

基于局部域的影响力最大化算法

沈记全, 林 帅, 李志莹

(河南理工大学 计算机科学与技术学院, 河南 焦作 454003)

摘 要: 用户影响力度量是影响力最大化问题的核心, 与网络拓扑结构相关的影响力度量指标主要分为全局性指标和局部性指标, 其中全局性指标需要依靠网络完整拓扑结构计算节点影响力且时间复杂度较高, 局部性指标通常忽略或弱化了网络中的自环和多边现象, 导致对节点影响力的度量不全面, 限制信息最终传播范围。结合三度分隔原理, 提出基于局部域的影响力最大化算法。考虑网络中的自环和多边现象, 根据网络拓扑结构构建生成图。依据生成图划分每个节点对应的局部域, 使用节点在局部域内的影响力近似其在全局范围内的影响力, 并据此选择候选种子节点。计算候选种子加入种子集合后的重叠比因子, 根据重叠比因子决定是否将此候选种子节点选作种子节点, 控制种子集合的影响力重叠程度。在真实数据集上的实验结果表明, 与 MaxDegree、PageRank 等算法相比, 该算法能有效识别高影响力节点群体, 扩大信息传播范围, 且具有较低的时间复杂度。

关键词: 影响力最大化; 局部影响力; 全局影响力; 三度分隔; 影响力重叠

开放科学(资源服务)标志码(OSID):



中文引用格式: 沈记全, 林帅, 李志莹. 基于局部域的影响力最大化算法[J]. 计算机工程, 2022, 48(7): 22-28.

英文引用格式: SHEN J Q, LIN S, LI Z Y. Influence maximization algorithm based on local domain [J]. Computer Engineering, 2022, 48(7): 22-28.

Influence Maximization Algorithm Based on Local Domain

SHEN Jiquan, LIN Shuai, LI Zhiying

(College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454003, China)

[Abstract] Measures of user influence are at the core of the influence maximization problem. As these measures relate to network topology, they may be classified as global indicators or local indicators. Global indicators rely on the complete network topology to calculate the influence of nodes, and incur a high time complexity. In contrast, local indicators generally ignore or weaken the self-loop and multilateral phenomena in the network, resulting in an incomplete measurement of the influence of nodes, which affects the final dissemination range of information. Using the principle of three degrees of separation, an influence maximization algorithm based on local domain is proposed in this study. This algorithm first constructs a generated graph by making full use of the edges, starting at and ending with the same interconnected nodes. Then, it generates a local domain for each node and approximates each node's global influence, calculated by the topology of its local domain. Finally, the algorithm conducts seed selection by considering the influence of each node, as well as the influence overlap ratio factor of a seed set. Experimental results on real datasets show that, compared with algorithms such as MaxDegree and PageRank, the proposed algorithm can effectively identify high-influence node groups, improve the scope of information dissemination, and incur lower time complexity.

[Key words] influence maximization; local influence; global influence; three degrees of separation; influence overlap

DOI: 10.19678/j.issn.1000-3428.0062160

0 概述

随着智能移动设备的普及和5G网络的应用, 社交网络发展迅速。影响力最大化问题作为社交网络分析领域的热点研究问题, 旨在从社交网络中寻找 k 个最具影响力的节点, 并最大化以这些节点作为种

子节点的最终信息传播范围, 被广泛应用于在线广告^[1-2]、病毒营销^[3]、专家发现^[4]、个性化推荐^[5]等任务。用户影响力度量是影响力最大化问题的核心。网络拓扑结构^[6]是影响力度量的重要依据。与网络拓扑结构相关的影响力度量指标可以分为全局性指标和局部性指标。介数中心性^[7]、接近中心性^[8]、离

基金项目: 国家自然科学基金面上项目(61972134); 河南省科技攻关计划重点项目(192102210123)。

作者简介: 沈记全(1969—), 男, 教授、博士, 主研方向为智能信息处理、计算机控制、应用系统集成; 林 帅、李志莹, 硕士研究生。

收稿日期: 2021-07-22 修回日期: 2021-10-08 E-mail: linshuai19825@163.com

心中心性^[9]、流介数中心性^[10]、Katz中心性^[11]、连通介数中心性^[12]等都属于全局性指标。全局性指标需要依靠网络完整拓扑结构,在整个网络范围内计算节点影响力,因此时间复杂度很高。度中心性^[13]、半局部中心性^[14]、特征向量中心性^[15]等都属于局部性指标。局部性指标仅依据局部范围的拓扑信息计算节点的影响力,与全局性指标相比时间复杂度较低。然而,影响力传播具有三度分隔特性,即社交网络中相距三度是强连接,强连接可以引发行为,相距超过三度是弱连接,弱连接只能传递信息。节点有自环^[16]特性,拥有自环的节点比没有自环的节点自身活跃度更高,两个节点之间还可能存在多边,这些因素都与节点的影响力密切相关。现有的局部性指标往往忽略这些因素,导致对影响力的度量不够全面,从而影响信息的最终传播范围。

本文结合三度分隔原理^[17],提出用节点在局部范围内的影响力近似其在全局范围内影响力的算法。根据网络拓扑结构构造生成图,依据生成图划分每个节点对应的局部域,根据节点在对应局部域内的影响力筛选候选种子节点。计算每个节点与种子集合的重叠比因子,并据此决定候选种子节点是否能成为种子节点。通过在真实数据集上的实验结果以验证本文算法的正确性和有效性。

1 相关知识

1.1 影响力最大化问题

社交网络可用有向图 $G=(V, E)$ 表示,其中, V 是节点集合且 $V=\{v_1, v_2, \dots, v_n\}$, E 是边集合且 $E=\{(v_i, v_j) | v_i, v_j \in V\}$ 。 $(\forall v_i) v_i \in V, (\forall v_j) v_j \in V$, 如果 v_i 对 v_j 产生一次通信行为,则从 v_i 到 v_j 构成了一条有向边 $\langle v_i, v_j \rangle$ 。

影响力最大化问题是从网络 $G=(V, E)$ 中找到一个大小为 k 的节点集合 S , 最大化以这 k 个节点作为种子节点开始影响力传播后影响力的最终传播范围(即激活的节点数最多)。影响力最大化问题可表示如下:

$$S \leftarrow \operatorname{argmax} f(S^*) \{S^* | S^* \in V, |S^*| = k\} \quad (1)$$

其中: S^* 表示节点集合; $f(S^*)$ 表示影响力的传播范围。

1.2 影响力传播模型

线性阈值 (Linear Threshold, LT) 模型^[18-19]可以用来模拟影响力的传播过程。在 LT 模型中,节点只能处于激活状态或者非激活状态。激活状态的节点通过有向边试图激活处于非激活状态的邻居节点。对于非激活状态的节点,当所有激活状态的邻居节点对其影响力之和大于该节点的激活阈值时,该非激活节点就转为激活状态。当网络中不再有节点被激活(即由非激活状态转为激活状态)时,影响力的传播过程收敛。

图1给出了根据 LT 模型模拟影响力传播的过程,其中,灰色圆表示激活节点,白色圆表示非激活节点,有向边上数字表示箭尾节点对箭首节点的影响力。节点 v_1, v_2, v_5 和 v_6 为种子节点,在传播开始时处于激活状态;其他节点则处于非激活状态。整个传播过程经过3轮迭代收敛。

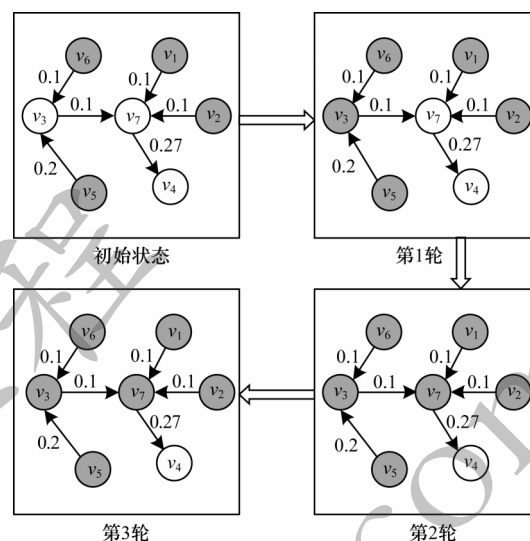


图1 LT模型上的信息传播示例

Fig.1 Example of information dissemination on LT model

2 影响力最大化算法

本文针对局部性度量指标构造生成图,根据影响力传播的三度分隔原理构建局部影响力度量模型,依据局部影响力度量模型设计基于局部域影响力的种子节点选择算法。

2.1 权重感知的生成图构造

边影响权重根据自环和多边属性计算。边权重反映了节点间的亲密程度,影响力的传播过程又受邻居间关系疏密的影响。因此,边影响权重的计算过程中应考虑自环和多边现象。

自环指的是一种特殊的环结构,这种环状结构只包含一个节点。社交网络中用户可能发起只有自己参与的社交活动,从而在对应的社交网络图中形成只有该节点参与的自环。自环多的节点活跃度较高,在信息传播过程中会主动地影响其他节点。本文引入自环因子度量自环对节点能力的影响。自环因子计算如下:

$$R_{v_i} = \sum \psi(v_i) \quad (2)$$

其中: R_{v_i} 表示节点 v_i 的自环因子; $\psi(v_i)$ 表示节点 v_i 每个自环的权重。在无权网络中, $\psi(v_i)$ 默认为1,此时节点自环因子等于节点的自环个数。

多边指的是社交网络图中两个节点间有多条边存在。在社交网络中,两个节点间的一次互动会在社交网络图中对应的节点间产生一条边。当两个节点间有多次互动时,它们之间就会有多条边。然而,随着边数的增加,图的存储和遍历成本也会增加。

在不影响图的存储和计算成本的前提下,本文引入多边因子,用以度量节点间存在的多边现象对信息传播过程的影响。多边因子随着两个节点间边的增多而增大。多边因子计算如下:

$$M_{v_i, v_j} = \sum_{i \neq j} \phi(v_i, v_j) \quad (3)$$

其中: M_{v_i, v_j} 表示节点 v_i 和 v_j 的多边因子; $\phi(v_i, v_j)$ 表示 v_i 到 v_j 一条边的权重。在无权网络中, $\phi(v_i, v_j)$ 默认为1,此时节点多边因子等于两节点之间的边数。

结合自环因子和多边因子,本文引入边影响权重,描述节点间的影响力。

定义1 给定一条边 (v_i, v_j) , v_i 对 v_j 的影响力即为该边的边影响权重。边影响权重计算如下:

$$\tau_{v_i, v_j} = M_{v_i, v_j} / \sum_{v_k \in N_{in}^D(v_j)} (M_{v_k, v_j} + R_{v_j}) \quad (4)$$

其中: $N_{in}^D(v_j)$ 是由出边直接指向 v_j 的节点构成的集合。

根据自环因子、多边因子和边影响权重构造生成图,并且生成图是有向带权图。

2.2 局部域影响力度量

影响力的全局性度量指标往往从拓扑结构的整体出发对节点的影响力进行全面准确的衡量,但是却存在计算复杂度高的问题。根据影响力传播遵循的三度分隔原理^[17],即节点的影响力在相距不超过三跳的邻居节点间传播时随着距离的增大而不断衰减,当传播距离超过三跳时几近消失。本文根据三度分隔特性,利用节点的影响力在特定的局部范围内的传播过程来近似其在全局范围内的传播过程。为了度量节点在特定局部范围内的影响力,引入最短距离、局部邻居以及局部域的概念,并结合局部域拓扑结构建立影响力度量模型。

定义2 节点间最短路径的长度即为节点间的最短距离。假设 $(\forall v_i) v_i \in V, (\forall v_j) v_j \in V$, 若 v_i 到 v_j 的最短路径为 $p(v_i, v_j)$ 且 $p(v_i, v_j) = \langle v_i, p_1, p_2, \dots, p_m, v_j \rangle$, 则节点 v_i 到 v_j 的最短距离 l_{v_i, v_j} 为 $(m+1)$ 。

定义3 给定节点 v_i , 若 v_j 是 v_i 的可达节点且到 v_i 的距离不大于 D , 则称 v_j 是 v_i 的局部邻居。

定义4 给定节点及其局部邻居构成的节点集合以及这些节点间的边构成的边集合共同组成此节点的局部域, 记为 A_L 。以节点 v_i 为中心的局部域记为 $A_L(v_i)$, 表示如下:

$$A_L(v_i) = (\tilde{V}, \tilde{E}) \quad (5)$$

其中: $\tilde{V} = \{v_i | v_i \cup N_L(v_i)\}$, $N_L(v_i)$ 表示节点 v_i 的局部邻居集; $\tilde{E} = \{\langle v_i, v_j, \tau_{v_i, v_j} \rangle | v_i \in \tilde{V}, v_j \in \tilde{V}\}$ 。

定义5 对于给定节点, 其局部域影响力为此节点为中心的局部域拓扑结构所决定的节点传播信息的能力, 记为 Cr 。 $(\forall v_i) v_i \in V, Cr(v_i)$ 计算如下:

$$Cr(v_i) = \sum_{v_j \in N_L(v_i)} 2^{-(l_{v_i, v_j} - 1)} \tau_{v_i, v_j} \quad (6)$$

其中: l_{v_i, v_j} 是节点 v_i 到节点 v_j 的最短距离; τ_{v_i, v_j} 是节点 v_i 到节点 v_j 的最短路径中最后一条边的边影响权重。

以8节点网络中节点 v_1 为例说明局部域影响力的计算过程, 如图2所示, 其中圆圈内部字母及其数字下标表示节点, 有向边上的数值代表两节点间的边影响权重。假设 $D=3$, 节点 v_1 的局部域由节点 v_1 与其局部邻居节点 $\{v_2, v_3, v_4, v_5\}$ 以及它们之间的边构成。

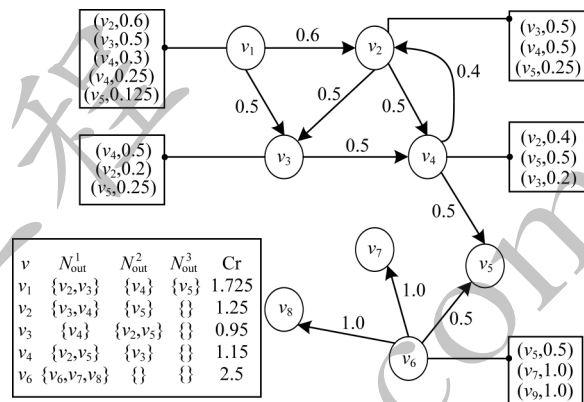


图2 计算节点的Cr值示例

Fig.2 Example of computing Cr values of nodes

2.3 影响力重叠检测

节点之间可能存在影响力重叠现象, 导致多个节点的共同影响力小于各节点影响力之和。为了保证影响力的传播范围最大, 在选择种子节点时应考虑节点之间可能存在的影响力重叠现象。权衡影响力重叠检测成本和消除影响力重叠后的收益, 本文允许影响力重叠, 但是应避免影响力过度重叠, 并利用式(7)判定给定节点间是否存在影响力重叠。

$$\text{overlap}(v_i, v_j) = \begin{cases} 1, & \frac{|N_{out}^D(v_i) \cap N_{out}^D(v_j)|}{|N_{out}^D(v_i) \cup N_{out}^D(v_j)|} > \eta \\ 0, & \text{其他} \end{cases} \quad (7)$$

其中: η 表示过度重叠判定阈值; $N_{out}^D(v_i)$ 和 $N_{out}^D(v_j)$ 分别表示 v_i 和 v_j 的出边直接连接的节点构成的集合。在式(7)的基础上, 本文定义了重叠比因子, 判定一个集合中节点间影响力重叠的程度。

定义6 重叠比因子为给定集合中影响力过度重叠的节点在集合中所占的比例, 记为 ω 。给定集合 $C \subseteq V$, 该集合的重叠比因子计算如下:

$$\omega_c = \frac{1}{|C|} \sum_{i \neq j} \text{overlap}(v_i, v_j) \quad (8)$$

2.4 算法描述

首先根据Cr值构建候选种子节点列表, 并将该列表中第一个节点作为种子加入种子集合。然后依次从候选种子节点列表中选择一个节点, 并计算该节点加入种子集合后该种子集合的重叠比因子。若

该重叠比因子小于预定义的重叠比阈值,则将此节点加入种子节点集合,否则不能将此节点加入种子集合。重复上述过程,直至选出足够数量的种子节点。算法1给出了基于局部域的影响力最大化算法的伪代码。

算法1 基于局部域的影响力最大化算法

输入 $G'=(V', E'), k, \eta, \theta // \theta$ 为激活阈值
输出 S
1.初始化: $S \leftarrow \emptyset$
2. $CrList \leftarrow$ node list with descending order of Cr values
3. $S+ = CrList(0)$
4.for $i \leftarrow 1$ to $(|CrList| - 1)$ do
5. $flag \leftarrow true$
6. $\omega_s \leftarrow 0$
7.for every s in S do
8.if $l_{s, CrList(i)} \geq \theta$ then
9. $flag \leftarrow false$
10.else
11. $\omega_s += overlap(s, CrList(i)) / k$
12.if $\omega_s \geq \eta$ then
13. $flag \leftarrow false$
14.break
15.end if
16.end if
17.end for
18.if $flag == true$ then
19. $S+ = CrList(i)$
20.end if

21.if $|S| == k$ then
22.break
23.end if
24.end for
25.return

对列表中的每个候选种子节点按顺序挑选,在最坏情况下需要遍历所有候选种子节点。对于每个遍历到的候选种子节点,均要根据式(8)检测是否与某种子节点存在影响力过度重叠。给定种子节点数 k ,判定某个候选种子节点是否与某个种子节点存在过度重叠的时间复杂度为 $O(kd)$,其中 d 为节点平均出度。基于以上分析,从节点集合 V 中选出 k 个种子节点的时间复杂度在最坏情况下为 $O(|V|kd)$,也可表示为 $O(|E|k)$ 。除种子集合 S 外,本文算法对于每个节点还存储了 $flag$ 和 ω_s 两个变量,且最多为列表中所有节点均设置,因此算法的空间复杂度在最坏情况下为 $O(|V|)$ 。

3 实验与结果分析

实验选取5种对比算法,通过在6个不同规模的真实数据集上比较算法的影响力传播范围、算法运行时间等指标来验证本文算法的正确性和有效性。

3.1 数据集与实验环境

表1给出了选用的数据集信息,其中, $|V|$ 表示节点总数, $|E|$ 表示边总数, d 为节点平均出度。所有数据集均来自斯坦福大学的大型网络数据集(<http://snap.stanford.edu/data/>)。

表1 数据集设置
Table 1 Dataset setting

数据集	$ V $	$ E $	d	备注
Wiki-Vote	889	2 914	3	节点表示网站用户,有向边表示一个用户给另一个用户投过票
Fb-Forum	899	33 720	37	节点表示网站用户,有向边表示一个用户评论了另一个用户
Escorts-Dynamic	10 106	50 632	5	节点表示买方和卖方,有向边表示买方评价了卖方
P2p-Gnutella04	10 876	39 994	3	节点表示主机地址,有向边表示主机之间发送了文件
Cora	23 166	91 500	3	节点表示科学论文,有向边表示左节点引用了右节点
Digg-Reply	30 398	87 627	2	节点表示网站用户,有向边表示一个用户回复了另一个用户

实验选用MaxDegree算法、PageRank算法、ICGW算法、Closeness算法和IDD1算法5种对比算法。这些对比算法分别采用最大度策略^[13]、PageRank方法^[20]、约束贪婪方式下的影响力节点识别方法^[21]、Closeness方法^[8]和改进的度折扣启发式策略^[22]选择种子节点。所有算法均用Python实现,在相同的Windows平台上运行。该平台采用Intel Core i5 1.80 GHz处理器,配置8 GB内存空间。

实验采用的评价指标分别为影响力传播范围、算法运行时间和占用的内存空间大小。影响力传播范围用激活的节点数表示,该值越大越好,算法运行时间和内存空间占用则越小越好。

3.2 结果分析

在本文算法中,参数 η 、 θ 和 D 分别表示过度重叠判定阈值、线性阈值模型下的激活阈值和局部域范围。在本次实验中,取值分别为0.2、0.4和3。该部

分给出的所有结果都是相关算法独立运行 1 000 次计算的平均结果。

图 3 给出了各算法的影响力传播范围,其中 k 表示种子节点数。由图 3 可知:本文算法在所有数据集上的种子节点的影响力平均传播范围最大,其次是 IDD1 算法和 ICGW 算法;本文算法相比于 MaxDegree 算法、PageRank 算法和 Closeness

算法所选种子节点的影响力平均传播范围较小;本文算法在 Cora 数据集上表现最好,所选种子节点影响力平均传播范围分别是 IDD1 算法和 ICGW 算法的 1.77 倍和 1.47 倍;本文算法在 Escorts-Dynamic 数据集上表现最差,但影响力传播范围仍是 IDD1 算法和 ICGW 算法的 1.04 倍和 1.03 倍。

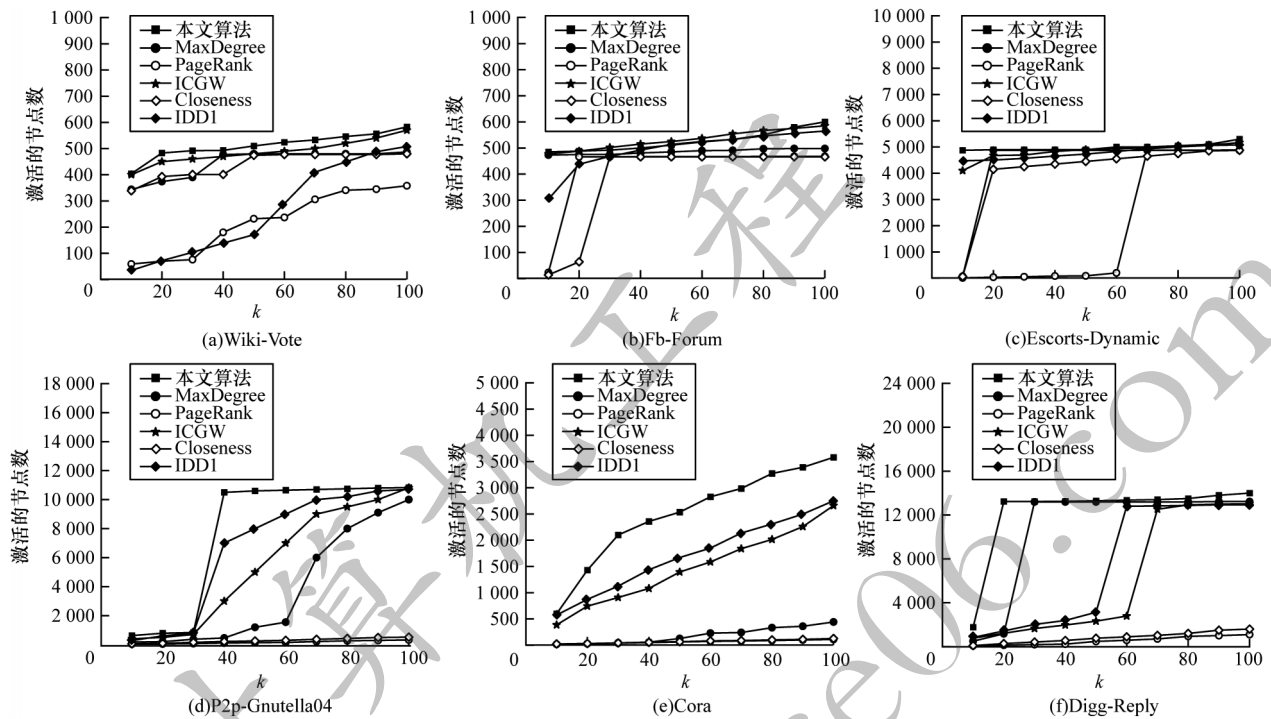


图 3 6种算法的影响力传播范围比较

Fig.3 Comparison of influence dissemination range among six algorithms

在种子节点数相同的情况下,本文算法的传播结果始终优于 IDD1 算法和 ICGW 算法。这是由于本文算法具备影响力重叠控制能力,因此在选择相同的种子时,它对节点的组合影响力更加敏感。在所有数据集上,本文算法的影响力传播范围在种子节点数从 10 到 50 阶段几乎都呈线性增长,这表明了本文算法识别高影响力节点群体的能力更强。

减少种子节点选择所消耗的时间是本文算法设计的另一个重要目标。图 4 给出了在 6 个数据集上各种算法选取 k 个种子节点所用的平均运行时间。总体来看,MaxDegree 算法平均运行时间最短,其次是本文算法、IDD1 算法、PageRank 算法和 ICGW 算法,Closeness 算法平均运行时间最长。由于 IDD1 算法和 ICGW 算法的影响力传播范围和本文算法最为接近,而其他 3 种算法的影响力传播范围则远小于本文算法,失去了比较意义,因此在评价算法的时间和空间性能时只与 IDD1 算法和 ICGW 算法比较。

由图 4 可知:本文算法在 P2p-Gnutella04 数据集

上表现最好,运行时间分别是 IDD1 算法和 ICGW 算法的 7% 和 8%;本文算法在 Cora 数据集上表现最差,但运行时间仍仅是 ICGW 算法的 17%;本文算法在 Wiki-Vote 数据集和 Cora 数据集上的运行时间多于 IDD1 算法,但在其他数据集上的运行时间都少于 IDD1 算法。因此,从所有数据集来看,本文算法要略优于 IDD1 算法。

通过实验评估本文算法的空间复杂度。算法在执行过程中占用的内存空间越多,其空间复杂度越高,反之亦然。在实验过程中,随机选择 $k=50$ 的运行过程,比较各算法的内存占用情况。表 2 给出了在 $k=50$ 时各算法在不同数据集上运行时占用的内存空间。由表 2 可知:在所有数据集上运行时,本文算法所占用的内存空间均小于 ICGW 算法;除了 Fb-Forum 数据集和 P2p-Gnutella04 数据集之外,本文算法的内存空间占用都低于 IDD1 算法。从整体来看,本文算法的内存空间占用情况要优于 IDD1 算法。

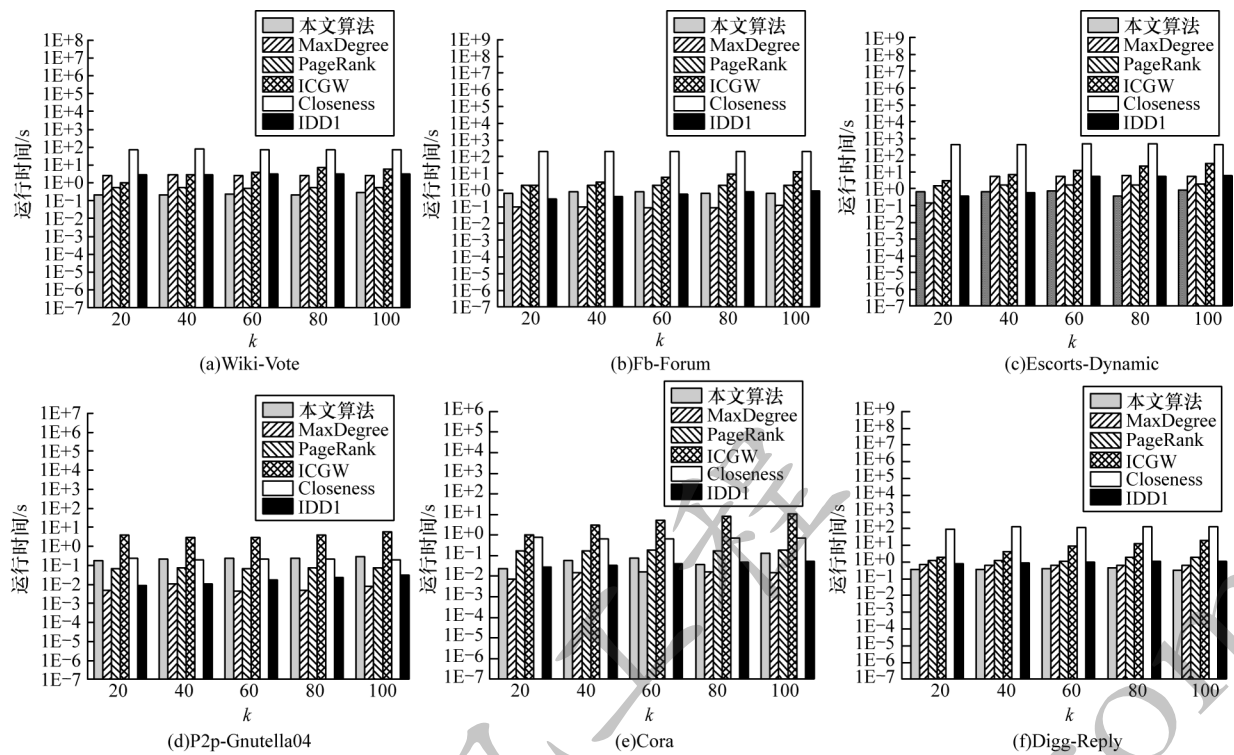


图 4 6 种算法的运行时间比较
Fig.4 Comparison of execution time among six algorithms

表 2 6 种算法在不同数据集上运行时所占用的内存空间

Table 2 Memory space consumed by six algorithms when running on different datasets						MB
算法	Wiki-Vote	Fb-Forum	Escorts-Dynamic	P2p-Gnutella04	Cora	Digg-Reply
本文算法	74.70	87.77	127.09	124.93	198.07	203.16
MaxDegree	75.90	88.56	127.91	125.96	197.71	203.30
PageRank	76.30	87.62	136.21	136.50	227.81	239.31
ICGW	75.40	87.80	129.96	127.15	202.24	208.30
Closeness	76.56	87.88	138.00	136.91	225.38	236.96
IDD1	74.90	87.41	127.82	125.85	201.43	207.10

4 结束语

针对度量用户影响力的全局性和局部性度量指标各自存在的局限性,本文利用节点在局部域内的影响力近似其在全局范围内的影响力,提出一种基于局部域的影响力最大化算法。构建以节点为中心的局部域影响力度量模型,依据影响力度量模型筛选候选种子节点集合。利用重叠比因子刻画候选种子节点与种子节点集合间的重叠程度,并根据重叠比因子决定是否将此候选种子节点选作种子节点。实验结果验证了该算法的正确性与有效性。下一步可将本文算法扩展应用于动态社交网络,利用其能高效准确选取高影响力种子节点的优势,设计并实现高影响力节点集合的增量式更新方法。

参考文献

[1] TANG X N, YANG C C. Ranking user influence in healthcare social media [J]. ACM Transactions on Intelligent Systems and Technology, 2012, 3(4): 1-21.

[2] HUANG J M, CHENG X Q, SHEN H W, et al. Exploring social influence via posterior effect of word-of-mouth recommendations [C]//Proceedings of the 5th ACM International Conference on Web Search and Data Mining. New York, USA: ACM Press, 2012: 573-582.

[3] BADASHIAN A S, STROULIA E. Measuring user influence in GitHub: the million follower fallacy [C]//Proceedings of the 3rd International Workshop on CrowdSourcing in Software Engineering. Washington D. C., USA: IEEE Press, 2016: 15-21.

[4] DOMINGOS P, RICHARDSON M. Mining the network value of customers [C]//Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2001: 57-66.

[5] 陶天一, 王清钦, 付聿炜, 等. 基于知识图谱的金融新闻个性化推荐算法[J]. 计算机工程, 2021, 47(6): 98-103, 114. TAO T Y, WANG Q Q, FU Y W, et al. Personalized recommendation algorithm for financial news based on knowledge graph[J]. Computer Engineering, 2021, 47(6): 98-103, 114. (in Chinese)

[6] WADHERA T. Brain network topology unraveling epilepsy and ASD association: automated EEG-based diagnostic model[J]. Expert Systems with Applications, 2021, 186: 115762.

[7] FREEMAN L C. A set of measures of centrality based on betweenness[J]. Sociometry, 1977, 40(1): 35.

[8] FREEMAN L C. Centrality in social networks conceptual

- clarification[J]. *Social Networks*, 1978, 1(3): 215-239.
- [9] HAGE P, HARARY F. Eccentricity and centrality in networks[J]. *Social Networks*, 1995, 17(1): 57-63.
- [10] FREEMAN L C, BORGATTI S P, WHITE D R. Centrality in valued graphs: a measure of betweenness based on network flow[J]. *Social Networks*, 1991, 13(2): 141-154.
- [11] KATZ L. A new status index derived from sociometric analysis[J]. *Psychometrika*, 1953, 18(1): 39-43.
- [12] ESTRADA E, HATANO N. Communicability in complex networks[J]. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 2008, 77(3): 036111.
- [13] BONACICH P. Factoring and weighting approaches to status scores and clique identification[J]. *The Journal of Mathematical Sociology*, 1972, 2(1): 113-120.
- [14] CHEN D B, LÜ L, SHANG M S, et al. Identifying influential nodes in complex networks[J]. *Physica A: Statistical Mechanics and Its Applications*, 2012, 391(4): 1777-1787.
- [15] BONACICH P. Power and centrality: a family of measures[J]. *American Journal of Sociology*, 1987, 92(5): 1170-1182.
- [16] 尼古拉斯·克里斯塔基斯, 詹姆斯·富勒. 大连接: 社会网络是如何形成的以及对人类现实行为的影响[M]. 北京: 中国人民大学出版社, 2013.
- CHRISTAKIS N A, FOWLER J H. *Connected: the surprising power of our social networks and how they shape our lives*[M]. Beijing: China Renmin University Press, 2013. (in Chinese)
- [17] 孔秀琼, 袁正中, 蔡萍. 带自环的无向网络的同步研究[J]. 闽南师范大学学报(自然科学版), 2019, 32(3): 26-30.
- KONG X Q, YUAN Z Z, CAI P. Synchronization of undirected complex network with self-loop[J]. *Journal of Minnan Normal University(Natural Science)*, 2019, 32(3): 26-30. (in Chinese)
- [18] KEMPE D, KLEINBERG J, TARDOS É. Maximizing the spread of influence through a social network [C]// *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, USA: ACM Press, 2003: 137-146.
- [19] 王怡, 梁循, 付虹蛟, 等. 社会网络中信息的扩散机理及其定量建模[J]. *中国管理科学*, 2017, 25(12): 147-157.
- WANG Y, LIANG X, FU H J, et al. Mechanism and quantity modeling of information diffusion in social networks[J]. *Chinese Journal of Management Science*, 2017, 25(12): 147-157. (in Chinese)
- [20] BRIN S, PAGE L. The anatomy of a large-scale hypertextual Web search engine[J]. *Computer Networks and ISDN Systems*, 1998, 30(1/2/3/4/5/6/7): 107-117.
- [21] ZHANG X H, LI Z Y, QIAN K, et al. Influential node identification in a constrained greedy way[J]. *Physica A: Statistical Mechanics and Its Applications*, 2020, 557: 124887.
- [22] 夏欣, 马闯, 张海峰. 基于改进的度折扣方法研究社交网络影响力最大化问题[J]. *电子科技大学学报*, 2021, 50(3): 450-458.
- XIA X, MA C, ZHANG H F. An improved degree discount approach for influence maximization in social networks[J]. *Journal of University of Electronic Science and Technology of China*, 2021, 50(3): 450-458. (in Chinese)

编辑 陆燕菲