

基于AdaBoost与BP神经网络的风速预测研究

柳玉^{1,2}, 郭虎全¹

(1 华北电力大学, 控制与计算机工程学院, 北京 102206; 2 新能源电力系统国家重点实验室, 北京 102206)

Wind Speed Prediction Based on AdaBoost and BP Neural Networks

LIU Yu^{1,2}, GUO Hu-quan¹

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 2. State Key Laboratory for Alternate Electrical Power System with Renewable Energy Sources, Beijing 102206, China)

ABSTRACT: This paper introduces an AdaBoost-based multi-neural network ensemble method for wind speed prediction. The result of the prediction by the ensemble method is theoretically and empirically proved to be superior to those by other methods. The AdaBoost algorithm is applied to the time series prediction. Based on the AdaBoost algorithm, back-propagation neural networks (BPNN) are generated; each for training on a random set of examples on wind speed data, then the results of each base learner will be combined to form the final hypothesis. The prediction error by this method is smaller than that by single BP neural network, and the analysis and simulation results suggest that the proposed approach results in better performance.

KEY WORDS: AdaBoost; BP neural network; short-term wind speed forecasting

摘要: 介绍了基于AdaBoost的多神经网络集成预测方法。集成方法的预测结果优于其他方法的预测结果,这一点在理论上和经验上已经得到证明。AdaBoost是适用于时间序列预测的集成方法。基于AdaBoost算法,采用多个BP神经网络训练随机生成的风速样本,再由多个训练结果生成最终的风速预测值。用该方法预测的误差低于用单一BP神经网络进行的预测,其分析和仿真结果表明了其优越性。

关键词: AdaBoost; BP神经网络; 短期风速预测

近年来,风力发电机组单机容量和大型并网风电场的发电总量迅速增长^[1-2],影响了电力系统的安全平稳运行^[3]。因此,进行风速预测很有意义。

短时期的预测需要基于对过去和现在的数据进行推广的统计技术。最简单的统计预测方法是持续预测法^[4],这种方法一般是将最后一步的测量数据作为下一步预测数据的输入部分,由于只考虑上一步测量值,模型的预测误差较大,预测结果不稳定;时间序列法^[5-6]利用历史测量值建立线性模型,

但是其低阶模型预测精度低,高阶模型参数整定难度大;卡尔曼滤波法^[7-8]是在假定已知噪声统计特性的情况下得出预测值,事实上噪声的统计特性难以估计;空间相关性法^[9]则是利用风场和邻近点的风速数据做预测,其预测效果较好,但原始数据收集量较大。

由于风速序列的高度非线性,如ARMA的线性模型不能进行很好地预测。神经网络^[10-11]则能够很好地逼近非线性函数。其处理非线性问题所表现出很强的优越性。然而神经网络也存在明显的不足,一些研究表明,对网络的配置和训练是NP问题,其自身缺乏严格的理论指导。如何最优地构造和训练一个神经网络,保证神经网络的泛化能力,更多的取决于使用者的经验和不断的试凑。1990年,Hansen和Salamon^[12]证明:通过训练多个神经网络并将其集成输出,会显著地提高神经网络的泛化能力。

本文用AdaBoost算法的思想集成多个神经网络的输出。即把BP神经网络作为弱预测器,通过反复的训练BP神经网络预测输出,最终得到由多个BP神经网络组成的强预测器。

1 预测模型

1.1 BP神经网络

BP神经网络即采用误差反向传播算法的多层前向神经网络。其重要特点是信号前向传递,误差反向传播。前向传递中,输入信号从输入层经隐含层逐层处理,直至输出层。相邻两层神经元完全互连,不相邻层无连接。每一层的神经元状态只影响下一层神经元状态,如果输出层得不到期望输出,则转入反向传播,根据预测误差调整网络权值和阈

基金项目: 国家重点基础研究发展计划项目(973项目)(2012CB215203)。

National Basic Research Program of China(973 Program)(2012CB215203)。

值,使BP神经网络预测输出不断逼近预测输出。其网络结构如图1所示。

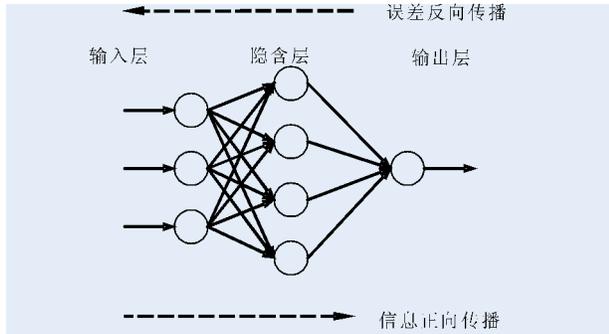


图1 BP神经网络结构

Fig.1 Structure of BPNN

1.2 AdaBoost算法

AdaBoost(Adaptive Boost)是Boosting算法的一种^[13],其主要思想是获取各学习样本的权重分布,最初所有权重被赋予相等的数值,但在训练过程中,这些样本权重被不断调整:预测精度低的样本权重得到加强,预测精度高的样本权重则被减弱。最终,弱预测器加强了对难以预测的样本的学习。这种思想源于Valiant提出的PAC(Probably Approximately Correct,可能近似正确)学习模型^[14]。这样,达到一定预测精度的弱预测器,经组合后形成的强预测器就具有很高的预测精度。由于AdaBoost算法不要求事先知道弱学习算法预测精度的下限而非常适用于实际问题中。

AdaBoost算法可描述为:

给定学习样本

$$(x_1, y_1), \dots, (x_m, y_m), x_i \in X, y_i \in Y, i=1, \dots, m$$

给定样本初始权重

$$D_1(i)=1/m, i=1, \dots, m \quad (1)$$

当 $t=1, 2, \dots, T$

- 1) 利用样本权重 D_t 训练弱学习器;
- 2) 获取弱学习器的预测函数 $h_t: X \rightarrow Y$,并用 $\varepsilon_t = Pr_{i \sim D_t}[h_t(x_i) \neq y_i]$ 表示对应的预测误差;

- 3) 选取 $\alpha_t = \frac{1}{2} \ln \left(\frac{1-\varepsilon_t}{\varepsilon_t} \right)$;

4) 更新样本权重:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & h_t(x_i) = y_i \\ e^{\alpha_t} & h_t(x_i) \neq y_i \end{cases} = \frac{D_t(i) \exp[-\alpha_t y h_t(x_i)]}{Z_t}$$

式中, Z_t 为归一化因子,以使 $\sum_{i=1}^m D_{t+1}(i)=1$ 。

输出最终的预测函数为

$$H(x) = \text{sign} \left[\sum_{t=1}^T \alpha_t h_t(x) \right] \quad (2)$$

2 实例分析

2.1 数据的预处理

根据GB/T18710-2002标准规定:风电场测风数据检验的内容包括数据的完整性检验和合理性检验,合理性检验包括合理性范围的检验、趋势性检验等。

根据GB/T18710-2002标准规定,合理性检验内容主要有:

1) 合理性检验范围:按标准规定,小时平均风速和风向的合理范围在0~40 m/s,0~360°,认为是合理的。超出范围的,认为是不合理的。

2) 趋势性的合理范围:按标准规定,1 h平均风速差的合理范围应小于6 m/s。

风速数据归一化

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad i=1, \dots, m \quad (3)$$

式中, \bar{x}_i 为归一化后的风速值; x_i 为原始风速; x_{\max} 、 x_{\min} 分别为原始风速序列中的最大值和最小值,单位为m/s。

2.2 风速预测

本文从某风电场小时级平均风速选取1 300组数据样本。由于不需要事先知道弱学习算法正确率的下限,对弱预测器,即BP神经网络结构可设置为6-6-1,过去6 h的风速数据作为输入,隐层结点数为6个,预测输出为下一小时的风速值。网络函数参考文献[15],即网络隐层神经元传递函数采用tansig函数,输出层采用purelin函数。训练步数定为50步。为了增加泛化能力,每个弱预测器的训练样本是从前700组风速数据中随机选择500组进行训练。取 $\varepsilon_t=1$,共训练生成不同权重下的10个BP神经网络弱预测器,最后由10个弱预测器组成一个强预测器对风速进行预测。测试样本则是按照时间序列对第701组样本后的50组样本进行测试,图2所示为AdaBoost-BP预测模型示意图,图3所示为原始风速序列。

由图4可知,用AdaBoost结合BP神经网络的预

测误差整体低于用弱预测器的BP神经网络的平均绝对误差。

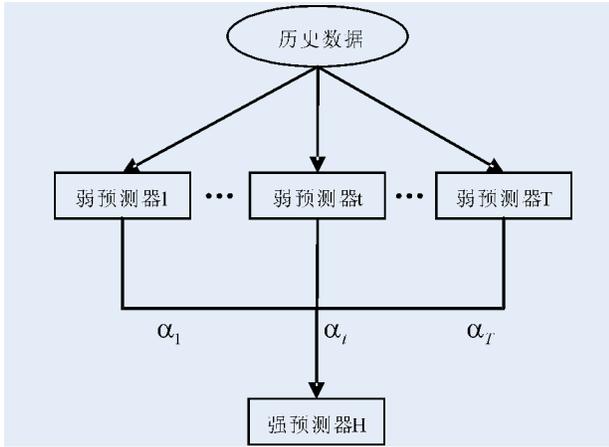


图2 AdaBoost-BP预测模型示意图

Fig.2 Model of AdaBoost-BP

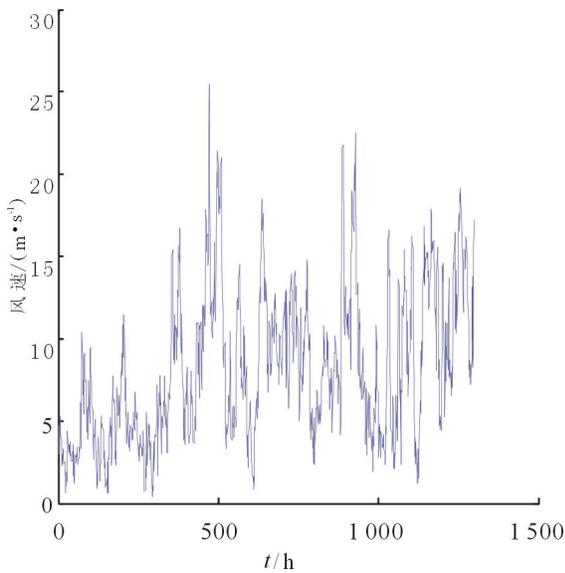


图3 原始风速序列

Fig.3 Wind speed series

图5、图6分别为权重最大和权重最小的BP神经网络与AdaBoost-BP的比较,从图上可以直观的看到,AdaBoost-BP预测结果还是优于单独的BP神经网络,尤其在峰值处AdaBoost-BP明显优于BP神经网络。

图7为AdaBoost-BP预测结果的相对误差,除个别点相对误差较大外,80%的测点的相对误差在20%以内。

本文对701组以后的50组样本做了测试,平均相对误差和平均相对方差AdaBoost-BP优于BP神经网络的平均值。按照时间顺序往后顺推100组,即100 h,并对预测误差做了对比,101-150组数据波

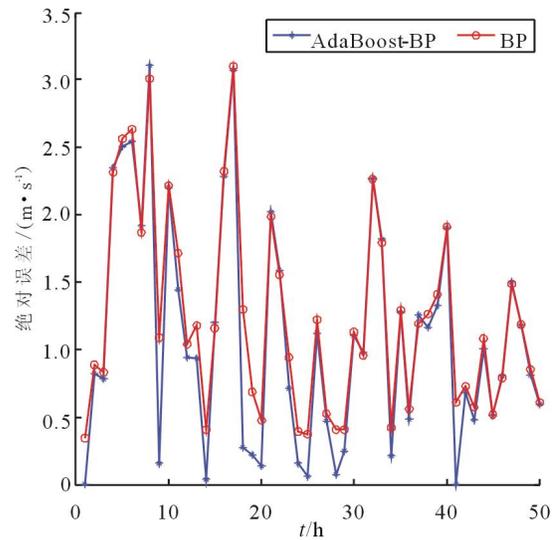


图4 AdaBoost-BP与BP绝对误差对比

Fig.4 AdaBoost-BP and BP absolute error comparison

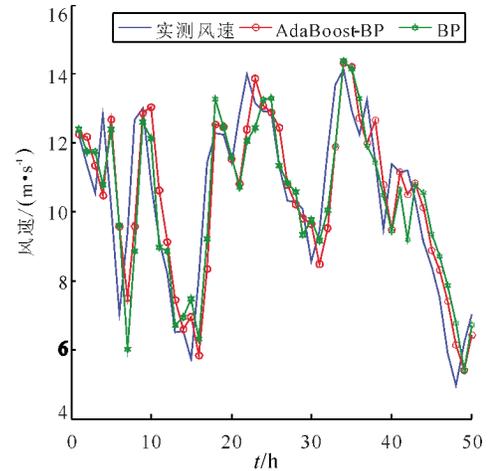


图5 最大权重预测结果图

Fig.5 Prediction with maximum weight

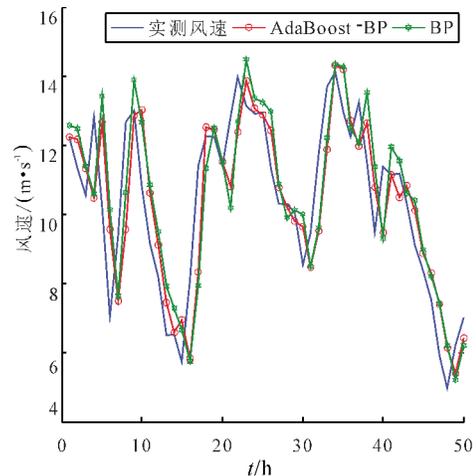


图6 最小权重预测结果图

Fig.6 Prediction with minimum weight

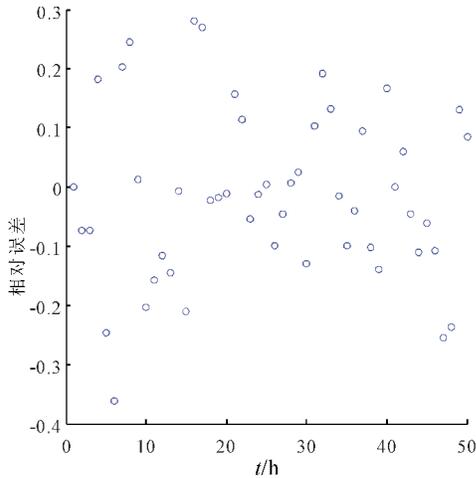


图7 相对误差

Fig.7 Relative error

动较小, 预测精度较高;201-250组数据的波动性较大, 预测结果相对不理想, 整体上随着时间往后推移, 预测精度在降低, 但是AdaBoost-BP优于BP神经网络。从表1结果可以看到AdaBoost-BP在横向统一数据预测结果对比上优于BP神经网络, 纵向时间推移结果上也优于BP神经网络, 有更好的泛化能力。

表1 预测误差结果对比

Tab.1 Error comparison

数据组	平均相对误差		平均相对误差方差	
	Ada Boost-BP	平均BP	Ada Boost-BP	平均BP
1-50	1.081	1.202	0.738	0.853
101-150	0.76	0.810	0.263	0.327
201-250	1.496	1.890	1.343	2.601
301-350	1.240	1.352	1.137	2.000
401-450	1.282	1.414	1.095	1.382
501-550	1.553	1.773	1.410	2.034

3 结论

本文用AdaBoost-BP模型预测风电场的风速, 对AdaBoost-BP模型与BP神经网络模型做了对比分析。

1) AdaBoost-BP的短期预测结果优于单个BP神经网络模型, 尤其在风速变化幅度较大时, 预测结果明显优于单个BP神经网络; 同时, 采用AdaBoost-BP模型的预测结果优于多个弱预测器的简单均值结果, 验证了AdaBoost的有效性。

2) 相对于BP神经网络而言, AdaBoost-BP提升了泛化能力。随着时间的推移, AdaBoost-BP模型的预测结果的优越性体现的更为明显, 除个别组风速数据的波动性较大, 使预测结果相对不理想。

在本文的基础上, 可采用多个不同结构的弱预测器有利于风电场的安全运行。

参考文献

- [1] 2009年世界风能报告 [EB/OL].http://www.wwindea.org/home/images/stories/worldwindenergyreport2009_cn.pdf.
- [2] 2009年中国风电装机容量统计 [EB/OL].<http://cwea.org.cn/upload/201006102.pdf>.
- [3] 傅旭, 李海伟, 李冰寒. 大规模风电场并网对电网的影响及对策综述 [J]. 陕西电力, 2010(1): 53-57.
FU Xu, LI Hai-wei, LI Bing-han. Review on influences of large-scale wind farms power systems and countermeasures [J]. Shaanxi Electric Power, 2010(1): 53-57 (in Chinese).
- [4] Alexiadis M, Dokopoulos P, Sahsamanoglou H, et al. Short-term forecasting of wind speed and related electrical power [J]. Solar Energy, 1998, 63(1): 61-68.
- [5] BROWN B G, KATZ R W, MURPHY A H. Time serial models to simulate and forecast wind speed and wind power [J]. Journal of Climate and Applied Meteorology, 1984, 23(8): 1184-1195.
- [6] 周培毅, 张新燕. 基于时间序列与支持向量机的风电场风速预测研究 [J]. 陕西电力, 2009(12): 1-4.
ZHOU Pei-yi, ZHANG Xin-yan. Study on wind speed forecasting of wind farm based on time series and support vector machine [J]. Shaanxi Electric Power, 2009, (12): 1-4 (in Chinese).
- [7] BOSSANYIE A. Short-term wind prediction using Kalman [J]. Wind Engineering, 1985, 9(1): 1-8.
- [8] 王耀南, 孙春顺, 李欣然. 用实测风速校正的短期风速仿真研究 [J]. 中国电机工程学报, 2008, 28(11): 94-100.
WANG Yao-nan, SUN Chun-shun, LI Xin-ran. Short-term wind speed simulation corrected with field measured wind speed [J]. Proceedings of the CSEE, 2008, 28(11): 94-100 (in Chinese).
- [9] ALEXIADIS M C, DOKOPOULOS P S, SAHSAMANOGLU H S. Wind speed and power forecasting based on spatial correlation models [J]. IEEE Transactions on Energy Conversion, 1999, 14(3): 836-842.
- [10] KARINIOTAKIS G N, STAVRAKAKIS G S, NOGARETE F F. Wind power forecasting using advanced neural networks models [J]. IEEE Transactions on Energy Conversion, 1996,