

带有同步预测的 WBAN 时序数据融合算法

王汝言, 翟美玲, 吴大鹏

(重庆邮电大学 宽带泛在接入技术研究所, 重庆 400065)

摘要: 提出一种带有同步预测的时序数据融合算法, 利用多分辨率分析特性对采集的原始数据进行预处理, 挖掘反映人体生理状态的本质特征, 进而采用同步预测机制在感知节点和汇聚节点处分别建立轻量级预测模型, 消除网内冗余数据的传输以降低能耗。结果表明所提出的融合算法具有较高的预测精度, 能够实现低开销的无线体域网时序数据融合。

关键词: 无线体域网; 数据融合; 多分辨率分析; 最小二乘支持向量机; 同步预测机制

中图分类号: TP393

文献标识码: A

Time series data aggregation algorithm with synchronous prediction for WBAN

WANG Ru-yan, ZHAI Mei-ling, WU Da-peng

(Broadband Ubiquitous Network Research Laboratory, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Due to the nonlinearity and nonstationarity of the physiological data sensed by WBAN, data aggregation cannot be effectively achieved according to the time-domain trend of data. Therefore, a novel time series data aggregation algorithm with synchronous prediction was proposed. By preprocessing the original sensing data with the multi-resolution analysis, the inherent characteristics of the physiological data can be obtained to establish a light-weight synchronous prediction model at both the sensor and sink. Numerical results show that the proposed aggregation algorithm can achieve a favorable prediction precision and a low energy consumption rate by eliminating the in-network data redundancy.

Key words: wireless body area network; data aggregation; multi-resolution analysis; LSSVM; synchronous prediction mechanism

1 引言

微型智能医学传感器、无线网络技术以及普适计算的日趋成熟促进了无线体域网 (WBAN, wireless body area network) 的快速发展。WBAN 是由可穿戴或植入式传感器节点构成的、以人体为中心的网络^[1,2], 可以随时随地监控人体生理状况, 在远程医疗^[3,4]、健康恢复^[5,6]、运动监控^[7,8]以及多媒体通信^[9,10]等领域具有广泛的应用价值。

WBAN 传感器节点尺寸很小, 电池容量低且难以替换, 因此如何高效利用有限能量来延长整个网络的生命期成为 WBAN 的关键技术。WBAN 的能量主要消耗在数据传输、信号处理和硬件操作 3 个方面, 其中数据传输过程耗能最大^[11~13]。同时, 节点采集的数据具有一定的时间冗余度, 冗余数据的传输加剧了通信能量的消耗。高效的数据融合算法能够显著减少冗余数据的传输, 降低数据冲突, 减轻网络拥塞, 有效节省能源开销, 显著延长网络生

收稿日期: 2014-07-17; 修回日期: 2015-02-18

基金项目: 国家自然科学基金资助项目 (61371097); 重庆市自然科学基金资助项目 (CSTC2013JJB40001, CSTC2013JJB40006); 重庆邮电大学青年自然科学基金资助项目 (A2012-93); 重庆市青年科技人才培养计划基金资助项目 (CSTC2014KJRC-QNRC40001)

Foundation Items: The National Natural Science Foundation of China (61371097); The Natural Science Foundation of Chongqing (CSTC2013JJB40001, CSTC2013JJB40006); The Foundation of Chongqing University of Posts and Telecommunications (A2012-93); Youth Talents Training Project of Chongqing Science & Technology Commission (CSTC2014KJRC-QNRC40001)

存时间^[14,15]。显然,对于 WBAN 来说,研究高效的数据融合策略对提高网络性能、充分利用网络资源十分关键。

数据融合技术是 WBAN 研究热点之一,目前,大部分的研究工作都是围绕着提高行为分类精确度而展开^[16,17]。文献[18]考虑了全局冲突情况下节点间感知信息的相互影响,利用动态冲突信誉值和新的组合规则模拟行为的不确定性,构建感知信息证据理论融合模型,用于解决环境辅助生活系统(AAL, ambient assisted living)中行为识别冲突问题。该融合方法提高了行为识别的有效性和准确性。文献[19]定义了一种新型的协作无线体域网体系结构(CBSN, collaborative body sensor networks)为多传感器数据融合以及普适计算的应用提供了一种灵活的框架。该框架能够有效支持实时数据处理,实现协作体域网对联合事件的精确分类和检测,解决了复杂应用中单一体域网不能满足系统要求的问题。

WBAN 数据融合在行为分类方面的研究^[16-19]普遍由汇聚节点执行数据分析和融合操作,需要建立在对大量数据进行训练的基础上才能确保较高的识别精度。WBAN 的汇聚节点资源丰富,能够对复杂的信号进行分析和处理,可以进行复杂的分类学习或模式识别。但是,在汇聚节点进行数据融合并不能减少 WBAN 网络内部的数据流量,并且为了确保行为分类的准确性,传感器节点的感知数据要全部发送给汇聚节点,反而加剧了数据传输能耗。

WBAN 节能策略仍然是中外研究者面临的一个挑战,降低数据传输能耗对延长网络生命周期至关重要。文献[20]采用了基于数据融合的功率优化技术,通过增大共享路径上传输数据总量,减少每比特的传输能量,从而降低 WBAN 的传输能耗。但是该数据融合只是通过设定发送时间限制不同节点的数据融合,并没有对冗余数据进行处理,因此无法适用于对数据实时性要求较高的 WBAN 应用场景。文献[21]介绍的无线传感器网络中基于自回归预测、移动平均预测、指数平滑预测的数据融合算法,文献[22]提出的无线传感器网络中基于灰色支持向量机预测模型的数据融合算法,都能降低网络能量消耗,但是预测算法不能很好地拟合体域网采集的非线性数据,且复杂的算法不适用于体域网。

本文提出一种适用于 WBAN 的基于小波变换多分辨率分析和最小二乘支持向量机的时序预测

数据融合算法(WT-LSSVM, wavelet transform and least squares support vector machine based algorithm)。首先,对感知的原始数据进行基于小波变换多分辨率分析的预处理,得到表征数据序列变化趋势的低频分量和表征数据序列波动性的高频分量。然后,节点和汇聚节点采用同步预测机制,根据各分量的特点分别建立轻量级预测模型。最后,将各分量的预测值重构得到最终预测结果,实现时序数据融合。在传感器节点处进行数据融合,不仅能满足 WBAN 应用对感知数据实时性的要求,而且能够减少网络内部冗余数据的传输,达到降低传输能耗、延长网络生命周期的目的。

2 数据预处理

目前, WBAN 主要应用于医疗领域,这种应用场景要求 WBAN 保持对突发事件产生的异常数据具有较高的敏感度。由于受到人体呼吸、心跳、自身活动等生理因素以及周围环境中电子电气设备的干扰, WBAN 设备会产生不可避免的噪声^[23],同时外界温度、湿度的动态变化也将影响感知数据的准确度。综合多种因素的影响, WBAN 采集的各种生理数据序列既具有非平稳性、非线性以及波动性的特点,又具有特殊的周期性,类似多个不同频率序列的叠加^[2]。直接对原始数据序列建立预测模型难以准确地描述数据序列的本质特性,不利于从感知数据中寻找具有生理意义或代表生理状态的特征值。因此,本文通过小波变换对 WBAN 所感知的原始数据序列进行预处理,以提高预测精度,降低传输数据冗余度。

小波变换通过伸缩和平移等运算对数据序列进行多尺度细化分析^[24],在时域和频域有效地表征数据序列的局部特征,能敏锐地探测正常生理参数中夹带的瞬间反常现象并展示其成分,可用于 WBAN 生理数据的分析,检测异常数据的产生。本文利用小波变换对数据序列进行逐步精细的多分辨率分析,得到较平滑的数据序列,便于历史数据特性的挖掘。

设 $\psi(t) \in L^2(\mathbb{R})$, $L^2(\mathbb{R})$ 表示定义在实数集 \mathbb{R} 上所有平方可积函数的集合。当 $\psi(t)$ 的傅里叶变换 $\psi(\omega)$ 满足式(1)所示的条件时,称 $\psi(t)$ 为一个基本小波或母小波。

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty \quad (1)$$

小波变换的基本原理就是将母小波 $\psi(t)$ 通过平移因子 b 和伸缩因子 a 进行伸缩和平移后产生一个函数族 $\{\psi_{a,b}(t)\}$, 如式(2)所示, 称其为小波基函数。

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), a, b \in R \quad (2)$$

应用小波变换对 WBAN 采集的原始数据序列进行多分辨率分析的实质就是将所采集的数据序列 $f(t)$ 表示为函数族的加权和。因此, 能量有限的原始数据序列函数 $f(t)$ 的连续小波变换定义为

$$W_f(a,b) = \int f(t)\psi_{a,b}(t)dt = |a|^{-\frac{1}{2}} \int_R f(t)\psi\left(\frac{t-b}{a}\right)dt \quad (3)$$

经过小波变换可以得到生理数据序列不同频段上的分量。节点可采用式(4)所示的重构算法对分解所得各系数进行单支重构, 得到不同频段上的概貌序列与细节序列以表征 WBAN 生理数据序列内在特性。

$$f(t) = \iint \frac{1}{a^2} W_f(a,b) \psi_{a,b}(t) da db \quad (4)$$

然而, 应用小波变换实现 WBAN 数据预处理需要解决 2 个问题: 小波基的选取和分阶层的确定。本文采用 db3 (daubechies3) 小波对 WBAN 感知数据进行处理, 其对异常数据比较敏感能突出生理数据中异常数据的特征, 尽量减少数据的失真, 以保证序列重构的精确度以及时频分析的良好性能。小波分解阶数决定了数据分解后频带的细节程度, 分解阶数越大, 则频率分量所包含的细节信息量越大, 合理地选取分解阶数能够显著提高数据处理效率。但过大的分解阶数会增加算法计算量, 甚至造成信号某些重要局部奇异性丢失。结合 WBAN 特点考虑, 本文对生理数据序列进行 3 级分解, 确保所提算法以较小的计算量得到比较平滑的子序列。通过 db3 小波基对生理参数序列进行 3 级分解, 得到低频逼近数据和各高频细节数据并分别进行单支重构, 得到原始数据序列的概貌序列和细节序列, 用于后期预测模型的建模及预测。本文通过小波变换对 WBAN 所感知的原始数据序列进行预处理, 通过逐步精细的多分辨率分析, 得到较平滑的数据序列。如此得到的子序列所包含的特征信息相对简单, 便于历史数据特性的挖掘, 降低了数据融合建模难度, 可节约融合开销。

3 轻量级预测模型

WBAN 为了达到人体对舒适度的要求, 节点尺寸非常小且具有有限的存储能力、计算能力、无线通信能力, 供电电池也难以替换和充电。因此要推广 WBAN 的应用系统, 必须设计轻量级、高效节能的数据融合算法, 以解决网络内大量感知数据消耗过多能量的问题, 增强 WBAN 的可用性。结合 WBAN 传感器节点的特点, 本文构建了基于最小二乘支持向量机 (LSSVM, least squares support vector machine) 的预测模型, 以实现 WBAN 的数据融合。

WBAN 感知的人体生理数据是非线性的。WBAN 实际应用场景中, 要求能灵敏地识别感知数据的变化, 且能精确识别并处理异常情况产生的突发数据。LSSVM 是在 SVM 的基础上发展而来, SVM 需要求解二次凸优化问题, 计算量为 $O(N^3)$, 其计算复杂度会随着样本集合大小 N 的增大而迅速增加。LSSVM 用等式约束替换 SVM 优化问题中的不等式约束问题, 通过求解线性方程组就可得到 LSSVM 模型, 其计算复杂度为 $O(N^2)$, 而且通过核函数的优化可以大大降低 LSSVM 的计算复杂度。因此, 该算法满足 WBAN 对预测模型低计算复杂度的要求。将 LSSVM 预测模型引入到 WBAN 实现时序数据融合, 既能对人体异常情况产生的非线性突发数据实现快速、准确预测, 又能有效地减少网络内的数据通信量, 达到节能的效果。

为了使提出的算法能够在趋势预测的基础上准确地逼近实际人体的感知数据, 需要对 LSSVM 进行训练。此处, WBAN 传感器节点将原始感知样本数据传递给 LSSVM 进行训练。设训练样本的输入输出对为 $(x_i, y_i), i=1, 2, \dots, n, x \in R^d, y \in R, x_i$ 表示 WBAN 实际感知数据, y_i 表示模型回归数据。由于 WBAN 感知数据在原时域空间存在线性不可分的问题, 通过式(5)所描述的由非线性映射 $\phi(\cdot)$ 构造的最优线性回归函数, 可将样本数据从原始空间非线性映射到高维特征空间, 寻求一个可以把样本划分且分类间隔最大的最优超平面。

$$f(x) = \omega^T \phi(x) + b \quad (5)$$

其中, ω 为高维特征空间的权值向量, b 为偏置量。通过上述非线性映射, 就把寻找最优分类线的问题转化成了寻求最优超平面的问题, 即将低维感知数据的非线性回归问题转化成了高维特征空间中的

线性回归问题。

再将 WBAN 感知数据的回归问题根据结构风险最小化 (SRM, structural risk minimization) 原则转化为有约束的二次优化问题, 如式 (6) 所描述。

$$\begin{cases} \min \Phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} c \sum_{i=1}^n \xi_i^2 \\ \text{s.t. } y_i = \omega^T \phi(x_i) + b + \xi_i \quad (i=1, 2, \dots, n) \end{cases} \quad (6)$$

其中, $c > 0$ 表示可调惩罚系数即模型平衡系数; ξ_i 表示回归生理参数值与实际感知值的误差, 用来调节超平面分类错误的点数, 使其尽量达到最小值。将式(6)有约束优化问题转化为无约束优化问题, 对应的拉格朗日函数为

$$L = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} c \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n a_i (\omega^T \phi(x_i) + b + \xi_i - y_i) \quad (7)$$

其中, a_i 表示拉格朗日乘子。由式(7)极值存在必要条件, 即 L 对各变量 ω, b, ξ_i, a_i 的偏导数为零, 可得如下约束条件

$$\begin{cases} \omega = \sum_{i=1}^n a_i \phi(x_i) \\ \sum_{i=1}^n a_i = 0 \\ a_i = c \xi_i \quad (i=1, 2, \dots, n) \\ \omega^T \phi(x_i) + b + \xi_i - y_i = 0 \quad (i=1, 2, \dots, n) \end{cases} \quad (8)$$

消除式(8)中变量 ω 和 ξ_i , 得到如式(9)所示的矩阵方程

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Omega + \mathbf{I}/\gamma \end{bmatrix}_{(n+1) \times (n+1)} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (9)$$

其中, $\mathbf{y} = [y_1, \dots, y_n]^T$, $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]^T$, \mathbf{I} 为 $n \times n$ 阶单位矩阵, Ω 为 $n \times n$ 阶矩阵, 且

$$\Omega_{ij} = \phi(x_i)^T \phi(x_j) \quad (10)$$

为了满足 WBAN 对算法低计算复杂度的要求,

本文选择参数较少的径向基 (RBF, radial basis function) 函数作为 LSSVM 的核函数, 由式 (11) 所示的内积核函数 $K(x, x_i)$ 替代最优分类平面里计算过程中的点积。

$$K(x, x_i) = e^{-\|x-x_i\|^2 / (2\sigma^2)} \quad (11)$$

进而, 通过最小二乘法得出参数 $[b, a]^T$ 的估计, 最后得到如式(12)所示的 LSSVM 预测模型。

$$y(x) = \sum_{i=1}^n a_i K(x, x_i) + b \quad (12)$$

数据预处理阶段所得的高频和低频分量分别按上述方式进行样本训练建立预测模型。完成训练后, 将不同频率的数据序列传递到相应预测模型的输入端, 依据式(12)所得的预测模型对系统下一时刻的时序数值进行单步预测, 将预测结果与下一时刻的实际感知数据进行比较来决定是否需要发送数据。当预测误差小于用户容忍误差阈值时, WBAN 传感器节点不发送数据, 减少了冗余数据的传输, 为 WBAN 网络节省了大量能量。本文的数据融合模型如图 1 所示。

4 同步预测数据融合机制

需要设计合理的融合机制确保远端服务器实时地接收 WBAN 的感知数据, 掌握人体生理状况。本文提出了同步预测机制, 在 WBAN 节点和汇聚节点同步建立相同的 LSSVM 预测模型, 并保持用于预测的数据序列同步, 即节点和汇聚节点在每个采样周期将相同数据序列输入相同的预测模型, 时时保持预测过程和预测结果的同步。节点采用小波变换和 LSSVM 混合模型对下一时刻的数据进行单步预测, 并将预测数据与下一时刻的实际感知数据进行比较。若预测误差在用户设定的容忍阈值范围内, 则认为该时刻采集的数据具有较强的时间冗余特征, 发送此类数据将造成不必要的能量消耗, 此

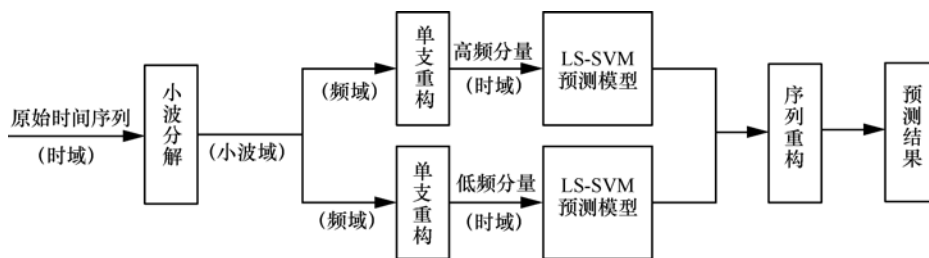


图 1 融合模型

时, 传感器节点将预测数据保存, 不需发送数据至其他节点; 若预测误差超出容忍阈值范围, 传感器节点则向汇聚节点发送实际感知数据。同时汇聚节点利用相同的预测数据序列对下一时刻数据预测后, 在一定时间内若没有收到节点发送的相应实测数据, 则表明预测成功, 并将该预测数据作为实际感知数据进行保存。

当 WBAN 节点在连续的感知周期内均成功预测下一刻的生理数据时, 汇聚节点在相应的连续周期中则无法接收到该节点发送的数据, 此时该传感器节点类似于失效节点。而 WBAN 的应用一般要求各节点采集的数据能够真实且实时地反映人体生理特征, 汇聚节点若长时间无法接收到节点发送的数据, 则难以实现实时监控, 不能对 WBAN 用户状态做出准确的分析和精确的判断。因此, 本文定义变量连续成功预测次数阈值 s 对连续成功预测次数进行限制, 当连续成功预测次数超过 s 时, 节点就发送实测数据给汇聚节点, 以证明节点的有效性。带有同步预测机制的融合过程具体如下。

预测之前, 汇聚节点根据应用要求定义全局阈值, 预测误差阈值 ε 及连续成功预测次数阈值 s , 并发送给相应的传感器节点。传感器节点产生 2 个数据序列, 实测数据序列 (A) 和预测值序列 (P_n), 分别存储实际测量值和用来预测的数据序列, 汇聚节点将用于预测的数据序列存储在 P_s , 且 $P_s = P_n$ 。初始阶段节点根据预测算法的要求进行数据采集, 将前 l 个感知周期采集的原始数据发送给汇聚节点, 作为预测模型的训练样本。此时节点 i 各序列值为 $A(i) = P_n(i) = P_s(i) = \{x_1, x_2, \dots, x_l\}$ 。

在第 $l+1$ 个感知周期, WBAN 中的传感器节点和汇聚节点利用相同的预测数据序列和模型预测。令 x_{l+1} 、 x'_{l+1} 分别代表实测数据和利用 P_n 得到的预测值, 汇聚节点由 P_s 得预测值 x'_{l+1} 。若 $\text{abs}(x'_{l+1} - x_{l+1}) < \varepsilon$, 则表明预测误差在用户可接受的阈值范围内, 预测成功, 那么传感器节点不发送实测值 x_{l+1} 给汇聚节点, 汇聚节点把预测值 x'_{l+1} 保存为实际感知数据, 保持其预测数据序列 P_s 与节点的 P_n 同步。此时, 节点 i 的各序列更新为 $A(i) = \{x_2, x_3, \dots, x_{l+1}\}$, $P_n(i) = \{x_2, x_3, \dots, x'_{l+1}\}$ 。

当预测失败或连续成功预测次数 r 超过预定值 s 时, 传感器节点就会发送实测数据 x_{l+1} 给汇聚节点, 然后各序列更新为 $A(i) = \{x_2, x_3, \dots, x_{l+1}\}$,

$$P_n(i) = \{x_2, x_3, \dots, x_{l+1}\}, \quad P_s(i) = \{x_2, x_3, \dots, x_{l+1}\}。$$

汇聚节点在连续的 s 个周期内收到 WBAN 传感器节点的实测数据, 则表明该节点处于正常工作状态, 否则就视其为失效节点。本文所提融合算法的整体过程如下。

Input:

$Y^{(0)}$: Original data sequence

ε : The threshold prediction error

r : The number of consecutive successful predictions

s : The threshold number of consecutive successful predictions

Output:

$\hat{y}(t+1)$: Next prediction data sequence

$y(t+1)$: Next sense data sequence

- 1) for (l : collection) do
- 2) WT data process
- 3) get a_j and $d_1, d_2, \dots, d_j // a_j$ 指低频近似序列; d_1, d_2, \dots, d_j 高频细节序列
- 4) end for
- 5) for each a_j and d_1, d_2, \dots, d_j do
- 6) LSSVM prediction // 分别进行预测
- 7) addition to reconstitution // 预测结果重构
- 8) get $\hat{y}(t+1)$
- 9) end for
- 10) $pe = \text{abs}(x'_{l+1} - x_{l+1})$
- 11) if $pe < \varepsilon$ then
- 12) if $r < s - 1$ then
- 13) save $\hat{y}(t+1)$
- 14) else go to Line 15)
- 15) else
- 16) send $y(t+1)$
- 17) end if

5 算法性能分析

5.1 预测模型评价指标

为了衡量本文所提融合算法中预测模型的预测精度, 即 WT-LSSVM 预测模型产生的预测值与实际感知值拟合程度的优劣, 采用如式 (13) 所示的均方根误差 (RMSE) 作为检验指标。其中, $\hat{y}(t)$ 为序列预测值, $y(t)$ 为序列实测值, N 为样本数量。RMSE 不存在预测误差正负抵消的问题, 能很好地描述预测值与实测值之间的误差情况, 对本文所提

预测模型的预测效果做出准确的整体性能评价。

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2} \quad (13)$$

5.2 性能分析

本文采用 MIT/BIH 数据库的心电信号数据对所提出的融合算法进行验证。MIT/BIH 数据库由美国麻省理工学院提供, 是国际上公认的可作为标准的心电数据库之一。该数据库共有 48 组心电图数据, 每个数据由后缀分别为 .atr、.hea、.dat 的 3 个文件组成。 .dat 文件同时存储一个记录的两道心电数据, 且不能直接读取, 要对数据格式进行转换。图 2 描述了 MIT 数据库 105.dat 的一段心电信号波形, 原始心电信号数据具有很大的波动性。

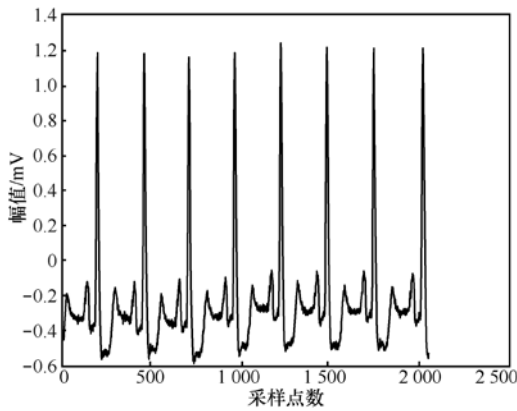


图 2 心电图原始信号

采用多分辨率分析对心电信号数据进行预处理, 有助于削弱环境及人体因素的影响。经 db3 小波基对原始数据序列进行 3 级分解, 得到如图 3 所示的低频

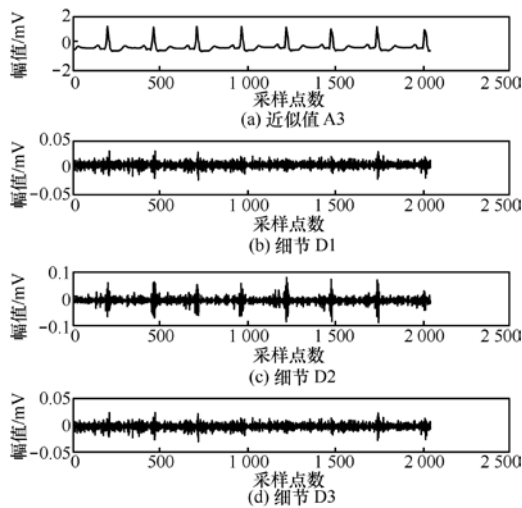


图 3 小波分解

逼近序列 a_j 和 高频细节序列 d_1, d_2, \dots, d_j 。得到的子序列改善了平滑度, 更利于特征信息的提取。

5.2.1 融合算法有效性验证

为验证本文所提算法的有效性, 取 105.dat 前 500 个心电图数据作为实验样本对所提的 WT-LSSVM 数据融合算法进行验证分析。预处理得到心电图低频逼近序列和各高频细节序列, 各取前 20 个数据作为训练样本, 通过数据训练分别建立 LSSVM 预测模型, 对余下的 480 个数据进行预测, 将各模型的预测值通过等权加和序列重构得最终预测结果。如图 4 所示, WT-LSSVM 的预测结果在变化趋势和数值上都很好地拟合了原始数据, 预测的均方根误差仅为 0.0529。原始数据经过预处理得到不同频段较平滑的数据序列, 便于预测模型的建立, LSSVM 能有效地处理生理数据的非线性问题, 达到较好的预测效果。

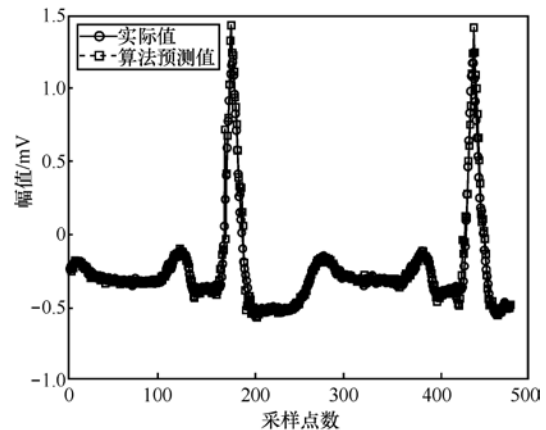


图 4 WT-LSSVM 预测效果

为进一步说明 WT-LSSVM 算法的有效性, 扩大预测样本空间, 对 2 028 个数据进行预测, 如图 5 所示, 预测结果仍然很好地拟合了实测值。

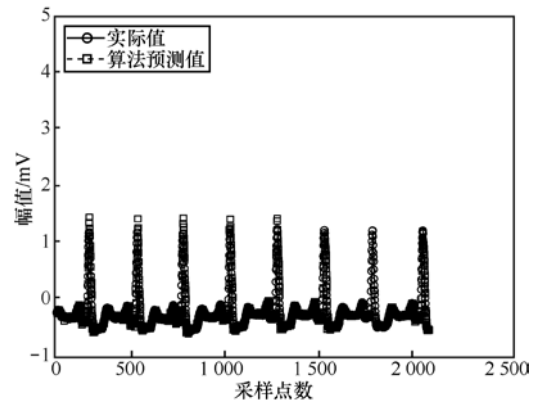


图 5 105.dat 的 WT-LSSVM 预测效果

从 MIT/BIH 数据库随机取另一组数据，如 121.dat，得到如图 6 所示的预测结果，预测均方根误差为 0.051 8，预测精度较高。综上，表明本文所提 WT-LSSVM 数据融合算法可以利用小样本数据进行训练，得到预测效果良好的模型，适用于 WBAN 的数据处理。

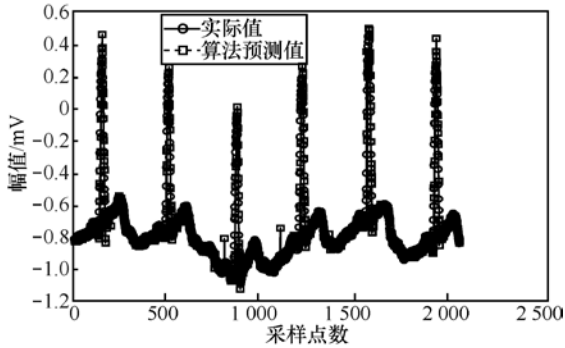


图 6 121.dat 的 WT-LSSVM 预测效果

5.2.2 融合算法准确性验证

预测误差阈值代表了该应用可接受的预测误差范围，阈值越小，可接受的误差范围越小，对预测精度的要求越高；阈值越大，可接受的误差范围越大，对算法预测精度的要求也相对较低。为验证 WBAN 中 WT-LSSVM 融合算法的准确性，本文设置不同的误差阈值，观察预测成功率。如图 7 所示，设置预测误差阈值取值为 0~0.05 mV，随着预测误差阈值的逐渐增大，预测成功率也相应的增大。当误差阈值 ε 在 0~0.01 范围内时，预测成功率随着阈值的增大快速上升， $\varepsilon=0.01$ 时预测成功率为 44.17%。 ε 超过 0.01 时，成功率上升趋势逐渐变得缓慢，因为心电信号微弱，数据取值本身就很小，误差值也就很小。直至 $\varepsilon=0.05$ 时，预测成功率为 88.75%。

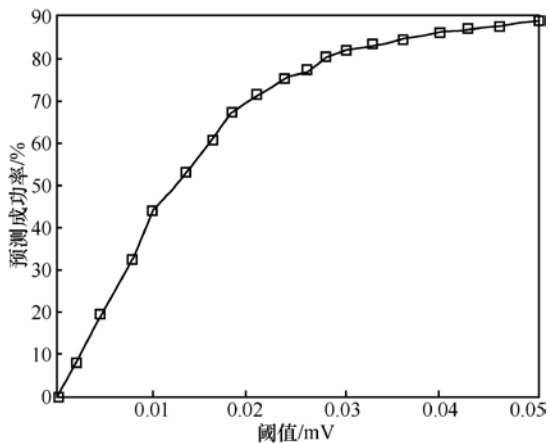


图 7 不同阈值时的预测成功率

图 8 给出了预测误差的累积概率分布。结果表明，预测误差小于 0.2 mV 的概率为 80.0%。可见，本文所提数据融合算法具有较高的准确度。

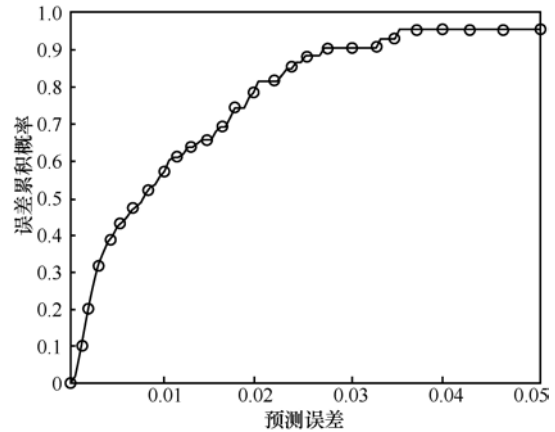


图 8 融合算法预测准确性验证

5.2.3 融合算法的能量有效性

本文采用基于预测的数据融合，减少冗余数据的传输为网络节省能量。定义通信能耗增益为

$$\eta = \frac{E_{DA}}{E} \quad (14)$$

E_{DA} 表示采用融合算法产生的通信能耗， E 为不采用融合算法时的通信能耗。由于该融合算法没有引起额外的通信开销，因此网内的通信能耗与数据通信量成正比关系，通过对比 2 种情况下网内通信量就可以计算网络能耗增益值。不同误差阈值时的能耗增益，如图 9 所示。误差阈值 ε 取值 0~0.05 范围时，通信能量消耗由 100% 降低到 18.96%，即可为传感器节点节省 81.04% 的能量。采用本文的数据融合算法，根据实际应用对数据精度的要求选取适当的阈值，则可以降低节点的通信能耗。

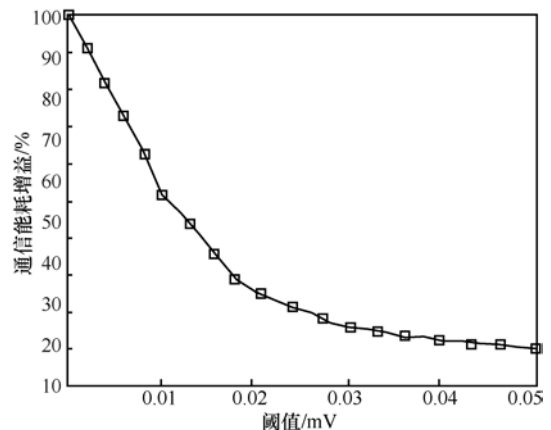


图 9 不同阈值时的通信能耗增益

5.2.4 融合算法对比分析

无线传感器网络基于预测的时间序列融合算法已经取得了很好的研究成果,为 WBAN 数据融合算法的研究提供了重要的参考和理论依据。为验证本文所提算法在 WBAN 中的实用性,利用 WT-LSSVM、文献[25]提出的 SVM、文献[26]提出的 LSSVM 及文献[27]提出的卡尔曼滤波(KF, Kalman filter) 4种数据融合算法,分别对 105.dat 前 500 个心电图数据进行训练、建模和预测,得到如图 10 所示的预测结果对比。各融合算法的均方根误差(RMSE)及平均每次预测运行时间如表 1 所示。

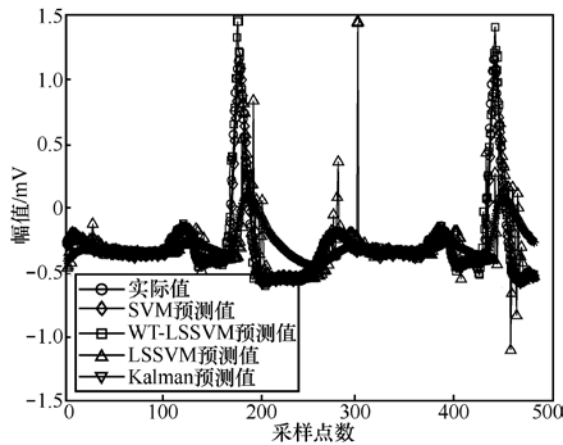


图 10 4种算法预测效果对比

表 1 4种算法预测结果分析

数据融合算法	每次预测耗时/s	RMSE	能耗增益/%
WT-LSSVM	0.838 4	0.052 9	57.708 3
LSSVM	0.305 5	0.363 3	69.583 3
SVM	4.922 6	0.044 7	53.750 0
Kalman	0.001 5	0.265 6	84.583 3

图 10 中, WT-LSSVM 的预测结果能很好地拟合原始数据序列的数值变化趋势,并能准确地逼近实测值大小。设置误差阈值 $\varepsilon=0.01$ 时, LSSVM 的预测成功率为 30.83%, SVM 为 46.46%, Kalman 仅为 18.33%, 而 WT-LSSVM 为 44.17%。

WT-LSSVM 要对 4 个不同频率分量分别进行建模及预测, 而 LSSVM 只需进行一次建模。因此表 1 中, WT-LSSVM 每次预测耗时相对较长, 为 LSSVM 的 2.36 倍。但是经过多分辨率分析所得的子序列更加平滑, 降低了建模难度, 提高了模型预测精度, 其均方根误差降低了 6.91 倍。WT-LSSVM

的均方根误差与 SVM 相近, 但预测耗时却明显缩短, 因为 SVM 计算复杂度较大耗时长。本文采用的 WT-LSSVM 是对 SVM 的改进, 降低了算法的计算复杂度, 提高了求解速度及预测精度, 更适用于能量和计算能力都受限的 WBAN。Kalman 在预测过程中, 通过状态方程的迭代递推出滤波器的最佳状态变量, 能够实现对线性动态系统的最佳估计。对于非线性生理数据, Kalman 预测只能拟合实测数据的变化趋势, 不能准确地逼近实测值, 对实测值大小的拟合情况很差, 其均方根误差为 0.265 6, 预测误差较大, 预测精度不能满足 WBAN 应用要求。

如表 1 所描述, 4 种基于预测的融合算法都能减少网络内部不必要的数据传输, 不同程度地降低数据通信产生的能量消耗。误差容忍范围为 0~0.01 mV 时, WT-LSSVM 算法产生的能量消耗仅为原始状态的 57.7%, 为 WBAN 节省了 42.3% 的通信能量。虽然 SVM 算法节省的能量高达 46.2%, 但是 SVM 计算复杂, 数据处理的能量消耗远远高于 WT-LSSVM。

因此, 本文所提的 WT-LSSVM 数据融合算法更加适用于能量和计算能力都受限的 WBAN。

6 结束语

WBAN 由于受到外界温度、湿度、人体运动状态等多种因素的影响, 所采集的数据多具有非平稳性、非线性的特点, 利用小波变换对原始数据序列进行多尺度细化分析能够挖掘其本质特性, 提高对原始数据序列的预测精度。经小波变换处理后, 对重构后的高低频分量分别进行 LSSVM 预测, 完成数据融合。此算法不仅能很好地解决小样本和非线性等问题, 并且满足 WBAN 传感器节点对低算法复杂度的要求。实验表明本文所提的 WT-LSSVM 数据融合算法能够较好地拟合原始数据序列, 大大减少 WBAN 网内数据通信量, 降低传感器节点的能量消耗。

参考文献:

- [1] ANZAI D, KATSU K, CHAVEZ-SANTIAGO R, *et al.* Experimental evaluation of implant UWB-IR transmission with living animal for body area networks[J]. IEEE Transactions on Microwave Theory and Techniques, 2014, 62(1):183-192.
- [2] IMTIAZ S A, CASSON A, RODRIGUEZ-VILLEGAS E. Compression in wearable sensor nodes: impacts of node topology[J]. IEEE Transactions on Bio-medical Engineering, 2014, 61(4):1080-1090.

- [3] PANTELOPOULOS A, BOURBAKIS N G. Prognosis—a wearable health-monitoring system for people at risk: methodology and modeling[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2010, 14(3): 613-621.
- [4] JUNNILA S, KAILANTO H, MERILAHTI J, *et al.* Wireless, multi-purpose in-home health monitoring platform: two case trials[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2010, 14(2): 447-455.
- [5] CORCHADO J M, BAJO J, TAPIA D I, *et al.* Using heterogeneous wireless sensor networks in a telemonitoring system for healthcare[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2010, 14(2): 234-240.
- [6] CHENG S H, HUANG C Y. Coloring-based inter-WBAN scheduling for mobile wireless body area networks[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2013, 24(2): 250-259.
- [7] GHASEMZADEH H, JAFARI R. Physical movement monitoring using body sensor networks: a phonological approach to construct spatial decision trees[J]. *IEEE Transactions on Industrial Informatics*, 2011, 7(1): 66-77.
- [8] BÄCHLIN M, TRÖSTER G. Swimming performance and technique evaluation with wearable acceleration sensors[J]. *Pervasive and Mobile Computing*, 2012, 8(1): 68-81.
- [9] WANG W, WANG C, ZHAO M. Resource optimized TTSH-URA for multimedia stream authentication in swallowable-capsule based wireless body sensor networks[J]. *IEEE Journal of Biomedical and Health Informatics*, 18(2): 404-410.
- [10] ALI K A, MOUFTAH H T. Wireless personal area networks architecture and protocols for multimedia applications[J]. *Ad Hoc Networks*, 2011, 9(4): 675-686.
- [11] KIMURA N, LATIFI S. A survey on data compression in wireless sensor networks[A]. *International Conference on Information Technology: Coding and Computing[C]*. Las Vegas, Nevada, USA: IEEE Press, 2005.8-13.
- [12] 刘浩, 李伟民, 李晓丽. 无线体征监测系统的低功耗调度方法[J]. *计算机应用*, 2012, 32(3): 839-842.
- LIU H, LI W M, LI X L. Low-power scheduling scheme for wireless physiological monitoring system[J]. *Journal of Computer Applications*, 2012, 32(3): 839-842.
- [13] HUANG X, SHAN H, SHEN X. On energy efficiency of cooperative communications in wireless body area network[A]. *Wireless Communications and Networking Conference (WCNC)[C]*. 2011.1097-1101.
- [14] YUAN F, ZHAN Y, WANG Y. Data density correlation degree clustering method for data aggregation in WSN[J]. *IEEE Sensors Journal*, 2014, 14(4): 1089-1098.
- [15] ROY S, CONTI M, SETIA S, *et al.* Secure data aggregation in wireless sensor networks: filtering out the attacker's impact[J]. *IEEE Transactions on Information Forensics and Security*, 2014, 9(4): 681-694.
- [16] GHASEMZADEH H, JAFARI R. Physical movement monitoring using body sensor networks: a phonological approach to construct spatial decision trees[J]. *IEEE Transactions on Industrial Informatics*, 2011, 7(1): 66-77.
- [17] WANG Z, JIANG M, HU Y, *et al.* An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensors[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2012, 16(4): 691-699.
- [18] SEBBAK F, CHIBANI A, AMIRAT Y, *et al.* An evidential fusion approach for activity recognition in ambient intelligence environments[J]. *Robotics and Autonomous Systems*, 2013, 61(11): 1235-1245.
- [19] FORTINO G, GALZARANO S, GRAVINA R, *et al.* A framework for collaborative computing and multi-sensor data fusion in body sensor networks[EB/OL]. <http://www.researchgate.net/publication/261373973>.
- [20] GHASEMZADEH H, JAFARI R. Data aggregation in body sensor networks: A power optimization technique for collaborative signal processing[A]. *The 7th Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)[C]*. 2010.1-9.
- [21] 回春立, 崔莉. 无线传感器网络中基于预测的时域数据融合技术[J]. *计算机工程与应用*, 2007, 43(21): 121-125.
- HUI C L, CUI L. Forecast-based temporal aggregation in wireless sensor networks[J]. *Computer Engineering and Applications*, 2007, 43(21): 121-125.
- [22] KANG J, TANG L, ZUO X, *et al.* GMSVM-based prediction for temporal data aggregation in sensor networks[A]. *Wireless Communications, Networking and Mobile Computing, WiCom'09 5th International Conference on[C]*. 2009.1-4.
- [23] 宫继兵, 王睿, 崔莉. 体域网 BSN 的研究进展及面临的挑战[J]. *计算机研究与发展*, 2010 (5): 737-753.
- GONG J B, WANG R, CUI L. Research advances and challenges of body sensor network(BSN)[J]. *Journal of Computer Research and Development*, 2010 (5): 737-753.
- [24] CHENG F H, CHEN Y L. Real time multiple objects tracking and identification based on discrete wavelet transform[J]. *Pattern Recognition*, 2006, 39(6): 1126-1139.
- [25] BANERJEE T P, DAS S. Multi-sensor data fusion using support vector machine for motor fault detection[J]. *Information Sciences*, 2012, 217: 96-107.
- [26] MOSHOUBI D, PANTAZI X E, KATERIS D, *et al.* Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier[J]. *Biosystems Engineering*, 2014, 117: 15-22.
- [27] WEI G, LING Y, GUO B, *et al.* Prediction-based data aggregation in wireless sensor networks: combining grey model and Kalman filter[J]. *Computer Communications*, 2011, 34(6): 793-802.

作者简介:



王汝言 (1969-), 男, 湖北浠水人, 博士, 重庆邮电大学教授、博士生导师, 主要研究方向为泛在网络、全光网络理论与技术、多媒体信息处理等。

翟美玲 (1988-), 女, 河北保定人, 重庆邮电大学硕士生, 主要研究方向为无线体域网、数据融合算法。

吴大鹏 (1979-), 男, 黑龙江大庆人, 博士, 重庆邮电大学教授、硕士生导师, 主要研究方向为泛在无线网络、无线网络服务质量管理等。