

DSCAU:非均衡负载无线传感器网络的基于支配集的分簇数据收集算法^①

奎晓燕^② 张士庚^③ 王建新

(中南大学信息科学与工程学院 长沙 410083)

摘 要 针对无线传感器网络(WSNs)在负载不均衡即各节点数据量不相同情况下的数据收集问题进行了研究,提出了 WSNs 在负载不均衡下的新的基于支配集的分簇数据收集算法——DSCAU。运用 DSCAU 时,每个节点对自身剩余能量、节点邻居数量、自身和邻居产生的数据量等情况进行综合考虑来选举候选簇首。为避免正式簇首过多,候选簇首根据自身邻居被其他候选簇首覆盖的数量,以反比概率成为正式簇首。同时为了均衡簇首的能量开销,对簇的规模进行了限制。理论分析和仿真实验表明,DSCAU 在多跳情况下能有效延长网络生命周期,并且能保证所有节点均加入簇,从而提高成簇算法的覆盖率。

关键词 非均衡负载,无线传感器网络(WSNs),支配集,分簇,数据收集

0 引言

无线传感器网络(wireless sensor networks, WSNs)综合了传感器技术、嵌入式计算技术、分布式信息处理技术和通信技术,能够协作地实时监测、感知、采集网络分布区域内的各种环境或监测对象的信息,并对这些信息进行处理^[1,3]。WSNs 在工业、农业、军事、医疗卫生等多个领域均有广泛的应用。由于 WSNs 的传感器节点能量有限,通信和计算能力较弱,如何使所有节点自组织地形成连通的网络来进行数据收集,以及如何节省能量以延长网络寿命,是目前无线传感器网络研究中的热点。目前 WSNs 进行数据收集多采用分簇算法,但已有的分簇算法普遍只考虑单跳网络而忽视多跳网络,且仅考虑网络负载均衡即网络中节点产生的数据量相同的情况。而实际部署的无线传感器网络通常是多跳的,且负载不均衡,即处于网络中不同部分的节点产生的数据量往往不相同。针对这种情况,本研究提出了一种适用于非均衡负载无线传感器网络的基于支配集的分簇数据收集算法(dominating set based clustering algorithm for data gathering in WSNs with unbalanced traffic load, DSCAU),并通过理论分析和

仿真实验验证了其有效性。

1 数据收集的分簇技术

利用分簇技术来进行数据收集是目前广泛采用的方法^[4]。根据簇首产生的方式,可以分为基于概率的分簇算法和基于支配集的分簇算法。其中,基于概率的分簇算法是在分簇开始时,使网络中的一部分节点同时以一定概率成为簇首;基于支配集的分簇算法是在分簇开始后,使每个时间步产生少量簇首,若干个时间步后,将产生足够的簇首来使网络中的所有节点均能加入簇。基于概率的分簇算法主要是面向单跳网络,而基于支配集的算法主要是面向多跳网络。

(1) 基于概率的分簇算法

低能耗自适应分簇分层(LEACH)^[5]是最早提出的基于概率的分簇数据收集算法。为了克服 LEACH 中能量低的节点成为簇首时容易过早死亡的缺点,混合的低能耗分布式分簇方法 HEED^[6]对网络中选择簇首的过程进行了改进。能量感知的无线传感器网络数据收集协议 EADEEG^[7]为了克服 HEED 的缺点,根据节点自身的能量与邻居的平均能量之间的比值来确定节点成为簇首的概率。但

^① 863 计划(2009AA112205),国家自然科学基金(61103203)和中南大学博士后科学基金(2011QNZT039)资助项目。

^② 女,1980 年生,博士生,讲师;研究方向:无线传感器网络,移动自组织网络;E-mail: xykui@csu.edu.cn, kuixiaoyan@126.com

^③ 通讯作者,E-mail: zgzhang@csu.edu.cn

(收稿日期:2012-03-14)

是,所有节点均以概率成为簇首,可能会使得网络中产生的簇首数量较少,导致其余节点需通过长距离通信传送数据到簇首,增大了能量耗费。能量感知的分布式分簇算法 BPEC^[8]对 EADEEG 进行了改进,提出以节点本身的剩余能量与邻居节点的平均能量之间的比值作为节点竞争簇首的主要参数,以节点的“度”作为节点竞争簇首的辅助参数。分布式定向分簇算法 DDC^[9]针对 LEACH 中节点负载不相同的问题,提出了一种基于预测的分簇方法。分簇算法 CDAT^[10]考虑了如何在保证网络覆盖率的条件下延长网络生命周期,但其产生的簇首数量较多,增加了节点的能量耗费。分簇数据收集算法 EECT^[11]改进了 LEACH 算法,使能量越高而产生数据量越少的节点有越高的概率成为簇首,其余节点则加入离自己最近的簇首。根据我们掌握的情况,EECT 是目前唯一一个针对节点产生不同数据量进行分簇数据收集的算法。但是,EECT 主要面向单跳网络,且没有解决簇首分布不均匀以及节点能量消耗不均匀的问题。

(2) 基于支配集的分簇算法

能量约束支配集(ECDS)算法^[12]通过构造能量约束支配集来进行分簇,从而有效地保存节点能量。常数近似率的连通支配集算法 FT-CDS-CA^[13]针对网络中节点容易被损坏而死亡的情况,提出构造具有高度容错能力的支配集来进行分簇。但是,FT-CDS-CA 没有对数据收集过程中节点的能量消耗进行有效的均衡和优化。分布式最小路由代价的连通支配集算法 MOC-CDS^[14]为了减少数据收集过程中网络总的通信能耗,提出构建一个最小路由代价支配集来进行分簇。在这个支配集中,任意 2 个节点间都可以通过最短路径来传输数据,使得数据能以最少的跳数传输到目的地,有效地减少了数据传输的通信能耗。

面向单跳网络的协议假定网络中任意两个节点之间可以相互通信。而在实际应用中,节点的通信半径是有限的,而且网络部署的区域广大,因此节点很难做到能与网络中的任意一个节点通信。因此,面向多跳网络的协议逐渐受到研究者的重视。

此外,目前已有的工作普遍假设节点在每轮产生的数据量是一样的。这里,轮是指从所有传感器节点收集一次数据并传送到 Sink 节点的过程,不管它持续的时间是多少^[4]。但是,在实际应用中,由于节点类型有多种,不同的节点往往产生不一样的数据和传送大小不等的数据量。因此,目前的工作

还很难应用在实际环境中。

为了提高算法的实用性,针对实际应用中的特点,本文提出了基于支配集的分簇数据收集算法 DSCAU。该算法能使节点在通信半径有限、节点产生数据量不相同的情况下完成数据收集,并能有效地均衡节点的能量消耗以延长网络生命周期。这里,网络生命周期定义为网络中第一个节点死亡的时间。支配集是指网络中的一个节点集合,网络中每个节点如果不是这个集合中的节点,则肯定是这个集合中的节点的邻居。基于支配集进行分簇,就是让支配集中的节点担任簇首,而其余节点可以在一跳范围内与至少一个簇首进行通信,减少了网络中节点多跳传输数据的能耗。根据我们掌握的情况,DSCAU 是第一个面向多跳数据不均衡网络的分簇数据收集协议。

2 系统模型与分析

与文献[15]类似,本文假设网络部署在一个环境非常恶劣的地方(如原始森林或战场),在这个环境中无法固定地部署 Sink 节点且人类或车辆均无法进入。因此,数据只能由飞机(或无人机)来进行采集,即飞机作为一个移动 Sink 来收集数据。假设飞机飞行的高度在 H 以内,当飞机飞到节点上方时,节点将以半径 H 发送自己的数据给飞机(移动 Sink)。由于 H 的值一般较大(150m),因此节点发送数据给飞机将会耗费较多的能量。为了能尽量节省节点的能量,我们采用分簇的方案,即将网络中的节点划分为若干个簇,每个簇有一个簇首,簇首将接收簇内节点发送来的数据,与自身数据一起发送给飞机。这样,就能有效减少远距离发送数据的节点数量,使节点的能量得到有效保存。

与 LEACH 一样,本文提出的数据收集算法也是按轮运行。在每一轮,网络中的节点将被划分为若干个簇,每个簇内有一个簇首和若干簇内节点。簇内节点感知到数据后,将数据发送给自己的簇首。簇首将所有簇内节点发送来的数据连同自己产生的数据一起发送给移动 Sink。

2.1 网络模型

假设网络部署在一个面积为 $M \times M$ 平方米的正方形区域,在区域中随机地散布了 n 个节点,所有节点组成一个连通的网络。每个节点 $v_i (i = 1, 2, \dots, n)$ 有一个唯一的 ID 号。网络具有如下性质:

(1) 节点部署后不再移动;

- (2) 节点的初始能量是异构的;
- (3) 节点不知道自己的地理位置;
- (4) 每个节点产生的数据量可以不相同。

2.2 通信模型和数据汇聚模型

假设每个节点 v_i 在每轮产生的数据量为 c_i bit, 而且每个节点可以通过信息交换获知自己邻居的数据量。假设数据不具有相关性, 每个簇首接收邻居发来的多个数据包后, 连同自己的数据一起发送给 Sink 节点。为了提高算法的适用性, 节点应尽量简单, 从而降低制造成本, 容易被各种应用接纳。为了实现这一目标, 本文假设节点仅具有 2 级可调功率:

(1) 当节点 v_i 担任簇内节点时, 采用一个较小的半径 r_n 进行通信, 此时 v_i 的通信能耗 $E_i(c_i, r_n)$ 满足自由空间的能量衰减模型, 即:

$$E_i(c_i, r_n) = c_i E_{elec} + c_i \epsilon_{fs} r_n^2 \quad (1)$$

这里, E_{elec} 代表发射电路消耗的能量 (单位为 nJ/bit), ϵ_{fs} 表示自由空间模型中功率放大所需的能量 (单位为 pJ/bit/m²)。

(2) 当节点 v_i 担任簇首时, 将以 $r_c = H$ 进行通信。此时, 节点的通信能耗 $E_i(k_i, r_c)$ 满足多路径衰减的能量衰减模型, 即:

$$\begin{aligned} E_i(k_i, r_c) &= k_i E_{elec} + k_i \epsilon_{mp} r_c^4 \\ &= k_i (E_{elec} + \epsilon_{mp} r_c^4) \end{aligned} \quad (2)$$

这里, $k_i = \sum_{v_j \in A} c_j + c_i$ 表示簇首 v_i 发送的数据量, A 表示 v_i 所在簇的簇内节点集合, ϵ_{mp} 表示多路径衰减模型中功率放大所需的能量 (单位为 pJ/bit/m⁴)。

(3) 每个节点接收 c bit 数据需要耗费的能量为

$$E_r(c) = c E_{elec} \quad (3)$$

2.3 相关定义

为了更好地介绍算法思想, 在描述问题前, 对即将使用的一些基本概念作如下定义:

定义 1 节点 v_i 的邻居是指与 v_i 之间的距离小于 r_n 的那部分节点, v_i 的邻居集合用 $N(v_i)$ 表示, 而 $|N(v_i)|$ 表示 v_i 的邻居数量。

定义 2 节点 v_i 的未覆盖邻居是指 v_i 的邻居中没有加入簇的节点, 而 $uncover(v_i)$ 表示 v_i 未覆盖邻居的集合, $|uncover(v_i)|$ 表示 v_i 未覆盖邻居的数量。

定义 3 节点 v_i 的簇首容量 $head_capacity(v_i)$ 为节点成为簇首时能进行数据收集的最小轮数。由于簇首在一轮中将接收所有簇内节点发送来的数据, 连同自己的数据一起发送给 Sink 节点, 因此当所有未覆盖邻居均成为簇内节点并向簇首发送数据时, 簇首消耗的能量为

$$\sum_{v_j \in uncover(v_i)} c_j E_{elec} + \left(\sum_{v_j \in uncover(v_i)} c_j \right) (E_{elec} + \epsilon_{mp} r_c^4)$$

+ c_i) ($E_{elec} + \epsilon_{mp} r_c^4$)。假设 v_i 的能量为 $E(v_i)$, 则可推出节点 v_i 的簇首容量 $head_capacity(v_i)$ 为

$$head_capacity(v_i) = \left\lfloor \frac{E(v_i)}{\sum_{v_j \in uncover(v_i)} c_j E_{elec} + \left(\sum_{v_j \in uncover(v_i)} c_j + c_i \right) (E_{elec} + \epsilon_{mp} r_c^4)} \right\rfloor \quad (4)$$

定义 4 节点平均度 D_n 表示网络中所有节点的度的平均值, 即 $D_n = (\sum_{i=1}^n D_i) / n$, 其中 $D_i = |N(v_i)|$ 表示节点 v_i 的度。节点平均度主要用于衡量网络中节点的密度, 我们将在实验环节利用它来指导网络场景的设置。由于在网络场景设置之前, 无法确定网络规模, 也无法知道节点具体的度, 因此本文将采用节点均匀分布时的密度来近似估算节点平均度, 如: 在边长 $M = 1000\text{m}$ 的正方形区域有 $n = 480$ 个节点, 节点的通信半径 $r_n = 100\text{m}$, 则节点平均度为 $n\pi r_n^2 / M^2 \approx 15$ 。

2.4 问题描述

为了保证网络每个节点均能加入簇, 且算法具有较高的可扩展性和能量有效性, 基于支配集的分布式分簇数据收集算法应该具备如下特点:

(1) 算法的实现是分布式的, 无需采集完网络中全部节点的信息后再集中进行计算。

(2) 节点仅需知道自己邻居节点的信息, 无需获取邻居以外的其它节点的情况。

(3) 分簇的结果能使每个非簇首节点在通信半径内与至少一个簇首直接进行通信。

(4) 分簇过程耗费的能量较少, 节点可以保存能量用于数据收集。

(5) 节点在数据收集过程中应负载均衡, 从而延长网络生命周期。比如说, 能量高的节点承担的数据转发量大些, 反之则小些。

本文主要研究如何采用分布式的方法在网络中构造支配集, 由支配集中的节点担任簇首, 然后让其他节点选择向某个簇首传送数据。组成支配集的节点不仅应该具有更高的能量, 而且在网络中分布均匀以便其他节点能加入簇。算法的目标是实现网络中节点的能量消耗均衡, 从而有效延长网络生命周期。

3 DSCAU 的设计

3.1 簇生成算法

采用 DSCAU 时, 簇首的确定需要在若干个时间步内完成。在每个时间步, 每个没有加入簇的节

```

在每个时间步开始时:
1. if ( $v_i$  没有加入簇)
2.   if ( $|uncover(v_i)| == 0$ )  $v_i$  成为正式簇首;
3.   else {将簇首容量  $head\_capacity(v_i)$  封装为一个
      消息  $Capacity(head\_capacity(v_i), v_i)$  广播给自己的
      邻居;  $temp = false$ ;
4.   end
在接收到第一个  $Capacity$  消息时:
1. 等待并接收其他邻居发来的  $Capacity$  消息;
2. for (每个发来  $Capacity$  消息的邻居  $u$ )
3.   if ( $head\_capacity(u) > head\_capacity(v_i)$ )
4.      $temp = true$ ;
5.   end
6. end
7. if  $temp == true$  //  $v_i$  成为候选簇首;
8.   广播一个宣布自己成为候选簇首的消息
      $Candidate(v_i)$  给自己的邻居;
9.   end
在接收到第一个  $Candidate$  消息时:
1. 等待并接收自己邻居发来的  $Candidate$  消息并统计
   候选簇首的数量  $candidateNum_i$ ;
2. 将  $candidateNum$  封装为一个消息  $Num\_of\_candidate(candidateNum_i, v_i)$  中并广播给自己的邻居;
在接收到第一个  $Num\_of\_candidate$  消息时:
1. if ( $v_i$  是候选簇首)
2.   等待并接收其他邻居发来的  $Num\_of\_candidate$ 
   消息;
3.   找到集合  $\{candidateNum_j | j \in uncover(v_i)\}$  的中
   值  $m$ ;
4.   以概率  $p_i = 1/m$  成为正式簇首并发布一个
    $Clusterhead(v_i)$  消息给自己的邻居;
5.   end
在接收到第一个  $Clusterhead(v_i)$  消息时:
1. if ( $v_i$  不是正式簇首)
2.   等待并接收自己邻居发来的  $Clusterhead$  消息;
3.   将所有发来  $Clusterhead$  消息的正式簇首按  $head\_capacity$ 
   值以降序排序;
4.   for (每个正式簇首  $z$ )
5.     if (以概率  $p_i = head\_capacity(z) / \sum_{k \in B} head\_capacity(u)$ 
       选择加入  $z$  的簇) //  $B$  是所有发来  $Clusterhead$ 
       消息的正式簇首的集合
6.       广播一个  $Join(z, v_i)$  消息通知  $z$ ;
7.       break;
8.     end
9.   end
10.  end
在接收到一个  $Join(v_i, v_j)$  消息时:
1. if ( $v_i$  承担的簇成员数量  $N_i < K$ )
2.   接纳  $v_j$  为自己的簇成员并广播一个  $Confirm(v_i, v_j)$ 
   消息通知  $v_j$ ;
在接收到一个  $Confirm(v_j, v_i)$  消息时:
1.  $v_i$  标记自己已经加入簇, 不再进行广播;

```

图1 DSCAU 的详细描述

点将根据簇首容量来确定自己是否能成为候选簇首,也就是说,当节点的簇首容量不小于所有未覆盖邻居的簇首容量时,该节点成为候选簇首。对于任意一个候选簇首,将根据自己的邻居被其他候选簇首覆盖的情况以反比概率成为正式簇首。图1是算法的详细描述。

在图1中,算法在每个节点 v_i 上分布地运行。当簇首 v_i 接收到消息 $Join(v_i, v_j)$ 时,如果 v_i 承担的簇成员数量 $N_i < K$,则接纳 v_j 为自己的簇成员并广播一个 $Confirm(v_i, v_j)$ 消息通知 v_j ,即 v_i 承担的簇成员数量 N_i 最多为 K 。这里,约束簇首承担的簇成员数量小于等于 K 是因为算法面向的是数据不汇聚的应用,如果簇首承担的簇成员数量太多,会导致簇首通信时能耗过大而过早死亡。

3.2 性能分析

在本节,我们将对 DSCAU 的性能进行分析。

定理1 DSCAU 执行完毕后,网络中的簇首集合一定是一个支配集。

证明:通过反证法来进行证明。假设在 DSCAU 执行得到的分簇结果中,有若干个节点没有加入簇,即这些节点不是簇首,它们的邻居中也没有任何一个簇首。这部分节点组成一个未加入簇的节点集合 U 。根据 DSCAU 可知,对于任何一个没有加入簇的节点 $v_i \in U$,将执行以下2个步骤:

(1)如果 v_i 的邻居已经全部加入簇,即 v_i 没有未覆盖邻居,则 v_i 直接成为正式簇首。

(2)如果 v_i 的邻居中存在未覆盖邻居,则 v_i 与自己的未覆盖邻居竞争成为正式簇首。如果 v_i 没有成为正式簇首,而它有一个或若干个邻居成为正式簇首,则 v_i 加入该邻居的簇。另一方面,如果 v_i 成为簇首,则它的未覆盖邻居将加入 v_i 的簇。

因此,集合 U 中的节点最终将全部加入簇(成为簇首或簇内成员)。由此可见,DSCAU 执行完毕后有若干节点没有加入簇的假设不成立。由于 DSCAU 执行完毕后网络中的每个节点如果不是簇首,则肯定是某个簇的簇内节点,因此网络中的簇首集合一定是一个支配集。得证。

4 仿真实验和性能分析

本节将对提出的算法与已有工作进行对比,进而验证 DSCAU 算法的性能。假设节点的通信半径 r_n 为 100m,簇首的通信半径 $r_c = 2r_n$,飞机飞行的高度为 $H = 2r_n$ 。假设 n 个节点随机散布在 $10r_n \times 10r_n$ 。

的正方形区域内。假设节点的能量水平异构,每个节点的能量将在 $[1J, 2J]$ 的范围内随机分配。节点在进行通信时的能耗参数为: $E_{elec} = 50nJ/bit$, $\varepsilon_{fs} = 13pJ/bit/m^2$, $\varepsilon_{mp} = 0.0013pJ/bit/m^4$ 。假设在每一轮中节点 v_i 产生一个大小为 c_i bit 的数据包,而 c_i 是从 $[200, 300]$ 中随机选取的一个整数。选择分簇算

法 EECT 作为对比,通过改变节点的密度和网络的规模来考察算法在不同网络条件下的性能。

4.1 分簇的效果

在 $n = 480$ (节点平均度为 15)的网络上分别运行 DSCAU 和 EECT,观察网络中簇的分布情况。实验结果如图 2 所示。

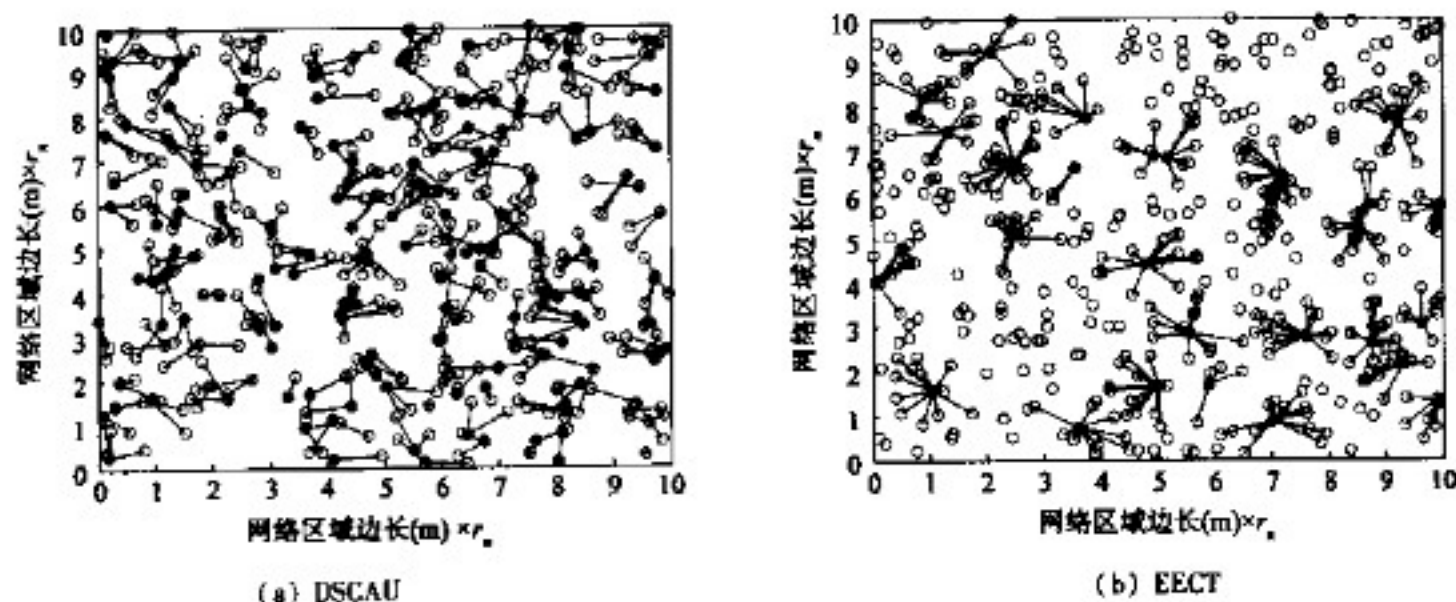


图 2 DSCAU 和 EECT 分簇对比

从图 2 可以看到,DSCAU 可以在节点通信半径有限的情况下,生成分布较为均匀的簇首,使网络中所有节点均能加入簇,此外簇的规模得到了有效的控制。而 EECT 是根据一定的概率生成簇首,这样不仅生成的簇首数量较少,而且簇首分布不均匀,导致很多节点无法加入簇。

4.2 网络生命周期对比

在本节,将分别在 2 种不同的网络条件下测试算法所获得的网络生命周期。假设节点发送的消息大小为 32bit。在通常情况下,节点在网络中密度不

会太大,这体现为节点的邻居数量(平均度)通常不大于 30。因此,我们设计了以下 2 种网络场景来测试算法的性能。在第 1 种网络场景中,设置网络的区域为 $10r_s \times 10r_s$,节点的总数分别为 $n = 320, 480, 640, 800, 960$,此时对应的节点平均度分别为 10, 15, 20, 25, 30。在第 2 种网络场景中,设置节点平均度固定为 15,网络区域分别设置为 $8r_s \times 8r_s, 10r_s \times 10r_s, 12r_s \times 12r_s, 15r_s \times 15r_s$,此时网络中节点总数分别为 $n = 306, 480, 688, 1075$ 。实验结果如图 3 所示。

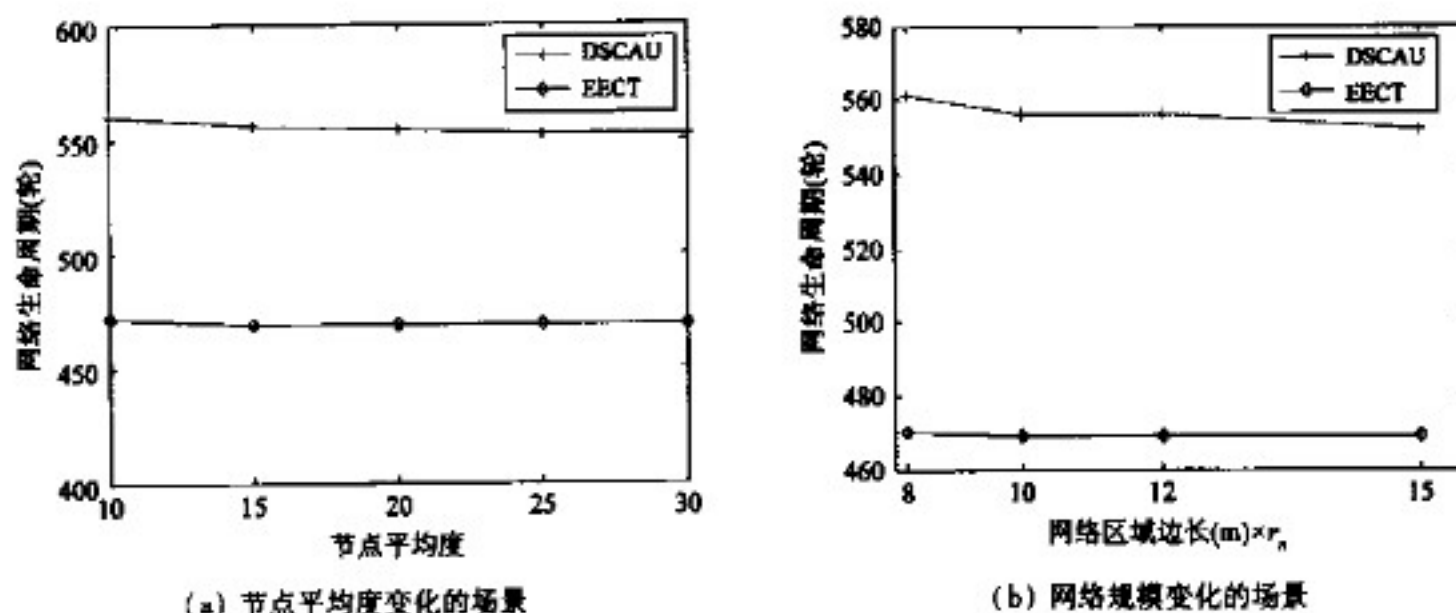


图 3 网络生命周期对比

从图3可以看到,在不同的网络条件下,DSCAU均获得了比EECT更高的网络生命周期。这是因为DSCAU很好地均衡了节点的能量消耗,让容量高的节点作为簇首进行长距离通信,而其余节点则选择加入某个簇进行短距离通信。

从图3(a)可以看到,随着网络密度逐渐加大,DSCAU的网络生命周期基本保持在同一水平,这是因为根据我们的测试,当 $K=2$ 时,算法的性能最好,此时无论网络密度多大,簇内节点数量最多是3(即1个簇首和2个簇成员),这样网络内的节点负载比较均衡且簇首的负担较轻。对于EECT,网络密度增大并没有让它出现明显变化,这是因为无论网络条件如何变化,EECT生成的簇首数量均较少,在多跳条件下造成较多节点无法加入簇。这些节点只能以长距离通信发送数据给Sink,导致这部分节点的能量耗费过大而过早死亡。因此,在不同网络条件下,EECT的网络生命周期均最低。

从图3(b)可以看到,如果网络中节点数量和网络区域大小均发生变化,而节点的平均度不变时,DSCAU所获得的网络生命周期变化并不大。EECT的网络生命周期在不同网络条件下均最低,这同样是因为EECT无法保证所有节点均能加入簇,大量无法加入簇的节点不得不自己发送数据给Sink,使得能量最少的节点容易快速死亡,大大降低EECT的网络生命周期。

5 结论

本文研究了如何利用支配集分簇来有效地进行数据收集。针对多跳网络中节点产生数据量不相同的条件下数据收集面临的难点问题提出了一个新的分布式分簇算法DSCAU。使用DSCAU时,每个节点首先将根据自身能量、邻居数量、自身和邻居产生的数据量等情况来选举候选簇首。如果一个节点发现自己成为簇首时能够比邻居节点承担更多轮次的数据收集,则宣布自己成为候选簇首。然后,每个候选簇首根据自己邻居被其他候选簇首覆盖的情况,以反比概率成为正式簇首,一方面避免正式簇首数量过多,一方面使得簇首在网络中分布均匀。DSCAU在网络中每个节点上分布地运行,不需要节点邻居以外的其它信息,具有较高的可扩展性。理论分析和仿真实验均表明,DSCAU能获得比目前已有算法更长的网络生命周期。

参考文献

- [1] Estrin D, Girod L, Pottie G, et al. Instrumenting the world with wireless sensor networks. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, USA, 2001. 2033-2036
- [2] 李建中,高宏. 无线传感器网络的研究进展. 计算机研究与发展, 2008, 45(1): 1-15
- [3] 郜帅,张宏科. 时延受限传感器网络移动 Sink 路径选择方法研究. 电子学报, 2011, 39(4): 742-747
- [4] Akyildiz L, Su W, Sankarasubramanian Y et al. A survey on sensor networks. *IEEE Communications*, 2002, 40(8): 102-114
- [5] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, USA, 2000. 3005-3014
- [6] Younis O, Fahmy S. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, 2004, 3(4): 366-379
- [7] 刘明,曹建农,陈贵海等. EADEEG: 能量感知的无线传感器网络数据收集协议. 软件学报, 2007, 18(5): 1092-1109.
- [8] 周新莲,吴敏,徐建波. BPEC: 无线传感器网络中一种能量感知的分布式分簇算法. 计算机研究与发展, 2009, 46(5): 723-730
- [9] 刘新华,方敏,旷海兰等. 一种基于负载均衡的无线传感器网络分布式定向分簇算法. 计算机研究与发展, 2009, 46(12): 2044-2052
- [10] 杨军,张德运,张云翼等. 基于分簇的无线传感器网络数据汇聚传送协议. 软件学报, 2010, 21(5): 1127-1137
- [11] Wei D, Navaratnam P, Gluhak A, et al. Energy-efficient clustering for wireless sensor networks with unbalanced traffic load. In: Proceedings of the IEEE Wireless Communications and Networking Conference, Sydney, Australia, 2010. 250-256
- [12] Albath J, Thakur M, Madria S. Energy constrained dominating set for clustering in wireless sensor networks. In: Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, Australia, 2010. 812-819
- [13] Kim D, Wang W, Li X Y, et al. A new constant factor approximation for computing 3-connected m-dominating sets in homogeneous wireless networks. In: Proceedings of the 29th IEEE Conference on Computer Communica-

- tions, San Diego, USA, 2010. 1-9
- [14] Ding L, Gao X F, Wu W L, et al. Distributed construction of connected dominating sets with minimum routing cost in wireless networks. In: Proceedings of 2010 IEEE 30th International Conference on Distributed Computing Systems, Genova, Italy, 2010. 448-457
- [15] Mhatre V, Rosenberg C, Kofman D, et al. A minimum cost heterogeneous sensor network with a lifetime constraint. *IEEE Transactions on Mobile Computing*, 2005, 4(1): 4-15

DSCAU: a dominating set based clustering algorithm for data gathering in wireless sensor networks with unbalanced traffic load

Kui Xiaoyan, Zhang Shigeng, Wang Jianxin

(School of Information Science and Engineering, Central South University, Changsha 410083)

Abstract

Based on the considerations that existing data clustering algorithms for wireless sensor networks (WSNs) mainly pay attention to single-hop networks while disregarding multi-hop networks, and they usually assume that every node in the networks produces the same amount of data, the data gathering for WSNs with unbalanced traffic load was studied, and a novel dominating set based clustering algorithm for data gathering in WSNs with unbalanced traffic load, called the DSCAU, was proposed. In the use of the DSCAU, when electing the tentative cluster head, each node takes its remainder energy, its traffic load, the number of its neighbors, and the traffic loads of its neighbors into consideration. The tentative cluster head will become the final cluster head with a probability inversely proportional to the numbers of other tentative cluster heads that cover its neighbors. Furthermore, the size of clusters is restricted to balance the energy consumption among different cluster heads. The theoretical analyses and simulation results show that the DSCAU can effectively prolong the network lifetime in multi-hop WSNs, meanwhile guaranteeing that all the nodes in the networks can join a cluster.

Key words: unbalanced traffic load, wireless sensor networks (WSNs), dominating set, clustering, data gathering