

传感器网络中基于数据压缩的汇聚算法*

谢志军¹, 王雷²⁺, 林亚平^{2,3}, 陈红¹, 刘永和³

¹(中国人民大学 信息学院,北京 100872)

²(湖南大学 软件学院,湖南 长沙 410082)

³(Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, USA)

An Algorithm of Data Aggregation Based on Data Compression for Sensor Networks

XIE Zhi-Jun¹, WANG Lei²⁺, LIN Ya-Ping^{2,3}, CHEN Hong¹, LIU Yong-He³

¹(College of Information, Renmin University of China, Beijing 100872, China)

²(College of Software, Hu'nan University, Changsha 410082, China)

³(Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, USA)

+ Corresponding author: Phn: +86-731-8821932, E-mail: wanglei_hn@hn165.com

Xie ZJ, Wang L, Lin YP, Chen H, Liu YH. An algorithm of data aggregation based on data compression for sensor networks. *Journal of Software*, 2006,17(4):860-867. <http://www.jos.org.cn/1000-9825/17/860.htm>

Abstract: Considering the characteristics and location information of nodes in sensor networks, a modified directed transfer model of sensor networks and a new distributed data aggregation model based on “area” are proposed. On the basis of these new models, a novel mixed entropy data compression algorithm based on interval wavelet transforming is proposed for sensor network, according to the characteristics of data in sensor networks and the good performances of wavelet transforming in compression of the data stream. Theoretical analyses and simulation results show that, the above new methods can compress the data stream and reduce the energy costs of nodes in data transferring efficiently for sensor networks. So, it can prolong the lifetime of the whole networks to a greater degree when the above new methods are deployed with those traditional DC (data centric) routing algorithms such as DD (directed diffusion) protocol for sensor networks.

Key words: sensor network; location information; data aggregation; interval wavelet transforming; entropy

摘要: 结合传感器网络的节点特性和位置信息,提出了一种基于连通支配集的传感器网络定向传播模型,以及一种基于“域”的分布式数据汇聚模型 DDAM(distributed data aggregation model).DDAM 把传感器网络按“域”划分来构建连通核,传感节点只需在连通核中寻径,因而可明显减少寻径时间复杂度并且具有更好的分布性;然后在该定向传播与数据汇聚模型基础上,考虑传感器网络的数据特性及小波变换在流数据压缩方面的良好性能,提出了一种基于区间小波变换的混合熵数据压缩方法.理论分析和实验仿真结果表明:对比传统的 DC 算法-DD 路由算法相结合的算法,新算法能对传感器网络中的流数据进行有效压缩,可更大程度地降低传感器节点

* Supported by the National Natural Science Foundation of China under Grant No.60273017 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2002AA4Z3420 (国家高技术研究发展计划(863)); the Key Project of Chinese Ministry of Education under Grant No.02036 (国家教育部科学技术研究重点基金项目)

Received 2004-12-29; Accepted 2005-09-06

数据传输的能耗,从而进一步延长整个网络的生命周期。

关键词: 传感器网络;位置信息;数据汇聚;区间小波变换;熵

中图法分类号: TP393 文献标识码: A

集信息传感、数据处理、GPS(global positioning system)定位及网络通信功能于一体的传感器在环境与军事监控、地震与气候预测、地下、深水以及外层空间探索等许多方面都具有广泛的应用前景^[1-5],而外界环境的不确定性经常导致需要布置成百上千这样的传感器协同工作.因此,对由大规模传感器(传感节点)构成的传感器网络(sensor network)的研究正引起研究人员的广泛关注,被认为是 21 世纪的一项挑战性研究课题^[6].

作为一种新型无线网络,与传统无线网络相比,传感器网络具有以下特性:(1) 节点分布稠密,数量众多;(2) 节点一般无人照管,能量有限且补充困难;(3) 节点的计算能力和存储空间有限.由此可知,在传感器网络中,节点无法维护任何全局信息.因此,传统无线网络中的路由协议不能直接用于传感器网络中的路由,有必要研究适合于传感器网络特性的新路由协议.目前提出的传感器网络路由协议主要可分为基于地址的路由协议(address-centric,简称 AC)和基于数据的路由协议(data centric protocol,简称 DC)两种.典型的基于 AC 的路由协议有 SAR(sequential assignment routing)^[7]等.基于 AC 的路由协议都没有考虑数据汇聚问题,任意源节点各自将数据独立发送到 Sink.而研究表明:传感器网络中,通信开销要远大于数据计算开销^[8].因此,人们在考虑数据汇聚的基础上提出了 DC 算法.代表性的基于 DC 的路由协议有:DD(directed diffusion)^[9],GITDC(greedy incremental tree data centric protocol)^[10],LEACH(low energy adaptive clustering hierarchy)^[11],GEAR(geographic and energy aware routing)^[12],DDCH(distributed data centric hierarchical routing algorithm)^[13]等.采用 DD 模型的 DC 算法显然可以节省路径的 hop 数目,因此而节约节点的传输能耗.上述的 DC 算法虽然比 AC 算法更节能,更适合传感器网络能量不足的特点,但由于只考虑了利用数据汇聚来降低能耗而没有考虑数据压缩,因此,其节能性无法得到更进一步的提高.与传统无线网络相比,在传感器网络中基于节能的考虑,传感器节点在空闲时经常是关闭的,只有当接到任务指令或监测对象出现时传感器节点才开始工作并产生传感数据.因此,其数据具有如下特点:阵发性、数据到达持续性、数据量的大小有不可预知性、其数据量随任务的变化而变化等.由此可知,传感器网中的数据具有流数据的特点,因而可以用流数据的处理方法来管理传感器网络中的数据.

由于小波变换对流数据具有良好的压缩性能,而传感数据具有流数据的特点,因此在数据汇聚的基础上,再利用小波变换来对传感数据进行压缩,显然可以进一步提高 DC 算法的节能性.因此,本文综合数据汇聚与数据压缩的思想,首先基于传感器节点的位置信息及连通支配集^[14]的思想,提出了一种基于连通支配集的传感器网络定向传播模型,以及一种基于“域”概念的分布式数据汇聚模型 DDAM(distributed data aggregation model).对比以往的数据汇聚模型,DDAM 首先提出一种“域”的概念,并且按域对传感器网络来划分成连通核,因而具有更好的分布性.同时,传感节点只需要在连通核中进行路由,所以能明显降低寻径的开销;然后在数据汇聚的基础上,结合传感数据的流数据特点,提出了一种基于区间小波变换的混合熵数据压缩算法,更加有效地压缩网络中的数据量,以实现进一步提高 DC 算法节能性的目的.理论分析和仿真实验结果表明:与传统的 DC 路由算法相比,新方法在有效提高数据信息熵的同时,可以更大程度地降低传感器节点的数据传输能耗.

1 传感器网络中基于“域”的数据汇聚模型

1.1 基于连通支配集的定向传播模型

定义 1(简单连通无向图). 设图 $G=(V,E)$, G 称为简单连通无向图,当且仅当图 G 满足以下两个条件: (1) G 为无自圈的、连通的无向图;(2) G 中任意两个节点之间最多有一条边.

定义 2(支配集). 图 G 的节点集 $S \subseteq V$ 为支配集,当且仅当节点集 S 满足以下条件: $\forall p \in V \Rightarrow p \in S$ 或 p 为 S 中的某个节点 q 的邻节点. S 中的节点称为支配点,图 G 的节点中不属于 S 中的节点称为从节点.

定义 3(连通支配集). 给定一个图 $G=(V,E)$,图 G 的节点集 $S \subseteq V$ 为满足如下条件的节点集合:由 S 导出的子图是连通图,且 S 是图 G 的一个支配集,则称 S 为连通支配集.若 S 为满足上述条件的最小节点集合,则称 S 为最

小连通支配集.

假定在传感器网络中各节点具有相同的有效通信距离.称两个节点是相邻的,即存在一条通信链路,当且仅当这两个节点在彼此有效的通信距离之内.假定相邻节点之间的链路是对称的,则传感器网络的拓扑结构可以看作是一个简单连通无向图 $G=(V,E)$,其中, V 为所有节点构成的顶点集合, E 为所有链路构成的边集合.在传感器网络中,各传感器节点利用传感部件采集被监测对象的原始数据,经过处理器部件处理后,通过无线网络传输到一个核心节点(core node);核心节点对数据进行汇聚,然后再对汇聚数据进行数据压缩,最后将汇聚压缩后的数据送到中心节点(sink node);Sink 对数据再次汇聚压缩,然后经 Internet 或卫星传输到用户数据处理中心.

在定向传播模型中,传感器网络中的节点不采用地址作为标识 ID,而是以节点可以提供的数据作为寻址依据.即 Sink 在网络中广播以某种数据格式构成的消息,询问它所感兴趣的监测数据,这种消息简称为兴趣.与这种兴趣匹配的节点(称为源节点)响应这种查询(称为事件),并回送监测数据给 Sink.

在传统的定向传播模型中,由于在 Sink 节点到传感器节点的兴趣发布过程中就建立了用于数据回传的梯度,因此不需要源节点对 Sink 再次寻径.这样虽然降低了部分寻径开销,但同时也降低了数据汇聚的性能,增加了数据传输的开销.研究表明:在传感器网络中,数据传输的开销为主要开销.为此,为了在寻径开销和数据传输开销中找到一个更好的平衡点,以进一步降低网络的整体开销,延长网络的生命周期,我们对上述定向传播模型进行了改进.首先利用分布式的方法形成整个网络的一个连通支配集,如图 1 所示,其中连通支配集由核心和网关节点构成;然后,源节点在回送监测数据给 Sink 时,仅在连通支配集中进行寻径.显然,这样可以有效提高源节点的路由效率.

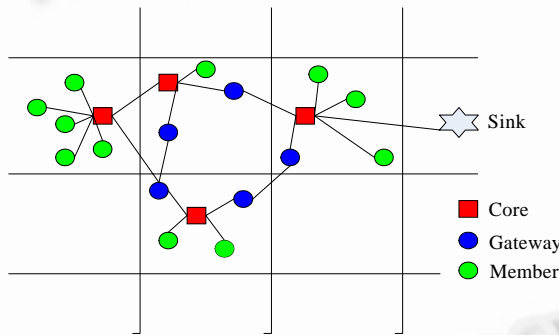


Fig.1 Connected dominant set based directed transfer model
图 1 基于连通支配集的定向传播模型

1.2 基于“域”的数据汇聚模型

定义 4(核). 图 G 的节点集 $C \subseteq V$ 为核,当且仅当节点集 C 满足以下条件: $\forall p \in V \Rightarrow p \in C$ 或 p 为 C 中的某个节点 q 的邻节点.

定义 5(连通核). 给定一个图 $G=(V,E)$,若图 G 的节点集 $C \subseteq V$ 为满足如下条件的节点集合:由 C 导出的子图是连通图,且 C 是图 G 的一个核,则称 C 为连通核.

定义 6(域). 任意传感节点 p ,假定其对 Sink 节点 R 的相对坐标为 (x,y) ,Sink 节点 R 及各传感节点具有相同的有效通信半径 r ,则称 p 属于域 (m,n) ,当且仅当如下公式成立: $m = \lceil x \lfloor \frac{r}{\sqrt{2}} \rfloor \rceil, n = \lceil y \lfloor \frac{r}{\sqrt{2}} \rfloor \rceil$. 其中,“ \lceil ”为除法运算符,“ \lfloor ”为取大于等于的整数运算符.

例:若 $r=15$, A 的地理坐标为 $(40,40)$,则 $A_x=4, A_y=4$,故 A 属于域 $(4,4)$;假定 B 的地理坐标为 $(35,37)$,则 $B_x=4, B_y=4$,故 B 同样属于域 $(4,4)$.

由定义 6 易知下述定理 1、定理 2 成立:

定理 1. 假定 Sink 节点 R 及各传感节点具有相同的有效通信半径 r ,则任意传感器网络存在满足定义 6 公式的唯一域划分.

定理 2. 同一域中的各传感节点互为邻节点.

证明:由定义 5 易知,每个“域”为一个长、宽均为 $\frac{r}{\sqrt{2}}$ 的正方形区域,该区域内任意两个节点之间的距离最

大为 $\sqrt{\left(\frac{r}{\sqrt{2}}\right)^2 + \left(\frac{r}{\sqrt{2}}\right)^2} = r$,即任意两个节点均在彼此的有效通信半径之内.因此,同一“域”中的各传感节点互为邻节点.

对任意传感节点 p ,假定其地理坐标为 (x,y) .假定 Sink 节点 R 及各传感节点具有相同的有效通信半径 r (communication radius),则任意传感节点 p 可以依据定义 6 计算出自己所属的“域”.同时,根据定理 1 和定理 2 的域划分规则,我们可以给出以下分布式数据汇聚模型 DDAM(我们规定 DDAM 操作于定义 1 给定的简单连通无向图 G):

分布式数据汇聚模型自动生成算法 DDAM.

- 1) 每个节点 p 利用 GPS 计算自己的地理坐标、剩余能量,并利用定义 6 计算出自己所属的域;
- 2) 每个节点 p 周期性地和其所有邻节点交换如下信息:(i) p 的当前状态 S_p (分为核心、网关或成员 3 种状态);(ii) p 的剩余能量 E_p ; (iii) p 所属的域 G_p ; (iv) p 的地理坐标.通过此操作,每个节点可以获知自己邻节点的状态、剩余能量、所属域及到邻节点的直线距离等信息;
- 3) 初始时,图 G 中的 Sink 节点的状态为核心,而所有传感器节点的状态均为成员.在每个周期内,每个节点 p 根据自己邻节点的状态 S_p 、能量 E_p 及所属域 G_p 的信息,按如下规则来计算自身新的状态:

p 所属域 G_p 内无核心,则 G_p 内的各节点依据剩余能量 E_p 的大小随机选举一个具有最大剩余能量的节点作为本域中的核心;

否则,若 p 不是核心或网关,且 p 与其他域中的节点相邻,则令 p 为网关;

- 4) 任意某个节点产生传感数据后,将数据直接传送给自己所属域中的核心.

由上述算法描述可知:DDAM 是完全分布式的,每个节点只需要知道其邻节点的有关信息.依据上述算法描述,易证以下定理 3 成立:

定理 3. 若图 $G=(V,E)$ 是简单连通无向图,则由算法 DDAM 得到的节点集 $\psi=\{p|p \text{ 为核心或网关且 } p \in V\}$ 是图 G 的一个连通核.

证明:首先,由定理 2 以及算法 DDAM 步骤 3) 中的可知:图 G 中的每个节点要么为核心,要么它的邻节点中至少存在一个核心,即与某个核心相邻.因此,集合 ψ 是图 G 的一个核.下面,用归纳法证明 ψ 是连通的.

设 p, q 为 ψ 中的任意两个核心节点,即 $p, q \in \psi$.由于假定传感节点具有相同的有效通信半径,故为表述方便,我们不妨设均为 1 个单位长度,因此 p, q 之间的距离即为 p, q 之间的最短路径的长度,记为 $d(p, q)$.由于图 G 是连通的,因此 $d(p, q) < R$ (R 为实数),也即 $d(p, q)$ 为有限整数.

- 1) (i) 若 $d(p, q)=1 \Rightarrow p, q$ 相邻,故它们直接可达 $\Rightarrow \psi$ 是连通的;

(ii) 若 $d(p, q)=2$,即存在 G 中的路径 (p, r, q) .由于 $d(p, q)=2$,因此 p, q 不相邻,故由定理 2 可知 p, q 不属于同一域;又因为 r 为 p, q 的邻节点,由算法 DDAM 步骤 3) 中的可知,若 r 不为核心,则 r 一定为网关;故 $r \in \psi \Rightarrow \psi$ 是连通的;

(iii) 若 $d(p, q)=3$,即存在 G 中的路径 (p, r_1, r_2, q) .由于 $d(p, r_2)=2$,因此 p, q 不相邻,故由定理 2 可知, p, r_2 不属于同一域;又因为 r_1 为 p, r_2 的邻节点,由算法 DDAM 步骤 3) 中的可知:若 r_1 不为核心,则 r_1 定为网关,故 $r_1 \in \psi$.由于 $d(r_1, q)=2$,同理可得 $r_2 \in \psi \Rightarrow \psi$ 是连通的.

- 2) 假设 $d(p, q)=m(m>3)$ 时, ψ 是连通的;

3) 对 $d(p, q)=m+1$,即存在 G 中的路径 $(p, r_1, r_2, \dots, r_m, q)$.由 ψ 是图 G 的一个核 $\Rightarrow r_2$ 为 ψ 中的核心或 r_2 与 ψ 中的某个核心 r 相邻 $\Rightarrow d(p, r) \leq 3$.由证明的步骤 1) 可知:节点 p, r 通过 ψ 可达,而 $d(r, q) \leq m$,由归纳假设可知 r, q 通过 ψ 可达 $\Rightarrow \psi$ 是连通的.

由定理 3 可知:由 DDAM 得到的节点集 ψ 构成了一个连通核.所以,可以把传感器网络按“域”划分来构建连

通核,传感节点只需要在连通核中进行路由,这样可以有效降低寻径的开销.

2 基于区间小波变换的数据压缩算法

由多分辨分析(MRA)理论,针对分析空间的正交尺度函数 ϕ 及其对应的小波函数 ψ ,可得如下快速 Mallat 分解算法:

$$c_{j-1,k} = \sum_{m \in Z} a_{m-2k} c_{j,m} \tag{2.1}$$

$$d_{j-1,k} = \sum_{m \in Z} b_{m-2k} c_{j,m} \tag{2.2}$$

及其重构算法:

$$c_{j,m} = \sum_{k \in Z} h_{m-2k} c_{j-1,k} + \sum_{k \in Z} g_{m-2k} d_{j-1,k} \tag{2.3}$$

其中,重构算法是分解算法的逆运算, j 为尺度; $c_{j,m}$ 和 $d_{j,m}$ 分别为尺度 j 上的高频(逼近)系数和低频(细节)系数, $\{h_k\}$ 与 $\{g_k\}$ 称为重构滤波器,而 $\{a_k\}$ 和 $\{b_k\}$ 称为分解滤波器.

在利用 Mallat 算法对传感数据进行小波分解时,式(2.1)和式(2.2)中的 $c_{j,m}$ 表示输入的传感数据信号;而 $c_{j-1,k}$ 和 $d_{j-1,k}$ 表示分解后的信号,其分解过程如图 2 所示.式(2.3)用于对分解后信号的重构,将分解后的信号还原成初始传感数据信号,其重构过程如图 3 所示.

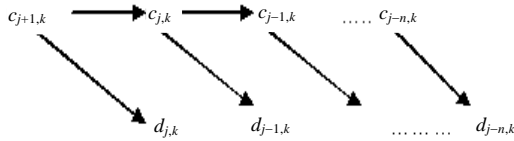


Fig.2 Decomposition process of sensing data

图 2 传感数据的分解过程

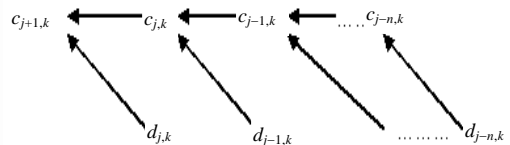


Fig.3 Reconstruction process of sensing data

图 3 传感数据的重构过程

利用上述快速 Mallat 算法进行 3 次小波分解后的空间-频率结构如图 4 所示.



Fig.4 Space-Frequency structure graph of decomposed sensing data

图 4 传感数据被分解后的空间-频率结构图

即原始传感信号被分解为 4 个部分,其中, H_0, H_1, H_2 是高频部分, L_2 是低频部分.原始信号分解后,能量集中在只占少量空间的低频部分 L_2 ,高频部分大量的系数很小,这就为传感数据的压缩提供了很好的机会.传统的小波定义在实轴 R 上,而实际传感信号常常囿于一个有限的区域 K .此时,用 $L^2(R)$ 上的小波来处理,便存在逼近空间 $L^2(R)$ 与传感信号空间 $L^2(K)$ 不匹配的问题,从而产生“边界效应”,导致处理效果不理想.从逼近论角度来看,解决边界效应的方法有两类:一是改造信号空间 $L^2(K)$,使之与小波空间 $L^2(R)$ 相适应,例如进行边界延拓等;二是直接采用适合信号空间的小波,即与信号空间匹配的 $L^2(K)$ 上的所谓区间小波,常常缩约为 $L^2([0,1])$ 区间小波^[15,16].边界延拓实现简单,却有可能在端点带来信号或其导数的不连续性,而且导致必须使用“太多”的小波,增加了运算量.而区间小波避免了这些不足,是传感信号压缩的良好的选择.

定义 8(信息熵). 信息熵又简称为熵(entropy),是随机变量 $I(a_j)$ 的数学期望值,记:

$$H(X) = \sum_{j=1}^m p_j \cdot I(a_j) = -\sum_{j=1}^m p_j \cdot \log p_j ,$$

单位为 bit/字符,其中 p_j 是字符 a_j 出现的概率.

在一些传感器网络中,数据常被简单地表示成属性-值对.例如,一个监测动物的传感器节点产生如下数据:

type=four-legged animal (动物的类型)
instance=elephant (名称)

location=[125,220] (节点的位置)
intensity=0.6 (信号的强度)
confidence=0.85 (可信度)
timestamp=01:20:40 (时间)

若在某一时间内此传感器节点一直在跟踪某个对象,例如 elephant.不难看出,得到的传感数据冗余度很大.若将此数据看成一个信源,可知其信息熵非常低,这就为传感数据进行压缩提供了可能.另外,传感数据有流数据的特点,而传感器的计算能力特别是能量极为有限,对此,如果能够将传感数据进行压缩,则能大量减少数据量,进而减少能耗.我们设计的基于区间小波变换的混合熵编码的流程如图 5 所示.

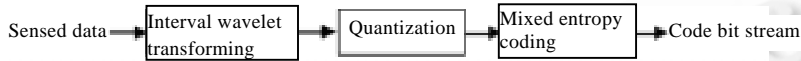


Fig.5 Procedure graph of mixed entropy coding based on interval wavelet

图 5 基于区间小波变换的混合熵编码的流程

小波变换的特点是:传感数据经变换后,绝大部分能量集中在低频系数上,高频部分的大量系数的值趋于 0.小波变换具有多尺度空间-频率分辨能力和时频局域化的特点,对压缩阵发性的流数据非常有效.传感器接收到的数据可以看成是一维信号,采用区间小波进行压缩,无须信号延拓,减少了运算量,且解码时能够完全重构.

量化阶段对小波变换后得到的高频系数和低频系数进行阈值处理,然后根据量化级将小波系数映射到某个整数区间.阈值处理有硬阈值和软阈值两种方法,其共同点是采用阈值来过滤掉值趋于 0 的小波系数,即将某一阈值下的系数略去,只保留那些能量较大的小波系数.对于传感数据的压缩来说,优化算法的运算时间以节省系统能量是至关重要的;从数据处理方面来看,又要使阈值量化后的小波系数能量损失最小.因此,需根据小波变换后系数的实际分布采用“自适应”阈值.

对于量化后的小波系数,其高频部分的大部分系数为 0.游程编码是通过形成串的字符、串的长度及串的位置来进行编码的方法,实现简单,其压缩率取决于数据流中的重复字符出现的次数和平均游程长度,因而在高频部分进行游程编码可以很好地压缩效能.对于低频部分的系数,则采用经典的 Huffman 编码.Huffman 编码依据字符出现概率来构造平均长度最短的异字头码字来进行编码的方法,也称为最佳编码.将这样两种编码方式有机地结合起来,形成混合熵编码,算法简单又有理想的压缩效果.

3 实验仿真

为了具体分析、比较新方法的节能性,我们将新方法应用于 DD 算法进行了模拟实验,并将实验结果与传统 DD 算法的节能性进行了比较.我们的仿真平台采用 SENSE(sensor network simulator and emulator)^[17]——一个专门用来模拟传感器网络的环境的平台.在 SENSE 中,被模拟的传感器网络分为 4 个层次:应用层(application)、网络层(networking)、MAC 层和物理层(physical).我们的算法建立在网络层,在该层可以模拟 AODV(ad-hoc on demand distance vector routing)和 DSR(dynamic source routing).我们首先根据给定的有效通信距离 CR (communication radius),在一个 100×100 的区域内,利用坐标数据随机生成具有 100 个节点的连通图,并对每个节点随机生成 50 个单位以内的能量.然后,对传感器网络中两种典型情况进行了模拟:情况 1 随机抽取 10% 的节点(1 个 Sink 节点,9 个 Source 节点)进行仿真实验;情况 2 随机抽取 1 个 Sink 节点,再随机选取占整个区域 10% 大小的子区域(即一个 10×10 的区域)中的节点作为 Source 节点进行仿真实验.在新方法中,各节点首先利用 DDAM 进行数据汇聚,在数据汇聚过程中的网关选择上也以剩余能量的大小作为选择优先度;然后,核心节点再利用基于区间小波变换的混合熵压缩方法进行数据压缩.由于通过连通核的构造可以有效减小路由搜索区域,因此在具体实验中,我们可仅在连通核中使用 DD 算法进行路由.在模拟实验中,实验数据均为执行新老 DD 算法各 1 000 次,选取其平均值为最后结果.具体实验结果如图 6~图 9 所示.

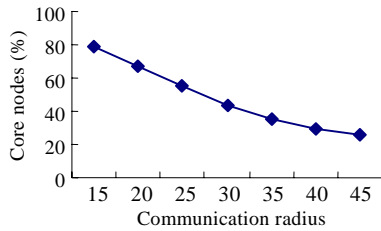


Fig.6 Node proportion in connected core of DDAM in case 1

图6 情况1中DDAM的连通核中节点数比例

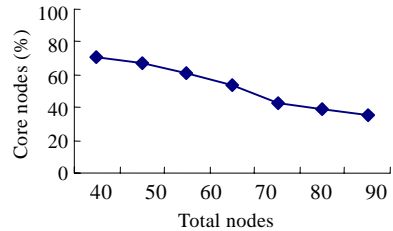


Fig.7 Node proportion in connected core of DDAM in case 2

图7 情况2中DDAM的连通核中节点数比例

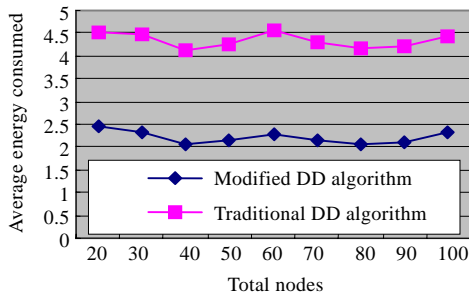


Fig.8 The comparison of energy consuming case 1

图8 情况1中新、老DD算法的节点耗能比较

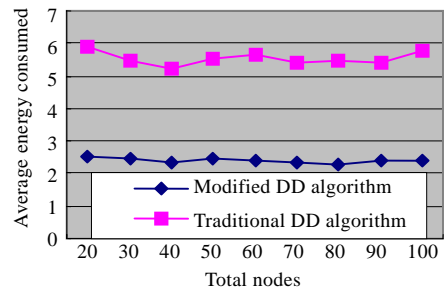


Fig.9 The comparison of energy consuming case 2

图9 情况2中新、老DD算法的节点耗能比较

图6和图7分别给出了实验情况1和 $CR=35$ 时情况2条件下,运用DDAM算法后,连通核中的节点数占总节点数比例的实验数据.由图6、图7可知,连通核的大小和 CR 、网络的节点稠密程度成反比,其大小仅为网络节点总数的20%~80%.因此,连通核的构造可有效降低算法的寻径时间复杂度.由图8、图9可知,利用基于区间小波变换的混合熵压缩方法进行数据压缩,可进一步降低DD算法的节点耗能,从而延长整个网络的生命周期.

4 结 语

基于传感器节点的位置信息,提出了基于连通支配集的传感器网络定向传播模型.综合考虑传感器网络的拓扑与数据特性,结合数据汇聚与数据压缩思想,提出了一种基于“域”概念的分布式数据汇聚模型,并在数据汇聚的基础上,结合小波变换对流数据的良好压缩特性,提出了一种基于区间小波变换的混合熵数据压缩方法,并将新方法结合传统的DC算法-DD路由算法进行了比较仿真实验.实验结果表明:与传统的DC算法-DD路由算法相比,新算法可进一步提高DC路由算法的节能性能.

References:

- [1] Xie ZJ, Wang L, Chen H. Subnets based distributed data-centric hierarchical ant routing for sensor networks. In: Liu JN, Connors A, eds. IEEE Int'l Conf. on Wireless Communications, Networking and Mobile Computing. Wuhan: IEEE Computer Society, 2005. 895-900.
- [2] Agre J, Clare L. An integrated architecture for cooperative sensing networks. IEEE Computer, 2000,33(5):106-108.
- [3] Ren FY, Huang HN, Lin C. Wireless sensor networks. Journal of Software, 2003,14(7):1282-1291 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1282.htm>
- [4] 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>
- [5] Srinivasan V, Chiasserini CF, Nuggehalli PS, Rao RR. Optimal rate allocation for energy-efficient multipath routing in wireless ad hoc network. IEEE Trans. on Wireless Communications, 2004,3(3):891-899.
- [6] Estrin D, Govindan R, Heideman J, Kumar S. Next century challenges: Scalable coordination in sensor networks. In: Heidemann J, ed. Proc. of the 5th Annual ACM/IEEE Int'l Conf. on Mobile Computing and Networking. Seattle: ACM Press, 1999. 263-270.

- [7] Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. A survey on sensor networks. *IEEE Communications Magazine*, 2002, 40(8):102-114.
- [8] Pottie G, Kaiser W. Wireless sensor networks. *Communications of the ACM*, 2000,43(5):51-8.
- [9] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In: Johnson DB, ed. *Proc. of the ACM MobiCom*. Boston: ACM Press, 2000. 56-67.
- [10] Krishnamachari B, Estrin D, Wicker S. Modelling data-centric routing in wireless sensor networks. In: Stojmenovic I, Olariu S, eds. *Proc. of the IEEE Infocom*. New York: IEEE Computer Society, 2002. 2-14.
- [11] Heinzelman W, Kulik J, Balakrishnan H. Negotiation based protocols for disseminating information in wireless sensor networks. *ACM Wireless Networks*, 2002,8(2-3):169-185.
- [12] Yu Y, Govindan R, Estrin D. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical Report, UCLA/CSD-TR-01-0023, UCLA Computer Science Department, 2001.
- [13] Lin YP, Wang L. A distributed data-centric clustering hierarchical routing algorithm for sensor networks. *ACTA ELECTRONICA SINICA*, 2004,32(11):1883-1889 (in Chinese with English abstract).
- [14] Peng W, Lu XC. A new distributed approximate algorithm in minimum connected dominating sets. *Chinese Journal of Computers*, 2001,24(3):254-258 (in Chinese with English abstract).
- [15] Gao XP, Zhang B. Interval-Wavelets neural networks (I)——Theory and implements. *Journal of Software*,1998,9(3):217-221 (in Chinese with English abstract).
- [16] Gao XP, Zhang B. Interval-Wavelets neural networks (II)——Properties and experiment. *Journal of Software*,1998,9(4):245-250 (in Chinese with English abstract).
- [17] Chen G, Branch J, Pflug MJ, Zhu L, Szymanski B. SENSE: A sensor network simulator and emulator. In: Schmidt D, ed. *Proc. of the 2nd Int'l Conf. on Pervasive Computing*. Vienna: Springer-Verlag, 2004. 249-267.

附中文参考文献:

- [3] 任丰原,黄海宁,林闯.无线传感器网络. *软件学报*,2003,14(7):1282-1291. <http://www.jos.org.cn/1000-9825/14/1282.htm>
- [4] 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>
- [13] 林亚平,王雷.传感器网络中的分布式数据汇聚层次路由算法. *电子学报*,2004,32(11):1883-1889.
- [14] 彭伟,卢锡诚.一个新的分布式最小连通支配集近似算法. *计算机学报*,2001,24(3):254-258.
- [15] 高协平,张钹.区间子波神经网络()——理论与实现. *软件学报*,1998,9(3):217-221.
- [16] 高协平,张钹.区间子波神经网络()——理论与实现. *软件学报*,1998,9(4):245-250.



谢志军(1974 -),男,湖南湘乡人,博士生,主要研究领域为传感器网络,数据流.



陈红(1965 -),女,教授,博士生导师,CCF高级会员,主要研究领域为数据流,传感器网络,数据流.



王雷(1973 -),男,博士,副教授,主要研究领域为计算机网络.



刘永和(1974 -),男,博士,副教授,主要研究领域为计算机网络.



林亚平(1955—),男,教授,博士生导师,主要研究领域为计算机网络.