

基于三级邻居的复杂网络节点影响力度量方法

杨书新* 梁文 朱凯丽

(江西理工大学信息工程学院 赣州 341000)

摘要: 已有的节点影响力度量方法均存在一定的局限性。该文基于三度影响力原则, 综合考虑局部度量的适宜层次及大规模网络的可扩展性, 提出一种基于3级邻居的节点影响力度量方法(TIM)。该方法将节点2, 3级具有传播衰减特性的邻居视为整体, 用于度量节点的影响能力。利用传染病模型及独立级联模型, 在3个真实数据集验证了该方法的有效性。实验结果表明, 基于3级邻居的节点影响力度量方法在影响力一致性、区分度、排序性等指标中表现优越, 且能够有效求解影响力最大化问题。

关键词: 复杂网络; 节点影响力; 影响力最大化

中图分类号: TP39

文献标识码: A

文章编号: 1009-5896(2020)05-1140-09

DOI: [10.11999/JEIT190440](https://doi.org/10.11999/JEIT190440)

Measurement of Node Influence Based on Three-level Neighbor in Complex Networks

YANG Shuxin LIANG Wen ZHU Kaili

(Faculty of Information Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: There are some limitations in the existing metric methods for measuring node influence. A measurement method of node influence with three-level neighbors is proposed, which is based on the principle of three-degree influence, and considering the appropriate level of local measurement and the scalability of the large-scale network. Firstly, the neighbors with propagation attenuation characteristics in the second and third level of a node are regarded as a whole, which is used to measure the influence of the node. Then, an algorithm for measure called Three-level Influence Measurement (TIM) is proposed. Finally, in order to validate the effectiveness of the algorithm, the experiments on three datasets are conducted by using susceptible-infected-recovered model and independent cascade model. The experimental results show that the proposed algorithm is superior in consistency of influence, discrimination, sorting performance and other evaluation indexes. Furthermore, the TIM is applied to effectively solve the problem of maximizing influence.

Key words: Complex networks; Influence of nodes; Influence maximization

1 引言

识别网络中具有影响力的节点是社交网、交通网、蛋白质交互网等复杂网络中的一个重要研究内容, 其研究有助于人们解决谣言传播的控制、新产品的营销推广、重要蛋白质的发现等实际问题^[1]。

围绕网络中如何识别具有重要影响力的节点, 国内外研究者提出了不同的方法。这些方法在不同数据集表现的效果不尽相同, 其差异的本质就是节点影响力的度量标准的互异。每个节点在网络中有不同的特征属性, 如邻居节点数量、邻居节点之间的拓扑结构、整个网络经过该节点之间的最短路径数量等。从节点拓扑结构特征属性信息利用的角度来看, 已有的度量标准可归为基于网络局部属性的度量、基于网络全局属性的度量、基于网络位置属性的指标、基于随机游走的度量、基于社团结构的度量等^[2,3]。

通过分析节点影响力度量标准的研究成果, 可以看出已有方法主要量化节点的直接、间接或二者混合的邻域结构信息, 如基于 k 个步长的随机游走方法^[4]、基于邻居三角结构和邻居度的方法^[5]、相

收稿日期: 2019-06-17; 改回日期: 2020-02-02; 网络出版: 2020-02-20

*通信作者: 杨书新 yimuyunlang@sina.com

基金项目: 国家自然科学基金(61662028), 江西省教育厅科学技术研究项目基金(GJJ170518), 江西省研究生创新专项资金项目(YC2018-S331)

Foundation Items: The National Natural Science Foundation of China (61662028), The Scientific Technology Research Foundation of the Education Department of Jiangxi Province (GJJ170518), The Special Foundation of Postgraduate Innovation of Jiangxi province (YC2018-S331)

对距离的指标^[6]以及多项指标的混合评价方法^[7]等。节点的影响扩散不是无休止的，它存在衰减及有效范围的概念。网络中信息在扩散的同时，时间在推移，拓扑结构在演进，传播效果会衰弱，这些因素共同导致影响的有效扩散只能在有限层级内发生。2013年，哈佛大学Christakis和加利福尼亚大学Fowler^[8]认为信息在3度强连接内高效传播，节点不仅对邻居节点有影响，还能对邻居的邻居节点以及邻居的邻居的邻居节点有影响，并基于此提出3度影响力原则。3度影响力和6度分隔是社会网络的两个重要理论。3度影响力原则将影响力的有效扩散限定在3级传播内，3度以外的影响传播可以被忽略。关于3度传播有很多例证，如转发微博与原创微博的行为90%在3度传播内发生、导致肥胖的行为、恶性事件、吸烟与戒烟的行为、幸福感的传播均成簇出现且均在3度传播范畴内发生^[9,10]。基于3度影响力原则，研究者提出了线性衰减度中心算法^[11](Linear Decrease Degree Centrality, LDDC), 3-layer局部度中心^[12]算法。LDDC及3-layer算法的本质均是3级邻居的节点数量累加。此类度量方法或掩盖各级邻居间紧密程度的信息，或抹去不同网络对影响扩散的敏感性差异。

节点影响力的研究虽卓有成效，但这些方法各有优缺点，在时间复杂度和准确性方面还需深入研究。为了同传统图结构的“度”概念区分，本文用“级”替代3度影响力原则的“度”。综合考虑局部方法的适宜度量层级与网络拓扑的差异性，本文提出一种新的基于3级邻居的节点影响力度量法(Three-level Influence Measurement, TIM)，并通过实验证明了该方法的有效性。第2节介绍TIM方法；第3节在不同数据集中，通过多种指标验证方法的有效性；第4节为总结及未来工作。

2 基于3级邻居的影响力度量方法

给定社交网络 $G=(V, E, p_{u,v})$, V 和 E 分别表示网络节点集合和边集合， $p_{u,v}$ 表示节点 u 对 v 的激活概率。将具有影响衰减特性的2, 3级邻居视作一个整体，记为 M_{23} 。若2级邻居被激活则 M_{23} 中的节点均可能被激活，此时 M_{23} 激活概率与节点 u 对2级邻居的激活概率近似。基于此推导出 M_{23} 激活概率为

$$\eta = \sum_{w \in N(v)} \sum_{v \in N(u)} p_{u,v} p_{v,w} \quad (1)$$

对于任意节点 u , M_{23} 的节点数量可能是1级邻居节点数量的若干倍。以图1为例，对于节点 u , V , W , X 分别为节点 u 的1~3级邻居。根据图1，标记为 W , X 的节点数量是标记为 V 的节点数量的3倍。

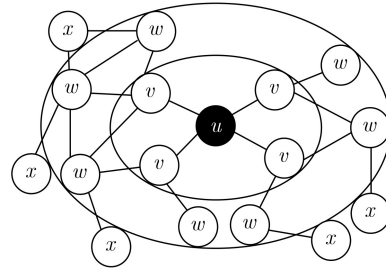


图1 3级影响传播示例

显然在数量关系上，节点 u 对间接邻居的影响大于其对直接邻居的影响，这与传播层级深入影响逐级衰减的特点相左。因此需放大1级影响，调节节点 u 对1级邻居和对 M_{23} 的影响差距。利用函数 $y=e^x$ 当 $x>0$ 恒有 $y>1$ 的特性放大1级传播的影响力的值，并引入参数 θ 调整影响差距。TIM度量节点影响力的计算公式为

$$\text{TIM}(u) = \theta \cdot e^{\sum_{v \in N(u)} p_{u,v}} + \eta \cdot |M_{23}| \quad (2)$$

其中， $|M_{23}|$ 表示节点 u 的2, 3级邻居节点数量。

算法1给出了节点TIM值的计算步骤。

算法1 TIM度量方法

输入: $G=(V, E, P)$ /* P 表示传播概率*/
 输出: 每个节点的TIM度量值

```

(1) function:  $F(\cdot)$  /*1级邻居的层序遍历函数*/
(2) for each  $u$  in  $V$  do:
(3)  $\text{TIM}(u) = 0, x=0, l=0$  /* $l$ 为集合的长度*/
(4) for each  $v$  in  $F(u)$  do:
(5)  $x += p(u, v)$ 
(6) end for
(7)  $\text{TIM}(u) = \theta \cdot \exp(x)$ 
(8) for each  $v$  in  $F(u)$  do:
(9) for each  $w$  in  $F(v) \setminus \{u\}$  do:
(10)  $l = \text{getSize}(\{F(w), w\} \setminus \{F(u)\})$ 
(11)  $\text{TIM}(u) += p(u, v) \times p(v, w) \times l$ 
(12) end for
(13) end for
(14) end for

```

算法1给出了节点TIM值的计算步骤。针对节点集 V 中的全部节点，TIM方法将节点度量值分为两部分计算。第1部分(4)–(7)行计算节点 u 对其1级邻居的影响度量值；第2部分(8)–(13)行，计算节点 u 对1, 2级邻居的传播概率同节点 u 对2, 3级邻居节点数量的乘积，并将该值与第1部分度量值累加，得到节点 u 的TIM度量值。其中，第(10)行 $\text{getSize}(\cdot)$ 函数获取集合的长度。TIM方法遍历每个节点的2, 3级邻居，每级平均访问 $\langle k \rangle$ 个节点，时间复

杂度为 $O(n\langle k \rangle^2)$, $\langle k \rangle$ 表示网络节点度的平均数, n 是网络节点数量。

3 实验分析

实验利用易感染-患病-痊愈(Susceptible Infected Recovered, SIR)模型和独立级联(Independent Cascade, IC)模型,从精度、一致性、区分度、排序性、影响力最大化效果等方面分析TIM方法的有效性。SIR和IC是经典的信息传播模型,广泛应用在信息传播研究中。在SIR模型中,节点有3种状态:易感染态、患病态和免疫态。易感染态节点有概率 λ 被患病节点传染,患病节点有 $1/k_e$ 概率从病态痊愈,免疫态节点不会再患病。对于给定网络,每次仅设定单个节点为传染源,在有限的传播步长 t 内,评价度量结果优劣的参考值 $F(t)$ 是该节点传播后的患病节点与免疫节点的和,每个节点的 $F(t)$ 值重复 10^3 次后取均值。IC模型模拟的是信息及联传递的过程,激活态节点以概率 $p_{u,v}$ 激活邻居节点,该过程一直延续至不再有新的节点被激活。

3.1 实验条件

本文实验程序采用Python编写,计算机配置为3.7 GHz的Intel Xeon E5-1620 v2 CPU,运行内存16 GB。实验设定SIR模型传播概率 $\lambda=0.02$,患病节点治愈概率 $1/k_e$ 为网络节点度均值的倒数,传播步长 $t=7$ 。表1给出了实验网络数据集的基本特征,3组数据集均下载自斯坦福大学SNAP网站。其中,p2p-Gnutella08数据集是完全分布式通信关系网,CA-HepTh数据集是科研合作关系网,WiKi-Vote数据集是来自维基百科的管理员投票网。通过预处理,3组数据采用无向无权图表示。

3.2 参数 θ 的确定

若以节点的 $F(t)$ 为参考度量值,TIM度量值与 $F(t)$ 值的关系可视为预测值与标准值的关系。为使参数 θ 求得的TIM度量值尽可能逼近SIR评价标准,实验利用统计学习方法的模型评估函数观察预测值与度量值之间的误差,误差越小则对应的 θ 值更适宜。模型评估函数表达式为

$$\left. \begin{aligned} R(\theta) &= \frac{1}{n} \sum_{u \in V} L(F(t), \text{TIM}(u)) \\ L(F(t), \text{TIM}(u)) &= [F(t) - \text{TIM}(u)]^2 \end{aligned} \right\} \quad (3)$$

表1 网络数据集基本特征

	p2p-Gnutella08	CA-HepTh	WiKi-Vote
节点数	6301	9877	7115
边数	20777	51971	100762
平均度	6.595	5.264	28.324
聚类系数	0.015	0.600	0.209

$R(\theta)$ 表示TIM方法的度量误差,该值越小越好。以p2p-Gnutella08数据集为例, θ 取不同值时,图2给出了其对应的误差情况。根据图2实验结果, θ 等于 10^3 时对应的误差值最小,此时TIM方法度量结果与SIR模型对应的评价结果具有更明显的一致性。 θ 等于1000时,其他两个数据集的实验表现也具有明显的一致性,因此全文参数 θ 的值均设为1000。

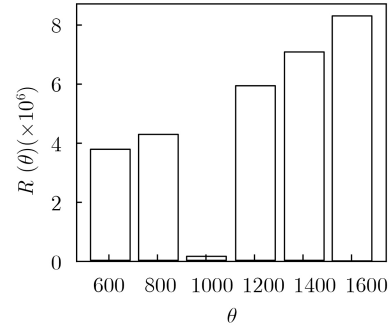


图2 参数 $R(\theta)$ 对应的直方图

3.3 精度提高比

已有的3级度量方法,如LDDC^[11],3-Layer算法^[12]本质上均为1~3级邻居节点数量的累加。LDDC算法在求和的步骤中对2,3级邻居添加了对应的衰减系数,但比3个层级邻居数量的直接累加,其效果并无明显差距。因此,本文舍去了已有3级度量算法具有针对性的参数,统一表示为3级邻居数量的直接累加和,并简称为朴素方法(Plain Measurement, PM)。基于SIR模型的节点影响力仿真结果 $F(t)$,本节设计精度提高比函数 $\varepsilon(k)$ 来对比朴素法和TIM方法。精度提高比的函数表达式为

$$\varepsilon(k) = \frac{\sum_{u \in V_{\text{Top}-k}} (\text{Skew}_u^{\text{PM}} - \text{Skew}_u^{\text{TIM}})}{\sum_{u \in V_{\text{Top}-k}} \text{Skew}_u^{\text{TIM}}} \times 100\% \quad (4)$$

其中, $\text{Skew}_u^{\text{PM}}$ 表示朴素法降序编号同 $F(t)$ 值降序编号的差的绝对值, $\text{Skew}_u^{\text{TIM}}$ 表示节点 u 的TIM值降序编号同 $F(t)$ 值降序编号的差的绝对值。以p2p-Gnutella08数据集的实验结果为例,编号367的节点 $F(t)$ 值最大,记编号1。该节点在朴素法和TIM方法降序编号分别为2,1,则 $\text{Skew}_{367}^{\text{PM}} = |1 - 2| = 1$, $\text{Skew}_{367}^{\text{TIM}} = |1 - 1| = 0$ 。偏差值越小表示节点影响力度量方法与基于SIR模型的仿真结果越相近,说明度量结果的精度越高。因TIM方法具有随机性,取其重复运行 10^3 次的均值作为节点影响力的最终值。表2列出了TIM方法对Top- k 节点度

量结果的精度提高情况。根据表2的实验结果，与朴素法相比，TIM方法整体度量精度有所提高。尤其在幂律网络中，以SIR模型结果为参考，TIM方

法Top-100节点的度量精度平均提高了2.6倍，效果显著。

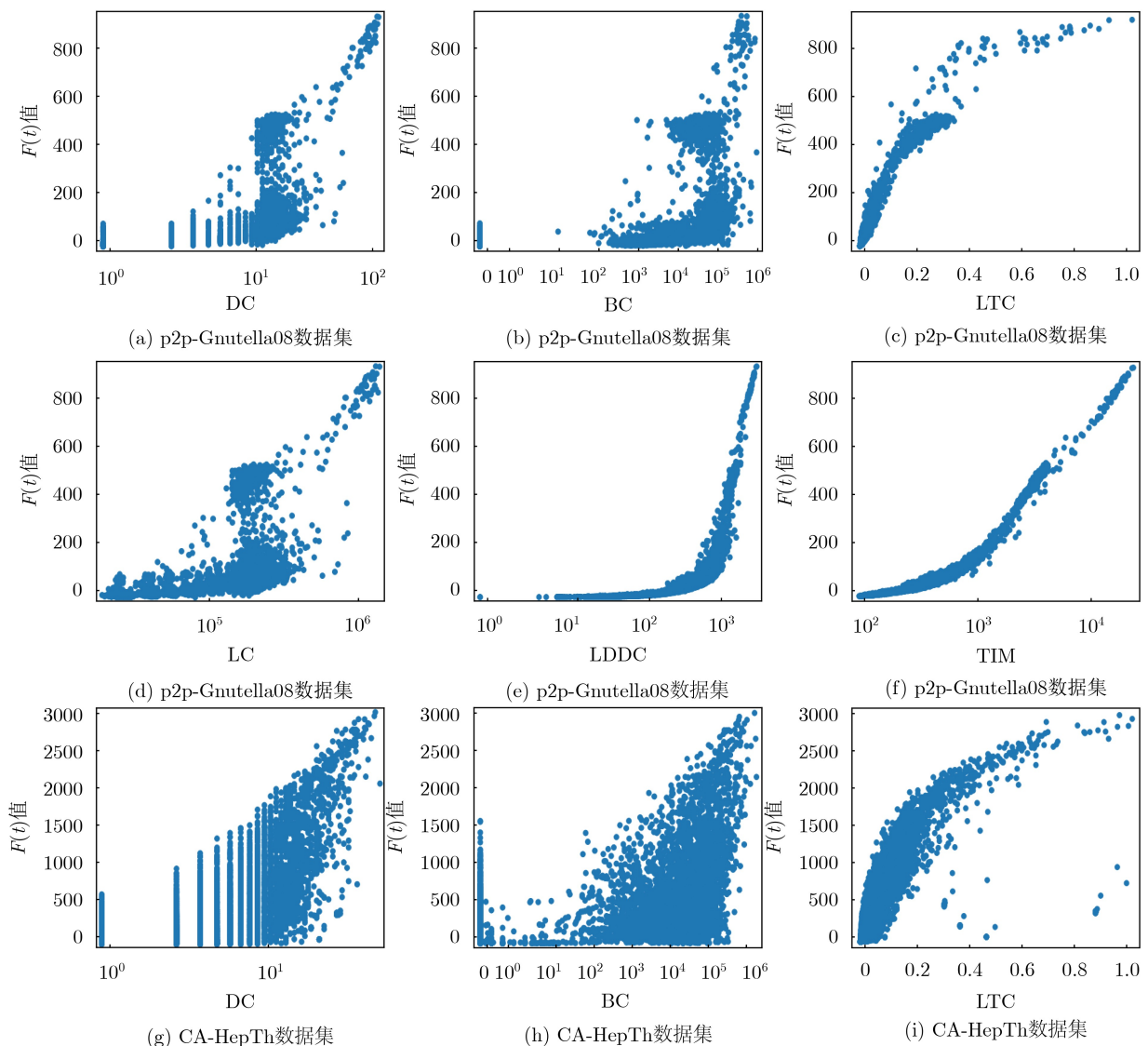
3.4 影响力一致性

针对影响力一致性评价指标，本节选取局部中心性(Local Centrality, LC)^[13]、LDDC^[11]、度中心性(Degree Centrality, DC)^[4]、介数中心性(Betweenness Centrality, BC)^[14]、三角中心性(Local Triangle Centrality, LTC)^[5]方法对比实验。其中，LC方法统计的是多级邻居数量的累加和，LTC算法统计的是节点所处3元闭包结构的数量。本节实验保留算法原思想，并未限制LC算法的度量层级。LC, LTC, LDDC, DC均属局部度量方法；BC方法具有鲁棒性高、不适用于大规模网络的特点，如果能找到比BC效果好、时间复杂度低的方法，那么该方法的适用范围将更加广泛。

以SIR模型的仿真结果 $F(t)$ 为参考，图3给出了各对比方法与 $F(t)$ 的相关性结果。其中纵坐标为

表2 精度提高比(%)

	p2p-Gnutella08	Wiki-Vote	CA-HepTh
Top-10	10.00	133.33	36.00
Top-20	22.58	280.00	16.46
Top-30	15.87	278.69	5.41
Top-40	2.20	291.60	16.09
Top50	7.56	272.59	3.32
Top-60	15.61	286.55	11.71
Top-70	13.77	263.78	10.25
Top-80	9.44	323.90	21.49
Top-90	2.71	241.53	7.13
Top-100	1.47	229.58	5.86
Avg	10.12	260.16	13.37



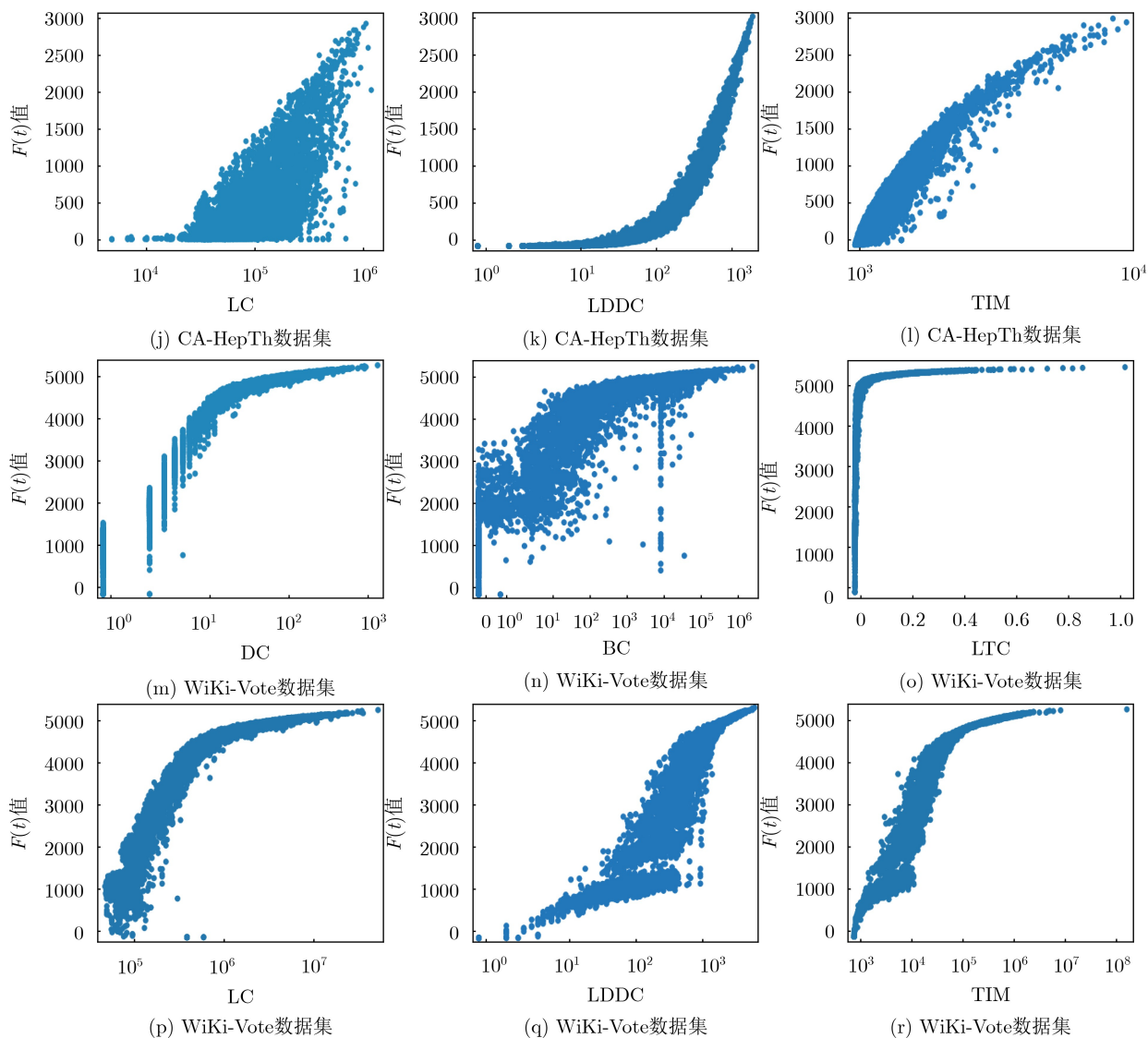


图3 影响力一致性实验结果

$F(t)$, 横坐标为各方法的度量值。良好的正相关性在点状图中应当表现为向上倾斜的曲线。图3实验结果表明LTC, LDDC及TIM在3个数据中均表现出较强的一致性, 其中TIM算法的一致性尤为明显。以CA-HepTh数据集DC, BC方法为例的实验结果频繁出现“竖条”图案, 通过观察, 该现象是部分节点度量值相同而节点对应的 $F(t)$ 值并不相同造成的, 即此方法度量值的区分度偏低。3.5节实验将对比各方法度量值的区分度。

3.5 度量值区分度

本节实验采用度量值区分度函数 $\text{disc}(\cdot)$ 评估TIM方法的对节点影响力的鉴别性能, $\text{disc}(\cdot)$ 函数表达式为

$$\text{disc}(\cdot) = \frac{\text{MaxDesOrder}}{n} \quad (5)$$

其中, n 为网络节点总数, MaxDesOrder为基于节

点度量值的最大排序序号, $\text{disc}(\cdot) \in (0, 1]$ 。 $\text{disc}(\cdot)$ 值越接近1, 则度量方法的区分度越大。实验先降序排列某方法的节点度量值并编号, 令度量值相等的节点具有相同编号, 再根据式(5)计算 $\text{disc}(\cdot)$ 值。以CA-HepTh数据集为例, 共计10个节点的DC值等于38, 值“38”第1次出现的节点排序序号为18, 则剩下的9个节点排序序号也是18。根据表3给出的区分度实验结果, TIM在6种度量方法中区分力最强, LC方法其次。BC, LDDC算法在不同数据集的区分度表现不稳定。本节实验对节点影响力一致性实验中DC, BC, LC实验图像的散点“抱团”现象及DC, LTC实验图像的“竖条”现象做出了解释。

3.6 排序性

排序性是指度量方法对Top- k 节点正确排序的能力, 各方法的实验过程如下: 建立各节点 $F(t)$ 值

表3 区分度实验结果

方法	p2p-Gnutella08	Wiki-Vote	CA-HepTh
DC	0.01206	0.04216	0.00557
LTC	0.05055	0.05205	0.04454
BC	0.71861	0.64216	0.40376
LC	0.85129	0.80689	0.72161
LDDC	0.60705	0.60899	0.32672
TIM	0.98905	0.99874	0.91789

与度量值的2元组, 降序排列节点度量值再为每组数对添加正序编号, 对Top- k 节点的 $F(t)$ 取平均, 将正序编号及其对应的 $F(t)$ 值建立排序折线图, 实验结果见图4。良好排序性能在图像中应表现为向下倾斜的曲线。图4实验结果表明LDDC及TIM的排序性能在不同数据集中总能表现良好。在p2p-Gnutella08和Wiki-Vote数据集中, TIM排序表现要优于LDDC算法。

3.7 影响力最大化问题求解效果

本节实验基于IC模型求解影响力最大化的应用问题。影响力最大化问题^[15]是在给定网络及传播模型中找到由 k 个种子构成的集合, 使该集合最终的影响收益尽可能的多。评价影响力最大化算法的性

能主要从运行时间及影响范围两方面判断。运行时间是算法模拟整个传播过程所需的时耗, 该值越小越好; 影响范围是在相同种子数和同一传播模型下算法最终激活的个体数, 该值越大越好。

本节选取当前具有代表性的算法开展对比实验, 包括: 度启发^[16](DegreeHeuristic, DH), 度折扣^[16](DegreeDiscount, DD), 强连通分量启发^[17](Strongly Connected Components heuristic, SCC), k 核覆盖算法^[18](k Core Coverage Algorithm, CCA), 改进贪心^[16](NewGreedy, NG), 随机(Random)及密度中心性^[19](DensityCentrality, DeC)共7种算法, 实验结果见图5—图7, 其中CCA算法的距离参数 d 设为2, 故图5—图7中CCA算法统一标记为CCA(2)。各算法的最终激活节点数是模拟了 10^4 次传播过程后确定的, 每个数据集在0.001~0.050间取3个不同的传播概率 p 值, 避免特定概率下具有偶然性的实验结果。根据图5—图7的实验结果, TIM算法在人工交互网络和幂律分布网络表现优异, 在科研合著网络(CA-HepTH数据集)表现次优。科研合著网络存在大规模的完全连通子图, 而一般图数据所包含的连通子图极少。TIM方法选拔出的部分种子处于完全连通子图中心, 其传播能力

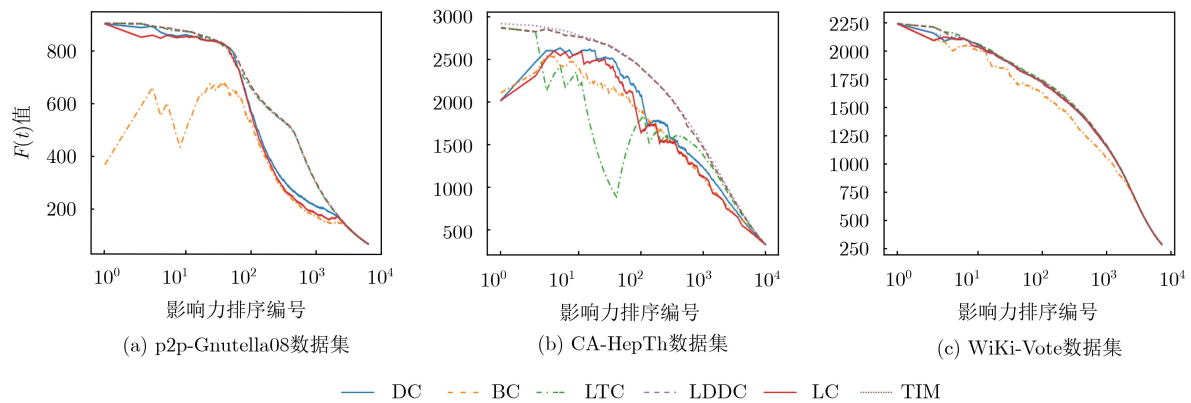


图4 排序性能

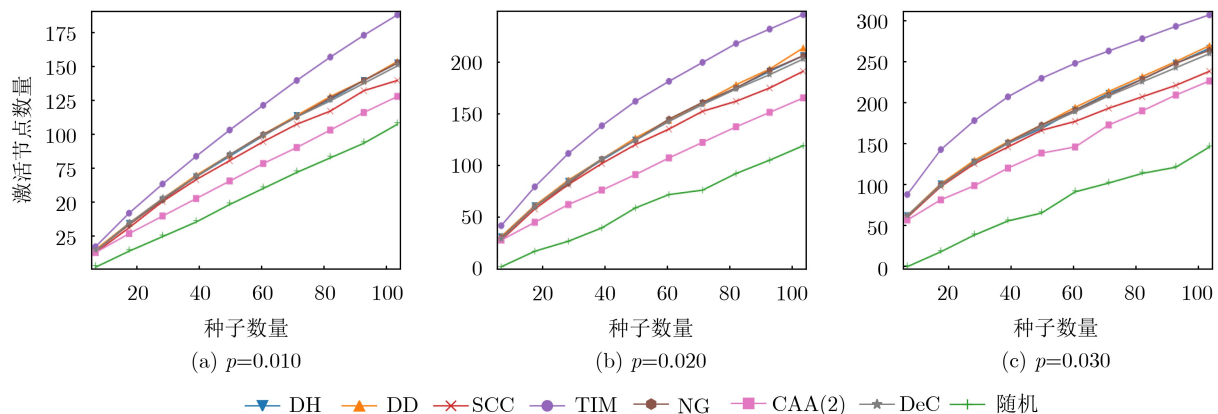


图5 p2p-Gnutella08数据集实验结果

很难表现, 因此图6(b)出现了影响传播乏力的现象。图5—图7的实验结果表明, 随着传播概率和种子数量的增加, TIM方法对应种子的影响效果显著提高。根据影响最大化的评价指标, TIM方法在影响力覆盖范围方面存在明显优势。

TIM是一种迭代式方法, 需反复运算得到稳定排序后确定Top- k 节点。基于反复的实验观察, 3组数据集最少迭代20次时方能得到唯一的Top- k

组合。因此, 本节实验中TIM方在各数据集的迭代次数为20次。表4给出了不同数据集中各算法的运行时间。根据表4的实验数据, DH和DD算法的运行时间最短, NG算法的运行时间最长。TIM方法尽管经过了20次迭代, 但其运行时长仍小于NG算法, 且最终影响收益要高于NG算法。综合影响效果和运行时耗两个指标, 基于独立级联模型, TIM方法能够有效挖掘影响力最大化问题的关键节点。

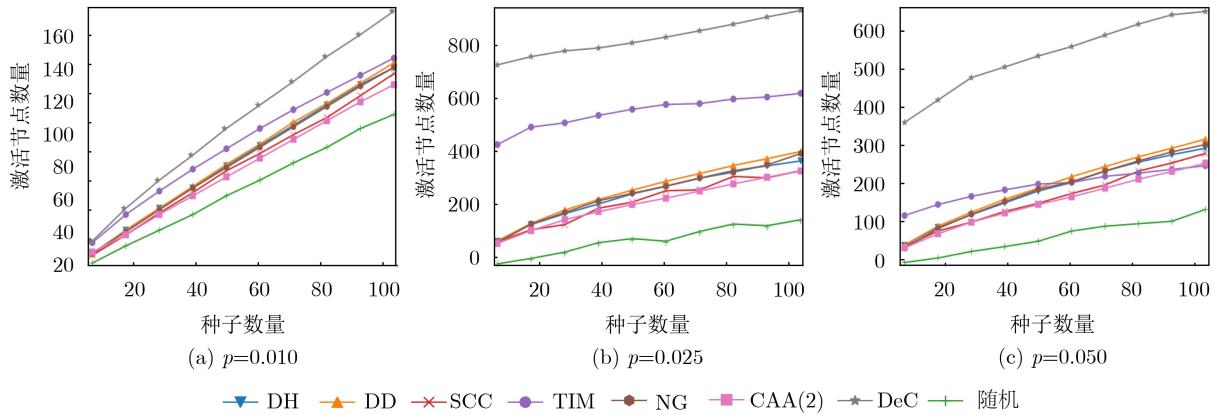


图6 CA-HepTH数据集实验结果

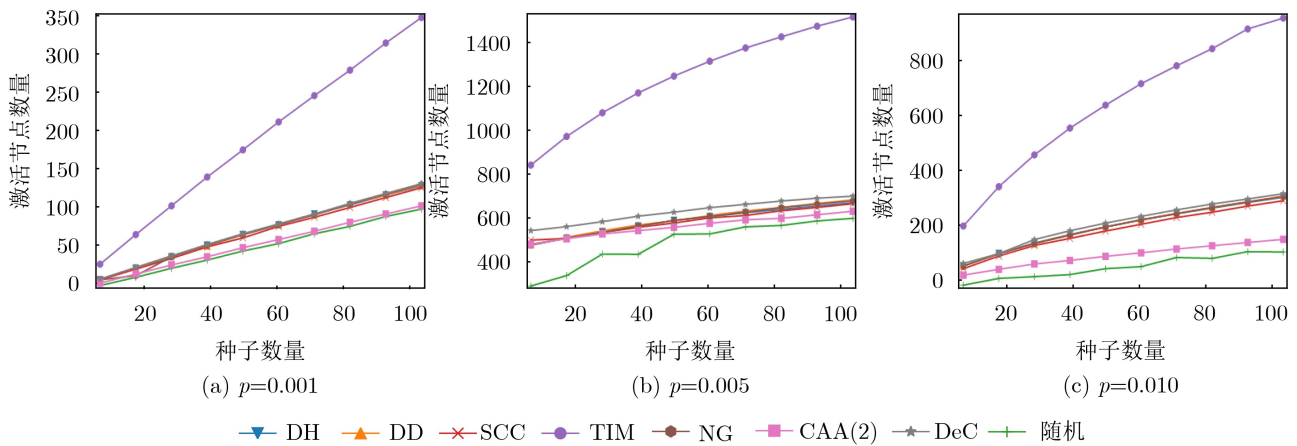


图7 WiKi-Vote数据集

表4 运行时间 (s)

数据集	传播概率	DH	DD	随机	SCC	TIM	NG	CCA(2)	DeC
p2p-Gnutella08	$p=0.010$	0.022	0.027	0.002	0.025	0.258	0.406	0.041	0.046
	$p=0.020$	0.028	0.029	0.003	0.036	0.278	0.415	0.043	0.058
	$p=0.030$	0.038	0.034	0.005	0.038	0.320	0.423	0.045	0.063
CA-HepTH	$p=0.010$	0.014	0.017	0.0016	0.019	0.194	0.573	0.054	0.030
	$p=0.025$	0.021	0.021	0.0025	0.027	0.239	0.645	0.062	0.059
	$p=0.050$	0.038	0.045	0.0062	0.034	0.325	0.792	0.076	0.161
WiKi-Vote	$p=0.001$	0.141	0.136	0.011	0.157	0.517	1.921	0.434	0.412
	$p=0.005$	0.249	0.265	0.026	0.261	0.651	1.947	0.481	0.526
	$p=0.010$	0.793	0.776	0.480	0.772	1.153	2.434	0.975	1.001

4 结束语

基于3度影响力原则, 本文所设计的TIM影响力度量方法将1级传播与2, 3级传播累加计算。实验表明TIM与具有代表性的方法对比, 在各项评价指标中表现优异, 能有效求解影响力最大化问题, 且具有较低的时间复杂度。TIM在近似评估节点影响力时考虑3级范围内直接和间接邻居数量, 但有研究表明同1级邻居节点的相互连接会影响节点的传播, 因此未来工作将结合同1级邻居之间连接等因素做进一步分析。

参考文献

- [1] 刘阳, 季新生, 刘彩霞. 一种基于边界节点识别的复杂网络局部社区发现算法[J]. 电子与信息学报, 2014, 36(12): 2809–2815. doi: [10.3724/SP.J.1146.2013.01955](#).
LIU Yang, JI Xinsheng, and LIU Caixia. Detecting local community structure based on the identification of boundary nodes in complex networks[J]. *Journal of Electronics & Information Technology*, 2014, 36(12): 2809–2815. doi: [10.3724/SP.J.1146.2013.01955](#).
- [2] 田晶, 方华强, 刘佳佳, 等. 运用复杂网络方法分析城市道路网的鲁棒性[J]. 武汉大学学报: 信息科学版, 2019, 44(5): 771–777. doi: [10.13203/j.whugis20150334](#).
TIAN Jing, FANG Huaqiang, LIU Jiajia, *et al.* Robustness analysis of urban street networks using complex network method[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(5): 771–777. doi: [10.13203/j.whugis20150334](#).
- [3] 王凯, 刘树新, 于洪涛, 等. 基于共同邻居有效性的复杂网络链路预测算法[J]. 电子科技大学学报, 2019, 48(3): 432–439. doi: [10.3969/j.issn.1001-0548.2019.03.020](#).
WANG Kai, LIU Shuxin, YU Hongtao, *et al.* Predicting missing links of complex network via effective common neighbors[J]. *Journal of University of Electronic Science and Technology of China*, 2019, 48(3): 432–439. doi: [10.3969/j.issn.1001-0548.2019.03.020](#).
- [4] 韩忠明, 刘雯, 李梦琪, 等. 基于节点向量表达的复杂网络社团划分算法[J]. 软件学报, 2019, 30(4): 1045–1061. doi: [10.13328/j.cnki.jos.005387](#).
HAN Zhongming, LIU Wen, LI Mengqi, *et al.* Community detection algorithm based on node embedding vector representation[J]. *Journal of Software*, 2019, 30(4): 1045–1061. doi: [10.13328/j.cnki.jos.005387](#).
- [5] 韩忠明, 陈炎, 李梦琪, 等. 一种有效的基于三角结构的复杂网络节点影响力度量模型[J]. 物理学报, 2016, 65(16): 168901. doi: [10.7498/aps.65.168901](#).
HAN Zhongming, CHEN Yan, LI Mengqi, *et al.* An efficient node influence metric based on triangle in complex networks[J]. *Acta Physica Sinica*, 2016, 65(16): 168901. doi: [10.7498/aps.65.168901](#).
- [6] 杨青林, 王立夫, 李欢, 等. 基于相对距离的复杂网络谱粗粒化方法[J]. 物理学报, 2019, 68(10): 100501. doi: [10.7498/aps.68.20181848](#).
YANG Qinglin, WANG Lifu, LI Huan, *et al.* A spectral coarse graining algorithm based on relative distance[J]. *Acta Physica Sinica*, 2019, 68(10): 100501. doi: [10.7498/aps.68.20181848](#).
- [7] 林冠强, 莫天文, 叶晓君, 等. 基于TOPSIS和CRITIC法的电网关键节点识别[J]. 高电压技术, 2018, 44(10): 3383–3389. doi: [10.13336/j.1003-6520.hve.20180925030](#).
LIN Guanqiang, MO Tianwen, YE Xiaojun, *et al.* Critical node identification of power networks based on TOPSIS and CRITIC methods[J]. *High Voltage Engineering*, 2018, 44(10): 3383–3389. doi: [10.13336/j.1003-6520.hve.20180925030](#).
- [8] CHRISTAKIS N A and FOWLER J H. Social contagion theory: Examining dynamic social networks and human behavior[J]. *Statistics in Medicine*, 2013, 32(4): 556–577. doi: [10.1002/sim.5408](#).
- [9] CHRISTAKIS N A and FOWLER J H. Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives[M]. New York: Little, Brown and Company, 2009: 30–117.
- [10] 许小可, 胡海波, 张伦, 等. 社交网络上的计算传播学[M]. 北京: 高等教育出版社, 2015: 163–198.
XU Xiaoke, HU Haibo, ZHANG Lun, *et al.* Computational Communication in Social Networks[M]. Beijing: Higher Education Press, 2015: 163–198.
- [11] 陈晓龙. 社会网络影响力最大化算法及其传播模型研究[D]. [硕士论文], 哈尔滨工程大学, 2016: 20–22.
CHEN Xiaolong. Research on influence maximization and diffusion model in social networks[D]. [Master dissertation], Harbin Engineering University, 2016: 20–22.
- [12] 王俊, 余伟, 胡亚慧, 等. 基于3-layer中心度的社交网络影响力最大化算法[J]. 计算机科学, 2014, 41(1): 59–63. doi: [10.3969/j.issn.1002-137X.2014.01.009](#).
WANG Jun, XU Wei, HU Yahui, *et al.* Heuristic algorithm based on 3-layer centrality for influence maximization in social networks[J]. *Computer Science*, 2014, 41(1): 59–63. doi: [10.3969/j.issn.1002-137X.2014.01.009](#).
- [13] CHEN Duanbing, LÜ Linyuan, SHANG Mingsheng, *et al.* Identifying influential nodes in complex networks[J]. *Physica A: Statistical Mechanics and its Applications*, 2012, 391(4): 1777–1787. doi: [10.1016/j.physa.2011.09.017](#).
- [14] FREEMAN L C. A set of measures of centrality based on betweenness[J]. *Sociometry*, 1977, 40(1): 35–41. doi: [10.2307/3033543](#).
- [15] LI Jianxin, LIU Chengfei, XU J X, *et al.* Personalized

- influential topic search via social network summarization[C]. The 33rd IEEE International Conference on Data Engineering, San Diego, USA, 2016: 1820–1834. doi: [10.1109/ICDE.2017.15](https://doi.org/10.1109/ICDE.2017.15).
- [16] CHEN Wei, WANG Yajun, and YANG Siyu. Efficient influence maximization in social networks[C]. The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 2009: 199–208. doi: [10.1145/1557019.1557047](https://doi.org/10.1145/1557019.1557047).
- [17] KEMPE D, KLEINBERG J, and TARDOS É. Maximizing the spread of influence through a social network[C]. The 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, USA, 2003: 137–146. doi: [10.1145/956750.956769](https://doi.org/10.1145/956750.956769).
- [18] 曹玖新, 董丹, 徐顺, 等. 一种基于k-核的社会网络影响最大化算法[J]. 计算机学报, 2015, 38(2): 238–248. doi: [10.3724/SP.J.1016.2015.00238](https://doi.org/10.3724/SP.J.1016.2015.00238).
- CAO Jiuxin, DONG Dan, XU Shun, *et al.* A k-core based algorithm for influence maximization in social networks[J]. *Chinese Journal of Computers*, 2015, 38(2): 238–248. doi: [10.3724/SP.J.1016.2015.00238](https://doi.org/10.3724/SP.J.1016.2015.00238).
- [19] IBNOULOAFI A and EL HAZITI M. Density centrality: Identifying influential nodes based on area density formula[J]. *Chaos, Solitons & Fractals*, 2018, 114: 69–80. doi: [10.1016/j.chaos.2018.06.022](https://doi.org/10.1016/j.chaos.2018.06.022).
- 杨书新: 男, 1979年生, 副教授, 研究方向为社会网络分析、生物信息学.
- 梁 文: 男, 1994年生, 硕士生, 研究方向为复杂网络、计算传播学.
- 朱凯丽: 女, 1994年生, 硕士生, 研究方向为隐私保护、推荐系统.