

基于 Bayes 序贯估计的无线传感器网络数据融合算法

张书奎^{①②} 崔志明^① 龚声蓉^① 孙涌^① 方巍^{①②}

^①(苏州大学计算机科学与技术学院 苏州 215006)

^②(江苏省计算机信息处理技术重点实验室 苏州 215006)

摘要: 移动代理被认为是无线传感器网络中解决数据融合的有效方法,但代理访问节点的次序以及总数对算法有较大影响,为此该文提出一种基于 Bayes 序贯估计的移动代理数据融合算法.该算法通过构造特定数据结构的报文,在多跳环境中由 Bayes 序贯估计调整梯度向量,据此动态决定移动代理的访问路径,使移动代理有选择地在传感器节点之间移动,且在节点处由移动代理对数据进行融合,将多余的感知数据剔除,而不是把原始数据传输到 Sink 节点.理论分析和模拟实验表明,该算法有较小的能量消耗和传输延时.

关键词: 无线传感器网络; 移动代理; 数据融合; Bayes 序贯估计

中图分类号: TP393

文献标识码: A

文章编号: 1009-5896(2009)03-0716-06

A Data Fusion Algorithm Based on Bayes Sequential Estimation for Wireless Sensor Network

Zhang Shu-kui^{①②} Cui Zhi-ming^① Gong Sheng-rong^① Sun Yong^① Fang Wei^{①②}

^①(School of Computer Science and Technology, Soochow University, Suzhou 215006, China)

^②(Jiangsu Key Laboratory of Computer Information Processing Technology, Soochow University, Suzhou 215006, China)

Abstract: Mobile Agent(MA) is more suitable for wireless sensor networks than the C/S model in data fusion. In MA based data fusion, the order of nodes visited along the route by MA has a significant impact on the algorithm efficiency and life time of wireless sensor networks. This paper proposes a Mobile Agent Data Fusion (MADF) algorithm based on bayes sequential estimation for wireless sensor networks. By designing data packet and data table with specific structure, and considers MA in multihop environments and adopts gradient of Bayes sequential estimation to dispatch MA. MA accounts for performing data processing and making data aggregation decisions at nodes rather than bring data back to a central processor (sink), redundant sensory data will be eliminated. Theoretical analysis and experimental results show that the proposed scheme is able to provide less energy consumption and network delay compared to directed diffusion schemes.

Key words: Wireless sensor networks; Mobile Agent; Data fusion; Bayes sequential estimation

1 引言

无线传感器网络(WSNs)的主要作用是对特定环境进行监测,将感知到的数据发送到基站进行进一步的处理.由于应用的需求,传感器节点需要高密度部署,而节点只有有限的计算能力、受限的电力供给以及较小的网络传输带宽等,这决定了WSNs难以传输大量的感知数据.如何有效地融合这些感知数据以减少网络流量,提高资源的利用率,同时也为提高故障容忍的能力,延长网络生命周期,一直是传感器网络研究的重要课题.

面对这些挑战,许多研究工作在延长网络寿命,提高传感器节点的监测精度,支持容错处理等(如传感器故障以及电池电源耗尽)方面作出了不懈的努力,其中大部分处理方法是

基于传统的客户/服务器计算模式^[1,2].在这种模式中,每个传感器节点发送它的感知数据到处理中心或者Sink节点,这会导致所传输的数据超过网络负载的阈值.例如,两个彼此接近的传感器在感知数据时,它们很可能有多余的或共同的数据.因此,在布署稠密的传感器网络中,为减少感知数据的传输量而进行数据融合和定向扩散(DD)就是一个非常重要的过程.对于数据融合,各个节点之间的协同就是非常重要的了^[3],但有效协同的关键是传感器低层数据处理的整合以及在相邻节点间进行一致的本地数据交换^[4],WSNs的这种特性提出了一些关于协同处理的重要议题,包括能量的效率、可量测性、可靠性等.对于定向扩散,它的一个重要特征就是每一个传感器节点都可以被应用感知,也就意味着不同节点之间存在着联系,而不仅仅是简单的发送.

Mobile Agent(MA)被认为是传感器网络中非常有效的数据融合和定向扩散方法^[5-8],可以解决传输感知数据过载问题^[5].Qi 等人在文献[2]提出了基于移动代理的分布式传感器网络(Mobile Agent-based Distributed Sensor

2008-01-11 收到, 2008-09-29 改回

国家自然科学基金(60673092,60873116)和教育部科研重点项目(207040),江苏省自然科学基金(BK2008161),江苏省重大科技支撑与自主创新项目(BE200844)和苏州大学科研预研基金资助课题

Network, MADSN)进行信号协作与信息处理。通过传输软件实体,即移动代理 MA 到传感器节点,有选择地在传感节点之间移动,MA 周期性访问目标区域节点,利用本地资源完成数据采集,融合感知数据,进行定向扩散等,逐步获得精确的全局目标信息,从而达到消除冗余,将大量感知数据减少或转换为更小数据的目的。然而,如何部署 MA 以及 MA 在目标节点如何迁移对能量消耗有较大影响^[2],找出一个最佳的节点访问次序是 NP-完全问题^[3]。本文提出一种基于 Bayes 序贯估计的移动代理数据融合算法(MADF),通过构造特定数据结构的报文和数据表,在多跳环境中由 Bayes 序贯估计来调整梯度向量,动态决定 MA 的访问路径,使 MA 有选择地在传感节点之间移动,并且在节点处由 MA 对数据进行融合处理,将多余的感知数据剔除,而不是把原始数据传输到 Sink 节点,从而减少了能量消耗,延长了网络寿命。为评估该算法,本文作了形式化分析。另外,通过设置恰当的参数,一系列的仿真实验结果表明,在节点之间的传输延迟、传输效率、能量消耗等方面,本文所提出的移动代理数据融合算法比直接的定向扩散有更好的性能表现。

2 协同移动代理数据融合算法

文献[11]提出了一种基于遗传算法的解决方案,尽管可以通过遗传算法得到一个整体最优解,但这对于资源受限的传感器节点来说却不是最佳方法^[2-4]。本文采用一种基于 Bayes序贯估计来调整梯度向量,以便MA能够动态地选择合适的迁移路径。

2.1 算法概述

图 1 显示了 MA 在 WSNs 环境中的操作过程,当 Sink 收到应用程序分配的任务请求时,它广播查询信息包,其中包含了感知任务、兴趣区域描述以及其他信息。如果某一节点发现它满足查询兴趣,就声明自己是一个目标源节点。每一个目标源节点通过探测生成一个响应。这样, Sink 从各源节点处接收到大量的响应数据包,然后由 Bayes 序贯估计决定出 MA 的访问节点队列。MA 的操作从 Sink 调度 MA 时开始,在 MA 生成报告后返回 Sink 结束。在大部分的情况下,每个目标源节点会在一定间隔时间内周期性地产生感知数据,这就意味着相同的 MA 需要被多次运行。因此,当 MA 在到达第 1 个节点时,它会被存储,之后,创建一个 Create-MA-Timer,被用来触发下一轮 MA 的迁移,并以此来收集相关信息源发来的数据。显然,连续两轮之间的时间间隔应该和 Create-MA-Timer 的值设置相一致。这个属性也可定义为从 MA 在 FirstNo 收集数据到在 LastNo 收集数据之间的间隔。然后 MA 继续访问其他的源节点,直到它到达最后一个节点。这个过程将会被重复,在最后一轮结束的时候,任务被完成。此算法有如下优点:

首先,当收集到的融合数据沿着被加强的路径返回 Sink 节点时,达到了通过 MA 减少传输数据量的目的。被 MA 收

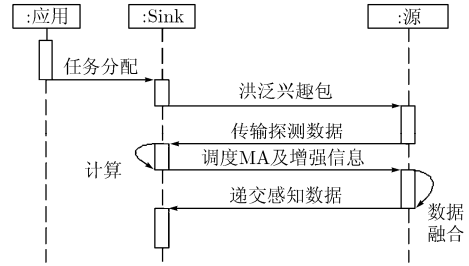


图1 MA操作序列示意图

集到的数据 S 的大小可以通过以下公式计算^[10]:

$$\left. \begin{aligned} S_{ma}^1 &= R_1 \\ S_{ma}^2 &= R_1 + (1 - \rho) \times R_2 \\ &\dots \\ S_{ma}^i &= S_{ma}^{i-1} + (1 - \rho) \times R_i \\ &= R_1 + \sum_{k=2}^i (1 - \rho) \times R_k \end{aligned} \right\} \quad (1)$$

这里 ρ 是融合率, S_{ma}^i 是当 MA 离开传感器节点 i 时的融合数据量, R_i 是将会被 ρ 融合的数据量。在 MA 从一个节点到另外一个节点的转移期间,它在融合感知数据的同时消除冗余数据。被融合的数据使用下式计算:

$$S^i = \sum_{k=1}^i R_k \quad (2)$$

其次,能量消耗将被减少。在以客户/服务器为基础的传感器网络计算模式中,目标区域的所有的源节点以一个确定的时间间隔向 Sink 发送数据;在以 MA 为主的方法中,MA 在传输运行处理代码的同时还要考虑源节点访问的次序。

2.2 MA 数据包格式

MA 信息包的结构如图 2 所示。SinkID 和 MA_SeqNum 用来识别一个数据包。每当 Sink 调度一个新的 MA 数据包时,都会使 MA-SeqNum 递增。FirstNo 和 LastNo 分别是将会被 MA 访问的第 1 和最后一个节点,它们分别指明了 MA 收集数据的开始和结束。RoundIdx 是当前轮的索引,它的值在 Sink 发起第 1 轮时被定义为 1,在接下来的轮中被 FirstNo 增加。LastRoundFlag 指明当前轮是否是此次任务的最后一轮,这个标记通过 FirstNo 来设定。当一个有着 LastRoundFlag 标记的 MA 到达一个源节点时,它能使系统资源在相关的操作执行完毕后得到释放。

当一个 MA 迁移时,它能改变某些可变属性。NextNo 指出下一个将要被访问的源节点,NextHop 指出下一个要访



图 2 MA 包结构

问的节点, 这个节点可能是一个中间节点, 也可能是一个目标源节点。如果 NextHop 与 NextNo 相同, 这就意味着下一跳就是当前的目标节点。Visiting-Sequence 包含了当前轮中尚未被访问的目标源节点的 ID, 它并不包含任何关于 Source-Visiting 的次序信息, 因为当 MA 到达一个目标源节点时 (LastNo 除外), NextNo 是动态决定的。当一个 MA 被创建时 Visiting-Sequence 初始包含了所有目标源节点的 IDs。当 MA 访问某个目标源节点以后, Visiting-Sequence 中相应的 IDs 被删除, ToSink 标志将会被设置以显示 MA 已访问过的节点。NextNo, NextHop, Visiting-SequenceList, 以及 SinkFlag 指明了 MA 动态迁移的路线。有效载荷部分包括两个数据字段, 一个是用来处理感知数据的 Processing Code; 另一个是将被传输的融合数据。Aggregated Data 的初始大小是 0, 随着 MA 从一个节点到另一个节点的迁移, Aggregated Data 的大小也随之改变。

2.3 Bayes 序贯估计

本文采用 Bayes 序贯估计来确定将要选择的下一个节点。已知节点的度量值为 z_1, z_2, \dots, z_n , 设定对要评估的目标节点 x 的后验概率为候选节点的信任值, 其密度函数为 $p(x | z_1, \dots, z_n)$, 每个新的传感器节点度量值 z_i 与当前的估计值 $p(x | z_1, \dots, z_n)$ (称为信任状态) 有关联, 共同形成一个关于被访问目标新的信任状态估计值 $p(x | z_1, \dots, z_n)$, 选择一个将被访问的传感器节点 i 的问题就是在最小代价方面提供最佳估计值问题, 也是一个优化问题, 这个优化问题的目标函数与信息获取和资源消耗代价有关, 定义如下:

$$M(p(x | z_1, \dots, z_f)) = \alpha \cdot \varphi_{\text{utility}}(p(x | z_1, \dots, z_f)) - (1 - \alpha) \varphi_{\text{cost}}(z_f) \quad (3)$$

这里 φ_{utility} 是所获取信息的有效性度量值, 其值 $0 < \varphi_{\text{utility}} \leq 1$, φ_{cost} 是通信及其他资源消耗的代价, α 是信息有效性与资源消耗代价的权重。根据当前的信任估计来选择被访问的下一个传感器节点, 这个节点可能是中继节点, 此时, 其信任估计值在 TTL 内驻留在该节点, 并且最终将其信任信息传输到首节点。

首先, 目标函数 $M(\cdot)$ 反映了传感器节点 j 所感知数据的有效性, 例如, 在有限的节点感知范围内, 感知数据的有效性通过与节点的接近程度来衡量。第二, 获得信息的代价通过链路带宽、传输延迟、节点电池电力消耗等节点的度量值来测量。在 Sink 节点固定的情况下, 这是从节点 j 请求数据, 获得数据, 将数据返回到 Sink 节点构成信任状态值。在这种情况下, 转发数据从节点 i 到节点 j 是能量消耗的主要方面, 通信代价是节点 i, j 间距离的函数, 此时目标函数演变为如下形式:

$$M(x_j) = -\alpha(x_j - x'_r)^t \cdot \sum^{-1}(x_j - x'_r) - (1 - \alpha)(x_j - x'_r)^t(x_j - x_i) \quad (4)$$

这里, $x'_r, \sum^{-1}, x_j, x_i$ 分别是目标位置、协方差、已查询过的节点位置及正在查询的节点位置。我们的任务是确定出哪个

传感器含有最有用的信息, 并且传输这些信息到查询节点。注意到, 当 x'_r, \sum^{-1} 值传递到下一个节点时, 所有的路由决定是由本地完成的。随着网络的拓扑变化, 在路由动态改变期间, 由 Sink 广播查询中获知的信息可调整信任值。当用户查询在 Sink 进入传感器网络, 从 Sink 广播查询中获知的信息, 查询转向潜在事件的区域, 图 3 给出了 Bayes 序贯估计来确定的节点访问次序示例。

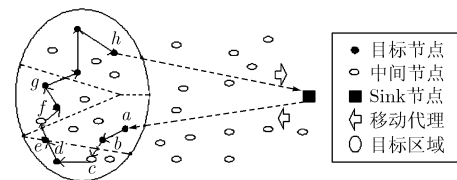


图3 目标节点被访问次序示意图

2.4 协同路由

开始时, Sink 广播兴趣通告, 一旦目标源节点接收到了相应的消息, 就会通过多种路径向 Sink 发送监测到的数据。如果 Sink 有多个前一节点, 它会选择一个最邻近节点来接受数据。因此, Sink 就要加强它的首选邻节点, 以此类推要加强其首选的前一跳节点等。源节点周期性地发送额外的探测消息来修改 Bayes 序贯估计, 以适应网络变化(如节点失效, 能源耗尽, 位置变动等)而导致的偏差。

传感器网络的最终的目标是监测事件的发生^[11]。如果监测源节点数量足够大, 满足事件监测的需要, Sink 会停止发送任何探测信号流, 因而所有的源节点或节点中的一个子集被 MA 选择性地访问。在所有将被访问的源节点中, Sink 会选择一个初始节点和一个最后节点, 之后 Sink 会产生一个有着图 2 结构数据包的 MA, 并将之调度到第 1 个节点。与此同时, Sink 会加强到达最后一个节点的路径。当 MA 到达第 1 个节点时, 它被存储其中。我们把整个任务划分成轮, 每一轮都需要 MA 访问选定了的节点, 并将之传送给 Sink。从第 1 个节点到最后一个节点, MA 会携带数据结果沿着加强了的路径回到 Sink。在第 1 轮中, 除了 MA 从一个节点到另一个节点动态收集数据之外, 它也将处理代码 Processing Code 复制到每个节点的存储器中去。在每一轮开始时, 第 1 个节点会构造另一个 MA 并将其调度以初始化新一轮。由于在一轮以后, 每一个节点都已存储了处理代码, 因而以后的轮中, MA 不再包含处理代码。任务结束后, 源节点将以前的处理代码丢弃。

目标源节点的探测消息使传感器节点依据 Bayes 序贯估计设立 ToSourceEntry, 这是一种目标源节点的梯度向量, 用来为 MA 的漫游选路服务的。ToSourceEntry 的值随着探测消息的扩散而减小。如果生命周期 Time-To-Live(TTL)

的值为零, 则传感器节点将不再设立任何 ToSourceEntry。在所有传感器的邻节点中, 只有最先转发探测消息源节点, 即 Bayes 序贯估计值高的节点, 会被选择成为传感器节点的下一跳。在图 4 中, 节点 A, B, C, D 是目标探测节点。ToSourceEntry 被 A, B, C, 12, 和 D 设立, 如表 1。

表1 ToSourceEntry(SeqNum=5)设置

	A		B		C		D	
	Next hop	Cast (ms)	Next hop	Cast (ms)	Next hop	Cast (ms)	Next hop	Cast (ms)
A	-	-	B	4.46	B	8.24	B	16.32
B	A	4.47	-	-	C	4.43	C	12.89
C	B	8.16	B	4.32	-	-	12	8.52
12	C	9.65	C	7.56	C	4.86	D	5.08
D	12	14.15	12	12.67	12	8.73	-	-

基于Bayes序贯估计的ToSourceEntries, MA的一条迁移路径与下列3种因素密切相关。

(1)选择 FirstNo 和 LastSrc。根据式(1), 在 FirstNo 时 MA 最小, 在 LastNo 时 MA 最大。因此为了减少总的通讯开销, FirstNo 应该选择最远的节点, 而 LastNo 应该选择最近的。在本文中, 第 1 个(最后一个)向 Sink 发送探测信息的节点是 FirstNo(LastNo), Sink 会加强到 LastNo 的路径。

(2)决定源节点访问的次序。除了 FirstNo 和 LastNo 被 Sink 选定之外, 其他节点的访问次序是在目标传感器的访问序列中动态生成的。例如, 在图 4 中 MA 到达节点 A 时节点会在 ToSourceEntry(如表 1)的基础上选择最近的探测节点。因为反应时间最短的节点为 B, 所以它就是最近的节点并会成为 NextNo。

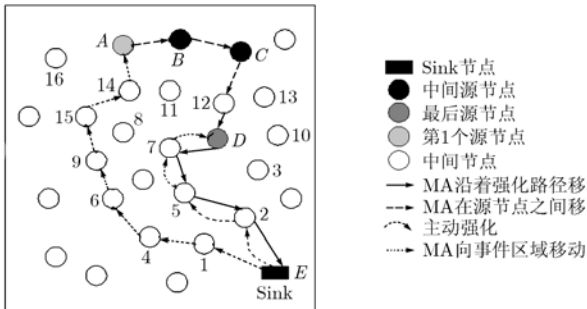


图 4 MA 协同路由

(3)寻找节点下一跳 MA 路由。Sink 调度始, MA 在节点之中迁移, 从 FirstNo 开始的迁移方式就像一个加强的信号被转发, 它的下一跳节点将会通过当前节点的 ToSourceEntry 决定。MA 最终会通过加强的路径回到 Sink, 例如如图 4 中的路径 D-7-5-2-E。

2.5 性能分析

设传感器网络中有 n 个节点 $S_k(1 \leq k \leq n)$, 每个节点成功完成 MA 指派任务的概率为 p , MA 在节点 k 处理数据需要消耗的能量为 E_k , 在节点 k, i 之间移动所消耗的能量为 E_{kj} , MADF 以最小化期望时间(以跳数的形式)和能耗来成功完成任务。不失一般性, 作出以下假设, 以简化 MADF 模式: (1)目标静止不动。(2)对于所有 MA 迁移步骤, 有 $E_{kj} = E$ 。(3)对于所有传感器节点, 有 $E_k = e$ 。

完成任务的能量消耗或者访问所有节点均失败, 形成路由由 $R = \langle S_1, S_2, \dots, S_n \rangle$ 表示, 那么

$$E_R = E + e + p_1E + \sum_{i=2}^n \left(\prod_{j=1}^{i-1} (1 - p_i)(E + e + p_iE) \right) + \prod_{j=1}^n (1 - p_j)E \quad (5)$$

这里, S_1 是被访问消耗的能量 E , 到达目的区域后, 不论成功或失败, 能量 E 都要消耗, 任务成功完成的概率是 p_1 , 接着 MA 返回到 0 节点, 同时消耗能量 E 。然而, MA 以概率 $(1 - p_1)$ 迁移到节点 S_2 , MA 从节点 S_1 迁移到 S_2 的消耗能量是 $(1 - p_1)E$ 。同理, MA 从第 $i-1$ 个节点迁移到第 i 个节点的消耗能量是 E , 随后, MA 在节点 i 以概率 p_i 被成功接收并返回 0 节点, 同时消耗能量 E 。因此, 在节点 S_i 累积的能耗是 $\prod_{j=1}^{i-1} (1 - p_i)(E + e + p_iE)$ 。当所有节点都访问失败, 就是最后一轮, 则须返回原始节点 0, 消耗能量 E 。则

$$H_{opR} = 1 \cdot p_1 + \sum_{i=2}^n \left(i \cdot \prod_{j=1}^{i-1} (1 - p_j)p_i \right) + (n + 1) \prod_{j=1}^n (1 - p_j) \quad (6)$$

采用下面的等式表示成功率:

$$p_k = 1 - \left(D_{esire} - \sum_{i=0}^{\text{hop}} I_k \right) / I_{\max} \quad (7)$$

式中 I_{\max} 是一个传感器所能提供的最大信息。 H_{opR} 是 MA 迁移经过的总的节点数, 得出如下定理:

定理 1 如果节点以 $I_k(k = 1, 2, \dots, n)$ 的降序访问, 即, $I_1 > \dots > I_k > \dots > I_n$ 。MADF 将得到的最佳路径。

证明 考虑两个相邻节点的路由顺序, 例如 k 和 $k+1$, 称这个新的路径为 R' , 只受到第 k 和 $(k+1)$ 交换影响, 那么能耗的差异是

$$E_R - E_{R'} = (E + e) \prod_{j=1}^{k-1} (1 - p_j)(p_{k+1} - p_k) \quad (8)$$

跳数是

$$H_{opR} - H_{opR'} = \prod_{j=1}^{k-1} (1 - p_j)(p_{k+1} - p_k) \quad (9)$$

由于 $p_k > p_{k+1}$, R 是一个规模较小能源消耗和跳数较小的较佳路线。这表明, 如果路径上第 k 个节点比第 $(k+1)$ 个节点使用代理完成任务可能性小, 可以通过交换它们来减少消耗能量和跳数。

从式(7)可以看出, 成功的可能性与 MA 累积的总信息

量直接有关。MA 携带的总信息越多, 将越有可能在当前传感器节点完成任务, 成功的概率也就越高。因此, MADF 最佳路径是减少信息获取的序列。 证毕

定理 2 如果对于所有传感器节点有 $p_k = p$, 那么在大量节点存在时, MA 移动的跳数 H_{opR} 为 $1/p$ 。

证明

$$H_{opR} = 1 \cdot p + \lim_{x \rightarrow \infty} \left[\sum_{i=2}^n \left(i \cdot \prod_{j=1}^{i-1} (1-p_j) \right) p + (n+1) \prod_{j=1}^n (1-p) \right] \\ = \sum_{i=1}^{\infty} i(1-p)^{i-1} p + \lim_{x \rightarrow \infty} (n+1) \prod_{j=1}^n (1-p) = 1/p \quad (10)$$

证毕

定理表明, 要在每个传感器节点上提高成功的可能性, 可以通过减少 MA 的跳数来达到。

3 模拟实验和评价

为了展示 MADF 的性能, 选择基于客户机/服务器的计算模式(DD)与 MADF 比较。使用 NS2 为仿真平台。每项任务以恒定比特率传输数据包 1 packet/s。Sink 和所有传感器节点都是静态的。实验中的基本参数值如表 2 所示, 它对所有实验是共用的。每个实验运行 60 次, 结果取其平均数。

表 2 仿真实验参数设置

参数	值
网络规模	800m×800m
网络拓扑	随机部署
传感器节点数	1500
MAC 层数据速率	1Mbps
节点通信范围	60m
感知数据间隔时间	0.33

对 3 项性能指标进行评价: 包交换率 u , 它是源节点产生的数据量与这些数据被传输到 Sink 节点数据量的比; 分组成功转发到 Sink 节点的能量消耗 e , E_{total} 表示全部能量消耗, 包括传输, 接收, 处理等过程 n_{data} , 标示传输到 Sink 的分组数量, 则 $\alpha = E_{total} / n_{data}$; 端到端分组平均延迟 T_{ete} , T_{dd} 和 T_{ma} 分别表示在 DD 和 MADF 情况下平均端到端延迟。

虽然实验结果受多参数的影响, 但有一组重要的参数要选择, 在缺省情况下, 对所有传感器节点成功地完成任务的概率($p=0.6$), 任务持续的时间(T_{task}), 数据压缩比($r=0.8$), MA 访问延迟($\tau=9ms$), 融合率($q=0.2$), 每一个传感器感知数据的大小($S_{data}=1kB$), 信息有效性与资源消耗代价的权重 $\alpha=0.35$ 。如果设定 $q=0$, 意味着数据融合未进行。每组只有一个参数(如 T_{task} , r, q 和 S_{data})改变, 其他参数不变。以下是实验结果。

在图 5 所示的实验中, 将 T_{task} 从 10s 增加到 600s, 能量消耗 e 随着 T_{task} 增加而减少。当 T_{task} 很小(低于 60s), MADF 比 DD 具有更高的 e 值, 这是因为 MADF 消耗能量(E_q)迁

移 MA 代码从 Sink 到目标区域。 T_{task} 值小时, n_{data} 也值小, 但 e 值大, 但是当 T_{task} 超出 100s, $r=0.8$, $u=0.2$ 时, MADF 比 DD 具有较低的 e 值。因此, 作为 MA 从一个源节点迁移到另一个节点所消耗代价的补偿, 源节点应该处理足够长的数据流。

在图 6 所示的实验中, 只改变 MA 存取延迟(τ), τ 由 0s 至 0.05s, T_{dd} 是常量, 因为改变 τ 不影响 DD。由于当 MA 访问每个源节点时, 由于存取延迟 τ 的影响, 该值设置为一个较大的值时, τ 将导致 T_{ma} 的较快增加。当 τ 超越 0.045s, 并且 $r=0.8$, $u=0.2$ 时, MADF 比 DD 有较大的端到端延迟。这说明 τ 的值依赖于 MA 系统的中间环境。

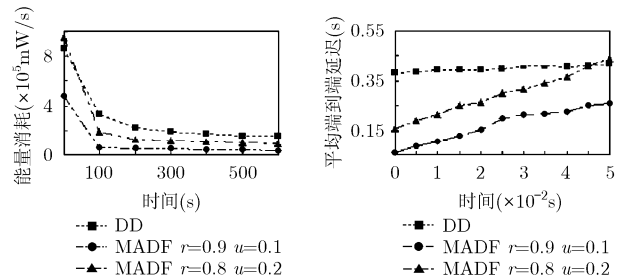


图5 成功传输分组的能量消耗

图6 端到端传输延迟

在图 7 所示的实验中, 改变了传感器(S_{data})传输数据的大小, 由 0.5kB 增加到 2kB 每次增加 0.25kB, 其它参数不变。结果显示, 在融合率方面 MADF 性能始终优于 DD。另外, 在 MADF 中, 每一轮只以单一数据流发送。相比之下, DD 从单个源节点以多数数据流形式发送, 这样一来, 拥塞在 DD 中比在 MADF 中更有可能发生。当 S_{data} 增加时, 拥塞更为严重。

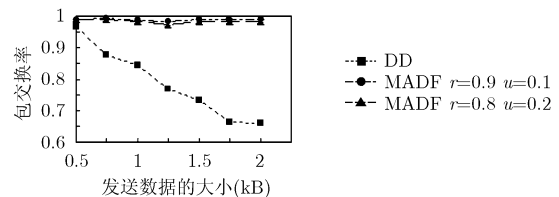


图7 包交换率

通过实验可以看出, MADF 表现出更好的一致性和相对较高的可靠性, 比 DD 更低的能量消耗, 这与分析结果是一致的。同时也说明在 MADF 中应恰当地选择 r 和 q 。注意 r 和 q 的值取决于应用的类型, 在决定采用 MADF 进行数据融合之前, 应首先研究应用的特征。

4 结束语

在无线传感器网络中, 当目标节点稠密部署并伴随有大量有冗余的感知数据生产的环境中, 各个节点传输所有感知数据不仅浪费网络带宽, 而且消耗很多电池能量。基于 MA

的分布式传感器网络来进行信号及信息的协作处理, 被认为是解决这些问题的有效方法。本文提出了一种新型的移动代理数据融合算法, 该算法通过 Bayes 序贯估计来调整梯度向量, 据此判断在多跳环境中来动态决定 MA 的访问路径, 从而有选择地在目标节点之间移动, 同时在目标源节点上利用本地资源完成数据融合。采用这一算法, 可以有效地完成 MA 从 Sink 到目标源节点、目标源节点之间、从源节点到 Sink 的迁移过程, 另外多余的感知数据将被剔除。通过形式化分析和仿真实验, 结果表明, 在端到端的传输延迟, 包交换率, 以及成功传输分组的能量消耗等方面, 都有较好的性能表现。

参 考 文 献

- [1] Akyildiz I F and Weilian S, *et al.* A survey on sensor networks[J]. *IEEE Communications Magazine*, 2002, 40(12): 102-114.
- [2] Qi H, Y Xu, and Wang X. Mobile-agent-based collaborative signal and information processing in sensor networks[J]. *Proc. IEEE*, 2003, 91(8): 1172-1183.
- [3] Chang Jiun-jian, Hsiu Pi-cheng, and Kuo Tei-wei. Search-oriented deployment strategies for wireless sensor networks [C]. 10th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC'07), Santorini Island, Greece, May 2007: 164-171.
- [4] Akkaya K and Younis M. A survey on routing protocols for wireless sensor networks[J]. *Ad hoc Networks*, 2005, 3(3): 325-349.
- [5] Joe I. A path selection algorithm with energy efficiency for wireless sensor networks[C]. 5th ACIS International Conference on Software Engineering Research, Management & Applications (SERA 2007), Busan, South Korea, August 2007: 419-423.
- [6] Chen M, Kwon T, and Choi Y. Data dissemination based on mobile agent in wireless sensor networks[C]. Proceedings of the IEEE Conference on Local Computer Networks 30th Anniversary (LCN05), Sydney, Australia, 2005: 1-2.
- [7] Lee Min-gu and Lee Sunggu. Data dissemination for wireless sensor networks[C]. 10th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC'07), Santorini Island, Greece, May 2007: 172-180.
- [8] Wook C and Das S. A novel framework for energy-conserving data gathering in wireless sensor networks[C]. Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM05), Miami, USA, 2005: 1985-1996.
- [9] Wu Q, Rao N S V, and Barhen J, *et al.* On computing mobile agent routes for data fusion in distributed sensor networks[J]. *IEEE Trans. on Knowledge and Data Engineering*, 2004, 16(6): 740-753.
- [10] Hu Haifeng and Yang Zhen. Mobile-agent-based information-driven multiresolution algorithm for target tracking in wireless sensor networks[C]. Qingdao, China, August 2007: 521-525.
- [11] Shakshuki E, Xing Xinyu, and Malik H. Mobile agent for efficient routing among source nodes in wireless sensor networks[C]. Third International Conference on Autonomic and Autonomous Systems (ICAS'07), Athens, Greece, June 2007: 39-44.

张书奎: 男, 1966年生, 博士, 副教授, 研究领域为无线传感器网络、分布式计算、网络安全。

崔志明: 男, 1961年生, 教授, 博士生导师, 研究领域为分布式计算、Deep Web 数据集成、数据挖掘。

龚声蓉: 男, 1966年生, 博士, 教授, 研究方向为无线网络、多媒体技术、智能信息处理。

孙 涌: 男, 1958年生, 博士, 副教授, 研究方向为无线网络、分布式计算、智能信息处理技术。

方 巍: 男, 1975年生, 博士生, 主要研究方向为 Deep Web 数据集成、智能信息处理技术。