











按照惯例, 针对人工网络, 已知其真实社区结构, 实验中仅比较各算法的标准互信息值; 对于真实的网络, 我们比较不同算法的模块度值<sup>[34]</sup>.

标准互信息 (Normalized Mutual Information, *NMI*)<sup>[35]</sup>常用来评价算法的划分结果与网络真实划分结果的吻合程度. 对于网络的两种划分 *A* 和 *B*, 它们之间的互信息定义如下:

$$NMI = \frac{-2 \sum_{i=1}^{C_A} \sum_{j=1}^{C_B} C_{ij} \log(C_{ij}N/C_i C_j)}{\sum_{i=1}^{C_A} C_i \log(C_i/N) + \sum_{j=1}^{C_B} C_j \log(C_j/N)} \quad (8)$$

其中, *N* 表示节点个数. *C<sub>A</sub>* 表示划分 *A* 中网络的社区数目, *C<sub>B</sub>* 表示划分 *B* 中网络的社区数目, *C<sub>i</sub>* 表示第 *i* 行元素和, *C<sub>j</sub>* 表示第 *j* 列元素之和, *C<sub>i,j</sub>* 表示划分 *A* 中的社区 *i* 和划分 *B* 中的社区 *j* 两者之间的交集含有的节点数目. 如果 *A* = *B*, 则 *NMI* 值为 1, 如果 *A* 划分和 *B* 划分完全不同, 则 *NMI* 值为 0.

### 3.1 人工网络比较实验

为了分析 DJaya 算法的优化性能, 我们在扩展 GN Benchmark 上进行了实验分析. 该人工网络的社区结构已知, 它含有 128 个节点, 共划分为 4 个不同的社区, 每个社区有 32 个节点, 社区中每个节点的平均度数为 16, 如表 2. 网络中的混合参数 *u* 主要用于确定某一社区中任意一个节点与其他社区的节点之间共享边的数量, 当混 *u* < 0.5 时, 网络中的社区结构较为明显, 但随着 *u* 的增大, 网络中的社区结构将变得越来越模糊, 性能一般的算法难以检测出网络中存在的社区结构, 从而可以区分不同算法的性能, 比较结果如图 4.

表 2 人工网络的参数 (GN benchmark)

节点数	128
边数	2048
平均节点数	16

由图 4 可见, 当 *u* < 0.35 时候, 各个算法的 *NMI* 值达到了 1, 都能准确找出社区结构; 当 *u* = 0.4, BGLL、ACO 算法的 *NMI* 开始下降; 当 *u* = 0.45, LGA、IDDE、DJaya 算法能找到和真实社区完全一样的社区结构, 相比其它算法, 它们较为稳定; 当 *u* = 0.5, 此时网络的社区结构已经非常模糊, DJaya 算法仍能准确划分接近 80% 的社区结构, 表明 DJaya 算法的一定优越性, 总体来说 DJaya 算法在人工网络上优于其它算法.

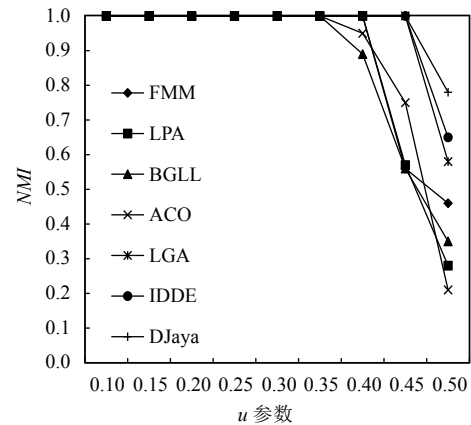


图 4 各种算法划分精度的比较

### 3.2 真实网络比较实验

真实的复杂网络与人工网络相比, 有不同的拓扑属性. 这里采用了 4 个被广泛使用的真实网络作为实验数据集, 其点数和边数见表 3. Karate<sup>[36]</sup>网络是通过对一个美国大学空手道俱乐部进行观测而构建出的一个社会网络. Dolphins<sup>[37]</sup>是由 Lusseau 等使用长达 7 年的时间观察新西兰 Doubtful Sound 海峡 62 只海豚群体的交流情况而得到的海豚社会关系网络. Polbooks<sup>[6]</sup>是由 Amazon 上销售的美国政治相关书籍页面上建立起来的网络. Football<sup>[1]</sup>网络是根据美国大学生足球联赛而创建的一个复杂的社会网络.

表 3 4 个真实网络数据的参数

网络	节点数	边数
Karate	34	78
Dolphins	62	159
Polbooks	105	441
Football	115	613

表 3 中, football 真实社区的划分效果与 DJaya 算法划分的效果图分别如图 5 和图 6.

如图 5, football 对应的真实社区有 12 个, 图 6 采用 DJaya 算法划分的社区则有 10 个, 可以看出 DJaya 算法合并了一些较小的社区, 保持了大部分社区划分是和真实社区一致, 文献[2]认为这种划分是合理的.

表 4 给出了各种算法在 4 个标准真实网络算例上所得的模块度值. 对于算例 Karate, CC-GA、LGA、DJaya 都达到 0.4198 最优值<sup>[38]</sup>; 在算例 Dolphins 上 DJaya 算法取得了最高的模块度值; 在算例 polbooks

上, DJaya 和 LGA 都达到了该数据集的理论最高值 0.5272<sup>[39]</sup>; 在算例 football上, LGA、IDDE、DJaya 都表现良好, 达到了最优值<sup>[39]</sup>。综上所述, 在 4 个算例上, DJaya 算法展现出较明显优势。

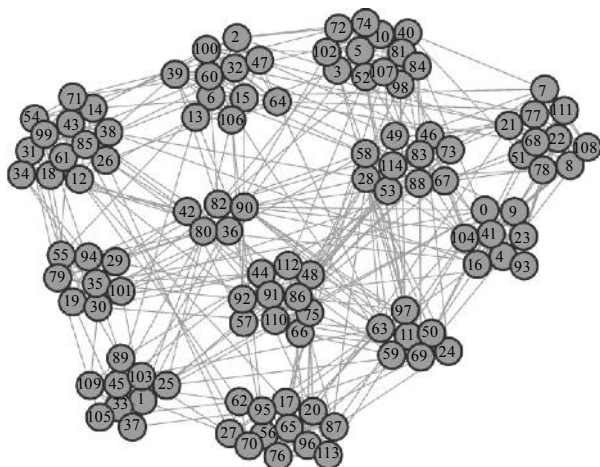


图 5 真实的 football 的社区 (12 个社区)

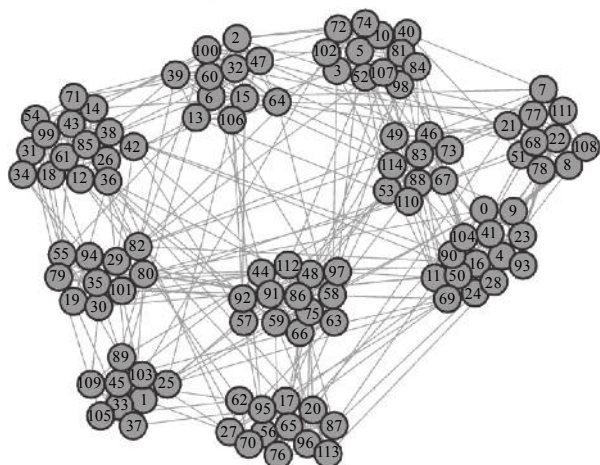


图 6 DJaya 划分的 football 社区 (10 个社区)

#### 4 总结

本文提出的 DJaya 算法用于求解复杂网络社区发现问题, 该算法采用局部邻接表示法的编码方式对种群进行社区表示, 并用标签传播思想进行初始化, 然后借鉴连续空间 Jaya 算法的靠近最优解、远离最差解的基本思想, 先判断种群个体的倾向性, 针对靠近最差解采用标签传播思想更新种群, 靠近最优解采用局部模块度函数更新种群, 最后采用贪心选择策略生成下一代种群。DJaya 算法实现简单, 容易理解, 具有全局搜索能力和局部开发能力。通过在真实网络和人工网络上

的实验对比发现, 该算法具有较好的优化能力和鲁棒性, 达到了较高的社区发现精度且自动确定社区个数。该算法适用于非重叠、静态社区, 考虑到模块度作为优化函数的局限性<sup>[39]</sup>, 今后将改进算法的目标函数, 或建立多目标优化函数, 同时考虑扩展算法的适应范围, 增大数据规模, 应用在重叠、动态社区发现问题。

表 4 各种算法在不同网络中的模块度对比

类型	算法	Karate	Dolphins	Polbooks	Football
贪心启发式算法	FMM	0.4013	0.5194	0.5168	0.5996
	LPA	0.3536	0.4860	0.5039	0.5881
	BGLL	0.4005	0.5164	0.5059	0.6018
	DJaya	<b>0.4198</b>	<b>0.5283</b>	<b>0.5272</b>	<b>0.6046</b>
智能元启发式算法	ACO	0.4198	0.5271	0.5271	0.6045
	IDDE	0.4198	0.5282	0.5272	<b>0.6046</b>
	CC-GA	<b>0.4198</b>	0.5263	0.5252	0.5632
	LGA	<b>0.4198</b>	0.5280	<b>0.5272</b>	<b>0.6046</b>
	DJaya	<b>0.4198</b>	<b>0.5283</b>	<b>0.5272</b>	<b>0.6046</b>

#### 参考文献

- Girvan M, Newman MEJ. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99(12): 7821–7826. [doi: 10.1073/pnas.122653799]
- Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Physical Review E*, 2004, 69(2): 026113. [doi: 10.1103/PhysRevE.69.026113]
- Kernighan BW, Lin S. An efficient heuristic procedure for partitioning graphs. *The Bell System Technical Journal*, 1970, 49(2): 291–307. [doi: 10.1002/j.1538-7305.1970.tb01770.x]
- Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Physical Review E*, 2004, 70(6): 066111. [doi: 10.1103/PhysRevE.70.066111]
- Newman MEJ. Fast algorithm for detecting community structure in networks. *Physical Review E*, 2004, 69(5): 066133.
- Newman MEJ. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(23): 8577–8582. [doi: 10.1073/pnas.0601602103]
- Xing Y, Meng FR, Zhou Y, et al. A node influence based

- label propagation algorithm for community detection in networks. *The Scientific World Journal*, 2014, 2014: 627581.
- 8 Van Dongen S. Graph clustering by flow simulation [Ph. D. thesis]. Utrecht: University of Utrecht, 2000. 1–173.
- 9 Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, 2008, 105(4): 1118–1123. [doi: [10.1073/pnas.0706851105](https://doi.org/10.1073/pnas.0706851105)]
- 10 Jia GB, Cai ZX, Musolesi M, *et al.* Community detection in social and biological networks using differential evolution. *Proceedings of the 6th International Conference on Learning and Intelligent Optimization*. Paris, France. 2012. 71–85.
- 11 Cai Q, Gong MG, Ma LJ, *et al.* Greedy discrete particle swarm optimization for large-scale social network clustering. *Information Sciences*, 2015, 316: 503–516. [doi: [10.1016/j.ins.2014.09.041](https://doi.org/10.1016/j.ins.2014.09.041)]
- 12 Tasgin M, Herdagdelen A, Bingol H. Community detection in complex networks using genetic algorithms. arXiv: 0711.0491, 2006.
- 13 金弟, 刘杰, 杨博, 等. 局部搜索与遗传算法结合的大规模复杂网络社区探测. *自动化学报*, 2011, 37(7): 873–882.
- 14 张英杰, 龚中汉, 陈乾坤. 基于免疫离散差分进化算法的复杂网络社区发现. *自动化学报*, 2015, 41(4): 749–757.
- 15 Duch J, Arenas A. Community detection in complex networks using extremal optimization. *Physical Review E*, 2005, 72(2): 027104. [doi: [10.1103/PhysRevE.72.027104](https://doi.org/10.1103/PhysRevE.72.027104)]
- 16 Guimerà R, Amaral LAN. Functional cartography of complex metabolic networks. *Nature*, 2005, 433(7028): 895–900. [doi: [10.1038/nature03288](https://doi.org/10.1038/nature03288)]
- 17 Blondel VD, Guillaume JL, Lambiotte R, *et al.* Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008, 2008(10): 10008. [doi: [10.1088/1742-5468/2008/10/P10008](https://doi.org/10.1088/1742-5468/2008/10/P10008)]
- 18 Pizzuti C. Evolutionary computation for community detection in networks: A review. *IEEE Transactions on Evolutionary Computation*, 2018, 22(3): 464–483. [doi: [10.1109/TEVC.2017.2737600](https://doi.org/10.1109/TEVC.2017.2737600)]
- 19 Cai Q, Ma LJ, Gong MG, *et al.* A survey on network community detection based on evolutionary computation. *International Journal of Bio-Inspired Computation*, 2016, 8(2): 84–98. [doi: [10.1504/IJBIC.2016.076329](https://doi.org/10.1504/IJBIC.2016.076329)]
- 20 Liu J, Abbass HA, Tan KC. Evolutionary community detection algorithms. In: Liu J, Abbass HA, Tan KC, eds. *Evolutionary Computation and Complex Networks*. Cham: Springer, 2019. 77–115.
- 21 Talbi EG. A parallel genetic algorithm for the graph partitioning problem. *Proceedings of the 5th International Conference on Supercomputing*. New York, NY, USA. 1991. 312–320.
- 22 Bui TN, Moon BR. Genetic algorithm and graph partitioning. *IEEE Transactions on Computers*, 1996, 45(7): 841–855. [doi: [10.1109/12.508322](https://doi.org/10.1109/12.508322)]
- 23 Tasgin M, Herdagdelen A, Bingol H. Community detection in complex networks using genetic algorithms. arXiv: 0711.0491, 2007.
- 24 Gog A, Dumitrescu D, Hirsbrunner B. Community detection in complex networks using collaborative evolutionary algorithms. *Proceedings of the 9th European Conference on Artificial Life*. Lisbon, Portugal. 2007. 886–894.
- 25 Pizzuti C. Ga-Net: A genetic algorithm for community detection in social networks. *Proceedings of the 10th International Conference on Parallel Problem Solving from Nature*. Dortmund, Germany. 2008. 1081–1090.
- 26 Gong MG, Fu B, Jiao LC, *et al.* Memetic algorithm for community detection in networks. *Physical Review E*, 2011, 84(5): 056101.
- 27 Chang HH, Feng ZR, Ren ZG. Community detection using ant colony optimization. *Proceedings of 2013 IEEE Congress on Evolutionary Computation*. Cancun, Mexico. 2013. 3072–3078.
- 28 Said A, Abbasi RA, Maqbool O, *et al.* CC-GA: A clustering coefficient based genetic algorithm for detecting communities in social networks. *Applied Soft Computing*, 2018, 63: 59–70. [doi: [10.1016/j.asoc.2017.11.014](https://doi.org/10.1016/j.asoc.2017.11.014)]
- 29 Rao RV. Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations*, 2016, 7(1): 19–34.
- 30 Gao KZ, Yang FJ, Zhou MC, *et al.* Flexible job-shop rescheduling for new job insertion by using discrete Jaya algorithm. *IEEE Transactions on Cybernetics*, 2019, 49(5): 1944–1955. [doi: [10.1109/TCYB.2018.2817240](https://doi.org/10.1109/TCYB.2018.2817240)]
- 31 Gao KZ, Zhang YC, Sadollah A, *et al.* Jaya, harmony search and water cycle algorithms for solving large-scale real-life urban traffic light scheduling problem. *Swarm and Evolutionary Computation*, 2017, 37: 58–72. [doi: [10.1016/j.swevo.2017.05.002](https://doi.org/10.1016/j.swevo.2017.05.002)]
- 32 Hruschka ER, Campello RJGB, Freitas AA. A survey of evolutionary algorithms for clustering. *IEEE Transactions on*



- Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2009, 39(2): 133–155. [doi: [10.1109/TSMCC.2008.2007252](https://doi.org/10.1109/TSMCC.2008.2007252)]
- 33 Lancichinetti A, Fortunato S, Radicchi F. Benchmark graphs for testing community detection algorithms. *Physical Review E*, 2008, 78(4): 046110. [doi: [10.1103/PhysRevE.78.046110](https://doi.org/10.1103/PhysRevE.78.046110)]
- 34 Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 2007, 76(3 Pt 2): 036106.
- 35 Danon L, Díaz-Guilera A, Duch J, *et al.* Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005, 2005(9): 09008. [doi: [10.1088/1742-5468/2005/09/P09008](https://doi.org/10.1088/1742-5468/2005/09/P09008)]
- 36 Zachary WW. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 1977, 33(4): 452–473. [doi: [10.1086/jar.33.4.3629752](https://doi.org/10.1086/jar.33.4.3629752)]
- 37 Lusseau D. The emergent properties of a dolphin social network. *Proceedings of the Royal Society B: Biological Sciences*, 2003, 270(S2): S186–S188.
- 38 Aloise D, Cafieri S, Caporossi G, *et al.* Column generation algorithms for exact modularity maximization in networks. *Physical Review E*, 2010, 82(4): 046112. [doi: [10.1103/PhysRevE.82.046112](https://doi.org/10.1103/PhysRevE.82.046112)]
- 39 Fortunato S, Barthelemy M. Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(1): 36–41. [doi: [10.1073/pnas.0605965104](https://doi.org/10.1073/pnas.0605965104)]