

# 基于链接率的 Ad Hoc 自适应按需加权分簇算法

杜国勇, 束永安

(安徽大学 计算机科学与技术学院, 安徽 合肥 230601)

**摘要:** Ad Hoc 网络通过分簇算法来实现网络分层, 以支持高效的资源管理和路由策略。稳定的分簇算法可以减少网络的计算和通信开销。为了提高分簇算法的稳定性, 文中提出了一种基于链接率的自适应按需加权分簇算法 (AOWLR)。该算法将节点的运动轨迹抽象为直线形和圆弧形两种, 并引入邻居节点的平均链接率预测值作为衡量簇稳定性的一个重要标准。NS2 仿真结果证明, 同 AOW 相比, AOWLR 算法的负载均衡因子较高, 连通统治集 (CDS) 更新频率和节点充当簇头的公平性指数 (HFI) 较低。

**关键词:** Ad Hoc; 分簇; 链接率

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2014)01-0093-05

doi: 10.3969/j.issn.1673-629X.2014.01.024

## An Adaptive On-demand Weighting Clustering Algorithm Based on Linking Rate in Ad Hoc Networks

DU Guo-yong, SHU Yong-an

(Institute of Computer Science and Technology, Anhui University, Hefei 230601, China)

**Abstract:** Ad Hoc networks use clustering algorithm to achieve hierarchical network to support effective resource management and routing strategies. A stable clustering algorithm can reduce the network overhead in calculation and communication. In order to improve the stability of clustering algorithm, a novel adaptive on-demand weighting clustering algorithm is proposed which is based on the linking rate. The node trajectories are abstracted to arcs and lines. The forecast of the average linking rate of neighboring nodes is used as the standard to measure the cluster's stability. Simulation results show that the AOWLR has higher loading balance factor, lower updating frequency of cluster domination set (CDS) and lower head fairness index (HFI) compared with the AOW.

**Key words:** Ad Hoc; clustering; linking rate

### 0 引言

Ad Hoc 网络是一组带有无线收发装置的移动终端组成的一个多跳的临时性自治系统, 它采用无中心结构, 各移动终端兼有路由和转发功能, 可以通过无线连接构成任意的网络拓扑并独立工作<sup>[1]</sup>。一般应用中, Ad Hoc 网络采用平面结构或层次结构, 但在大规模网络中特别是要求满足一定的 QoS 需求时, 通常采用层次结构。

Ad Hoc 网络层次结构通过分簇算法来构造, 即网络节点被划分为若干个簇, 簇内节点通过簇头交换信息, 簇间节点通过簇头或网关交换信息。分簇算法的好坏直接影响着路由协议的性能, 一个好的分簇算法,

可以减少计算和通信开销, 能够较好地支持资源管理和路由策略。

文中在现有的自适应按需加权分簇算法的基础上加以改进, 通过预测节点间的链接率来确定簇结构的稳定性, 可以有效减少统治集更新次数, 从而减少算法的执行频率。

### 1 相关概念

文中把 Ad Hoc 网络抽象为无向图  $G = (V, E)$ , 其中  $V$  表示边集合,  $E$  表示节点集合。某一时刻节点  $i$  的坐标记为  $(x_i, y_i)$ , 速度记为  $v_i$ , 运动方向记为  $\theta_i$ 。用  $C$  表示网络中的簇头数, 用网络的统治集 (Connected

收稿日期: 2013-03-22

修回日期: 2013-06-25

网络出版时间: 2013-11-12

基金项目: 中国博士后科学基金资助项目 (20080440706)

作者简介: 杜国勇 (1985-), 男, 硕士研究生, 研究方向为无线 Ad Hoc 网络路由协议; 束永安, 副教授, 研究方向为无线网络及下一代网络体系结构。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20131112.1636.021.html>

Domination Set, CDS) 表示网络中簇头集合, 用  $U$  表示单位时间内统治集更新次数。定义负载平衡因子<sup>[1]</sup>(Load Balance Factor, LBF) 为:  $LBF =$

$$\frac{n_c}{\sum_{i=1}^n (x_i - u)^2}$$

其中,  $n_c$  表示网络中簇的数目;  $x_i$  表示簇  $i$  的成员节点数;  $u$  表示簇头的平均度数,  $u = \frac{N - n_c}{n_c}$ ,

$N$  为网络中节点个数。定义簇头公平性指数<sup>[1]</sup>(Head Fairness Index, HFI) 为:  $HFI = E\{|[t_i - E(t_i)]|\}$ ,  $i \in V$ 。其中,  $t_i$  表示节点  $i$  充当簇头的总时间;  $E(t_i) =$

$$\frac{\sum_{i \in V} t_i}{N}$$

表示节点  $i$  充当簇头的平均时间。

## 2 分簇算法的相关研究

研究人员已提出多种节点分簇算法。最小 ID 分簇算法 (LOWID)<sup>[2]</sup> 通过广播 Hello 消息使各节点获得其邻居节点 ID, 该算法选取相邻节点中具有最小 ID 号的节点作为簇头, 其一跳邻居节点成为其簇成员节点, 不再参与接下来的簇头选举, 重复上述操作直到所有节点选完为止。这种算法的优点是算法简单、收敛快, 缺点是选择最小 ID 号的节点作为簇头会加快该节点能量的消耗, 没有考虑负载平衡等因素。

最高节点度分簇算法 (HIGHD)<sup>[3]</sup> 的目标在于减少簇的数目。通过广播 Hello 消息使各节点获得其邻居节点的度数, 选取该节点和相邻节点中度数最高的节点作为簇头, 若有多个节点的度数处于相同的最高值, 则选取 ID 最小的节点作为簇头, 其一跳邻居节点成为其簇成员节点, 不再参加接下来的簇头选举, 重复上述操作直到所有节点选完为止。这种算法的优点在于网络中簇的数目较少, 减少了分组投递时延, 缺点是簇的规模往往很大, 簇成员数不受限制, 簇头负担过重, 没有考虑负载平衡。

上述算法中, 簇头的选举只考虑了某方面的因素。自适应按需加权分簇算法<sup>[4-5]</sup>(Adaptive On-demand Weighing clustering algorithm, AOW) 通过权重因子来选取簇头, 计算公式为  $W_n = c_1 D_n + c_2 P_n + c_3 M_n + c_4 T_n$ , 其中  $c_1 + c_2 + c_3 + c_4 = 1$  且  $0 < w_1, w_2, w_3, w_4 < 1$ ,  $D_n, P_n, M_n, T_n$  分别表示节点度数与理想度数之差、节点与其邻居节点的距离和、节点的移动性和节点的能量损耗。簇结构比较灵活, 并可以根据具体的环境对各因素设置不同的权值, 提高了节点的公平性。

文中在 AOW 的基础上提出了一种基于预测链接率的自适应按需加权分簇算法 (Adaptive On-demand Weighing clustering algorithm based on Linking Rate,

AOWLR)。该算法考虑了最佳链接度和能量状态及簇的稳定性, 使簇头的选择更加合理, 并且具有较小的通信开销。

## 3 基于预测链接率的自适应分簇算法 (AOWLR)

AOWLR 与 AOW 相比, 改进了权值计算公式, 给出了具体的度数和能量的计算, 并且预测将来一段时间的链接情况代替计算移动性作为簇稳定性的重要标准, 使得出的簇结构更加稳定, 减少了运行分簇算法的频率。AOWLR 的权重计算公式定义为:  $W = w_1 D + w_2 S + w_3 E + w_4 \frac{1}{P}$ , 其中  $w_1 + w_2 + w_3 + w_4 = 1$  且  $0 < w_1, w_2, w_3, w_4 < 1$ ;  $D$  表示节点的度数与理想度数之差;  $S$  表示节点与其邻居节点距离的平均值;  $E$  表示节点的能量损耗;  $P$  表示节点与其邻居节点的平均链接率。

### 3.1 AOWLR 中簇头度数、距离、剩余能量计算

$D_i$  表示节点  $i$  的度数与理想度数之差。簇头的邻居节点太多则会拥塞, 太少则浪费带宽, 因此簇头的邻居节点数应该适中。根据文献[6] 可得, 簇头的最佳节点度数应该为  $d_{ideal} = \frac{w_1}{w_2} \sqrt{N}$ 。其中,  $w_1, w_2$  分别表示簇内和簇间的通信带宽;  $N$  表示整个网络的节点个数。可得  $D_i = |d_i - d_{ideal}|$ 。其中,  $d_i$  为节点  $i$  的度数。AOWLR 选取与其理想度数的距离值最小的节点为簇头。

$S_i$  表示节点  $i$  与其邻居节点距离的平均值, 即  $S_i = \frac{\sum_{j=1}^k S_{ij}}{k}$ 。其中,  $k$  为节点  $i$  的邻居节点个数,  $S_{ij}$  表示节点  $i$  与其第  $j$  个邻居节点的距离。由于有效范围内节点间距离越近, 信号衰减越少, 通常节点间距离的四次方和节点发射功率成正比<sup>[7]</sup>, 因此选择与邻居节点平均距离值小的节点作为簇头可以保证发射成功并节约能量。

$E_i$  表示节点  $i$  的能量损耗, 节点的能量损耗分为两个部分, 作为簇头的能量损耗和作为簇成员的能量损耗。根据文献[8] 可得:

$$(1) \text{ 若状态转换前节点 } i \text{ 处于簇头状态, 则 } E_i = E_{iold} + t_{iold} \times d_{iold} \times e_m$$

其中,  $t_{iold}$  为节点  $i$  转换状态 (状态被定义为节点作为簇头或簇成员的性质不变且度数不变的情况) 时前一状态的持续时间;  $d_{iold}$  为节点  $i$  在前一状态时的度数;  $e_m$  表示节点作为簇成员时单位时间消耗的能量,  $e_m$  可以通过统计获得该值;  $E_{iold}$  为节点  $i$  在前一状态初始

时的能量损耗。

(2) 若状态转换前节点  $i$  处于簇成员状态,则  $E_i = E_{iold} + t_{iold} \times e_m$ 。

### 3.2 AOWLR 中预测节点的链接率模型

$P_i$  表示节点  $i$  与其邻居节点的平均链接率。为了减少分簇开销和簇头的频繁改变,通过 GPS 定位来获得 Ad Hoc 网络各节点的坐标、移动速度、移动方向,选择与邻居节点平均链接率最大的节点作为簇头。

#### 3.2.1 圆弧运动轨道半径的计算

Ad Hoc 网络中节点的运动具有随机性,但由于需要提供某种服务,因此节点的运动仍具有一定的规律性,文中通过节点已有的历史位置来预测节点的运动轨迹,可将节点运动的轨迹抽象为直线型和圆弧形两种<sup>[9]</sup>,其中直线型为圆弧形运动时半径较大时的近似情况。

定义 1:对于规则运动节点,已知轨道上的历史位置两点  $A$  和  $B$ , $r$  为节点间传输的有效距离, $\lambda$  为可调节系数(这里直接用 1),由节点  $A$  和节点  $B$  计算出轨道半径  $R$ ,若满足条件  $R \leq \lambda r$ <sup>[9]</sup> 则认为此节点的运动轨道为圆弧轨道,否则近似看作直线轨道。

对于规则运动节点,已知轨道上两点  $A$  和  $B$ ,由图 1 可得节点规则运动轨道半径  $R_{AB}$  为:

$$R_{AB} = \frac{l_{AB}}{2\sin\theta}$$

其中, $l_{AB} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}$ ;  $\theta$  为  $AB$  向量与  $v_B$  的夹角,则有:

$$R_{AB} = \frac{l_{AB}}{2} \times \sqrt{1 + \left(\frac{1 + k_{AB}k_B}{k_{AB} - k_B}\right)^2} \quad (1)$$

其中, $k_{AB} = \frac{y_B - y_A}{x_B - x_A}$ ;  $k_B = \tan\theta_B$ ,  $\theta_B$  为  $B$  节点运动方向。

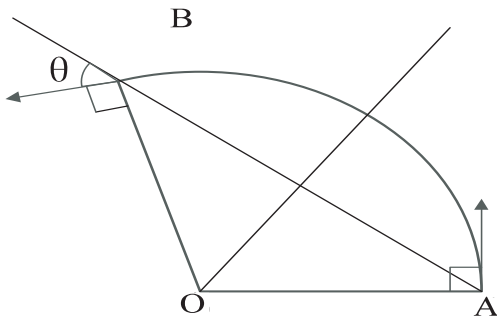


图 1 模拟节点的规则运动情况

定义 2:对于无规则运动节点,随机选择历史轨迹中时间间隔为  $\Delta t$  的两点,根据公式(1)计算出此  $\Delta t$  时间的轨道半径,按照上述的方法共选择  $m$  次,得到其对应的半径系列为  $\{R_1, R_2, \dots, R_m\}$ 。已知  $P$  为此半径系

列中满足半径  $R \leq \lambda r$  的概率, $P^1$  为可调节系数(这里直接用 0.6),若  $P \geq P^1$ ,则认为此节点的运动轨道为圆弧轨道,否则近似认为节点运行轨道为直线轨道。

说明: $\Delta t$  足够小,可近似看作  $\Delta t$  时间内节点的运动为规则运动; $m$  作为统计值个数不能太小。

分析节点运动轨迹,通过 GPS 获得  $m$ (这里直接用 10) 个历史间隔  $\Delta t$  时间的节点系列  $\{a_1, a_2, \dots, a_{m+1}\}$ ,对系列中连续两点  $a_i$  和  $a_{i+1}$ ( $i$  从 1 到  $m$ ) 根据定义 1 和公式(1) 计算半径,得到半径系列  $\{R_1, R_2, \dots, R_m\}$ ,根据定义 2 判断节点运动模式,若为直线运动,则不进行任何处理;若为弧形运动,则选取此半径系列中满足  $R_i \leq \lambda r$  的半径构成新系列  $\{R_1^1, R_2^1, \dots, R_n^1\}$ ,则节点不规则弧形运动轨道半径  $R$  为:

$$R = \frac{1}{n} \sum_{i=1}^n R_i^1$$

#### 3.2.2 不同运动模式下预测节点的链接率

##### (1) 直线运动轨迹预测。

已知节点运动的预测起始点  $(x_1, y_1)$ ,速度为  $v_1$ 、方向为  $\theta_1$ ,时间为  $t$  的直线运动轨迹为:

$$\begin{cases} x = x_1 + v_1 t \cos\theta_1 \\ y = y_1 + v_1 t \sin\theta_1 \end{cases} \quad (2)$$

##### (2) 弧形运动轨迹预测。

已知节点运动的预测起始点  $A(x_A, y_A)$ 、历史节点  $B(x_B, y_B)$ 、轨道半径  $R$ 、 $A$  的运动方向  $\theta_A$ ,根据图 2 得两点坐标:

$$\begin{aligned} O \text{ 点} & \begin{cases} x_o = x_A - R \sin\theta_A \\ y_o = y_A + R \cos\theta_A \end{cases} \\ O^1 \text{ 点} & \begin{cases} x_{o^1} = x_A + R \sin\theta_A \\ y_{o^1} = y_A - R \cos\theta_A \end{cases} \end{aligned} \quad (3)$$

已知  $l_{BO}$  表示节点  $B$  与  $O$  点的距离, $l_{BO^1}$  表示节点  $B$  与  $O^1$  的距离。若  $l_{BO} > l_{BO^1}$ ,则圆心为  $O^1$ ,否则为  $O$ 。

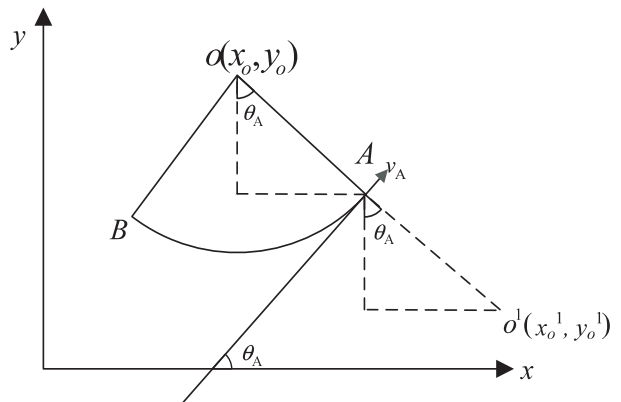


图 2 计算节点的移动轨道圆心

##### (3) 平均链接率的计算。

取一个足够大的时间值  $t_{old}$  作为计算链接率的预测总时间,根据文献[9 - 10] 可得两节点链路链接率:

(a) 根据公式(2),两直线轨道节点*i*和*j*的预测最大保持时间为

$$T_{\max} = \frac{-(ab + cd) + \sqrt{(a^2 + c^2)r^2 - (ad - bc)^2}}{a^2 + c^2}$$

其中, $a = v_i \cos \theta_i - v_j \cos \theta_j$ ;  $b = x_i - x_j$ ;  $c = v_i \sin \theta_i - v_j \sin \theta_j$ ;  $d = y_i - y_j$ ;  $r$ 为节点间传输的有效距离。则预测链接率为:

$$P = \begin{cases} \frac{T_{\max}}{t_{\text{std}}}, T_{\max} \text{ 值确定} \\ 1, T_{\max} \text{ 值为无穷大} \end{cases} \quad (4)$$

(b) 根据公式(3),两圆轨道节点*i*和*j*的上限间距为:

$$D_{\max} = \sqrt{(x_{oi} - x_{oj})^2 - (y_{oi} - y_{oj})^2} + R_i + R_j$$

其中, $(x_{oi}, y_{oi})$ 和 $(x_{oj}, y_{oj})$ 分别为*i*节点和*j*节点的运动轨道圆心坐标; $R_i$ 和 $R_j$ 分别为*i*节点和*j*节点的运动轨道半径。

则预测链接率为:

$$P = \begin{cases} 1, D_{\max} \leq r \\ \frac{r}{D_{\max}}, D_{\max} > r \end{cases} \quad (5)$$

(c) 根据公式(2)和(3),圆轨道节点*i*与直线轨道节点*j*预测估计上限间距为:

$$D_{\max} =$$

$\sqrt{(x_j + v_j t \cos \theta_j - x_{oi})^2 + (y_j + v_j t \sin \theta_j - y_{oi})^2} + R_i$   
要保持链接需要  $D_{\max} \leq r$  有解,若有解为  $\{t \mid 0 \leq t \leq t^1\}$ ,若无解则近似认为链接率为0。则预测链接率为:

$$P = \begin{cases} 0, t^1 \text{ 不存在} \\ \frac{\int_0^{t^1} (r - D_{\max}) dt}{\int_0^{t^1} (r - D_{\max}) dt + \int_{t^1}^{t_{\text{end}}} (D_{\max} - r) dt}, t^1 \text{ 存在} \end{cases} \quad (6)$$

根据公式(4)~(6)得节点的平均链接率为:

$$P = \sum_{i=1}^n \frac{p_i}{n}$$

其中, $n$ 为节点的邻居个数; $p_i$ 为该节点与其第*i*个邻居的链接率。

### 3.3 AOWLR 具体描述

(1) 初始化,每个节点*i*通过周期性交互“Hello”消息获得邻居节点个数作为自己的度数,记为 $d_i$ ;

(2) 计算各节点*i*的度数与其理想度数 $d_{\text{ideal}}$ 的差值,即 $D_i = |d_i - d_{\text{ideal}}|$ ;

(3) 计算各节点*i*与其邻居节点距离的平均值 $S_i$ ;

(4) 计算各节点*i*能量损耗 $E_i$ ;

(5) 计算各节点*i*与其邻居节点的平均链接率 $P_i$ ;

(6) 确定权重系数,根据权重计算公式计算出每个节点*i*的权重 $W_i$ ,然后将 $W_i$ 放置在周期性的“Hello”消息中广播出去;

(7) 选取相邻节点中 $W_i$ 值最小的节点作为簇头,若有多个节点的权值 $W_i$ 处于相同的最小值,则选取其中ID较小的节点作为簇头并广播“Hello”消息,通知其他节点自己是簇头。第一次收到此消息的邻居节点作为其成员节点并广播“Hello”消息,通知其他节点自己是该簇头的成员节点,不再参加接下来的簇头选举。孤立节点自动成簇;

(8) 对上一步剩余的节点重复步骤(2)~(7),直到所有节点都位于簇中。

## 4 仿真结果

文中采用 NS2<sup>[11]</sup>进行仿真实验。在此模拟环境中不考虑噪声、分组传输差错和冲突,当节点到达区域边界时,它将向区域内反射。仿真实验采用30个节点随机均匀分布在一个 $100 * 100$ 单位面积的区域,运动模型为随机运动模型<sup>[12]</sup>,仿真时间为1000单位时间,节点移动方向范围为 $[0, 2\pi]$ ,移动速度范围为 $[0, 5]$ ,节点初始能量为1000单位能量,簇成员节点单位时间消耗的能量为0.1单位能量,簇头的理想度数为8,定义2中的 $\Delta t$ 、 $m$ 取值分别为0.1、15,预测总时间 $t_{\text{std}}$ 为20单位时间,AOWLR权值系列 $\{w_1, w_2, w_3, w_4\}$ 的取值分别为 $w_1 = 0.7, w_2 = 0.2, w_3 = 0.05, w_4 = 0.05$ ,LOWID、HIGHD、AOW均采用上述相同的参数设置。文中通过节点间传输的有效距离 $r$ 取不同的值获得不同的 $C$ 、 $U$ 、LBF和HFI来评价LOWID、HIGHD、AOW、AOWLR的性能。

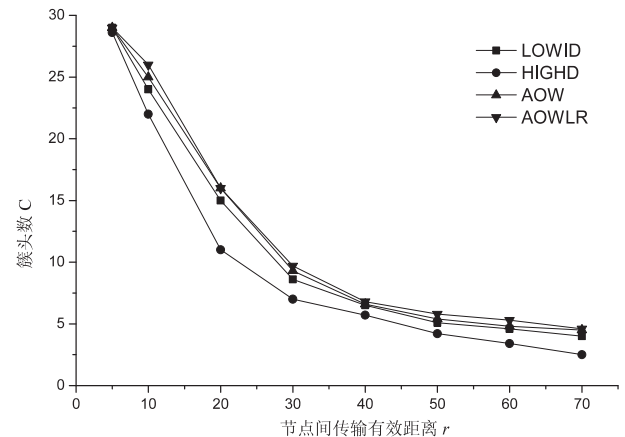


图3 簇头数随传输有效距离的变化

图3表明,在相同条件下4种分簇算法获得的簇头数均随节点间传输有效距离的增大而减小,这是因为簇中节点的个数将随传输有效距离的增大而增大,从而导致簇的个数减少。当传输有效距离很小时,基本上每个节点独立成簇,簇头数达到最大值;当传输有

效距离很大时,整个网络近似成单个簇。

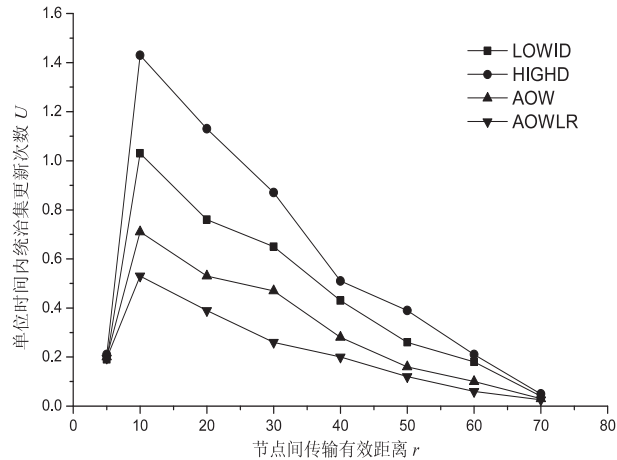


图 4 统治集更新频率随传输有效距离的变化

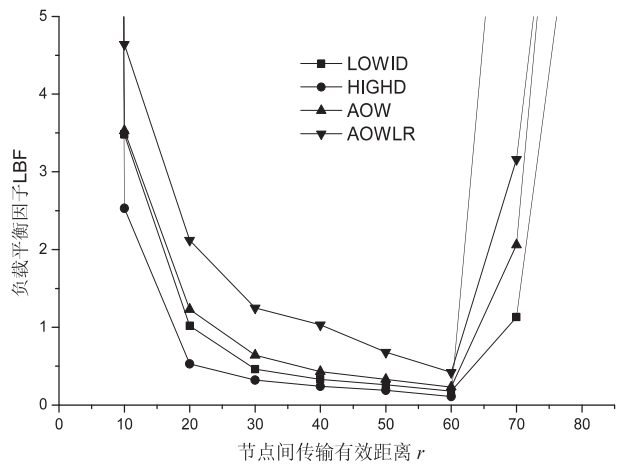


图 5 LBF 随传输有效距离的变化

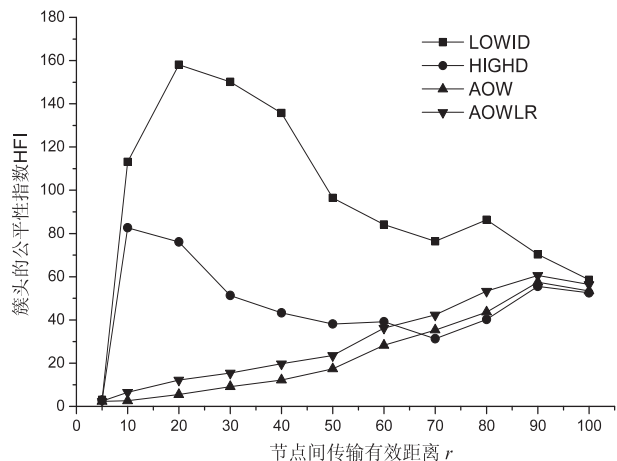


图 6 HFI 随传输有效距离的变化

由图4可知,节点间传输有效距离超过10时,随着传输有效距离的增大簇的规模不断增大,同时簇结构将更加稳定,因此单位时间内统治集更新次数  $U$  将随着传输有效距离的增大而减小。当传输有效距离很小时,各节点近似成簇,簇结构基本不会更新而导致  $U$  值很小;当传输有效距离很大时,整个网络近似成单个簇而导致  $U$  值很小。在相同条件下,4 种分簇算法中 AOWLR 能够预测链接率,具有更好的稳定性,故  $U$  值

相对其他算法较小。

图 5 表明,在相同条件下 4 种分簇算法的 LBF 随着节点间传输有效距离的增大先减小后增大。当传输有效距离很小时,每个节点近似成簇,负载非常均衡,因此 LBF 值非常大;当传输有效距离很大时,整个网络近似成单个簇,此时得到一个完全平衡的系统,因此 LBF 值很大。AOW 与 AOWLR 综合考虑了多种因素,让簇结构在尽量稳定情况下簇头均匀分布,有较好的负载平衡性,故 LBF 值相对其他算法较大。

由图 6 可知,LOWID 倾向于选择 ID 较小的节点充当簇头,HIGHD 倾向于选择度数最大的节点充当簇头,这两个算法使得各节点作为簇头的时间偏差较大,因此 HFI 值较大。AOW 和 AOWLR 根据多种因素选择簇头,使得各节点公平性地充当簇头,因此其 HFI 值较 LOWID 和 HIGHD 小。4 种分簇算法下,当传输有效距离很小时,每个节点近似成簇,非常公平,因此 HFI 值很小,AOW 和 AOWLR 随着传输有效距离的增大,簇头数减少,公平性降低,HFI 值增大。而 LOWID 和 HIGHD 公平性指数先增大后减小。当传输范围很大时,4 种分簇算法下的簇结构均近似为单个簇,簇结构基本相同,即最终收敛于这个相同的 HFI 值。

### 5 结束语

文中针对 AOW 稳定性不足的问题,细化了节点度数和能量的计算方法,提出了预测节点链接率的改进型自适应加权分簇算法。通过预测节点的运动轨迹获得节点间的链接时间,进而计算节点与其邻居节点的平均链接率。该算法获得的簇结构较 AOW 稳定,减少了算法的执行频率。

### 参考文献:

- [1] 郑少仁,王海涛,赵志峰,等. Ad Hoc 网络技术[M]. 北京:人民邮电出版社,2005.
- [2] Lin C R, Gerla M. A distributed architecture for multimedia in dynamic wireless networks[C]//Proc of IEEE GLOBECOM. [s. l.]: [s. n.], 1995: 1468-1472.
- [3] Parekh A K. Selecting routers in Ad-Hoc wireless networks [C]//Proceedings of the SBT/IEEE international telecommunications symposium. [s. l.]: [s. n.], 1994.
- [4] 王海涛,田 畅,郑少仁. 一种新型的 Ad Hoc 网络分簇算法及其性能仿真[J]. 系统仿真学报,2003, 15(2): 193-197.
- [5] 吴 迪,刘英学,冯永新,等. Ad Hoc 网络中一种基于权值的分簇算法[J]. 小型微型计算机系统,2006, 27(2): 202-206.
- [6] Xu Kaixin, Hong Xiaoyan, Gerla M. An Ad Hoc network with

表2 实验结果统计表1

词本身特征种类	系统识别出的新词	识别正确的新词	召回率 P/%	正确率 R/%
T(B)	118	28	28	23.68
T(B+Context)	115	36	36	31.36
T(B+MI)	113	35	35	30.86
T(B+Context+MI)	116	41	41	35.26
T(B+FF)	110	33	33	29.96
T(B+Context+FF)	99	40	40	40.76
T(B+MI+FF)	93	36	36	38.62
T(B+Context+MI+FF)	96	43	43	45.12

其他条件相同,选取全部词特征,不同的核函数进行实验,结果见表3。

表3 实验结果统计表2

核函数种类	惩罚因子	系统识别出的新词数	识别出正确的新词数	召回率 P/%	正确率 R/%
径向基核函数(RBF)	$C_+ = 0.0001$ $C_- = 0.3$	96	43	43	45.12
多项式核函数	$C_+ = 0.0001$ $C_- = 0.3$	112	38	38	33.93
Sigmoid核函数	$C_+ = 0.0001$ $C_- = 0.3$	139	28	28	20.15

通过实验发现,采用径向基核函数(RBF)时新闻识别系统的召回率和正确率最优。

### 3 结束语

该方法研究的重点是支持向量机结合约束条件进行分类来识别新闻,从得到的召回率来看,达到了预期的目标,因此该方法对于中文新闻识别的召回率的提高有一定的促进作用,通过实验分析,得到以下结论:文中选取的语料具有普遍性,根据支持向量机的小样本训练对大样本具有良好分类的特性以及实验数据分析得出该方法对广泛的语料的新词识别具有较好的兼容性;通过对分类图像的分析可以看出在实验的图像

(上接第97页)

mobile backbones[C]//Proc of IEEE international conference on communications. [s.l.]:[s.n.],2002:3138-3143.

[7] ElBatt T A, Krishnamurthy S V, Connors D, et al. Power management for throughput enhancement in wireless Ad-Hoc networks[C]//Proc of IEEE international communications. [s.l.]:[s.n.],2000:1506-1513.

[8] 钟明洋,符云清,肖磊,等.基于稳定度的Ad Hoc网络分簇算法[J].计算机工程,2012,38(17):98-101.

[9] 林要华,胡华平.基于轨道预测的自适应Ad Hoc分簇算法[J].计算机工程与科学,2010,32(2):27-30.

中最终新闻识别的点距离分界线也就是支持向量的距离较小,这就说明该方法还有提升的空间。

实验表明采用径向基核函数时具有较好的召回率和正确率,下一步工作重点是选取更加适合的支持向量机核函数,考虑使用混合核函数进行实验,以达到最佳的分类效果,同时可以考虑建立动态词典,在掌握全局信息的基础上进行新闻识别以提高系统的效率。

### 参考文献:

[1] 钱秋银,张正兰.一种基于多分类SVM的相关反馈图像检索方法[J].计算机技术与发展,2009,19(8):65-68.

[2] 吕红良.基于大规模语料库的中文新闻识别[D].大连:大连理工大学,2008.

[3] 徐亮.中文新闻识别研究[D].大连:大连理工大学,2008.

[4] Jeroslow R, Wang J. Solving propositional satisfiability problems[M]//Annals of mat hepatics and artificial intelligence. [s.l.]:Springer,1990.

[5] 崔世起.中文新闻检测与分析[D].北京:中国科学院,2006.

[6] 秦浩伟,步丰林.一个中文新闻识别特征的研究[J].计算机工程,2004,30:369-370.

[7] 闫友彪,陈元琰.机器学习的主要策略综述[J].计算机应用研究,2004(7):4-10.

[8] 邹纲,刘洋,刘群,等.面向Internet的中文新闻语检测[J].中文信息学报,2004,18(6):1-9.

[9] Nie Jianyun. Unknown word detection and segmentation of Chinese using statistical and heuristic knowledge[J]. Communications of COLIPS,2008,5(1&2):47-57.

[10] 徐远方,李成城.基于SVM和词特征的新词识别研究[J].计算机技术与发展,2012,22(5):134-136.

[11] Li Chengcheng, Xu Yuanfang. Based on support vector and word features new word discovery research[C]//Proceedings of 2012 IEEE international conference on computer science and automation engineering. [s.l.]:[s.n.],2012:166-168.

[10] Su W, Lee Sung-Ju, Gerla M. Mobility prediction in wireless networks[C]//Proc of the 21st century military communications conference. [s.l.]:[s.n.],2000:491-495.

[11] 柯志亨,程荣祥,邓德隽. NS2仿真实验-多媒体和无线网络通信[M].北京:电子工业出版社,2009.

[12] Royer E M, Melliar-Smith P M, Moser L. An analysis of the optimum node density for ad hoc mobile networks[C]//Proc of IEEE international conference on date. [s.l.]:[s.n.],2001:857-861.

# 基于链接率的Ad Hoc自适应按需加权分簇算法

作者: [杜国勇](#), [束永安](#), [DU Guo-yong](#), [SHU Yong-an](#)  
作者单位: [安徽大学 计算机科学与技术学院, 安徽 合肥, 230601](#)  
刊名: [计算机技术与发展](#)

---

ISTIC

英文刊名: [Computer Technology and Development](#)

---

年, 卷(期): 2014(1)

本文链接: [http://d.wanfangdata.com.cn/Periodical\\_wjfz201401024.aspx](http://d.wanfangdata.com.cn/Periodical_wjfz201401024.aspx)