

基于 GPS 和神经网络的滑坡位移监测算法

张敏敏¹, 吕晓军², 贾新春¹, 杨波¹

(1. 山西大学 数学科学学院, 太原 030006;

2. 中国铁道科学研究院 电子计算技术研究所, 北京 100081)

摘要: 针对使用 GPS 接收机进行滑坡位移监测时, 单点 GPS 定位误差较大的问题, 提出了一种基于 GPS 和神经网络的滑坡位移监测算法; GPS 接收机在滑坡发生之前和滑坡发生之后测得的定位数据耦合在一起, 不是线性可分的; 采用具有非线性可分特性的神经网络, 把耦合在一起的定位数据分成两类, 一类属于未滑坡的 GPS 数据, 另一类属于发生滑坡的 GPS 数据, 避免了对 GPS 定位误差这一非线性非高斯问题进行准确建模的过程; 利用 GPS 接收机测得的样本训练集训练神经网络, 用训练后的神经网络模型来验证测试集的分类效果; 实测实验结果表明, 对于低精度的 GPS 接收机, 当滑坡程度分别达到 3 米、5 米、8 米时, 训练样本分类正确率分别是 95.85%、99.23%、99.99%, 测试样本分类正确率分别是 82.94%、89.44%、91.05%, 说明所提出的算法适用于较大程度的滑坡。

关键词: 滑坡; GPS 定位; BP 神经网络; 定位误差; 分类

Landslide Displacement Monitoring Algorithm Based on GPS and Neural Network

Zhang Minmin¹, Lü Xiaojun², Jia Xinchun¹, Yang Bo¹

(1. School of Mathematical Sciences, Shanxi University, Taiyuan 030006, China;

2. Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: Aiming at the problem of single-point GPS positioning error when using GPS receivers for landslide displacement monitoring, a landslide displacement monitoring algorithm based on GPS and neural network is proposed. Before the landslide and landslide occurred, the positioning data measured by GPS receivers are coupled together and are not linearly separable. The neural network with non-linear separable features is used to divide the coupled positioning data into two classifications, one belongs to the non-landslide GPS data and the other belongs to the landslide GPS data, which avoids the accurate modeling of GPS nonlinear and non-Gaussian positioning error. The sample training set measured by the GPS receiver is used to train the neural network, and the trained neural network model is used to verify the classification effect of the test set. The experimental results show that for the low-precision GPS receivers, when the landslide reaches 3 meters, 5 meters and 8 meters respectively, the correct rates of training samples classification are 95.85%, 99.23% and 99.99% respectively, and the correct rates of testing samples classification are 82.94%, 89.44% and 91.05% respectively, indicating that the proposed algorithm is suitable for a greater degree of landslide.

Keywords: landslide; GPS positioning; BP neural network; positioning error; classification

0 引言

我国各类地质灾害频发, 尤其是滑坡灾害影响最为严重, 而铁路沿线的滑坡更是危及旅客们的人身安全, 因此, 滑坡监测预警越来越受到研究者的关注。文献 [1-2] 对滑坡监测技术方法现状进行了分析, 按监测对象的不同, 滑坡监测可分为 4 大类: 即位移监测、物理场监测、地下水

监测和外部诱发因素监测。这 4 大类监测又可分为若干小类, 每类监测采取的方法手段不同, 使用的仪器不同, 获取的参数也不同。本文研究的是滑坡位移监测, 采用低精度 GPS 接收机测得滑坡测点在不同时刻的经纬度坐标数据, 根据经纬度坐标之间的距离来判断是否发生滑坡。因此, GPS 定位问题是滑坡监测中很关键的一个问题。

低精度 GPS 接收机单点定位精度在理想情况下一般能达到 10 米左右。GPS 定位误差主要有卫星误差、传播误差以及接受误差。目前, 差分 GPS (DGPS) 技术是减小定位误差的较好方法, 但此方法至少需要 2 台 GPS 接收机, 基准接收机需要通过 DGPS 数据链为动态用户提供差分改正数, 这在一定程度上限制了 DGPS 的使用范围。许多学者研究了在不改变硬件设备的前提下通过对 GPS 误差信号的建模来提高单点定位精度的方法。文献 [3-4] 利用 Kalman 滤波算法对 GPS 定位数据进行误差处理分析, 得到较小的估计误差。文献 [5] 研究了交互式多模型 Kalman

收稿日期:2017-12-25; 修回日期:2018-01-21。

基金项目:国家自然科学基金项目(U1334210, 61374059)。

作者简介:张敏敏(1992-),男,山西闻喜人,硕士研究生,主要从事 GPS 数据处理、模式识别、无线传感器网络方向的研究。

吕晓军(1968-),男,山西运城人,副主任研究员,主要从事交通信息工程及控制、模式识别与智能系统、无线传感器网络等方向的研究。

贾新春(1964-),男,山西大同人,博士,教授,博士生导师,主要从事网络化控制系统、模糊控制、无线传感器网络等方向的研究。

滤波 (IMMKF) 算法, 使用 2 个模型来描述 GPS 在不同环境噪声中可能的状态, 通过有效的加权融合进行系统状态估计, 验证了变噪声模型的 IMMKF 能很好地适应环境噪声的变化。文献 [6-8] 建立了适合 GPS 误差一阶差分序列的 AR 模型, 并利用扩展 Kalman 滤波算法对系统状态进行估计, 有效提高了定位精度。但是以上这些方法都假设 GPS 的观测噪声和过程噪声为独立不相关的高斯白噪声, 因此可以用 Kalman 滤波、AR 等方法来去除噪声, 提高定位精度。而在实际中, 观测误差服从的分布不是严格的高斯分布。因此, 处理 GPS 定位误差这一非线性非高斯问题是一个难点。文献 [9] 研究了基于马尔可夫链蒙特卡洛 (MCMC) 粒子滤波的 GPS 定位数据处理算法, 避免了噪声的高斯假设和非线性部分的线性化误差。文献 [10] 采用多个 GPS 接收机定位数据的平均, 相比于单个 GPS 接收机, 提高了定位精度。上述这些方法考虑的是 GPS 单点静态定位问题, 而本文的主要目的是利用 GPS 定位数据进行滑坡位移监测, 通过 GPS 接收机测得的经纬度数据与 GPS 接收机实际位置经纬度数据进行比较, 来判断是否发生滑坡, 即把 GPS 接收机测得的经纬度数据分成两类, 判断它是属于未滑坡的一类还是发生滑坡的一类, 这样滑坡监测问题转化为一个对测量的 GPS 数据进行二分类的问题。然而 GPS 接收机未移动情况下和移动之后测得的 GPS 数据耦合在一起, 不是线性可分的, 因此常规的线性分类器无法对 GPS 数据进行有效的分类。

神经网络不需要对所解决问题进行建模, 单个神经元有较简单却又反映非线性本质特征的非线性核, 通过这些基本的非线性核自组织复合, 使得它具有从训练数据中学习任意复杂的非线性映射的能力, 也包括实现复杂的非线性分类判别函数。使用神经网络方法避免了繁琐的常规建模过程, 且神经网络模型有良好的自适应和自学习能力、较强的抗干扰能力, 使得神经网络在模式识别的非线性分类器设计中应用甚广^[11]。神经网络具有“黑盒子”的特点, 只要事先确定神经网络的网络结构、神经元的传递函数和连接权值的学习算法, 并不需要关心网络最后实现的分类器的具体形式, 能够有效的解决非线性的分类问题。因此, 本文将 GPS 定位和神经网络结合起来进行滑坡位移监测。

本文针对滑坡位移监测中滑坡发生前后 GPS 定位数据不可分的问题, 使用具有非线性可分特性的神经网络, 把耦合在一起的定位数据分成两类。首先将 GPS 接收机测得的数据集分为训练集和测试集两类, 接着用训练集来训练神经网络, 最后用训练后的神经网络模型来验证测试集的分类效果, 提出了一种基于 GPS 和神经网络的滑坡位移监测算法。

1 GPS 定位误差采集

使用和芯星通 UM220 接收机连续采集接收机天线位置数据 (lon_i, lat_i) , $i=1, \dots, n$, 假设天线实际位置为:

$$(\overline{lon}, \overline{lat}) = (\frac{1}{n} \sum_{i=1}^n lon_i, \frac{1}{n} \sum_{i=1}^n lat_i), \text{ 根据所获得的}$$

数据与实际位置比较, 得到 GPS 定位误差 $(\Delta lon_i, \Delta lat_i)$, $i=1, \dots, n$ 。GPS 卫星绕地球的运转周期是 12 小时, 地球自转周期是 24 小时, 考虑到白天和夜晚环境因素的影响, 可以认为 GPS 接收机测量数据的变化周期是 24 小时。图 1 是单点静态 GPS 接收机从 2017 年 8 月 24 日 20 时到 2017 年 8 月 25 日 20 时采集的数据误差, 采样间隔为 1 秒。

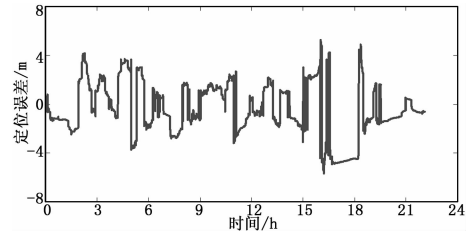


图 1 单点 GPS 定位误差

图 1 反映了当 GPS 接收机不发生移动时, GPS 接收机测得的位置数据与 GPS 接收机实际位置之间的距离误差相对较大 (例如, 当测得距离为 4 米时, 无法判断是 GPS 接收机移动导致的还是测得的距离误差)。图 2 是测得的两个 GPS 接收机之间的相对定位误差。

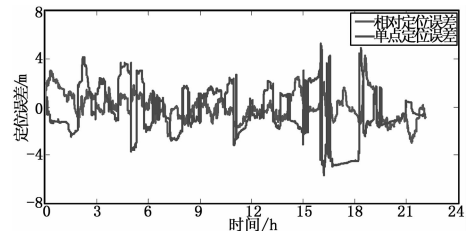


图 2 GPS 定位误差

两个 GPS 接收机之间的相对定位消除了接收机共有的卫星误差和传播误差, 所以与单点定位相比, 相对定位的误差减小, 但只是略微减小, 达不到期望的定位精度。因此, 为了避免对 GPS 定位误差进行准确的建模, 不用把 GPS 接收机测得的经纬度数据转化为距离, 而是考虑利用神经网络把 GPS 定位数据分成两类, 首先用 GPS 定位数据样本训练神经网络, 再利用训练后的神经网络模型来判断最新测得的 GPS 数据是属于未滑坡的一类还是发生滑坡的一类。

2 基于 GPS 和神经网络的滑坡位移监测算法

基于 GPS 和神经网络的滑坡位移监测算法流程如图 3 所示。首先将 GPS 接收机测得的数据集分为训练集和测试集两类, 然后分别对两组数据集进行归一化处理, 接着用训练集来训练神经网络, 最后用训练后的神经网络模型来验证测试集的分类效果^[12]。

2.1 神经网络的模型结构

多层感知器神经网络实现的是从 m 维输入 x 到 n 维输出 y 的一个映射。它由多个采用 Sigmoid 型传递函数的神经元节点连接而成, 这些神经元节点分层排列, 每一层的神经元接受来自前一层的信号, 经过加工后又传递给后一层。

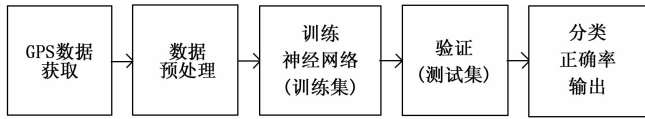


图 3 基于 GPS 和神经网络的滑坡位移监测算法流程

多层感知器神经网络的第一层是输入层, 每个节点对应于 x 的每一维, 节点本身并不完成任何处理, 只是把每一维的信号“分发”到后一层的每个节点。最后一层是输出层, 如果 y 是一维, 则输出层只有一个节点。在输入层和输出层之间的各层都被称作“隐含层”。这种神经网络实现的是从输入层到输出层的映射, 把一个样本特征向量的各分量分别输入到网络输入层的各个对应节点上, 经过在网络上从前向后一系列加工运算, 在输出端得到相应的输出值(或输出向量)。相关研究证明, 任意一个从 x 到 y 的非线性映射, 都存在一个适当结构的三层前馈神经网络能够以任意的精度来逼近它。所以本文考虑带有一个隐含层的多层感知器神经网络作为分类器。

所设计的分类器目的是要对 GPS 接收机未移动情况下测得的位置数据与 GPS 接收机移动之后测得的位置数据进行分类, 所以选择 GPS 接收机测得的经度数据和纬度数据作为神经网络的输入。因此, 输入层取 2 个神经元, 输出层取 1 个神经元。隐含层神经元数目可由经验公式 $h = \sqrt{m+n} + a$ 获得, 其中 h 为隐含层神经元数目, m 为输入层神经元数目, n 为输出层神经元数目, a 为 1~10 之间的调节常数。隐含层神经元数目增加, 可以提高精度, 但带来更大的计算量, 故隐含层神经元数目选为 8。

本文采用如图 4 所示的网络结构形式, 包括经度数据和纬度数据组成的 2 维矢量的输入层; 具有 Sigmoid 型传递函数、8 个神经元组成的一个隐含层; 以及 1 个神经元组成的线性输出层。

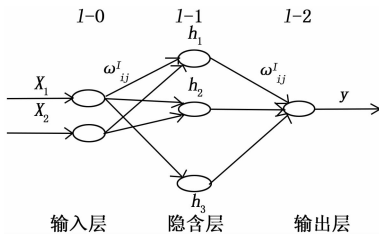


图 4 多层感知器神经网络的模型结构

2.2 神经网络的学习算法

多层感知器神经网络采用反向传播 (BP) 算法训练其网络权值, 其核心思想是梯度下降法, 即以训练样本被错分的程度为目标函数, 训练中每次出现错误时便使权系数朝着目标函数相对于权系数的负梯度方向更新, 直到目标函数取得极小值, 即没有训练样本被错分。

针对本文所采用的多层感知器神经网络结构, 对图 4 神经网络的各个变量约定如下:

输入向量是 2 维, $x = [x_1, x_2]^T$ 。用上标 l 代表神经元节点所在的层, 输入层记 $l = 0$, 隐含层记 $l = 1$, 输出层记

$l = 2$ 。第 l 层神经元个数记为 n_l , 第 l 层第 i 个神经元的输出记作 x_i^l , 输入层节点个数为 $n_0 = 2$, 隐含层神经元个数为 $n_1 = 8$, 输出层节点个数为 $n_2 = 1$, 即网络输出记为 y 。第 l 层的权值都用 ω_{ij}^l 表示, 其中上标 l 表示所在层, 下标 ij 表示所连接的两个节点, ω_{ij}^l 就表示第 $l-1$ 层的节点 i 连接到第 l 层的节点 j 的权值。BP 算法是迭代计算各个权值, 在下面的算法描述中, 用 $\omega_{ij}^l(t)$ 表示在 t 时刻(第 t 步迭代)时权值 ω_{ij}^l 的取值。

BP 算法的具体步骤如下:

- 1) 用小随机数进行权值初始化, 设训练时间 $t = 0$ 。
- 2) 从训练集中得到一个训练样本 $x = [x_1, x_2]^T \in \mathbf{R}^2$, 记它的期望输出是 $d \in \mathbf{R}$ 。

- 3) 计算在 x 输入下当前神经网络的实际输出

$$y = f\left(\sum_{j=1}^8 \omega_{jk}^{l=2} f\left(\sum_{i=1}^2 \omega_{ij}^{l=1} x_i\right)\right), \quad k = 1 \quad (1)$$

其中: $f(\cdot)$ 是 Sigmoid 函数

$$f(\alpha) = \frac{1}{1 + e^{-\alpha}} \quad (2)$$

- 4) 从输出层开始调整权值, 做法是:

对第 l 层, 用下面的公式修正权值

$$\omega_{ij}^l(t+1) = \omega_{ij}^l(t) + \Delta\omega_{ij}^l(t), \quad j = 1, \dots, n_l, \quad i = 1, \dots, n_{l-1} \quad (3)$$

其中: $\Delta\omega_{ij}^l(t)$ 为权值修正项

$$\Delta\omega_{ij}^l(t) = \eta \delta_j^l x_i^{l-1} \quad (4)$$

η 是学习步长。

对于输出层 ($l = 2$), δ_j^l 是当前输出与期望输出之误差对权值的导数

$$\delta_j^l = y_j(1 - y_j)(d_j - y_j), \quad j = 1 \quad (5)$$

对于隐含层 ($l = 1$), δ_j^l 是输出误差反向传播到该层的误差对权值的导数:

$$\delta_j^l = x_j^l(1 - x_j^l) \sum_{k=1}^{n_{l+1}} \delta_k^{l+1} \omega_{jk}^{l+1}(t), \quad j = 1, \dots, n_l \quad (6)$$

5) 在更新全部权值后对所有训练样本重新计算输出, 计算更新后的网络输出与期望输出的总误差。检查算法终止条件, 若满足条件, 则算法结束, 否则置 $t = t + 1$, 返回步骤 (2)。具体流程如图 5 所示。

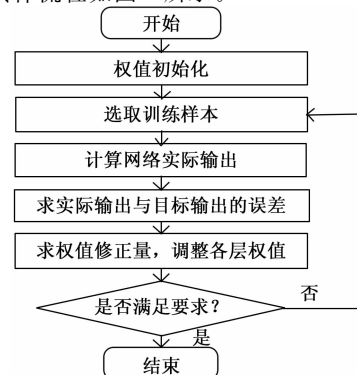


图 5 BP 神经网络流程图

3 实验分析

为了验证 BP 神经网络对 GPS 接收机测得的位置数据的分类性能, 采用两个 GPS 接收机来模拟滑坡发生前后 GPS 接收机的位置, 通过改变两个接收机之间的距离来模拟滑坡程度。图 6 显示了相距 5 米的两个 GPS 接收机随时间变化的位置图。图 7 显示了两个 GPS 接收机平面位置的散点图。

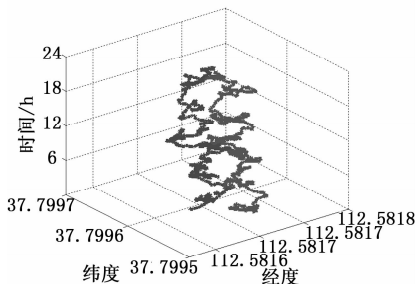


图 6 两个 GPS 接收机随时间变化的位置图

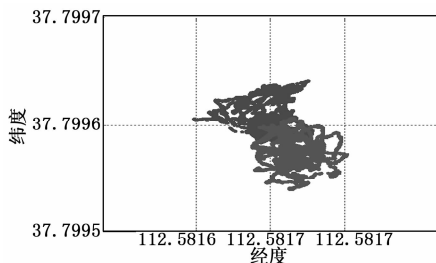


图 7 两个 GPS 接收机平面位置图

从图 6 中可以看出 GPS 接收机测得的位置数据是围绕其真实位置连续变化的, 而从图 7 中可以看出相距 5 米的两个 GPS 接收机, 它们测得的 GPS 位置数据之间有重合, 对于这些重合的数据, 很难判断它们是属于哪个 GPS 接收机, 即在滑坡监测过程中, 对于较小程度的滑坡, 很难通过 GPS 接收机测得的位置数据判断是否发生滑坡。

图 7 中两个 GPS 接收机测得的二维位置数据是线性不可分的, 但从图 6 中可看出在高维空间, 两组数据可通过一个超平面实现分类, 所以本文采用 BP 神经网络来对两组数据进行非线性分类。选取 GPS 接收机连续观测 48 小时的数据作为样本集, 其中前 24 小时的数据作为训练集, 后 24 小时的数据作为测试集。表 1 给出了 GPS 接收机不同距离之间数据的分类效果。

表 1 GPS 接收机不同距离之间样本分类正确率的比较

两个 GPS 接收机 的距离/m	训练样本分类 正确率/%	测试样本分类 正确率/%
8	99.99	91.05
5	99.23	89.44
3	95.85	82.94
2	88.31	70.87
1	84.68	61.17

到的, 从表 1 中可看出当 GPS 接收机相距 8 米时, 训练样本分类正确率是 99.99%, 测试样本分类正确率是 91.05%。即可以认为当滑坡程度达到 8 米时, GPS 接收机监测出滑坡的概率是 91.05%。随着 GPS 接收机之间的距离不断减小, 训练样本分类正确率和测试样本分类正确率都随之降低。当 GPS 接收机之间距离减小至 1 米时, 训练样本正确分类率是 84.68%, 而测试样本分类正确率只有 61.17%, 即当滑坡程度是 1 米时, GPS 接收机监测出滑坡的概率只有 61.17%。这种情况主要是因为低精度 GPS 接收机本身的测量精度较低, 如果使用较高测量精度的 GPS 接收机, 则滑坡位移监测精度会更高。

4 结束语

本文提出了一种基于 GPS 和神经网络的滑坡位移监测算法, 来解决使用 GPS 接收机进行滑坡位移监测时, 由于单点 GPS 定位误差较大导致滑坡发生前后 GPS 定位数据不可分的问题。利用 BP 神经网络非线性可分的特性, 将耦合在一起的 GPS 定位数据分成两类, 避免了对 GPS 定位误差的常规建模过程。最后实验结果表明, 对于低精度的 GPS 接收机, 当滑坡程度达到 8 米时, 测试样本分类正确率为 91.05%, 验证了算法的有效性。但是随着滑坡程度逐渐减小, 测试样本分类正确率也随之降低, 这说明本文所研究的算法也是依赖 GPS 接收机测量精度的, 如何减小这依赖程度, 还有待进一步研究。

参考文献:

- [1] 李亚玲, 苏海峰, 李铭全. 滑坡监测方法研究综述 [J]. 西南公路, 2015 (4): 153-156.
- [2] 唐亚明, 张茂省, 薛强, 等. 滑坡监测预警国内外研究现状及评述 [J]. 地质论评, 2012, 58 (3): 533-541.
- [3] 杜晓辉, 任章. 基于卡尔曼滤波的 GPS 静态定位精度分析 [J]. 全球定位系统, 2008, 33 (5): 47-51.
- [4] 刘婵媛, 陈国光. 基于 GPS 的卡尔曼滤波技术研究 [J]. 弹箭与制导学报, 2006, 26 (4): 110-112.
- [5] 陆建山, 王昌明, 宋高顺, 等. 基于卡尔曼滤波的交互式多模型 GPS 定位方法研究 [J]. 兵工学报, 2011, 32 (6): 770-774.
- [6] 关桂霞, 朱虹, 关永, 等. GPS 定位误差分析与状态估计 [J]. 计算机工程, 2008, 34 (1): 236-237.
- [7] 朱虹, 关永, 田健仲, 等. 单点 GPS 定位误差建模研究 [J]. 微计算机信息, 2008, 24 (16): 206-208.
- [8] 刘娣, 薄薄明, 邹卫军. 基于时间序列的 GPS 误差建模及单点定位精度研究 [J]. 兵工学报, 2009, 30 (6): 825-828.
- [9] 王尔申, 蔡明, 庞涛. MCMC 粒子滤波的 GPS 定位数据处理算法 [J]. 数据采集与处理, 2013, 28 (2): 213-218.
- [10] Schrader D K, Min B C, Matson E T, et al. Real-time averaging of position data from multiple GPS receivers [J]. Measurement, 2016, 90: 329-337.
- [11] 张学工. 模式识别 [M]. 北京: 清华大学出版社, 2010.
- [12] 陈煌琼. 基于神经网络的滑坡预测及其控制研究 [D]. 武汉: 华中科技大学, 2013.

表 1 中的结果是对每种情况进行 10 次实验之后平均得