



传感器网络中具有负载平衡的移动协助数据收集模式^{*}

石高涛¹, 廖明宏²

¹(天津大学 计算机科学与技术学院, 天津 300072)

²(哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

Movement-Assisted Data Gathering Scheme with Load-Balancing for Sensor Networks

SHI Gao-Tao¹, LIAO Ming-Hong²

¹(School of Computer Science and Technology, Tianjin University, Tianjin 300072, China)

²(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

+ Corresponding author: Phn: +86-451-86413213, E-mail: kyle_shi@hit.edu.cn

Shi GT, Liao MH. Movement-Assisted data gathering scheme with load-balancing for sensor networks. *Journal of Software*, 2007, 18(9): 2235-2244. <http://www.jos.org.cn/1000-9825/18/2235.htm>

Abstract: Data gathering is the basic function of the sensor networks. However, the existing gathering schemes are almost based on the architecture with a static base station which results in the quick death of nodes around the base station. The reason is that the sensor nodes located near a base station have to relay data for a large part of the network and thus deplete their batteries very quickly. This paper discusses how to use the mobile base station for data gathering with load-balancing. A data gathering scheme MADG (movement-assisted data gathering), which makes use of the mobile base station for data collection, is presented. In this scheme, the base station moves in a stationary annularity area exploited for data buffering. The gathered data are firstly forwarded into the buffering area and then collected by the mobile base station. It is theoretically proved that the location $\sqrt{2}R/2$ away from the center is the optimal location for minimizing the energy consumption for transmitting data and that there exists a location which can make the maximal node load minimize. This paper then considers the optimum location jointing the energy consumption and load-balancing based on above analyses. Compared with the static base station scheme and the existing mobile base station scheme, MADG reduces the load by over 95% and 80%, respectively.

Key words: sensor network; mobile base station; data gathering; load-balancing; network lifetime

摘要: 数据收集是无线传感器网络的一个基本功能。然而, 现有的数据收集模式大都是基于静止基站的网络结构, 导致基站周围的节点由于担负着网络内的所有负载而快速死亡, 成为网络性能的瓶颈。研究如何利用移动基站收集数据来达到负载平衡, 提出了一个利用移动基站协助数据收集的模式 (movement-assisted data gathering, 简称 MADG), 它将基站移动区域设置为缓冲区, 首先将数据沿最短路径传输到缓冲区内, 然后在基站移动的过程中进行数据收集。证明了缓冲区位置设置在距离中心 $\sqrt{2}R/2$ 时数据传输总能耗最少, 并证明了存在一个缓冲区位置使得最大节点负载最小化, 进而确定了同时考虑到能源消耗和负载平衡的基站移动区域。理论分析和实验结果表明, 提出

* Supported by the Key Program of the National Natural Science Foundation of China under Grant No. 60533110 (国家自然科学基金重点项目); the Harbin Institute of Technology Foundation of China under Grant No. HIT 2002.74 (哈尔滨工业大学校基金)

Received 2006-03-10; Accepted 2006-07-26

的数据收集模式在很大程度上降低了网络节点的最大负载,并且减少了数据传输能源中的消耗,分别比固定基站和同类工作的最大网络负载降低 95%和 80%以上。

关键词: 传感器网络;移动基站;数据收集;负载均衡;网络生命周期

中图分类号: TP393 文献标识码: A

无线传感器网络的主要功能是对周边环境信息进行采集和管理,并将感知到的数据发送到基站进行进一步的处理^[1,2]。传感器节点上的资源非常有限,因此,如何提高能源的使用效率从而延长网络生命周期成为主要的研究问题。传统的数据收集模式都假设基站是静止不动的,这样,基站周围的节点将担负网络内的所有负载,成为网络性能的瓶颈,不但降低了网络的吞吐量,而且这些节点将快速死亡,从而减少了网络的寿命。

为了平衡基站周围的负载,有以下几种方法:第一,在基站周围部署更多的节点轮流工作,或者部署能力更强的节点。这种方法并不能从根本上解决负载均衡问题;第二,引入移动中继来分担基站周围节点的负载,如文献[3],这种方法可以有效延长网络的寿命,但负载不平衡问题依然存在;第三种,通过移动基站可以平衡这些节点的负载。由于在移动基站的过程中,负责向基站转发消息的节点不停地发生变化,因此可以直接进行负载均衡。Luo^[4]等人提出了移动基站来进行负载均衡,并且认为基站沿着网络周围移动是最佳的移动策略。

我们受文献[4]的启发,提出了一个移动协助的数据收集模式 MADG(movement-assisted data gathering)。该模式让基站有目的地移动,并在移动的过程中进行数据收集。与文献[4]不同,本文将基站移动区域设置为缓冲区,首先将数据沿最短路径传输到缓冲区内,然后在基站移动的过程中进行数据收集。我们给出了更加有效的基站移动策略,该策略同时考虑到数据收集过程中的能源消耗和负载均衡。此外,MADG 可以同时支持基于事件驱动的数据收集和查询驱动的数据收集。

本文第 1 节给出目前有关数据收集的研究情况,并对这些工作进行简单评述。第 2 节给出本文工作的网络模型和问题描述。MADG 模式的设计和性能分析将在第 3 节中详细讨论,论述缓冲区位置选择和数据路由的问题。第 4 节给出实验结果。最后是本文的结论。

1 相关工作

数据收集是传感器网络的主要任务,在早期的研究中主要侧重于基站固定的情况,所有感知节点在网络初始化时,记录下基站的位置以及到达基站的路径,每当探测到数据信息时,就通过多跳的方式向基站发送。比如,文献[5]建立以基站为根的路由树,当探测到满足查询的消息后,沿着预先维护的路径进行传输。LEACH^[6], HEED^[7]和 DEEG^[8]协议把网络分成若干簇,当感知到数据时首先发送到簇头,由簇头转发给基站。然而,这些方法在数据的传输过程中都要经过基站周围的节点,导致这些节点的死亡速度将比其他节点快得多。

避免这一问题的直观方法是移动基站,使负责向基站转发的节点经常变化,将网络负载分摊在不同的节点上。文献[4]详细地分析了这一问题,并提出了一个将数据路由和基站移动相结合的模式来达到负载均衡。由于基站不停地移动,这种方式只能支持基于查询的数据收集,同时,网络内的负载并没有降至最低。文献[9,10]也采用了移动基站的方法来最大化网络寿命,但不同的是,这些工作将基站移动问题模型化为线性规划问题,找到基站的最优移动计划和特定的停留点。

另一种依靠基站移动来收集数据的模式是节点通过单跳的方式直接与基站通信。在这种模式中,当基站移动到该节点附近时,节点才将探测到的数据传输给基站。Shah^[11]等人提出了数据骡子(data mules)的模式,探测到数据的节点首先将数据缓存起来,当有 mules 移动到该节点附近时,数据传输给 mules,然后由 mules 在移动到后方基站后卸掉数据。这种方式的好处是节点所拥有的负载较少,但是数据却无法保证及时传输到基站。与之类似的模式有文献[12,13]。

此外,Wang^[3]提出了一种移动中继的方法来收集数据的模式,其中基站是固定不动的,若干分布在基站周围(两跳之内)的中继节点有规律地沿圆环移动,所有两跳外的节点在探测到数据时首先传输给中继节点,然后由中继节点直接传输给基站。这种方法通过引入中继节点来分担基站周围节点的负载,从而解决了网络瓶颈问题。

但这种方法只能有限地延长网络寿命(4倍),并且两跳之外的节点负载平衡问题依然存在。

本文提出的模式同样采用移动协助的收集策略,与现有工作不同,本模式结合了现有两种工作方式(数据骡子和移动基站),先将数据传输到特定区域,然后基站在该区域内移动时再将数据收集起来。因此,这种方式同时支持基于事件驱动和查询驱动的数据收集。此外,本文同时考虑了负载平衡和数据传输过程中的能源消耗问题,找出基站最优的移动区域,给出一个负载更加平衡的路由策略。

2 网络模型及问题描述

本节给出 MADG 所基于的网络模型和基本的假设。与文献[3,4]类似,本文假设 N 个传感器网络节点以密度为 ρ 的泊松分布随机分布在一个半径为 R 的圆形区域内,其中圆心为 O 。这些节点具有强连通性质,也就是任意节点之间都存在至少一条路径。假设每个节点的半径为 r ,这里 $r \ll R$ 。所有节点以同样的频率 λ 向基站传输数据。此外,为了便于描述模式思想,假定每个节点都知道自己的位置和圆心 O 的位置以及区域半径 R 。我们将在模式描述部分给出当节点不具有圆心位置和半径信息时的解决方法。

假定基站位于该网络区域内,负责收集节点感知到的数据,其上的能源被认为是无限的。使用与文献[1]同样的符号,令 $load_n$ 表示节点 n 传输数据消耗的能源。 \overline{load}_n 表示节点 n 在一段时间内的平均能源消耗。每单位数据传输需要耗能表示为 ϵ 。在一段时间内,由于网络内节点将数据传输到基站平均总的能源消耗表示为 \overline{E}_{total} 。为了更好地节省能源,平衡负载,需要将路由策略与基站移动策略相结合。我们的目的是通过选择合适的基站移动策略和路由策略使每个节点上的最大负载最小化,同时,能够使数据向基站传输过程中的能源消耗最小化。因此,问题可以形式化为

$$\text{Minimize } load_N \equiv \max_{v \in N} \overline{load}_n(M, R) \quad (1)$$

$$\overline{E}_{total}(M, R) \quad (2)$$

Constraints: M constraints, R constraints.

这里, M 和 R 分别指基站的移动策略和节点的路由策略。在本文中,移动策略主要是指基站在什么位置移动。从对以上问题的分析中可以看出,本文所要解决的问题与文献[4]类似,但不同的是,这里的问题引入了另外一个因素,即数据传输能源消耗 \overline{E}_{total} 。本文不仅要考虑节点的负载,而且还要考虑数据在传输过程中总的能源消耗,从而更加节省能源。

3 移动协助的数据收集 MADG

3.1 基本思想

在基站移动的数据收集模式中,如果使探测到的数据直接传输(多跳的方式)给不停变动位置的基站,则需要经常在全网通知基站位置,从而引入很大的负载。MADG 也是依靠移动的基站来完成数据收集的,不同的是, MADG 不是将探测到数据的节点直接与基站进行通信,而是首先将数据传输到一个特定的区域内,我们称其为缓冲区。基站在该区域内来回移动,并且在移动的过程中将该区域内缓存的数据收集起来。

在 MADG 中,缓冲区是一个内径为 R_m ,外径为 R_n 的圆环,如图 1(a)中阴影区域所示,基站就在该区域内周期性地移动(移动速度可以根据应用需要调整)。当节点感知到数据后,它首先将数据转发到缓冲区,等待基站收集。由于在一个长时段内,网络内的任何位置都以相同的概率发生事件,因此,为了达到负载平衡的目的,一种启发式的路由策略是数据沿最短路径传输到缓冲区。当网络密度较高时,数据相当于沿着圆直径方向传输到缓冲区域。这样,直观上看,沿最短路径向缓冲区传输数据将所有负载充分地分摊在外围节点上。因此,这种路由策略的负载平衡能力要比直接与基站进行通信的性能更好。

总的来看, MADG 的特点在于:

- (1) 数据不直接与基站通信,而是先传输到一个特定的缓冲区内;
- (2) 基站在固定的区域内移动,并在移动中收集数据;

- (3) 模式综合考虑了负载平衡和数据传输能耗;
- (4) MADG 可以同时支持事件驱动的数据收集和查询驱动的数据收集;
- (5) 时间延迟具有上界,最大延迟为缓冲区周长/基站移动速度.

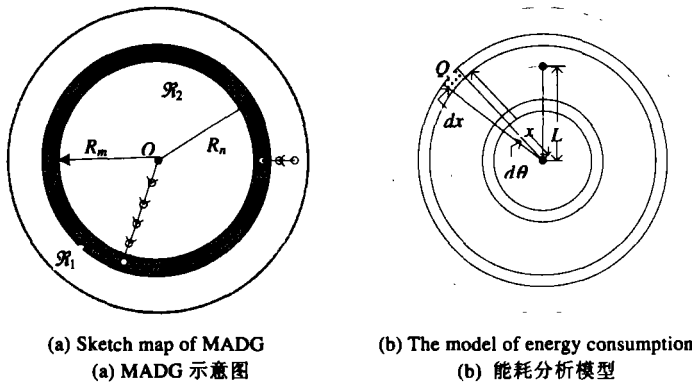


Fig.1 MADG and the model of energy consumption
图 1 MADG 及能耗分析模型

3.2 缓冲区域选择

本节将基于最短路径路由策略讨论如何选择缓冲区的位置能够使能耗和负载平衡达到最理想的情况,然后基于这些分析给出缓冲区的优化位置.

3.2.1 最小能耗缓冲区

由于数据需要首先传输到缓冲区中,因此,缓冲区位置的选择将影响数据传输的能源消耗.为了简化分析,我们把缓冲区看成一个与网络中心距离为 L 的圆,如图 1(b)所示.下述定理给出了 L 的最优值.

定理 1. 当缓冲区位于距圆心 $\sqrt{2R}/2$ 处时数据传输能耗最小.

证明:如图 1(b)所示,缓冲区位于距圆心 L 处,因此,把整个网络区域划分成 \mathcal{R}_1 和 \mathcal{R}_2 两部分,分别是与圆心距离大于 L 的区域和小于 L 的区域.由于 $r \gg R$,因此,缓冲区可以近似看成是一个很窄的区域.图 1(b)画出了该区域的中心线.对于 \mathcal{R}_1 ,在距圆心为 x 、宽为 dx 的圆环上,取夹角 $d\theta$ 的一小段扇环,如图中的阴影区域 Q ,这一区域内的节点个数为 $2x\rho dx \times \rho$.由于其中的节点到缓冲区的数据传输是沿着最短路径到达 Q ,也即沿着到圆心的方向传输数据,那么在高密度网络中,每个节点到达 Q 的跳数近似为 $(x-L)/r$.假设事件发生的频率为 λ ,那么阴影部分的节点传输到缓冲区的能耗为 $2x\rho dx \times (x-L)/r \times \lambda \epsilon$.为了计算总能耗,对其积分有

$$\int_L^R \int_0^{2\pi} (2x\rho\lambda\epsilon(x-L)/r) d\theta dx = 2\pi\rho\lambda\epsilon/r \times \int_L^R x(x-L) dx = \pi\rho\lambda\epsilon(2R^3 - 3R^2L + L^3)/3r \quad (3)$$

同样,对于位于 \mathcal{R}_2 内的节点,其总的能耗为

$$\int_0^L \int_0^{2\pi} (2x\rho\lambda\epsilon(L-x)/r) d\theta dx = 2\pi\rho\lambda\epsilon/r \times \int_0^L x(L-x) dx = \pi\rho\lambda\epsilon L^3/3r \quad (4)$$

因此,总的的数据传输能耗为 $\pi\rho\lambda\epsilon(2R^3 - 3R^2L + 2L^3)/3r$,要使总的传输能耗最少,即令 $f(L) = -3R^2L + 2L^3$ 取得最小值, $f(L)$ 在 $L = \sqrt{2R}/2$ 时取得最小值. \square

假定每单位数据传输的能耗不变,图 2 画出了总的能源消耗在不同的传输半径下随缓冲区位置变化的情况.其中, $R=10, \rho=1, \lambda=1, \epsilon=1$.

根据定理 1,如果将数据缓冲区设置为半径 $L = \sqrt{2R}/2$ 的圆周上,则能够将数据传输带来的能耗降为最低.选择 $\sqrt{2R}/2$ 作为缓冲区的位置只能使总能耗达到最小,而并不一定使每个节点上的负载达到最小.因此我们在第 3.2.2 节中分析网络负载并在第 3.2.3 节给出更加优化的缓冲区位置.

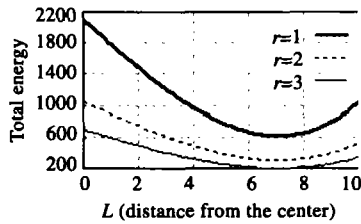


Fig.2 Total energy varies with the location of buffer

图 2 总能耗随缓冲区位置变化情况

3.2.2 最小化最大负载缓冲区

本节讨论如何选择缓冲区能够使网络中节点的最大负载最小化.定理 2 说明存在一个缓冲区,使得网络中的最大负载最小化.

定理 2. 存在一个缓冲区位置使得网络中的最大负载最小化,并且该位置与λ(数据频率)、ε(单位数据能耗)无关.

证明:首先分析网络区域内节点上的负载情况.对负载的分析需要考虑两种情况:第 1 种情况是距离圆心大于 L 的区域 \mathcal{R}_1 ,如图 3(a)所示;第 2 种情况是在 L 之内的区域 \mathcal{R}_2 ,如图 3(b)所示.对于位于 \mathcal{R}_1 内的节点,由于数据是沿着最短路径传到缓冲区中的,因此这种情况类似于数据传输到圆心.因此,我们直接使用文献[4]给出的负载分析模型,如图 3(a)所示.因此,对于任意一个距圆心距离为 d 的节点 n,经过它的平均负载为

$$\overline{load}_n^1 = \frac{(S_1 + S_2)\rho\lambda\epsilon}{S_2\rho} \approx \frac{\beta(R^2 - d^2)\lambda\epsilon}{\pi r^2} + \lambda\epsilon \tag{5}$$

其中, $\beta = 2\arcsin(r/d)$, S_1 是图中的灰色区域, S_2 是斜线区域.限于篇幅,更加详细的模型推导可参见文献[4].

当节点位于 \mathcal{R}_2 时,由于是通过最短路径传输,它的负载模型如图 3(b)所示,节点 n 上的平均负载为

$$\overline{load}_n^2 = \frac{(S_1 + S_2)\rho\lambda\epsilon}{S_2\rho} = \frac{\lambda\rho\epsilon\beta(d+r)^2/2}{(\beta(d+r)^2/2 - \beta d^2/2)\rho} = \frac{(d+r)^2\lambda\epsilon}{r^2 + 2dr} \tag{6}$$

由式(5)和式(6)可以看出, \overline{load}_n^1 是 d 的递减函数,而 \overline{load}_n^2 则是 d 的递增函数.并且 \overline{load}_n^1 和 \overline{load}_n^2 分别在 $d=R$ 和 $d=0$ 时取得最小值 $\epsilon\lambda$,因此, \overline{load}_n^1 和 \overline{load}_n^2 两条曲线必然相交.假设相交于 L' ,且 $0 < L' < R$.可以断定,当 $d < L'$ 时, $\overline{load}_n^1 > \overline{load}_n^2$;当 $d > L'$ 时, $\overline{load}_n^1 < \overline{load}_n^2$.因此可以说,当缓冲区位于 L' 时,网络中的最大负载最小化.同样,由式(5)和式(6)可知, L' 只与 r 和 R 有关. □

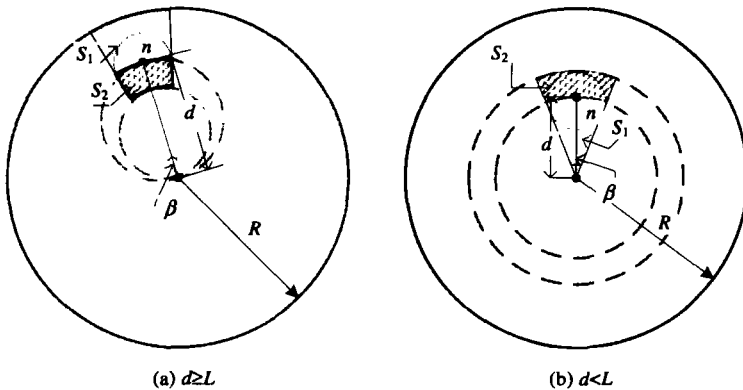


Fig.3 The model of load analysis

图 3 负载分析模型

根据定理 2 的分析,令 $\overline{load}_n^1 = \overline{load}_n^2$,计算得到 L' 值,即可找到使得整个网络中的最大负载最小化的缓冲区位置.由于节点能耗受节点负载的影响,因此,最小化最大负载能够提高网络寿命.此外,根据式(5)和式(6),节点上

的平均负载与网络密度无关.因此,这种模式具有一定的可扩展性.

3.2.3 能耗和负载结合

前两节分别给出了能耗最少和最大负载最小的缓冲区位置.本节讨论将总能源消耗和节点负载结合时的缓冲区位置.

图 4 是当 $L = \sqrt{2}R/2$ 时 MADG 的负载分布情况,其中 $R=10, r=1, \lambda=1, \epsilon=1$, 此时总能耗最少.在图 4(a)中,顶端的锯齿是由于两种路由方式的负载不平滑造成的.图 4(b)比较清楚地显示了这一点.从图中可以看出,缓冲区附近的节点的负载最高.更进一步地,所选择的 L 越小,网络的最大负载越大,若增大 L 值可以进一步降低缓冲区附近的负载.当两个负载曲线恰好相交时,网络中的最大负载降到最低.因此,在以上网络配置中,将 L 增大至 7.5 附近,能够使缓冲区附近的负载降低到最小.图 5 画出了这种情况下的网络负载分布.当 $R=10$ 时,可以将数据缓冲区设置在距离圆心为 7.1~7.5 附近,此时,数据传输能耗和节点负载分布同时达到比较理想的状态.

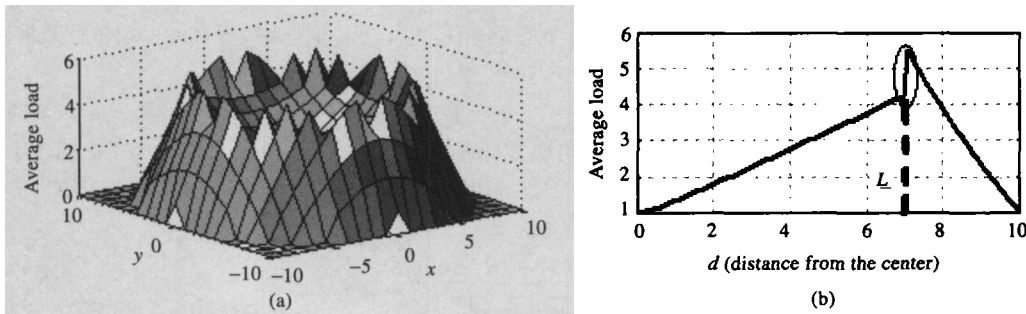


Fig.4 Load distribution when $L = \sqrt{2}R/2$

图 4 $L = \sqrt{2}R/2$ 时负载分布情况

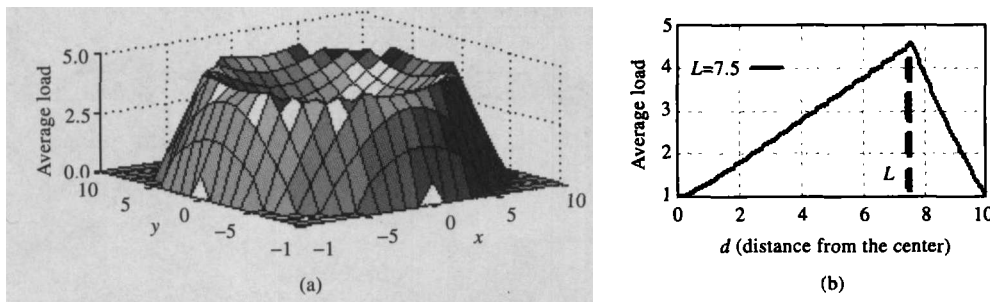


Fig.5 Load distribution of MADG when $L=7.5$ and $R=10$

图 5 MADG 的负载分布情况,其中 $L=7.5$

事实上,当 $L = \sqrt{2}R/2$ 时, \overline{load}_n^1 和 \overline{load}_n^2 并不相等.表 1 给出了随着网络半径 R 的变化 L' 和 $\sqrt{2}R/2$ 之间的关系比较.可以看出,这两个位置非常接近.这一巧合促使我们这样设置缓冲区:选择与圆心距离位于区间 $[\sqrt{2}R/2, L']$ 周围的区域作为缓冲区,并控制基站在此区域内移动.从表中还可以看出,在实际情况下(r 和 R 的比例更加实际),该缓冲区的宽度一般不超过节点的半径,比如当 $r=10, R=100$ 时.这样,缓冲区的数据在基站到达时可以直接发送给基站.

文献[4]给出了基站在网络内随机移动和基站围绕圆周运动两种数据收集模式,其最大负载分别是 28 和 23 左右,因此, MADG 比文献[4]中的最大负载分别降低 84% 和 80%.在相同的网络中,基站固定在圆心时节点最大负载在 100 左右.与之相比, MADG 将最大负载降低了 90%,从而将网络生命周期延长了 500% 以上.

Table 1 Comparisons of location between the min energy and min max load

表 1 最小能耗位置和最小最大负载位置对照

R	Min energy location	Min max load location			
		r=1	r=3	r=10	r=15
10	7.1	7.5	7.8	-	-
20	14.2	15.2	15.4	-	-
50	35.3	37	38	-	-
100	71	75	75.5	75.9	76.7
150	106	112	113.2	113.3	114

3.3 数据路由

在 MADG 中,由于缓冲区将网络覆盖区域分成了 \mathcal{R}_1 和 \mathcal{R}_2 两部分,因此,这两个区域内的节点在传输数据时方向恰好相反.一般情况下,每个节点至少知道自己的位置信息.然而,仅有自身位置信息是无法判定传输方向的.下面分 3 种情况来描述如何具体实现数据传输.为简单起见,假设节点位于极坐标系中,圆心为原点.

3.3.1 节点有圆心位置和圆半径信息

当节点同时有圆心位置和圆半径信息时,这种情况下的数据转发节点的选择最为简单.假设探测到数据的节点 S_d 的位置为 (r_d, θ_d) ,如果 S_d 位于区域 \mathcal{R}_1 中,那么该节点保存一个节点集 \mathcal{N}_1 ,这里,对 S_d 的任一邻居节点 S_i 当且仅当满足条件(7)和条件(8)时, $S_i \in \mathcal{N}_1$.

$$|\theta_i - \theta_d| \leq \Delta\theta \quad (7)$$

$$r_i < r_d \quad (8)$$

其中, $\Delta\theta$ 是一个常数.如果 S_d 位于区域 \mathcal{R}_2 中,那么该节点保存一个节点集 \mathcal{N}_2 ,这里,对 S_d 的任一邻居节点 S_i 当且仅当满足条件(7)和条件(9)时, $S_i \in \mathcal{N}_2$.

$$r_i > r_d \quad (9)$$

这样,每当节点 S_d 探测到数据时,如果 S_d 位于区域 \mathcal{R}_1 中,它就在邻居节点集 \mathcal{N}_1 中选择距离 (R_n, θ_d) 最近的节点 S_i 作为下一跳,然后将数据传输给 S_i .如果 S_d 位于区域 \mathcal{R}_2 中,它就在邻居节点集 \mathcal{N}_2 中选择距离 (R_m, θ_d) 最近的节点 S_i 作为下一跳,然后将数据传输给 S_i .就这样,数据依次传输下去,直到 S_i 位于缓冲区中,也就是说,当 $R_m \leq r_i \leq R_n$ 时,数据不再传输.同样,如果 S_d 本身就在缓冲区内,那么数据也不传输.

3.3.2 节点只有圆心位置信息

当节点只有圆心位置信息时,它无法判断自己是位于 \mathcal{R}_1 和 \mathcal{R}_2 中的哪个区域内,从而也就无法选择数据传输方向,为此需要采取一些辅助措施.由于基站的移动位置在网络部署时已经固定下来,因此,该基站所经过的区域就是数据缓冲区.为此,可以令基站在第 1 周的移动过程中通知它所经过的节点位于缓冲区中.当这些节点接到通知消息时,它们将这些消息转发到 \mathcal{R}_1 区域内,即转发给极径大于 R_n 的节点.这样,整个区域内的节点分成 3 种:直接接收到基站通知的节点、接到转发通知的节点和没有接到任何消息的节点.这 3 种节点分别位于缓冲区、 \mathcal{R}_1 和 \mathcal{R}_2 中,从而每个节点确定了自己所在的区域.有了这些信息后,节点就可以按照第 3.3.1 节中的路由选择方法进行转发数据.

3.3.3 节点只有自身位置信息

在这种情况下,每个节点不知道网络的半径和圆心在哪里,因此,以上两种方式都无法完成数据转发.与第 3.3.2 节类似,这里同样使用基站来完成路由信息创建.基站在第 1 周移动的过程中向其所经过的节点发出消息,该消息中包含一个跳数信息.这些节点接到消息后将跳数初始化为 0,然后广播给所有邻居节点,这些邻居节点记录下消息来源,并将消息经历的跳数加 1 后继续广播下去.如果节点收到的消息跳数比历史消息跳数要小,那么该节点更新消息来源,然后将跳数加 1 后继续广播,否则直接将消息丢弃.当这一过程结束后,所有节点都记录着到达缓冲区节点的最短路径的上一跳节点,从而在探测到数据时直接沿建立的路由转发数据.

4 实验评价

本节给出我们所提出的数据收集模式的性能评价实验结果,同时与其他收集模式进行比较.由于主要关注

上层收集模式对节点负载的影响,因此,以下实验忽略了物理层和 MAC.如果不作特殊说明,以下实验都是在一个半径 $R=100$ 单位的圆内随机分布 3 000 个传输半径为 10 的节点上进行的.其中, $R_m=71,R_n=75$.数据传输率和单位能源消耗归一化为 1.

实验主要比较 3 种移动协助数据收集模式的负载分布和能源消耗情况,这 3 种数据收集模式分别是随机移动 RM(random movement)、绕圆周移动 PM(peripheral movement)、静止基站(static-base model)和第 3.2 节提出的 MADG 移动模式.其中,前两种模式由文献[4]中提出,RM 策略假设基站以相同的概率出现在网络内的任意位置,PM 策略使基站沿着网络区域的外围移动,并且感知到的数据都直接向基站传输.SM 模式将基站固定在圆心位置.实验假设每个节点都具有圆心位置和圆半径信息.由于直接计算每个节点上的平均负载比较麻烦,为此,以下实验数据都是 10 个不同的拓扑中运行(默认每次运行 3 600 个时间单位)后的实际负载的平均值.

4.1 数据传输能耗

本节考察数据缓冲区的位置对数据传输总能耗的影响,实验将 L 从 10 变化到 100,每次递增 5,统计每一步的能源消耗.图 6(a)给出了 MADG 的总传输能耗随着缓冲区位置的变化情况.可以清楚地看到,实验显示的结果与第 3.2.1 节中总能耗的分析相吻合,最小总能耗的缓冲区位置在 70~75 之间.

实验也考察了 3 种不同策略的总传输能耗,实验在 10 个不同的拓扑中分别运行 3 种算法,统计结果如图 6(b)所示.可以看出,MADG 的总能耗总是比其他两种策略要小.此外,实验中随机移动比沿外围移动总能耗要小,这是因为当基站沿外围移动时,数据传输平均距离比基站随机移动时的平均距离要小.可见,尽管 PM 和 RM 模式与固定基站模式相比能够更好地负载平衡,但是其代价是数据走了较长的路径,消耗了更多的能源.

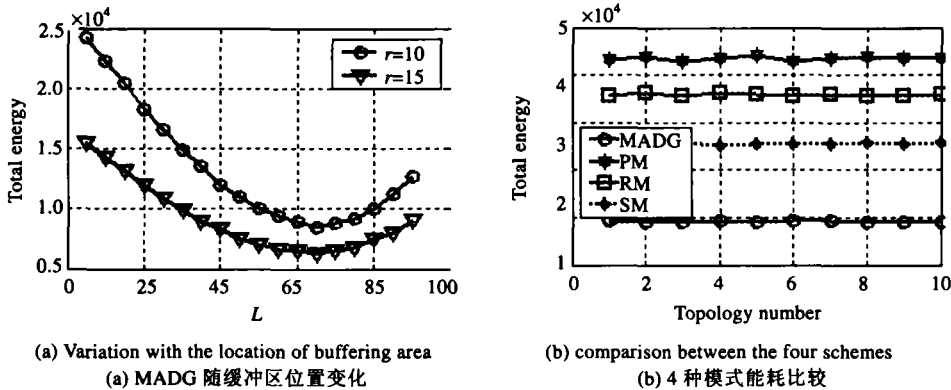


Fig.6 Comparison of the total energy consumption

图 6 总能源消耗比较

4.2 负载分布

本节验证网络负载的分布情况.实验将网络区域划分成 10 个围绕圆心等宽度的圆环,计算落入每个圆环内节点的平均负载.图 7 给出了在不同的移动策略下节点的负载情况.从图中可以看出,随机移动的负载平衡情况最差,沿外围移动的负载情况次之,而本文所提出的 MADG 中的移动策略在第 3.1 节的路由支持下使得最大节点负载降到最低,平衡情况最优.此外,从图中可以看出,随机移动和沿外围移动策略的负载随着距圆心距离的增大而减小,这与文献[4]中的实验分析相吻合,而 MADG 的负载分布趋势也与本文的分析相吻合,在 70~80 之间的区域承担的负载最大.由于在仿真实验中网络拓扑不规则,并且基站移动也不是连续的,因此,图 7 并不像理论分析的那样连续.

4.3 生命周期比较

为了更加清楚地比较不同模式的节能效果,本节比较了不同收集模式下的网络寿命.网络寿命定义为第 1 个节点死亡的时刻^[4].实验将每个节点的初始能量设置为 75,在 10 个不同的网络拓扑中运行算法.图 8 给出了实

验结果.可以看出,静止基站模式的第1个节点死亡时间最短,随机移动基站模式次之,MADG则极大地延长了网络的生命周期.此外,我们在实验中发现,当节点的初始能源设置较小时,PM模式的第1个节点死亡时刻比RM模式要小,这与PM数据传输能源消耗过大有关.

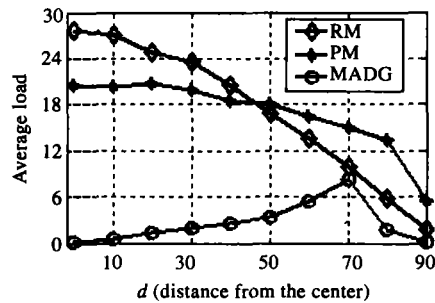


Fig.7 Comparison of load under different move strategies

图7 不同移动策略下的负载比较

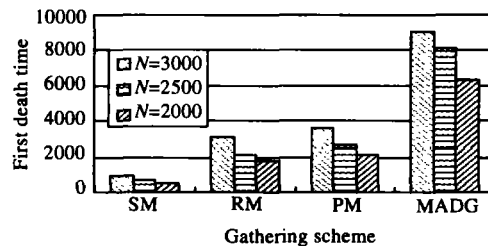


Fig.8 Comparison of the first death time under different gathering schemes

图8 不同模式下第1个节点死亡时刻

5 结论与进一步的工作

本文研究了利用移动基站收集数据来解决静止基站的网络瓶颈问题.综合考虑了数据传输能耗和节点负载平衡,提出了一个数据收集模式 MADG,将数据先传输到一个特定的区域,并使基站在该区域内移动收集数据.给出了数据传输总能耗最少和最小化最大负载的缓冲区位置.理论证明了当缓冲区位于距离圆心 $\sqrt{2}R/2$ 时能使数据传输总能耗达到最小,并且存在一个缓冲区位置使网络内节点上最大负载最小化.最后分析了同时有利于能耗和负载平衡的缓冲区位置.此外,给出了一个启发式的路由策略及其详细的实现过程.理论分析和实验结果都表明,与现有工作相比,MADG在能耗和负载平衡方面都达到最优,分别比固定基站和同类工作的最大网络负载降低95%和80%以上,生命周期延长了500%以上.

本文研究了当网络呈圆形分布时的网络负载平衡和能源消耗问题.实际的网络可能会呈不规则分布,对于这种情况,一种替代方案是找到该网络区域的内接圆,然后再根据文中的定理设置缓冲区.此时,性能并不能达到最优.对更好方法的探索将是我们今后的研究工作之一.

致谢 感谢瑞士 EPFL 计算机与通信科学学院 Jun Luo 博士在本文实验方面提供的建议.

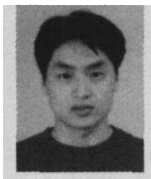
References:

- [1] Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. Wireless sensor networks: A survey. *Computer Networks*, 2002,38(4): 393-422.
- [2] Li JZ, Li JB, Shi SF. Concepts, issues and advance of sensor networks and data management of sensor networks. *Journal of Software*, 2003,14(10):1717-1727 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1717.htm>

- [3] Wang W, Srinivasan V, Chua KC. Using mobile relays to prolong the lifetime of wireless sensor networks. In: Porta TL, Lindemann C, eds. Proc. of the 11th Annual Int'l Conf. on Mobile Computing and Networking. New York: ACM Press, 2005. 270–283.
- [4] Luo J, Hubaux JP. Joint mobility and routing for lifetime elongation in wireless sensor networks. In: Znati T, ed. Proc. of the 24th IEEE INFOCOM. Washington: IEEE Computer Society, 2005. 1735–1746.
- [5] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In: Pickholtz R, ed. Proc. of the ACM/IEEE Int'l Conf. on Mobile Computing and Networks. Boston: ACM Press, 2000. 56–67.
- [6] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless microsensor networks. In: Sprague RH, ed. Proc. of the 33rd Annual Hawaii Int'l Conf. on System Sciences. Washington: IEEE Computer Society, 2000. 3005–3014.
- [7] Younis O, Fahmy S. Heed: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks. IEEE Trans. on Mobile Computing, 2004,3(4):660–669.
- [8] Liu M, Gong HG, Mao YC, Chen LJ, Xie L. A distributed energy-efficient data gathering and aggregation protocol for wireless sensor networks. Journal of Software, 2005,16(12):2106–2116 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/16/2106.htm>
- [9] Gandham S, Dawande M, Prakash R, Venkatesan S. Energy-Efficient schemes for wireless sensor networks with multiple mobile base stations. In: Kero TEF, ed. Proc. of the IEEE GLOBECOM. Washington: IEEE Computer Society, 2003. 377–381.
- [10] Wang ZM, Basagni S, Melachrinoudis E, Petrioli C. Exploiting sink mobility for maximizing sensor networks lifetime. In: Sprague RH, ed. Proc. of the 38th Hawaii Int'l Conf. on System Sciences. Washington: IEEE Computer Society, 2005. 287–295.
- [11] Shah R, Roy S, Jain S, Brunette W. Data mules: Modeling a three-tier architecture for sparse sensor networks. In: Ekici E, ed. Proc. of the IEEE Workshop on Sensor Network Protocols and Applications. Piscataway: IEEE Computer Society, 2003. 30–41.
- [12] Kansal A, Somasundara A, Jea DD, Srivastava MB, Estrin D. Intelligent fluid infrastructure for embedded networks. In: Banavar GS, Zwaenepoel W, eds. Proc. of the 2nd Int'l Conf. on Mobile Systems, Applications and Services. New York: ACM Press, 2004. 111–124.
- [13] Zhao W, Ammar M, Zegura E. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In: Murai J, ed. Proc. of the 5th ACM MobiHoc. New York: ACM Press, 2004. 187–198.

附中文参考文献:

- [2] 李建中, 李金宝, 石胜飞. 传感器网络及其数据管理的概念、问题与进展. 软件学报, 2003, 14(10):1717–1727. <http://www.jos.org.cn/1000-9825/14/1717.htm>
- [8] 刘明, 龚海刚, 毛鸞池, 陈力军, 谢立. 高效节能的传感器网络数据收集和聚合协议. 软件学报, 2005, 16(12):2106–2116. <http://www.jos.org.cn/1000-9825/16/2106.htm>



石高涛(1978—),男,河南鹿邑人,博士,讲师,主要研究领域为无线传感器网络,网络嵌入式系统.



廖明宏(1966—),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为新型操作系统,实时与嵌入式计算,对等计算.