

基于 TinyOS 的多跳式传感器网络

马娅婕¹, 尹首一², 田翔川³

(1. 武汉科技大学信息科学与工程学院, 武汉 430081; 2. 清华大学微电子学研究所, 北京 100084;

3. Inforsense 软件技术有限公司, 伦敦 W6 7JP)

摘要: 针对城市环境大气污染的监测, 设计一种基于 TinyOS 的多跳式传感器网络。该网络采用层次化结构, 基于 TinyOS 操作系统, 实现传感器网络的多跳式通信。为对大气污染数据进行分析, 将一种分布式 k-mean 聚类算法应用于该传感器网络的数据挖掘。仿真实验表明, 该系统能有效地实现数据的多跳传输和分布式数据挖掘。

关键词: 传感器网络; 多跳式; TinyOS 系统; 分布式数据挖掘

Multi-hop Sensor Network Based on TinyOS

MA Ya-jie¹, YIN Shou-yi², TIAN Xiang-chuan³

(1. Information Science and Engineering College, Wuhan University of Science and Technology, Wuhan 430081;

2. Institute of Microelectronics, Tsinghua University, Beijing 100084; 3. Inforsense Ltd., London W6 7JP)

【Abstract】 A multi-hop sensor network based on TinyOS is designed. The system is based on hierarchical network architecture. The software design principle based on TinyOS is introduced, which can realize the multi-hop communication. A distributed k-mean algorithm is designed for the data mining in this sensor network to analyse the air pollution data. According to simulation results, this system can efficiently realize the multi-hop data communication and distributed data mining.

【Key words】 sensor network; multi-hop; TinyOS; distributed data mining

1 概述

传感器网络能通过各类集成化的微型传感器实时监测、感知和采集各种环境或监测对象的信息, 并通过随机自组织无线通信网络以多跳中继方式将所感知信息传送到用户终端, 因此, 传感器网络在军事国防、工农业、城市管理、生物医疗、环境监测、抢险救灾、反恐反恐、危险区域远程控制等许多领域都有重要的科研价值和实用价值^[1]。

针对城市环境大气污染的监测, 本文设计一种多跳式传感器网络, 并在 TinyOS 操作系统上实现数据采样、通信及分布式数据处理等功能, 可对大气污染成分进行采样和分析, 以了解交通工具以及人类行为对环境的影响程度, 对保护环境遏制污染具有重大的意义。该系统具有以下特点: (1) 层次化的网络结构: 包括移动节点构成的移动子网和固定节点构成的固定网格; (2) 多跳式通信模式: 固定节点采用多跳式的通信方法进行数据传输和交换; (3) 按需传输数据: 移动采样节点收到轮询信号后发送采样信息给固定节点; (4) 分布式的数据处理: 采样数据在固定网格上进行分布式的数据挖掘。

2 总体介绍

2.1 网络结构

该传感器网络的功能主要是对城市大气污染的监测, 因此, 针对城市环境的特点, 将该传感器网络设计分为 2 层结构: 由移动传感器节点(Mobile Sensor Node, MSN)组成的移动子网和由固定传感器节点 (Stationary Sensor Node, SSN)组成的网格环境, 其网络分布如图 1 所示。其中, MSN 可安装在各种移动设备(如手机、PDA 等)和交通工具上, 主要负责数据采样的功能; SSN 利用城市的建筑特征进行安装, 形成

网络结构, 实现数据的搜集和分布式处理。

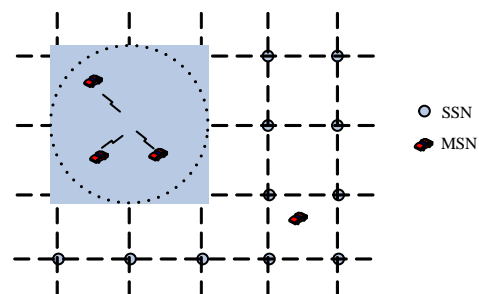


图 1 城市大气污染监测传感器网络结构

MSN 搜集到的数据经过简单的处理(如 A/D 转换、去噪等)后传输给 SSN, 在 SSN 处进行保存、更新、挖掘和分析, 并根据要求传输给中心节点(data sink)。这样设计不仅可以充分利用城市道路和建筑的现实环境, 而且有助于数据的分布式存储和处理。

2.2 基于 TinyOS 的设计原理

TinyOS 是美国加州大学伯克利分校专门针对现代无线传感器的特性与需求设计的开放源代码的嵌入式操作系统, 其特点是体积小、结构高度模块化、基于组件的架构方式、低功耗等^[2]。TinyOS 将无线传感器的功能模块化, 并用组件的形式实现和包装, 组件之间通过系统定义或用户自定义的

作者简介: 马娅婕(1974 -), 女, 副教授、博士, 主研方向: 传感器网络, 拓扑设计, 数据挖掘, 流量工程; 尹首一, 讲师、博士; 田翔川, 工程师

收稿日期: 2008-01-20 **E-mail:** yjiema@googlemail.com

接口进行联系。

根据组件的不同功能，基于 TinyOS 的传感器单元软件框架自顶向下抽象为 Main 组件(为所有程序的入口)、应用/用户组件(用户根据传感器的功能需求自定义的组件)、行为组件(例如操纵声音、信号灯等动作、传感器感知及采样控制、各种通信协议的执行等)和硬件模块抽象组件(包括传感器芯片的各种硬件驱动程序)。

本文所涉及的是用户自定义组件的功能设计，并实现与上层 Main 组件以及下层组件之间的接口链接。各种下层功能可根据项目的总体需求，选择合适的系统组件并将它们链接起来。由于涉及到大气污染数据的采样和数据的无线传输和处理，因此在设计的时候用到的基本系统组件包括时钟、采样、A/D 转换、无线通信等；硬件平台选用加州大学伯克利分校主持开发的低功耗无线传感器节点 Mica；用户自定义组件需要实现的功能则包括：

(1)MSN：设定采样周期，接收和缓存采样数据，接收 SSN 的轮询信息，发送采样数据给 SSN；

(2)SSN：发送轮询信息，从 MSN 接收采样数据，与其他 SSN 交换采样数据，执行分布式数据挖掘算法，将数据和挖掘结果发送给中心节点。

3 多跳式传感器网络的设计

3.1 多跳式路由协议

多跳路由协议选用RFC3561(Ad hoc On-demand Distance Vector (AODV) Routing)^[3]，该协议是一种按需距离向量路由协议，采用多跳式路由策略，具有实现简单、防止环路、维护方便、省电及扩展性强的特点。本设计采用系统提供的 AODV组件实现该路由协议，并利用AODV组件提供的发送、接收接口实现轮询和采样数据的收发、数据挖掘所需信息的多跳式交换等功能。

3.2 传感器单元的软件设计

(1)MSN 的组件选择及链接

传感器单元 MSN 所使用的组件、组件提供和使用的接口，以及各组件间的链接关系如图 2 所示。根据上一节所定义的系统功能要求，MSN 所需要的组件包括 Main 组件、应用组件 MSN、采样所需定时组件 TimerC、采样及 A/D 转换组件 DemoSensorC 和传输采样数据和多跳路由所需组件 AODV。

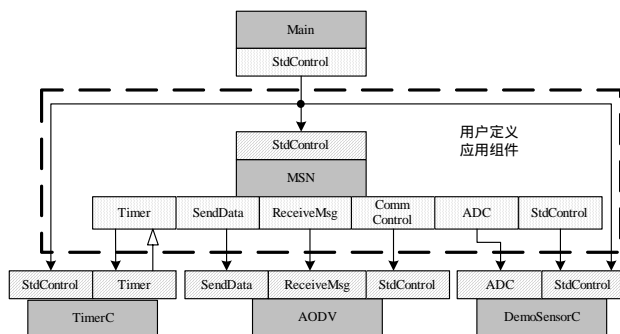


图 2 MSN 结构

其中，灰色阴影框表示组件；斜线框表示某组件为上层组件提供的接口；黑点框表示某组件与下层组件通信所使用的接口；实心箭头线表示命令的通信方向是从线的尾部到箭头方向；空心箭头线表示事件的通信方向是从线的尾部到箭头方向；虚线框中的组件及接口为用户自定义的应用组件，

而其他组件和接口则由系统提供。

(2)SSN 的组件选择及链接

传感器单元 SSN 所使用的组件、组件提供和使用的接口，以及各组件间的链接关系如图 3 所示。由图 3(各框和箭头表示的含义同图 2)可知，SSN 所需要的组件包括 Main 组件、应用组件 SSN、轮询所需定时组件 TimerC、接收采样数据和多跳路由所需组件 AODV。

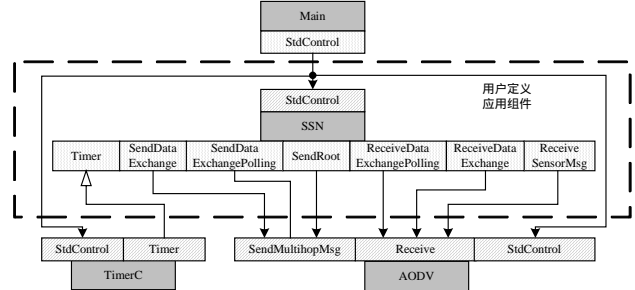


图 3 SSN 结构

(3)数据处理流程

1)数据处理流程包括：采样大气污染数据，并转换成数字信号；暂存数据；定时器到时，将采样的数据(一个或者多个)发送给 SSN。

2)SSN 对数据的处理流程包括：从 MSN 接收采样数据；每个 SSN 利用随机算法(该随机算法不在本文的讨论范围之内)选取一定数量的 SSN 作为信息交换节点集合，定时发送轮询信息给这些 SSN 节点以请求数据信息；每个 SSN 接收信息交换节点传送的数据信息；每个 SSN 执行分布式数据挖掘算法。

3.3 数据挖掘算法

考虑到该传感器网络需要对不同的大气污染源和污染等级进行数据挖掘，以确定大气污染的不同程度，因此，笔者设计一种 k-means 分布式聚类算法。算法描述如下：

(1)节点 SSN_i 利用随机算法选取一定数量的SSN作为信息交换的节点集合 S_i ，并随机选取 k 个本地数据作为 k 个初始类中心 $C_{i,j}^0$ ($j=0, 1, \dots, k-1$)；

(2)计算每个本地数据与 k 个中心欧式距离，将数据分配到距离最小的中心作为一类，并用 $(C_{i,j}^0, Num_{i,j})$ 作为本地类的描述，其中 $C_{i,j}^0$ 为SSN节点 i 的第 j 个类的类中心； $Num_{i,j}$ 为SSN节点 i 的第 j 个类的成员数量总和；

(3)节点 SSN_i 定时发送轮询信息给 S_i 中的节点以请求数据信息；

(4)某节点 SSN_n 接收到轮询信息后，将本地数据描述 $(C_{n,j}^0, Num_{n,j})$ 发送给轮询节点；

(5)若节点 SSN_i 接收到了全部数据交换节点回应的数据描述信息，则利用这些信息计算新的类中心 $C_{i,j}^1$ ；

(6)比较新的类中心 $C_{i,j}^1$ 与初始类中心 $C_{i,j}^0$ 之间的偏移量，若偏移量小于预先定义的阈值，则算法结束；否则用 $C_{i,j}^1$ 代替 $C_{i,j}^0$ ，并重复执行步骤(2)~步骤(6)。

4 仿真及结果分析

为考察多跳式路由和分布式数据挖掘的运行情况，在 OMNet++ 软件^[4]上进行仿真实验。实验主要关注路由信息、轮询信息以及数据信息在各节点间的收发状况，并获得分布式聚类算法的计算结果。实验定义 18 个节点，其拓扑如图 4 所示。在拓扑结构中，包括 12 个 SSN 节点：tic[0]~tic[11]，6 个 MSN 节点：tic[12]~tic[17]，数据在网络中可延拓双向传输。聚类所使用的测试数据集为某区域某一天 8 点~18 点

(下转第 276 页)