

多传感器的 BPNN 和 SVM 多源 异构数据融合算法

王晓琪[†], 陈颖聪, 谢敏敏, 张嘉慧, 蔡上

(广东电网有限责任公司 梅州供电局, 广东 梅州 514021)

摘 要:多传感器的多源异构数据融合处理时,大量的冗余数据及复杂的非线性可分空间导致能耗较大,为此,提出了 BP 神经网络和支持向量机的多源异构数据融合算法。以数据关系构建约束条件,利用 BP 神经网络算法建立数据清洗模型,判定节点变量的活跃程度,优化数据输入;建立数据集合,提取数据特征向量;利用支持向量机泛化能力强、凸优化的特点,获取特征的最优分类超平面,获得非线性可分多源数据集转化为高维线性可分空间的最优决策值,输出结果。实验结果表明,该算法融合多源异构数据的能量消耗小、延迟低,融合效果好。

关键词:BP 神经网络;支持向量机;多源异构数据;数据清洗;数据融合

中图分类号:TP393

文献标识码:A

Multi Sensor Heterogeneous Data Fusion Algorithm Based on BPNN and SVM

WANG Xiaoqi[†], CHEN Yingcong, XIE Minmin, ZHANG Jiahui, CAI Shang

(Meizhou Power Supply Bureau of Guangdong Power Grid Co., Ltd., Meizhou, Guangdong 514021, China)

Abstract: In the process of multi-sensor multi-source heterogeneous data fusion processing, a large number of redundant data and complex nonlinear separable space lead to high energy consumption. Therefore, a multi-source heterogeneous data fusion algorithm based on BP neural network and support vector machine is proposed. Based on the data relationship, the constraint conditions were established, and the BP neural network algorithm was used to establish the data cleaning model, and the activity degree of node variables was determined to optimize the data input. To set up data set and extract data feature vector; Based on the support vector machine's strong generalization ability and convex optimization, the optimal classification hyperplane of the features is obtained, and the optimal decision value of the nonlinear separable multi-source data set is obtained into the high-dimensional linear separable space. Experimental results show that this algorithm has low energy consumption, low delay and good fusion effect.

Key words: BP neural network; support vector machine; multi source heterogeneous data; data cleaning; data fusion

通常,一次性全部上传多个传感器的数据至云端,大量的多源异构数据会降低处理效率。由此,需要将多源异构数据转换成分布式数据^[1]上传。但是,由于其数据属性形式呈现多模态,存在噪声

构成复杂的非线性可分空间,导致转化过程复杂,因此需要对不同传感器设备采集的多源异构数据^[2]实施融合处理。

文献[3]提出基于 BP 神经网络构建数据融合

模型,将环境影响因子作为输入变量,期望变化值作为输出数据,进行相关性计算,完成模型的求解。但该算法缺少对数据的清洗处理,导致数据融合效果差。文献[4]提出利用数据之间的结构稀疏性以及时间关联性,建立正则化融合模型,利用梯度下降优化算法求解。但该算法构建模型过程中计算量较大,导致数据融合延迟高。文献[5]提出根据数据定位结果,结合神经网络算法构建数据融合模型,再使用遗传算法优化模型。但该算法在优化模型时易陷入局部最优,融合能量消耗大。

为此,本文提出 BP 神经网络(back propagation neural network, BPNN)和支持向量机(support vector machine, SVM)的传感器设备多源异构数据融合算法。利用 BP 神经网络对数据进行清洗,初步排除噪声干扰,消除冗余数据;利用 SVM 算法进行数据融合运算,借助其泛化能力强且稳定的优势,寻找并训练最优分类面,避免陷入局部最优等情况发生,确保分类的有效性,从而使构建的多源异构数据融合结果更为可靠。

1 基于 BP 的多源异构数据清洗算法

利用机器学习算法中的 BP 神经网络构建多源异构数据清洗处理模型,去除多传感器设备上传数据的大量冗余数据。其机理是先通过设定多源异构数据的相似性约束关系,实现数据内部结构的关联性,不仅可以提高后期 BP 神经网络内部数据清洗运行的速度,还能够提升数据清洗工作效率。BP 神经网络中最基本的分类器是分类回归树,主要就是在上述关联性的约束下,通过对多源异构数据结果进行相互整合与分类,获取更为精准的多源异构数据关系,排除无用数据关系,从而完成数据清洗。

1.1 获取多源异构数据的相似度关系约束

多个传感器设备上传的原始多源异构数据以数据表的形式存在,设定数据表为 b ,数据总量为 p ,以此建立数据表集合 $B = \{b_1, b_2, \dots, b_p\}$,数据属性集合为 $\varphi = \{e_1, e_2, \dots, e_g\}$,使用曼哈顿距离算法^[6-7]获取多源异构数据之间的相似度,过程如下式所示:

$$l_{b_{i\varphi}, b_{j\varphi}} = \sum_{\varphi=1}^S |b_{i\varphi} - b_{j\varphi}| \quad (1)$$

式中,多源异构数据 i 与数据 j 之间的相似度用 $l_{b_{i\varphi}, b_{j\varphi}}$ 表示;样本数据属性用 φ 表示;数据 i 的属性特征标记 $b_{i\varphi}$ 形式;数据 j 的属性特征标记 $b_{j\varphi}$ 形

式;数据维度标记 S 形式。基于上述计算结果可知,数据之间距离越大,相似度就越小。

依据上述分析结果,设定数据之间的相关系数为 ρ_{b_i, b_j} ,制定数据的参数阈值 λ ,并将其作为数据清洗模型约束条件,过程如下式所示:

$$\lambda = \begin{cases} \lambda > \rho_{b_i, b_j}, & b_i, b_j \text{ 不相关} \\ \lambda \leq \rho_{b_i, b_j}, & b_i, b_j \text{ 相关} \end{cases} \quad (2)$$

基于上述构建的约束条件,构建数据清洗模型。

1.2 基于 BPNN 的数据清洗模型

基于上述建立的约束条件,利用 BP 神经网络^[8-9]构建数据清洗模型,将传统神经网络中模型的中间层实施限制处理。设定多源异构数据清洗模型为 T ,模型中包含 N 层节点,具体模型结构如图 1 所示。

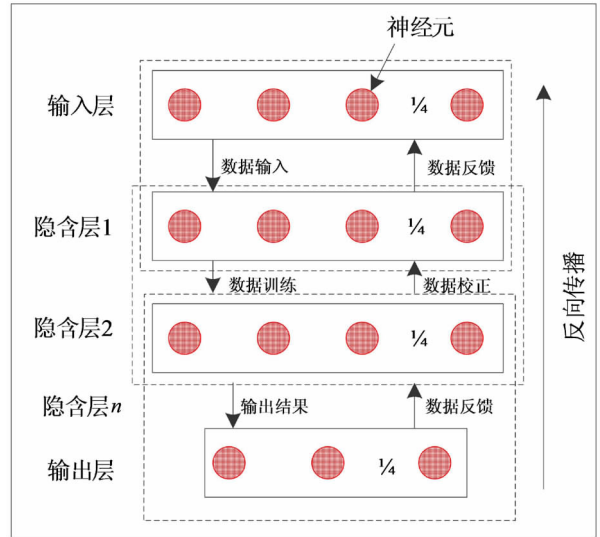


图 1 多源异构数据神经网络模型

BPNN 数据清洗模型^[10]在反向传播时,无监督玻尔兹曼机(Restricted Boltzmann machine, RBM)作为模型的主要部件,是解决多源异构数据冗余数据去除以及特征提取的关键。

在多源异构数据的清洗模型中,每一层的 RBM 结构中都存在可视层以及隐含层数据,设定输入层数据节点有 $c = \{\alpha_1, \alpha_2, \dots, \alpha_c\}$ 个,隐含层节点有 $d = \{\beta_1, \beta_2, \dots, \beta_d\}$ 个,以此构建多源异构数据的清洗模型的层级之间的节点权重矩阵,过程如下式所示:

$$\xi = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1d} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2d} \\ \vdots & \vdots & & \vdots \\ \omega_{c1} & \omega_{c2} & \cdots & \omega_{cd} \end{bmatrix} \quad (3)$$

式中,构建的节点权重矩阵描述成 ξ 形式;矩阵向量标记 ω_{ai} 形式。基于构建的节点权重矩阵结合 RBM 能量函数,获取多源异构数据的节点能量函数,过程如下式所示:

$$A(\alpha_i, \beta_j) = \sum_{i=1}^c \alpha_i \delta_i - \sum_{j=1}^c \beta_j \gamma_j - \sum_{i,j=0}^c \alpha_i \beta_j \omega_{ij} \quad (4)$$

式中,节点数据 α_i 的偏置系数用 δ_i 表示;节点数据 β_j 的偏置系数用 γ_j 表示;数据权重标记 ω_{ij} 形式;常数用 i, j 表示;模型中输入层数据的节点数量用 c 表示;获取的数据节点能量函数描述成 $A(\alpha_i, \beta_j)$ 形式。

依据构建的节点能量函数对模型输入层节点以及隐含层节点的数据分布概率展开测试,测试结果如下式所示:

$$Q(\alpha_i, \beta_j) = \frac{\epsilon^{-A(\alpha_i, \beta_j)}}{\sum_{i,j=0}^c \epsilon^{-A(\alpha_i, \beta_j)}} \quad (5)$$

式中,数据能量函数的归一化系数描述成 $\sum_{i,j=0}^c \epsilon^{-A(\alpha_i, \beta_j)}$ 形式;分布因子描述成 ϵ 形式;数据分布概率用 $Q(\alpha_i, \beta_j)$ 表示。

基于上述获取的数据分布概率能够有效获取模型节点变量的活跃程度,优化清洗模型。一般来说,将多源异构数据样本输入清洗模型中,模型中间层数据节点之间状态相互独立,假定隐含层节点 β_j 为 1 时,获取数据分布概率的约束条件,过程如下式所示:

$$Q(\beta_j = 1 | \alpha_i, \omega_{ij}) = \eta(\varphi_j + \sum_{i,j=1}^c \alpha_i \cdot \omega_{ij}) \quad (6)$$

式中,建立的约束条件用 $Q(\beta_j = 1 | \alpha_i, \omega_{ij})$ 表示;二值型 Sigmoid 函数标记 η 形式;数据状态系数表述成 φ_j 形式。基于上述计算结果可知,当 Sigmoid 函数^[11] 自变量取值区间在 $[0, 1]$ 时,数据的概率分布函数会呈现单一的连续变化状态。

利用上述计算流程激活神经网络模型隐含层节点后,输入 β_j 对模型实施反向传播,实现模型的训练。根据训练结果中抓取的节点能量异常和分布离散的数据,进行去除处理,完成多源异构数据的清洗^[12]。

2 SVM 的多源异构数据融合算法

基于上述数据清洗结果,对多传感器设备的多

源异构数据特征展开提取处理。SVM 是一种以统计为基础的二分类方法。该方法的基本思路是:在训练集合内,通过训练求出一个最小的边界函数,从而使每个数据之间的最短距离达到最大。其具有结构风险最小化原理,具有很好的泛化能力,并且是一个凸优化问题,只要确定局部最优解就一定是全局最优解,避免了陷入局部最优的问题。使用 SVM 算法获取特征的最优分类超平面,构建数据融合模型,最后通过建立的数据融合模型,实现多源异构数据的融合。

2.1 提取多源异构数据特征

基于上述数据清洗结果,更新多源异构数据集合为 $Y_{ij}(t)$, 样本数据量为 M 个,数据状态为 t , 使用均值标准化方法对数据实施标准化处理,结果如下式所示:

$$Y_{ij}^*(t) = \frac{Y_{ij}(t)}{Y_j} \quad (7)$$

式中,数据的标准化结果标记 $Y_{ij}^*(t)$ 形式;数据均值标记 $\bar{Y}_j$ 形式;常数标记 i, j 形式。基于上述数据得到标准化处理结果,提取多源异构数据 i 在第 j 个属性时的特征向量,计算结果如下式所示:

$$\varphi(Y_{ij}) = \sum_{t=1}^T Y_{ij}^*(t) / T \quad (8)$$

式中,多源异构数据的特征向量用 $\varphi(Y_{ij})$ 表示;数据属性记录时期用 T 表述。

2.2 多源异构数据融合算法

利用 SVM 算法^[13] 通过核函数加快收敛速度,控制空间的维数,将数据特征非线性向量空间映射到线性可分空间中,获取最优的分类超平面,构建多源异构数据的数据融合模型。

设定获取的数据特征向量为输入向量 y_i , 核函数为 $K(\cdot)$, 特征向量集为 y , 输出向量为 u_i , 以此获取数据的最优超平面,过程如下式所示:

$$u = \text{sgn} \sum m_i = \tau_i u_i K(y_i, y) + \vartheta \quad (9)$$

式中,获取到最优超平面用 u 表示;加权系数用 τ_i 表示;偏置项用 ϑ 表示;特征向量数量用 m 表示;常数标记 i 形式。

基于获取的最优超分类面,将多源异构数据划分成 h 组数据,且不同组别的数据来自不同的传感器。设定数据的实际值为 u , 训练输出为 U , 泛化参数为 χ , 以此建立 SVM 的多源异构数据的数据融合模型^[14-15], 如下式所示:

$$\begin{cases} J = \min \frac{1}{2} \sum_{i,j=1}^g (l_i^* - l_i)(l_j^* - l_j) K(y_i, y_j) \\ \quad + \chi \sum_{i=1}^g (l_i^* - l_i) \sum_{j=1}^g u_i (l_i^* - l_i) \\ \text{s. t. } (l_i^* - l_i) = 0 \end{cases} \quad (10)$$

式中,构建的数据融合模型用 J 表示;数据 i 、 j 的最优解标记 l_i^* 、 l_j^* 形式;最优解均值标记 l_i^* 、 l_j^* 形式;约束条件标记 s. t. $(l_i^* - l_i)$ 形式;传感器数量用 g 表示。

基于确定的数据最优解,建立数据的决策函数,完成模型求解,实现数据的融合,过程如下式所示:

$$f(x) = \sum_{i=1}^g (\bar{y}_g - l_g) K(y_i, y) + \vartheta \quad (11)$$

式中,构建的决策函数用 $f(x)$ 表示;传感器 g 的数据均值集合标记 $\bar{y}_g$ 形式;数据最优解标记 l_g 形式。多源异构数据融合流程如图 2 所示。

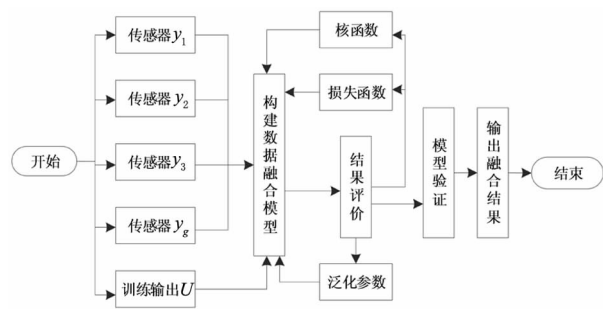


图 2 多源异构数据融合流程图

最后基于上述数据融合流程,将完成训练处理的数据录入数据融合模型,根据核函数、损失函数和泛化参数,对所构建数据融合模型进行验证,输出融合结果。在进行多源数据融合时,需要同时将相关数据库的数据进行封装存储和查询,以便随时进行业务处理、检索查询和在线优化等操作。对完成上述处理的多源异构数据,按照数据的特征进行归类,通过此种方式,即可实现对多源数据特征的数据清洗处理,可以减少重复记录多源异构数据,快速处理空缺记录,有效提高数据的质量,完成高效的数据清洗,实现多源异构数据的融合。

3 实验

3.1 实验准备

为了解决大量的冗余数据及复杂的非线性可分空间导致能耗较大的问题,验证多传感器的

BPNN 和 SVM 多源异构数据融合算法的整体有效性,需要对此方法进行测试。以某变电站室内传感器采集的数据为测试样本,包括湿度传感器、气体传感器、温度传感器,将 2022 年 1 月 1 日至 4 月 30 日的数据作为训练集,将 2022 年 5 月 1 日至 5 月 31 日的数据作为测试集。不同的参数组合影响着多源异构数据融合算法的测试效果,同时也是 BP 神经网络算法和 SVM 算法研究的难点。主要通过仿真实验,在 Matlab 平台上构建 BP 神经网络算法和 SVM 算法,依据每个参数对 BP 神经网络算法和 SVM 算法的最优路径的影响,确定每个参数较为合理的取值范围,因此设置 BP 神经网络的训练误差为 0.05,最大迭代次数为 40000,训练次数为 5000,权重衰减次数为 0.001,步长为 0.02。设置 SVM 算法的处罚因子为 1,高斯核函数宽度参数为 0.5,距离中值为 5,加权系数为 2。

在仿真平台上,输入湿度传感器、气体传感器、温度传感器采集的数据,将众多输入数据参数类别及其可能取到的值放在一起进行融合,当所有可能的组合被检索一遍后,得到最优参数组合即为最佳值,由此进行参数调优。由于数据之间呈现多源异构特性,数据分布如图 3 所示。

分析图 3 可知,通过仿真平台采集湿度传感器、气体传感器、温度传感器三组多源异构原始数据,得到的原始数据是不规律分布的。根据传感器多源异构数据分布情况,对该数据进行 BPNN 清洗处理,结果如图 4 所示。

清洗处理过程中,判断数据的来源与渠道,在此过程中若多源异构数据存在丢失或异常情况,则需要清洗去除异常情况。完成处理后,将数据录入计算机,使用本文设计算法提取多源异构数据集合中的特征信息。在此基础上,选择不同的链路得到数据在终端的融合处理结果。通过图 4 可以看出,本文的清洗算法可以很好地排除数据干扰,数据优化效果明显。

3.2 实验结果及分析

采用所提算法对上述传感器多源异构数据进行融合处理,在进行 BP 神经网络清洗后,利用 SVM 算法获得最优融合结果,设定错误项的惩罚系数为 3。实验中,核函数分别采用了‘RBF’:径向核函数/高斯核、线性核函数、多项式核函数,得到了不同的分类结果。多项式核函数的阶数 3,只对多项式核函数有用,默认值为 3。多源异构传感

数据分类结果如图 5 所示。

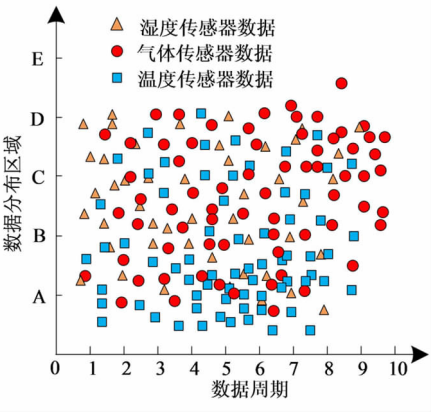
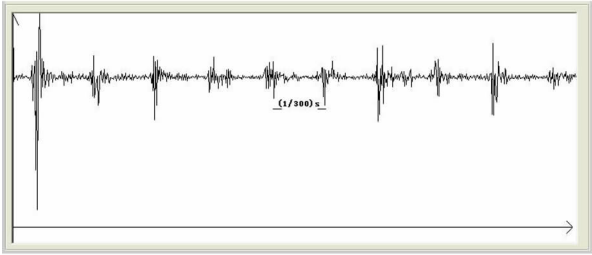
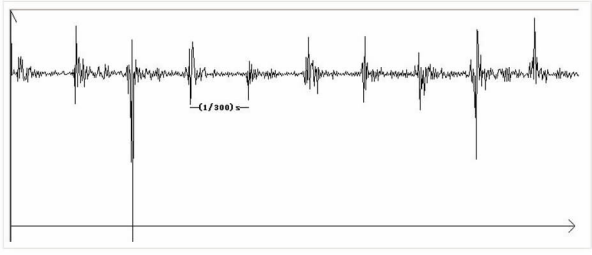


图 3 传感器多源异构原始数据分布图



(a) 原始多传感信号混合数据



(b) 完成清洗后的数据

图 4 数据清洗结果

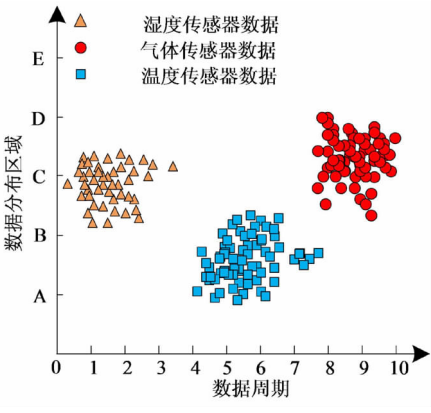


图 5 所提算法的多源异构数据分类效果图

根据所得清洗后的传感数据结果,在终端调用后台 sklearn 库数据,实时统计并记录数据融合情况,分析不同算法下能量消耗、融合延迟、包交换率

以及融合误差测试结果。

在数据融合过程中,数据融合算法的融合性能是检测算法优劣的关键,为体现性能的优越性,选取数据融合过程中的能量消耗、融合延时、包交换率测试及融合误差四个方面进行测试,将所提算法基于 BP 神经网络的滑坡监测多源异构数据融合算法(文献[3]算法)、基于时间光滑正则化的序列诊疗数据融合算法(文献[4]算法)进行对比,因文献[5]算法和文献[3]算法具有相同的融合算法和实验条件,在一定程度上具有相似的验证效果,所以此处省略与文献[5]的对比。

(1) 能量消耗测试。

应用了数据融合算法后,节点传播的生命长度将大大延长,随着传播节点数量的增多,在多个节点上传播数据会消耗大量的电量,融合数据后,节点的能耗将大幅度降低。采用所提算法、文献[3]算法以及文献[4]算法开展数据融合时,对上述 3 种算法在数据融合时产生的能量消耗展开测试,测试结果如图 6 所示。

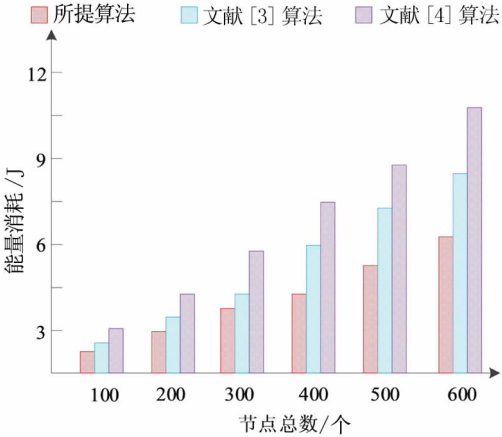


图 6 不同算法的能量消耗测试结果

分析图 6 实验数据可知,随着数据融合时间的不断增加,3 种数据融合算法检测出的融合能量消耗均出现了不同程度的上升。但是,所提算法检测出的数据融合能量消耗是 3 种算法中最低的,当节点总数为 600 时,达到了最高,为 6.4 J。

(2) 融合延迟测试。

数据在融合过程中会产生延迟现象,延迟时间越长,说明数据融合算法的融合性能越差,反之则越好。采用所提算法、文献[3]算法以及文献[4]算法开展数据融合时,测试上述 3 种算法在数据融合时产生的延迟,结果如图 7 所示。

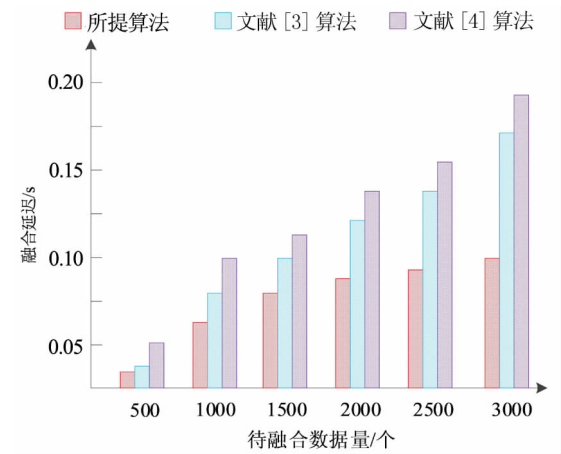


图 7 不同算法的融合延迟测试结果

分析图 7 可知,随着待融合数据量的增加,3 种算法检测出的数据融合延迟都出现了不同程度的上升。其中,所提算法在数据融合时,产生的融合延迟低于文献[3]算法以及文献[4]算法的测试结果,当待融合数据量为 3000 个时,延迟时间达到最高,为 0.1 s。

(3)包交换率测试。

基于上述测试结果,选取包交换率作为数据融合效果检测指标,数据融合过程中,检测出的数据包交换率越高,说明数据的融合效果越好,反之则越差。测试数据在所提算法、文献[3]算法以及文献[4]算法下开展数据融合时的融合效果,结果如图 8 所示。

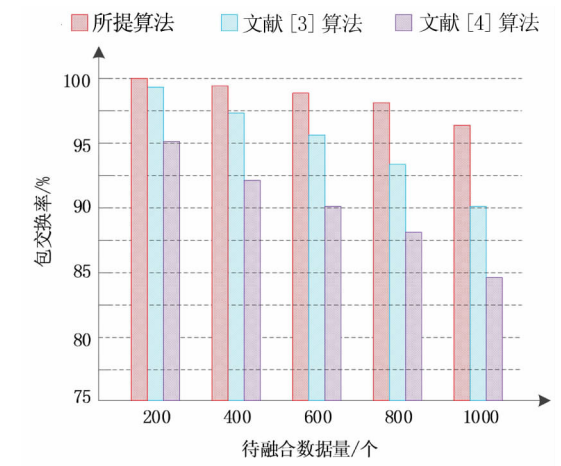


图 8 不同算法的包交换率测试结果

分析图 8 实验数据可知,随着待融合数据的增加,3 种数据融合算法检测出的包交换率都呈现不同程度的下降。但是,所提算法在数据融合时检测出的包交换率是 3 种算法中最高的,最低为 96%。

(4)融合误差测试。

以多传感器采集的实验数据来源为根据,设定传感器数据分析的需求协方差值,对比不同算法的追踪需求误差,结果如图 9 所示。

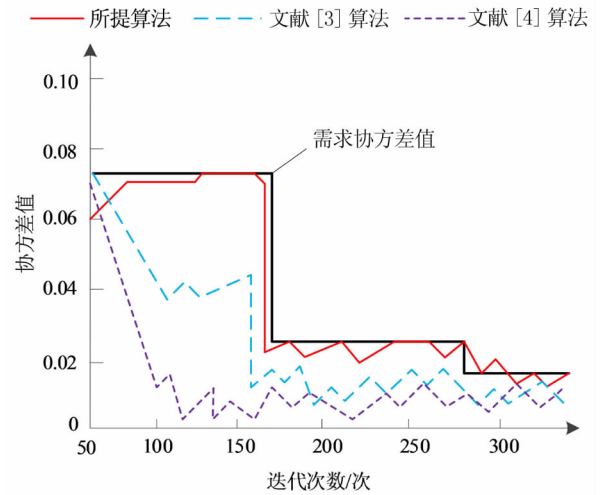


图 9 不同算法的融合协方差测试结果

分析图 9 实验数据可知,不同算法均可以满足设定的需求协方差值,但是文献[3]的数据融合算法追踪融合需求误差能力较低,浪费了数据融合处理的精度,文献[4]的数据融合误差一直处于较为稳定的状态,没有根据需求的精度进行自适应调整,可见在传感器数据处理中应用效果较差,而所提算法可以持续地保持追踪需求协方差值,具有较好的融合精度。

综上所述,所提算法在数据融合时,能量消耗低、融合延迟低、包交换率高,且可保持较好的融合精度,说明所提算法在数据融合时的融合性能好。这主要是因为所提算法在数据融合前,使用 BP 神经网络算法对数据实施了清洗处理,去除了冗余数据的干扰,增强了数据融合的效果,提升了包交换率。并在此基础上,利用 SVM 算法获得了最优超分类面,避免陷入局部最优,大大减少了循环运算量,所以所提算法在数据融合时产生的能量消耗及融合延迟均低于其他两种数据融合方法。

4 结 论

随着云端计算技术使用范围的增加,多源异构数据融合算法就变得尤为重要。针对传统数据融合算法中存在的问题,提出应用 BPNN 和 SVM 的多传感器设备的多源异构数据融合算法。该算法

以 BP 神经网络进行数据清洗,构建多源异构数据集,提取数据特征;结合 SVM 算法构建数据融合模型,通过模型输出完成数据的融合。通过上述研究,得到如下结论。

(1)本文算法检测出的数据在节点总数为 600 时达到最高,融合能量消耗为 6.4 J。

(2)本文算法在数据融合时,产生的融合延迟低于其他算法的测试结果,当待融合数据量为 3000 个时,延迟达到最高,为 0.1 s。

(3)本文算法在数据融合时检测出的包交换率较高,均为 95%以上。

(4)本文算法可以持续地保持追踪需求协方差值,具有较好的融合精度。

但是该算法在提取数据特征时,还存在一定问题,今后会针对该项缺陷继续优化该数据融合算法,具体内容如下。

(1)在进行数据特征提取时,未考虑由原始输入形成的新特征,忽略了提取过程中的复杂性,导致提取速度较慢,所以在下一步的工作中需要重点研究此问题;

(2)多源异构数据网络由一个节点代表一个实体,该物体的有向边代表实体间的关联。在此基础上,今后的研究工作可以强化实体间的关联,提取数据特征,同时快速处理空缺记录,提高多源异构数据融合速度。

参考文献

[1] 张新宇. 基于分布式云的医疗数据共享系统[J]. 信息技术, 2021(8):59-63+69.

[2] 侯慧,于士文,肖祥,等. 基于空间多源异构数据的台风下输电杆塔风险评估[J]. 电力系统自动化,2020,44(10):127-134.

[3] 王智伟,王利,黄观文,等. 基于 BP 神经网络的滑坡监测多源异构数据融合算法研究[J]. 地质力学学报,2020,26(4):575-582.

[4] 郑毅,胡祥培. 基于时间光滑正则化的序列诊疗数据融合方法[J]. 管理科学,2021,34(4):61-70.

[5] 刘宇,惠鸿飞,路永乐,等. 基于遗传算法优化神经网络的多源信息融合室内定位方法[J]. 中国惯性技术学报,2020,28(1):67-73.

[6] 窦家维,葛雪,王颖因. 保护隐私的曼哈顿距离计算及其推广应用[J]. 计算机学报,2020,43(2):352-365.

[7] 李浩君,张鹏威,郭海东. 基于种群曼哈顿距离的自适应多目标粒子群优化算法[J]. 计算机集成制造系统,2020,26(4):1019-1032.

[8] 郭杰,苏路,朱海浪,等. 基于 IBAS-BP 神经网络的短期风电功率预测[J]. 信息技术,2021(10):38-43.

[9] 韩正新,乔耀华,吕守国,等. 基于 HC 显著算法与 BP 神经网络的机器视觉系统[J]. 信息技术,2020,44(3):54-58.

[10] 姜山,丁治明,朱美玲,等. 面向时空图建模的图小波卷积神经网络模型[J]. 软件学报,2021,32(3):726-741.

[11] VERDOOLAEGE G, KAYE S M, ANGIONI C, et al. The updated ITPA global H-mode confinement database: description and analysis[J]. Nuclear Fusion, 2021, 61(7):76-105.

[12] 时珉,尹瑞,胡傲宇,等. 基于滑动标准差计算的光伏阵列异常数据清洗办法[J]. 电力系统保护与控制,2020,48(6):108-114.

[13] 杨滨峰,赵杰. 一种新的物联网传感器数据融合方法[J]. 信息技术,2022(5):13-17+24.

[14] HASHIM H A, ELTOUKHY A. Landmark and IMU data fusion:systematic convergence geometric nonlinear observer for SLAM and velocity bias[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 31(3):1-10.

[15] GHAZIPOUR F, MAHJOURI N. A multi-model data fusion methodology for seasonal drought forecasting under uncertainty: application of Bayesian maximum entropy [J]. Journal of Environmental Management, 2022, 304(15):114-120.