

DOI: 10.13718/j.cnki.xdzk.2023.08.004

杨智翔, 许小可, 肖婧. 基于高阶信息局部策略的符号网络社区检测 [J]. 西南大学学报(自然科学版), 2023, 45(8): 31-47.

# 基于高阶信息局部策略的符号网络社区检测

杨智翔, 许小可, 肖婧

大连民族大学 信息与通信工程学院, 辽宁 大连 116600

**摘要:** 现有符号网络社区检测方法中, 局部搜索策略作为符号网络社区检测算法的重要组成部分, 可加速算法收敛速度, 但符号网络局部搜索策略大多仅利用连边及节点等低阶结构信息, 忽略了可挖掘符号网络更深层、更丰富信息的高阶结构. 为提升现有符号网络社区检测的局部搜索策略性能, 提出了一种基于符号模体的局部搜索策略, 设计了一种基于符号模体进行社区迁移的新方法, 将传统社区编号在二元组之间的迁移扩展到了三元组, 综合利用符号网络低阶和高阶拓扑结构信息来优化节点的结构平衡性, 提升算法收敛速度和检测性能. 在模型网络和实证网络上的实验表明, 设计的局部搜索策略相对于现有算法表现出更高的精度和质量.

**关键词:** 社区检测; 局部搜索; 模块度优化; 符号模体;

符号网络

中图分类号: TP391

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2023)08-0031-17

## Community Detection in Signed Network Based on High-Order Information Local Search Strategy

YANG Zhixiang, XU Xiaoke, XIAO Jing

College of Information and Communication Engineering, Dalian Minzu University, Dalian Liaoning 116600, China

**Abstract:** In existing signed network community detection methods, the local search strategy serves as an important component of the signed network community detection algorithm and can accelerate the convergence rate of the algorithm. However, most of the local search strategies in signed networks only utilize low-order structural information such as edges and nodes, thereby, neglecting the deeper and richer high-order structure that can be mined from signed networks. To enhance the performance of the existing signed network community detection local search strategy, this article proposes a local search strategy based on signed motifs. A new method based on signed motif for community migration was designed,

收稿日期: 2023-06-07

基金项目: 国家自然科学基金项目(62173065); 辽宁省自然科学基金项目(2020-MZLH-22).

作者简介: 杨智翔, 硕士, 主要从事复杂网络社区检测研究.

通信作者: 肖婧, 副教授, 硕士研究生导师.

which extends the traditional community numbering migration between binary tuples to ternary tuples. By comprehensively utilizing the low-order and high-order topological structural information of the signed network, the structural balance of nodes is further optimized, thereby, improving the convergence rate and detection performance of the algorithm. Experiments on synthetic and real-world networks demonstrated that the designed local search strategy exhibited good accuracy and quality.

**Key words:** community detection; local search; modularity optimization; signed motif; signed network

现实世界中很多复杂系统都可以用复杂网络模型来表示,例如:社交网络<sup>[1]</sup>、生物网络<sup>[2]</sup>、协作网络<sup>[3]</sup>等.符号网络作为复杂网络中一种特殊的网络,是指边具有正或负符号属性的网络,其中正边和负边分别表示积极和消极的关系.例如:神经元之间存在促进和抑制关系<sup>[4]</sup>;社交关系中信任和不信任<sup>[5-6]</sup>.因此,符号网络可以更好地描述节点之间的关系,符号网络相关研究具有重要的实际意义.

由于负边的存在,使得符号网络中的社区检测和无符号网络中的社区检测具有明显区别.在无符号网络中,社区结构定义为一组节点,这些节点在簇内具有密集连接,在簇之间具有稀疏连接.对于符号网络,社区结构不仅由连接密度定义,也由连接的符号定义.也就是说,社区内部的正链接密集而负链接稀疏,社区之间的负链接密集而正链接稀疏.然而,在现实符号网络中,往往社区内部存在一些负链接,同样社区间也存在一些正链接,这也使符号网络社区检测面临巨大的挑战.

目前针对符号网络社区检测问题的方法大致分为如下几类:基于谱聚类的方法<sup>[7-8]</sup>、基于概率生成模型的方法<sup>[9]</sup>、基于模块度优化的方法<sup>[10]</sup>等.其中模块度优化是最典型的复杂网络社区检测方法之一,由于其高精度和强可移植性等优点被广泛应用.模因算法作为解决组合优化问题的一个非常强大的工具,适合于解决模块度优化的 NP 难问题,已被普遍应用于符号网络社区检测中.它的灵感来源于自然系统模型,该模型将种群的进化适应与个体在其成员生命周期内的学习相结合<sup>[11]</sup>.模因算法是一种基于群体的元启发式算法,它融合了不同的搜索策略,即全局搜索策略和局部搜索策略,在解决优化问题方面,模因算法已被证明比传统进化算法更高效<sup>[10]</sup>.

局部搜索策略作为模因算法中重要组成部分之一,它能极大地提高全局搜索算法寻找全局最优解的速度,并辅助提升全局最优化性能.近年来,局部搜索策略常被学者们用于加速全局算法的收敛.如 Li 等<sup>[10]</sup>提出并比较了 EA-SN 和 CSA-SN 两种进化算法,以及 EAHC-SN 和 CSAHC-SN 两种模因算法.模因算法在进化算法基础上使用了爬山策略,仅对最优种群个体进行局部搜索操作,每次迭代都使用贪婪算法对每个节点选择最优适应度,实验表明模因算法优于进化算法. Amelio 等<sup>[12]</sup>提出了一种结合遗传算法和局部搜索的符号网络多目标社区检测方法,其局部搜索策略在爬山策略的基础上增加了仅将节点移动至正相连的邻居节点社区的限定条件,若符号模块度增加则保留移动,反之则不移动. Che 等<sup>[13]</sup>提出了一种新颖的模因算法(MACD-SN),其中结合了一种新的局部搜索算法,能够以一定的概率跳出局部最佳结果,并快速接近全局最优值.

上述方法仅简单地利用节点和连边等低阶结构信息进行局部搜索操作,忽略了网络中广泛存在且具有统计意义的高阶组织结构,如网络模体<sup>[14-15]</sup>等,因此对于提升符号网络社区检测性能相对有限.模体是复杂网络中的基本拓扑和功能单元<sup>[14-16]</sup>,而符号模体作为一种连边具有正或负符号属性的特殊模体,可展现出符号网络节点间更丰富且有意义的拓扑连接关系,对于局部搜索操作具有很强的指导性意义.

针对上述问题,本研究提出了一种基于符号模体的三元组局部搜索策略(Triad Local Search, TLS).设计了一种基于符号模体进行社区迁移的新方法,将传统社区编号在二元组之间的迁移扩展到了三元组,综合利用符号网络低阶和高阶拓扑结构信息,进一步优化节点的结构平衡性,提升算法收敛速度.为验证局部策略的有效性,TLS 算法同差分进化策略构建成模因算法 DE-SN,实验结果表明 TLS 算法在社区划分的精确性和质量性表现出一定优势.将 TLS 算法迁移至元启发式优化算法为共生生物搜索算法和遗传

算法的模因算法中, 在模型网络和实证网络上均验证了新算法的通用性.

## 1 基础概念和数据集

### 1.1 符号网络定义

符号网络社区检测旨在挖掘社区内部正链接密且社区之间负链接密集的社区划分结构. 本研究主要考虑无向符号网络. 一个符号网络可以抽象为一个图  $G = (V, E)$ , 其中  $V = \{v_1, v_2, \dots, v_n\}$  代表节点(顶点)集合,  $E = \{(v_i, v_j) | v_i, v_j \in V \cap i \neq j\}$  代表边(链接)集合. 图  $G$  还可以表示成加权邻接矩阵  $A = (A_{ij})_{n \times n}$ , 其中  $n$  代表节点数目, 若节点  $v_i$  和  $v_j$  之间存在正边(负边)相连, 则  $A_{ij} = 1 (A_{ij} = -1)$ ; 若节点  $v_i$  和  $v_j$  之间不存在连边, 则  $A_{ij} = 0$ . 图  $G$  的一个社区划分可表示为  $C = \{C_1, C_2, \dots, C_k\}$  且  $C_i \subset V$ , 其中  $k$  代表社区数目, 表示由一组节点组成的第  $i$  个社区. 符号网络上的社区检测问题可以定义为公式(1)的形式<sup>[17]</sup>.

$$\begin{cases} A_{ij} > 0, (v_i, v_j) \in E \wedge (v_i \in C_p) \wedge (v_j \in C_p) \\ A_{ij} < 0, (v_i, v_j) \in E \wedge (v_i \in C_p) \wedge (v_j \in C_q) \end{cases} \quad (1)$$

式中:  $p, q = 1, 2, \dots, k$ , 且  $p \neq q$ .

### 1.2 评价指标

#### 1.2.1 符号模块度 $Q_s$

在无符号网络社区检测中, Newman 等<sup>[18]</sup>定义的模块度函数是最常用的指标之一, 但由于符号网络中存在负边, 因此原始的模块度  $Q$  无法直接应用于符号网络. Gómez 等<sup>[19]</sup>基于正边和负边的贡献, 改进了模块度定义, 使其扩展到符号网络. 本研究中, 将符号模块度  $Q_s$  作为目标函数, 符号网络模块度  $Q_s$  定义为:

$$Q_s = \frac{1}{2w^+ + 2w^-} \sum_i \sum_j \left[ w_{ij} - \left( \frac{w_i^+ w_j^+}{2w^+} - \frac{w_i^- w_j^-}{2w^-} \right) \right] \delta(C_i, C_j) \quad (2)$$

式中:  $w^+$  和  $w^-$  分别为图  $G$  中所有正链接和负链接的权重总和;  $w_{ij}$  为节点  $i$  和节点  $j$  之间连边的权重;  $w_i^+$  ( $w_j^+$ ) 为节点  $i$  ( $j$ ) 的正权重总和;  $w_i^-$  ( $w_j^-$ ) 为节点  $i$  ( $j$ ) 的负权重总和;  $C_i$  ( $C_j$ ) 为节点  $i$  ( $j$ ) 所属社区,  $i, j = 1, 2, \dots, k$ ;  $\delta(C_i, C_j)$  为罚函数, 若节点  $i$  和节点  $j$  同属于一个社区, 则  $\delta(C_i, C_j) = 1$ , 否则为 0.

#### 1.2.2 归一化互信息

归一化互信息(Normalized Mutual Information, NMI)<sup>[20]</sup>可用于评估算法检测结果的精确性. 假设  $A$  和  $B$  是一个网络的两个划分, 那么归一化互信息定义为:

$$NMI(A, B) = \frac{-2 \sum_{i=1}^{C_A} \sum_{j=1}^{C_B} C_{ij} \log\left(\frac{C_{ij} N}{C_i C_j}\right)}{\sum_{i=1}^{C_A} C_i \log\left(\frac{C_i}{N}\right) + \sum_{j=1}^{C_B} C_j \log\left(\frac{C_j}{N}\right)} \quad (3)$$

式中:  $N$  为网络的节点数量;  $C$  为混淆矩阵;  $C_{ij}$  为同时隶属于  $A$  中的第  $i$  个社区和  $B$  中第  $j$  个社区的节点数量;  $C_A$  和  $C_B$  分别为划分  $A$  和  $B$  中的社区数量;  $C_i$  和  $C_j$  分别为  $C$  中第  $i$  行和第  $j$  列元素的和.  $NMI$  取值范围为  $[0, 1]$ , 越接近 1, 表示两个划分越相似.

#### 1.2.3 节点 Frustration 指数

Frustration 指数<sup>[21-22]</sup>是目前衡量符号网络结构平衡最广泛的指标之一, 它统计的是社区内部的负链接数量与社区之间的正链接数量的总数, 表示社区的不平衡程度, Frustration 指数越低代表社区划分越稳定. Frustration 指数(Frustration of Node,  $F_N$ )可定义为:

$$F_N(v_i) = \sum_j \left[ \frac{|A_{ij}| - A_{ij}}{2} \delta(C_i, C_j) + \frac{|A_{ij}| + A_{ij}}{2} (1 - \delta(C_i, C_j)) \right] \quad (4)$$

式中:  $v_i$  为对象节点  $i$ ;  $C_i$  为节点  $i$  所属社区;  $\delta(C_i, C_j)$  为罚函数;  $F_N(v_i)$  表示节点  $i$  的连边中在社区内部的负链接数量与社区之间的正链接数量的总数, 它代表节点  $i$  的不平衡程度, 值越小节点  $i$  越稳定, 当  $F_N=0$  时, 节点  $i$  达到结构平衡。

### 1.3 数据集

#### 1.3.1 模型网络

选择 Yang 等<sup>[17]</sup>设计的随机生成符号网络模型, 生成实验使用的合成符号网络, 该模型可以描述为:

$$SG(c, n, k, p_{in}, p_-, p_+) \quad (5)$$

式中:  $c$  为网络中社区的数量;  $n$  为每个社区包含的节点数量;  $k$  为每个节点的度;  $p_{in}$  为每个节点连接同一社区中其他节点的概率;  $p_-$  为负链接出现在社区内的概率;  $p_+$  为正链接出现在社区之间的概率. 为模拟现实世界中不平衡网络结构, 并且控制生成的社区内部和社区之间的连边概率相同, 因此  $p_{in}$  取值为 0.5, 本研究中生成了以下 3 种不同的不平衡网络对各局部策略进行评估。

①  $SG(4, 32, 32, 0.5, 0, p_+)$ :  $p_+$  以 0.1 为间隔在  $[0, 1]$  范围内取值. 该网络由 128 个节点构成, 分成 4 个社区, 每个社区包含 32 个节点. 每个节点的度为 32, 分别以 0.5 的比率连接社区内部和社区外部的节点. 社区内部负链接出现的概率为 0, 社区之间相连以  $p_+$  比率为正链接。

②  $SG(4, 32, 32, 0.5, p_-, 0)$ :  $p_-$  以 0.1 为间隔在  $[0, 1]$  范围内取值. 该网络社区之间正链接出现的概率为 0, 社区内部相连以  $p_-$  比率为负链接。

③  $SG(4, 32, 32, 0.5, p_-, p_+)$ :  $p_-$  和  $p_+$  分别以 0.1 为间隔在  $[0, 0.5]$  范围内取值. 该网络社区内部相连以  $p_-$  比率为负链接, 社区之间相连以  $p_+$  比率为正链接。

#### 1.3.2 实证网络

为检验算法在现实世界网络中的可行性及性能优势, 选择 2 个社交网络和 4 个生物网络作为实验数据. 社交网络包括描述斯洛文尼亚议会关系的网络(Slovene Parliamentary Party, SPP)<sup>[23]</sup>和描述新几内亚高地的网络(Gahuku-Gama Subtribes, GGS)<sup>[24]</sup>. 生物网络分别为 EGFR 网络<sup>[25]</sup>(表皮生长因子受体途径网络)、Macrophage 网络<sup>[26]</sup>(巨噬细胞网络)、Yeast 网络<sup>[15]</sup>(酵母网络)和 Ecoli 网络<sup>[27]</sup>(大肠杆菌的基因调控网络). 以上实证网络都从 SNAP<sup>[28]</sup>中获得, 具体信息如表 1 所示。

表 1 实证网络信息

网络	节点数/个	边数/条	正边数/条	负边数/条
SPP	10	45	18	27
GGS	16	58	29	29
EGFR	329	779	515	264
Macrophage	678	1 425	947	478
Yeast	690	1 080	860	220
Ecoli	1 461	3 215	1 879	1 336

## 2 算法描述

本节首先介绍现有符号网络中具有代表性的局部搜索算法, 再重点阐述提出的三元组局部搜索策略 TLS. 局部搜索策略是模因算法优于传统进化算法的主要原因之一, 以其收敛速度快、扩展性强等优势, 常被用于提高求解效率和精度. 如图 1 所示的模因算法流程包括生成初始种群、计算所有个体适应度值、元启发式优化算法、更新个体适应度值和局部搜索策略, 本研究提出了一种三元组局部搜索策略(Triad Local Search, TLS)应用于模因框架, 旨在提升其在符号网络社区检测的性能. 为了展示提出策略的优越性, 在模因框架中选取了差分进化算法<sup>[29]</sup>(Differential Evolution, DE)作为元启发式优化算法, 与 TLS 策

略构成  $DE-SN_{TLS}$  框架进行实验验证. 同时, 又在模因框架中选取了共生生物搜索<sup>[30]</sup> (Symbiotic Organisms Search, SOS) 和遗传算法<sup>[31]</sup> (Genetic Algorithm, GA) 作为元启发式优化算法, 与 TLS 策略构成  $SOS-SN_{TLS}$  和  $GA-SN_{TLS}$  框架, 验证本研究提出策略的通用性.

### 2.1 基于低阶信息的局部搜索算法

现有符号网络上代表性的局部搜索算法有以下几种: 爬山算法<sup>[32-33]</sup>、利用正边信息的改进爬山算法<sup>[12]</sup> 以及模因算法  $MACD-SN$ <sup>[13]</sup> 中具有概率跳出局部最优的局部搜索策略.

#### 2.1.1 爬山算法

爬山算法 (Hill Climbing, HC) 是一种局部择优的启发式方法, 核心思想是贪婪地寻找每个节点局部最优适应度, 缺点是极易陷入局部最优. 它是一种经典的局部搜索算法, 被广泛使用于无符号网络社区检测和符号网络社区检测, 如 Gong 等<sup>[33]</sup> 将爬山策略与算法相结合实现无符号网络社区检测, Li 等<sup>[10]</sup> 在其基础上将爬山策略应用于符号网络.

HC 算法的详细说明如图 2 所示, 9 个节点对应 6 个社区的结构示例图为 EGFR 网络提取的一个子图. 在一次局部搜索操作过程中, 节点 A 被选作中心节点 (对象节点), 其对应的社区编号为 1. 初始子图中, 中心节点有一个负边相连的邻居节点 B 和 4 个正边相连的邻居节点 {C, E, G, H}, 其社区编号分别对应为 {2, 3, 4, 5, 6}, 故节点 A 有 4 条不平衡连边 (图中红色连边), 即节点 A 初始的  $F_N$  值为 4. 在 HC 操作中, 首先将中心节点所有邻居节点的社区作为候选社区 (满足条件的邻居社区), 然后分别计算节点 A 归属于每个候选社区后的符号模块度, 最后选择符号模块度最大的候选社区作为迁移社区. 图中当中心节点的社区为 6 时取得  $Q_S$  最大值为 0.143, 因此社区编号 6 作为迁移社区从迁移节点 (被选中的候选节点) 迁移至中心节点, 即实现了社区编号在二元组之间的迁移.

HC 操作后中心节点 A 的社区编号更改为 6, 节点 A 的  $F_N$  值降至 3,  $Q_S$  从 0.087 优化至 0.143, 达到 HC 操作的最优结果. HC 算法未利用符号信息, 而是贪婪地寻找符号模块度最大的迁移社区, 此操作极易使算法陷入局部最优, 并且 HC 算法缺少跳出局部最优的操作.

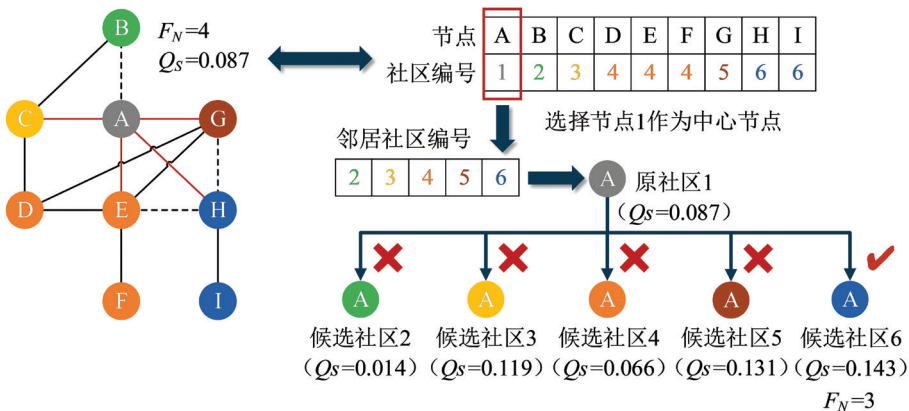


图 2 HC 算法在 EGFR 网络子图上的详细说明

#### 2.1.2 改进爬山算法

Amelio 等<sup>[12]</sup> 提出算法中使用的局部搜索策略是在爬山策略的基础上, 增加了仅将节点移动至正相连

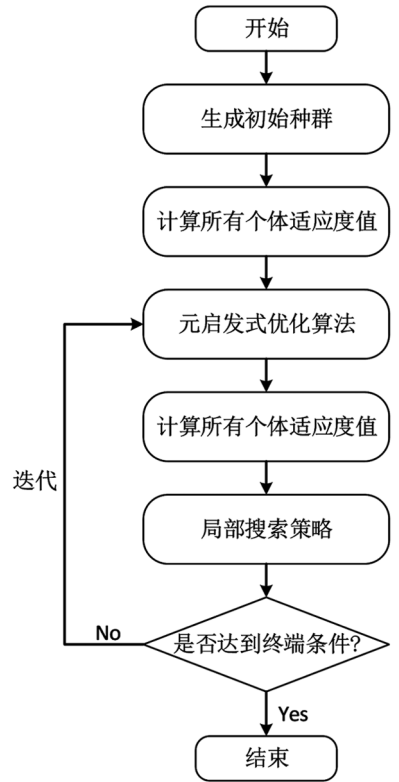


图 1 模因算法流程图



的邻居节点社区的限定条件. 改进爬山策略(Hill Climbing Using Positive, HC-P)简单利用了符号信息, 使爬山策略更适用于符号网络.

HC-P 算法详细说明如图 3 所示, 唯一与 HC 算法操作不同之处在于中心节点只考虑将正边相连的邻居社区 {3, 4, 5, 6} 作为候选社区, 而负边相连的邻居社区 2 不做考虑, 最后从候选社区中选择符号模块度最大社区作为迁移社区. HC-P 操作后中心节点 A 的社区同样更改为 6, 节点 A 的  $F_N$  值也降至 3,  $Q_s$  从 0.087 优化至 0.143, 达到 HC-P 操作的最优结果. HC-P 算法利用符号信息对候选社区进行筛选, 相较于 HC 算法的效率和性能有所提升, 但 HC-P 算法并未对爬山策略极易陷入局部最优的缺陷进行改进.

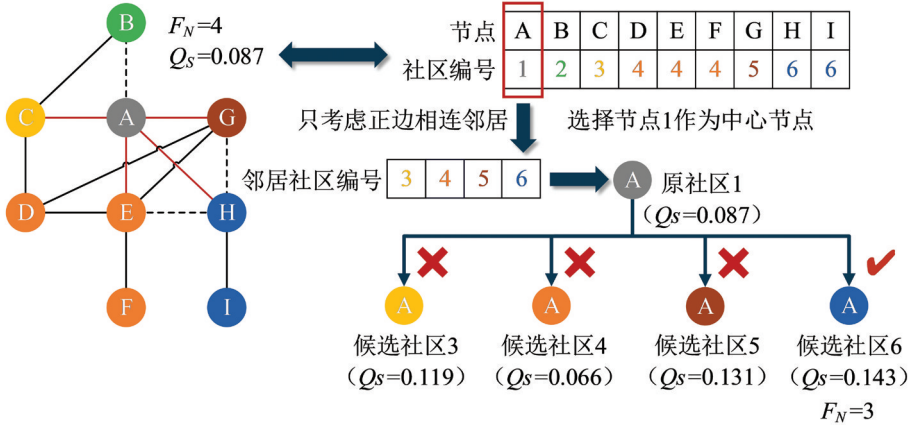


图 3 HC-P 算法在 EGFR 网络子图上的详细说明

### 2.1.3 MACD-SN 中的局部搜索策略

Che 等<sup>[13]</sup>提出的符号网络社区检测算法中使用了局部搜索策略(Local Seek, LS), 使用作者自定义的算子(社区失衡度  $CID$ , 表示同社区内节点的平均 Frustration 值)作为局部优化目标. LS 算法结合了符号信息, 并引入概率  $\rho$ , 一定概率接受较差的个体, 有利于算法跳出局部最优解.

LS 算法的详细说明如图 4 所示, 在同一个子图上, 节点 A 初始的  $F_N$  值为 4. 首先, LS 算法将正边相连的邻居社区 {3, 4, 5, 6} 作为候选社区, 若无正边相连的邻居社区, 才会将负边相连的邻居社区作为候选社区. 其次, 选择  $CID$  最小的候选社区 6 再对比前后  $Q_s$ , 若  $Q_s$  增大或者存在一定概率  $\rho$  情况下, 则将该候选社区作为迁移社区; 反之, 不进行社区迁移, 中心节点社区保持不变. 由于节点 A 的社区改为 6 时,  $Q_s$  增大至 0.143, 因此, LS 操作后节点 A 的社区更换为 6, 且  $F_N$  值为 3, 达到 LS 操作的最优结果.

但对于真实社区划分为社区间正边数量较多的网络, 受局部优化目标影响, LS 算法容易将正边相连不同社区的节点划分成同一社区, 从而导致社区划分质量不佳.

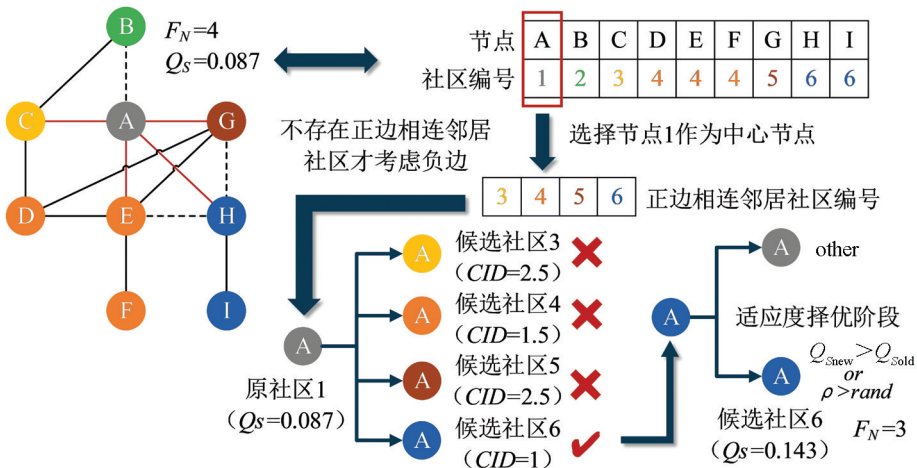


图 4 LS 算法在 EGFR 网络子图上的详细说明

### 2.2 基于高阶信息的局部搜索算法

上一小节中的 3 种基于低阶信息的局部搜索算法, 社区只在二元组之间进行迁移, 即从迁移节点迁移至中心节点, 这种迁移方式的效率不高, 每次操作只对中心节点进行社区调整, 中心节点邻域内的不合理连边调整缓慢. 此外, 单纯关注局部目标来选择局部优化最佳的候选社区(候选节点)作为迁移对象, 往往会陷入局部最优的社区划分.

为解决上述问题, 提出一种基于符号模体的三元组局部搜索算法(Triad Local Search, TLS), 可以有效提高收敛速度和局部性能. 核心思想为通过综合利用符号网络低阶和高阶拓扑结构信息, 优化节点的不平衡程度, 尽可能达到节点的结构平衡, 从而提升社区划分质量. TLS 算法对符号网络高阶和低阶信息的利用, 分别体现在基于三阶符号模体进行三元组社区迁移和依据低阶连边符号信息选择候选节点. 三元组社区迁移是将迁移节点迁移至中心节点的社区编号再传递给其他候选节点, 为更准确表示该迁移模式并探索更深层的规律, 本研究使用三阶符号模体表示三元组社区迁移, 并从中挖掘高阶结构信息, 进一步指导局部搜索操作.

下面详细介绍 TLS 算法具体实现. 研究中主要考虑无向的三阶符号模体, 如图 5 所示一共有 7 个三阶的无向符号模体, 其中  $S_1, S_2$  和  $S_3$  只有两条连边. 符号模体上的社区迁移需要满足以下 2 个条件:

- 1) 根据正链接和负链接的符号属性, 社区迁移只发生在节点间的正链接, 而社区不可以负链接进行迁移.
- 2) 模体中的社区迁移只能通过中心节点进行, 两个都为非中心节点, 无法进行社区迁移.

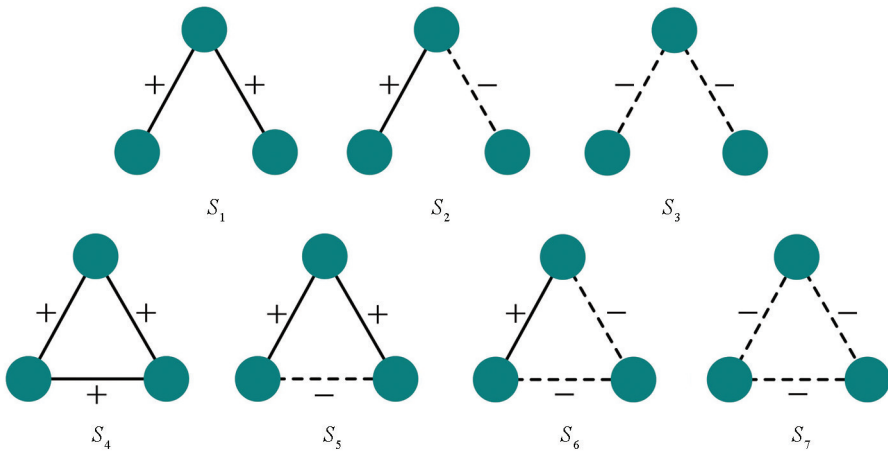


图 5 无向的三阶符号模体

图 5 中的 7 种符号模体依据中心节点位置的不同, 将会有 12 种不同的符号模体. 然而, 研究发现并非 12 种符号模体都能够满足符号模体的社区迁移条件, 只有图 6 中的 3 种符号模体满足迁移条件, 其中红色节点表示中心节点, 将该模体命名为中心节点符号模体, 并做如下定义:

**定义 1** 中心节点符号模体(Central Node Signed Motif,  $S_C$ ). 满足中心节点职能的符号模体称为中心节点符号模体, 即中心节点符号模体可以让中心节点实现社区信息的接收和传递.

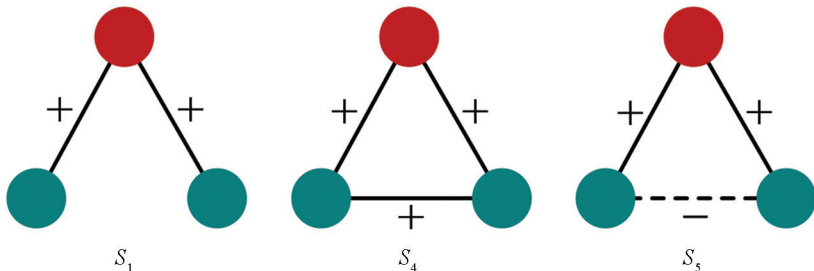


图 6 中心节点符号模体

由于非中心节点符号模体不满足社区迁移条件, 即不参与社区迁移, 因此在 TLS 算法中只需要关注 3 种中心节点符号模体以提高计算效率. 算法 1 为 TLS 算法伪代码, 主要步骤如下:

1) 随机选择目标个体的若干个节点进行优化, 获得中心节点正边相连的邻居节点(作为候选节点)及对应社区编号(第 1 行至第 4 行);

2) 中心节点在候选节点中随机选择一个节点作为迁移节点, 并将迁移节点的迁移社区传递给中心节点, 实现二元组之间的社区迁移(第 5 行至第 6 行);

3) 寻找所有包含中心节点和迁移节点的中心节点符号模体, 为加快代码执行效率, 可将寻找模体的操作简化为求中心节点正边相连且不同社区的邻居节点. 在进行社区迁移时进行条件判断, 若传递后的符号模块度大于原符号模块度或在一定阈值  $\rho$  内, 则确认传递, 邻居节点社区编号替换为中心节点的社区编号, 从而实现了社区编号在三元组(模体)之间的迁移. 否则, 传递失败, 邻居节点社区编号保持不变(第 7 行至第 16 行).

#### 算法 1 TLS 算法

输入: 目标个体  $X$ , 邻接矩阵  $A$ , 阈值参数  $\rho$

输出: 优化后新个体  $X_{new}$

1.  $n_{rand} = \text{random}(1, n)$
2. **for**  $N_c = 1$  **to**  $n_{rand}$  **do**
3.      $\text{neighbors}_\rho = \text{get neighbors positive}(N_c)$
4.      $\text{commID}(\text{neighbors}_\rho) = \mathbf{X}[\text{neighbors}_\rho]$
5.      $N_m = \text{random}(\text{neighbors}_\rho)$
6.      $\text{commID}(N_c) = \text{commID}(N_m)$
7.     **for each**  $np \in \text{neighbors}_\rho$  **do**
8.         **if**  $\text{commID}(np) \neq \text{commID}(N_c)$  **then**
9.              $Q_{Sold} = \text{compute } Q_S$
10.             $Q_{Snew} = \text{compute } Q_S$
11.            **if**  $Q_{Snew} > Q_{Sold}$  **or** random number  $r \leq \rho$  **then**
12.                 $\text{commID}(np) = \text{commID}(N_c)$
13.            end if
14.         end if
15.     end for
16. end for
17. 输出优化后新个体  $X_{new}$

TLS 算法的详细说明如图 7 所示, 同样以 EGFR 网络子图为例, 在一次随机选择若干个节点进行局部搜索操作时, 节点 A 被选作中心节点, 其对应的社区编号为 1. 在节点 A 进行 TLS 操作时, 初始的  $F_N$  值为 4. 首先, 检查中心节点所有邻居节点, 可以发现有一个负边相连的邻居节点 B 和 4 个正边相连的邻居节点 {C, E, G, H}, 其社区编号分别对应为 {2, 3, 4, 5, 6}. 其次, 将正边相连的邻居节点作为候选节点, 并随机选择一个候选节点 E 作为迁移节点, 迁移节点 E 将迁移社区 4 传递给中心节点 A, 此时节点 A 的  $F_N$  值减小至 3,  $Q_S$  降为 0.066. 然后, 寻找所有包含中心节点和迁移节点的中心节点符号模体, 并逐个进行社区迁移. 如图有 3 个满足条件的符号模体, 分别为  $S_1(E, A, C)$ ,  $S_4(E, A, G)$  和  $S_5(E, A, H)$ . 三元组社区迁移是否成功在于迁移后的符号模块度是否增大, 并存在概率  $\rho$  接受较差情况. 因为不同模体的判别顺序不影响最终结果, 所以先对  $S_1(E, A, C)$  进行判别, 发现节点 C 社区更改为 4 后,  $F_N$  值减小至 2,  $Q_S$  增大至 0.092, 符合判别条件, 即接受社区迁移. 再对  $S_4(E, A, G)$  进行判别, 节点 G 社区更改为 4 后,  $F_N$  值减小至 1,  $Q_S$  增大至 0.184, 接受社区迁移. 最后对  $S_5(E, A, H)$  进行判别, 若节点 H 的社区更改



为 4, 则  $Q_s$  降至 0.002, 假设此时未能达到概率  $\rho$ , 即不接受这个较差的情况, 那么最终 TLS 操作将以  $Q_s$  为 0.184 的个体输出。

上述操作不难发现, TLS 算法与传统的局部策略截然不同. TLS 算法不依赖局部目标随机选择候选社区(候选节点)方式, 有助于扩大解空间, 避免陷入局部最优解. 虽然随机选择候选社区会获得更差的社区结构, 如图 7 中  $Q_s$  不升反降. 但在 TLS 操作后期, 以节点的结构平衡( $F_N=0$ )为优化方向的三元组社区迁移, 能够加速中心节点及邻域收敛至的最优社区划分. 以 EGFR 网络子图为例, TLS 算法能够获得 3 种对比算法都无法获得的社区划分, 并且相较于三者的社区划分质量更优。

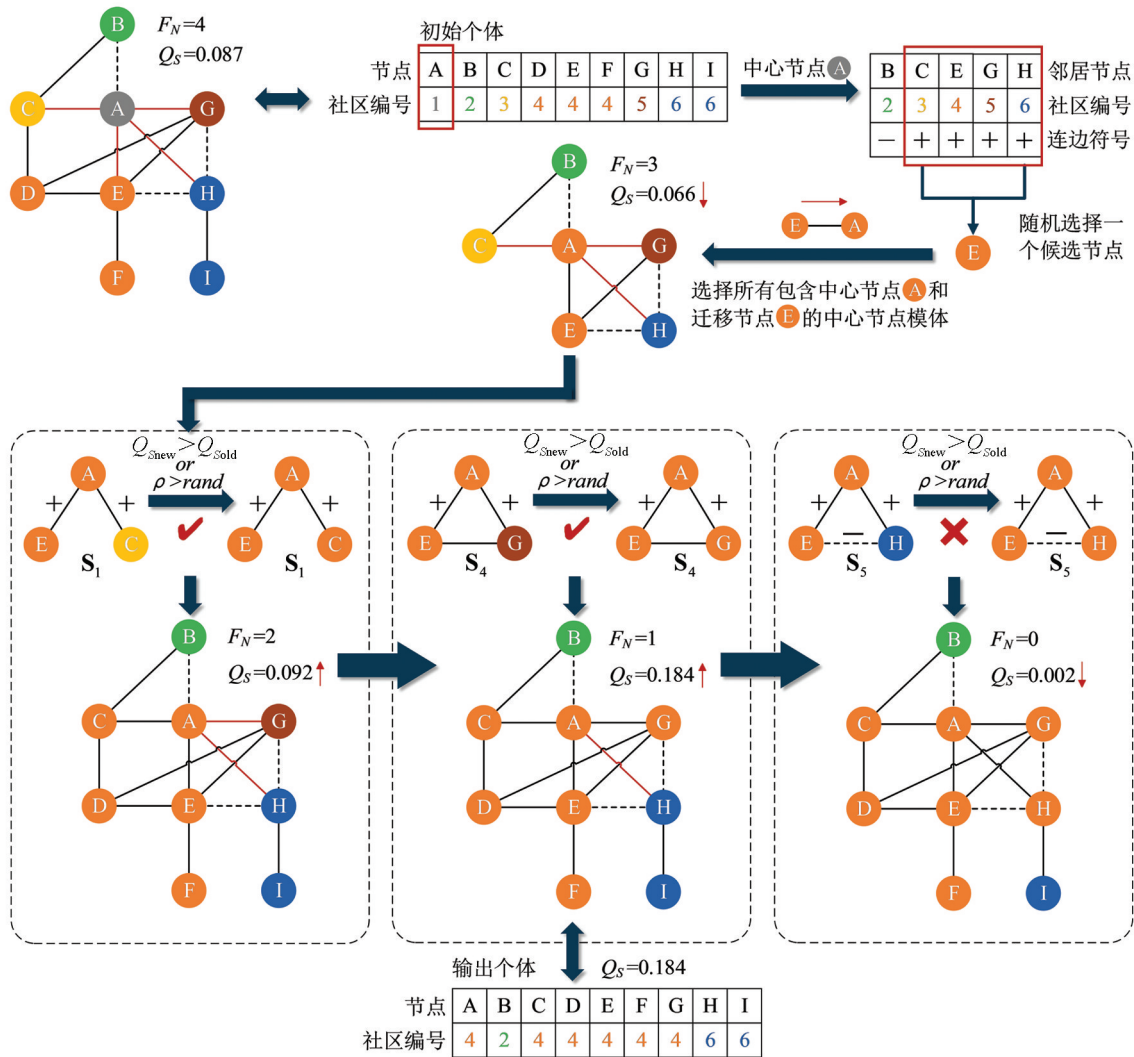


图 7 TLS 算法在 EGFR 网络子图上的详细说明

### 2.3 算法复杂度分析

HC 算法、HC-P 算法和 LS 算法的平均时间复杂度都为  $O(n * d)$ , 其中  $d$  代表网络  $G$  中节点的平均度. 在最差的情况下, 网络  $G$  为完全图, 因此 3 种算法的时间复杂度为  $O(n^2)$ . TLS 算法平均时间复杂度同样为  $O(n * d)$ , 最差时间复杂度为  $O(n^2)$ , 但不同于其他 3 种算法, 并不是对一个目标个体的所有节点都进行局部操作, 而是随机选择若干个节点, 实际运行效率会更高。

## 3 实验结果与分析

本节将对提出 TLS 算法的性能进行验证, 与现有典型的局部搜索算法进行对比, 分别从不同正负边比例的模型网络和实证网络的实验来检验分析。

### 3.1 实验设置

#### 3.1.1 对比策略

为检验 DE-SN 算法在符号网络中的性能,并重点验证提出的三元组局部搜索算法 TLS 与其他局部策略相比更具优势,将 DE-SN 算法中的局部策略分别替换为 HC 算法、HC-P 算法、LS 算法,从而可以得到  $DE-SN_{HC}$ ,  $DE-SN_{HC-P}$ ,  $DE-SN_{LS}$  3 种对比算法。

#### 3.1.2 参数设置

上述搜索策略参数设置均采用原文献中的推荐值,其中 DE 算法中变异因子  $F$  取值为 0.9,交叉因子  $CR$  取值为 0.3. DE-SN 算法设置种群规模  $NP$  为 100,最大迭代次数  $t_{max}$  为 100. 为保持种群个体多样性及探索能力,防止算法过早收敛,仅选择模块度值较为优异的前 10% 个体作为目标个体进行局部操作. 另外,以下实验中 LS 局部搜索策略参数  $\rho$  使用原文推荐值 0.15.

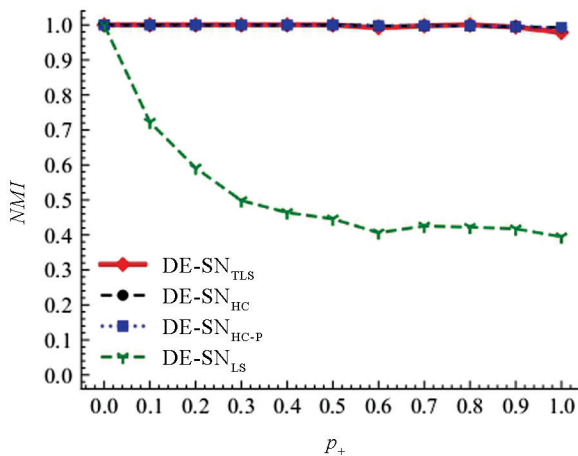
### 3.2 模型网络性能测试

首先,为检验社区之间的正边对 DE-SN 算法检测性能的影响,在  $SG(4, 32, 32, 0.5, 0, p_+)$  网络集合上,对 DE-SN 算法与各对比算法进行性能测试. TLS 算法的阈值参数  $\rho$  取值为 0 进行测试. 所有算法在模型网络上的实验结果如图 8a 所示,图中数据点代表各算法分别在 11 个  $SG(4, 32, 32, 0.5, 0, p_+)$  网络上独立运行 20 次得出社区划分  $NMI$  值的平均值。

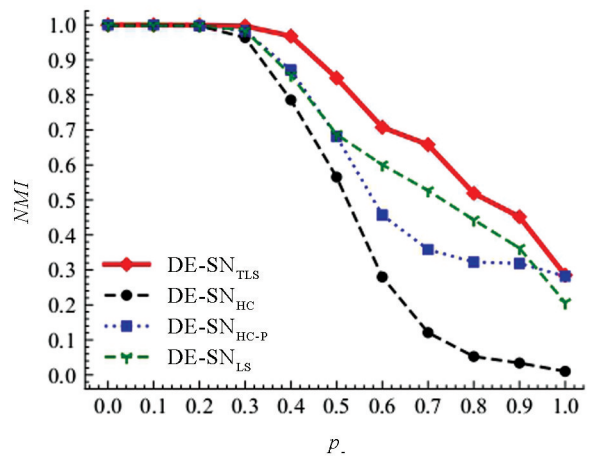
从图 8a 所示结果中可知,当  $p_+$  取值为  $[0, 0.5]$  时,  $DE-SN_{TLS}$ ,  $DE-SN_{HC}$  和  $DE-SN_{HC-P}$  算法在  $SG(4, 32, 32, 0.5, 0, p_+)$  网络上均能够稳定地检测得到全局最优社区划分,对应  $NMI$  平均值均达到 1.0. 当  $p_+$  取值为  $[0.6, 1]$  时,  $DE-SN_{HC}$  和  $DE-SN_{HC-P}$  算法  $NMI$  平均值接近 1.0.  $p_+$  为 0.8 时,  $DE-SN_{TLS}$  的  $NMI$  值为 1. 然而  $DE-SN_{LS}$  算法的性能表现相对不佳,当  $p_+ \leq 0.3$  时,  $DE-SN_{LS}$  算法的性能逐渐下降. 当  $p_+ > 0.3$  时,性能趋于相对平衡.  $DE-SN_{TLS}$  算法在  $SG(4, 32, 32, 0.5, 0, p_+)$  网络上的检测精度显著优于  $DE-SN_{LS}$  算法,且与  $DE-SN_{HC}$  和  $DE-SN_{HC-P}$  算法检测精度大致相同。

其次,为检验社区内部的负边对 DE-SN 算法检测性能的影响,在  $SG(4, 32, 32, 0.5, p_-, 0)$  网络集合上,同样进行各 DE-SN 算法的测试及对比分析. TLS 算法的阈值参数  $\rho$  取值为 0.15 进行测试. 图 8b 中数据点代表各算法分别在 11 个  $SG(4, 32, 32, 0.5, p_-, 0)$  网络上独立运行 20 次所得社区划分  $NMI$  值的平均值。

从图 8b 所示结果可以看出,  $DE-SN_{TLS}$  在  $SG(4, 32, 32, 0.5, p_-, 0)$  网络上能够获得相对最优的  $NMI$ , 与其他对比算法相比具有较为明显的性能优势.  $DE-SN_{HC}$  在  $p_- > 0.2$  时,  $NMI$  值下降趋势最为明显,且逐渐趋于 0.  $DE-SN_{TLS}$ ,  $DE-SN_{HC-P}$  和  $DE-SN_{LS}$  在  $p_- > 0.3$  时,所检测得  $NMI$  值开始逐渐下降,而  $DE-SN_{TLS}$  整体检测性能下降最为缓慢,对于任意  $p_-$  的取值,都能够优于其他算法。



a.  $SG(4, 32, 32, 0.5, 0, p_+)$



b.  $SG(4, 32, 32, 0.5, p_-, 0)$

图 8 模型网络上 4 种 DE-SN 算法的  $NMI$  检测结果

最后, 为检验社区之间的正边和社区内部的负边同时存在对 DE-SN 算法检测性能的影响. 在  $SG(4, 32, 32, 0.5, p_-, p_+)$  网络集合上, 进行各 DE-SN 算法的测试及对比分析. 此时, TLS 算法的阈值参数  $\rho$  取值为 0 进行测试. 各图中数据点代表各算法在 36 个网络上独立运行 20 次所得社区划分  $NMI$  值的平均值.

如图 9 所示, 每张热力子图的每一小块表示对应  $p_-$  和  $p_+$  取值, 算法 20 次检测得到  $NMI$  均值. DE-SN<sub>TLS</sub> 与 DE-SN<sub>HC</sub> 和 DE-SN<sub>HC-P</sub> 算法检测结果十分相近, 但仔细观察可以发现, 在  $(p_-, p_+) = (0.4, 0)$ ,  $(p_-, p_+) = (0.4, 0.1)$ ,  $(p_-, p_+) = (0.5, 0)$  和  $(p_-, p_+) = (0.5, 0.1)$  4 个块的检测精确度要显著高于 DE-SN<sub>HC</sub> 和 DE-SN<sub>HC-P</sub> 算法检测精确度. 然而, DE-SN<sub>LS</sub> 的检测结果相对较差, 仅在  $(p_-, p_+) = (0, 0)$  和  $(p_-, p_+) = (0.1, 0)$  时, 检测得  $NMI$  值为 1.0, 在  $(p_-, p_+) = (0.2, 0)$  和  $(p_-, p_+) = (0.3, 0)$  时, 检测得  $NMI$  值接近 1.

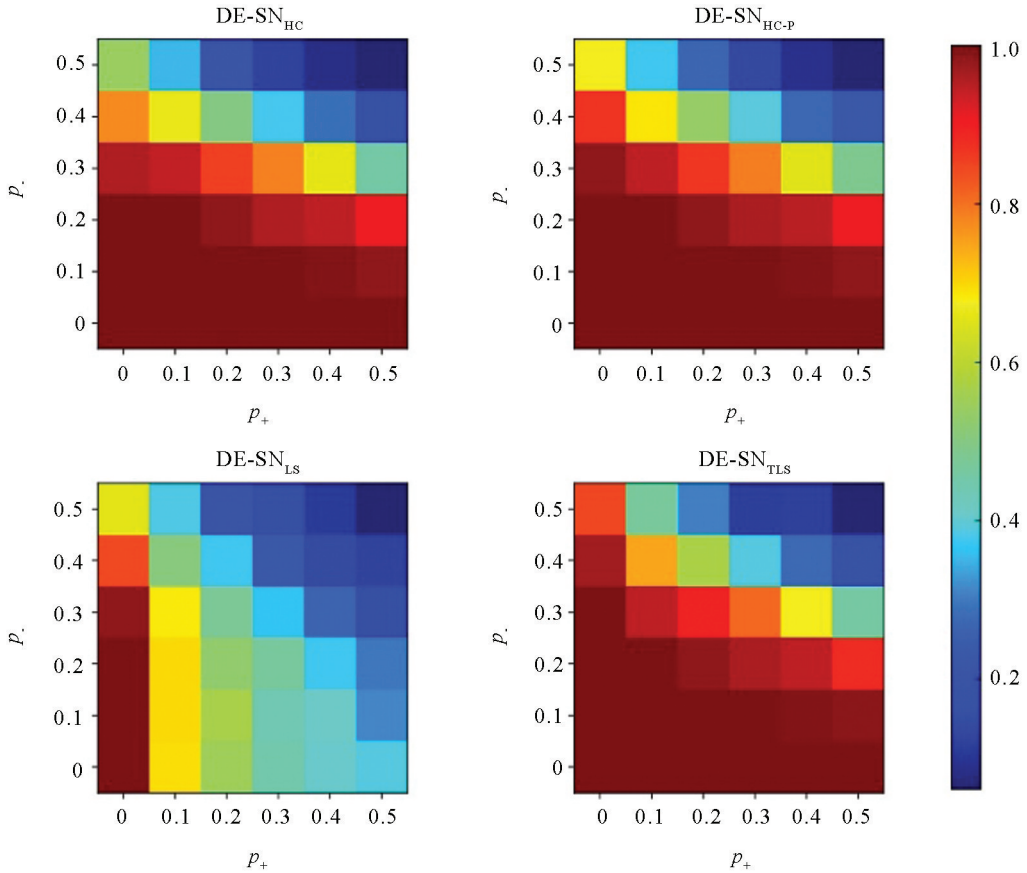


图 9  $SG(4, 32, 32, 0.5, p_-, p_+)$  网络上 4 种 DE-SN 算法的  $NMI$  检测结果

以上模型网络的实验表明, TLS 算法相较于其他代表性局部搜索算法, 在精确性和稳定性上能够表现出一定优势.

### 3.3 实证网络性能测试

将 DE-SN 算法在表 1 所示 2 个具有真实社区划分的符号社交网络和 4 个没有真实社区划分的符号生物网络上进行测试, 并与算法进行对比, 社交网络和生物网络实验结果分别如表 2 和表 3 所示. 此时, TLS 算法的阈值参数  $\rho$  取值为 0, 同时为确保生物网络检测尽可能达到收敛, 迭代次数  $t_{\max}$  取值为 200. 表 3 中数据为各算法在每个网络上独立运行 20 次, 所得社区划分对应符号模块度  $Q_S$  的平均值和  $NMI$  的平均值.

表 2 符号社交网络上各 DE-SN 算法检测结果

网络	评价指标	DE-SN <sub>HC</sub>	DE-SN <sub>HC-P</sub>	DE-SN <sub>LS</sub>	DE-SN <sub>TLS</sub>
SPP	$Q_{S_{avg}}$	0.454 7	0.454 7	0.454 7	0.454 7
	$NMI_{avg}$	1	1	1	1
GGS	$Q_{S_{avg}}$	0.431 0	0.431 0	0.431 0	0.431 0
	$NMI_{avg}$	1	1	1	1

表 3 中的实验结果表明, DE-SN<sub>TLS</sub> 能够在社交网络上达到最优社区划分的同时, 能在所有生物网络上获得比对比算法更优的  $Q_S$  平均值. TLS 算法相比较于其他典型代表性的局部搜索算法, 具有较好的质量和稳定性.

表 3 符号生物网络上各 DE-SN 算法检测结果

网络	评价指标	DE-SN <sub>HC</sub>	DE-SN <sub>HC-P</sub>	DE-SN <sub>LS</sub>	DE-SN <sub>TLS</sub>
EGFR	$Q_{S_{avg}}$	0.294 5	0.294 8	0.280 2	<b>0.303 6</b>
	$Q_{S_{max}}$	0.300 9	0.302 1	0.284 7	<b>0.309 2</b>
Yeast	$Q_{S_{avg}}$	0.558 2	0.558 2	0.565 1	<b>0.598 2</b>
	$Q_{S_{max}}$	0.562 9	0.560 5	0.567 4	<b>0.601 8</b>
Macrophage	$Q_{S_{avg}}$	0.327 6	0.327 8	0.315 8	<b>0.333 8</b>
	$Q_{S_{max}}$	0.329 9	0.330 4	0.322 7	<b>0.336 7</b>
Ecoli	$Q_{S_{avg}}$	0.391 1	0.385 9	0.339 5	<b>0.398 6</b>
	$Q_{S_{max}}$	0.394 3	0.390 3	0.342 4	<b>0.401 9</b>

## 4 通用性和鲁棒性检验

为深入验证 TLS 算法可提升符号网络社区检测的局部搜索性能, 并检验 TLS 算法在不同优化策略的模因算法上的通用性. 基于模因算法框架, 选择共生生物搜索和遗传算法分别作为模因算法中的元启发式算法进行全局搜索, TLS 算法作为局部搜索策略, 构建 SOS-SN 算法和 GA-SN 算法 2 个新的模因算法. 同上一节中的对比算法、参数设置和测试数据集对各算法进行实验.

### 4.1 SOS-SN 算法检验分析

SOS-SN 算法在模型网络上的实验结果如图 10a 和 10b 所示. 从图 10a 的实验结果发现, 对于任意  $p_+$  的取值, SOS-SN<sub>TLS</sub> 与 SOS-SN<sub>HC</sub> 和 SOS-SN<sub>HC-P</sub> 大致相同, 当  $p_+$  取值为  $[0, 0.5]$  时, 3 种算法检测得  $NMI$  均值都为 1.0, 当  $p_+$  取值为  $[0.5, 1]$  时, 3 种算法取得  $NMI$  均值近似 1.0. 而 SOS-SN<sub>LS</sub> 的检测性能依旧随着  $p_+$  的增大而整体呈现下降趋势. 从图 10b 的实验结果发现, 对于不同  $p_-$  取值, SOS-SN<sub>TLS</sub> 依旧能够保持相对最优的检测精度.

如图 11 检验社区之间的正边和社区内部的负边同时存在对 SOS-SN 算法检测性能的影响, SOS-SN<sub>TLS</sub> 依旧能够明显的优于 SOS-SN<sub>LS</sub> 检测结果, 并且在  $(p_-, p_+) = (0.4, 0)$ ,  $(p_-, p_+) = (0.4, 0.1)$ ,  $(p_-, p_+) = (0.5, 0)$  和  $(p_-, p_+) = (0.5, 0.1)$  4 个块的检测精确度依旧高于 SOS-SN<sub>HC</sub> 和 SOS-SN<sub>HC-P</sub> 法检测精确度.

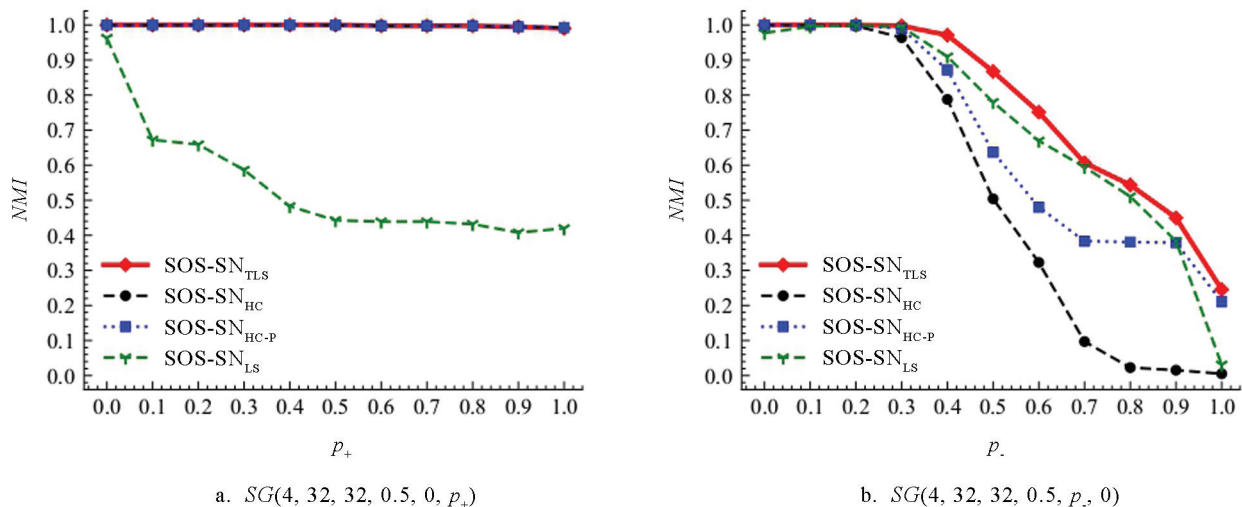


图 10 模型网络上 4 种 SOS-SN 算法的 NMI 检测结果

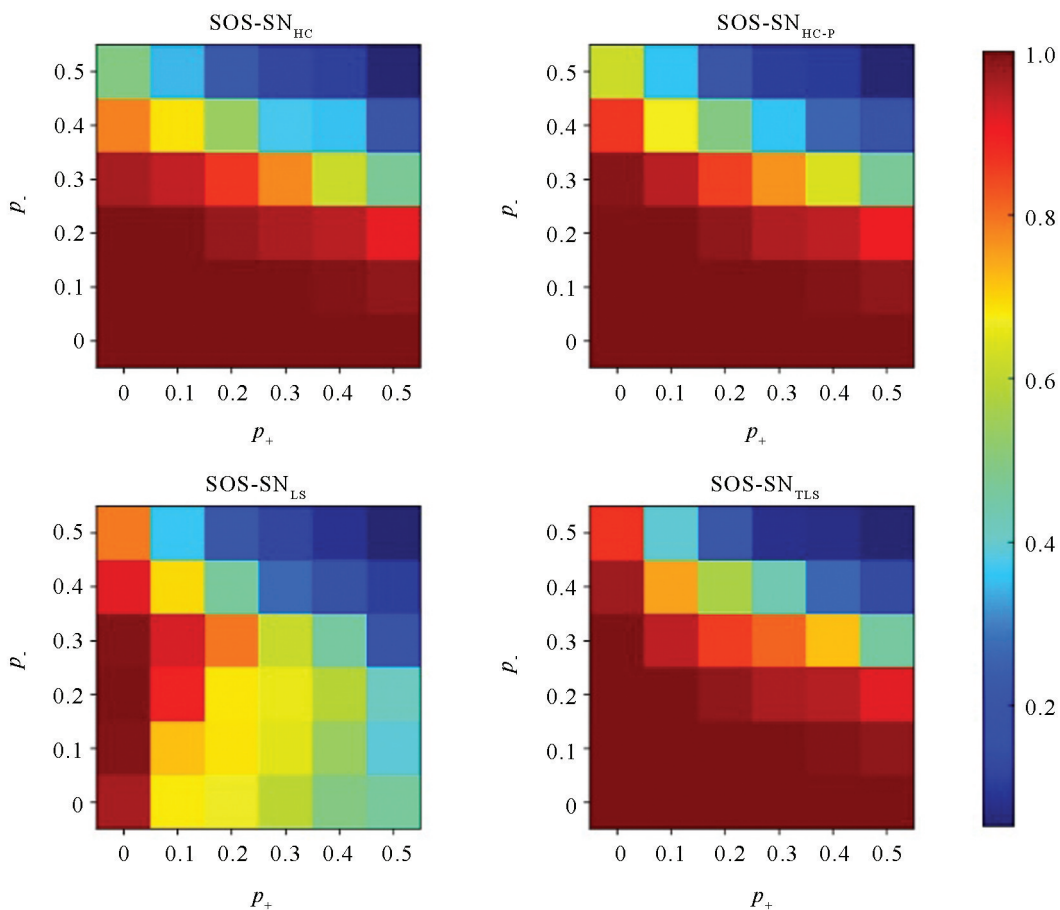


图 11  $SG(4, 32, 32, 0.5, p_-, p_+)$  网络上 4 种 SOS-SN 算法的 NMI 检测结果

在实证网络上对 SOS-SN 算法进行性能检验, 由于 SPP 网络和 GGS 网络相对简单, 所有算法都能获得最优社区划分, 故不再展示. 如表 4 所示为各个 SOS-SN 算法在符号生物网络上的检测结果, TLS 算法在生物网络上与对比算法相比依旧表现出明显的优势.



表 4 符号生物网络上各 SOS-SN 算法检测结果

网络	评价指标	SOS-SN <sub>HC</sub>	SOS-SN <sub>HC-P</sub>	SOS-SN <sub>LS</sub>	SOS-SN <sub>TLS</sub>
EGFR	$Q_{Savg}$	0.296 8	0.294 5	0.310 3	<b>0.327 5</b>
	$Q_{Smax}$	0.301 9	0.301 4	0.318 4	<b>0.334 3</b>
Yeast	$Q_{Savg}$	0.542 3	0.535 2	0.581 0	<b>0.638 7</b>
	$Q_{Smax}$	0.573 1	0.548 1	0.587 7	<b>0.642 7</b>
Macrophage	$Q_{Savg}$	0.325 2	0.321 8	0.361 7	<b>0.362 5</b>
	$Q_{Smax}$	0.327 9	0.325 4	0.365 6	<b>0.367 3</b>
Ecoli	$Q_{Savg}$	0.386 5	0.379 7	0.325 8	<b>0.417 2</b>
	$Q_{Smax}$	0.390 3	0.383 8	0.332 4	<b>0.419 4</b>

#### 4.2 GA-SN 算法检验分析

GA-SN 算法在模型网络上的实验结果如图 12a 和 12b 所示. 从图 12a 的实验结果发现, GA-SN<sub>TLS</sub>, GA-SN<sub>HC</sub> 和 GA-SN<sub>HC-P</sub> 在 GA-SN 算法上的实验结果十分相近. 而 GA-SN<sub>LS</sub> 的检测性能依旧较差, 当  $p_+$  取值为  $[0, 0.5]$  时, 检测精度缓慢下降, 但当  $p_+ > 0.5$  时, GA-SN<sub>LS</sub> 的检测性能下降明显. 从图 12b 的实验结果发现, GA-SN<sub>TLS</sub> 的检测性能依旧能够保持相对最优. 如图 13 为 4 种算法的热力图所示, GA-SN<sub>TLS</sub> 依旧能够在大多数的区块上优于其他算法的检测结果. 同样在符号生物网络上对 GA-SN 算法进行性能检验. 如表 5 符号生物网络上, 除了在 Macrophage 网络上, GA-SN<sub>LS</sub> 的质量性检测结果要优于 GA-SN<sub>TLS</sub>. 在其他生物网络上, GA-SN<sub>TLS</sub> 的检测结果都要优于其他对比算法.

根据以上所有实验结果表明, 在 SOS-SN 算法和 GA-SN 算法中, TLS 算法辅助全局搜索算法的社区检测性能依旧普遍优于其他对比局部搜索策略. 本研究提出的 TLS 算法可以与不同全局搜索算法相结合, 并辅助算法获得更优的社区划分结果, 从而验证了 TLS 算法拥有较好的通用性和鲁棒性.

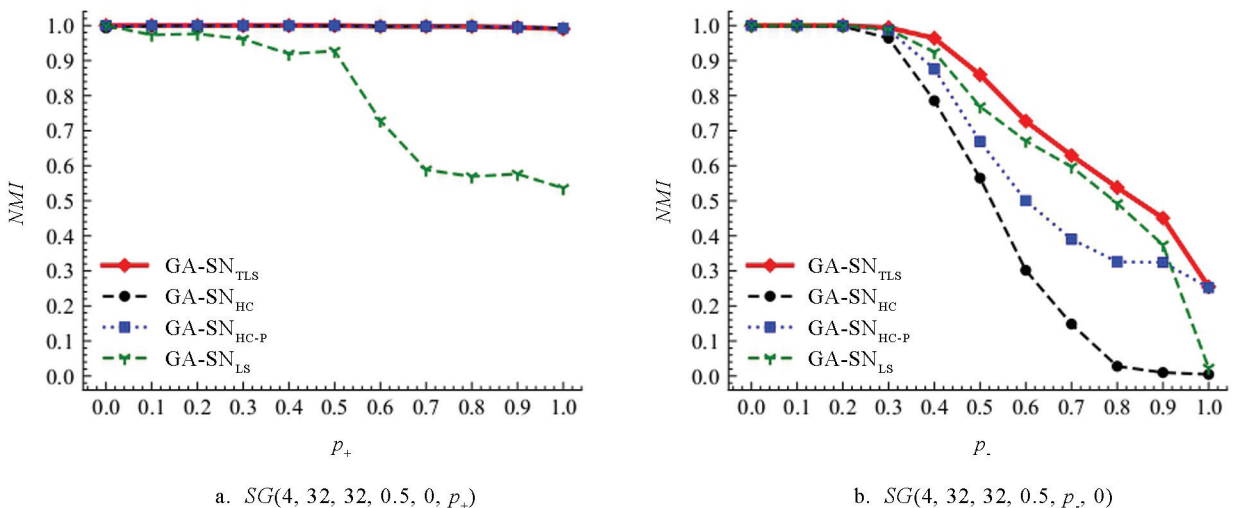


图 12 模型网络上 4 种 GA-SN 算法的 NMI 检测结果

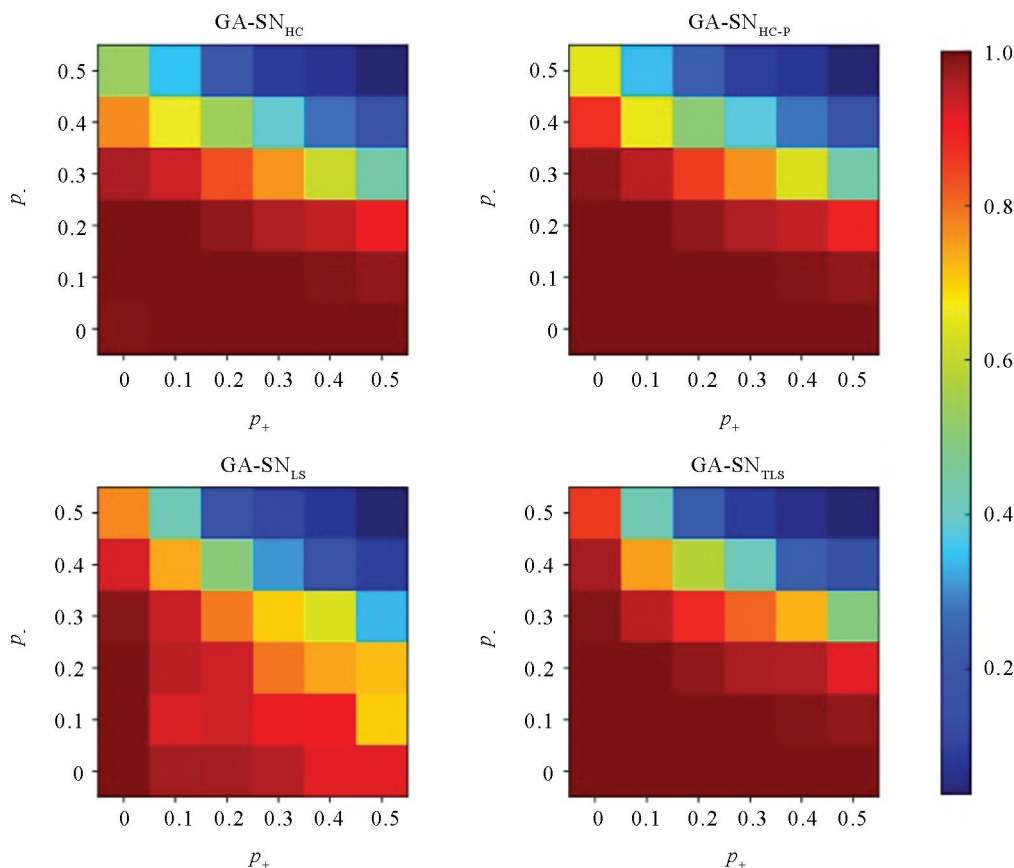
图 13  $SG(4, 32, 32, 0.5, p_-, p_+)$ 网络上 4 种 GA-SN 算法的 NMI 检测结果

表 5 符号生物网络上各 GA-SN 算法检测结果

网络	评价指标	GA-SN <sub>HC</sub>	GA-SN <sub>HC-P</sub>	GA-SN <sub>LS</sub>	GA-SN <sub>TLS</sub>
EGFR	$Q_{Savg}$	0.300 7	0.296 4	0.319 4	<b>0.329 3</b>
	$Q_{Smax}$	0.310 4	0.304 3	0.330 5	<b>0.334 4</b>
Yeast	$Q_{Savg}$	0.542 5	0.536 4	0.598 6	<b>0.638 2</b>
	$Q_{Smax}$	0.567 0	0.554 8	0.607 3	<b>0.643 2</b>
Macrophage	$Q_{Savg}$	0.323 0	0.323 8	<b>0.370 2</b>	0.357 8
	$Q_{Smax}$	0.327 8	0.325 7	<b>0.374 7</b>	0.360 9
Ecoli	$Q_{Savg}$	0.378 6	0.374 8	0.319 4	<b>0.417 4</b>
	$Q_{Smax}$	0.390 1	0.380 2	0.330 5	<b>0.421 6</b>

## 5 结论

本研究提出了一种基于符号模体的三元组局部搜索算法 TLS. 该算法采用了一种基于符号模体进行社区迁移的新方法, 将传统社区编号在二元组之间的迁移扩展到了三元组, 综合利用了符号网络低阶和高阶拓扑结构信息, 进一步优化节点的结构平衡性, 提升算法收敛速度和检测性能. 为验证提出的局部搜索算法社区检测性能, 将 TLS 算法同差分进化策略构建成模因算法 DE-SN, 在模型网络 and 不同规模及特性的实证网络上进行了检验, 并分别将爬山策略 HC、改进爬山策略 HC-P 和 MACD-SN 算法中使用的局部搜

索方法 LS 作为对比局部策略进行对比分析。实验结果表明, TLS 算法能够在模型网络上获得更加精确和稳定的检测结果, 并且在实证网络上同样表现出相对较好的质量性和稳定性。此外, 通过将 TLS 算法迁移至元启发式优化算法为共生生物搜索算法和遗传算法的模因算法中, 分别构建了 SOS-SN 算法和 GA-SN 算法, 在模型网络和实证网络上均验证了 TLS 算法的通用性。

本研究提升了现有符号网络社区检测的局部搜索策略性能, 拓展了模体理论的应用场景, 有助于加深对网络结构和功能特性的理解。未来研究中将扩展到有向符号网络和加权符号网络等, 使提出的局部搜索算法能够适应更多不同类型复杂网络应用需求。

### 参考文献:

- [1] HEIDER F. Attitudes and Cognitive Organization [J]. *The Journal of Psychology*, 1946, 21(1): 107-112.
- [2] NEWMAN M E J. Detecting Community Structure in Networks [J]. *The European Physical Journal B*, 2004, 38(2): 321-330.
- [3] XIA Z Y, BU Z. Community Detection Based on a Semantic Network [J]. *Knowledge-Based Systems*, 2012, 26: 30-39.
- [4] PARISIEN C, ANDERSON C H, ELIASMITH C. Solving the Problem of Negative Synaptic Weights in Cortical Models [J]. *Neural Computation*, 2008, 20(6): 1473-1494.
- [5] CHEN Y R, KANG Q M, DUAN W Q, et al. An Iterated Local Search Algorithm for Community Detection in Signed Networks [J]. *International Journal of Modern Physics C*, 2022, 33(8): 2250105.
- [6] AREF S, NEAL Z. Detecting Coalitions by Optimally Partitioning Signed Networks of Political Collaboration [J]. *Scientific Reports*, 2020, 10(1): 1506.
- [7] KUNEGIS J, SCHMIDT S, LOMMATZSCH A, et al. Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization [C] // *Proceedings of the 2010 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2010: 559-570.
- [8] CHIANG K Y, HSIEH C J, NATARAJAN N, et al. Prediction and Clustering in Signed Networks: A Local to Global Perspective [J]. *The Journal of Machine Learning Research*, 2014, 15(1): 1177-1213.
- [9] ZHAO X H, YANG B, LIU X Y, et al. Statistical Inference for Community Detection in Signed Networks [J]. *Physical Review E*, 2017, 95(4): 042313.
- [10] LI Y D, LIU J, LIU C L. A Comparative Analysis of Evolutionary and Memetic Algorithms for Community Detection from Signed Social Networks [J]. *Soft Computing*, 2014, 18(2): 329-348.
- [11] MOSCATO P. On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts-Towards Memetic Algorithms [J]. *Caltech Concurrent Computation Program*, 1989, 826(10): 79-158.
- [12] AMELIO A, PIZZUTI C. An Evolutionary and Local Refinement Approach for Community Detection in Signed Networks [J]. *International Journal on Artificial Intelligence Tools*, 2016, 25(4): 1650021.
- [13] CHE S W, YANG W, WANG W. A Memetic Algorithm for Community Detection in Signed Networks [J]. *IEEE Access*, 2020, 8: 123585-123602.
- [14] BENSON A R, GLEICH D F, LESKOVEC J. Higher-Order Organization of Complex Networks [J]. *Science*, 2016, 353(6295): 163-166.
- [15] MILO R, SHEN-ORR S, ITZKOVITZ S, et al. Network Motifs: Simple Building Blocks of Complex Networks [J]. *Science*, 2002, 298(5594): 824-827.
- [16] HUANG J Y, HOU Y N, LI Y S. Efficient Community Detection Algorithm Based on Higher-Order Structures in Complex Networks [J]. *Chaos: an Interdisciplinary Journal of Nonlinear Science*, 2020, 30(2): 023114.
- [17] YANG B, CHEUNG W, LIU J M. Community Mining from Signed Social Networks [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2007, 19(10): 1333-1348.

- [18] NEWMAN M E J. Analysis of Weighted Networks [J]. *Physical Review E*, 2004, 70(5): 056131.
- [19] GÓMEZ S, JENSEN P, ARENAS A. Analysis of Community Structure in Networks of Correlated Data [J]. *Physical Review E*, 2009, 80(1): 016114.
- [20] LANCICHINETTI A, FORTUNATO S, RADICCHI F. Benchmark Graphs for Testing Community Detection Algorithms [J]. *Physical Review E*, 2008, 78(4): 046110.
- [21] DOREIAN P, MRVAR A. A Partitioning Approach to Structural Balance [J]. *Social Networks*, 1996, 18(2): 149-168.
- [22] HARARY F. On the Measurement of Structural Balance [J]. *Behavioral Science*, 1959, 4(4): 316-323.
- [23] FERLIGOJ A, KRAMBERGER A. An Analysis of the Slovene Parliamentary Parties Network [J]. *Developments in Statistics and Methodology*, 1996, 12: 209-216.
- [24] READ K E. Cultures of the Central Highlands, New Guinea [J]. *Southwestern Journal of Anthropology*, 1954, 10(1): 1-43.
- [25] ODA K, MATSUOKA Y, FUNAHASHI A, et al. A Comprehensive Pathway Map of Epidermal Growth Factor Receptor Signaling [J]. *Molecular Systems Biology*, 2005, 1(1): 1-17.
- [26] ODA K, KIMURA T, MATSUOKA Y, Et Al. Molecular Interaction Map of a Macrophage [J]. *AfCS Research Reports*, 2004, 2(14): 1-12.
- [27] SHEN-ORR S S, MILO R, MANGAN S, et al. Network Motifs in the Transcriptional Regulation Network of *Escherichia Coli* [J]. *Nature Genetics*, 2002, 31(1): 64-68.
- [28] LESKOVEC J, KREVL A. SNAP Datasets; Stanford Large Network Dataset Collection [EB/OL]. <http://snap.stanford.edu/data>.
- [29] JIA G B, CAI Z X, MUSOLESI M, et al. Community Detection in Social and Biological Networks Using Differential Evolution [C] //International Conference on Learning and Intelligent Optimization. Berlin, Heidelberg: Springer, 2012: 71-85.
- [30] XIAO J, WANG C, XU X K. Community Detection Based on Symbiotic Organisms Search and Neighborhood Information [J]. *IEEE Transactions on Computational Social Systems*, 2019, 6(6): 1257-1272.
- [31] PIZZUTI C. GA-Net: A Genetic Algorithm for Community Detection in Social Networks [C] //Proceedings of the 10th International Conference on Parallel Problem Solving from Nature; PPSN X, 2008: 1081-1090.
- [32] ZHANG C, HEI X H, YANG D D, et al. A Memetic Particle Swarm Optimization Algorithm for Community Detection in Complex Networks [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2016, 30(2): 1659003.
- [33] GONG M G, FU B, JIAO L C, et al. Memetic Algorithm for Community Detection in Networks [J]. *Physical Review E*, 2011, 84(5): 056101.