

一种基于 ARMA 的 WSN 非均衡分簇路由算法

韩志杰^{1,2,3}, 王汝传^{1,3}, 凡高娟¹, 肖 甫¹

(1. 南京邮电大学计算机学院, 江苏南京 210003; 2. 河南大学计算机与信息工程学院, 河南开封 475004;
3. 苏州大学计算机与技术学院, 江苏苏州 210006)

摘 要: 针对无线传感器网络中分簇路由算法中存在的“热区”问题, 提出了一种基于虚拟区域划分的非均衡簇路由算法. 算法将簇划分的任务交由能量无限制的汇聚节点完成, 使得靠近汇聚节点的内层簇的规模小于外层簇的规模. 在簇的结构中引入了主、从簇头节点, 从而实现了分布式簇头选举工作, 同时在分簇过程中避免了每个阶段的能量消耗. 将 ARMA 预测模型引入到主簇头节点的更换过程中, 从而避免了主簇头因为能量完全消耗而死亡, 也避免了因为主簇头死亡而造成网络分割, 降低网络的生存时间, 利用 NS2.31 仿真平台对基于虚拟区域划分的非均衡簇路由算法进行了仿真验证, 结果表明与传统路由算法相比, 该算法延长了 WSN 的生存时间, 有效提高了 WSN 网络健壮度.

关键词: 无线传感器网络; 路由; ARMA

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2010) 04-0865-05

An Unequal Clustering Algorithm in Wireless Sensor Networks Based on ARMA

HAN Zhi-jie^{1,2,3}, WANG Ru-chuan^{1,3}, FAN Gao-juan¹, XIAO Fu¹

(1. College of Computer, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China;
2. School of Computer and Information Engineering, Henan University, Kaifeng, Henan 475004, China;
3. School of Computer Science & Technology, Soochow University, Suzhou, Jiangsu 210006, China)

Abstract: Proposes unequal clusters of routing algorithm based on divided virtual region for “thermal region” in the wireless sensor network. The routing algorithm gives the task of clusters divided to sink node with non-limited energy, therefore, the scale of inside cluster layer which near the sink node is smaller than outer layer cluster. To realized the distributional election work of cluster head and avoided energy consumption in each wheel, we take main and vice cluster head into the structure of clusters. The ARMA forecast model has been intruduced into the switch process of main cluster head, therefor it has not only avoided main cluster head dies because of energy consumes completely, but also avoided the network divided as well as the short life of the network caused by the death of main cluster head. Compared with the tradition routing algorithm, the simulation result in the NS2.26 indicated that this routing algorithm prolonged the life of WSN.

Key words: wireless sensor networks (WSN); route; ARMA

1 引言

路由算法在无线传感器网络中都起着至关重要的作用, 采用何种路由算法决定了最终的数据路由路径, 直接影响网络的整体性能, 针对传感器网络能量受限、资源受限的自身特点, 设计一种性能好的路由协议是非常具有挑战性的^[1,2].

在以分簇方式组织的传感器网络中, 传感器节点的角色分为两种: 簇头和簇成员. 簇头作为簇的中心负责

簇结构的建立, 收集簇成员的数据, 经融合处理后发送给汇聚节点. 由于簇头距离汇聚节点的距离一般较远, 有研究表明, 采用通过簇头组成的骨干网实现多跳路由的方式更有利于节约能量^[3]. 然而这种做法带来了一个能量消耗不均衡的问题: 靠近汇聚节点的簇头节点由于需要转发大量来自其它簇的数据而负担过重, 过早耗尽自身能量而失效, 造成网络分割, 降低网络存活时间. 研究者称这个问题为“热区”(Hot spots)问题^[4].

在传统的同构网络分簇协议中, 通过周期性地重新

收稿日期: 2008-11-15; 修回日期: 2009-11-19

基金项目: 国家自然科学基金 (No. 60573141, 60773041); 江苏省自然科学基金 (No. BK2008451); 国家高科技 863 项目 (No. 2007AA01Z404, 2007AA01Z478); 南京市高科技项目 (No. 2007 软资 106); 现代通信国家重点实验室基金 (No. 9140C1105040850); 江苏省博士后基金 (No. 0801019C); 江苏高校科技创新计划项目 (No. CX08B-085Z, CX08B-086Z)

选择簇头(如 LEACH^[5]、HEED^[6]),可以平衡簇头与簇成员节点之间的能量消耗,但不能解决上述的“热区”问题.以节点的剩余能量为依据选择簇头的方法(如 EECS)也不能完全解决这个问题,因为它只是在网络的局部比较节点剩余能量,无法从整体上协调节点的能量消耗以解决“热区”问题.可以认为上述两种方法都是以被动的方式来均衡网络中节点的能量消耗,即在网络中出现能量消耗不均衡之后方采取措施来均衡能量消耗.

在以周期性选择簇头的分簇协议中,往往是基于平衡簇头与簇成员节点之间的能量消耗这一出发点考虑的,通常是所有传感器节点动态的形成多个簇,且经过一段时间的数据传输之后,会进行重新分簇.这样一来,在出现“热区”问题的网络中就会存在两种情况:

(1)整个网络主动进行簇头选举工作,即存在簇头节点即将能量耗尽之前发出重新选举簇头的请求,然后进入新一轮簇头选举阶段.但是处于“热区”中的簇头节点能量耗费较高,当进行新一轮簇头选举时,处于非“热区”中的簇头也许此时能量还很充裕.因此,这种情况下进行这个网络的簇头选举工作对于那些处于非“热区”中的能量充裕的簇头来说是不公平的.

(2)整个网络被动进行簇头选举工作,即网络运行一段时间之后,不管网络是否需要重新选择簇头都会进行一次新的簇头选举工作.此时处于“热区”和非“热区”的簇头节点之间也会由于消耗能量的不同产生矛盾,当两次簇头选举工作的间隔较长时,处于“热区”中的簇头节点会因为能量消耗过大而提前进入失效状态,从而造成网络分割,降低网络存活时间;当两次簇头选举工作的间隔较短时,整个网络会因为频繁的进行簇头选举工作而消耗过多的能量,从而造成网络生存时间缩短.

以上两种情况都会造成节点(簇头节点和普通节点)宝贵能量的白白耗费.

通过对传感器节点能量消耗情况和无线传感器网络路由协议的研究,我们提出一种基于 ARMA 流量预测的单元格带管理节点的分簇路由协议 ACRP(ARMA-Based Cluster Routing Protocol),该协议的主要思想是利用可补充能量的汇聚节点采用集中式方式将传感器节点覆盖的区域固定划分为许多单元格,每个单元格作为一个簇.在每个簇中设置一个管理节点,管理节点基于 ARMA 模型的流量预测,以分布式方式实现每个簇中簇头的重新选举,以减少成簇的开销,实现簇头的均衡能量消耗,同时增强网络的健壮性,延长 WSN 网络生存时间.

2 ACRP 协议

假设传感器节点随机分布在一个正方形监测区域

内,并且该传感器网络具有以下性质:

(1)以汇聚节点为中心建立无线传感器网络环境,节点均匀分布于汇聚节点四周,以汇聚节点为原点建立坐标系.

(2)汇聚节点具有较强的计算、存储能力,且能量能无限制.

(3)所有传感器节点部署后静止不动,并且具有相同的初始能量值,具有相似的处理通信能力,地位是对称平等的.

(4)所有节点都是同构的,具备数据融合的功能.

(5)根据接收者的距离远近,节点可以自由调节其发送功率以节约能量消耗.

(6)所有节点可以获知自己的坐标位置.

2.1 更换簇头节点

簇头更新存在的问题有网络中每个簇的簇头是否有权利发起更换簇头的请求;每个簇的簇头在什么样的情况之下才能发起更换簇头的请求;簇头的更换是集中进行还是分布进行等.比如在簇头与汇聚节点之间通信采用多跳方式(即通过簇头组成的骨干网实现的多跳路由)更有利与节约能量,但是这种做法也带来了一个能量消耗不均衡的问题:在这种所有传感器节点的数据都发送到汇聚节点的“多对一”数据传输模式中,靠近汇聚节点的簇头由于需要转发大量来自其它簇的数据而负担过重,在该节点担当簇头的一轮中,会因为更换簇头的时间未到,或者是没有单独更换簇头的权利(即分布式簇头更替,每个簇的簇头更替是根据自身情况进行),而过早耗尽自身能量而失效.

本文基于 ARMA 流量预测模型应用到了主簇头节点的更换过程中,提出了一种基于 ARMA 流量预测机制的主簇头节点更换策略.

流量模型必须有可管理的参数个数,参数估计必须简单.有许多随机模型用于描述网络流量,通信网络中的流量模型可以分为稳定的和不稳定的两种,稳定的流量模型又分为两类:短相关和长相关.短相关模型包括马尔可夫过程和自回归模型(AR),自回归滑动平均模型(ARMA),以及自回归求和滑动平均模型(ARIMA);长相关流量模型包括:F-ARIMA 和 Fractional Brownian Motion,我们采用 ARMA 模型,它不能建立长相关的模型,因此只能进行短期实时预测.

假设数据流量序列为 $X_0, X_1, \dots, X_i, \dots, X_n$,我们这里采用 ARMA(2,1)来进行预测,其模型为

$$\begin{aligned} \varphi(B)X_i &= \theta(B)a_i \\ \varphi(B) &= 1 - \varphi_1 B - \varphi_2 B^2 \\ \theta(B) &= 1 - \theta_1 B \end{aligned} \quad (1)$$

其中 B 是后移算子, a_i 是白噪声. $\varphi_1, \varphi_2, \theta_1, \sigma_a^2$ (白噪声

方差)为估计参数. ARMA 相关矩估计方法主要有最小二乘估计方法,最大似然估计,最大熵估计等等,考虑到传感器节点的计算能力,这里我们采用最小二乘估计方法可解出 $\hat{\varphi}_1, \hat{\varphi}_2, \hat{\theta}_1, \hat{\sigma}_a^2$.

判断时间序列的稳定性,其稳定性条件为 $\hat{\varphi}_1 + \hat{\varphi}_2 < 1, \hat{\varphi}_2 - \hat{\varphi}_1 < 1, |\hat{\varphi}_2| < 1$, 如果满足此条件则视为平稳序列. 则得出 ARMA 拟和模型为:

$$X_t = \hat{\varphi}_1 X_{t-1} + \hat{\varphi}_2 X_{t-2} + a_t - \hat{\theta} a_{t-1} \quad (2)$$

进一步地,我们利用逆函数法进行一步预测, ARMA 的逆函数记为 $I_1, I_2, \dots, I_n$,

$$I_1 = \hat{\varphi}_1 - \hat{\theta}_1, I_2 = \hat{\varphi}_2 - I_1 \hat{\theta}_2, I_3 = I_j \hat{\theta}_1 \dots (j > 3) \quad (3)$$

则一步预测模型如式(14)所示,其中 m 为 X_t 之前 m 次观测值,可根据预测精度的要求取值.

$$X_t(1) = \sum_{j=1}^m I_j X_{t+1-j} \quad (4)$$

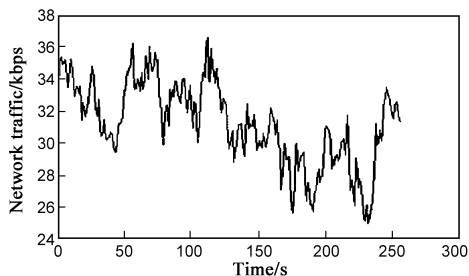
其多步预测模型为:

$$\hat{X}_t(l) = \hat{\varphi}_1 \hat{X}_t(l-1) + \hat{\varphi}_2 \hat{X}_t(l-2) \quad (5)$$

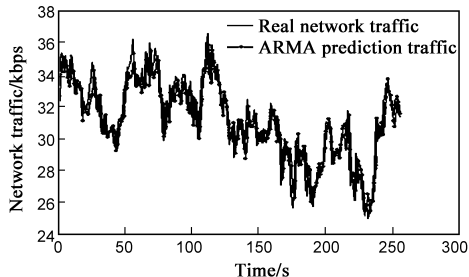
在基于 ARMA 模型的主簇头节点更换算法中,主簇头节点根据数据流量的预测,一经发现数据流量将超过本节点负载的电池能量消耗,则立即发送更替主簇头节点的请求.

主簇头节点更换算法的主要思想,引入预测机制,主簇头节点在预测其自身能量将要消耗殆尽前发出更替主簇头节点的请求,从而避免了主簇头因为能量完全消耗而死亡,也避免了因为主簇头死亡而造成网络分割,降低网络的生存时间.

为判断模型的预测精度,我们利用 Matlab7.0 进行了仿真验证.如图 1(a)所示为一个 WSN 中主路由上的



(a) 原始 WSN 流量



(b) ARMA 一步预测

图1 基于 ARMA 的 WSN 流量预测

某一传感器节点的真实网络流量,从任意时刻起在 256s 内,每秒采样数据流量 1 一次,利用本模型可计算出其估计值,其一步预测的比较图如图 1(b)所示.从图中可以看出,预测曲线和真实网络流量曲线比较吻合,说明 ARMA 预测模型能够有效预测 WSN 网络流量,作为一种短时相关预测模型,ARMA 预测模型的一步预测取得了较好的预测效果,且模型简单,运算为线性运算,复杂度低,能够作为一种实时网络流量预测模型预测 WSN 网络流量.

2.2 路由机制

在某一普通节点完成数据监测任务后,将会传送监测数据至主簇头节点,然后主簇头节点收集簇内监测节点采集数据后,进行数据融合,再将融合数据发送到汇聚节点.

普通节点为了能够将数据发送到主簇头节点,将需要维护自己邻居节点的简单的路由信息表以及主簇头节点和从簇头节点的相关信息.当自身的能量消耗超过预先设定的层次阈值 Δ_E 时,该节点会将其新的能量层次值进行广播(NOP 消息),其邻节点收到 NOP 消息后,便更新自己路由信息表中的数据.

在本文提出的簇内数据传递规则中,簇内节点并不像 LEACH 算法那样采用单跳方式直接传送采集数据到主簇头节点,它采用多跳路由的方式将数据传送给主簇头节点.现在我们以图 2 中的普通节点 4 为例,来说明如何实现以多跳的方式将数据传递给主簇头节点.

第一步:节点 4 查找自己的路由信息表,获取节点 3 和节点 2 的信息.

第二步:设节点 4 和节点 3 的距离为 d_1 ,节点 4 和节点 2 的距离为 d_3 ,节点 3 和节点 2 的距离为 d_2 .假设 $d_3^2 < \delta(d_1^2 + d_2^2)$,那么节点 4 将数据传递给节点 2.

第三步:节点 2 检查自距离主簇头节点的距离 l ,发现 $l < 3$,则直接将数据传递给主簇头节点.采用如下步骤进行处理:

■ 主簇头节点 ◆ 从簇头节点 ● 普通节点

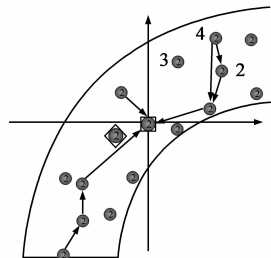


图2 簇内数据传递规则示意图

(1)如果 $l < 3$, (z, l) 节点直接传送数据到主簇头节点,否则转入下一步.

(2)计算节点 (z, l) 和节点 $(z, l-1)$ 的距离 $d_1, (\theta,$

$f, z, l)$ 和节点 $(z, l-2)$ 的距离 d_3 , 节点 $(z, l-1)$ 和节点 $(z, -2)$ 的距离 d_2 . 如果 $d_3^2 > \delta(d_1^2 + d_2^2)$, δ 是根据实际网络情况确定, 则传送数据到节点 $(z, l-1)$, 节点 $(z, l-1)$ 将本身采集数据和所接收数据进行数据融合, 然后依据该数据传递规则继续发送融合数据到下一跳节点, 否则转入下一步.

(3) 如果 $l < 3$, 则等同于第一步, 否则以当前节点为源节点, 转入(2)操作.

总之, 节点 (z, l) 在传送数据到主簇头节点的时候, 会依次判断是否把节点 (z, t) , $t < l$ 作为下一跳节点. 如果是, 则节点 (z, l) 传送采集数据到节点 (z, t) , 节点 (z, t) 将采集数据与接收数据进行融合计算, 进行下一步传送; 否则, 节点直接传送数据到主簇头节点.

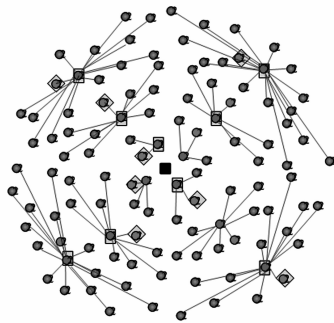
在网络运行的稳定阶段也即数据传输阶段, 主簇头节点负责收集簇内普通节点采集的数据, 然后对数据进行融合计算, 然后通过多跳路由方式将融合数据通过邻居主簇头节点转发至汇聚节点. 具体过程描述如下:

假设需要传输数据的主簇头节点为 $(0, 2, 0, 0)$, 记为 MCH_i , 距离汇聚节点的距离记为 D_i ; 记主簇头节点 MCH_i 相邻 m 个主簇头节点集合为 $S_n = \{MCH_j, \dots, MCH_m\}$, 主簇头节点 $MCH_i, MCH_j, \dots, MCH_m$ 和汇聚节点之间的距离分别为 $D_i, D_j, \dots, D_k$.

具体路由过程可描述如下:

第一步: 设存在主簇头节点集合 $S_v = \{MCH_v | MCH_v \in S_n \& D_v < D_i\}$, 如果集合 S_v 为空, 则主簇头节点直接

■ 汇聚节点 □ 主簇头节点 ◇ 从簇头节点 ● 普通节点



■ 汇聚节点 □ 主簇头节点 ◇ 从簇头节点 ● 普通节点

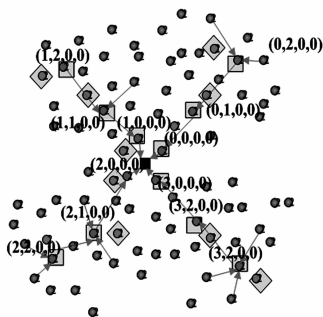


图3 簇间路由示意图

传送融合数据到汇聚节点, 否则转到下一步).

第二步: 在集合 S_v 中选择能量水平值最大的主簇头节点, 如果能量水平值最大的节点是唯一的, 则将该节点作为下一跳路由节点, 否则从多个能量相等的主簇头节点中挑选 D_v 最小的主簇头节点作为下一跳.

第三步: 将该主簇头节点作为路由起始节点, 重复上述操作.

这样, 在每个采集信息的普通节点和汇聚节点之间就形成了一条稳定的路径, 如图6所示.

3 仿真实验

为了评价本文提出算法的性能, 本文在 NS2 仿真软件中对算法进行仿真验证. 仿真环境具体配置如下:

在 $500\text{m} \times 500\text{m}$ 的正方形区域内, 假设汇聚节点的坐标为 $(0, 0)$, 即坐标原点. 以汇聚节点为原点建立坐标系, 则 200 个传感器节点均匀的分布在汇聚节点周围. 并且所有的传感器节点以固定的频率采集数据, 即节点以固定的频率发送采集数据, 带宽为 20KB , 时延为 0.1s , 且数据包的大小是固定的, 均为 2000bit , 所有节点的初始能量值相等, 均为 0.25J , 每层簇头个数

$$\text{actual_}N_{CH(i)} = \alpha \frac{(r_i + r_{i-1})}{r_i - r_{i-1}}, 4 < \alpha < 3\pi, \text{ 仿真中 } \alpha \text{ 取 } 2\pi,$$
 此处 π 为周期律, $\epsilon_{\text{amp}} = 0.0013\text{pJ/bit/m}^2$, $E_{\text{elec}} = 50\text{nJ/bit}$, $\beta = 6$.

3.1 网络的存活节点数

在传感器网络中, 传感器节点一般采用电池供电且不可更换, 因此, 降低能量消耗, 延长网络生命周期是设计网络时要考虑的重要因素之一, 网络的存活节点数表明了随着时间的推移, 仍然存活的节点的总数, 是体现路由协议是否属于能源有效性协议的一个重要指标.

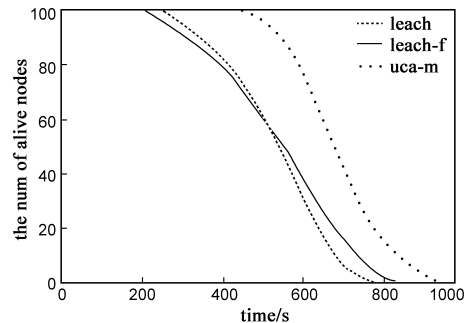


图4 随时间变化网络中生存节点个数

图4表示随时间推移网络中处于存活状态的节点个数. 根据仿真结果, LEACH算法在240s时出现第一个死亡节点, LEACH-F算法在200s时出现第一个死亡节点; 本文提出的UCA-M算法的第一个死亡节点推迟在430s时出现, 因为UCA-M算法利用非均衡簇结构消除了多条路由算法中的“热区”问题, 网络中簇头节点能

量消耗相对平衡,此外,算法还引入了主、从簇头的概念,使网络中簇头的更换可以分布式的进行,从而整个网络的性能得到了良好的改善.从上图,在 430s 到 920s 之间,UCA-M 算法节点死亡的速度明显比前两个算法要快,这是因为在 UCA-M 算法前期,尽量是网络能量均衡消耗;而到了网络运行的后期阶段,由于所有节点的剩余能量水平都很低,因此节点也会在这一阶段相继死亡,网络质量快速下降.

3.2 网络的负载均衡度

在分簇的传感器网络中,簇头扮演着极其重要的角色,簇头能量消耗的快慢直接影响着整个网络的性能好坏.本文主要以网络中簇头节点的负载均衡度来衡量整个网络的负载.具体计算公式如下:

$$LBF = \frac{1}{1 + \frac{1}{m} \sum_{i=1}^m (k_i - \bar{\phi})^2} \quad (6)$$

其中, k_i 表示第 i 个簇头的负载, $\bar{\phi}$ 表示平均负载.

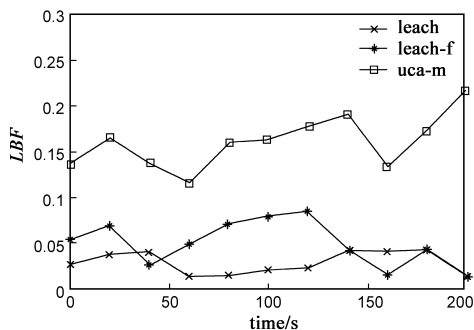


图5 LBF

从图 5 中我们可以很明显的看出, UCA-M 算法的 LBF 比 LEACH 算法和 LEACH-C 算法高,说明了 UCA-M 算法中簇头节点的负载比较均衡.这是因为 UCA-M 算法中,距离汇聚节点近的簇头节点所管理的成员节点要少于距离汇聚节点远的簇头节点,这就可以让距离汇聚节点近的那些簇头节点节省一部分能量用于转发来自外层的监测数据,从而减少了由于瓶颈因素而造成的簇头节点负担过重而提前进入死亡状态. LEACH-C 算法的 LBF 则介于 UCA-M 算法和 LEACH 算法之间,其簇头的选择仅考虑了节点的位置和能量,没有考虑簇的规模问题.

3.3 汇聚节点接收数据量

从图 6 中可以看出, UCA-M 算法中汇聚节点在整个网络生命周期内接收到的数据包总数是最多的.此外,当网络中节点未进入死亡周期的时候,三种算法汇聚节点接收到的数据包总量都随时间线性增长,并且 LEACH 算法线性增长的速度大于 LEACH-F 和 UCA-M 算法,这是因为 UCA-M 算法采用的是多跳路由策略,而 LEACH 和 LEACH-F 算法是簇头直接给汇聚节点发送数

据,在 UCA-M 算法中,由于数据需要经过多个中间簇头节点的转发,所以会存在一定时间的延迟;当节点进入死亡周期后,在单位时间内汇聚节点接收到的数据包都越来越少,直到最后所有节点死亡以后,汇聚节点不再接收数据包,数据包总数达到最大值,不再变化.

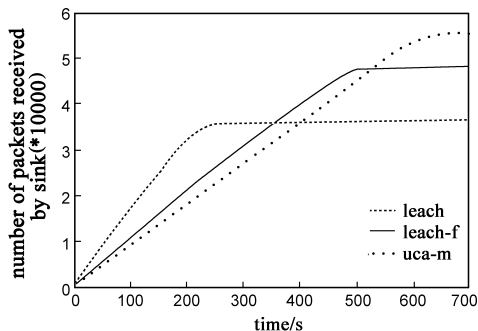


图6 Sink节点接收包数量

4 总结

本文针对无线传感器网络路由协议设计的要求:能量高效、可扩展性、鲁棒性和快速收敛性,基于 ARMA 预测模型主动选举簇头节点,提出了一种带管理节点的分簇路由协议——TCBCRP 协议,详细描述了该协议的工作过程.并在 NS2.31 平台对路由协议进行了仿真,仿真结果表明协议路由具有丢包率低,延迟小特点,全网节点消耗也比较低.

参考文献:

- [1] Fang Q, Zhao F, Guibas Lj. Light weight sensing and communication protocols for target enumeration and aggregation [A]. Proceedings of the 4th ACM International Symposium On Mobile Ad Hoc Networking & Computing [C]. Annapolis, Maryland: ACM Press, 2003. 165 - 176.

(下转第 893 页)

作者简介:

韩志杰 男,1979 年生,河南周口市人,博士研究生,研究方向为计算机软件理论、计算机网络、信息安全与对等计算.

E-mail: hanzhijie@126.com

王汝传 男,1943 年生,安徽合肥人,教授、博士生导师.主要研究方向是计算机软件、计算机网络和网络、对等计算、信息安全、无线传感器网络、移动代理和虚拟现实技术等.

E-mail: wangrc@njupt.edu.cn

凡高娟 女,1983 年生,河南周口人,南京邮电大学计算机学院博士研究生,主要研究方向是无线传感器网络.

E-mail: fangaojuan@126.com

肖甫 男,1980 年生,湖南邵阳人,博士,南京邮电大学通信与信息工程流动站博士后.主要研究领域:卫星网络通信、无线自组织网络、多媒体技术等. E-mail: xiaofu@njupt.edu.cn

(上接第 869 页)

- [2] Younis M, Youssef M, Arisha K. Energy-aware routing in cluster-based sensor networks[A]. Proceedings of the 10th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems[C]. Fort Worth: IEEE Computer Society, 2002. 129 – 136.
- [3] V Mhatre, C Rosenberg. Design guidelines for wireless sensor networks: communication, clustering and aggregation[J]. Ad Hoc Network, 2004, 2(1): 45 – 63.
- [4] C Li, M Ye, G chen, J Wu. An energy-efficient unequal clustering mechanism for wireless sensor networks[A]. Proceedings of the 2th IEEE International Conference on Mobile Ad-hoc and Sensor Systems[C]. Washington, DC: IEEE, 2005. 597 – 604.
- [5] Handy M J, Haase M, Timmermann D. Low energy adaptive clustering hierarchy with deterministic cluster-head selection[A]. Proc of the 4th IEEE Conf on Mobile and Wireless Communications Networks[C]. Stockholm: IEEE Communications Society: 2002. 368 – 372.
- [6] Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks[J]. IEEE Transactions on Mobile Computing, 2004, 3(4): 366 – 379.