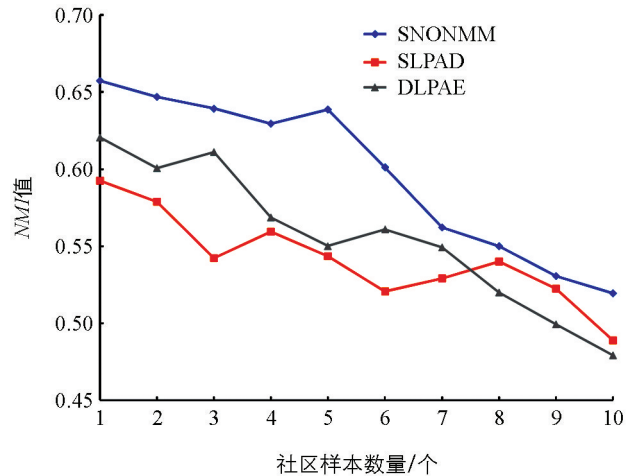


图 6 各种算法在人工合成网络中的 NMI 值

- for Predicting Social Relationships [J]. *Journal of Computational Science*, 2021, 53: 101358.
- [4] IRIBARREN J L, MORO E. Affinity Paths and Information Diffusion in Social Networks [J]. *Social Networks*, 2011, 33(2): 134-142.
- [5] 许进, 杨扬, 蒋飞, 等. 社交网络结构特性分析及建模研究进展 [J]. *中国科学院院刊*, 2015, 30(2): 216-228.
- [6] LI M, WANG X A, GAO K, et al. A Survey on Information Diffusion in Online Social Networks: Models and Methods [J]. *Information*, 2017, 8(4): 118-138.
- [7] WANG F, WANG H Y, XU K. Diffusive Logistic Model towards Predicting Information Diffusion in Online Social Networks [C] //2012 32nd International Conference on Distributed Computing Systems Workshops. Macau: IEEE, 2012.
- [8] 莫小梅, 沈浩, 俞定国. 基于复杂网络的全球新闻流动模式分析 [J]. *西南大学学报(自然科学版)*, 2020, 42(12): 15-24.
- [9] DA SILVA A R, RODRIGUES R F, DA FONSECA VIEIRA V, et al. Influence Maximization in Network by Genetic Algorithm on Linear Threshold Model [C] //International Conference on Computational Science and Its Applications. Cham: Springer, 2018.
- [10] ORIEDI D, MUSUMBA G. Modeling Online User Relationships: a Review [C] //2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC). Suzhou: IEEE, 2022.
- [11] ZHOU F, JIAO J R, LEI B Y. A Linear Threshold-Hurdle Model for Product Adoption Prediction Incorporating Social Network Effects [J]. *Information Sciences*, 2015, 307: 95-109.
- [12] LI J X, PENG W, LI T, et al. Social Network User Influence Sense-Making and Dynamics Prediction [J]. *Expert Systems With Applications*, 2014, 41(11): 5115-5124.
- [13] 杜诗雨, 齐佳音. 在线社交网络结构对信息传播的影响研究——以北京邮电大学某实验室为例 [J]. *北京邮电大学学报(社会科学版)*, 2013, 15(6): 1-7.
- [14] 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.
- [15] 任薇, 阮淇昱, 韩孟凯, 等. 一种基于语义推理的网络社区发现模型 [J]. *西南大学学报(自然科学版)*, 2022, 44(9): 178-186.
- [16] 昌阳, 马慧芳. 基于染色随机游走的可重叠社区发现 [J]. *计算机工程与科学*, 2022, 44(5): 834-844.
- [17] VAN LIERDE H, CHOW T W S, CHEN G R. Scalable Spectral Clustering for Overlapping Community Detection in Large-Scale Networks [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2020, 32(4): 754-767.
- [18] 江旭晖, 沈英汉, 李子健, 等. 社交知识图谱研究综述 [J]. *计算机学报*, 2023, 46(2): 304-330.
- [19] 顾天凯, 王朝坤, 楼昀恺. 时序网络中短时社区搜索方法研究 [J]. *计算机学报*, 2022, 45(2): 334-353.
- [20] 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.
- [21] ASTON N, HERTZLER J, HU W. Overlapping Community Detection in Dynamic Networks [J]. *Journal of Software Engineering and Applications*, 2014, 7(10): 872-882.
- [22] LIU K, HUANG J B, SUN H L, et al. Label Propagation Based Evolutionary Clustering for Detecting Overlapping and Non-Overlapping Communities in Dynamic Networks [J]. *Knowledge-Based Systems*, 2015, 89: 487-496.

责任编辑 夏娟