

数据集中,CDNEV 算法的表现无论在准确率、召回率单个指标上,还是在综合评价指标 $F1$ 上,均优于对比算法。我们在 4 个数据集上,比较了算法的平均指标。

Table 2 Real networks used in the experiment

表 2 实验中用到的真实网络

网络	缩写	$ V $	$ E $	标记网络
Zachary's karate club ^[42]	Karate	34	78	√
Dolphin social network ^[43]	Dolphin	62	159	√
American college football ^[8]	Football	115	616	√
Books about US politics ^[10]	Polbooks	105	441	√
Email communicatin network ^[44]	Email	1 133	5 451	×
com-DBLP ^[57]	DBLP	317 080	1 049 866	√

Table 3 Comparisons of CDNEV and baseline algorithms' community detection on real networks

表 3 CDNEV 算法与其他算法在真实网络中划分结果的比较

Data		GN	FG	IM	SG	WT	LV	LE	LPA	GCE	LFM	SK	DW	N2V	CDNEV
Karate	P	0.971	0.971	0.971	1.000	0.941	0.971	1.000	0.971	0.971	0.882	0.941	0.971	0.971	1.000
	R	0.516	0.722	0.761	0.563	0.411	0.545	0.533	0.761	0.972	0.908	0.956	0.972	0.972	1.000
	$F1$	0.674	0.828	0.853	0.720	0.572	0.698	0.695	0.853	0.972	0.895	0.948	0.972	0.972	1.000
Dolphins	P	0.984	0.935	0.968	0.968	0.952	0.935	0.952	0.968	0.871	0.935	0.935	0.645	0.968	0.984
	R	0.619	0.619	0.478	0.489	0.603	0.452	0.406	0.549	0.438	0.817	0.803	0.548	0.819	0.985
	$F1$	0.745	0.745	0.639	0.650	0.738	0.610	0.569	0.700	0.583	0.872	0.864	0.593	0.887	0.984
Football	P	0.835	0.574	0.930	0.878	0.870	0.870	0.626	0.800	0.852	0.713	0.922	0.270	0.852	0.922
	R	0.954	0.992	0.929	0.968	0.972	0.977	0.887	0.977	0.917	0.976	0.883	0.267	0.887	0.925
	$F1$	0.890	0.727	0.930	0.921	0.918	0.920	0.734	0.879	0.883	0.824	0.902	0.268	0.869	0.923
Polbooks	P	0.857	0.838	0.848	0.848	0.848	0.848	0.848	0.848	0.762	0.838	0.887	0.819	0.838	0.848
	R	0.763	0.769	0.596	0.557	0.761	0.627	0.599	0.763	0.746	0.793	0.879	0.776	0.746	0.944
	$F1$	0.807	0.802	0.700	0.672	0.802	0.721	0.702	0.803	0.754	0.815	0.883	0.797	0.789	0.893

$F1$ 指标值最低的是 DW 算法,为 0.657 5,最高的是 LFM 算法,为 0.851 5,CDNEV 算法平均 $F1$ 则为 0.95,明显高于其他算法。相对于其他算法,CDNEV 算法的 $F1$ 指标平均提高了 19%。

在一些特殊情况下,如 SG 算法和 LE 算法,在 Karate 数据集上的准确率也可以达到 100%,但这些算法是将一个大社团中的节点划分到多个子社团,这种“过度划分”导致它们的召回率很低,所以最终的 $F1$ 值较低。由第 3.5 节易知,由于 SG 算法是基于贪心迭代算法的一种改进,因此,SG 算法相较于 CDNEV 算法,其时间复杂度较高。CDNEV 算法与其他算法在 E-mail 网络上的模块度比较可见表 4。

Table 4 Comparisons of CDNEV and baseline algorithms' modularity on E-mail networks

表 4 CDNEV 算法与其他算法在 E-mail 网络上的模块度比较

Data	GN	FG	IM	SG	WT	LV	LE	LPA	GCE	LFM	SK	DW	N2V	CDNEV
M	0.532	0.506	0.52	0.579	0.53	0.54	0.488	0.532	0.436	0.127	0.527	0.513	0.548	0.553

DBLP 网络是标记网络,且网络规模较大,各算法在社团划分上的 ARI 指标见表 5,CDNEV 算法的 ARI 指标与 IM、N2V 等算法基本持平,高于其他算法。验证了 CDNEV 算法在较大规模网络中的可扩展性。

Table 5 Comparisons of CDNEV and baseline algorithms' ARI on DBLP networks

表 5 CDNEV 算法与其他算法在 DBLP 网络上的 ARI 指标比较

Data	GN	FG	IM	SG	WT	LV	LE	LPA	GCE	LFM	SK	DW	N2V	CDNEV
ARI	0.314	0.353	0.396	0.412	0.372	0.393	0.362	0.401	0.339	0.218	0.407	0.357	0.413	0.413

综合无标记和有标记网络上的实验结果可知,相较于其他算法,CDNEV 算法在真实网络中的划分效果优于其他算法。

4.3 模拟网络数据集实验

为了进一步评估这些算法的性能,我们采用 LFR^[58]人工合成网络标准网络来进行实验,LFR 网络中节点的度分布及社团的规模分布均为幂律分布,使其更接近真实网络。本次实验主要通过设置以下参数来生成所需的

模拟复杂网络.

模拟网络节点总数为 n ; 模拟网络的节点平均度为 $k(P)$; 模拟网络节点最大度为 $k_{\max}(P_{\max})$. 另外, 还需要通过实验分析、比较拓扑混合参数 μ 的作用, μ 表示模拟网络中社团内节点与社团外部节点连接的边数占节点总边数的比例, μ 越大, 说明网络结构越不明显; 具有指数分布形式的度分布的参数 ε_1 , 模拟网络节点度分布服从幂指数为 ε_1 的幂律分布; LFR 网络的社区规模服从指数分布, 其参数为 ε_2 ; 最小社区规模为 c_{\min} , 指定最小社区的节点数; 最大社区规模为 c_{\max} , 指定最大社区的节点数.

按照文献[59]中的实验设计建议, 对 LFR 基准网络的参数设置如下.

- (1) 网络规模 n 取值为 1 000;
- (2) 最小社团规模 c_{\min} 取值为 10 或 20;
- (3) 混合参数 μ 从 0.05 变化到 0.7, 间隔为 0.05.

(4) 我们保持其他参数不变, 即节点的平均度 $k(P)$ 为 20; 最大度 $k_{\max}P_{\max}$ 为 2.5 倍 $k(P)$; 最大社团规模 c_{\max} 为 5 倍 c_{\min} ; 节点度与社团规模的幂律分布指数分别为 $\varepsilon_1=-2, \varepsilon_2=-1$.

我们通过设置 LFR, 模拟不同参数, 进行了 3 组实验.

实验 1: 设 $n=1000, c_{\min}=10, c_{\max}=50$, 混合参数 μ 为变化量, 各算法对应的 $F1$ 值如图 1 所示.

实验 2: 设 $n=1000, c_{\min}=20, c_{\max}=100$, 混合参数 μ 为变化量, 各算法对应的 $F1$ 值如图 2 所示.

实验 3: 设 $\mu=0.65, c_{\min}=10, c_{\max}=50$, 社团规模 n 为变化量, 各算法对应的 $F1$ 值如图 3 所示.

对实验 1 和实验 2, 我们生成了 2 个网络规模相同、但网络结构不同的 LFR 模拟网络, 通过分析不同算法随着混合参数 μ 取不同值的划分效果. 为了更进一步说明 CDNEV 算法和其他算法的性能差异, 实验 3 中我们固定 $\mu=0.65$, 把网络规模从 1 000 按间隔为 1 000 逐步提升到 10 000, 然后对比分析不同算法性能随着网络规模变化的情况.

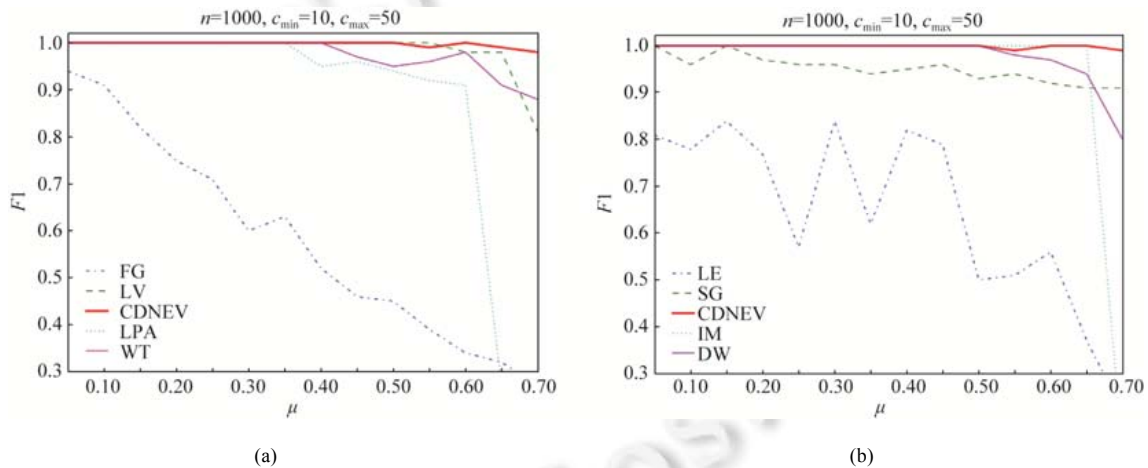


Fig.1 The variation of different algorithms' $F1$ with different μ on base networks, $n=1000, c_{\min}=10, c_{\max}=50$

图 1 不同算法在基准网络上 $F1$ 值随 μ 的变化结果, $n=1000, c_{\min}=10, c_{\max}=50$

图 1 表示在 $n=1000, c_{\min}=10, c_{\max}=50$; 图 2 表示在 $n=1000, c_{\min}=20, c_{\max}=100$, 分别在混合参数 μ 取不同值时, 得到对应的 $F1$ 值. 图 1 和图 2 的 2 个子图是 CDNEV 和不同算法的比较曲线. 图 1 和图 2 的横轴表示 μ 值, 纵轴表示 $F1$ 值.

从图 1 和图 2 可以看出, 在混合参数 μ 取值较小时 ($\mu < 0.2$), 10 种算法所表现出来的 $F1$ 值差别不明显, 特别是在 $c_{\min}=20, c_{\max}=100$ 的情况下. 但是, 随着混合参数 μ 值的增大, 网络中社区结构开始变得模糊化, 不同算法之间的性能差异开始变大. 由图 1 和图 2 可知, CDNEV 对应的 $F1$ 值一直高于除 WT 以外的其他算法, 并且与 WT 的

差距最多不超过 0.02.基于随机游走的 WT 算法的时间复杂度主要依赖于网络中边的数目,时间复杂度为 $O(mn^2)$,而 CDNEV 的时间复杂度仅为 $O(n\log n)$,因此本文提出的算法效率远高于 WT 算法.

图 3 给出了 CDNEV 算法和其他算法在不同网络规模下的 $F1$ 指标变化曲线.

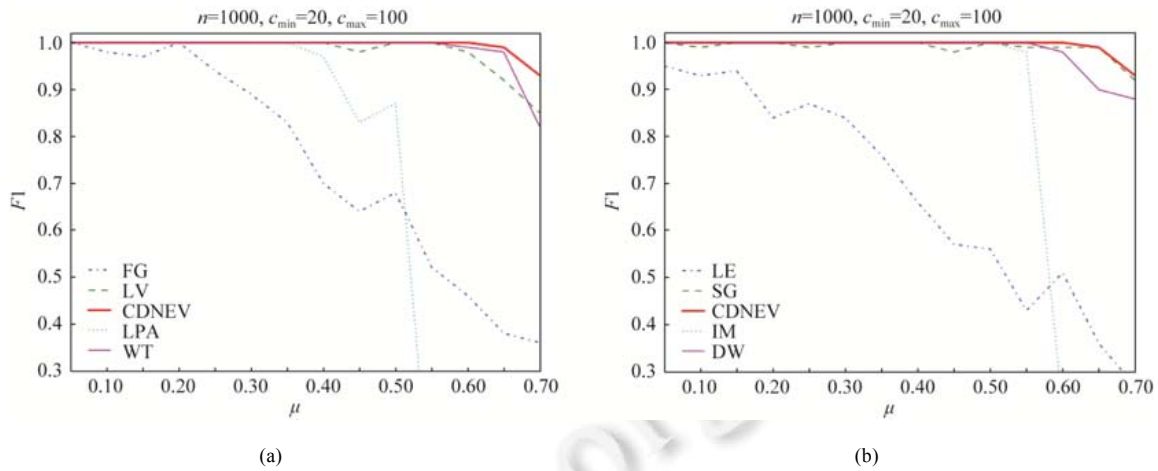


Fig.2 The variation of different algorithms' $F1$ with different μ on base networks, $n=1000, c_{\min}=20, c_{\max}=100$

图 2 不同算法在基准网络上 $F1$ 值随 μ 的变化结果, $n=1000, c_{\min}=20, c_{\max}=100$

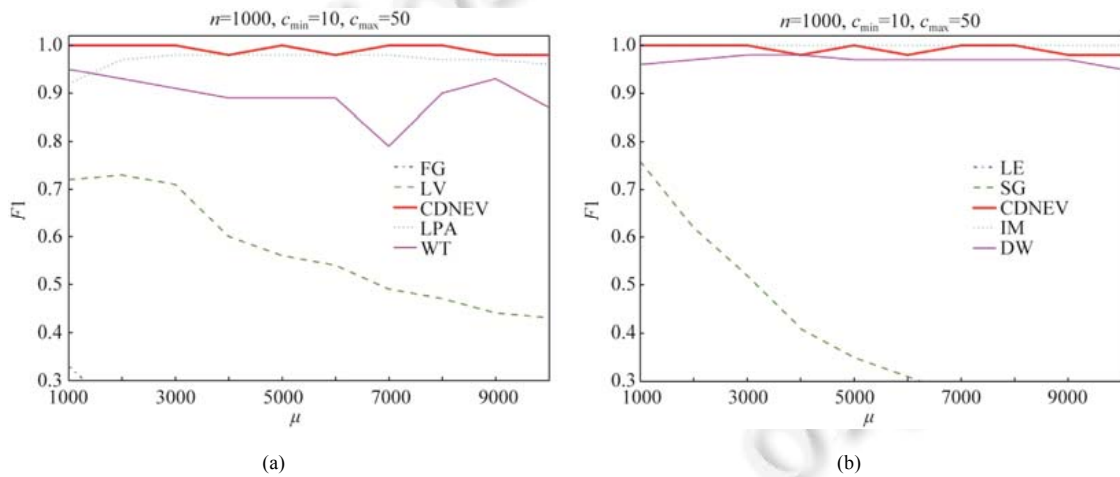


Fig.3 The variation of different algorithms' $F1$ with different network scales on base networks

图 3 不同算法在基准网络规模变化下的 $F1$ 值

如图 3 所示,可以发现,随着网络规模的扩大,不同算法的性能差别同样逐渐加大.原本在网络规模为 1 000 时表现最好的 WT 算法,随着网络规模的扩大,性能逐渐变差.而本文所提出的 CDNEV 算法依然保持较高的水准,IM 算法和 CDNEV 算法在网络规模扩大时性能差异不明显,IM 算法和 CDNEV 算法优于其他算法,但是 IM 算法的时间复杂度为 $O(n(m+n))$,也高于 CDNEV 算法的 $O(n\log n)$,结果说明,CDNEV 算法能够应用于大规模网络上的社团划分.

随着模拟网络节点和边的数量的增加,网络结构更加复杂,导致算法在真实网络和模拟网络上的结果有所下降.但从实验结果中可以清晰地看出,随着网络规模的扩大,所有算法的性能都在下降,CDNEV 算法表现得相对于其他算法依然有一定优势.

4.4 节点向量的有效性分析

为了深入分析 CDNEV 算法生成向量的有效性,我们对 Karate 网络中的节点进行启发式随机游走,然后采用节点向量表示算法对每个节点生成一个二维的分布式向量.

为了验证节点二维分布式向量的效果,我们用二维向量做散点图,如图 4 所示.图 4 中的每个点代表一个节点,不同的颜色表示节点所属社团.

从图 4 可以看出,虽然只用了二维向量,放弃了部分维度上的信息,但用二维向量作为节点距离,然后对节点聚类,依然能够区分社团结构.从散点图可以看出,只有 1 个蓝色的节点靠近红色的社团、1 个红色的节点靠近蓝色的社团,社团之间的重叠区域较少.随着维度的增加,社团内部的节点距离将会更加紧密,这样保证了 CDNEV 算法生成的节点分布式向量能够有效地表示节点在社团结构上的特性.

综合真实网络和模拟网络上的实验结果分析,CDNEV 算法具有简单、有效的特点,适合于在大型复杂的网络上进行社团划分.

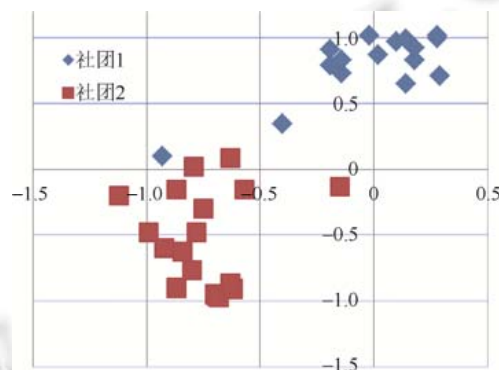


Fig.4 Relationship of 2D vector scatter diagram and community structure

图 4 二维向量散点图与社团结构的关系

5 总结与未来工作

本文提出一种结合自然语言处理方法与聚类方法的社团划分算法 CDNEV. CDNEV 算法首先通过启发式的随机游走算法生成节点上下文序列,然后将节点序列类比为自然语言处理中的词序列,使用 SkipGram 算法学习能够代表每个节点的节点特征向量,在学习过程中使用 Hierarchical Softmax 方法加快特征向量的学习速率,最后依据网络中的核心节点,使用 K-Means 聚类方法得到最后复杂网络的社团结构. CDNEV 算法与其他经典的社区发现算法在多种真实网络和模拟基准网络数据集上进行了比较,实验结果表明,算法在大多数数据集上具有一定优势,不论在有标记网络中的 $F1$ 值,还是在无标记网络中的模块度值都处于较高水准,同时,算法复杂度较小,执行效率高,说明 CDNEV 算法能够适用于大规模复杂网络的社团划分任务,而且能够保持较高的精度.

在大型网络上存在重叠社团结构,如何构造满足重叠等复杂社区结构的节点向量是一个值得研究的方向. CDNEV 模型目前仅使用了网络的结构信息,在未来的工作中,还将引入节点上的文本标记等其他信息.另外,研究算法的并行化处理,使得算法能够应用于大规模网络也是值得研究的方向.对于不同的社团划分算法,还应考虑统一开发的实验语言与运行平台,以便更好地测试不同算法在实际应用中的执行效率.

References:

- [1] Newman MEJ, Watts DJ. Renormalization group analysis of the small-world network model. *Physics Letters A*, 1999,263(4): 341-346.

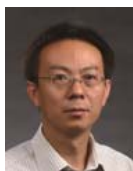
- [2] Shao F, Jiang GP. Optimal traffic routing strategy based on community structure. *Acta Physica Sinica*, 2011,60(7):078902 (in Chinese with English abstract).
- [3] Perozzi B, Al-Rfou R, Skiena S. Deepwalk: Online learning of social representations. In: *Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. ACM, 2014. 701–710.
- [4] Bedi P, Sharma C. Community detection in social networks. *Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery*, 2016,496-500(3).
- [5] Fortunato S, Hric D. Community detection in networks: A user guide. *Physics Reports*, 2016,659:1–44.
- [6] Newman MEJ. Modularity and community structure in networks. *Proc. of the National Academy of Sciences*, 2006,103(23): 8577–8582.
- [7] Shiga M, Takigawa I, Mamitsuka H. A spectral clustering approach to optimally combining numerical vectors with a modular network. In: *Proc. of the 13th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. ACM, 2007. 647–656.
- [8] Fu LD, Gao L, Ma XK. A centrality measure based on spectral optimization of modularity density. *Science China Information Sciences*, 2012,42(5):550–560 (in Chinese with English abstract).
- [9] Riolo MA, Newman MEJ. First-principles multiway spectral partitioning of graphs. *Journal of Complex Networks*, 2014,2(2): 121–140.
- [10] Liu R, Feng S, Shi R, *et al.* Weighted graph clustering for community detection of large social networks. *Procedia Computer Science*, 2014,31:85–94.
- [11] Wang Z, Chen Z, Zhao Y, *et al.* A community detection algorithm based on topology potential and spectral clustering. *The Scientific World Journal*, 2014,2014:329325. [doi: 10.1155/2014/329325]
- [12] Jin H, Wang S, Li C. Community detection in complex networks by density-based clustering. *Physica A: Statistical Mechanics and its Applications*, 2013,392(19):4606–4618.
- [13] Lin W, Kong X, Yu PS, *et al.* Community detection in incomplete information networks. In: *Proc. of the 21st Int'l Conf. on World Wide Web*. ACM, 2012. 341–350.
- [14] Gennip Y, Hunter B, Ahn R, *et al.* Community detection using spectral clustering on sparse geosocial data. *SIAM Journal on Applied Mathematics*, 2013,73(1):67–83.
- [15] Kernighan BW, Lin S. An efficient heuristic procedure for partitioning graphs. *Bell System Technical Journal*, 1970,49(2): 291–307.
- [16] Newman MEJ. Fast algorithm for detecting community structure in networks. *Physical Review E*, 2004,69(6):066133.
- [17] Pan Y, Li DH, Liu JG, *et al.* Detecting community structure in complex networks via node similarity. *Physica A: Statistical Mechanics and its Applications*, 2010,389(14):2849–2857.
- [18] Zanjani AAH, Darooneh AH. Finding communities in linear time by developing the seeds. *Physical Review E*, 2011,84(3):036109.
- [19] Wang XY, Zhao ZX. Partitioning community structure in complex networks based on node dependent degree. *Acta Physica Sinica*, 2014,63(17):178901 (in Chinese with English abstract).
- [20] Clauset A. Finding local community structure in networks. *Physical review E*, 2005,72(2):026132.
- [21] Lancichinetti A, Fortunato S, Kertész J. Detecting the overlapping and hierarchical community structure in complex networks. *New Journal of Physics*, 2009,11(3):033015.
- [22] Lee C, Reid F, McDaid A, *et al.* Detecting highly overlapping community structure by greedy clique expansion. *arXiv Preprint arXiv:1002.1827*, 2010.
- [23] Xu X, Yuruk N, Feng Z, *et al.* SCAN: A structural clustering algorithm for networks. In: *Proc. of the ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. ACM, 2007. 824–833.
- [24] Huang J, Sun H, Han J, *et al.* SHRINK: A structural clustering algorithm for detecting hierarchical communities in networks. In: *Proc. of the ACM Conf. on Information and Knowledge Management, CIKM 2010*. Toronto: DBLP, 2010. 219–228.
- [25] Girvan M, Newman MEJ. Community structure in social and biological networks. *Proc. of the National Academy of Sciences*, 2002, 99(12):7821–7826.
- [26] Blondel VD, Guillaume JL, Lambiotte R, *et al.* Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008,(10):P10008.

- [27] Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 2007,76(3):036106.
- [28] Sun H, Liu J, Huang J, *et al.* CenLP: A centrality-based label propagation algorithm for community detection in networks. *Physica A: Statistical Mechanics & Its Applications*, 2015,436:767–780.
- [29] Tsourakakis C, Gkantsidis C, Radunovic B, *et al.* Fennel: Streaming graph partitioning for massive scale graphs. In: *Proc. of the 7th ACM Int'l Conf. on Web Search and Data Mining*. ACM, 2014. 333–342.
- [30] Yuan C, Chai Y. Method for local community mining in the complex networks. *Acta Automatica Sinica*, 2014,40(5):921–934 (in Chinese with English abstract).
- [31] Albert R, DasGupta B, Mobasher N. Topological implications of negative curvature for biological and social networks. *Physical Review E*, 2014,89(3):032811.
- [32] Gan WY, Nan HE, Li DY, *et al.* Community discovery method in networks based on topological potential. *Ruan Jian Xue Bao/Journal of Software*, 2009,20(8):2241–2254 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3318.htm> [doi: 10.3724/SP.J.1001.2009.03318]
- [33] Wu F, Huberman BA. Finding communities in linear time: A physics approach. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2004,38(2):331–338.
- [34] He D X, Xu Z, Zuo W, *et al.* Community mining in complex networks—Clustering combination based genetic algorithm. *Acta Automatica Sinica*, 2010,36(8):1160–1170 (in Chinese with English abstract).
- [35] Okamoto H. Local detection of communities by attractor neural-network dynamics. In: *Artificial Neural Networks*. Springer Int'l Publishing, 2015. 115–125.
- [36] Zhang ZY. Community structure detection in social networks based on dictionary learning. *Science China (Information Sciences)*, 2011,41(11):1343–1355 (in Chinese with English abstract).
- [37] Tang L, Liu H. Scalable learning of collective behavior based on sparse social dimensions. In: *Proc. of the 18th ACM Conf. on Information and Knowledge Management*. ACM, 2009. 1107–1116.
- [38] Tang J, Qu M, Wang M, *et al.* Line: Large-scale information network embedding. In: *Proc. of the 24th Int'l Conf. on World Wide Web*. Int'l World Wide Web Conferences Steering Committee, 2015. 1067–1077.
- [39] Cao S, Lu W, Xu Q. GraRep: Learning graph representations with global structural information. In: *Proc. of the 24th ACM Int'l on Conf. on Information and Knowledge Management*. ACM, 2015. 891–900.
- [40] Wang DX, Cui P, Zhu WW. Structural Deep Network Embedding. In: *Proc. of the 22nd ACM SIGKDD Int'l Conf.* 2016. 1225–1234. [doi: 10.1145/2939672.2939753]
- [41] Li J, Ritter A, Jurafsky D. Learning multi-faceted representations of individuals from heterogeneous evidence using neural networks. *arXiv Preprint arXiv:1510.05198*, 2015.
- [42] Zhou N, Zhao WX, Zhang X, *et al.* A general multi-context embedding model for mining human trajectory data. *IEEE Trans. on Knowledge & Data Engineering*, 2016,28(8):1945–1958.
- [43] Yang C, Liu Z, Zhao D, *et al.* Network representation learning with rich text information. In: *Proc. of the 24th Int'l Joint Conf. on Artificial Intelligence*. Buenos Aires, 2015. 2111–2117.
- [44] Chen J, Zhang Q, Huang X. Incorporate group information to enhance network embedding. In: *Proc. of the ACM Int'l on Conf. on Information and Knowledge Management*. ACM, 2016. 1901–1904.
- [45] Tang J, Qu M, Mei Q. PTE: Predictive text embedding through large-scale heterogeneous text networks. In: *Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. ACM, 2015. 1165–1174.
- [46] Bengio Y, Ducharme R, Vincent P, *et al.* A neural probabilistic language model. *The Journal of Machine Learning Research*, 2003, 3:1137–1155.
- [47] Mikolov T, Chen K, Corrado G, *et al.* Efficient estimation of word representations in vector space. *arXiv Preprint arXiv:1301.3781*, 2013.
- [48] Mikolov T, Sutskever I, Chen K, *et al.* Distributed representations of words and phrases and their compositionality. In: *Advances in Neural Information Processing Systems*. 2013. 3111–3119.

- [49] Mnih A, Hinton GE. A scalable hierarchical distributed language model. In: Advances in Neural Information Processing Systems. 2009. 1081–1088.
- [50] Morin F, Bengio Y. Hierarchical probabilistic neural network language model. In: Proc. of the Int'l Workshop on Artificial Intelligence and Statistics. 2005. 246–252.
- [51] Bottou L. Stochastic gradient learning in neural networks. Proc. of Neuro-Nimes, 1991,91(8).
- [52] Chen Q, Wu TT, Fang M. Detecting local community structures in complex networks based on local degree central nodes. Physica A: Statistical Mechanics and Its Applications, 2013,392(3):529–537.
- [53] Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. Physical Review E, 2006,74(3):036104.
- [54] Rosvall M, Axelsson D, Bergstrom CT. The map equation. The European Physical Journal Special Topics, 2010,178(1):13–23.
- [55] Reichardt J, Bornholdt S. Statistical mechanics of community detection. Physical Review E, 2006,74(1):016110.
- [56] Pons P, Latapy M. Computing communities in large networks using random walks. In: Computer and Information Sciences-ISCIS 2005. Berlin, Heidelberg: Springer-Verlag, 2005. 284–293.
- [57] Yang J, Leskovec J. Defining and evaluating network communities based on ground-truth. Knowledge & Information Systems, 2015,42(1):181–213.
- [58] Lancichinetti A, Fortunato S, Radicchi F. Benchmark graphs for testing community detection algorithms. Physical Review E, 2008, 78(4):046110.
- [59] Lancichinetti A, Fortunato S. Community detection algorithms: A comparative analysis. Physical Review E, 2009,80(5):056117.

附中文参考文献:

- [2] 邵斐,蒋国平.基于社团结构的负载传输优化策略研究.物理学报,2011,60(7):078902.
- [8] 付立东,高琳,马小科.基于社团检测的复杂网络中心性方法.中国科学(信息科学),2012,42(5):550–560.
- [19] 王兴元,赵仲祥.基于节点间依赖度的社团结构划分方法.物理学报,2014,63(17):178901.
- [30] 袁超,柴毅.复杂网络的局部社团结构挖掘算法.自动化学报,2014,40(5):921–934.
- [32] 淦文燕,赫南,李德毅,王建民.一种基于拓扑势的网络社区发现方法.软件学报,2009,20(8):2241–2254. <http://www.jos.org.cn/1000-9825/3318.htm> [doi: 10.3724/SP.J.1001.2009.03318]
- [34] 何东晓,周翔,王佐,等.复杂网络社区挖掘——基于聚类融合的遗传算法.自动化学报,2010,36(8):1160–1170.
- [36] 张忠元.基于字典学习的网络社团结构探测算法.中国科学(信息科学),2011,41(11):1343–1355.



韩忠明(1972—),男,山西吕梁人,博士,教授,CCF 专业会员,主要研究领域为社会网络,数据挖掘,大数据处理.



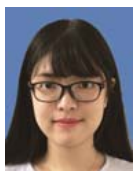
郑晨辉(1994—),女,学士,主要研究领域为社交网络挖掘.



刘雯(1992—),男,学士,主要研究领域为社交网络挖掘.



谭旭升(1990—),男,学士,主要研究领域为社交网络挖掘.



李梦琪(1993—),女,学士,主要研究领域为深度学习,自然语言处理.



段大高(1976—),男,博士,副教授,CCF 专业会员,主要研究领域为异构数据挖掘,大数据处理,社会网络.