

# 基于两阶段聚类的机会社会网络路由算法

张涓舒, 王慧强, 冯光升, 吕宏武, 温秀秀

(哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

**【摘要】**为提升机会社会网络路由过程中消息投递率、降低消息平均时延,对其消息转发过程进行了研究,提出一种基于两阶段聚类分析的机会社会网络路由算法。以分组路由策略为基础,通过两阶段聚类分析方法降低簇划分过程对节点资源的需求,并分别为簇内/间消息设计转发策略,优化了消息转发与中继节点选取的过程。此外,在聚类分析的过程中引入事件链分析的方法,深入挖掘节点间的内在社会关联,提高簇划分的准确性。仿真结果表明,在大规模复杂网络环境中该算法能够提高投递率5%~10%,降低投递时延10%以上,而在资源不足的情况下也能够获得接近80%的投递率。

**关键词** 事件链; 聚类; 机会社会网络; 路由; 社会网络分析

**中图分类号** TP393 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2017.04.021

## An Opportunistic Social Networks Routing Based on Two-Step Clustering

ZHANG Yu-shu, WANG Hui-qiang, FENG Guang-sheng, LÜ Hong-wu, and WEN Xiu-xiu

(College of Computer Science and Technology, Harbin Engineering University Harbin 150001)

**Abstract** In order to increase the delivery rate and reduce the transmission delay of messages in the process of routing, an opportunistic social networks routing based on two-step clustering is presented. The method of two-step clustering is used to reduce the resource requirements for nodes, and different forwarding strategies are applied to inter-message and intra-message respectively, aiming to optimize the process of message forwarding and relay node choosing. Besides, the chain of events is used in clustering to analyze internal social relationship between nodes, which can improve the accuracy of clustering. Our evaluations show that the protocol can lead to a 5~10% increase in delivery ratio and a 10% at least decrease in delivery delay in large complex networks, and get about 80% of messages delivered in networks with insufficient resources of computing and storage.

**Key words** chain of events; clustering; opportunistic social network; routing; social network analysis

机会网络(opportunity network)<sup>[1]</sup>是移动自组织网络的演化,是一种缺乏持续端到端连接的新型网络体系结构,利用节点移动带来的相遇机会实现信息传递。由于缺乏持续稳定的端到端路径,且延迟较高,故机会网络采用存储-携带-转发形式的路由协议<sup>[2]</sup>。近年来,随着大量低成本、具有短距离无线通信能力的移动设备的广泛应用,以人为载体的移动设备通过相遇机会获得数据交换的情况逐渐增多。这种趋势使得在不具备基础网络设施的情况下,移动设备间通过自组织网络形式实现大规模数据通信成为可能<sup>[3]</sup>。同时,由于人类活动具有社会性,在由人携带的移动设备组成的机会社会网络中,节点的移动符合人类的社会活动规律<sup>[4]</sup>,使得上述机会网络中节点的行为具有社会网络的特征,被称为

机会社会网络(opportunistic social networks, OSN)<sup>[5]</sup>。与机会网络的传统节点运动模型不同,人类的社会关系比较稳定而且存在一定的依赖性,表现为节点的“聚集”现象<sup>[6]</sup>。在社会网络中将这有聚集趋势的节点称为“簇”,属于同一簇的节点相遇概率高、持续时间长;而分属不同簇的节点相遇概率较低、持续时间较短。在机会网络环境中,节点数据的交换需要利用节点间的相遇机会完成,但真实的社会网络通信中经常体现“小世界”现象<sup>[7-8]</sup>。因此,探索人类活动的社会关系能够使机会社会网络路由协议具有更高实用性和有效性。

因此,研究者将研究的重点逐渐转移到网络节点关系对路由性能的影响上<sup>[9]</sup>,以解决传统机会路由算法不适用大规模、复杂网络的问题<sup>[10]</sup>。文献[11]

收稿日期: 2016-01-21; 修回日期: 2016-12-15

基金项目: 国家自然科学基金(61370212, 61402127, 61502118); 教育部高等学校博士点基金优先发展领域项目(20122304130002); 中央高校基本科研业务费专项资金(HEUCF100601)

作者简介: 张涓舒(1986-), 男, 博士生, 主要从事机会社会网络方面的研究。

将分层路由技术引入到机会网络中,提出一种基于分层的机​​会路由机制。文献[12]采用基于连接的最大直径方式,文献[13]通过在线统计节点相遇概率的方式,为节点的簇划分以及信息交换和路由提供基础。另外,文献[14]依据小世界理论将活跃的中间节点定义为桥节点,通过桥节点有效的降低消息投递过程的传递次数。文献[15]针对PSN,引入中心度和社区的概念,运用社会学的相关理论优化路由策略。文献[16]将社会距离作为选取消息传递目标的度量,同时为降低动态社区划分方法的计算压力,引入静态的社会关系图。

本文以聚类思想为基础,借鉴现有路由算法的优点,针对现有基于聚类的路由方法普遍存在的计算资源需求大、聚类结果不准确等问题,提出了一种基于两阶段聚类分析的机会社会网络路由算法。该方法依据节点的社会属性,对其进行聚类分析,将关系密切、接触频繁的节点划分到同一簇中,对簇内/簇间消息采用不同的转发策略,以达到提高路由效率和降低网络流量的效果。同时,为降低聚类分析对网络计算资源的占用,该过程采用基于相遇统计的简单聚类和基于数据交换的层次聚类两阶段完成。第一阶段聚类分析对计算能力要求较低,能够保障低计算资源网络环境下的基本路由需求;第二阶段聚类分析在计算资源充足的节点上进行,能够提升聚类结果的准确性,有效提高路由效率。

## 1 两阶段聚类算法

将物理或抽象对象的集合分成由类似对象组成的多个类的过程被称为聚类,由聚类所生成的簇是一组数据对象的集合。路由算法中,将移动节点作为聚类分析的对象,由聚类所生成的簇是一组节点的集合,这些节点与同一簇中的节点彼此关系紧密、相遇频繁,与其他簇中的节点联系较少。通常情况下,聚类算法需要依据一定规则反复迭代至算法收敛,这一过程需要消耗大量计算资源。机会社会网络中节点的计算能力与资源的分散分布的结构不能满足常规聚类算法的需求,因此在路由算法采用基于事件触发的聚类模式。第一阶段基于相遇统计的简单聚类设定节点相遇为聚类算法的触发事件,节点相遇时交换相遇列表并更新聚类信息;第二阶段基于数据交换的层次聚类设定节点交互为聚类算法的触发事件,仅在节点间有消息传递时才进行聚类信息的更新。

### 1.1 基于相遇概率统计的第一阶段聚类算法

机会社会网络中节点间不存在稳定的端到端连接,因此需要利用节点相遇的历史信息估算节点相遇概率,并以此作为网络拓扑的粗略描述。对于拥有 $n$ 个节点的机会社会网络 $N$ ,节点相遇矩阵 $M$ 是一个 $n \times n$ 概率矩阵,其中第 $i$ 行第 $j$ 列的元素是节点 $i$ 和节点 $j$ 的相遇概率 $m_{ij}$ ,代表对两节点在未来一段时间内相遇概率的预测。节点相遇概率按照规定的周期 $T$ 更新,更新规则应满足如下条件:如果在上一更新周期 $T$ 内节点 $i$ 与节点 $j$ 相遇,相遇概率 $m_{ij}$ 相应增大;反之, $m_{ij}$ 减小。据此提出节点间相遇概率公式:

$$m_{ij} = \begin{cases} (1-\alpha)m_{ij} + \alpha & \text{节点相遇} \\ (1-\alpha)m_{ij} & \text{没有相遇} \end{cases} \quad (1)$$

式中, $\alpha \in (0,1)$ ,当 $\alpha$ 取值较小时,相遇矩阵需经过较长时间达到稳定状态;反之,相遇矩阵变化迅速,容易出现反复震荡的情况。因此,需要根据不同应用场景进行选择。同时,由式(1)可知,如果节点频繁相遇,节点间相遇概率会随时间不断变大,反之会不断减少。相遇概率可以表示节点间的联系密切度,据此估计未来一段时间内节点的相遇情况,从而为机会社会网络中消息的转发提供依据。

基于节点间的相遇概率,定义节点 $i$ 与包含节点 $j$ 的簇 $C_j$ 的相遇概率 $m_{iC_j}$ ,簇 $C_i$ 与簇 $C_j$ 的相遇概率 $m_{C_iC_j}$ 如下所示:

$$m_{iC_j} = \frac{1}{|C_j|} \sum_{k=1}^{|C_j|} m_{ik} \quad (2)$$

$$m_{C_iC_j} = \frac{1}{|C_i| |C_j|} \sum_{k=1, k \neq l}^{|C_i|} \sum_{l=1}^{|C_j|} m_{kl} \quad (3)$$

式中, $|C_i|$ 、 $|C_j|$ 分别表示簇 $C_i$ 和簇 $C_j$ 所包含的节点数。

根据机会社会网络的特性,分簇算法基于节点的相遇事件触发。当节点 $i$ 与节点 $j$ 相遇时,通过计算相应的节点与簇间相遇概率以及簇与簇间的相遇概率,确定簇的划分与节点的加入/退出操作。

当节点 $i$ 与节点 $j$ 不属于同一簇时,如果节点 $i$ 与簇 $C_j$ 的相遇概率 $m_{iC_j}$ 大于等于分簇阈值 $\varepsilon_0$ ,且节点 $j$ 与簇 $C_i$ 的相遇概率 $m_{jC_i}$ 小于分簇阈值 $\varepsilon_0$ ,则将节点 $i$ 加入簇 $C_j$ 。

当节点 $i$ 与节点 $j$ 不属于同一簇,且 $m_{iC_j} \geq \varepsilon_0$ , $m_{jC_i} \geq \varepsilon_0$ 时,如果簇 $C_i$ 与簇 $C_j$ 的相遇概率 $m_{C_iC_j}$ 小于分簇阈值 $\varepsilon_0$ ,则比较节点 $i$ 与簇 $C_j$ 相遇概率 $m_{iC_j}$

与节点  $j$  与簇  $C_i$  相遇概率  $m_{jC_i}$  的大小; 若  $m_{iC_j} > m_{jC_i}$ , 则将节点  $i$  加入簇  $C_j$ , 否则将节点  $j$  加入簇  $C_i$ 。

当节点  $i$  与节点  $j$  不属于同一簇, 且  $m_{iC_j} \geq \varepsilon_0$ ,  $m_{jC_i} \geq \varepsilon_0$ ,  $m_{C_i C_j} \geq \varepsilon_0$  时, 将簇  $C_i$  与簇  $C_j$  合并。

当节点  $i$  与节点  $j$  属于同一簇时, 先同步节点  $i$  与节点  $j$  所属簇  $C$  的成员列表, 计算节点  $i$  与簇  $C$  除去节点  $i$  后形成的新簇  $C - \{i\}$  的相遇概率  $m_{i(C-\{i\})}$ 。若  $m_{i(C-\{i\})} < \varepsilon_0$ , 则将节点  $i$  从簇  $C$  中除去。对节点  $j$  的处理方法与节点  $i$  相同。

## 1.2 基于事件链分析的第二阶段聚类算法

在动态网络分析算法中, 将两个或两个以上的节点之间发生的连接称为事件, 机会网络中将节点间通信定义为事件。通常情况下, 机会社会网络中节点的活动都带有明显的社会性, 事件的发生大多具有明显的周期性规律, 因此能够形成具有紧密时间、空间关联事件集合。可以认为属于同一簇的节点间的链接往往具有明显、稳定的时空关联, 将这类事件集合称为事件链。事件与事件链是节点间内在关联和社会性的具体体现, 利用机会社会网络的这一特点, 对其进行基于事件链的深度聚类分析, 能够更精确地确定簇成员, 剔除无用节点, 进一步提高路由算法的效率。

机会社会网络中事件间的联系与事件发生的位置有关, 距离较近的事件间产生联系并组成事件序列的概率较大, 相反距离较远的事件间产生联系的概率较小。事件有多个节点之间的连接产生, 因此事件发生的位置为包含参与事件各个节点的区域, 一般用事件的中心点和半径表示。事件  $W$  的中心点坐标记为  $(x_W, y_W)$ , 根据式(4)进行计算。

$$\begin{cases} x_W = \frac{1}{n} \sum_k x_k \\ y_W = \frac{1}{n} \sum_k y_k \end{cases} \quad (4)$$

式中,  $n$  为参与事件的节点数;  $(x_k, y_k)$  为第  $k$  个节点参与事件的坐标。

令  $d_{WV} = \sqrt{(x_W - x_V)^2 + (y_W - y_V)^2}$  表示事件  $W$  和事件  $V$  之间的距离。若事件  $W$  和事件  $V$  之间存在联系表示为  $W \sim V$ ; 若不存在联系表示为  $W \not\sim V$ 。 $P(W \sim V)$  表示事件  $W$  和事件  $V$  之间存在联系的概率, 可记为  $p_{WV}$ 。文献[17]提出了一种根据距离估算事件间存在联系概率的方法, 用  $p_{WV}^L$  表示:

$$p_{WV}^L = \frac{1}{1 + e^{(d_{WV} - \alpha)}} \quad (5)$$

式中,  $\alpha$  是一个常量, 但文献[17]并未对其进行详细的说明。

为适应机会社会网络中事件联系的多样性, 对式(5)做出两点改动。首先假设事件与其他事件产生联系的能力各不相同。用半径  $r_W$  表示事件  $W$  的这种能力。 $r_W$  越大, 事件  $W$  能够与越多的事件接触并取得联系。一般认为事件包含的节点越多越容易与其他事件产生联系, 定义包含节点数为  $\delta_W$  的事件  $W$  的半径为  $c(\delta_W + 1)$ 。其中  $c$  为参数, 与公式应用的环境有关。常数 1 为确保半径非零。

引入  $r_{WV}$  代替式(5)中的常量  $\alpha$ ,  $r_{WV}$  表示事件  $W$  和事件  $V$  半径中较大的一个。因此可得:

$$r_{WV} = c(\max(\delta_W, \delta_V) + 1) \quad (6)$$

式中,  $\delta_V$  表示事件  $V$  包含的节点数。

第二个改动是添加核函数  $K(d_{WV})$  作为  $p_{WV}$  的权重参数。改变后的计算公式使得产生联系时距离小于  $r_{WV}$  的事件  $W$  和事件  $V$  有较高概率  $p_{WV}$ 。考虑核函数在  $d_{WV} = r_{WV}$  时的连续性, 定义  $K(d_{WV})$  如下:

$$K(d_{WV}) = \begin{cases} (1 - (d_{WV}/r_{WV})^2)^2 & d_{WV} \leq r_{WV} \\ 0 & d_{WV} > r_{WV} \end{cases} \quad (7)$$

同时增加常数  $\rho$ , 表示距离较远的事件之间也有小概率存在联系。因此, 重新定义事件  $W$  和事件  $V$  之间存在联系的概率  $p_{WV}$  如下:

$$p_{WV} = \frac{1}{1 + e^{(d_{WV} - r_{WV})}} K(d_{WV}) + \rho(1 - K(d_{WV})) \quad (8)$$

等价于:

$$p_{WV} = \begin{cases} p_{WV}^L K(d_{WV}) + \rho(1 - K(d_{WV})) & d_{WV} \leq r_{WV} \\ \rho & d_{WV} > r_{WV} \end{cases} \quad (9)$$

$p_{WV}$  仅能表示事件  $W$  和事件  $V$  形成事件链的可能性大小, 不能单独用于确定事件链。为此还需引入事件相似度的概念。对于给定的两个事件  $W$  和  $V$ , 定义事件相似度函数  $\text{sim}(W, V)$ 。 $\text{sim}(W, V)$  只表示事件  $W$  和事件  $V$  所包含的节点集合的相似程度, 与事件发生的时间与位置无关。若事件  $W$  包含的节点数为  $\delta_W$ , 事件  $V$  包含的节点数为  $\delta_V$ , 其中两事件都包含的节点数为  $\delta_{W \cap V}$ , 可将  $\text{sim}(W, V)$  简单定义为:

$$\text{sim}(W, V) = \frac{2\delta_{W \cap V}}{\delta_W + \delta_V} \quad (10)$$

选定阈值  $\beta$  以及参数  $\sigma$ , 定义事件链判别式  $\Delta$ :

$$\Delta = \text{sim}(W, V) + \sigma p_{WV} \quad (11)$$

若  $\text{sim}(W, V) + \sigma p_{WV} \geq \beta$ , 则  $W \sim V$ ; 若  $\text{sim}(W, V) + \sigma p_{WV} < \beta$ , 则  $W \not\sim V$ 。其中阈值  $\beta$  在区间  $(0, 1)$  上取

值, 取值越大聚类算法获得的簇越紧凑, 每个簇包含的节点越少; 反之, 阈值  $\beta$  取值越小, 聚类算法获得的簇越松散, 每个簇包含的节点越多。根据判别式可知当  $\text{sim}(W, V) = 1$ , 即事件  $W$  与  $V$  所包含节点相同时有  $\text{sim}(W, V) + \sigma p_{wv} < \beta$ , 使得  $W \sim V$  成立。 $p_{wv}$  通过事件发生位置间的距离, 间接地预测事件所包含的节点间的内在联系, 因此在判别式中权重小于  $\text{sim}(W, V)$ , 即参数  $\sigma$  的取值区间为  $(0, 1)$ 。

对于任意事件  $W$  和事件  $V$ , 通过计算能够确定两者是否属于同一事件链。按照事件链的定义, 位于同一事件链上的事件所包含的节点有较高的相遇的概率, 因此应属于同一簇; 相反, 不在同一事件链上的事件所包含的节点, 即缺少相互连接的公共节点, 又没有时间、空间上的内在联系, 彼此相遇的概率较低, 因此应分属不同的簇。第二阶段的聚类分析是基于第一阶段聚类分析的结果, 同时充分利用了事件链的相关性质。当事件  $W$  和事件  $V$  位于同一事件链, 即  $W \sim V$  成立时, 若事件  $W$  包含节点所属的簇  $C_w$  与事件  $V$  包含节点所属的簇  $C_v$  相同, 不对簇  $C_w / C_v$  进行操作; 若事件  $W$  包含节点所属的簇  $C_w$  与事件  $V$  包含节点所属的簇  $C_v$  不同时, 需合并簇  $C_w$  与簇  $C_v$ 。当事件  $W$  和事件  $V$  不位于同一事件链, 即  $W \not\sim V$  成立时, 若事件  $W$  包含节点所属的簇  $C_w$  与事件  $V$  包含节点所属的簇  $C_v$  不同, 不对簇  $C_w$  与簇  $C_v$  进行操作; 若事件  $W$  包含节点所属的簇  $C_w$  与事件  $V$  包含节点所属的簇  $C_v$  相同, 需对簇  $C_w / C_v$  进行分割。

## 2 基于中心度的分簇消息转发机制

经过两阶段聚类分析, 整个网络被分割成为若干簇, 每个簇可以看作一个规模较小的机会社会网络。聚类分析中分簇决策的依据是节点间的内在联系, 所以同簇节点的关系紧密、相遇概率较高, 而分属不同簇节点的关系松散、相遇概率较低。因此按照消息携带节点和目标节点是否属于同一个簇, 可以将机会社会网络中消息转发分成两种情况, 消息携带节点与目标节点属于同一个簇的称为簇内消息转发, 属于不同簇的称为簇间消息转发。为提高路由协议的效率, 需要针对这两种情况中节点间关系的特点, 分别提出相适应的策略。

### 2.1 簇内消息转发策略

簇内的消息转发策略采用改进的Spray and Wait路由算法, 消息在中间节点传递的过程中副本配额的分配比例依据各节点与目的节点的相遇概率决

定。当源节点  $s$  有消息  $M$  需要发送到目的节点  $d$  时, 如果节点  $s$  与节点  $d$  属于同一个簇, 则节点  $s$  首先确定消息  $M$  的副本配额为  $L_0$ , 即在簇  $C_s$  中最多存在  $L$  个消息  $M$  的副本。在消息转发过程中, 当携带消息  $M$  的中间节点  $i$  (副本配额为  $L$ ) 与节点  $j$  相遇时, 节点  $i$  将  $\varepsilon L$  份副本转发给节点  $j$ 。其中:

$$\varepsilon = \frac{m_{jd}^2}{m_{id}^2 + m_{jd}^2} \quad (12)$$

每次转发节点  $i$  最多转移  $L-1$  份副本, 当其只携带一份副本时, 节点将一直保存该副本直到与目的节点相遇或消息失效。

### 2.2 簇间消息转发策略

在社会网络中, 中心度高的节点不仅与同簇节点关系紧密, 与部分异簇节点间也有较高的相遇概率。这类节点能够成为不同簇节点之间联系的桥梁, 充分利用可以有效提高簇间消息转发效率。节点  $i$  中心度  $\text{cen}_i$  的一般测量式为:

$$\text{cen}_i = \sum_j r_{ij} (\alpha + \beta \text{cen}_j) \quad (13)$$

由于机会社会网络中, 节点中心度测量只能涉及与其相连(有相遇历史)的节点, 具有局域性, 因此可将式(13)简化为:

$$\text{cen}_i = \sum_{j \in C_i} m_{ij} \text{cen}_j \quad (14)$$

当携带簇间消息  $M$  的节点  $i$  与节点  $j$  相遇时, 如果节点  $i$  与节点  $j$  属于同一簇, 仅当  $\text{cen}_i < \text{cen}_j$  时, 节点  $i$  将消息转发给节点  $j$  并删除消息  $M$ ; 如果节点  $i$  与节点  $j$  属于不同簇, 节点  $i$  将消息转发给节点  $j$ , 并保留消息  $M$ 。在传递过程中, 需要将消息传递到达的簇记录在消息中, 以免消息在簇间的重复传递。

## 3 性能分析

### 3.1 仿真环境

本文利用MATLAB平台对MIT Reality Trace数据集<sup>[18]</sup>进行基于事件驱动的仿真。MIT Reality Trace共包括94个节点, 近110 K条接触记录。节点间的接触间隔基本符合指数分布, 同时存在明显社会特征。

在仿真过程中, 节点按照一定概率产生目的地地址随机的数据包。数据包大小在 $[1, 100]$  kB区间均匀分布。数据集的前1/4数据作为仿真的预热, 以使聚类算法的分簇结果达到稳定状态, 仿真结果从数据集的后3/4部分仿真中获得。

选取PRoPHET、Clustering和BUBBLE作为对比算法。在参数选择上, 除特别列出的专有参数采用

算法的惯用取值外, 其余参数与本文算法相同。

### 3.2 仿真结果及分析

#### 1) 投递率与投递时延

投递率与投递时延是机会社会网络较为重要的评价标准, 其中投递率是指成功送达目的节点的消息数在所有产生的消息中所占的比例, 投递时延是指消息从源节点产生到送达目的节点所需时间的平均值。结合应用环境特点和算法特性, 本文针对网络运行时间的变化对协议性能的影响进行仿真。

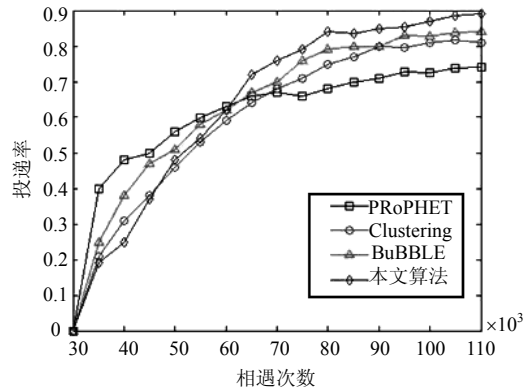
图1a表示仿真过程中投递率随接触次数增加的变化趋势。实验中将每个消息的生存时间(TTL)设定为9个月, 并且不限制节点的缓存大小, 即在仿真过程中不会出现报文被丢弃的情况。在仿真开始阶段PRoPHET的投递率迅速上升, 较早地达到了稳定状态。其他算法的投递率逐渐上升, 在中期达到稳定状态。最终本文算法投递率最高, 而PRoPHET投递率最低, Clustering和BUBBLE投递率基本相当。

PRoPHET基于历史相遇记录预测未来一段时间的节点相遇概率, 所需要的信息少, 使用的分析方法简单, 因此能在仿真过程中较早地达到稳定阶段, 但最终投递率不如其他算法。基于聚类分析的Clustering和基于社区的BUBBLE的核心都是将节点划分成为较小的集合, 降低网络节点间关系的复杂性。仿真过程中此类算法需要较长的预热过程, 在收集足够的信息使得聚类分析或社区划分稳定后, 算法的投递率会达到较高的水平。本文算法在聚类分析的基础上, 利用社会分析学中事件链理论深入探索节点间的内在联系, 因此相对于其他算法获得了更高的投递率。

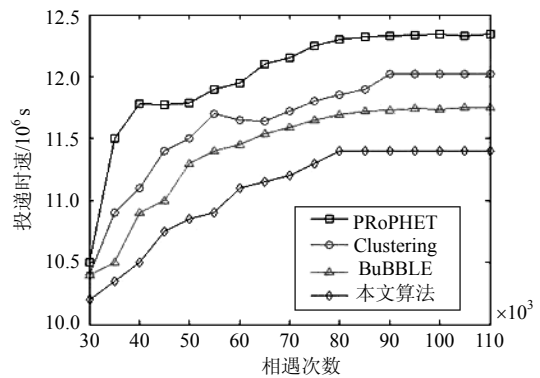
图1b表示仿真过程中投递时延随接触次数增加的变化趋势。仿真开始阶段, 各个算法投递时延较小。随仿真进程推进, 投递时延逐渐增加, 直到仿真进行到50 K条记录后趋于平稳。由图中可以看出, 在各算法的投递时延达到稳定后, PRoPHET的投递时延最大, Clustering其次, BUBBLE稍好于Clustering, 而本文算法获得了最低的投递时延。

PRoPHET只考虑节点间的相遇概率, 不具备分析节点间相互关系的能力, 不能给出优化的消息传递路径, 因此具有最高的投递时延。Clustering与BUBBLE通过聚类分析与社区划分能够区分节点间关系的疏密, 减少传递过程中的转发次数, 因此获得相对较低的传递时延。BUBBLE因为引入了中心度的概念, 有效利用中心度高的活跃节点传递消息, 所以投递时延优于Clustering。本文算法采用事件链

分析方法, 充分利用节点间的内在联系, 降低了消息无效转发的几率, 因此获得了最低的投递时延。



a. 相遇次数对投递率的影响



b. 相遇次数对投递时延的影响

图1 算法投递率与投递时延变化趋势

#### 2) 转发效率

转发效率是成功递交的消息数与网络中消息转发总次数的比值。转发效率能够体现路由算法的优化程度, 具有较高转发效率的算法能够在网络状态相同的情况下成功送达更多的消息, 获得较高投递率。图2表示仿真过程中算法转发效率的变化趋势。

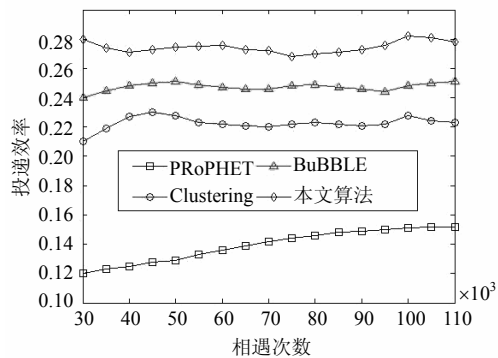


图2 算法转发效率对比

从图中可以看出, 本文算法转发效率最高, 接近0.3, 即消息从产生到送达目的节点平均需要3~4次传递。而Clustering与BUBBLE转发效率均低于

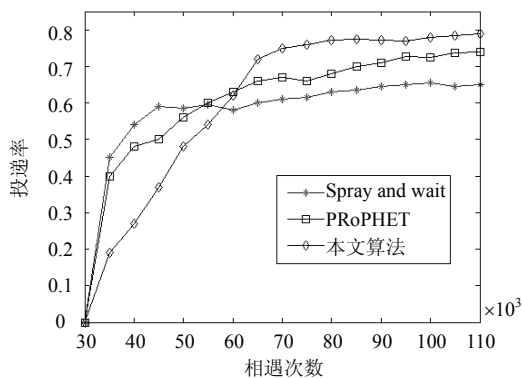
0.25, 消息成功送达需4次以上的传递。表现最差的PRoPHET的转发效率在0.15以下, 平均消耗7次以上转发才能将消息送达。通过分析可知, 本文提出的算法在投递率相同的情况下转发次数少于其他算法, 因而产生的副本数量、占用的节点缓存较少, 能够有效节省网络资源, 提高网络的利用率。

### 3) 算法效率

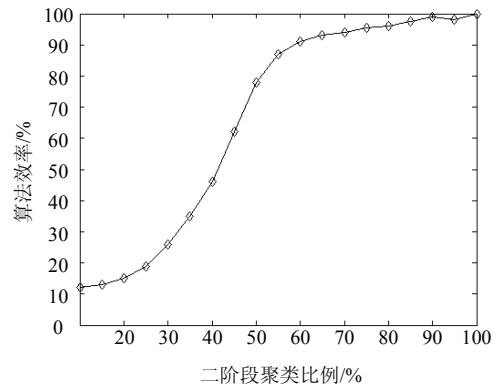
受到网络移动性的影响, 机会社会网络的节点通常计算能力较弱, 进而制约了路由算法在实际使用中的效果。本文算法在节点分析阶段采用两阶段聚类设计, 第一阶段聚类算法设计简洁, 占用计算资源少, 能确保算法的基本效率要求, 而第二阶段聚类分析, 能够有效提高路由效率。在节点计算能力较弱的情况下, 本文算法可放弃第二阶段聚类过程。本组实验中本文算法只执行第一阶段的聚类分析, 图3a表示实验中本文算法与PRoPHET、SW的投递率变化趋势。从图中可以看出, 在计算资源不足的情况下本文算法依然获得了最高的投递率。

图3b表示网络中节点计算能力对本文路由消费效率的影响。为模仿计算资源受限的情况, 图中横坐标表示仿真中各节点进行第二阶段聚类分析的比例, 纵坐标表示算法的投递率在计算资源受限情况下与正常情况下的比值。从图中可以看出, 当网络中节点的计算资源能够满足略高于40%的节点具有第二阶段聚类分析能力时, 本文算法的效率就能够基本达到计算资源充足情况下的80%左右。

综合上述实验可知, 在采用真实世界中人类移动轨迹的移动模型的仿真场景中, 本文算法能够提高机会社会网络的路由能力, 相比其他算法取得了较高的投递率与较低的投递时延, 并且具有较高的转发效率, 节省了节点的缓存空间。



a. 计算能力不足时相遇次数对投递率的影响



b. 二阶段聚类比例对算法效率的影响

图3 计算能力对算法的影响

## 4 结束语

本文通过对机会社会网络特点的分析, 提出了一种基于两阶段聚类分析的机会社会网络路由算法, 以解决机会社会网络中网络规模增大、节点计算资源匮乏引起的路由算法效率低下的问题。

根据机会网络中节点在运动与交互过程中体现出的社会特性, 采用聚类分析方法将网络划分成为若干规模较小的网络, 降低了网络规模对路由算法效率的影响; 聚类分析过程分为两阶段实现, 减少了路由算法对节点计算资源的需求。实验结果表明, 在大规模复杂网络环境中, 本文算法能够有效提供消息投递率、降低投递时延, 同时能够降低路由算法对网络计算、存储资源的消耗。

## 参考文献

- [1] PELUSI L, PASSARELLA A, CONTI M. Opportunistic networking: Data forwarding in disconnected mobile ad hoc networks[J]. IEEE Communications Magazine, 2006, 44(11): 134-141.
- [2] 熊永平, 孙利民, 牛建伟, 等. 机会网络[J]. 软件学报, 2009, 20(1): 124-137.  
XIONG Yong-ping, SUN Li-min, NIU Jian-wei, et al. Opportunistic networks[J]. Journal of Software, 2009, 20(1): 124-137.
- [3] 徐佳, 李千目, 张宏, 等. 机会网络中的自适应喷雾路由及其性能评估[J]. 计算机研究与发展, 2015, 47(9): 1622-1632.  
XU Jia, LI Qian-mu, ZHANG Hong, et al. Performance evaluation of adaptive spray routing for opportunistic networks[J]. Journal of Computer Research and Development, 2015, 47(9): 1622-1632.
- [4] CONTI M, MORDACCHINI M, PASSARELLA A, et al. A semantic-based algorithm for data dissemination in opportunistic networks[C]//Self-Organizing Systems. Berlin Heidelberg: Springer, 2014: 14-26.
- [5] PIETILÄNEN A K, DIOT C. Dissemination in opportunistic

- social networks: the role of temporal communities[C]// Proceedings of the Thirteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing. [S.l.]: ACM, 2012: 165-174.
- [6] 牛建伟, 周兴, 刘燕, 等. 一种基于社区机会网络的消息传输算法[J]. 计算机研究与发展, 2015, 46(12): 2068-2075.  
NIU Jian-wei, ZHOU Xing, LIU Yan, et al. A message transmission scheme for community-based opportunistic network[J]. Journal of Computer Research and Development, 2015, 46(12): 2068-2075.
- [7] 王慧强, 胡海婧, 朱金美, 等. 面向DTN感染路由协议的缓存管理算法[J]. 电子科技大学学报, 2015, 44(3): 403-409.  
WANG Hui-qiang, HU Hai-jing, ZHU Jin-mei, et al. Message-period-based buffer management algorithm for Epidemic routing protocol of DTN[J]. Journal of University of Electronic Science and Technology of China, 2015, 44(3): 403-409.
- [8] 聂旭云, 杨炎, 刘梦娟, 等. 基于节点能力模型的容迟网络路由算法[J]. 电子科技大学学报, 2013, 42(6): 905-910.  
NIE Xu-yun, YANG Yan, LIU Meng-juan, et al. Capability model based routing strategy in DTN[J]. Journal of University of Electronic Science and Technology of China, 2013, 42(6): 905-910.
- [9] 张振京, 金志刚, 舒炎泰. 基于节点运动预测的社会性DTN高效路由[J]. 计算机学报, 2013, 36(3): 626-635.  
ZHANG Zhen-jing, JIN Zhi-gang, SHU Yan-tai. Efficient routing in social DTN based on nodes' movement prediction[J]. Chinese Journal of Computers, 2013, 36(3): 626-635.
- [10] 郑啸, 罗军舟, 曹玖新, 等. 面向机会社会网络的服务广告分发机制[J]. 计算机学报, 2012, 35(6): 1235-1248.  
ZHENG Xiao, LUO Jun-zhou, CAO Jiu-xin, et al. Service advertisement dissemination in opportunistic social networks[J]. Chinese Journal of Computers, 2012, 35(6): 1235-1248.
- [11] AHMED S, KANHERE S S. Cluster-based forwarding in delay tolerant public transport networks[C]//IEEE Conference on Local Computer Networks. [S.l.]: IEEE, 2007: 625-634.
- [12] WHITBECK J, CONAN V. HYMAD: Hybrid DTN-MANET routing for dense and highly dynamic wireless networks[J]. Computer Communications, 2010, 33(13): 1483-1492.
- [13] DANG H, WU H. Clustering and cluster-based routing protocol for delay-tolerant mobile networks[J]. IEEE Transactions on Wireless Communications, 2010, 9(6): 1874-1881.
- [14] DALY E M, HAAHR M. Social network analysis for routing in disconnected delay-tolerant manets[C]// Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing. [S.l.]: ACM, 2007: 32-40.
- [15] HUI P, CROWCROFT J, YONEKI E. Bubble rap: Social-based forwarding in delay-tolerant networks[J]. IEEE Transactions on Mobile Computing, 2011, 10(11): 1576-1589.
- [16] JAHANBAKHS K, SHOJA G C, KING V. Social-greedy: a socially-based greedy routing algorithm for delay tolerant networks[C]//Proceedings of the Second International Workshop on Mobile Opportunistic Networking. [S.l.]: ACM, 2010: 159-162.
- [17] HOFF P D, RAFTERY A E, HANDCOCK M S. Latent space approaches to social network analysis[J]. Journal of the American Statistical Association, 2002, 97(460): 1090-1098.
- [18] EAGLE N, PENTLAND A. Reality mining: Sensing complex social systems[J]. Personal and Ubiquitous Computing, 2006, 10(4): 255-268.

编辑 蒋晓