

基于物联网和 PCA 支持向量机的交通流量预测系统

王永皎, 郭力争

(河南城建学院 计算机科学与工程学院, 河南 平顶山 467036)

摘要: 为了解决已有交通流量监测系统存在的数据采集分散、车辆识别度低、实时性差和流量预测误差大等问题, 设计了一种基于物联网技术和最小二乘支持向量机 (Least Squares Support Vector Machine, LSSVM) 的交通流量预测系统; 首先, 描述了系统原理和部署模型, 然后对系统的硬件即车载传感器节点和 Sink 节点进行了设计, 同时对系统的软件流程进行了描述, 通过在监控中心执行 PCA 主成分分析方法实现对采集数据提取独立主成分, 消除无关冗余数据, 在此基础上采用 LSSVM 实现道路交流流量预测; 最后, 在十字路口布置实验环境, 实验结果表明: 文章方法能实时精确地实现交通流量预测, 与其它方法相比, 具有拟合精度高和的泛化能力强的优点, 具有很强的实用性。

关键词: 交通流量预测; 支持向量机; 物联网; 主成分分析

Prediction System for Traffic Flow Based on Internet of Things and PCA Support Vector Machine

Wang Yongjiao, Guo Lizheng

(School of Computer Science, University of Urban Construction, Pingdingshan 467036, China)

Abstract: In order to solve the given traffic flow monitoring system existing the problems such as data collection dispersion, low vehicle identification, low in-time performance and big prediction error, a traffic flow prediction based on Internet of things and LSSVM (Least Squares Support Vector Machine) was proposed. Firstly, the principle and deployment model of system was described. Then the hardware of system includes vehicle sensor and Sink node was designed, and the system software was also introduced. The monitoring center executed the PCA method to extract the independent main information for the rude data, and then the LSSVM was operated to predict the traffic flow for the next time. Finally, the method in this paper was simulated in the environment of crossing road, and the simulation result shows: the proposed method can in time and accurately predict the traffic flow, and compared with other methods, it has the advantages of high fitting accuracy and generalizing ability. Therefore, it has big practicability.

Keywords: traffic flow prediction; support vector machine; Internet of things; principal component analysis

0 引言

交通系统^[1-2]是支持社会经济发展和人民生活质量提高的基础设施, 对交通网络流量^[3-4]进行实时动态的预测和统计是提高交通运输效率的可行方法。现有的通过车载自组网^[5]将移动计算与智能交通相结合进行交流流量的预测与分析, 具有可靠性和实时性不高的缺点。

物联网^[6-7]由 Kevin Ashton 于 1998 年提出, 能实现物体与物体以及人与物体之间的通信。因此, 采用基于物联网的交通系统能克服车载自组网所具有的可靠性和实时性不高的缺点, 能对具有复杂性、非线性、非结构化的城市交通状态进行实时分析和预测, 实现真正动态、实时和有效的交通控制和管理。

因此, 本文设计了一种基于物联网和 PCA 支持向量机的智能交通流量预测系统, 首先描述了系统总体结构, 然后对系统软硬件进行了设计, 提出了一种采用 PCA^[8-9]主成分分析法对采集数据的属性简约方法对传感器发送的数据进行降维, 并采用支持向量机^[10-11]对交通流量^[12]进行预测, 通过实验证明

了文中系统能有效地预测城市交通流量, 具有重要的意义。

1 系统原理

文中基于物联网和 PCA 支持向量机的交通流量监控系统采用 Zigbee 协议组成树状网络拓扑结构, 系统主要由三部分组成: 车载传感器节点、Sink 节点和控制中心组成, 文中系统在道路上的部署模型如图 1 所示。

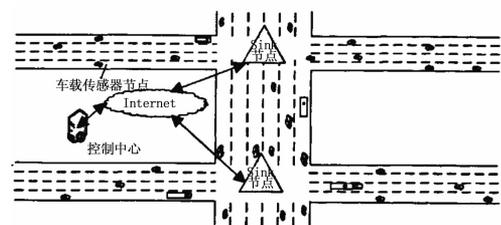


图 1 系统道路部署模型

其中, 车载传感器节点位于在道路行驶的每辆汽车上, 将其采集的车辆时速信息、位置信息 (通过读取 GPS 设备值) 等打包成数据包, 每隔一段时间 t , 车辆传感器向周围车辆广播自身的数据包, 当车辆传感器节点收到其它节点的数据包后, 对信息进行融合从而预测一个交通流量的初值, 并周期性地发送入网请求, 以申请加入网络, 将流量融合信息通过网络发送给 Sink 节点, 同时接收从 Sink 节点发送的各道路交

收稿日期: 2014-01-17; 修回日期: 2014-03-17。

基金项目: 河南省重点科技攻关项目 (132102210478)。

作者简介: 王永皎 (1977-), 女, 河南新乡人, 副教授, 博士, 主要从事人工智能、图像处理领域等方向的研究。

通流量当前值和预测值，为其选择道路通行提供有力的依据。

Sink 节点分布在各个道路交通十字路口处，主要功能是接收由各车载传感器发送的各道路交通流量信息和传达由控制中心发送的数据，Sink 节点在接收了车载传感器发送的数据后，对其进行融合，最终将交通流量的融合结果通过互联网发送给控制中心；同时接收控制中心发送的流量估计值，将其发送给车载传感器。

控制中心在接收了 Sink 节点发送的交通流量后，对各条道路对应的各条流量信息进行主成分分析，以降低数据属性维数，然后采用支持向量机对道路交通流量进行预测，将各道路的交通流量的预测结果通过 Sink 节点发送给车载传感器。

2 硬件设计

系统硬件设计主要包含车载传感器节点的设计和 Sink 节点的设计

2.1 Sink 节点硬件设计

Sink 节点的硬件设计如图 2 所示。

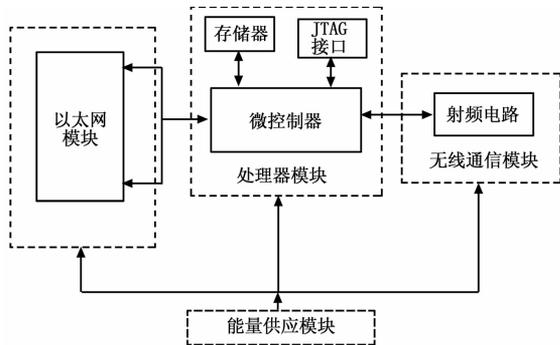


图 2 Sink 节点硬件设计框图

微控制器采用 S3C2410，具有功耗低、性能高和价格低等优点，其 CPU 为 16/32 位的 RISC 精简指令 ARM920T 处理器，同时具备扩展 I/O 接口、ADC 接口、MMC 接口、JTAG 接口以及其它接口电路等。

以太网模块采用芯片 DM9000，其支持 IEEE802.3x 全双工流量控制和半双工流量控制，兼容 3.3 V 和 5 V 的输入输出电压，支持处理器的字节或双字节读写方式，具备 4K 的静态随机存储器。

无线通信模块采用满足 IEEE802.15.4 协议标准的片上系统芯片 CC2430。

2.2 车载传感器节点硬件设计

车载传感器节点的硬件设计如图 3 所示。

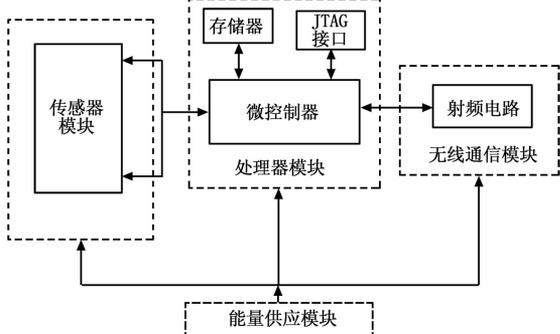


图 3 车载传感器节点硬件设计框图

车载传感器节点的微控制器和无线通信模块采用满足 IEEE802.15.4 协议标准的具有功耗低、集成性高和成本低的优点的片上系统芯片 CC2430。

传感器模块采用 MOD-EL2420 微机械电容式三轴加速度传感器，该传感器融合了氨阻尼、G-LOGGERT、输出数字脉冲序列化、校准精度高和抗电干扰性强的优点。

3 软件设计

3.1 Sink 节点软件设计

Sink 节点的主要软件流程描述如下。

(1) 维护网络拓扑结构：当有新的车载传感器节点即新的车辆加入时，为车载传感器分配唯一 ID；当有车载传感器节点离开时，将该节点 ID 从管理节点表中删除；

(2) 实时监测车载传感器节点是否有数据发送，当有数据到来时，接受并融合由车载传感器发送的数据，并最终将其发送到监控中心；

(3) 接收监控中心发送的交流流量当前和预测值，并将其发送到 WSN 中的车载传感器，最终通知车主道路流量情况。

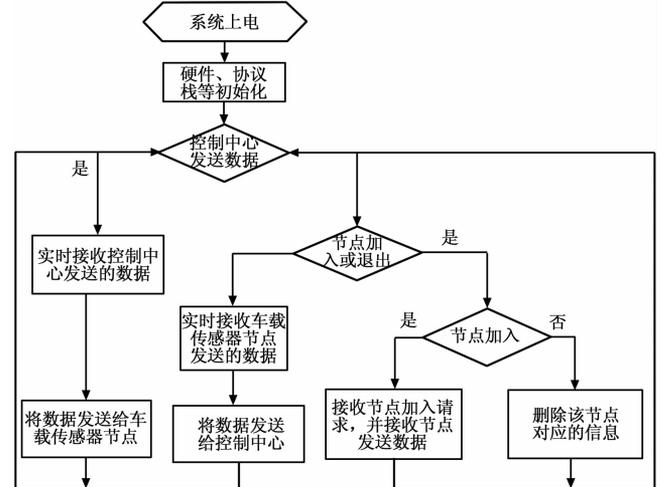


图 4 Sink 节点软件功能图

3.2 车载传感器节点软件设计

车载传感器节点的主要软件流程可以描述为：

(1) 在经过十字路口时监测到空闲的无线通信信道时，通过无线自组织的方式加入到网络中，并被分配唯一全局 ID；

(2) 在中断到来时，周期性地采集周围车载传感器节点即车辆信息，结合自身信息进行有效融合后，发送给 Sink 节点；

(3) 接收由 Sink 节点发送的数据，并通过广播和 LED 显示的方式提示用户流量预测值。

4 监控中心流量预测

4.1 PCA 主成分分析

监控中心收到的来自 Sink 节点发送的道路交通流量是来自多个车载传感器发送的数据，而各车载传感器采集的数据往往对应了道路网的各断面交通信息，而这些信息又存在着相互重叠的情况。

因此，为了提高预测精度和效率，首先采用 PCA 主成分分析法对各断面交通流量变量进行降维，得到一组新的、独立的能反应断面流量的综合变量。

假设在时刻 t ，道路对应的断面为 m ，断面交通流量数据

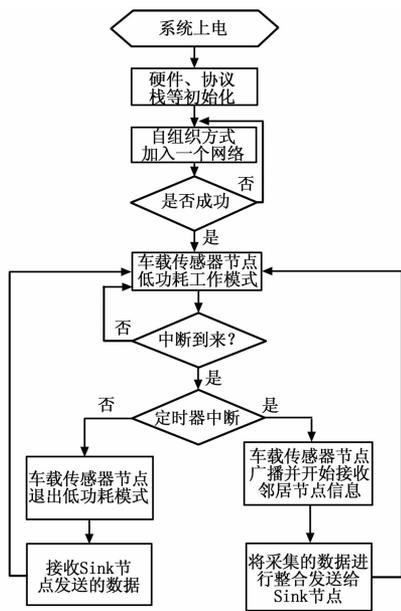


图 5 车载传感器节点软件功能图

长度为 n , 采用 $S_i(t)$ 表示时刻 t 对应的第 i 个断面的交通流量变量, 即:

$$S = \begin{bmatrix} S_1(1) & S_1(2) & \cdots & S_1(n) \\ S_2(1) & S_2(2) & \cdots & S_2(n) \\ \vdots & \vdots & & \vdots \\ S_m(1) & S_m(2) & \cdots & S_m(n) \end{bmatrix} \quad (1)$$

采用 PCA 主成分分析法提取采集数据主要成分的过程可以描述为:

(1) 首先计算各断面交通流量变量的均值为:

$$\bar{S}(t) = \frac{1}{m} \sum_{i=1}^m S_i(t) \quad (2)$$

(2) 将每个交通流量变量减去均值得到 $n \times n$ 的协方差矩阵, 如下所示:

$$S = \frac{1}{n} \sum_{i=1}^M (x_i - \bar{x})(x_i - \bar{x})^T \quad (3)$$

(3) 求解协方差矩阵 S 的特征值 λ_i 以及对应的正交归一化的特征向量 z_i , 并计算求取的各特征值的累积贡献率 α :

$$\alpha = \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^D \lambda_i} \quad (4)$$

(5) 选取贡献率 α 大于预先设定的阈值 th 的特征值对应的特征向量来构造特征空间 \sum , 并将矩阵 S 投影到特征空间 \sum 上, 得到的矩阵 A 即为所求主成分矩阵:

$$A = \sum^T X \quad (5)$$

4.2 基于 LSSVM 的交通流量预测

最小二乘支持向量机 (Least Squares Support Vector Machine, LSSVM) 是在支持向量机的基础上发展起来的, 将传统支持向量机中的不等式约束改为等式约束, 将误差平方和损失函数作为训练样本集的经验损失函数, 因此, 将二次规划问题转换为线性方程组的求解问题。

假设样本数据 $\{x_i, y_i\} (1 \leq i \leq n)$, 其中 n 为样本数, $x_i \in \mathbf{R}^d$ 为输入向量, $y_i \in \mathbf{R}$ 为对应的输出, $\varphi(x)$ 为将样本从低维映射到高维的空间的非线性函数, 则可以通过下式对非线性函数进行估计:

$$y(x) = w^T \varphi(x) + b \quad (6)$$

式中, w 为超平面的法向量, b 为偏差, 在考虑函数复杂度和拟合误差的基础上, 上式可以转化为求最小化 $\|w^T w\|$ 的问题, 即:

$$\begin{cases} \min \frac{1}{2} w^T w + \frac{1}{2} c \sum_{i=1}^N \xi_i^2 \\ s. t. y_i = w^T x_i + b + \xi_i, \quad 1 \leq i \leq n \end{cases} \quad (7)$$

其中: c 为惩罚参数, ξ_i 为松弛变量, 采用拉格朗日乘子法对其进行求解, 建立式 (7) 对应的优化问题为:

$$\begin{cases} \max W(\lambda) = \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i,j=1}^n \lambda_i \lambda_j y_i y_j K(x_i, x_j) \\ s. t. \sum_{i=1}^n \lambda_i y_i = 0 \quad 0 \leq \lambda_i \leq c, i = 1, 2, \dots, n \end{cases} \quad (8)$$

式 (8) 中, λ_i 为拉格朗日乘子, $K(x_i, x_j) = \varphi(x_i) \varphi(x_j)$ 为核函数, 因此 LSSVM 对函数的估计可以表示为:

$$f(x) = \text{sign}(\lambda_i y_i k(x_i, x_j) + b) \quad (9)$$

因此, 通过式 (9) 则可以在输入交通流量当前值的前提下, 获得下一时刻的交通流量的估计值。

5 仿真实验

在某几条道路及十字路口布置实验环境, 对文中基于物联网和 PCA 支持向量机的智能交通流量预测系统进行仿真, 监控中心在收到 Sink 节点数据后进行交通流量预测, 将监测 30 天得到的 1 200 个采集数据进行仿真, 将其中的 600 个数据作为训练样本, 剩余的 600 个数据作为测试样本, 文中的参数设置如下: PCA 贡献率阈值 th 为 0.9, 惩罚参数 c 为 2, ξ_i 为松弛变量为 0.8, 采用核函数为径向基函数:

$$k(x_i, x_j) = \exp(-\|x_i - x_j\| / 2\sigma^2) \quad (10)$$

式中, 宽度参数 σ^2 取值为 200, 其取值是通过 30 次实验的观察得到的最优训练效果时对应的值。

监控中心通过 PCA 主成分分析法将采集的数据经过降维, 获得独立不相关的数据后, 然后通过 LSSVM 进行交通流量估计, 并与文献 [13] 对比, 得到的结果如图 6~7 所示。

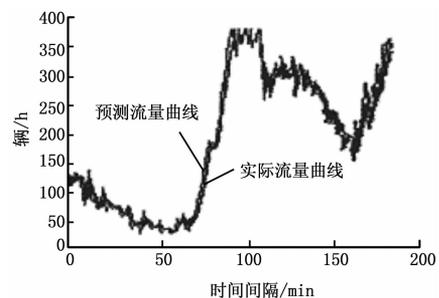


图 6 文献 [13] 方法预测与实际流量对比

从图 7 中可以看出, 文中方法能较好地拟合实际交通流量, 具有较好的预测能力, 且与文献 [13] 方法相比, 文中方法误差较小、拟合能力更强, 同时具有较强的泛化能力。

(下转第 2233 页)

护进程,调用 socket () 函数建立 TCP 套接字,然后调用 bind () 函数将该套接字和本地端口(即 8000 或 8001)绑定在一起,再调用 listen () 函数使套接字做好侦听准备,之后等待客户端连接请求。当客户端调用 connect () 函数请求与服务器建立连接时,服务器调用 accept () 函数来接受连接。连接一旦建立,客户端和服务器之间就可以通过调用 recv () 函数和 send () 函数来发送和接收数据。当客户端发送数据时,服务器调用数据接收函数 recv () 接收数据,然后判断发送方式,调用 BU-61580 或 DEI116 的发送数据函数将数据内容写入相应的寄存器或存储器内,启动发送;当客户端读取数据时,服务器将存储器或寄存器地址传递给数据读取函数,然后把读取到的数据经以太网打包后发送给客户端^[10]。待数据传送结束后,双方调用 closesocket () 函数关闭套接字。

5 结论

LXI 总线自推出以来就广泛受到国内外专家学者的关注,近年来发展迅速,显示出其组建测试系统的众多优点。本文设计实现了 LXI 总线的 1553B 及 ARINC429 通信模块。硬件方面基于 SOPC 技术将 IP 核、接口、外设、微处理器等集成到 FPGA 芯片上,设计了模块的以太网接口、1553B 接口和 ARINC429 接口。软件方面,基于 $\mu\text{C}/\text{OS II}$ 和 LwIP 协议实现了 TCP/IP 协议,从而实现了 1553B 及 ARINC429 总线的协议任务。LXI 总线的 1553B 及 ARINC429 通信模块的设计实现,满足了越来越多武器装备对 1553B 及 ARINC429 总线的测试

需求,推动 LXI 总线在国内的研究和应用。

参考文献:

- [1] Agilent Technologies. 转向 LXI 的十个好理由 [EB/OL]. <http://www.agilent.com>, 2006.
- [2] Department of Defense. MIL-STD-1553B. Aircraft Internal Time Division Command/Response Multiplex Data Bus [S]. Washington D. C: Department of Defense, 1978.
- [3] 冯金富,刘安,左伟,等. 模型驱动的开放式结构悬挂物管理系统设计方法 [J]. 空军工程大学学报:自然科学版, 2010, 11 (4): 1-5.
- [4] Altera. Cyclone II FPGA 系列简介 [EB/OL]. <http://www.altera.com.cn/products/devices/cyclone2/overview/cy2-overview.html>, 2005.
- [5] Adam Dunkels. Design and Implementation of the LwIP TCP/IP Stack [M]. 2001, 4-35.
- [6] 刘金龙,史忠科. 基于 FPGA 的以太网与 1553B 网络接口设计 [J]. 测控技术, 2010, (10): 84-87.
- [7] 张田甜,刘亚斌. 基于 PXI 总线的多功能 ARINC429 通信接口电路设计 [J]. 计算机测量与控制, 2009, 17 (3): 534-535.
- [8] 阙大顺,王近涛. LwIP 协议在 $\mu\text{C}/\text{OS-II}$ 系统上的移植和实现 [J]. 舰船电子工程, 2006, 26 (4): 89-91.
- [9] 王泉崇. 嵌入式 Web 服务器 [D]. 杭州:浙江大学, 2005.
- [10] 陈普然. 基于嵌入式 MPU 核的 1553B 总线接口控制器的设计 [D]. 南京:南京航空航天大学, 2003.

(上接第 2215 页)

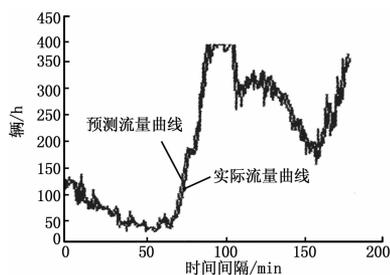


图7 文中方法预测与实际流量对比

6 结语

为了实现对具有复杂性、非线性、非结构化等特征的城市交通状态进行实时分析和预测,提出了一种基于物联网和 PCA 支持向量机的智能交通预测系统。首先,介绍了系统原理和总体结构,然后对系统的软硬件进行了详细设计,提出了一种在监控中心采用 PCA 对采集数据进行降维,提取其中独立不相关主成分的方法,并采用 LSSVM 对交流流量值进行预测。仿真实验证明了文中方法具有较高的预测精度,是一种有效可行的方法。

参考文献:

- [1] 刘小洋,伍民友. 车联网:物联网在城市交通网络中的应用 [J]. 计算机应用, 2012, 32 (4): 900-904.
- [2] Row S J. IntelliDrive: safer smarter greener [EB/OL]. [2010-04]. <http://www.fhwa.dot.gov/publications/publicroads/>

10julaug/04. cfm.

- [3] 蒋新华,陈宇,朱铨,等. 交通物联网的发展现状及趋势研究 [J]. 计算机应用研究, 2013, 30 (8): 2256-2261.
- [4] Baskar L D, De Schutter B, Hellendoorn J, et al. Traffic control and intelligent vehicle highway systems: a survey [J]. Intelligent Transport Systems, 2011, 5 (1): 38-52.
- [5] 李丽君,刘鸿飞,杨祖元,等. 车用自组网信息广播 [J]. 软件学报, 2010, 21 (7): 1620-1634.
- [6] Elson J, Girod L, Estrin D. Fine-grained Network Time Synchronization Using Reference Broadcasts [A]. Fifth Symposium Operating Systems Design and Implementation, Boston [C]. Massachusetts, USA, 2001: 1-2.
- [7] 吴振强,周彦伟,马建峰. 物联网安全传输模型 [J]. 计算机学报, 2011, 34 (8): 1351-1364.
- [8] 郁雪,李敏强. 基于 PCA-SOM 的混合协同过滤模型 [J]. 系统工程理论与实践, 2010, 30 (10): 1850-1854.
- [9] 赵鑫,汪维家,曾雅云,等. 改进的模块 PCA 人脸识别新算法 [J]. 计算机工程与应用, 2013, (26): 1-5.
- [10] Osuna E, Freund R. Training support vector machines: an application to face detection [A]. Proc of computer vision and pattern recognition, IEEE Computer Society [C], 1997: 130-136.
- [11] 庄新妍. 遗传优化支持向量机在电力负荷预测中的应用 [J]. 计算机仿真, 2012, 29 (3): 348-350.
- [12] 唐新来,李春贵,王萌,等. 基于混沌粒子群算法的神经网络短时交通流预测 [J]. 2010, 18 (8): 1893-1895.
- [13] 李建武,陈森发,黄鹏. 基于粗集理论和支持向量机的道路网短时交通流量预测 [J]. 计算机应用研究, 2010, 27 (10): 3683-3690.