

黄海松,秦志远,张 慧. 基于农业物联网的农作物生长监测数据融合研究[J]. 江苏农业科学,2017,45(21):241-243,251.  
doi:10.15889/j.issn.1002-1302.2017.21.068

# 基于农业物联网的农作物生长监测数据融合研究

黄海松,秦志远,张 慧

(贵州大学现代制造技术教育部重点实验室,贵州贵阳 550025)

**摘要:**针对农业无线传感器网络数据融合精度低、传感器节点能量有限的问题,结合农作物生长环境监测的特点,提出了一种适用于农作物生长监测数据融合的架构和算法。将无线传感器网络划分成多个固定层级,减少数据传输量并逐层降低误差。仿真试验结果表明,该算法融合结果比算术平均法和单一自适应加权算法更接近真实值。该算法能够减少数据传输量,且具有很好的可拓展性。

**关键词:**农业物联网;格拉布斯准则;分层数据融合;自适应加权;数据融合

**中图分类号:** TP393.02;S126 **文献标志码:** A **文章编号:** 1002-1302(2017)21-0241-03

农业物联网有力地推动了信息化和农业现代化的融合,对精细农业的应用和发展起到了重要作用<sup>[1]</sup>。农作物生长环境变化缓慢、实时性要求低,应用环境更为恶劣,因而农业物联网的应用特点和环境与工业物联网有本质区别<sup>[2]</sup>,应根据农业生产环境特点应用物联网技术。农业物联网中信息传输是一个至关重要的环节,农业信息传输广泛采用无线传感器技术<sup>[3]</sup>。农业物联网具有监测面积大、数据传输距离远等特点,因此信息传输要充分考虑到农业物联网的特殊性。

相比工业物联网,农业物联网要求传感器节点拥有更长的使用寿命,但是无线传感器网络中传感器节点大都用电池供电且连续工作寿命有限。减少数据传输量和数据处理量是延长无线传感器节点寿命的有效途径。如果将不经处理的数据直接发送给上位机处理,不仅浪费能量,缩短传感网络整体寿命,而且不利于数据利用,甚至会造成网络拥塞和瘫痪。

## 1 无线传感器网络数据融合技术

数据融合技术可以有效去除冗余信息,提高数据准确度和精度,节省传感器节点能量,达到延长网络寿命的目的<sup>[4]</sup>,因此能够有效地解决以上问题。无线传感器网络数据融合技术是众多科技工作者的研究热点<sup>[5-7]</sup>。刘凯等为提高多传感器检测系统预警精度,提出了在数据层、特征层、决策层分别进行数据融合的多传感器分层数据融合模型<sup>[5]</sup>。王华东等提出单个传感器数据融合,再进行传感器间数据加权自适应融合的方法,但在去除粗大误差时计算方法较为复杂,且传感器间数据融合全部在同一节点完成,使该节点负担较重<sup>[6]</sup>。Chen 等提出了一种基于簇的数据自适应融合方法,实现数据在空间和时间上的自适应融合<sup>[7]</sup>。无线传感器节点和路由

器的能量有限且处理能力较弱,因而数据融合算法应当尽量简洁,易于实现<sup>[8]</sup>。自适应加权数据融合算法无需传感器系统先验知识,依靠传感器采集的测量数据即可实现较高精度的融合估计<sup>[9]</sup>,且计算方法简便。

结合上述研究及农作物生长监测的特点,本研究借鉴分簇路由协议的思想及其周期性采集数据的特点,提出一种适用于农作物生长监测的农业物联网数据融合的架构及算法。该算法虽然增加了数据融合次数,但是减少了数据传输量,在 Berkeley Mote 中,1 bit 数据传输的能源消耗可以执行 800 条指令<sup>[10]</sup>。

## 2 农作物生长监测数据融合研究

农作物生长监测具有环境变化缓慢、实时性要求不高的特点。根据经验,北方温室温度每 10 min 变化 1℃,且温室植物对温湿度变化敏感度不高<sup>[11]</sup>,其他温室或农田环境也存在类似情况。据此本研究提出的分层自适应加权数据融合架构如图 1 所示,其中基站下设  $p$  个子层节点,每个子层节点下设  $m$  个普通成员节点,每个普通节点单位时间内采集  $k$  个数据。图 2 为本研究提出的算法流程示意图,首先对单位时间内单个传感器节点采集的数据进行最优估计,然后子层节点对该层内传感器节点估计值进行自适应加权融合,最后基站将子层节点融合值再次进行自适应加权融合,获得单位时间内的测量数据最优融合值。

### 2.1 初始测量数据预处理

估计算法是建立在可靠的测量初值基础上的,在单个节点的多次测量中,不能保证每次测量值都是正确的,因而首先需要去除粗大误差。粗大误差(又称疏失误差)是指在测量过程中,偶尔产生的某些不应有的反常因素造成的测量数值超出正常测量误差范围的小概率误差。无线传感器网络进行数据采集时,受外界干扰不可避免地出现粗大误差。含有粗大误差的数据会干扰试验结果,甚至歪曲试验结论。

剔除测量数据中的粗大误差,可以提高数据融合的自适应速度和精确度。常用剔除粗大误差的方法有分布图法、拉伊达法则、格拉布斯准则等。分布图法借助中位数、四分位数等剔除粗大误差,但过程较为繁琐。测量次数趋于无穷大时才可使用拉依达准则<sup>[12]</sup>,测量次数有限时拉依达准则就不很

收稿日期:2016-05-23

基金项目:贵州省自然科学基金(编号:黔科合 J 字[2015]2043 号);

贵州省重大基础研究项目(编号:黔科合 JZ 字[2014]2001)。

作者简介:黄海松(1977—),女,贵州大方人,教授,博士生导师,主要从事制造业信息化、物联网技术相关研究。E-mail:huang\_h\_s@126.com。

通信作者:秦志远,硕士研究生,主要从事物联网技术、数据融合相关研究。E-mail:Zhiyuan\_Qin1992@163.com。

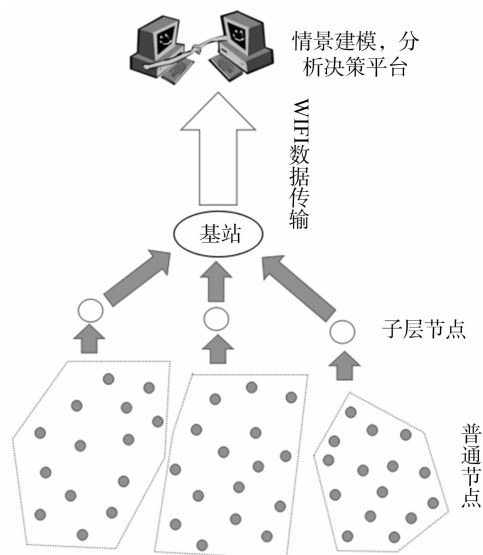


图1 数据分层自适应加权融合架构示意

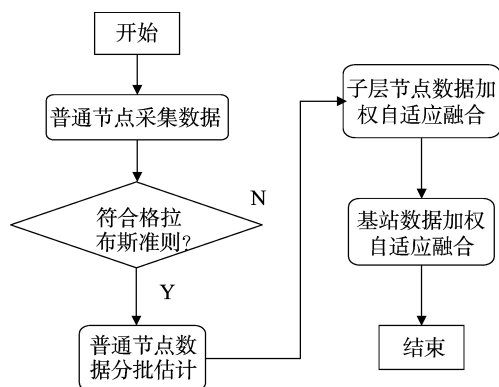


图2 数据分层自适应加权融合算法流程示意

可靠。实践证明,在检验数量较少的数据时,格拉布斯准则剔除疏失误差的准确性较高,为保证实时性,单位时间采集数据次数不能过多,因此本研究选用格拉布斯准则剔除粗大误差。

设某节点单位时间内的测量数据分别为  $x_1, x_2, x_3, \dots, x_i, \dots, x_k$  则此节点测量数据算术平均值:

$$\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i。$$

第  $i$  个测量值的残余误差为:

$$V_i = x_i - \bar{x}。$$

对应的该组数据标准差为:

$$\sigma = \sqrt{\frac{1}{k-1} \sum_{i=1}^k V_i^2}。$$

在表 1 中可查出格拉布斯统计量的临界值  $g_0(n, \alpha)$ , 即  $p[g \geq g_0(n, \alpha)] = \alpha$  (显著水平  $\alpha$  一般取 0.05 或 0.01, 即置信度为 95% 或 99%), 为提高测量精度, 取置信度为 95%。

若第  $i$  个测量数据  $x_i$  满足如下条件, 即

$$|V_i| \geq g_0(n, \alpha) \sigma,$$

则将  $x_i$  剔除。

将剩余数据重复以上过程, 直到所有数据满足要求为止。

## 2.2 单个节点数据分批估计

对单个传感器节点单位时间内采集的多个数据进行分批

表 1 部分常用格拉布斯准则临界值

$n$	$\alpha$		$n$	$\alpha$	
	0.05	0.01		0.05	0.01
4	1.463	1.492	14	2.371	2.659
5	1.672	1.749	15	2.409	2.705
6	1.822	1.994	16	2.443	2.747
7	1.938	2.097	17	2.475	2.785
8	2.032	2.231	18	2.504	2.821
9	2.110	2.323	19	2.532	2.854
10	2.176	2.410	20	2.557	2.884
11	2.234	2.485	21	2.580	2.912
12	2.285	2.550	22	2.603	2.939
13	2.331	2.607	23	2.624	2.963

估计, 减少数据发送次数, 节省节点能量, 提高数据精确度。对于单个节点数据去除疏失误差后, 将测量数据分为  $n$  组。第  $j$  组数据可分别表示为  $x_{j1}, x_{j2}, x_{j3}, x_{j4}, \dots, x_{jn}, 4 \leq n_j \leq 7$ , 即每组至少有 4 个数据, 至多有 7 个, 此处考虑计算量与精准性的平衡。

第  $j$  组平均值为:

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ji}。$$

对应的方差为:

$$\sigma_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2。$$

单一节点数据同属于一批测量数据, 可认为近似服从正态分布, 由分批估计理论可得到第  $i$  个节点的最优融合方差<sup>[13]</sup>为:

$$\sigma_i^2 = \left( \sum_{j=1}^n \frac{1}{\sigma_j^2} \right)^{-1}, i = 1, 2, \dots, m。$$

其中  $\sigma_i^2$  越小表明该节点测量数据融合之后的精度越高<sup>[14]</sup>。

由各组方差和平均值能够计算得知第  $i$  个节点融合值为

$$x_i = \left( \sum_{j=1}^n \frac{1}{\sigma_j^2} \right)^{-1} \sum_{j=1}^n \frac{1}{\sigma_j^2} \bar{x}_j, i = 1, 2, \dots, m。$$

假设将测量数据分为 2 组, 则该节点的方差可估计为

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}。$$

该节点数据的估计值为

$$x = \frac{\sigma_1^2 x_2 + \sigma_2^2 x_1}{\sigma_1^2 + \sigma_2^2}。$$

## 2.3 子层节点数据加权自适应融合

经过上述步骤, 得到单个节点的测量数据的估计值。成员节点将数据估计值发送给相应的子层节点, 由子层节点融合后转发给基站, 减少了数据传输距离, 同时进一步提高了数据准确度。设每个子层有  $m$  个节点, 单个节点数据融合值记为  $x_i$ , 方差为  $\sigma_i^2$ , 对每个子层内节点数据进行层内自适应加权融合。根据权值最优分配原则计算各节点估计值在组内的最优权值  $W_i$ , 然后对  $x_i$  自适应加权融合处理, 计算得到子层节点数据融合最优值和对应的方差。

设某子层内第  $i$  个节点的权值为

$$W_i = \left( \sigma_i^2 \sum_{i=1}^m \frac{1}{\sigma_i^2} \right)^{-1} \text{ 且 } \sum_{i=1}^m W_i = 1。$$

第  $q$  个子层节点融合值和方差分别为

$$Y_q = \sum_{i=1}^m W_i x_i, q = 1, 2, \dots, p;$$

$$\sigma_q = \sum_{i=1}^m W_i \sigma_i, q = 1, 2, \cdots, p。$$

依据上述计算过程,可以得到各个子层节点的融合值和方差,为下一步在基站位置数据融合提供数据支持。

2.4 基站级数据加权自适应融合

基站接收各个子层节点数据的融合值和方差,再次融合数据,降低误差。设每个基站下设  $p$  个子层,基站级数据融合过程与子层融合过程基本相同。由上个步骤得到的子层数据方差  $\sigma_q$ ,可计算获知各个子层的自适应加权因子  $W_q$ ,结合子层数据融合值  $Y_q$ ,最终可求得该时段内测量数据最优值  $Z$ 。

第  $q$  个子层自适应加权因子为

$$W_q = \left( \sigma_q^2 \sum_{q=1}^p \frac{1}{\sigma_q^2} \right) \text{ 且 } \sum_{q=1}^p W_q = 1。$$

可得基站层级最终融合值为

$$Z = \sum_{q=1}^p W_q Y_q。$$

3 改进算法测试分析

为验证本研究算法的有效性,在实验室内选定 1 个监测区域,由 15 个温度传感器节点和 3 个子层路由节点构成无线传感网络,温度传感器节点编号 1~15,每个节点单次试验采集 10 次数据,在不同温度下进行 5 次试验,基准真值分别为 13.5℃、17.2℃、20.8℃、23.7℃、25.0℃。3 个子层节点标号为 1~3,每个子层节点连接 5 个普通节点,分别对应 1~5 号、6~10 号、11~15 号传感器节点。

3.1 试验数据处理

以第 5 次试验数据为例展示试验数据处理过程,表 2~表 4 分别为 1~5 号、6~10 号、11~15 号温度传感器节点融合值、方差和权值,表 5 为 1~3 号子层测量数据融合值、方差和权值。依据上面公式进行计算,得出第 5 次试验的最终数据融合值为 25.24℃。

表 2 1~5 号传感器相关数据

标号	1	2	3	4	5
融合值	25.554 3	24.613 1	24.241 1	25.055 3	25.789 1
方差	0.378 9	0.575 6	0.601 5	0.417 6	0.431 4
权值	0.245 5	0.161 6	0.154 6	0.222 7	0.215 6

表 3 6~10 号传感器相关数据

标号	6	7	8	9	10
融合值	25.743 9	25.656 4	25.023 4	25.330 8	24.721 1
方差	0.149 1	0.207 1	0.812 7	1.052 2	0.371 3
权值	0.408 7	0.294 3	0.075 0	0.057 9	0.164 1

表 4 11~15 号传感器相关数据

标号	11	12	13	14	15
融合值	25.048 1	24.101 8	24.941 1	24.490 1	25.366 6
方差	0.972 9	1.175	0.608 4	0.530 6	0.407 8
权值	0.130 8	0.108 3	0.209 1	0.239 8	0.312 0

表 5 子层节点相关数据

标号	1	2	3
融合值	25.137 0	25.472 4	24.888 8
方差	0.465 0	0.304 7	0.636 2
权值	0.307 0	0.468 6	0.224 4

3.2 对比与分析

分别使用本研究算法、算术平均法和自适应加权数据融合方法对试验数据进行处理,结果见表 6。为直观展示 3 种计算方法融合值与真值近似程度,将真值标准化为零,3 种计算方法得到的融合值以相对差值的形式在图 3 中表示。在第 5 次试验中,算术平均值比自适应加权融合值更接近基准真值,但此种“接近”有较大不确定性,其余试验中自适应加权融合结果比算术平均法结果更接近真实值。总体来看,自适应加权融合方法优于算术平均法,本研究方法优于分层自适应加权融合方法且更稳定,主要原因是分层自适应加权融合方法最初没有剔除粗大误差,导致计算偏差较大。

表 6 3 种算法数据融合值与真值对比

试验标号	1	2	3	4	5
真值(℃)	13.5	17.2	20.8	23.7	25.0
本研究算法(℃)	13.41	17.34	20.93	23.78	25.24
算术平均法(℃)	13.75	16.84	20.38	23.92	24.84
自适应加权(℃)	13.62	16.97	21.01	23.27	25.38

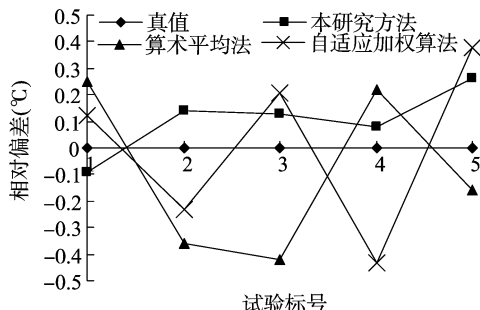


图3 3 种算法融合值与真值相对偏差比较

4 结束语

本研究结合农作物生长监测环境特点提出的改进分层自适应加权融合架构与算法,将无线传感器网络划分成多个固定层级,周期性采集数据并融合,得到被测数据单位时间内的最优融合值。首先对单个传感器节点数据进行估计,然后对每一层分别进行自适应加权融合。该算法相比以往算法融合精度高,稳定性好,简单易行,在满足数据融合实时性要求的同时,提高了测量数据精准度,虽然增加了计算量,但是减少了数据传输量,理论上可以节省无线传感器节点的能量,延长无线传感器网络的寿命。

该方法具有较好的可拓展性,可根据无线传感网络规模大小调节网络分层,适当增加或减少网络层数。将该算法用于其他类型数据融合时,如湿度、CO<sub>2</sub> 含量、光照度等,可根据被测量的实时性要求和精确度要求,设定传感器采集数据周期以及周期时间内的数据采集量。下一步的研究重点是在融合多类型数据,为研究基于农业物联网的农作物生长情景建模奠定基础。

参考文献:

[1]秦怀斌,李道亮,郭理. 农业物联网的发展及关键技术应用进展[J]. 农机化研究,2014(4):246-248.

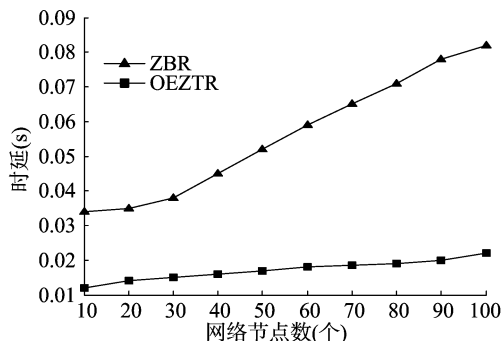


图8 平均节点时延对比图示

#### 4 系统测试及分析

为了验证本研究提出的 ZigBee 路由优化算法(OEZTR 算法)的可行性,将改进的路由算法应用到温室大棚在线监测系统中。ZigBee 节点的组网方式为树型网络,远程监测中心上位机软件采用 VB.NET 语言编写而成,系统的测试界面如图 9 所示。系统各个节点(终端节点、路由节点及协调器节点)均采用电池供电,通过设置对比测试试验,即 ZigBee 节点采用经典路由算法与节点采用 OEZTR 算法,结果表明,节点的电池寿命在同等条件下可延长 5 d 左右。



图9 远程监测中心上位机显示界面

#### 5 结束语

为了实现温室大棚的信息化和智能化,本研究设计了一种基于 ZigBee 技术的温室大棚远程监测系统,利用短距离、低功耗的 ZigBee 技术,利用机会路由算法改进的 ZigBee 路由协议,实现对温室大棚的各种环境参数(如温湿度、光照度、二氧化碳浓度等环境参数)进行实时监控和智能控制,使温室大棚反季节农作物处于最佳的生长环境,有利于反季节农作物的优质高产。下一阶段的工作是尽可能地优化 ZigBee 路由协议,以更好地满足温室大棚远程监测系统对实时性要求高的特点。

#### 参考文献:

- [1] 滕志军,李国强,王中宝,等. 基于 ZigBee 的温室大棚远程监测系统[J]. 农机化研究,2012,34(4):148-151.
- [2] Zhu Y L, Song J J, Dong F Z. Applications of wireless sensor network in the agriculture environment monitoring[J]. International Workshop on Automobile, 2011(16):608-614.
- [3] Su T M, Chen P. Intelligent energy monitoring system based on ZigBee-Equipped smart sockets[J]. Department of Electronic and Computer Engineering, 2011(23):73-81.
- [4] 李建坡,钟鑫鑫,徐 纯. 无线传感器网络静态节点定位算法综述[J]. 东北电力大学学报,2015,35(2):73-82.
- [5] 张 猛,房俊龙,韩 雨. 基于 ZigBee 和 Internet 的温室群环境远程监控系统设计[J]. 农业工程学报,2013,29(增刊1):171-176.
- [6] 杨 玮,吕 科,张 栋,等. 基于 ZigBee 技术的温室无线智能控制终端开发[J]. 农业工程学报,2010,26(3):198-202.
- [7] 陈智伟,苏维均,于重重,等. 基于 WSNs 的农业温室监控系统的设计[J]. 传感器与微系统,2011,30(7):82-84,87.
- [8] 孙增友,周 池. 基于能量和距离的 WSN 自适应分簇算法[J]. 东北电力大学学报,2016,36(1):82-86.
- [9] 蔡 镔,毕庆生,李福超,等. 基于 ZigBee 无线传感器网络的农业环境监测系统研究与设计[J]. 江西农业学报,2010,22(11):153-156.
- [10] 纪建伟,左仲善,邹秋滢. 基于 ZigBee 与嵌入式技术的农业远程监测系统的设计[J]. 中国农机化,2010(6):83-86,94.
- [11] 融合算法[J]. 农业工程学报,2012,28(23):160-166.
- [9] 高峰,俞 立,王 涌,等. 无线传感器网络作物水分状况监测系统的上位机软件开发[J]. 农业工程学报,2010,26(5):175-181.
- [10] 邱 爽,吴 巍. 无线传感器网络数据融合算法研究[J]. 武汉理工大学学报,2008,30(7):119-122.
- [11] 陈春玲,崔 琳,许童羽,等. 日光温室无线传感器多数据融合技术研究[J]. 沈阳农业大学学报,2016,47(1):86-91.
- [12] 樊雷松,强 彦,赵涓涓,等. 无线传感网中基于 BP 神经网络的数据融合方法[J]. 计算机工程与设计,2014,35(1):62-66.
- [13] 张立勇,张 利,李 丹. 基于最优分组原则的多传感器分组加权融合算法[J]. 仪器仪表学报,2008,29(1):200-205.
- [14] 蒋 婷,滕召胜,顾红艳,等. 基于 EMD 与分批估计的动态称量快速融合方法[J]. 仪器仪表学报,2015,36(6):1406-1414.
- [2] 常 超. 精准农业中 WSN 渐进融合算法研究[D]. 重庆:重庆大学,2012.
- [3] 同诗扬,祁广云,衣淑娟. 基于农业物联网的改进型双门限频谱检测算法研究[J]. 农机化研究,2016(1):53-56.
- [4] 陈正宇,杨 庚,陈 蕾,等. 无线传感器网络数据融合研究综述[J]. 计算机应用研究,2011,28(5):1601-1604.
- [5] 刘 凯,郭勇义,吴世跃. 煤矿安全监测多传感器分层数据融合模型研究[J]. 工矿自动化,2014,40(6):45-50.
- [6] 王华东,王大羽. 一种改进的多无线传感器数据分批估计自适应加权融合算法[J]. 传感技术学报,2015,28(8):1239-1243.
- [7] Chen H, Mineno H, Mizuno T. Adaptive data aggregation scheme in clustered wireless sensor networks[J]. Computer Communications, 2008,31(15):3579-3585.
- [8] 熊迎军,沈明霞,陆明洲,等. 温室无线传感器网络系统实时数据

(上接第 243 页)