

基于SW-SVR的畜禽养植物联网异常数据实时检测方法

段青玲^{1,2} 肖晓琰¹ 刘怡然¹ 张璐¹

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 北京市农业物联网工程技术研究中心, 北京 100097)

摘要: 畜禽养植物联网由于工作环境恶劣、网络传输故障等因素容易产生异常感知数据,为保证数据质量,根据畜禽养植物联网数据流周期性、时序性等特点,提出了一种基于滑动窗口与支持向量回归(Sliding window and support vector machines for regression, SW-SVR)的异常数据实时检测方法。首先根据畜禽物联网数据流特征周期以及采样频率确定滑动窗口尺寸;然后通过SVR模型预测畜禽养植物联网数据流中某一时刻传感器测量值;最后计算预测区间,根据实际测量值是否落入该区间判断是否异常并对异常数据进行置换处理。采用畜禽养植物联网环境数据进行试验,结果表明:所提滑动窗口计算方法得到的窗口尺寸预测的MAPE为0.1884,畜禽养植物联网异常数据检测率达98%,能够有效检测和处理畜禽养植物联网数据流中的异常数据。

关键词: 异常数据检测; 畜禽养植物联网; 滑动窗口; 支持向量回归

中图分类号: TP274+.2; TP393.03 **文献标识码:** A **文章编号:** 1000-1298(2017)08-0159-07

Anomaly Data Real-time Detection Method of Livestock Breeding Internet of Things Based on SW-SVR

DUAN Qingling^{1,2} XIAO Xiaoyan¹ LIU Yiran¹ ZHANG Lu¹

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. Beijing Engineering Research Center of Agricultural Internet of Things, Beijing 100097, China)

Abstract: Due to bad work environment and network transmission failure, it is easy to generate abnormal sensory data in livestock breeding Internet of things system. In order to ensure the quality of sensory data, according to the characteristics of sensory data flow such as periodicity, temporality, infinity, etc., a method was proposed based on sliding window and support vector machines regression (SW-SVR) for livestock breeding Internet of things abnormal sensory data detection in real time. Firstly, the sliding window size was decided according to the characteristic period and sampling frequency of data flow from livestock breeding Internet of things system, and the history data within sliding window was selected as the input value of prediction model. Then, the sensor estimated measurement value at certain moment in livestock breeding Internet of things system was predicted by using SVR model. Finally, the prediction interval (PI) was calculated, and the abnormal sensory data was identified if the sensor actual measurement data fell out of the PI. The abnormal data would be replaced by the predictive data. The abnormal sensory data detection method was tested by data flow from real livestock breeding Internet of things system. Experiment results showed that the mean absolute percent error value of prediction with window size calculated by the sliding window method was 0.1884. The correct detection rate of abnormal data based on SVR model with radial basis function kernel (RBF kernel) achieved 98%, which had higher accuracy compared with BP neural network (BPNN) method. Abnormal data can be effectively detected and treated in livestock breeding Internet of things system.

Key words: anomaly data detection; livestock breeding internet of things; sliding window; support vector machines for regression

引言

当前物联网已经成为畜禽养殖生产获取数据的

重要手段之一。在畜禽养植物联网中,传感器按照时间序列连续不断地采集温湿度、光照、有害气体浓度等畜禽生长环境数据,并以数据流的形式传输至

数据处理系统,具有显著周期性、实时性、无穷性等特征^[1-3]。畜禽物联网的工作环境通常较恶劣,由于传感器设备的暂时性故障、数据传输错误等原因产生异常数据^[4-6],使得数据质量急剧下降,无法保证设备的自动控制和数据的有效分析。

异常值的概念由 Edgeworth 首次定义,是指“显然严重偏离样本集合中其他观测值的观测值”^[7]。异常数据检测是指从数据集中找出明显区别于其他数据的数据^[8]。根据异常检测的数据来源,异常数据检测技术可分为离线检测和在线检测。离线异常数据检测主要采用最近邻、聚类、神经网络分类器、支持向量机分类器等机器学习^[9-13]和统计方法^[14-16]对静态样本数据进行异常检测,离线异常数据检测的对象是收集的静态数据,不适用于畜禽物联网中实时数据流异常检测。在线异常数据实时检测主要基于预测对实时数据流进行异常检测,何欢^[17]采用递归神经网络检测概念漂移方法进行异常数据检测,着重于检测数据中的隐藏状态变化;HILL 等^[18]利用 BP 神经网络模型计算预测值及置信区间,与实际测量值比较进行异常分类。苑进等^[19]通过自回归高斯模型计算大棚内温湿度预测误差带进行异常分类,但高斯过程回归模型受初值和协方差函数影响较大。SVR 算法以其训练样本小、泛化能力强的特点在数据预测领域广受关注^[20],并成功应用在时间序列预测研究中^[21-24],目前在异常检测领域也有所研究^[25-29],但应用较少。基于预测的异常数据实时检测方法,一般选取固定滑动窗口确定预测模型的输入项,是预测准确性和运行效率的重要影响因子,而滑动窗口尺寸选择因素复杂,物联网中数据采集频率、特征周期各不相同,如畜禽物联网中氨气、温度采集时间间隔分别为 3 min 和 5 min,目前的研究中多是人为确定滑动窗口尺寸,预测结果具有随机性。本文提出基于 SW-SVR 的畜禽养殖物联网异常数据实时检测方法,选择窗口尺寸并根据滑动窗口内历史数据预测传感器测量值,计算预测区间,并与实际值对比进行异常数据分类与替换处理,以期满足畜禽养殖物联网数据流异常检测对准确度和运算效率的要求。

1 基于 SW-SVR 的畜禽养殖物联网异常数据检测方法

根据畜禽物联网中数据流周期性、实时性、无穷性的特点构建基于 SW-SVR 的畜禽养殖物联网异常数据检测方法,以有效保证数据的实时性和准确率,为设备的自动控制和有效的数据分析提供数据支持。

1.1 总体流程

畜禽养殖物联网异常数据实时检测处理流程如图 1 所示。其中针对畜禽养殖物联网异常数据实时检测处理流程中的符号定义如下:设 x_t 为畜禽养殖物联网中传感器在 t 时刻的实际测量值,称 $X = \{\dots, x_{t-1}, x_t, x_{t+1}, \dots\}$ 为畜禽养殖物联网中传感器采集的数据流,表示为按照一定时间次序组成的无限集合。设 q 为滑动窗口尺寸, x_{t-q} 为 t 时刻之前第 q 个畜禽物联网传感器测量值,称 $D_t = \{x_{t-q}, x_{t-q+1}, \dots, x_{t-1}\}$ 为 t 时刻滑动窗口数据集。通常选取 t 时刻之前的 q 个历史数据组成滑动窗口数据集,作为单步预测的输入项。若 D_t 表示 t 时刻滑动窗口数据集,称 $\bar{x}_t = \text{SVR}(D_t)$ 为通过 SVR 算法计算得到的 t 时刻预测值。

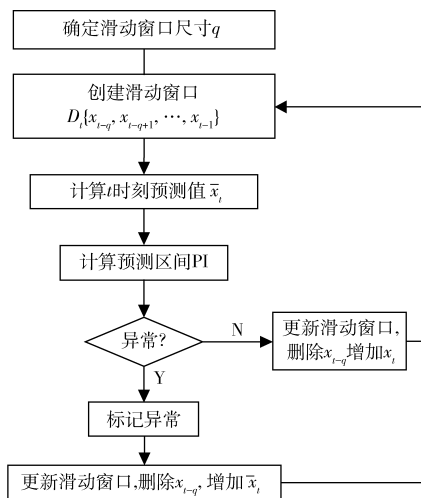


图 1 基于 SW-SVR 模型的异常数据检测方法

Fig. 1 Anomaly data detection method based on SW-SVR

畜禽养殖物联网异常数据实时检测处理流程包括以下步骤:

- (1) 基于采样频率和数据周期确定滑动窗口 q 。
- (2) 根据滑动窗口尺寸 q , 选择窗口内历史数据 (即 t 时刻其所在时间序列之前 q 个测量值) 组成滑动窗口数据集。
- (3) 采用 SVR 算法进行预测, 即计算数据流中 t 时刻的预测值。
- (4) 计算 t 时刻预测区间 PI, 与实际测量值对比判断是否异常, 并对异常数据进行置换处理, 剔除异常数据。
- (5) 更新滑动窗口, 重复步骤 (2) ~ (4)。

1.2 滑动窗口构建

畜禽养殖物联网数据流与静态数据不同, 具有实时性、无穷性等特点, 不可能用所有的数据集作为数据源, 因此需要创建滑动窗口, 以窗口内的历史数据作为预测模型的输入值。假设窗口尺寸为 q , 则保存最近一段时间的 q 个数据, 也称之为滑动窗口

数据集。

滑动窗口尺寸的选择会影响预测模型的准确性与时间效率,窗口尺寸越大,说明基于越多的历史数据来预测当前值,准确度越高,但同时耗费时间越长;因此选择滑动窗口需要均衡以上 2 个因素,选择合适的窗口尺寸。

根据畜禽物联网数据流中采集数据的特征周期以及数据采集的时间间隔,提出了滑动窗口尺寸 q 的计算方法,公式为

$$q = \text{ceil}(Q)$$

(1)

其中 $Q = \frac{T}{\Delta t} e^{-\beta \ln \sqrt{T/\Delta t}}$

式中 T ——畜禽养殖物联网数据流规律分布的特征周期

Δt ——畜禽物联网传感器采集数据时间间隔

$\frac{T}{\Delta t}$ ——畜禽物联网数据流一个特征周期的样本数量

$\text{ceil}(Q)$ ——上取整数函数,取不小于 Q 的最小整数

特征周期指对畜禽养殖物联网数据流进行分布规律分析得到的周期,如环境检测数据特征周期为 1 d;滑动窗口尺寸对预测准确性影响较大,与运行效率成反比;基于一个特征周期的样本预测效果最好,但样本量过大,耗时严重;窗口最小取 1 时,运行时间最短,但准确率较低。式(1)中一个特征周期的样本数量 $T/\Delta t$ 同时控制滑动窗口选择函数上升幅度和衰减速度,满足预测准确度和运行效率需求。

1.3 基于 SVR 的单步预测

采用 SVR 算法对畜禽养殖物联网数据流中的测量值进行预测^[30-31]。基于 SVR 的单步预测模型取 t 时刻之前长度为 q 滑动窗口内历史数据作为输入数据源,利用 SVR 模型对 t 时刻物联网传感器的测量值进行预估计算,输出第 t 个序列值。因此,SVR 模型输入结点的个数为 q ,输出结点个数为 1。公式为

$$\bar{x}_t = M(D_t) + R$$

(2)

式中 $M(D_t)$ ——SVR 模型

R ——经验风险,可以用不同损失函数来描述,本试验中采用 e 不敏感函数

SVR 模型中的主要参数包括复杂度参数 C 、学习算法以及核函数^[19],设置复杂度参数 C 为 1.0;核函数用于将原始数据映射到高维空间中进行非线性回归,一定程度上受数据集的影响,其中应用最广的为径向基核函数^[30],本文选取径向基函数和多项式函数作为回归预测的核函数进行对比。

1.4 异常数据识别与处理

在对滑动窗口内的历史数据进行单步预测后,需要利用模型残差确定概率为 p (也称置信水平)的预测区间(Prediction interval,PI)。假定新近的测量值落入预测区间的概率为 $p = 100(1 - \alpha)$,且模型残差符合均值为零的高斯分布,则预测区间表示为

$$P_t = \bar{x}_t \pm t_{\alpha/2, n-1} S \sqrt{1 + 1/n}$$

(3)

式中 P_t —— t 时刻的预测区间

$t_{\alpha/2, n-1}$ —— p 百分位数自由度为 $n - 1$ 的符合学生 t 分布概率分布函数

S —— n 个样本的标准偏差

采用该预测区间的优势在于区间宽度受经验模型误差影响,不受人为干预,保证了预测模型的客观性。

计算得到预测区间后,将 t 时刻畜禽物联网的实际测量值与预测区间进行比较,若 t 时刻的实际测量值落入对应的预测区间内,判定该实际测量值正常,否则判断该时刻实际测量值为异常数据。

针对判断为异常数据的实际测量值,需要将该测量值标识为异常数据,并且用 SVR 模型回归计算的预测值代替被识别为异常数据的实际测量值,作为计算下一时刻预测值的输入数据,更新滑动窗口。

$$x_t \rightarrow \bar{x}_t$$

$$D_t \{x_{t-q}, x_{t-q+1}, \dots, x_{t-1}\} \rightarrow D_{t+1} \{x_{t-q+1}, x_{t-q+2}, \dots, \bar{x}_t\}$$

式中 D_t ——预测 t 时刻测量值的数据源

D_{t+1} ——更新滑动窗口后预测 $t + 1$ 时刻测量值的数据源

基于滑动窗口的 SVR 算法以历史数据建立模型计算得到预测值,直接替代异常数据在算法实现方面更为简练,并且作为下一时刻预测模型的输入值预测准确度更高。

2 试验

试验采用科尔诺公司的 MOT300 - NH3 型氨气检测仪根据电化学原理采集氨气数据,并且配合信号发射器和接收器进行远距离无线传输。以 2016 年 4 月 8 日 00 时 00 分至 11 日 23 时 53 分在天津惠康种猪养殖场采集的氨气数据作为样本集,氨气传感器节点每隔 3 min 采集一次数据,持续 96 h,共采集到 1 920 个数据,其中包括噪声数据 60 个。

在确定滑动窗口尺寸和单步预测畜禽物联网传感器测量值阶段,采用模型评估参数平均绝对误差(MAE)、均方误差(MSE)、平均绝对百分比误差(MAPE)对基于径向基核函数的 SVR 模型(SVR - RBF)、基于多项式核函数的 SVR 模型(SVR - Poly)以及 BP 神经网络(BP neural network, BPNN)模型

的预测效果进行分析,公式为

$$MAE = \frac{1}{N} \sum_{t=1}^N |x_t - \bar{x}_t|$$

(4)

$$MSE = \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x}_t)^2$$

(5)

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{x_t - \bar{x}_t}{x_t} \right|$$

(6)

式中 t ——时间序列时刻

N ——样本个数

x_t —— t 时刻的实际测量值

$\bar{x}_t$ —— t 时刻的预测值

在畜禽物联网异常数据分类处理阶段,通过计算评估参数检测率(True positive rate,TPR)、误检率(Flase positive rate,FPR)对 SVR 模型以及 BP 神经网络模型在不同置信水平下的异常数据检测效果评估,公式为

$$T_{pr} = \frac{T_p}{T_p + F_n} \times 100\%$$

(7)

$$F_{pr} = \frac{F_p}{F_p + T_n} \times 100\%$$

(8)

式中 T_{pr} ——检测率,% F_{pr} ——误检率,%

T_p ——判断为异常实际上是异常值的个数

F_p ——判断为异常实际是正常值的个数

F_n ——判断为正常实际上是异常值的个数

T_n ——判断为正常实际上也是正常值的个数

2.1 滑动窗口选择

对连续一个月的氨气数据进行分析,氨气浓度具有明显周期性分布规律,且以天为周期,则单步预测时取特征周期 T 为 1 440 min; β 一般取值为 0.5、1、2^[2],本试验中取 0.5;样本采集时间间隔为 3 min。

根据试验采集数据的特征周期、采集间隔等参数,按照式(1)计算得到窗口尺寸为 24。滑动窗口的尺寸取值一般在 10~30 之间^[18],本文分别采用滑动窗口尺寸为 10、15、20、25、30、35、40 进行试验;采用基于径向基核函数的 SVR 模型分别在不同的窗口尺寸下对畜禽养植物联网中氨气样本数据预测,模型评估结果如图 2 所示。

由图 2 中不同窗口尺寸下 SVR 模型预测效果来看,随着窗口尺寸增大,MAE、MSE、MAPE 逐渐减小;平均运行时间呈递增趋势;MAE、MSE、MAPE 分别低于 0.4、0.35、0.2 时,窗口尺寸 q 为 25 时,平均运行时间最低,为 2.5 ms,为最佳窗口尺寸,与根据滑动窗口算法计算得到的窗口尺寸为 24 时的误差以及运行时间相差甚微,可视为结果基本一致,本文提出的滑动窗口选择算法能够满足畜禽养植物联网单步预测的准确性与时间效率要求。

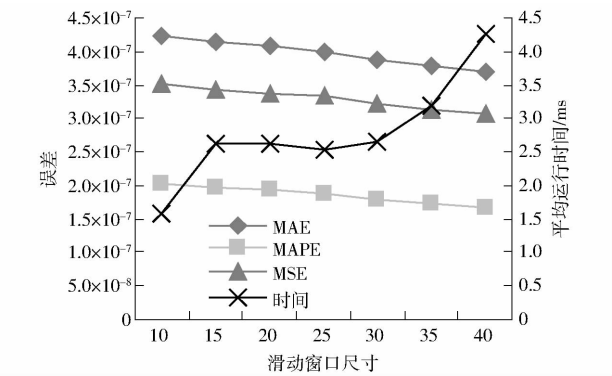


图 2 不同窗口尺寸模型评估结果

Fig. 2 Model evaluation results in different window sizes

2.2 基于 SVR 的单步预测结果分析

SVR 模型中滞后步长由滑动窗口尺寸确定,核函数分别选取径向基核函数和多项式核函数;BP 神经网络预测模型^[17]中学习率为 0.01,动量因子为 0.1,隐藏层个数为 1,每个隐藏层节点数设为 50。

试验中分别采用基于径向基核函数的 SVR 模型、基于多项式核函数的 SVR 模型、BP 神经网络模型对畜禽物联网中氨气样本数据进行预测,采用 MAE、MSE、MAPE、平均运行时间参数对 3 种预测模型进行效果评估,结果如表 1 所示。

表 1 不同预测模型评估结果

Tab. 1 Model evaluation results with different prediction models

预测模型	MAE	MAPE	MSE	平均运行时间/ms
SVR - RBF	0.400 3	0.188 4	0.335 2	2.52
SVR - Poly	0.492 0	0.225 0	0.499 0	10.85
BPNN	0.494	0.319 9	0.425 2	2.67

表 1 中针对种猪场氨气浓度单步预测评估结果可以看出,SVR 模型中多项式核函数的 MAE、MSE、MAPE 都比较高,平均用时过长,明显不适用于试验样本数据集的预测;与 BP 神经网络模型相比较,基于径向基核函数的 SVR 模型的 MAE、MSE、MAPE 相对较低,耗时较少,预测效果更好。

SVR 模型中多项式核函数平均运行时间明显较长,预测均方误差较高,在后续试验中主要分析 SVR - RBF 模型和 BP 神经网络模型对畜禽物联网中氨气样本数据的预测效果以及拟合度,其部分结果如图 3 所示。

图 3 中针对种猪场内氨气的单步预测结果中,基于 BP 神经网络模型的预测结果波动较大,明显受异常数据影响,容易出现过拟合效果;基于径向基核函数的 SVR 算法的单步预测结果相对平稳,受异常数据的影响较小,单步预测效果更好。

2.3 异常检测结果分析

采用基于径向基核函数的 SVR 模型和 BP 神经

网络模型对畜禽物联网中的氨气数据进行预测,并设置置信水平,计算其置信区间,置信水平 p 为

95% 和 99% 时 2 种模型的部分预测结果如图 4 所示。

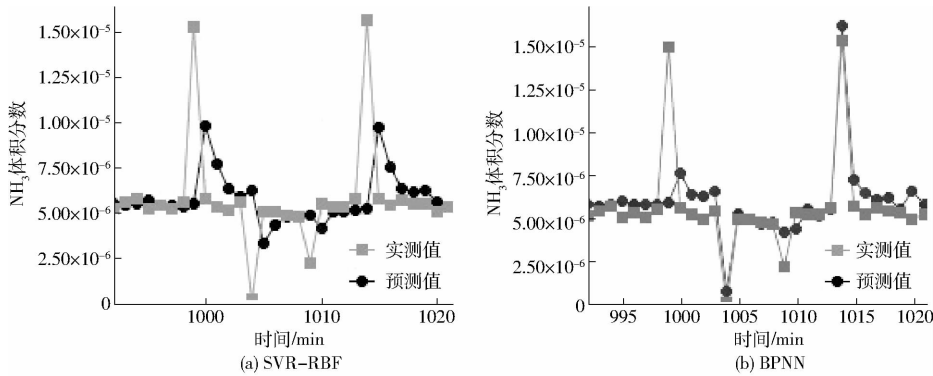


图 3 2 种预测模型单步预测氨气结果

Fig. 3 One-step-ahead prediction results for NH_3 based on SVR - RBF model and BP neural network model

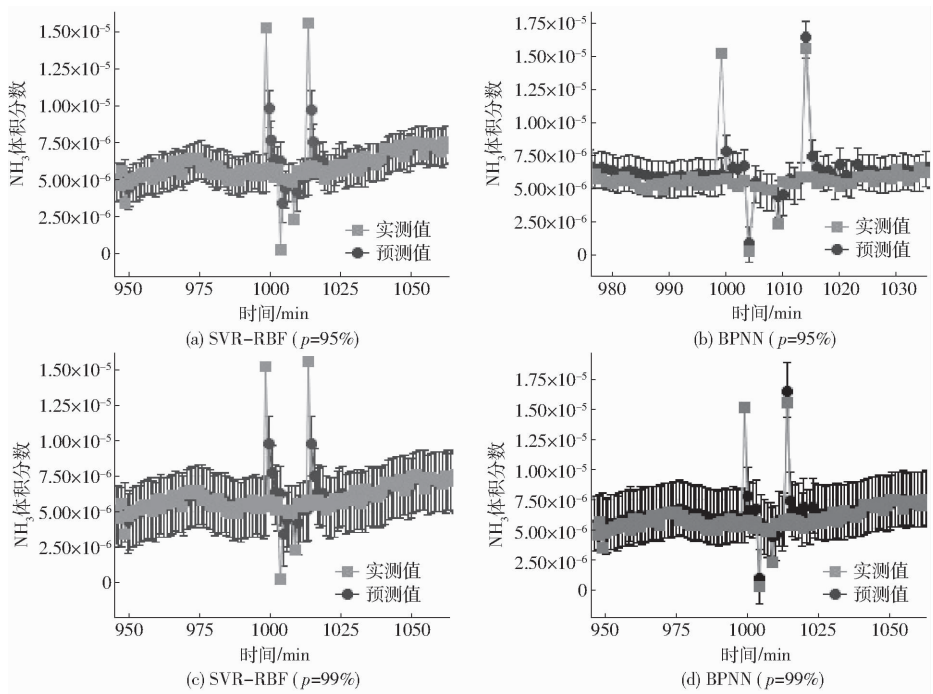


图 4 置信水平分别为 95%、99% 时 2 种模型预测氨气结果

Fig. 4 Predicted results for NH_3 based on SVR - RBF model and BP neural network model in different confidence levels of 95% and 99%

分别采用基于径向基核函数的 SVR 模型和 BP 神经网络模型对种猪场氨气样本数据进行异常检测,采用 TPR、FPR 参数对 2 种预测模型进行效果评估,结果如表 2 所示。

表 2 不同预测区间下检测率和误检率
Tab. 2 TPR and FPR with different PIs %

模型	95% PI		99% PI	
	TPR	FPR	TPR	FPR
BPNN	96. 67	8. 16	88. 60	4. 68
SVR - RBF	98. 33	3. 34	94. 50	2. 50

由图 4 中针对畜禽养殖中氨气数据预测结果以及表 2 中异常数据检测效果可以看出,置信水平由 95% 增长到 99%,检测率和误检率都有所下降,为

防止出现过拟合现象,取置信水平为 95% 较为合理;基于 SVR 模型的数据检测率达 98. 33%,同时误检率低于 BP 神经网络,检测效果较好。

依据滑动窗口计算方法,采用置信水平为 95%,核函数为径向基核函数的 SVR 模型对温度、粉尘、硫化氢等其他畜禽养殖环境因子进行异常检测试验,其中温度、粉尘采集间隔为 5 min,硫化氢采集间隔为 3 min,窗口尺寸分别选取 15、15、24,异常检测率和误检率均在 98% 和 3% 左右,能够满足畜禽养殖物联网数据准确度的要求。

综上所述,基于 SW - SVR 模型的异常数据检测方法针对畜禽养殖物联网中不同环境因子的数据特征周期和采集频率不同的情况,获取最为适合的

窗口尺寸,并且训练所需样本较少,预测拟合效果较好,适用于畜禽养殖物联网数据流中的实时异常数据检测。

3 结论

(1)提出了一种基于 SW-SVR 的异常数据检测方法,利用滑动窗口内的历史数据预测当前时刻数据,并计算预测区间,与实际测量值进行比对从而检测异常数据,实现在线检测畜禽养殖物联网数据

流中的异常数据,有效地保证了物联网传感器中的数据质量;通过对 SVR 算法中两种核函数进行比较,确定最为适合的核函数,达到较理想的预测效果;通过与 BP 神经网络算法比较,预测效果较好,达 98.33%,且耗时较短。

(2)提出了滑动窗口选择方法,依据畜禽物联网采集数据的特征周期以及采集数据频率,计算滑动窗口尺寸,符合预测准确性以及时间效率的要求。

参 考 文 献

- 熊本海,杨振刚,杨亮,等.中国畜牧业物联网技术应用研究进展[J].农业工程学报,2015,31(增刊1):237-246.
XIONG Benhai, YANG Zhengang, YANG Liang, et al. Review on application of Internet of Things technology in animal husbandry in China [J]. Transactions of the CSAE, 2015, 31(Supp. 1): 237-246. (in Chinese)
- 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融合算法[J].农业工程学报,2012,28(23):160-166.
XIONG Yingjun, SHEN Mingxia, LU Mingzhou, et al. Algorithm of real time data fusion for greenhouse WSN system [J]. Transactions of the CSAE, 2012, 28(23): 160-166. (in Chinese)
- 魏芳芳,段青玲,肖晓琰,等.基于支持向量机的中文农业文本分类技术研究[J/OL].农业机械学报,2015,46(增刊):174-179. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2015S029&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2015.S0.029.
WEI Fangfang, DUAN Qingling, XIAO Xiaoyan, et al. Classification technique of chinese agricultural text information based on SVM[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(Supp.): 174-179. (in Chinese)
- CHEN N, DUAN Q, WANG J, et al. Development of early-warning model for intensive pig breeding [M]. Computer and Computing Technologies in Agriculture VIII, Berlin: Springer International Publishing, 2014.
- BRANCH J W, GIANNELLA C, SZYMANSKI B, et al. In-network outlier detection in wireless sensor networks [J]. Knowledge and Information Systems, 2013, 34(1): 23-55.
- ZHANG Y, MERATNIA N, HAVINGA P J M. Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine [J]. Ad Hoc Networks, 2013, 11(3): 1062-1074.
- ANSCOMBE F J, GUTTMAN I. Rejection of outliers [J]. Technometrics, 1960, 2(2): 123-146.
- FREEMAN J. Outliers in statistical data (3rd edition) [J]. Journal of the Operational Research Society, 1995, 46(8): 1034-1035.
- TANG J, CHEN Z, FU W C, et al. A robust outlier detection scheme for large data sets [C] // Pacific-Asia Conference on Knowledge Discovery & Data Mining, 2002: 6-8.
- 孟静,吴锡生.一种基于聚类和快速计算的异常数据挖掘算法[J].计算机工程,2013,39(8):60-63,68.
MENG Jing, WU Xisheng. An outlier data mining algorithm based on clustering and rapid calculation [J]. Computer Engineering, 2013, 39(8): 60-63, 68. (in Chinese)
- 费欢,李光辉.基于 K-means 聚类的 WSN 异常数据检测算法[J].计算机工程,2015,41(7):124-128.
FEI Huan, LI Guanghui. Abnormal data detection algorithm for WSN based on K-means clustering [J]. Computer Engineering, 2015, 41(7): 124-128. (in Chinese)
- HUANG X, ZHAI G, SUI L, et al. Study on the detection of abnormal sounding data based on LS-SVM [J]. Acta Oceanol, 2010, 29(6): 115-120.
- NI Z, WANG F, HU X. Hypercube KNN-based adaptive anomaly detection for wireless sensor networks [C] // Wireless communication and sensor network: Proceedings of the International Conference on Wireless Communication and Sensor Network (WCSN), 2015: 649-657.
- ZHANG R, ZHOU M, GONG X, et al. Detecting anomaly in data streams by fractal model [J]. World Wide Web, 2015, 18(5): 1419-1441.
- PAN D, LIU D, ZHOU J, et al. Anomaly detection for satellite power subsystem with associated rules based on kernel principal component analysis [J]. Microelectronics Reliability, 2015, 55(9): 2082-2086.
- O'REILLY C, GLUHAK A, IMRAN M A, et al. Anomaly detection in wireless sensor networks in a non-stationary environment [J]. IEEE Communications Surveys & Tutorials, 2014, 6(3): 1613-1432.
- 何欢.基于概念漂移的异常检测技术研究[D].成都:电子科技大学,2015.
- HILL D J, MINSKER B S. Anomaly detection in streaming environmental sensor data: a data-driven modeling approach [J]. Environmental Modelling & Software, 2010, 25(9): 1014-1022.

- 19 苑进,胡敏,WANG Kesheng,等. 基于高斯过程建模的物联网数据不确定性度量与预测[J/OL]. 农业机械学报,2015,46(5):265-272. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150538&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2015.05.038.
YUAN Jin, HU Min, WANG Kesheng, et al. Uncertainty measurement and prediction of iot data based on Gaussian process modeling[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2015,46(5):265-272. (in Chinese)
- 20 NELLO C, JOHN S T. Support vector machines[M]. Cambridge: Cambridge University Press, 2000.
- 21 WANG X, PARDALOS P M. A survey of support vector machines with uncertainties[J]. Annals of Data Science, 2014, 1(3): 293-309.
- 22 SUJAY R N, DEKA P C. Support vector machine applications in the field of hydrology: a review[J]. Applied Soft Computing, 2014, 19(6):372-386.
- 23 YUKIMASA K, HIROSHI M. Sliding window-based support vector regression for predicting micrometeorological data[J]. Expert Systems with Applications 2016,59: 217-225.
- 24 曾绍华,魏延,唐远炎. 剔除支持向量回归中异常数据算法[J]. 重庆大学学报:自然科学版,2012,35(12):120-132.
ZENG Shaohua, WEI Yan, TANG Yuanyan. Algorithm of removing outliers in SVR[J]. Journal of Chongqing University: Natural Science Edition, 2012,35(12):120-132. (in Chinese)
- 25 李丹玲,陈平雁,周凤麒. 基于线性 ν -支持向量回归机的异常数据检测[J]. 数理统计与管理,2011,30(1):59-63.
LI Danling, CHEN Pingyan, ZHOU Fengqi. Outlier detection based on linear ν -SVRM[J]. Journal of Applied Statistics and Management,2011,30(1):59-63. (in Chinese)
- 26 汪旭颖. 基于支持向量回归机的油田异常井预警模型研究[D]. 大庆:东北石油大学,2015.
- 27 SONG S X, ZHANG A Q, WANG J M, et al. SCREEN: stream data cleaning under speed constraints[C]//Proceedings of the 2015 ACM SIGMOD Conference, 2015:827-841.
- 28 CAO L, YANG D, WANG Q, et al. Scalable distance-based outlier detection over high-volume data streams[C]//2014 IEEE 30th International Conference on Data Engineering (ICDE), 2014:76-87.
- 29 刘京,常庆瑞,刘森,等. 基于 SVR 算法的苹果叶片叶绿素含量高光谱反演[J/OL]. 农业机械学报,2016,47(8):260-265,272. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160834&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.08.034.
LIU Jing, CHANG Qingrui, LIU Miao, et al. Chlorophyll content inversion with hyperspectral technology for apple leaves based on support vector regression algorithm[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2016,47(8):260-265,272. (in Chinese)
- 30 孙德山. 支持向量机分类与回归方法研究[D]. 长沙:中南大学, 2004.
- 31 郭虎升,王文剑. 动态粒度支持向量回归机[J]. 软件学报,2013,24(11):2535-2547.
GUO Husheng, WANG Wenjian. Dynamical granular support vector regression machine[J]. Journal of Software,2013,24(11): 2535-2547. (in Chinese)