

基于物联网的农业自动化框架研究

山东省滕州市南沙河镇人民政府 任德志

摘要: 智慧农业能够有效地将现代科学技术与农业种植相结合,从而实现无人化、自动化、智能化管理。本文结合物联网技术对用于农作物和土壤分布式监测的农业自动化系统进行研究,并根据雾计算的设计原则,提出层次化的数据处理架构,详细描述了通过雾计算来利用本地节点计算资源的方法,并通过作物和土壤监测的实际现场数据对提出的框架进行了验证。

关键词: 物联网; 智慧农业; 雾计算

物联网是指通过互联网连接,物与物之间可以相互作用的共享网络。这使我们能够以前所未有的规模收集真实世界的的数据,并利用这些数据改进模型,这些模型能够帮助我们对技术、社会或环境过程的理解和预测。因此,需要新的数据处理基础设施来存储和检索以在线方式收集的信息,同时提供机制来运行基于这些数据的分析和控制算法。目前,除了传统的环境监测作为物联网设计的初始关键驱动之外,还包括智慧城市、工业和农业。

一、雾计算方法概述

雾计算就是本地化的云计算,是云计算的补充,其核心思想是指分散在网络中的边缘计算节点的网络化,并利用嵌入式设备上有限的计算资源来减少收集的大量数据,只向上游传输更高级别的信息段。考虑到巨大的异构性,边缘节点可以运行的处理基元范围从基本的阈值检测到更高级的异常检测和嵌入式学习算法。无线传感器网络(WSN)是一种能够部署雾计算系统的技术,其中成百上千的感知节点通过低功耗的无线信道自组织、引入和通信。与农业的情况一样,多跳通信网络可以覆盖大片区域,因为网络协议依靠集群头、网关和集线器作为中间的数据集中器。另一种定义是与传统的集中式和大规模云基础设施相反或互补的雾系统。云平台的复杂功能在领域层面上被分解为功能或空间上分布的实体,这些实体相互协作,以实现共同的监控、事件检测和控制。在精准农业用例中,这可以帮助实现优化的分布式灌溉或肥料用量方案,考虑当地属性和土壤的差异,微气候和作物的特殊性。在这种特定的场景中,需要将雾计算与云计算集成,这是因为当跨多个农场联合高级信息时,可以得出联合观测。

在文献《一种农业管理方法雾计算平台及系统》中,作者介绍了一种面向精准农业的雾计算框架。两层系统能够显著减少网络传输的数据量。减少计算负荷,最重要的是,与集中式处理相关的云计算成本被强调为雾方法的基本好处。在文献《基于6LoWPAN无线传感器网络的农业环境实时监控系统中》,作者提出了一种用于农村智能农业的混合物联网。通信网络采用6LoWPAN本地无线电作为现场接口,通过WiFi实现远程连接。采用6LoWPAN边界路由器和专用网关,从

现场级、中远距中继和云端保证网络的跨域集成,讨论了智慧农业应用对网络的吞吐量、延迟和移动性支持等方面的需求。这为结合传感和控制需求量化数据聚合潜力提供了很好的参考。在文献《聚类算法农业信息采集系统设计及试验研究》中提出了一种分布式计算体系结构,该体系结构将农业系统的基本组成部分如:作物、土壤、气候、水分和养分、能量等有机地结合在一起。消息系统围绕消息队列遥测传输(MQTT)进行标准化,以将传感器、执行器、通信节点、设备和子系统连接起来,设计了一种用于灌溉控制的决策树,并将其集成在边缘设备上进行现场决策。在顶层,云服务通过终端用户仪表盘提供数据,以提供高层决策支持。

二、基于雾计算的农业自动化系统架构

为了在精准农业中高效地收集和处理数据,我们设计了一个系统架构,如图1所示,其算法流程如图2所示。它由如下信息层和物理层组成:现场层、雾计算层、云计算层、数据显示层,通过跨层的上下游数据流和控制信息流连接。

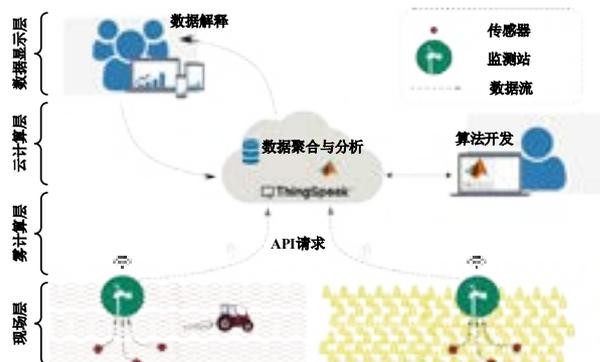


图1 基于雾计算的农业自动化系统架构

(一) 现场层

包括部署在精准农业应用中的实际传感器,以测量感兴趣的物理参数,这些参数包括空气温度、空气湿度、太阳辐射、不同深度的土壤温度、风速和降雨量;现场层还可以扩展以容纳智能执行器,例如用于灌溉或细粒度的养分剂量,以执行来自更高层系统的命令。

(二) 雾计算层

雾节点从传感器收集数据,并运行数据处理基元进行智能聚合,以减少网络流量和能源支出,主要想法是

在本地得出特定过程的基本模型特征，并以紧凑的形式发送到云端；在这一层也可以利用感知变量之间的相关性进行本地决策，从而完全避免了上层的成本和延迟的增加。

(三) 云计算层

数据流向一个共同的云平台；关于具体的实施，我们使用ThingSpeak平台，结合Matlab算法开发更高级别的处理程序；在云层，模型参数允许在需要时重建时间序列特征，同时考虑到固有的建模误差。

(四) 数据显示层

涉及前端软件系统，向终端用户或决策者展示数据分析的结果，在事件检测的情况下能够提供移动访问和及时警报，领域专家对过程的参数化也是在这一层实现的。

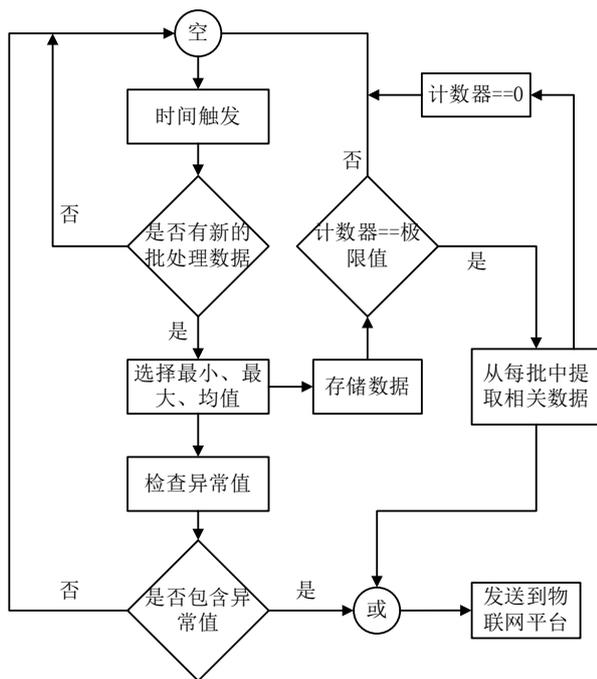


图2 雾计算算法流程图

根据信息的类型，现场测量结果以两种方式上传到物联网应用程序：事件和测量。请注意，大多数监测系统通常有一个主要的批处理程序，基本上包括在特定时间段内执行最小、最大和平均值。注意，批处理过程通常可用于大多数监控系统。通常认为这是进一步进行本地数据处理的起点。例如，批处理在30min内定义。一旦新的数据可用，则计算最小值、最大值和平均值。

对于每一批测量，都要执行异常值检查程序，考虑平均值周围测量值的数据方差的接受带宽。如果最小值或最大值超过阈值，该程序会输出一个事件。事件 E 被定义为：

$$E = \{e(x_i) \in Q, T_{\min} < x_i < T_{\max}\} \quad (1)$$

其中， x_i 表示第 i 次迭代的测量值， T_{\min} 和 T_{\max} 为阈值，其计算公式分别为：

$$T_{\min} = \text{mean}(1-w) \quad (2)$$

$$T_{\max} = \text{mean}(1+w) \quad (3)$$

其中， w 是接受带宽大小定义的权重。

聚合的数据集是基于不同的方法实现的。这些方法都是为了寻找相关的数据点，目的是减少数据集的大小，同时提供一个令人满意的初始数据的重建。就数据量而言，一种有效的方法是基于使用24h计算的最小和最大值提取。很明显，这种方法只适用于在一段时间内有规律的测量，在一天内没有明显的变化。这种方法适合的测量是土壤温度。相反，变化检测是一种适用于不规则形状的数据集的常用方法。这种方法是对发生趋势变化的数据点进行提取。

给定一组数据点 (x_i, y_i) , $i=1, \dots, n$ ，对每一对 x_i, x_{i+1} 的趋势 t_i 进行跟踪，我们可得出：

$$x_{i+1} - x_i > \delta \Rightarrow t(i) = 1 \quad (4)$$

$$x_{i+1} - x_i < \delta \Rightarrow t(i) = -1 \quad (5)$$

$$x_{i+1} = x_i \Rightarrow t(i) = 0 \quad (6)$$

如果 $t(i) \neq t(i+1)$ ，则意味着检测到了一个趋势变化。相应的数据点 x_{i+1} 被添加到相关数据集中。

三、数据聚合及实验验证

从传感器数据中提取高级信息的一种参考方法是符号聚合近似 (Symbolic aggregate approximation, SAX)。它通过将标签符号分配给时间序列的片段来进行操作，从而将其移植到统一的低维表示中。它属于导致非参数建模的时间序列数据挖掘技术家族。范围通过数据直方图或以统一方式识别。该方法提供了线性复杂性，并开辟了多种统计学习工具的使用和应用。通过定义段数和字母大小，SAX参数化非常重要，这会影响到结果的稳健性。

PAA确定了SAX定义的背景，其中符号属于PAA列出的聚合数值。若干离散事件模型可以合并所得到的聚合片段，例如马尔可夫模型，以便计算用于未来观察的观察模式的概率。根据PAA方法描述，从长度为 n 的时间序列 X 开始，将其近似为任意长度 $M \leq n$ 的向量 $\bar{X} = (\bar{x}_1, \dots, \bar{x}_M)$ ，其中 n 可被 M 整除。向量 \bar{x}_i 的每个元素由下式计算：

$$\bar{x}_i = \frac{M}{n} \sum_{j=n/M(i-1)+1}^{(n/M)i} x_j \quad (7)$$

首先将原始数据分成 M 个大小相等的帧，然后计算每个帧的均值，从而将时间序列的维数从 n 降为 M 个样本。通过将均值放在一起得到一个新的序列，这被认为是原始数据的PAA变换 (近似)。在计算复杂度方面，该方法可以将PAA变换的复杂度从 $O(NM)$ 降低到 $O(Mm)$ ，其中 m 为帧数作为调谐参数。两个近似的时间序列向量 \bar{X} 和 \bar{Y} 之间的距离度量定义为：

$$D_{PAA}(\bar{X}, \bar{Y}) = \sqrt{\frac{n}{M}} \sqrt{\sum_{i=1}^M (\bar{x}_i - \bar{y}_i)^2} \quad (8)$$

通过分析可知, PAA满足下界条件, 如公式(9)所示, 并能够确保不存在误判

$$D_{PAA}(\bar{X}, \bar{Y}) \leq D(X, Y) \quad (9)$$

此外, 基于云的应用程序通过基于插值机制的估计重建数据集。通常可采用三种插值方法: 常见的线性插值(也被称为分片线性插值)和两种密切相关的插值, 即三次样条和分段三次Hermite插值多项式(pchip)。

一般来说, 给定n个点 (x_i, y_i) , $i=1, \dots, n$, 其图形可通过n个点表示为 $P_n(x)$, 用拉格朗日形式表示为

$$P_n(x) = \sum_{i=1}^n \left(\prod_{j=1, j \neq i}^n \frac{x-x_j}{x_i-x_j} \right) y_i \quad (10)$$

公式(7)中可以用插值多项式的幂形式表示为:

$$P_n(x) = a_0 x^{n-1} + a_1 x^{n-2} + \dots + a_{n-1} x + a_n \quad (11)$$

其中 a_k 系数通过线性方程组计算:

$$\begin{bmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (12)$$

考虑到这一点, 通过首先计算分割差分生成分段线性插值

$$\delta_i := \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (13)$$

然后将插值函数构造为:

$$P(x) = y_i + \delta_i(x - x_i) \quad (14)$$

为了对提出的框架进行验证, 我们从一个实验研究农场安装的现场设备网络中收集实验数据。在长期监测数据集中, 选择一个涵盖一个月数据的样本进行分析。首先, 对数据进行预处理, 以消除缺失值和噪音, 并在30分钟的时间间隔内取平均值。

为了评估重建数据的一致性, 通过不同的估计算法(即不同的插值法), 进行拟合度统计。本文选择的统计指标为均方根误差(Rootmeansquareerror, RMSE), 它是对数据中随机成分的标准偏差的估计, 其统计结果如下:

表1 拟合度统计结果

	pchipRMSE	三次样条 RMSE	线性插值 RMSE
土壤温度	0.0943	0.1293	0.0876
太阳辐射度	0.2536	0.3782	0.2463

通过所研究的案例可以看出, 线性插值的拟合效果更好。在实际的应用中, 我们也可以选择pchip方法, 因为其结果相当接近, 并且它的表现更为直观, 尤其是能够更好地建模峰值, 并遵循基线周围的预期行为。

四、结束语

农业在一个国家的整体经济发展中占有重要地位。而物联网又在智能农业中发挥着至关重要的作用。本文充分结合物联网技术, 根据雾计算的设计原则, 提出包含现场层、雾计算层、云计算层、数据表示层四个部分的层次化的农业自动化系统框架。最后, 通过对实际数

据的分析, 验证了提出框架的有效性。

参考文献:

- [1] 杨永志, 高建华. 试论物联网及其在我国的科学发展[J]. 中国流通经济编辑部, 2010(2):46-49.
- [2] 杜志强. 我国工业物联网行业竞争分析及发展前景[J]. 市场调查信息: 综合版, 2022(20):3.
- [3] 李玥. 试析物联网技术在智慧城市建设中的运用[J]. 信息通信, 2022(004):235-237.
- [4] 孙海丽, 龙翔, 韩兰胜, 等. 工业物联网异常检测技术综述[J]. 通信学报, 2022(003):196-210.
- [5] 湛颖, 戴华珍, 龚文静等. 智慧农业物联网控制系统设计与研究[J]. 信息记录材料, 2022(004):195-197.
- [6] 张颖楣. 物联网精准农业问题及对策[J]. 南方农机, 2020, 51(2):81, 88.
- [7] 尹青山, 段成德. 一种农业管理方法雾计算平台及系统:CN107392398A[P]. 2017.
- [8] 王晓喃, 殷旭东. 基于6LoWPAN无线传感器网络的农业环境实时监控系统[J]. 农业工程学报, 2010(10):224-228.
- [9] 解蕾, 狄光智. 聚类算法农业信息采集系统设计及试验研究[J]. 农机化研究, 2020(9):113-116, 124.
- [10] 张豹, 应励志, 余宇峰. 基于趋势特征的时间序列符号聚集近似表示方法[J]. 计算机应用, 2022, 42(S01):123-129.
- [11] 崔利宏, 高小淞, 孟敏, 等. 构造二元四次Hermite插值公式的方法[J]. 吉林师范大学学报: 自然科学版, 2015, 36(1):63-65.