

DOI: 10.7652/xjtuxb201604007

采用影响力节点集扩展的局部社团检测

常振超, 陈鸿昶, 黄瑞阳, 于洪涛, 刘阳
(国家数字交换系统工程技术研究中心, 450002, 郑州)

摘要: 针对规模化网络中局部社团检测存在的对初始节点位置敏感、拓扑信息难以有效利用问题, 提出了一种采用影响力节点集扩展的社团检测 (IN-LCD) 方法。首先定义了节点的局部影响力指标, 通过该指标计算并构造了源节点附近的影响力节点子集, 然后从影响力节点子集开始, 以迭代更新的方式, 进行连续的社团扩张, 最后通过节点和社团相似性指标计算, 完成整个局部社团的获取。IN-LCD 方法从有效利用节点局部信息出发, 通过最具影响力节点集合进行社团扩展, 有效克服了局部社团检测对初始节点位置敏感的问题。在真实和人工网络数据集上的实验表明, IN-LCD 方法与已有的最佳局部社团检测方法相比, 识别性能提升了 5.3%, 更能有效应用于局部信息出发的社团检测场景。

关键词: 社团检测; 局部信息; 影响力节点; 识别性能

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2016)04-0041-07

A Local Community Detection Method Using Expansion of Influential Nodes Set

CHANG Zhenchao, CHEN Hongchang, HUANG Ruiyang, YU Hongtao, LIU Yang
(National Digital Switching System Engineering & Technological Research Center, Zhengzhou 450002, China)

Abstract: A local community detection algorithm based on influential nodes set (IN-LCD) is proposed to focus the problems that the local community detection in large-scale network is sensitive to the position of source nodes and the topology information is difficult to effectively use. A local influence index for nodes is defined, and a subset of influential nodes near the source node is calculated and constructed with the index. Then, the continuous expansion of the community is realized from the subset, and the whole local community is constructed through the calculation of the similarity index between nodes and community. The method uses the most influential nodes set to expand the community and effectively overcomes the sensitive problem of local community detection to initial node position. Experiments on real and artificial network data sets and a comparison with an existing local community detection method show that the recognition performance of the proposed IN-LCD is improved by 5.3%.

Keywords: community detection; local information; influential node; recognition performance

网络存在于人类生活的方方面面, 如电子邮件 网、在线社交网络等^[1]。社团通常指的是具备相似

收稿日期: 2015-10-25。 作者简介: 常振超(1987—), 男, 博士生; 陈鸿昶(通信作者), 男, 教授, 博士生导师。 基金项目: 国家自然科学基金资助项目(61171108); 国家重点基础研究发展计划资助项目(2012CB315901); 国家科技支撑计划资助项目(2014BAH30B01)。

网络出版时间: 2016-01-13 网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20160113.1951.002.html>

属性的节点组合^[2],是网络中最基础的组织单元和功能结构,对其进行检测是研究网络属性的最基本任务之一。当前已经有多种社团检测的方法,主要可以分为基于全局信息和基于局部信息2类^[3]。基于全局信息的方法,首先通过某种定义将整个网络视为一张图,进而分析整个图的全局信息,其在一些规模比较小的网络取得了较好的效果^[4-6],但这类方法由于需要获取全局信息,因此难以应对当前信息残缺且规模庞大的在线网络处理需求。例如,Facebook和Twitter包含了上亿个节点,Web网络包含了数十亿个网页,在这些规模巨大的复杂网路中检测社团将会非常耗时,且全局信息难以获取^[7]。基于局部信息的方法由于从局部角度出发不需要对网络全局结构进行分析,因此近年来引起了研究者的广泛关注^[8]。局部社团检测聚焦于研究子图结构,从节点的邻居信息出发,检测包含该节点的社团结构,并不需要提前预知整个网络的全局信息。然而,现有的局部社团检测算法识别性能较差,原因在于算法对初始节点的位置敏感,如果初始节点位于社区的核心位置(具备较高的影响力),那么社团检测的结果将会更好于初始节点位于边界位置(具备较低的影响力)的节点,因此还需要进一步提高现有的局部社团检测算法的鲁棒性。

基于此,本文提出了一种影响力节点集扩展的局部社团检测方法(local community detection based on influential nodes, IN-LCD)。该方法首先从给定的节点出发,寻找该节点邻近节点中最具影响力的节点,构造最具影响力的候选节点集,然后基于构造的节点集合,进行扩展以获取完整社团结构。本文方法从有效利用节点的邻居信息出发,很好地克服了局部社团检测对于初始节点位置敏感的问题。

1 局部社团检测

局部社团检测指的是从给定的源节点出发,检测网络中包含该节点所隶属的局部社团机构,下面将对局部社团检测方法进行回顾和总结。

1.1 相关工作

Bagrow等人首先提出了一种检测局部社团的方法^[9],该方法从源节点出发,不断添加连续的外壳上顶点。外壳指的是距离源节点特定距离范围内的节点,特定距离由测地线距离来进行度量。算法的性能取决于初始节点位置,当节点处于社团的边界位置时,算法效果将不会很好。后续研究通过定义指标来进行社团检测,Clauset等人提出了一种基于

R 测度的局部社团检测方法^[10]。Luo等人提出了基于模块度指标 M 的局部社区发现算法LWP^[11]。

近些年,局部社团检测方法开始了飞速的发展,Lancichinetti等人提出了适应度函数 F 指标以度量社区内部和外部的连接密度差异^[12],该算法简单易行,但初始节点的随机选取往往导致社区发现的不稳定,而适应度指标的相关参数也需要预先确定。Raghavan提出了一种基于标签传播(LPA)的社团检测方法^[13],该方法从信息传播的角度出发,每个节点将采用其邻居节点中数目最多的标号进行标记,拥有相同标签的节点归属于同一个社团。Chen等人提出了一种基于局部度中心的社团检测方法LMD^[14],首先发现局部度中心节点,围绕这些节点去发现局部社团,该方法能够达到最佳的识别性能。后续改进的种子集合扩展算法^[15],也取得了性能相当的识别效果。

1.2 问题分析

总结已有方法可知,基于 R 、 M 和 F 的方法对于初始节点的位置敏感,当初始节点位置位于边界时候,其发现效果往往不是很好。LPA方法在更新节点时,对每个标签做相同处理,忽略了节点的亲密程度。LMD方法能够有效地避免社区发现结果对于节点位置的敏感度,该方法从局部度中心节点替代源节点出发去发现局部社团,但局部度中心节点仅仅考虑了最近邻信息,而没有考虑次近邻信息和更远的拓扑信息,同时局部度中心节点的重要性没有进行排序。总之,现有算法的主要缺陷有:①对起始节点的位置敏感,当节点位于社团的边缘部分时,对于社团的发现精度不够;②难以有效利用拓扑信息,仅从邻居信息出发,缺失了部分拓扑信息的利用;③扩展方向难以精确控制,现有算法都是由内而外以单个节点的步骤进行扩展,缺乏有效的扩展指示。

2 局部影响力节点集扩展的方法

本文研究的无权无向网络用 $G(V, E)$ 来表示,其中 V 是网络中的节点集合, E 是网络中节点边的集合,给定源节点 s ,局部社团是指网络 $G(V, E)$ 中含有 s 的完整局部社团结构 C 。

2.1 影响力节点查找

研究者已经提出了很多种节点重要性指标^[16],主要有紧密中心性^[17]、介数中心性^[18]、度中心性^[19]指标等。由于这些指标针对的都是局部信息角度,因此需要提出一种能够有效应对局部社团发现的衡

量指标。基于全局信息的介数中心性指标和紧密中心性指标的算法复杂度较高,且在局部网络中社团的全局结构往往难以获取,不再适用。度指标从局部角度出发,仅仅考虑邻居信息的连接情况,没有对节点周围的环境(例如节点所处的网络位置、更高阶邻居等)进行更深入细致地探讨,因而在很多情况下用度指标来进行中心性测量是不够精确的。

基于对局部信息度指标的改进,LMD 方法考虑了度指标,相当于进行了一阶邻居范围的扩展,对于局部信息的利用效果好于度指标,但其未对邻居信息进行节点排序和局部信息的进一步挖掘。Chen 等人提出了一种半局部的中心性衡量指标^[16],利用了节点的 4 层邻居信息,相比于其他的中心度衡量方法,已经验证该方法能够很好地适用于异质分布的社会网络。本文借鉴文献[16]提出的半局部中心性指标,该指标从源节点的最近邻和次近邻信息出发,能够较为有效地识别出局部信息下节点的影响力,其定义如下。

定义 1 局部影响力指标 给定图 G 中的一个节点 v ,其局部中心性衡量指标 C_v 的定义为

$$C_v = \sum_{u \in \Gamma_v} Q_u$$

(1)

$$Q_u = \sum_{w \in \Gamma_u} N_w$$

(2)

式中: Γ_v 和 Γ_u 分别是节点 v 和节点 u 的最近邻节点集合; N_w 是节点 w 的最近邻节点度数总和。

下面用图 1 构造的网络对该定义的计算进行说明,如图 1 所示,构造了一个含有 20 个节点的网络社团,对该网络中所有节点的度数和影响力大小进行计算,结果如表 1 所示,其中度数最大的节点序号为 20,度数为 7,而最大影响力节点序号为 1。以节点 20 为例,计算的具体方法为:节点 20 包含有 7 个最近邻节点(序号为 13~19)以及一个次近邻节点(序号为 8),所以 $N_{20}=8$ 。由式(1)可知, $Q_{20}=54$,由式(2)可知,其节点中心度大小为 $C_{20}=Q_{13}+Q_{14}+Q_{15}+Q_{16}+Q_{17}+Q_{18}+Q_{19}=124$ 。

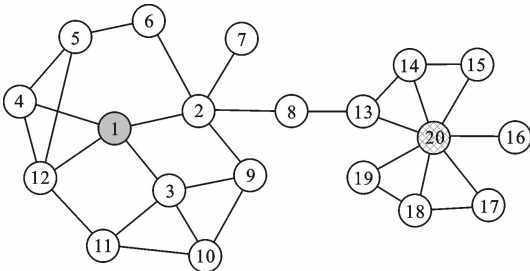


图 1 影响力节点分布示意图

图 1 中所有节点的度值(边数) K_v 、最近邻节点的度数总和 N_v 、次近邻节点度数总和 Q_v 、节点影响力 C_v 值在表 1 中给出。由表 1 可知,度最大的节点(节点序号 20),其中心性不一定为最大,而度数相对较小的节点,由于其位于局部网络的核心位置(节点序号 1),其邻居也有可能具备更多的连接关系,即其节点影响力反而大于高度数节点。该度量方式很好地反映了节点的局部信息影响力。

基于定义 1,本文设计了一种局部影响力节点集 H 构建算法,具体步骤如下:

- (1)初始化局部影响力节点集合 H ;
- (2)根据式(1)和式(2),计算源节点 s 的一阶邻居节点的局部影响力值,将其添加到 H 中,按照取值大小进行排序;
- (3)计算源节点邻居节点的局部影响力值,添加到 H 中,并更新序列。

表 1 图 1 中各节点的影响力值

节点序号	K_v	N_v	Q_v	C_v
1	4	11	33	128
2	5	11	40	105
3	4	7	32	100
4	3	7	25	87
5	3	6	23	73
6	2	8	17	63
7	1	5	11	40
8	2	8	20	64
9	3	8	24	94
10	3	6	22	77
11	3	7	21	85
12	4	8	31	102
13	3	9	24	96
14	3	8	24	92
15	2	7	16	76
16	1	7	8	52
17	2	7	15	74
18	3	7	22	82
19	2	7	15	74
20	7	8	52	124

注:黑体数据为度最小和度最大的节点的各项指标。

相关研究文献^[14]表明,当获取了给定节点附近的影响力节点集后,以该集合为种子候选集合,通过聚类策略围绕这些节点进行聚类,通常能够获得稳定的局部社区结构。因此,给出每个社团中影响力

最大的节点,这些节点具备较高的 C_v 值,能够获取较好的识别效果,这也在实验部分进行了验证。

2.2 局部社团检测

针对局部社团检测,需要引入新的度量指标,改进已有局部社团发现方法^[12,15],本文给出了如下局部社团中所需定义的指标。

定义2 节点适应度 已知某局部社区 C 和待合并的邻接节点 a ,则该节点对社区 C 的节点适应度指标 $f(a, C)$ 表示为

$$f(a, C) = S_{im}(a, C+a) f_{C+a} - S_{im}(a, C-a) f_{C-a} \quad (3)$$

式中: $S_{im}(\cdot)$ 表示子团相似度; f_{C+a} 和 f_{C-a} 表示在社区 C 中加入节点和删除节点 a 的聚类选择策略时的社区适应度值。

式(3)中前后两项的差值最终反映了节点 a 对于社区 C 的适应度贡献值。 $f(a, C)$ 越大,说明节点 a 与社区 C 的节点适应度越大,越有可能隶属于社区 C 。节点适应度表示节点对于某社区的适应度贡献值。

定义3 子团相似度 在网络 $G(V, E)$ 中,定义任意两节点子团 A 与 B 的相似度为

$$S_{im}(A, B) = V_s(A, B) L_s(A, B) \quad (4)$$

式中: $S_{im}(A, B)$ 表示子团 A 和 B 的相似度; $V_s(A, B)$ 表示集合 A 与 B 共享节点在二者所有节点集中所占的比例

$$V_s(A, B) = \frac{V(A) \cap V(B)}{V(A) \cup V(B)} \quad (5)$$

其中 $V(A)$ 表示集合 A 中所有的节点集合。 $L_s(A, B)$ 为采用余弦刻度来衡量的相似度指标

$$L_s(A, B) = \frac{\sum_{k \in N(A) \cap N(B)} E(A, k) E(B, k)}{\left(\sum_{k \in N(A)} E^2(A, k) \right)^{1/2} \left(\sum_{k \in N(B)} E^2(B, k) \right)^{1/2}} \quad (6)$$

式中: $E(A, k) = \sum_{i \in A} \omega_{ik}$ 表示集合 A 与某节点 k 的所有连接的边权之和; $L_s(A, B)$ 表示集合 A 与 B 共享连边的权重总和与二者所有边权重总和比值的归一化。

子团相似度用于将算法中不同的核心节点子团进行合并的判定。节点子团的相似性一般有2种衡量指标:节点重合度和连接相似性。结合上述2种指标,本文提出了一个子团相似度综合指标 $S_{im}(A, B)$ 。 $S_{im}(A, B)$ 的定义说明:当子团 A 和 B 地重合的节点越多,共有连接在二者的所有直接连接中的

比重越大,则说明二者联系越紧密,子团相似性越强,越有可能属于同一社区。

基于上述定义,局部社团检测算法包含2个阶段,且这2个阶段随着每一层的局部社团发现过程进行迭代计算,即社团的影响力节点是迭代构造,根据迭代处理的结果,重新构建和排序原有的影响力节点集合,直到局部社团指标不再满足,终止整个局部社团检测过程。首先根据2.2节算法1构造初始影响力节点集合 H ,然后基于影响力节点集合进行社团发现,主要步骤如下:

(1)初始化源节点 v_0 和 H 中的节点为原始社团集合,构造种子社团集 $C(s) = \{C_0, C_1, C_2, \dots, C_l\}$;

(2)从距离源节点 v 最近的且位于 H 中的节点依次开始扩张社团,如果不同节点的 C_v 一致,则随机选择节点进行扩展;

(3)根据2.3节提出的式(3),计算扩展节点对于其邻近节点子团的节点适应度,若 $f_{C(i)}^i > 0$,则将该节点加入对应的核心节点子团,得到更新后的 $C(i)$;

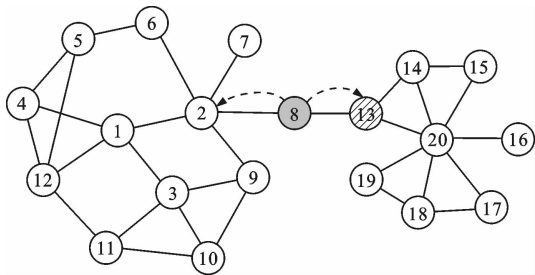
(4)根据式(4),将不同的核心节点获取的子团与给定节点所属的核心节点子团进行相似度判决,若相似度大于门限阈值 ξ (通常设定为0.5),则合并不同的节点子团,并更新社团集合 $C(s)$;

(5)重复步骤3和4,以遍历所有的核心节点子团,得到最终的给定节点 s 隶属的局部社区 C 。这2个阶段随着社团检测的进行是重复迭代的更新过程,指导所有归属于局部社团的节点添加进来为止。

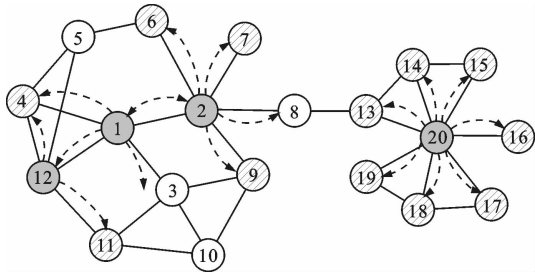
以图1中构造的网络为例,若给定的源节点序号为8,传统的局部社团检测方法是从节点8直接开始扩展社团,如图2a所示,首先将8的邻居加入到待发现社团中去,即节点2和13,然后考虑节点2和13的邻居。与之前的方法不同,本文方法的局部社团扩展过程如图2b所示,更新并找到局部影响力节点集合 $H = \{1, 20, 2, 12, \dots\}$,然后以 H 中的节点排序后的顺序代替原始节点进行扩展,能够较为快速和稳定地发现包含节点8的局部社团。

2.3 复杂度分析

节点集合的构造是动态更新过程,并非一次性完全构造完毕。首次构造影响力节点集合是从源节点的直接相连的邻居节点出发,构造最具影响力节点集合。添加进该集合中最具影响力节点之后,在新添加节点的邻居中重新计算并获取新的最具影响力集合节点,然后将新得到的集合添加到之前总的



(a)传统的局部社团扩展方法



(b)本文提出的社团扩展方法

图 2 社团扩展方法对比示意图

最具影响力集合,更新和排序总影响力节点集合。根据影响力大小排序,再次将新的节点合并入局部社团中去,反复迭代上述过程。本文方法的计算量主要包括 2 个方面:①单次迭代过程中,邻居节点的个数和影响力计算;②迭代的次数。在单次迭代中,由于计算的仅仅是单个节点的邻居,因此需要参与运算的节点数量很小,该部分运算复杂度也较低。第 2 部分运算量,迭代次数与网络中局部社团规模有关,无论何种局部社团检测算法,都需要遍历该局部社团所包含的节点。综上所述,本文所提方法的复杂度与其他的局部社团复杂度相比,相差不大。

3 实 验

为评估本文方法的有效性,在人工网络和真实网络中分别比较了 Clauset 方法^[10]、LWP 方法^[11]和 LMD 方法^[14]这 3 个局部社团发现算法与本文方法的性能。这 3 个算法中,Clauset 最早提出了局部信息下社团检测的概念,LWP 定义了局部模块速度 M 进行社团发现,LMD 是近 2 年性能较好的基于局部度中心节点的社团发现,LMD 和 IN-LCD 算法都是从原始节点进行跳转到最具影响力的节点子团进行扩展。在实验中,所采用的网络中的社团结构均为已知,具体信息详见下面小节的数据介绍。初始节点为随机节点,重复随机做实验 20 次,根据网络规模不同,初始节点数目为原始网络规模的 1/10 左右,且对于所有不同初始节点所得到的结果进行

求平均,以保证算法在实际结果中最终各项指标的准确性。

3.1 评价指标

为了评估不同算法的识别性能,本文使用了在局部社团发现中应用较为广泛的准确率(P_{re})、查全率(R_{ec})和综合指标(F)作为评估准则^[19]。其定义如下

$$P_{re} = \frac{|C_D \cap C_R|}{|C_D|} \tag{7}$$

$$R_{ec} = \frac{|C_D \cap C_R|}{|C_R|} \tag{8}$$

$$F = \frac{2P_{re}R_{ec}}{P_{re} + R_{ec}} \tag{9}$$

式中: C_D 是算法检测到的局部社团; C_R 是真实的包含源节点的局部社团。

由上述 3 个式子可知,准确率指的是正确划分的节点在所发现社团中的比例,查全率指的是正确划分的节点在真实网络中的比例, F 是二者的折中。性能较好的算法,应该具备较高的 F 。

3.2 人工网络

采用最常见的 LFR 网络作为人工合成网络来进行实验^[20],根据生成网络的节点数(1 000, 10 000)不同,将网络分为 N_1 和 N_2 2 类。在每一类网络中,节点的平均度(K_a)、最大度(K_m)、最小社团尺寸(C_{mi})和最大社团尺寸(C_{ma})的设置不同的取值, m_u 是网络中所有节点的外部度和内部度的比例, m_u 越大表示社团结构越不明显,其取值大小为 0.1~0.9,间隔为 0.1。因此,每一类网络包含有 9 组不同 m_u 值的网络。网络参数如表 2 所示。

表 2 LFR 网络设计参数

网络集合	N	K_a	K_m	C_{mi}	C_{ma}	m_u
N_1	1 000	20	50	50	200	0.1~0.9
N_2	10 000	40	100	50	200	0.1~0.9

在这 2 类网络上对 4 种算法分别进行实验,图 3 和图 4 是 4 种算法在网络节点为 1 000 和 10 000 时 F 指标随着 m_u 增大的结果。根据 LFR 数据上对 4 种算法进行的实验表明,随着网络规模的增大,各个算法的识别性能均呈现下降趋势。在同样的网络规模下,随着 m_u 值的增加,网络的结构越来越不明显,各个算法的识别性能均有所下降,但无论在哪种情况下,本文方法在识别性能上均比其余 3 种算法要好,这也验证了本文方法的有效性。

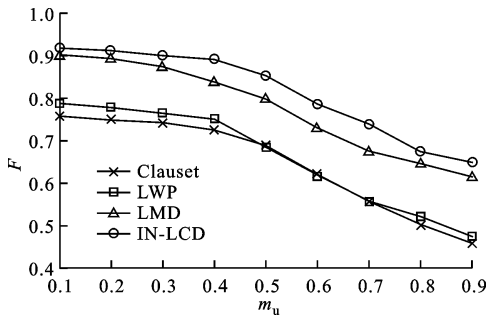


图 3 网络节点数为 1 000 时 4 种方法的实验结果

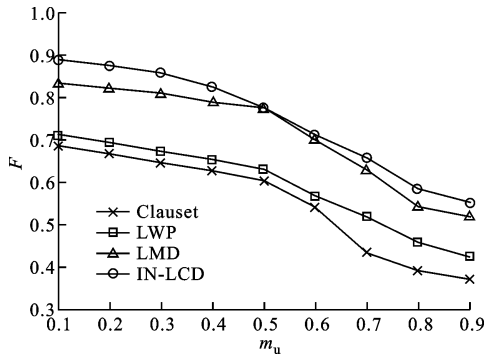


图 4 网络节点数为 10 000 时 4 种方法网络上的实验结果

3.3 真实网络

为了进一步验证本文方法的有效性,与其他算法在 3 种真实网络上进行了实验。这 3 个网络分别是搏击俱乐部网络(Karate)^[21]、美国大学橄榄球网络(Football)^[1]和在线社交网络社团网络(Facebook)^[15],网络的基本情况如表 3 所示。

表 3 真实网络基本信息

网络	节点数	边数	社团数
Karate	34	78	12
Football	115	616	2
Facebook	4 039	88 234	193

针对真实网络数据集的准确率 P_{re} 、查全率 R_{ec} 和综合测度 F 结果如表 4~表 6 所示。可以看出,在不同结构的网络中,算法的识别率和查全率相差很大,由于真实网络的社团结构较为复杂,因此对于算法鲁棒性的要求更高,综合 3 种指标,应以 F 的值来考量最终的算法性能。由表 4~表 6 可知,本文方法在不同的真实网络中均同样表现了较高的算法性能。分析可知:Clauaset 和 LWP 算法对于给定节点的位置敏感;LMD 算法仅仅考虑了节点的最近邻信息,所以当网络规模和网络中节点度数和尺寸的异质分布增大时,其算法有效性也降低;本文 IN-

LCD 方法由于确定了最近邻和次近邻节点的影响力大小,能够按照影响力大小次序从多个节点进行社团扩张,因此取得了较好的识别效果。

表 4 4 种方法在真实网络中的准确率比较

网络集	P_{re}			
	Clauset	LWP	LMD	IN-LCD
Karate	0.891	0.891	0.870	0.906
Football	0.810	0.775	0.775	0.845
Facebook	0.895	0.680	0.904	0.970

表 5 4 种方法在真实网络中的查全率比较

网络集	R_{ec}			
	Clauset	LWP	LMD	IN-LCD
Karate	0.740	0.806	0.890	0.890
Football	0.402	0.435	0.782	0.889
Facebook	0.276	0.560	0.615	0.615

表 6 4 种方法在真实网络中的综合指标比较

网络集	F			
	Clauset	LWP	LMD	IN-LCD
Karate	0.809	0.846	0.880	0.898
Football	0.537	0.557	0.782	0.866
Facebook	0.422	0.614	0.732	0.753

综合人工网络和真实网络数据集可知,IN-LCD 方法与已有的最佳局部社团检测方法相比, F 值更高,识别性能提升了大约 5.3%,即社团检测性能更佳。分析可知,本文方法由于充分利用了局部条件下的网络拓扑信息,能够很好地应对真实网络中节点度和社团尺寸的异质分布的情况,因此具备较高的识别性能,即能够从局部的角度出发,更为精确地检测局部社团结构。

4 结 论

针对局部社团检测中的初始位置敏感和扩展方向不确定的问题,本文提出了一种基于节点影响力集扩展局部社团检测方法,由于从最具影响力的节点集合出发能够较快且准确地发现局部社团的节点归属,因此本文所提局部社团检测方法有效地避免了原有方法对初始节点敏感的问题。与已有局部社团检测算法相比,IN-LCD 方法具有更高的识别性能,且能够很好地适用于当前在线网络全局信息难以获取的处理情况。下一步的工作将聚焦于多尺度

和重叠情况下局部社团检测的研究。

参考文献:

- [1] GIRVAN M, NEWMAN M E. Community structure in social and biological networks [J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 99(12): 7821-7826.
- [2] 杨建伟, 桂小林, 安健, 等. 一种信任关系网络中的社团结构检测算法 [J]. 西安交通大学学报, 2014, 48(12): 80-86.
YANG Jianwei, GUI Xiaolin, AN Jian, et al. Community structure detecting in trust relationship networks [J]. Journal of Xi'an Jiaotong University, 2014, 48(12): 80-86.
- [3] FORTUNATO S. Community detection in graphs [J]. Physics Reports, 2009, 486(3/4/5): 75-174.
- [4] NEWMAN M E. Fast algorithm for detecting community structure in networks [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2004, 69(6): 066133.
- [5] 陈国强, 王宇平. 采用离散粒子群算法的复杂网络重叠社团检测 [J]. 西安交通大学学报, 2013, 47(1): 107-113.
CHEN Guoqiang, WANG Yuping. Overlapping community detection of complex networks based on discrete particle swarm algorithm [J]. Journal of Xi'an Jiaotong University, 2013, 47(1): 107-113.
- [6] 卫红权, 陈鸿昶, 刘力雄, 等. 基于强度排序的通信社区检测算法 [J]. 通信学报, 2014, 35(10): 165-170.
WEI Hongquan, CHEN Hongchang, LIU Lixiong, et al. Communication community detection algorithm based on ranking of strength [J]. Journal on Communications, 2014, 35(10): 165-170.
- [7] COSCIA M, ROSSETTI G, GIANNOTTI F, et al. DEMON: a local-first discovery method for overlapping communities [C]//Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2012: 615-623.
- [8] AARON C, NEWMAN M E J, CRISTOPHER M. Finding community structure in very large networks [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2004, 70(6): 264-277.
- [9] BAGROW J P, BOLLT E M. Local method for detecting communities [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2005, 72(4): 046108.
- [10] AARON C. Finding local community structure in networks [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2005, 72(2): 254-271.
- [11] LUO F, WANG J Z, PROMISLOW E. Exploring local community structures in large networks [C]//Proceedings of 2006 IEEE International Conference on Web Intelligence. Piscataway, NJ, USA: IEEE, 2006: 233-239.
- [12] LANCICHINETTI A, FORTUNATO S, KERTESZ J. Detecting the overlapping and hierarchical community structure in complex networks [J]. New Journal of Physics, 2009, 11(15): 19-44.
- [13] USHA NANDINI R, REKA A. Near linear time algorithm to detect community structures in large-scale networks [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2007, 76(3): 233-239.
- [14] CHEN Q, WU T T, FANG M. Detecting local community structures in complex networks based on local degree central nodes [J]. Physica: A Statistical Mechanics & Its Applications, 2013, 392(3): 529-537.
- [15] MENG F, ZHU M, ZHOU Y, et al. Local community detection in complex networks based on maximum cliques extension [J]. Mathematical Problems in Engineering, 2014, 2014(1): 1-12.
- [16] CHEN D, LV L, SHANG M S, et al. Identifying influential nodes in complex networks [J]. Fuel & Energy Abstracts, 2012, 391(4): 1777-1787.
- [17] FREEMAN L C. Centrality in social networks conceptual clarification [J]. Social Networks, 1978, 1(3): 215-239.
- [18] FREEMAN L C. A set of measures of centrality based on betweenness [J]. Sociometry, 1977, 40(1): 35-41.
- [19] BONACICH P. Factoring and weighting approaches to status scores and clique identification [J]. Journal of Mathematical Sociology, 1972, 2(1): 113-120.
- [20] ANDREA L, SANTO F, FILIPPO R. Benchmark graphs for testing community detection algorithms [J]. Physical Review: E Statistical Nonlinear & Soft Matter Physics, 2008, 78(4): 561-570.
- [21] ZACHARY W W. An information flow model for conflict and fission in small groups [J]. Journal of Anthropological Research, 1977, 33(4): 473-480.

(编辑 刘杨)