

参考文献：

- [1]A. Hintz. Fingerprinting Websites using traffic analysis. In R. Dingledine and P. Syverson, editors, Proceedings of Privacy Enhancing Technologies Workshop (PET 2002), volume 2482 of Lecture Notes in Computer Science, pages 171-178. Springer-Verlag, Apr. 2002.
- [2]D. Wagner and B. Schneier. Analysis of the SSL 3.0 protocol. In Proceedings of the 2nd USENIX Workshop on Electronic Commerce (EC-96), pages 29-40. USENIX Association, Nov. 1996.
- [3]M. Liberatore and B. N. Levine. Inferring the Source of Encrypted HTTP Connections. In Proceedings of the 13th ACM

Conference on Computer and Communications Security (CCS 2006), pages 255-263, Alexandria, Virginia, USA, Oct. 2006. ACM Press.

[4]D. Herrmann, R. Wendolsky, and H. Federrath. Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naïve-Bayes Classifier. In CCSW '09: Proceedings of the 2009 ACM CCS Workshop on Cloud Computing Security, pages 31-42, Chicago, Illinois, USA, Nov. 2009. ACM Press.

[5]Andriy Panchenko, Lukas Niessen, Andreas Zinnen. Website Fingerprinting in Onion Routing Based Anonymization Networks.

基于 ZigBee 的温湿度监测系统设计

靳尧凯 戴贻康 焦运良

(华北计算机系统工程研究所 北京 100083)

摘要：过去温湿度监测大多采用的是有线的连接方式，这种连接方式具有成本较高，不适宜复杂环境，系统脆弱等缺点。随着传感器技术和无线通信技术的发展，温湿度测量逐步走向无线的方向，ZigBee 技术作为一种新型的短距离无线通信技术，在无线感知传输方面具有很大的优点，本文探究了 ZigBee 中各节点的入网行为，之后分别从硬件和软件部分进行了设计。硬件方面选择 CC2530 作为主控芯片，DHT11 作为传感器，软件方面则以 IAR 软件为基础，结合德州仪器公司推出的 Z-Stack 协议栈，分别进行了不同节点的软件设计，最后进行功能的测试，结果显示设计基本符合要求。

关键词：ZigBee 技术；温湿度；节点；CC2530

在工业生产、环境监测和日常生活中，要密切关注周围环境的变化，只有适宜的环境，才能发挥最高的生产效率，创造出最大的效益。

随着电子科学技术的不断发展，使得人们对无线通信的传输质量、传输距离和快捷性的要求越来越高^[1]。人们使用 FF 基金会现场总线，CAN 总线等有线电缆的连接方式进行数据传输的监测系统有很大的局限性，例如设备位置较为固定，现场布线麻烦，不适宜复杂的环境，网络脆弱，一个节点的失效可能导致整个系统瘫痪。随着微机电系统 (MEMS) 的发展，片上系统解决方案的成熟，无线传感已经开始走入寻常百姓家，发展越来越快，应用领域越来越广阔，它的发展和运用给人们的生活和生产带来了深远的影响。

目前，短距离无线通信技术慢慢成为研究的一个热点领域，它们包括无线局域网 (Wi-Fi)，蓝牙 (Bluetooth)，红外数据通信 (IrDA)，近场通信 (NFC)，超宽带通信 (UWB)，以及 ZigBee 技术^[2]。其中 ZigBee 技术以其成本低廉，低功耗，高可靠性正广泛应用于家庭和楼宇网络，工业控制，公共场所等领域^[3]。

1 ZigBee 技术研究

1.1 ZigBee 技术特点

知 ZigBee 协议模型主要由应用层 (APL) 和网络层 (NWK)，媒介访问控制层 (MAC) 和物理层 (PHY) 组成，其中 NWK 层和 APL 层是被 ZigBee 联盟定义的，PHY 层和 MAC 层是被 IEEE802.15.4 标准定义的^[4]。ZigBee 技术有以下几个特点。

(1) 功耗低：由于 ZigBee 技术多用于偏僻或者人们难以到达的地方，因此 ZigBee 设计之初就要求了低功耗，通常来说，一个 ZigBee 节点依靠两节干电池即可工作 6~24 个月^[5]。

(2) 成本低：由于 ZigBee 协议栈简单，工作流程清晰，开发和维护的成本较低，并且 ZigBee 协议对外公开，不需要收取专利费，因此具有较低的成本，适合大规模使用。

(3) 网络容量大：一个 ZigBee 网络最高可以有 65535 个数据模块，此外，ZigBee 有 27 个信道可以用来通信，而 TI 公司所有支持 ZigBee 底层协议的芯片 CC2530 在 2.4GHz 的频段上，有 16 个信道

进行通信，这就保证了一个区域可以有多个网络的存在。

(4) 自组织能力：ZigBee 模块上电后，协调器能够自行建网，自动分配网络地址，终端和路由器上电后能够自行请求加入网络。

(5) 可靠性高：网络通信采用了避免碰撞的机制，数据收发时不会产生冲突，并且预留了一些间隙专门为了需要固定带宽的通信，每一个带有明确地址的数据包在接收方接收后都会产生一个确认信息，数据可以重新发送，这些方法保证了数据在通信过程中不会丢失和出错。

1.2 ZigBee 的网络拓扑结构

ZigBee 组网灵活多变，能够组成星型、树型、和网格型的网络拓扑结构^[6]。一般是根据现场需要决定选择何种网络结构。星型结构是 ZigBee 网络中最简略的结构。一般由一个负责创建网络维护网络的协调器和一系列终端传感器节点组成。由于每个传感器只能和协调器传递信息，因此传感器之间若想传递信息，只能通过整个网络的核心——协调器，协调器将信息转发后就可以实现传感器之间的通信。星型拓扑网络的结构如图 1 所示。

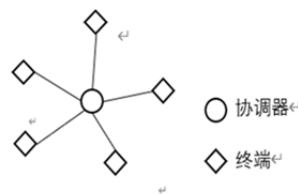


图 1 星型网络结构图

树型结构是由一个负责创建网络、维护网络的协调器，若干个数据转发的路由器和一系列采集现场数据的传感器组成的。协调器可以与路由器和传感器节点互相连接，下一级的路由器仍然可以无线连接更下一级的路由器和传感器。其结构图可表示为如图 2 所示。

树型网络的优点是可以进行中继路由，理论上经过路由，ZigBee 的传输距离可以达到很远，它的网络结构比星型结构复杂，但是也相对灵活。缺点是信息的路由只有一条通道可以选择。

网络结构：一个网络只能有一个协调器节点，因此网络结构也有一个网络核心协调器，同时还有众多的路由器和传感器节点，与树型网络的结构组成相类似，而且这种结构与树型网络结构也相似，但是节点之间的路由通道增加为多条。主要表现是路由器节点之间可以直接进行通信，由此节点与节点之间的通信便可以通过路由器作为桥梁来通信，当某个路由器节点因故障而失去作用，节点之间的通信便可以绕过该路由器节点，通过其他的路径进行通信。网格型的网络结构可以参考图 3 所示。

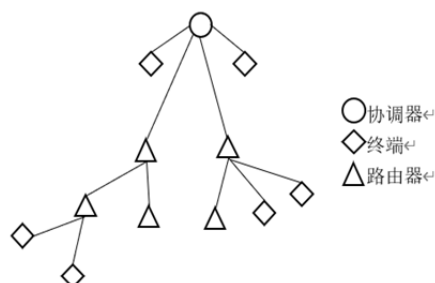


图 2 树型网络结构图

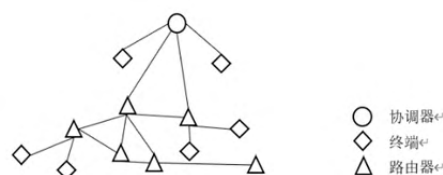


图 3 网格型网络结构图

2 系统总体设计方案

终端即为传感器，首先将各传感器节点分散放置在不同位置，传感器节点在上电工作后感知到周围的温度和湿度的电信号，并且经过放大，模拟/数字转换，经传感器节点内嵌的 8051 核的单片机读取传感器数字信息后，经过运算得出真实的温湿度值。

如果传感器已经入网，则将这些信息以无线的方式先发送给路由器，路由器再转发给协调器。协调器负责整个 ZigBee 监测系统网络的组建和维护，并且还需要接受来自路由器节点转发来的传感器感知到的信息。协调器上电后先组网，网络组建成功后，路由器和各传感器节点加入该网络中来。当传感器开始工作后，协调器开始接收路由器节点转发过来的信息，通过串口显示到 LCD 屏幕上。

2.1 系统硬件设计

2.1.1 传感器节点的硬件设计

传感器节点包括了传感器单元，电源供电单元，射频收发单元，数据处理单元^[7]。其结构图如图 4 所示。

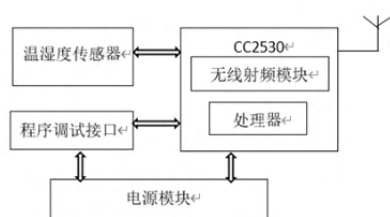


图 4 传感器节点硬件结构图

传感器模块用来感知周围的温湿度信息，并将采集到的温湿度的模拟电信号经过模拟/数字转换成可供 CC2530 读取的数字信号。CC2530 中的处理器单元负责控制处理信息的存储，I/O 口，控制射频模块的发送等，电源模块负责给整个节点供电。

本设计需要采集周围环境的温度和湿度信息，采用的国内奥松公司推出的集温度采集和湿度采集于一体的温湿度传感器 DHT11。

ZigBee 芯片具有自己的协议栈，我们不需要知道协议栈里具体程序的运行流程，只需将传感器采集到的信息，通过串口发送到 ZigBee 芯片。ZigBee 芯片能够将这些信息经过运算处理后以无线的

方式发送出去，接收器按照某种配置，能够实时、无误的接受这些数据或者命令。因此，要求该 ZigBee 芯片具有射频，数据处理，同时该芯片业必须支持 ZigBee 底层协议。选择 TI 公司推出的 CC2530 芯片作为该设计的主控芯片。CC2530 芯片成本较低，内置增强型的 8051 单片机，具有 8Kb 的 RAM，可在线下载，支持 Z-Stack 协议栈。CC2530 运行功耗较低，能够在一块干电池的供电下持续运行 6 个月以上。

2.1.2 协调器节点的硬件设计

协调器节点负责组建网络，维护网络，接收路由器转发来的传感器节点的数据，是整个网络的核心。本质上来说，终端节点，协调器节点，路由器节点的硬件部分一样，只是终端外接了一个传感器只需要下载不同功能的代码就可以实现不同的功能。因此协调器节点的硬件电路与传感器节点的硬件电路相同，只是少了传感器，这里不再赘述。

2.1.3 路由器节点的硬件设计

如果测量的范围过大，且障碍物过多，那么协调器与终端节点的数据发送可能会中断，因此设置一个路由器节点负责数据的转发。由于路由器节点也没有传感器，因此该节点与协调器节点的硬件相同，下载对应的代码即可实现路由功能。

2.2 系统软件设计

2.2.1 Z-Stack 协议栈

TI 公司除了推出其公司的 ZigBee 芯片，也公开了很完整实用的协议栈供开发者下载使用。ZigBee 协议的内层分为不同的两大部分^[8]，一部分是由 IEEE802.15.4 标准规范的物理层和媒体访问控制层，另一部分则是由 ZigBee 联盟规范制定的网络层和应用层，Z-Stack 就是将这些规范打包整理，并且用 C 语言呈现出来。

Z-Stack 相当于一个简易的操作系统，在协议栈中的文件夹 Zmain.c 是它的主函数，负责整个程序大框架的运行，它有两个重要的功能，一个是进行 ZigBee 的初始化操作，另一个是执行该操作系统，保证 ZigBee 的正常工作。

2.2.2 传感器（终端）节点软件设计

终端节点主要利用传感器来感知环境收集数据，然后将温湿度信息打包发送给路由器^[9]。传感器节点首先需要初始化并且尝试加入协调器已经创建成功的网络中，然后将自己设备绑定到父节点上，就可以将终端收集到的温湿度信息以无线通信的方式传送到路由器上。

先初始化，之后开始扫描信道，查看附近是否有可以加入的 ZigBee 网络，若发现有可以加入的网络，则终端会向协调器发送请求帧请求加入网络。若协调器收到该帧并且同意其入网，则会向传感器（终端）发送一个 16 位独一无二的地址。加入网络成功后进入休眠模式，当定时时间到，读取传感器的数值，存储在数据缓冲区，并且在主控芯片的控制下打包将数据发送给路由器，发送不成功再次发送，直到成功后再次进入休眠状态，等待下一次的周期性发送。

2.2.3 协调器节点软件设计

协调器决定了当前环境网络的存在与否，所以它处于非常重要的位置。在该设计中它的软件主要实现的功能包括创建该区域的 ZigBee 网络，同时负责其他节点的入网事项，在稳定工作后还需要接收相应的数据信息，还需要将收到的数据通过串口加上屏幕驱动程序在 LCD 上显示。协调器的软件工作流程如图 5 所示。

(1) 无线传感器网络的建立和加入。开发者不需要关注协议栈的细节部分，只需要知道数据的流动过程，并且使用协议栈提供的相关函数实现我们需要的功能即可。

(2) 数据的接收。在 ZigBee 网络的正常工作中，协调器会将收到的数据包存储起来，因此只需要通过相关函数找到存储起来的数据，就可以经过相应的驱动程序显示在 LCD 屏幕上。

2.2.4 路由器节点软件设计

本设计采用树型网络拓扑结构，在中间有一级数据的中继，这一功能由路由器实现。路由器的工作流程与终端的工作流程相似，上电后先进行软件和硬件的初始化操作，然后进行信道的扫描，选择一个信道，然后进行网络的搜索，搜索到合适的网络后请求加入。

路由器的唯一功能就是进行数据的转发，下载相应的代码后，设备以路由器的身份入网后，就可以实现数据的转发。

3 系统调试与结果

硬件：协调器节点，路由器节点、终端节点、仿真器、LCD 显示屏，PC 机。软件：IAR Embedded workbench，串口调试助手。

本设计采用树型网络，由终端节点采集温湿度数据，将采集到的数据发送给路由器，经过路由器转发给协调器，协调器收到温湿度信息后驱动 LCD 显示屏显示当前温湿度，并且在测试过程中可以打开

串口调试助手查看串口上发送的具体数据。

将相应的代码经过编译，环境配置后下载到对应的模块中，待各个模块初始化完成后，便可测得相应的温湿度信息。打开串口调试助手，可以看到在串口上的温湿度信息每隔 10 秒钟刷新一下。

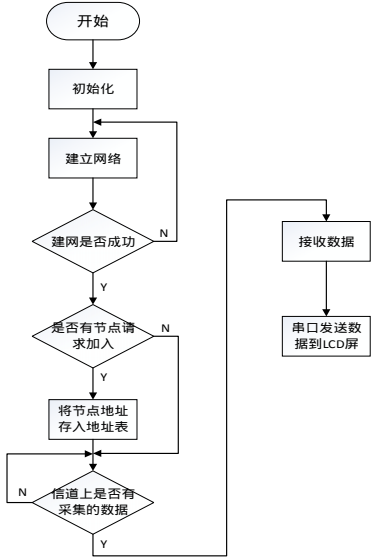


图 5 协调器软件工作流程

4 结束语

综合以上设计结果，可知，该设计基本符合要求。但本设计还存在以下的不足：本设计中只做了三个节点，完成了最基础的树型网络组建，但是往往在实际应用时，会有更多的节点参与进来。因此本设计可考虑增加传感器，增加路由器节点，设计成更为复杂的网状结构。另外本设计采用的芯片 RAM 只有 8K，若要设计更复杂的功能，可以考虑加片外 RAM 或者功能更为强大的芯片。

参考文献：

[1]陈斌.无线传感器网络的发展与应用[J].科技风, 2010, 40 (02): 190-191.
[2]陈龙.基于 LEACH 协议的无线传感器网络攻击检测技术研究[D].哈尔滨工业大学, 2009.
[3]王明福.采用功率控制的无线传感器网络 MAC 协议研究[D].合肥工业大学, 2007.
[4]刘起明. 基于 ZigBee 的智能家居环境监测系统设计与实现[D].华中师范大学, 2019.
[5]张林. 基于无线传感网的大棚数据采集系统设计[D].安徽理工大学, 2017.
[6]侯智超. ZigBee 技术在井口开采作业监测中的应用研究[D].吉林大学, 2016.
[7]高英明. 无线传感器网络能量管理技术与理论研究[D].大连理工大学, 2009.
[8]王岩. 基于 ZigBee 协议无线传感器网络的设计与实现[D].天津大学, 2007.
[9]钱蕾. 基于物联网的智能家居控制系统研究[D].华北理工大学, 2018.

基于机器学习的网站识别方案

刘天一 张汝嫻 袁艺 邢韦川 林清然 周延森 指导老师

（国际关系学院 北京 100091）

摘要：本文基于图片分类算法 NSFW 实现了色情网站识别，使用自然语言处理提取网页特征并通过 SVM 分类器训练得到博彩网站检测模型，实验表明其精度高于 95%，最后结合二者进行非法网站检测。
关键词：机器学习；非法网站检测；自然语言处理
基金项目：本文得到国际关系学院大学生学术支持计划项目资助

1 引言

近年来互联网技术发展迅速，网络资源丰富了生活。但与此同时，大量博彩、色情等非法网站的出现严重影响了网络环境，非法网站检测也成为保护网络安全的一环。

有关部门已开展大量专项整治行动，但由于网络建站成本低、活动隐蔽性强及监管难度大等特点，许多诸如色情、博彩等非法网站仍屡禁不止，对网络安全造成严重威胁，目前急需高效检测的技术手段。

非法网站检测目前主要有三类技术：黑名单、静态检测、动态检测。基于 URL 检测黑名单的技术简单且准确率较高，但灵活性较差，且人工维护黑名单耗时耗力。静态检测是目前主流，多数预先采用网络爬虫获取目标网站静态数据，通过机器学习提取特征构建分类识别模型^[2-4]，其关键是选取有效检测特征与构建适合的分类器。动态检测耗时长且难度较大，针对挂马类恶意网站识别率较高，而对于博彩、色情等网站效果欠佳。

相比已有的非法网站检测工作，我们针对不同类型非法网站采用不同手段，进而提高识别的准确性与效率。一方面，使用自然语言处理技术提取网页特征并利用 SVM 生成博彩网站识别与分类模型。另一方面，在已有 NSFW 色情图片识别平台基础上实现色情网站检测

模型。

2 非法网站系统构建

本文所设计的非法网站检测系统框架如图 1，输入指定 URL 后系统依次对其进行是否为博彩、色情网站的识别，从而对其是否为非法网站作出初步判断。

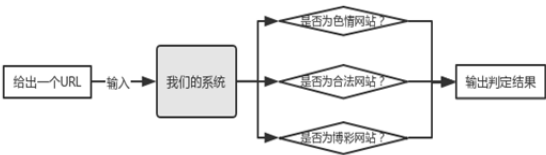


图 1 非法网站识别系统

2.1 基于 SVM 的博彩网站检测

博彩网站检测模型构建步骤如图 2。

- (1) 请求 URL 获取 HTML 文本；
- (2) 通过正则匹配从 HTML 指定标签中提取文本。由于各个国家对非法网站的界定不同，目前本模型只针对国内网站进行识别，因