

DOI:10.7652/xjtuxb201408013

采用链路聚类的动态网络社团发现算法

董哲, 伊鹏

(国家数字交换系统工程技术研究中心, 450002, 郑州)

摘要: 针对当前基于节点的动态网络社团结构发现算法难以发现稳定的社团结构的问题,提出了一种采用链路聚类的动态网络社团发现算法(LDC)。该算法首先从链路的角度得到网络的链路图结构;然后对比不同时刻的链路图结构,将动态网络中节点的添加与移除以及边的添加与移除等复杂的变化信息简化为链路添加和链路移除 2 种增量变化信息;再在前一时刻社团结构的基础上以改进的链路划分密度函数对增量变化信息中变化的链路进行处理,判断该链路是否加入到社团中从而得到最优的社团结构;最后将得到的链路社团转化成为最终的节点社团结构。实验结果表明,相比于当前基于节点的动态社团发现算法,LDC 算法能够有效地发现网络中结构稳定的社团结构,其模块度值和标准化互信息值至少提高了 0.19 和 0.13,且算法的运行效率要明显优于基于节点的动态社团发现算法。

关键词: 链路聚类;增量方法;社团发现;动态网络

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2014)08-0073-07

A Community Detection Algorithm for Dynamic Networks Using Link Clustering

DONG Zhe, YI Peng

(National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, China)

Abstract: A community detection algorithm for dynamic networks is proposed to overcome the limitation that the current node-based dynamic community detection algorithm is difficult to identify the stable community structure. The algorithm uses link clustering technique and gets a link graph structure of the network, and then the complex incremental information in the dynamic network such as addition and removing of nodes and edges are simplified into addition and removing of links. An improved link partition density function is proposed to process a link in the incremental information and decide whether the link should be joined into the community based on the existing community structure to get the optimal community structure. At the end, the algorithm transforms the optimal link community structure into a node-based community structure. Experimental results and comparisons with the node-based community detection algorithm show that the algorithm can get the stable community structure and the modularity and NMI can raise at least 0.19 and 0.13, respectively, and that the efficiency of the algorithm is superior to the dynamic node-based community detection algorithms.

Keywords: link clustering; increment method; community detection; dynamic network

收稿日期: 2014-01-03。 作者简介: 董哲(1989—),男,硕士生;伊鹏(通信作者),男,教授。 基金项目: 国家“973 计划”资助项目(2012CB315901,2013CB329104);国家“863 计划”资助项目(2011AA01A103,2011AA01A101);国家科技支撑计划资助项目(2011BAH19B01)。

网络出版时间: 2014-06-18

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20140618.1138.010.html>

社团作为复杂网络的重要特性受到了研究者的日益关注。社团将网络划分成许多由节点组成的群组,群组内部节点连接非常紧密而群组之间连接则较为稀疏。发现复杂网络中的社团结构具有很多重要的应用,如恐怖组织识别、网络舆情监控以及谣言的传播行为分析等。在动态网络中,由于其结构更加复杂,因此挖掘动态网络中的社团结构变得异常困难。

目前,关于复杂网络中的社团结构的研究已经取得了很多成果^[1-3]。在现实生活中,绝大多数的网络都具有动态特性,这些网络随着时间的推移不断发生演化。虽然,这些变化对于网络的局部结构产生的影响较小,但随着时间积累网络中的社团结构会发生明显的变化,因此需要对其社团进行重新发现。目前,已经提出了许多动态网络社团发现算法。如 Tang 等人提出一种类似于语义迭代分析过程的算法^[4],发现动态多模网络中社团演化情况,该算法能够得到局部最优的社团结构,但是却需要人为指定社团个数,不能自适应地得到最优的社团结构。文献[5]提出了分析动态网络中社团的 Social Cost 算法,将社团发现转化为图着色问题,但该算法只关注特定个体的演化关系,不能从全局的角度来得到社团结构。还有其他一些动态社团发现算法,如: Gong 等人提出一种基于局部搜索的多目标免疫算法^[6],但该算法复杂度非常大; Nguyen 等人提出一种自适应增量算法 QCA 算法^[7],该算法只对发生变化的节点进行处理,能够得到较优的社团,但节点的变化情况复杂多变,导致算法的复杂度较大;文献[8]提出一种主题模型来发现网络中的社团结构,但是其算法的复杂度较大;文献[9]提出一种简单的增量社团发现算法,该算法复杂度较小,但是其算法精度较差;文献[10]提出了一种基于优化算法的动态社团发现方法,该算法避免了网络中离群点对算法的影响,提高了算法精度,但是其复杂度较大。

大多数社团结构发现方法是基于节点聚类的算法,难以发现稳定的社团。在网络中当节点存在交互关系时认为两节点间存在链路且唯一。因此,文献[11]提出一种基于链路的 Link 算法,该算法发现网络中稳定的社团具有明显的优势。算法首先对网络中的链路进行聚类,将得到的链路社团通过节点链路关系映射为节点社团^[12],但该算法无法发现动态网络中的社团结构。

针对上述问题,本文提出一种基于链路聚类的增量社团发现算法(link-based dynamic community

detection algorithm, LDC)。该算法通过各个时刻的网络结构信息得到相关链路增量信息,再基于改进的链路划分密度函数(LD)对链路增量信息进行处理,以改进的链路模块度为目标函数,得到链路社团。为了验证算法的有效性,本文将其应用于人工动态网络和真实的动态网络中。实验证明,本文算法能够有效地发现具有典型动态信息网络中的多尺度稳定社团结构。

1 相关定义与指标

给定网络 $G(V, E)$, V 为节点集合, E 为边集合, C_i 为第 i 个社团节点集合。给出以下定义:

定义 1 节点时序图 S_t 表示 t 时刻个体间的关系, 动态网络 G 可以用节点时序图的集合表示, T 为时间段划分总数, $G = (S_1, S_2, S_3, \dots, S_t, \dots, S_T)$, $1 \leq t \leq T$, $\bigcup_{1 \leq t \leq T} S_t = G$ 。

定义 2 给定 S_t , 则链路图 L_G 是 S_t 中节点间边的关系构成的链路集合, 如图 1 所示。 $L_G = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ 。

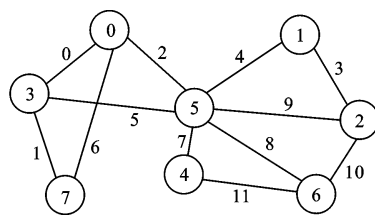


图 1 链路图

定义 3 给定链路 L_1 和 L_2 , $L_1 = \langle v_1, v_2 \rangle$, $L_2 = \langle v_3, v_4 \rangle$, 其中 v_1, v_2, v_3, v_4 为节点。若 $L_1 \cap L_2 = \langle v_1, v_2 \rangle \cap \langle v_3, v_4 \rangle = \emptyset$, 则链路 L_2 与 L_1 互为邻链路。

定义 4 给定 t 时刻链路社团 C_i, C_j 和链路 $L = \langle v_1, v_2 \rangle$, 若 $\{L | L \notin C_i, L \notin C_j, v_1 \in C_i, v_2 \in C_j, i \neq j\}$ 则 L 为桥链路。

1.1 链路划分密度

本文通过文献[11]中的链路划分密度 D 函数提出一种改进的链路划分密度函数 D_L 处理增量信息。假定一个网络有 M 条链路, $\{P_1, \dots, P_C\}$ 分别将网络划分成 C 个链路子集。定义社团的链路划分密度

$$D_c = [m_c - (n_c - 1)] / \left[\frac{n_c(n_c - 1)}{2} - (n_c - 1) \right] - [m_b - (n_b - 1)] / \left[\frac{n_b(n_b - 1)}{2} - (n_b - 1) \right] \quad (1)$$

式中: $m_c = |P_c|$ 和 $n_c = \left| \bigcup_{e_{ij} \in P_c} \{i, j\} \right|$ 为子集 P_c 中

链路数和节点数; m_b 表示社团之间的桥链路数; n_b $= |\bigcup_{e_{ij}, i \in P_C, j \in P_C} \{i, j\}|$ 表示社团之间的节点数。改进

的链路划分密度 $D_L = \sum_C \frac{m_C}{M} D_C$, 则

$$D_L = \frac{2}{M} \sum_C \left[m_C \frac{m_C - (n_C - 1)}{(n_C - 2)(n_C - 1)} - m_C \frac{m_b - (n_b - 1)}{(n_b - 2)(n_b - 1)} \right] \quad (2)$$

1.2 链路模块度

给出矩阵 $\mathbf{H}_{N \times M}$, N 为网络中节点总数, M 为网络中链路总数。如果链路 l 与节点 i 相关 ($i \in l$), 则 $\mathbf{H}$ 中的元素 $H_{il} = 1$, 否则为 0。令节点度 $k_i = \sum_l H_{il}$, $k_l = \sum_i H_{il}$, 表示与链路 l 相关的节点数, 则模块度为

$$Q_L = \frac{1}{W} \sum_{R \in \mathbf{P}} \sum_{l, m \in (1, |\mathbf{R}|)} \left[R_{lm} - \frac{k_l k_m}{W} \right] \quad (3)$$

式中: $\mathbf{R}$ 为链路邻接矩阵, 且 $R_{lm} = \sum_i H_{il} H_{im} (1 - \delta_{lm})$; $|\mathbf{R}|$ 为 $\mathbf{R}$ 的维数; $W = \sum_{l, m} R_{lm}$; $\mathbf{P}$ 为储存基于链路的社团划分结构所对应的邻接矩阵。

2 LDC 动态社团发现算法

给定网络 $G = (V, E)$, $C = \{C_1, C_2, \dots, C_k\}$, 令动态网络中的增量信息为 $\epsilon = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$ 。网络中的链路可分为 2 类: 社团内部的链路 (intra-community links, IL), 即边的 2 个端点均在该社团内; 社团之间的链路 (bridge-community links, BL), 即边的 2 个端点位于 2 个不同的社团内。对于 G 中的每一个社团 C , 当添加 IL 链路或者移除 BL 链路时使得社团的紧密性更强且网络结构更加清晰。相反, 移除 IL 链路和添加 BL 链路使得网络结构渐渐模糊。当 2 个社团之间不存在干扰关系, 或者干扰较小时, 链路的增加或者移除有可能形成一个新的社团结构。因此, 在更新社团结构时, 网络结构的细微变化将会导致其社团产生巨大的变化。从链路的角度出发, 通常随着时间的推移网络的变化其实就是链路的增加或移除。因此, 网络中的变化信息就是一系列简单事件的集合, 这些简单的事件包括新的链路的加入和已经存在的链路的移除等信息。

2.1 链路添加

当新增加一条链路 e 时, 有 2 种情况: ① 链路 e 完全处于社团 C_i 内部; ② 链路 e 处于社团 C_i 和 C_j 之间, $i \neq j$ 。对于情况①, 根据定理 1, 社团结构保持

不变。对于情况②, 由定理 2 可知, 如果桥链路 e 被划分进新的社团中, 则该社团必为 C_i 和 C_j 之一。推论 1 给出了对桥链路 e 进行划分的判断条件, 如果链路 i 或者 j 对应的链路划分密度变化值 Δd_i 和 Δd_j 均不满足, 则不对该链路进行操作, 具体见算法 1。

算法 1 添加边 new-link

步骤 1 输入新加入的链路 e 和 t 时刻链路社团结构 C_t ;

步骤 2 输出 $t+1$ 时刻链路社团结构 C_{t+1} ;

步骤 3 若 e 是社团内部链路, 则 $C_{t+1} \equiv C_t$, 否则 $k = \arg \max(\Delta d_i, \Delta d_j)$ 并将 e 加入到 C_k , 更新 C_{t+1} 。

定理 1 如果 C_i 是网络 G 中一个社团, 那么添加任何一条 IL 链路到 C_i , C_i 不会分解成更小的模块。

证明 假设增量信息 ϵ_i 表示要添加一条内部链路 e 到社团 C_i 中, 则 C_i 的链路划分密度为

$$D_C = \frac{m_C - (n_C - 1)}{n_C(n_C - 1)/2 - (n_C - 1)} - \frac{m_b - (n_b - 1)}{n_b(n_b - 1)/2 - (n_b - 1)} \quad (4)$$

令 D'_C 表示 e 加入到社团 C_i 时的链路划分密度, 则

$$D'_C = \frac{(m_C + 1) - (n_C - 1)}{n_C(n_C - 1)/2 - (n_C - 1)} - \frac{m_b - (n_b - 1)}{n_b(n_b - 1)/2 - (n_b - 1)} \quad (5)$$

令 $\Delta = D'_C - D_C$, 则 $\Delta > 0$, $D'_C > D_C$, 故当加入内部链路 e 到 C_i 中后, 社团更稳健。

定理 2 如果添加的链路位于社团 C_i 和 C_j 之间, 若该桥链路要被重新划分, 则社团 C_i 和 C_j 为首选。

证明 假设要添加的位于社团 C_i 和 C_j 之间的链路为 e , 则由于 e 的节点分别位于社团 C_i 和 C_j 中, 所以当把链路 e 加入到其他社团时, D_L 值并不变。对社团 C_i , 链路 e 未加入时, 相应链路前后对应的链路划分密度为

$$D_{L,i} = \frac{m_C - (n_C - 1)}{n_C(n_C - 1)/2 - (n_C - 1)} - \frac{m_b - (n_b - 1)}{n_b(n_b - 1)/2 - (n_b - 1)} \quad (6)$$

当加入链路 e 时, 其链路划分密度变为

$$D'_{L,i} = \frac{m_C - (n_C - 1)}{n_C(n_C - 1)/2 - (n_C - 1)} - \frac{(m_b + 1) - n_b}{(n_b + 1)n_b/2 - n_b} \quad (7)$$

令 $\Delta_1 = D'_{L,i} - D_{L,i}$, 则 $\Delta_1 > 0$, 故链路 e 加入社团 C_i 之后, C_i 的链路划分密度增大。同理, 令 $\Delta_2 = D'_{L,j} - D_{L,j}$, $\Delta_2 > 0$, C_j 的链路划分密度也增大。

综上所述, 如果添加的链路位于社团 C_i 和 C_j

之间时,则社团 C_i 和 C_j 为的首选。

推论 1 如果添加的桥链路 e 位于社团 C_i 和 C_j 之间,若满足条件 $\Delta_d = D_{L,i}(E+e) - D_{L,j}(E+e) + D_{L,j}(E) - D_{L,i}(E) > 0$,则桥链路 e 将被划分进社团 C_i 中,反之,桥链路 e 将被划分进社团 C_j 中。

证明 由定理 2 可知,如果添加的桥链路 e 位于社团 C_i 和 C_j 之间, C_i 和 C_j 为其首选,则有

$$\begin{aligned} \Delta_d &= \Delta_1 - \Delta_2 = (D'_{L,i} - D_{L,i}) - (D'_{L,j} - D_{L,j}) = \\ &D_{L,i}(E+e) - D_{L,j}(E+e) + D_{L,j}(E) - D_{L,i}(E) \end{aligned} \quad (8)$$

当 $\Delta_d > 0$ 时, $\Delta_1 > \Delta_2$,则桥链路 e 应被划分进社团 C_i 中;当 $\Delta_d < 0$ 时, $\Delta_1 < \Delta_2$,则桥链路 e 应被划分进社团 C_j 中。

2.2 链路移除

当某链路 e 被移除时,可分为 2 种情况:①该桥链路 e 处于社团 C_i 和 C_j 之间, $i \neq j$;②该链路 e 完全处于社团 C_i 内部。依据推论 3,对于情况①,移除桥链路时,社团不发生变化。

推论 3 如果链路 e 为社团 C_i 和 C_j 之间的桥链路,当移除该链路时社团 C_i 和 C_j 结构更加明显,整体社团结构不变。

证明 当移除的链路 e 是社团 C_i 和 C_j 之间的桥链路时,社团内部节点间的链路关系未发生变化,而链路社团之间的链路被移除时,社团间的连接关系变得更稀疏,网络中的社团结构更加健壮、明显。因此,整体社团结构不会发生变化。对于情况②,当移除的链路 e 是一条 IL 链路时,令 $S(e)$ 表示 e 的邻链路集合, $\forall l \in S(e)$,若 $C_k = \arg\max(D_{L,k}(l))$,则将链路 l 划分进社团 C_k 中,其中 $1 \leq k \leq N$, N 为当前时刻的链路社团总数,具体见算法 2。

算法 2 移除边 removal_link

步骤 1 输入被删除的链路 e 和 t 时刻链路社团结构 C_t ;

步骤 2 输出 $t+1$ 时刻链路社团结构 C_{t+1} ;

步骤 3 若链路 e 是一条 BL 链路,则 $C_{t+1} \equiv C_t$,否则 $C_k = \arg\max(D_{L,k}(l))$, $l \in S(e)$, $k \in (1, N)$,并将链路 l 加入到社团 C_k ,更新 C_{t+1} 。

算法 3 给出了本文 LDC 算法的具体步骤。

算法 3 LDC 算法

步骤 1 输入 $G = G_0 = (V_0, E_0)$,增量信息 $\epsilon = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$;

步骤 2 输出 t 时刻网络 G_t 的社团结构 C_t ;

步骤 3 将节点图 G_0 转化为链路图;

步骤 4 发现初始时刻链路图社团结构 C_0 ;

步骤 5 从初始时刻起,若 $e \in \text{new_link}(L(u, v))$,则 $\text{new_link}(C_t, L(u, v))$,否则 $\text{removal_link}(C_t, L(u, v))$;

步骤 6 将链路社团结构 C_t 映射成为节点社团,得到各个时刻节点社团结构。

2.3 LDC 算法的时间复杂度分析

如 LDC 算法所示,算法的步骤 5 虽然对所有时刻和变化的链路信息进行扫描,但之后的操作都是以步骤 5 作为判断条件的,如果条件不满足,就不会有后续操作。因此,运算中真正进行操作的遍数是 ϵ_t 的集合 m_t ,即所有与增量信息有关的链路的数目。 m_t 的大小与网络的时刻划分尺度有关,通过统计发现,网络中 m_t 的规模为 $O(m)$, m 为链路数。在步骤 5 添加边操作,每次循环有 $O(1)$ 个单元操作,移除边操作需要 $O(lk)$ 个单元操作,其中 l 为链路的邻链路数, k 为社团个数,因此每一遍需要时间 $O(lk)$ 。综合起来,总的算法时间复杂度为 $O(m \times lk)$ 。

3 仿真实验

为验证 LDC 算法的有效性,本节分别对人工动态网络和真实的动态社会网络数据集进行仿真实验。本文将 LDC 算法分别与 Link 算法^[11]、GaoCD 算法^[12]、QCA 算法^[7]及 MIEN 算法^[13]进行比较。Link 算法从链路的角度发现网络中社团;GaoCD 基于遗传算法,从链路的角度发现静态网络中重叠社团结构;QCA 算法是基于节点的增量动态网络社团发现算法;MIEN 算法通过压缩与解压缩,使用快速模块度算法更新网络中的社团。

3.1 人工网络

采用文献[14]的方法来生成人工数据集。该人工网络数据集共由 128 个节点组成,共 4 个社团,每个社团共有 32 个节点,如图 2 所示。同一社团内的节点对之间存在链路的概率为 p_{in} ,不同社团的节点间存在链路的概率为 p_{out} ,本文令 $p_{in} = 0.1935$, $p_{out} = 0.0208$ 。

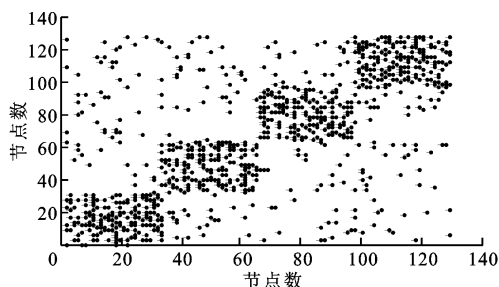


图 2 具有 128 个节点 4 个社团结构的人工网络数据集

为了生成动态人工网络,在第一个时刻之后的每个时刻,随机选择 10% 的节点离开所属的社团,随机地加入到其他 3 个社团中。本文按照此方法生成 10 个时刻的人工网络。

将 LDC 算法应用于该数据集,仿真结果如图 3 所示。由图 3a 可知,在人工动态网络数据集中,本文提出的 LDC 算法其模块度要明显优于其他算法。

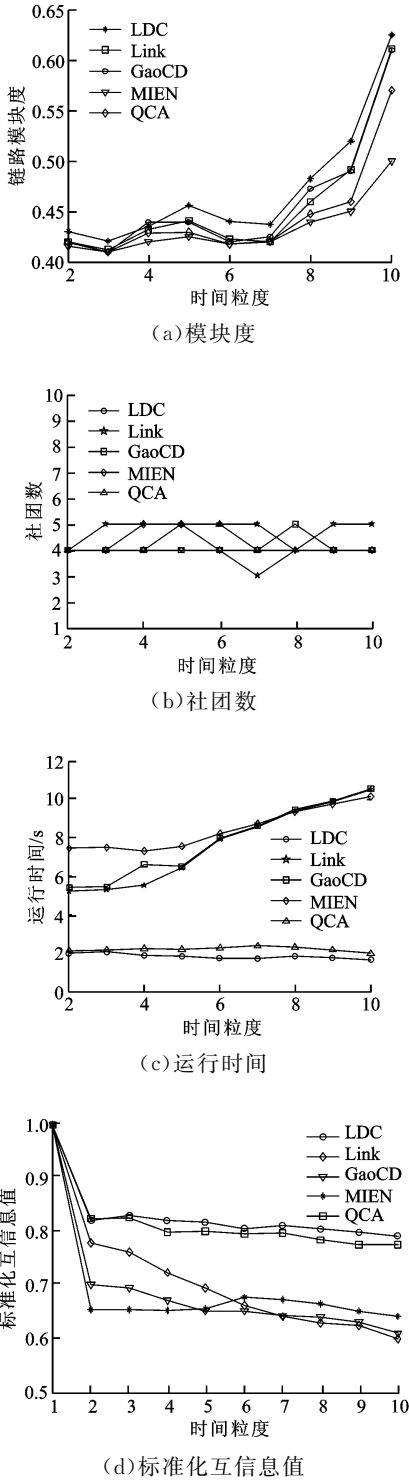


图 3 5 种算法处理人工数据集的性能对比

由图 3b 可以看出,LDC 算法发现的 4 个社团符合实际情况;由图 3c 可以看出,因为网络中的链路增量变化信息较少,所以 LDC 算法的运行时间相比于其他算法要少很多。标准化互信息值用来表征算法发现的社团结构与真实社团之间的差异,用于描述算法的准确性。由图 3d 可以看出,由于 LDC 算法只针对发生变化的链路信息进行处理,符合真实网络中社团的变化情况,所以其标准化互信息值要大于其他算法,提高了 0.15。因此,LDC 算法能够准确地发现网络中的社团结构。

3.2 真实网络

为了验证 LDC 算法在现实社会网络中的性能,本文采用了 2 种真实的社会网络数据集,分别是 ENRON 邮件网络^[5]和 VAST 数据集。ENRON 邮件网络数据集来自 ENRON 公司内部的邮件联系网络,包含来自 151 名用户邮件消息的网络数据集,其时间跨度为 1999 年 5 月到 2002 年 3 月;VAST 数据集来自 IEEE VAST 2008,是一个开放的竞赛项目的数据集,它包含一组共 400 人 10 天的通话数据。

采用 VAST 数据集,以每一天为一个时间粒度,对比本文 LDC 算法和其他 4 种算法处理人工数据集的性能,结果如图 4 所示。

由图 4a 可以看出,相比于其他算法,LDC 算法的链路模块度提高了 20%;由图 4b 可见,LDC 算法中发现 VAST 数据集中的社团更加稳定;由图 4c 可知,LDC 算法的效率较高,算法性能优越;由图 4d 可知,相比于其他算法,LDC 算法的标准化互信息值整体提高了 0.13,LDC 算法利用网络中的链路增量信息,能够发现更加准确的社团结构。

采用 ENRON 数据集,以每个月为一个时间粒度,对比本文 LDC 算法并和其他 4 种算法处理人工数据集的性能,结果如图 5 所示。

由图 5a 可以看出,LDC 算法能够更有效地发现社团结构,相比于其他算法,LDC 算法的链路模块度提高了 19%;由图 5b 中可以看出,相比于其他几种算法,LDC 算法得到的社团数目相对较多,这是由于该公司内部出现了一些变动,导致网络中的链路增量信息不完整导致了社团个数的增加,但同时说明 LDC 算法更倾向于发现具有多尺度、结构更加稳定的社团结构;图 5c 说明,LDC 算法的运行效率要优于其他算法;由图 5d 可知,LDC 算法的 NMI 值相比于其他算法要提高 0.15,且要比其他算法稳定。

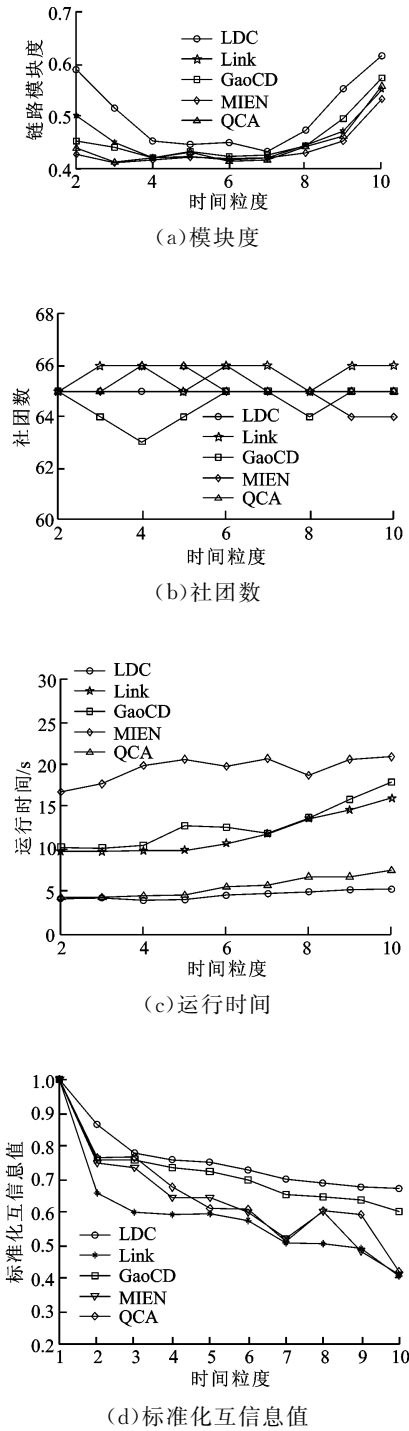


图 4 5 种算法处理 VAST 数据集对比

通过将 LDC 算法分别应用于人工网络和真实网络数据集中并和其他算法进行对比,由仿真结果可知,由于 LDC 算法是从链路的角度来发现动态网络中的社团,只对网络中的链路变化信息进行处理,简化了网络中的增量信息,相比于其他算法,LDC 算法更能够准确地发现网络中稳定的社团结构。

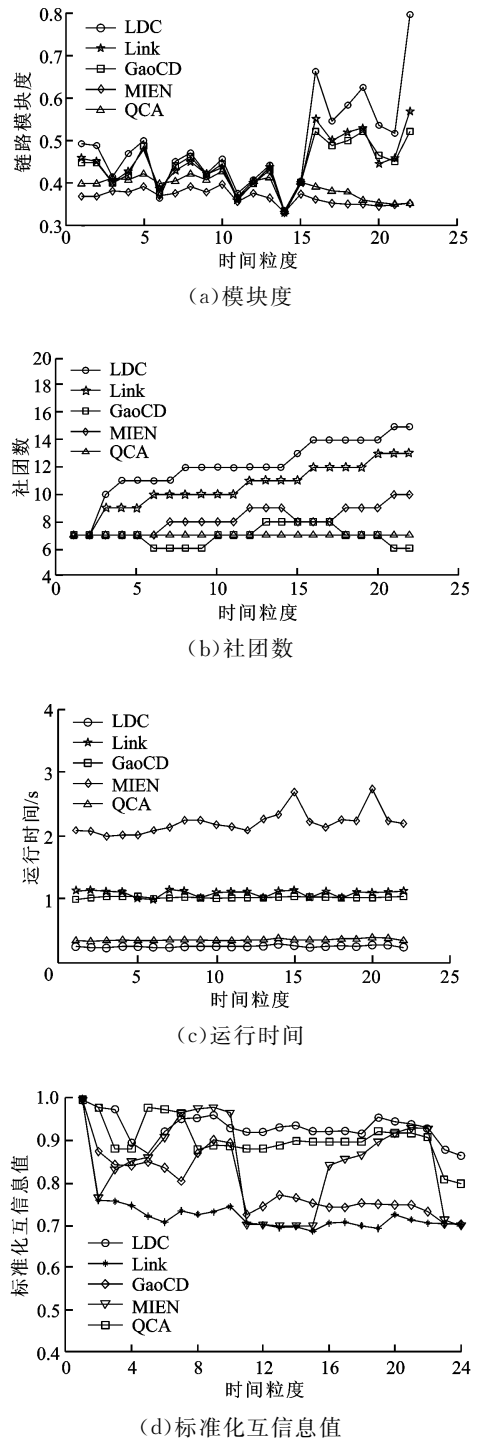


图 5 5 种算法处理 Enron 数据集性能对比

4 结 论

本文提出一种基于链路聚类的动态社会网络社团发现算法——LDC 算法,并提出了一种改进的社团划分密度,将复杂的动态网络中的变化信息简化为链路添加和链路移除 2 种增量信息,从链路的角度

度重新解决社团结构的发现问题。本文分别将LDC算法应用于人工动态网络和真实的动态社会网络中进行仿真实验,结果表明,与当前的动态网络社团发现方法相比,LDC算法能够快速准确地发现动态社会网络中稳定的多尺度社团结构。

参考文献:

[1] 汪小帆. 复杂网络中的社团结构分析算法研究综述 [J]. 复杂系统与复杂性科学, 2008, 5(3): 1-12.
WANG Xiaofan. An overview of algorithms for analyzing community structure in complex networks [J]. Complex Systems and Complexity Science, 2008, 5(3): 1-12.

[2] 李晓佳, 张鹏, 狄增如, 等. 复杂网络中的社团结构 [J]. 复杂系统与复杂性科学, 2008, 5(3): 19-42.
LI Xiaojia. ZHANG Peng, DI Zengru, et al. Community structure in complex networks [J]. Complex Systems and Complexity Science, 2008, 5(3): 19-42.

[3] 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.

[4] TANG L, LIU H, ZHANG J. Identifying evolving groups in dynamic multimode networks [J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(1): 72-85.

[5] TANTIPATHANANANDH C, BERGER-WOLF T, KEMNE D. A framework for community identification in dynamic social networks [C]// Proceedings of the 13th ACM International Conference on Knowledge discovery and Data Mining. New York, USA: ACM, 2007: 717-726.

[6] GONG M G, ZHANG L J, MA J, et al. Community detection in dynamic social networks based on multi-

objective immune algorithm [J]. Journal of Computer Science and Technology, 2012, 27(3): 455-467.

[7] NGUYEN N P, DINH T N, XUAN Y, et al. Adaptive algorithms for detecting community structure in dynamic social networks [C] // Proceedings IEEE INFOCOM. Piscataway, NJ, USA: IEEE, 2011: 2282-2290.

[8] LI D, DING Y, SHUAI X, et al. Adding community and dynamics to topic models [J]. Journal of Informetrics, 2012, 6(2): 237-253.

[9] 单波, 姜守旭, 张硕, 等. IC: 动态社会关系网络社区结构的增量识别算法 [J]. 软件学报, 2009, 20(1): 184-192.
SHAN Bo, JIANG Shouxu, ZHANG Shuo, et al. IC: incremental algorithm for community identification in dynamic social network [J]. Journal of Software, 2009, 20(1): 184-192.

[10] BASSETT D S, PORTER M A, WYMBS N F, et al. Robust detection of dynamic community structure in networks [EB/OL]. [2013-03-14]. <http://dx.doi.org/10.1063/1.4790830>.

[11] AHN Y Y, BAGROW J P, LEHMANN S. Link communities reveal multiscale complexity in networks [J]. Nature, 2010, 466(7307): 761-764.

[12] SHI C, CAI Y, FU D, et al. A link clustering based overlapping community detection algorithm [J]. Data & Knowledge Engineering, 2013, 87(1): 394-404.

[13] DINH T N, XUAN Y, THAI M T. Towards social-aware routing in dynamic communication networks [C] // Proceedings of IEEE 28th International Conference on Performance Computing and Communications. Piscataway, NJ, USA: IEEE, 2009: 161-168.

[14] YANG T, CHI Y, ZHU S, et al. Detecting communities and their evolutions in dynamic social networks: a Bayesian approach [J]. Machine Learning, 2011, 82(2): 157-189.

(编辑 刘杨)

(上接第 72 页)

[13] CAO Lijun, LIU Xiyin, WANG Zhiping, et al. The spatial outlier mining algorithm based on the KNN graph [J]. Journal of Software, 2013, 8(12): 3158-3165.

[14] YAN Sifeng, HAN Jiawei. Span: graph-based substructure pattern mining [C]// Proceedings of the 2002 IEEE International Conference on Data Mining.

Piscataway, NJ, USA: IEEE, 2002: 721-724.

[15] NOBLE C C, COOK D J. Graph-based anomaly detection [C]// Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2003: 631-636.

(编辑 武红江)