

一种低开销的机会网络社区检测算法

任 智, 黄希凯, 谭永银

(重庆邮电大学 移动通信技术重庆市重点实验室, 重庆 400065)

摘 要: 针对 CDTFS(Community Detection algorithm based on Threshold of Familiar Set)算法在社区检测阶段中存在通信冗余开销的问题,提出了一种低开销机会网络社区检测算法—NSDMDT(Network Structure Detection Mechanism based on Dynamic Threshold). NSDMDT 算法优化了新节点加入朋友集合时子图的发送顺序和社区关系子图更新机制,减少社区检测时网络中的通信开销,进一步提高了网络的性能. 理论分析和仿真结果表明在节点通信半径不同、网络中节点个数不同以及节点移动速率不同的三种情况下,NSDMDT 算法网络开销均要低于 CDTFS 算法,同时在与朋友节点的累积相遇持续时间所占比例上也优于 CDTFS 算法.

关键词: 机会网络;社区检测;子图;开销

中图分类号: TP393

文献标识码: A

文章编号: 1000—7180(2016)09—0060—04

A Low-overhead Community Detect Algorithm in Opportunistic Networks

REN Zhi, HUANG Xi-kai, TAN Yong-yin

(Chongqing Key Laboratory of Mobile Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: To solve the existing problem of redundant communication overhead of the community detection in the CDTFS algorithm, and a novel routing algorithm called NSDMDT (Network Structure Detection Mechanism based on Dynamic Threshold) is proposed. The algorithm optimizes the sending order when a new node joining familiar set and the mechanism of updating community relationship sub graph, reduces the network overhead in the network community detection and improves the performance of the network. Theoretical analysis and simulation results show that NSDMDT algorithm outperforms an existing CDTFS algorithm in terms of control overhead when the communication radius is not the same or when the number of nodes in the network is different, and the ratio of cumulative duration with friend nodes.

Key words: opportunistic networks; community detection; sub graph; overhead

1 引言

机会网络是一种特殊的移动自组织网络,它不要求网络时刻保持连通,能够在网络链路处于间歇性连接的状态下,为通信提供端到端的服务,并且具有网络拓扑分割以及变化快等特点^[1]. 随着移动终端设备高速发展,利用人所携带的移动设备进行消息传输也越来越普遍^[2]. 由于人与人之间存在比较

稳定的社会关系,将社会关系引用到机会网络当中,充分考虑节点的社会属性和移动特性,能更好地优化机会网络的消息传输机制^[3]

如何对机会网络中的社区结构进行检测,是机会网络数据转发算法需要考虑的一个重要的方面. 社区检测算法的优劣,影响到社区路由的网络性能. Newman 等人^[4]分析,网络中的社区是一些在社会特性上彼此相似的节点集合,这些节点与网络中的

其他节点有较明显的社会特性区别. 社区检测算法是为了发现网络中社会性相似的节点, 并利用社会性特点提高网络相关性能. 当前的社区结构检测按照划分时采用的策略不同, 可以分为集中式的社区检测算法和分布式的社区检测算法. 集中式的社区检测算法具有代表性的有 WNA 算法、图分割算法和 GN 算法^[5]. 分布式社区检测算法主要有 SIMPLE 算法、k-CLIQUE 和 MODULARITY 算法^[6]. 由于机会网络节点的移动性和随机分布性, 网络中的任何节点几乎不可能收集完全网络中的节点的信息, 而且将社区检测结果公布给每个节点也很难做到, 因此, 分布式的社区检测算法更适用于机会网络场景. 本文针对分布式社区检测算法 CDTFS^[7] 社区检测算法中存在通信冗余开销的问题, 提出了一种低开销的机会网络社区检测算法—NSDMDT (Network Structure Detection Mechanism based on Dynamic Threshold), 目的是减少社区检测时网络中的通信开销, 进一步提高网络的性能.

2 问题描述

CDTFS 算法是在基于移动节点相遇规律的研究上, 提出的一种分布式社区检测算法. 该算法旨在解决一般分布式社区检测算法 (如 k-CLIQUE、SIMPLE 算法等) 需要预先设置某个阈值的问题, 根据网络状况的变化, 动态地设定适应的条件阈值. 本文通过对该算法的研究, 发现在社区检测阶段, 存在以下的问题.

(1) 当两节点 V_i 和 V_j 相遇时, 若 V_j 满足条件, 是新加入到 V_i 的朋友集合 F_i 中的节点, 则 V_j 将关系图发送给 V_i . 然后, V_i 将收到的子图合并到本地关系图中, 算法为了对称性, 若一个节点将对方加入自己的朋友集合中, 则另一个节点也将该节点加入自己朋友集合并交换信息. 如果 V_j 满足加入 V_i 朋友集合的条件, V_i 还需要发控制消息告知 V_j , 此过程增加了 V_i 通知 V_j 发送的通信开销.

(2) 在新的时间窗口 T 中, 由于 V_j 在当前时间窗口的朋友集合 F_i 发生变化, 因此 V_j 在与 V_i 相遇时将当前的本地关系图发送给 V_i . V_i 根据 V_j 传递给自己的子图信息, 对自己的社区关系子图进行更新. 但是, 如果 V_j 只是在前面基础上增加了朋友节点, 在原算法中, V_j 会将所有的朋友节点集合发给 V_i , 包括以前的朋友节点, 这样, 重复发送上一个时间周期 T 内的朋友节点会增加了一些不必要的通信冗余开销.

3 NSDMDT 算法

为解决上述的 2 个问题, 本文提出了 NSDMDT 算法, 优化了社区关系图的更新过程, 有效地减少原算法在构建社区本地关系图中的通信开销, 从而达到降低网络通信开销, 提高网络性能的效果.

3.1 NSDMDT 算法的新机制

(1) 优化新节点加入朋友集合时子图发送顺序

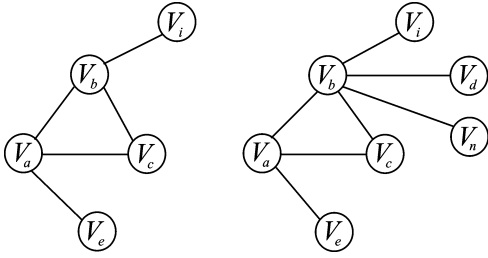
以节点 V_a 和 V_b 为例, 两个节点相遇时, 通过计算对方与自己的相遇持续时间和相遇次数, 判断是否将对方加入自己的朋友集合. 此时, 节点 V_a 若是根据条件, 要将 V_b 加入自己的朋友集合中, 则 V_a 主动将自己的本地关系图发送给 V_b , 而不是依据原算法需要 V_a 通知 V_b , 而根据算法提出的朋友关系对称性 V_b 收到 V_a 发送的本地关系图, 则知道 V_a 将自己加入到其本地关系子图中, V_b 将自己的关系图也发送给 V_a , 这样减少了通知对方发送子图的通信开销, 并减少通知对方并等待对方回应的通信时间.

(2) 优化朋友节点子图更新

以节点 V_a 和 V_b 为例, 如图 1 所示. 若 V_b 的朋友集合 F_b 发生了变化, 在新的时间窗口 T 中, 当节点 V_a 与 V_b 相遇时, V_b 要将当前的本地关系图发送给 V_a . 然后, V_a 根据由 V_b 引入的关系子图, 对自己的本地关系图进行更新. 此时有一种可能情况, V_b 只是在前一个时间周期 T 的基础上增加了朋友节点, 按照原算法, V_b 会将所有的朋友节点集合发给 V_a , 包括上一个周期 T 到当前 T 内没有变动的朋友节点, 重复发送朋友集合节点增加了一些通信冗余开销. 在 NSDMDT 算法中, 若出现这种情况, 上一个 T 内重复的邻居节点不再发送, 并用占 1 bit 的标志位来表示此种情况, 例如, 上一个周期 T 中 V_b 的朋友集合包括 V_a, V_c 和 V_i , 用 $F_b(t-1)$ 表示, 即 $F_b(t-1) = \{V_a, V_i, V_c\}$, 当前周期 V_b 朋友集合变动为 $F_b(t)$, V_b 的朋友集合包括 V_a, V_i, V_d, V_c 和 V_n , $F_b(t-1) = \{V_a, V_i, V_d, V_c, V_n\}$, 此时 V_b 向 V_a 发送的朋友关系子图是 $\{1, V_d, V_n\}$, 这样减少了冗余的通信开销, 同时一定程度上缩短了 V_a 的子图更新时间.

3.2 NSDMDT 算法操作

(1) 首先, 网络中各节点记录在时间周期 T 内与邻居节点的累积相遇持续时间和累计相遇次数, 以节点 i, j 为例, 分别用 $C_{ij}(T), d_{ij}(T)$ 表示两节点之间的累积相遇持续时间和累计相遇次数, 然后计算出 i 节点的邻居节点的相遇持续时间和次数的均值, 分别表示为



(a)更新前 (b)更新后

图1 优化朋友节点子图更新

$$\overline{C_i(T)} = \frac{1}{N} \sum_{k=1}^N C_{ik}(T) \quad (1)$$

$$\overline{d_i(T)} = \frac{1}{N} \sum_{k=1}^N d_{ik}(T) \quad (2)$$

(2) 将 i 节点的邻居节点中大于平均值的节点加入到相遇次数朋友集合 F_i^c 和相遇持续时间朋友集合 F_i^d 当中,相遇次数朋友集合的定义为

$$F_i^c = \{v_j \mid C_{ij} \geq \frac{1}{N} \sum_{k=1}^N C_{ik}(T) \mid \} \quad (3)$$

相遇持续时间的朋友集合定义为

$$F_i^d = \{v_j \mid d_{ij} \geq \frac{1}{N} \sum_{k=1}^N d_{ik}(T) \mid \} \quad (4)$$

(3) 最终的朋友集合 $F_i(T)$ 为相遇次数 F_i^c 和相遇持续时间 F_i^d 的交集,即节点选择相遇持续时间较长并且相遇次数较多的节点作为自己的朋友节点,朋友集合 $F_i(T)$ 为

$$F_i(T) = \{v_j \mid F_i^c \cap F_i^d \} \quad (5)$$

(4) 每个节点根据自己的朋友集合,会构造一个本地关系图,即哪些节点与自己是朋友关系.各节点在相遇时,会根据各节点的朋友集合来构造本地局部关系图.网络中各节点按照新机制将各自朋友集合中的节点加入到自己的本地关系图中,在图中,两个节点如果是朋友集合,则用一条边连接这两个节点来表示.节点在相遇时,互相交换彼此的朋友集合,并且根据节点间的朋友关系在互为朋友的节点之间建立一条边.

(5) 在社区节点之间建立的社区关系图的基础上,NSDMDT 算法利用已有研究提出的 FOCS 算法^[8],针对社区关系图中可能存在的一个或多个社区结构进行检测,进而完成对整个网络的社区检测.

3.3 NSDMDT 算法理论分析

为了确定 NSDMDT 算法能够减少原算法在社区检测阶段的通信开销,本文通过数学推导,对 NSDMDT 算法、CDTFS 算法的通信开销进行理论分析.

证明:假设上一个时间周期 T 中 V_b 的朋友集合包

括 V_a 和 V_c ,当前新的周期 T 内的朋友集合为 V_a, V_c, V_d, V_e, V_f ,则原算法需要传递的关系图为 $\{V_a, V_c, V_d, V_e, V_f\}$,假设控制包长度为 l .而新算法 NSDMDT 需传递的关系图为 $\{1, V_d, V_e\}$,假设控制包长度为 l_{new} ,则可得 $l_{\text{new}} < l$.假设整个社区检测过程中, V_a 和 V_b 发送矢量关系图为 N ,则由式(6)可知

$$\sum_{i=1}^N l_{\text{new}_i} < \sum_{i=1}^N l_i \quad (6)$$

针对网络中的其他节点,式(6)也同样满足,可见通过新机制 2,NSDMDT 算法能够减少原算法在社区检测阶段的通信开销.而由新机制 1,通过优化新节点加入朋友集合时的子图发送顺序,减少了节点之间的请求包的发送,进一步减少了原有的 CDTFS 算法的通信开销.

4 仿真分析

本文利用 ONE^[9] 仿真软件的开发环境进行算法对比仿真实验.在保证其他参量不变的情况下,首先改变网络场中节点的个数,从通信开销方面将所提 NSDMDT 算法及 CDTFS 算法进行对比.接下来,考虑节点通信范围大小,在节点通信半径不同的情况下,对比 NSDMDT 和 CDTFS 算法.然后考虑了节点移动速率不同时,两种算法的开销对比.本文也分析了两种算法中,各节点与朋友节点的累积相遇持续时间占与所有相遇节点的累积相遇持续时间之和的比例.仿真参数设置如表 1 所示.

表1 仿真参数设置

参数	数值
仿真场景/m ²	4 500×3 400
网络仿真时间/h	24
节点个数	40,60,80,100,120
节点通信半径/m	5,10,15,20,25
节点移动速度/(km/h)	2,4,6,8,10
节点缓存/Mbit	[5,50]
节点移动模型	Map Based Movement

(1) 节点个数不同

图 2 是不同节点个数对网络通信开销影响的仿真结果图.由图 2 可知,随着节点个数的增加,网络的通信开销会有所提高.这是因为网络中节点个数的增加,传递社会关系子图的数据包和包的大小会相应增加,使网络开销增加.而 NSDMDT 算法对 CDTFS 算法进行了改进,优化了朋友节点子图更新,同时优化新节点加入朋友集合时子图发送顺序,

避免了一些冗余的网络通信开销。

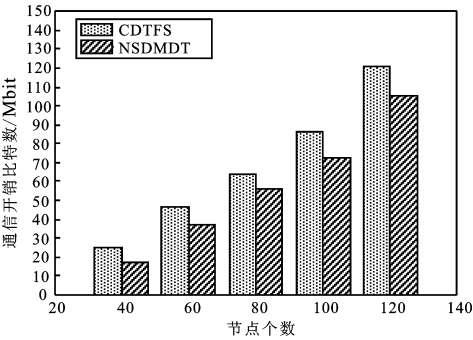


图 2 不同节点个数

(2)通信半径不同

图 3 是节点的不同通信半径对网络通信开销影响的仿真结果图. 由图 3 可知,随着节点通信半径的增加,网络的通信开销会有所提高. 这是因为网络中节点通信半径的增加,会使节点发现更多的邻居节点,从而其传递的社会关系子图的数据包和包的大小会相应增加,使网络开销增加. NSDMDT 算法增加了新机制,使得其在网络通信开销上低于 CDTFS 算法.

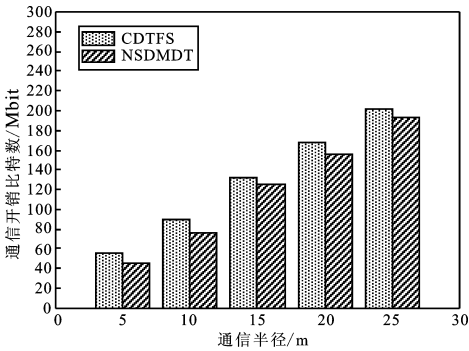


图 3 不同通信半径

(3)节点移动速率不同

图 4 是节点不同移动速率对网络通信开销影响的仿真结果图. 由图 4 可知,随着节点移动速率的增加,网络的通信开销会有所提高. 这是因为网络中节点移动速率的增加,会使节点能更快地遇见其他的邻居节点,并更快地在网络中运行,遇见更多的其他节点,从而其传递的社会关系子图的数据包和包的大小会相应增加,使网络开销增加. 而 NSDMDT 算法由于采用优化子图发送顺序和子图更新两种新机制,所以其在不同移动速度的情况下的网络通信开销也同样低于 CDTFS 算法.

(4)与朋友节点的累积相遇持续时间的比例

本文对比了各节点与朋友节点的累积相遇持续

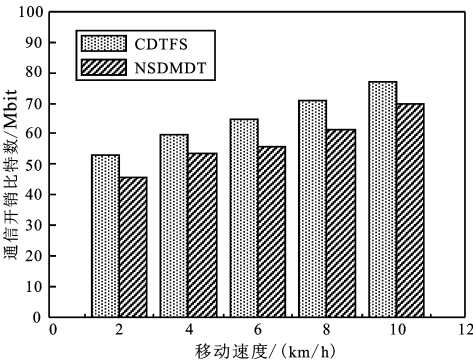


图 4 不同移动速率

时间与所有相遇节点的累积相遇持续时间之和的比例. 同样, NSDMDT 算法在此指标上也优于 CDTFS 算法. 如图 5 所示, NSDMDT 算法略高于 CDTFS 算法. 因为 NSDMDT 算法改进了新节点加入朋友集合时的子图发送顺序,减少了节点请求发送社会关系子图的过程,更及时地将节点加入到自己的朋友节点集合中,使节点与朋友节点的持续时间更准确.

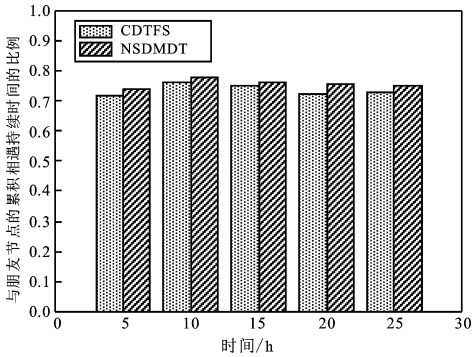


图 5 与朋友节点的累积相遇持续时间的比例

5 结束语

针对现有机会网络在社区检测阶段的网络通信开销存在冗余的问题,本文提出了 NSDMDT 算法. 该算法优化了朋友节点子图更新的过程,同时优化新节点加入朋友集合时子图发送顺序,避免了一些冗余的网络通信开销. 仿真结果验证了在网络中节点个数不同和节点通信半径不同、节点移动速率不同的这三种情况下, NSDMDT 算法在网络通信开销方面相对于原有的 CDTFS 算法得到降低,提高了网络性能. 同时,在各节点与朋友节点的累积相遇持续时间占与所有相遇节点的累积相遇持续时间之和的比例上, NSDMDT 算法优于 CDTFS 算法.

(下转第 69 页)

参考文献:

[1] J Kennedy, R C Eberhart. Particle swarm optimization [C]//Proc IEEE Int Conf Neural Networks. Perth, Australia;IEEE,1995: 1942-1948.

[2] XIA T, WANG X, LIANG S, et al. ANN trained by PSO with adaptive genetic operator and its application [J]. Journal of PLA University of Science and Technology: Natural Science Edition, 2011, 12 (1) : 70-74.

[3] 朱蓉,靳雁霞,范卫华. 融合优质粒子分布的粒子群优化算法[J]. 小型微型计算机系统, 2015, 35 (3): 576-580.

[4] Yannis Marinakis. An improved particle swarm optimization algorithm for the capacitated location routing problem and for the location routing problem with stochastic demands[J]. Applied Soft Computing, Volume 37, December 2015,37(12):680-701.

[5] 刘衍民,赵庆祯,牛奔. 基于自适应动态邻居和广义学习的改进粒子群算法[J]. 计算机应用,2010,30(10): 2578-2581.

[6] HUX, EBERHART R C. Multi-objective optimization using dynamic neighborhood particle swarm optimization [C]// Proceedings of IEEE Congress on Evolution Computation. Washington, DC: IEEE, 2002:

1677-1681.

[7] Wei Hong Lim, Nor Ashidi Mat Isa. Particle swarm optimization with increasing topology connectivity[J]. Engineering Applications of Artificial Intelligence, 2014,27(1):80-102.

[8] Mohammad Reza Bonyadi, Xiang Li, Zbigniew Michalewicz. A hybrid particle swarm with a time-adaptive topology for constrained optimization [J]. Swarm and Evolutionary Computation, 2014,18(10): 22-37.

[9] 鹿艳晶,马苗. 基于灰色粒子群优化的快速图像匹配算法[J]. 计算机工程与应用,2009,45(10):157-159.

作者简介:

靳雁霞 女,(1973-),博士,副教授. 研究方向为智能优化算法、虚拟现实技术.

张晓闻(通讯作者) 女,(1990-),硕士研究生. 研究方向为智能优化算法. E-mail :503212590@qq.com

银莉 女,(1991-),硕士研究生. 研究方向为智能优化算法.

张鑫 男,(1994-),硕士研究生. 研究方向为智能优化算法.

(上接第 63 页)

参考文献:

[1] 熊永平,孙利民,牛建伟,等. 机会网络[J]. 软件学报, 2009, 20(1): 124-137.

[2] Liu Li, Jing Yanfang. A survey on social-based routing and forwarding protocols in opportunistic networks[C]// Computer and Information Technology. Chengdu: IEEE Press, 2012: 635-639.

[3] WangYunsheng, Wu Jie. Social-tie-based information dissemination in mobile opportunistic social networks [C]//The Conference on International Symposium and Workshops. Madrid: IEEE Press, 2013: 1-6.

[4] Newman M E J. Modularity and community structure in networks[J]. Proceedings of the National Academy of Sciences, 2006, 103(23): 8577-8582.

[5] Fortunato S. Community detection in graphs[J]. Physics Reports, 2010, 486(3): 75-82.

[6] Hui P, Yoneki E, Chan S Y, et al. Distributed community detection in delay tolerant networks[C]// Proceedings of ACM/IEEE international workshop on

Mobility in the evolving internet architecture. New York: ACM Press, 2007: 716-722.

[7] 胡桐,郭忠文, Ludwig Wenning B, et al. 基于社团的移动容迟网络源路由算法[J]. 中国海洋大学学报(自然科学版), 2014, 44(3): 108-117.

[8] Clauset A, Shalizi C R, Newman M E J. Power-law distributions in empirical data [J]. SIAM review, 2009, 51(4): 661-703.

[9] Keranen A, Ott J, Karkkainen T. The ONE simulator for DTN protocol evaluation[C]// Proceedings of the International Conference on Simulation Tools and Techniques. Brussels: IEEE Press, 2009: 55-59.

作者简介:

任智 男,(1971-),博士后,教授. 研究方向为宽带自组网与无线通信.

黄希凯(通讯作者) 男,(1990-),硕士研究生. 研究方向为机会网络路由算法. E-mail:413575446@qq.com.

谭永银 男,(1989-),硕士研究生. 研究方向为基于博弈的机会网络路由算法.