

DOI: 10.7667/PSPC150294

优化遗传算法寻优的 SVM 在短期风速预测中的应用

颜晓娟¹, 龚仁喜², 张千锋³

(1. 钦州学院工程训练中心, 广西 钦州 535011; 2. 广西大学电气工程学院, 广西 南宁 530000;
3. 钦州学院科技处, 广西 钦州 535011)

摘要: 针对遗传算法存在的早熟和收敛慢的问题, 提出一种融合小生境算法、免疫算法的优化遗传算法。一方面通过疫苗因子引导初始种群的生成, 使个体具有某些优秀基因, 减少寻优时间, 并随数据的更新, 提出疫苗因子和参数寻优范围的自适应更新机制。另一方面在种群的进化过程中, 通过小生境遗传算法维护种群的多样性。实验结果表明, 将基于优化遗传算法寻优的 SVM 应用到短期风速预测中是可行的, 具有较高的预测精度和收敛速度。

关键词: 优化遗传算法; 短期风速预测; SVM; 参数寻优; 自适应更新

Application of optimization SVM based on improved genetic algorithm in short-term wind speed prediction

YAN Xiaojuan¹, GONG Renxi², ZHANG Qianfeng³

(1. Industrial Training Center, Qinzhou University, Qinzhou 535011, China; 2. School of Electrical Engineering, Guangxi University, Nanning 530000, China; 3. Office of Sci-tech, Qinzhou University, Qinzhou 535011, China)

Abstract: An improved genetic algorithm combined with niche algorithm and immune algorithm is proposed to solve the standard genetic algorithm for prematurity and slow convergence rate. On one hand, vaccine is introduced to guide the generation of initial population, so the individual has some excellent gens which will reduce the optimal time, and an adaptive updating mechanism of vaccine and parameter scope is proposed as the wind speed data renewed. On the other hand, niche genetic algorithm is adopted to keep the diversity of evolution population. Examples show that it is feasible to apply the proposed method in short-term wind speed prediction, with higher forecasting accuracy and convergence speed.

Key words: improved genetic algorithm; short-term wind speed prediction; SVM; parameter selection; adaptive updating

0 引言

风电在为社会提供清洁能源的同时, 也给电力系统的稳定运行带来了一系列新问题。由于风速, 风向的随机性, 风电机组的输出功率也具有随机性, 大规模风电并网后, 会对电力系统的电能质量带来很多不利影响, 也会给电力系统调度计划的制定带来困难^[1-4]。风速预测对降低风电并网的成本以及电力部门合理的制定调度计划具有重要意义。

风速预测是根据已有的数据样本对将来的风速值进行预测, 其预测精度的高低是关键问题。目前,

短期风速预测的开展一般是基于历史数据进行的, 其方法主要有: 卡尔曼滤波法^[5]、时间序列法^[6]、人工神经网络^[7-8]、支持向量机(SVM)^[9]等。其中, SVM 以结构风险最小化为目标, 预测结果一般比基于经验风险最小值的神经网络算法更为精确^[10], 另外, SVM 建立在小样本学习的基础上^[11], 因此受到了广泛的重视, 但其惩罚参数和核参数设置的好坏对 SVM 预测模型的性能具有很大影响。文献[12]利用遗传算法(GA 算法)对 SVM 的相关参数进行寻优, 有效提高了短期预测的精度, 但 GA 算法存在收敛速度慢和早熟收敛的问题。文献[13]采用小生境算法对 GA 算法进行优化, 建立 SVM 短期预测模型, 该方法改进了 GA 算法的早熟收敛问题, 提高了模型的预测精度, 但收敛速度慢的问题没有得到有效改进。

基金项目: 广西高校科学技术研究项目(KY2015YB312); 广西壮族自治区中青年骨干教师基础能力提升项目资助(基于IA-NGA 算法寻优的 SVM 短期风能预测研究)

本文采用优化的遗传算法对 SVM 的参数进行寻优, 利用遗传算法运行过程的并行性、寻优过程中的自组织、自适应机制, 结合免疫算法中疫苗因子的信息处理机制, 小生境算法对种群多样性的维护机制, 并针对风速数据是时序变化的特点, 对疫苗因子和参数寻优范围的自适应更新进行了研究, 给出了基于优化遗传算法寻优的 SVM 短期风速预测模型。实例结果表明, 基于该优化遗传算法的 SVM 风速预测精度更高, 预测时间更短, 证明了优化遗传算法寻优的 SVM 在短期风速预测中是可行的。

1 支持向量机

SVM 的基本思想是寻找 \mathbf{R}^n 上的一个实值函数 $g(x)$, 并能用 $y=g(x)$ 推断任一输入 x 所对应的输出值 y , 对于数据样本 $(x_1, y_1), (x_2, y_2) \cdots (x_m, y_m) (x_i \in \mathbf{R}^n, y_i \in \mathbf{R}, i=1, 2, \dots, m, m$ 为训练样本个数, \mathbf{R} 是实数集合), 其线性回归函数如式(1)所示。

$$y(x) = \langle \omega, \Phi(x) \rangle + b \quad (1)$$

其中: ω 是权值向量; $\Phi(x)$ 是二维空间中的 n 维向量 $x(x \in \mathbf{R}^n)$ 非线性映射到 Hilbert 空间的转换; b 为常数。

根据结构风险最小化原则, 将式(1)转化为

$$\begin{aligned} \min_{\omega, b, \xi_i^*} & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \\ \text{s.t.} & \begin{cases} (\langle \omega, \Phi(x_i) \rangle + b) - y_i \leq \varepsilon + \xi_i, i=1, \dots, m \\ y_i - (\langle \omega, \Phi(x_i) \rangle + b) \leq \varepsilon + \xi_i^*, i=1, \dots, m \\ \xi_i^* \geq 0, i=1, \dots, m \end{cases} \end{aligned} \quad (2)$$

其中: 松弛变量 ξ_i^* 的值表示对应点离群的远近; 惩罚参数 C 的大小代表回归模型对离群点的重视程度。 C 值的引入在模型的泛化性和精确性之间找到平衡, 其值选择的合理与否, 深刻影响着回归模型的泛化性能。

为了解式(2)表示的优化问题, 引入拉格朗日乘子向量, 将其转化为

$$\begin{aligned} \min_{a_i^* \in \mathbf{R}^{2m}} & \frac{1}{2} \sum_{i,j=1}^m (a_i^* - a_i)(a_j^* - a_j) K(x_i, x_j) + \\ & \varepsilon \sum_{i=1}^m (a_i^* + a_i) - \sum_{i=1}^m y_i (a_i^* - a_i) \\ \text{s.t.} & \begin{cases} \sum_{i=1}^m (a_i^* - a_i) = 0 \\ 0 \leq a_i^* \leq C, i=1, \dots, m \end{cases} \end{aligned} \quad (3)$$

其中: $\alpha^* = (\alpha_1, \alpha_1^*, \dots, \alpha_m, \alpha_m^*)^T$ 是拉格朗日乘子向

量; $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ 为核函数, 表示转换函数的内积值。

最后, 根据最优解 $\alpha^* = (\bar{\alpha}_1, \bar{\alpha}_1^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$ 构造决策函数如式(4)所示。

$$f(x) = \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i) K(x_i, x) + \bar{b} \quad (4)$$

其中: $\bar{\alpha}_i^*$ 、 $\bar{\alpha}_i$ 是位于开区间 $(0, C)$ 的 Lagrange 乘子分量。 \bar{b} 的值由位于 $(0, C)$ 间的 $\bar{\alpha}^*$ 分量确定, 若选择 $\bar{\alpha}_j$, 其值则如式(5)所示。

$$\bar{b} = y_j - \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i)(x_i \cdot x_j) \quad (5)$$

支持向量机常用的核函数有多项式核函数、高斯径向基(RBF)核函数, 以及 Sigmoid 核函数三种。由于 RBF 核函数直观地反映了两个数据向量间的距离, 所以该核函数的应用范围最为广泛。因此, 本文选用高斯径向基核函数作为 SVM 回归模型的核函数, RBF 核函数公式如式(6)所示。

$$K(x, x') = \exp(-\|x - x'\|^2 / \delta^2) \quad (6)$$

核函数确定以后, 利用 SVM 对风能进行回归拟合预测涉及到的参数就是惩罚参数 C 和核参数 δ 。

2 基于优化遗传算法的 SVM 参数寻优

2.1 小生境免疫遗传算法

遗传算法(Genetic Algorithm, GA)源于对生物进化机制的模拟研究, 是一种适合于复杂系统优化的一种随机全局搜索和优化算法。

小生境算法, 通过对小生境环境中不同物种间竞争与合作的模拟, 在某一给定的环境中, 比较种群内个体间的相似程度而对个体的适应度进行调整, 使算法达到维护种群多样性的目的。

免疫算法 IA(Immune Algorithm)中的疫苗因子是指从具体待求问题的先验知识中提取出的一种包含一个或多个连续基因的特征信息, 同时也可以将其看作是对待求的最佳个体匹配模式的一种估计。

标准遗传算法对问题的领域和种类没有限制, 是一种不依赖具体问题的并行搜索方法, 设置的参数也较少, 但缺点也较为明显: 1) 经过多次迭代后, 种群内适应度大的个体间会形成“近亲繁殖”, 产生早熟和多样性减少的现象; 2) 种群的初始值随机产生, 从开始到最后得到结果的时间较长, 效率低。而小生境算法在迭代的过程中通过对个体适应度进行调整, 能有效地避免适应度较大的个体迅速占据整个种群的情况发生; 通过免疫算法中的疫苗因子

对初始种群进行免疫接种，能有效减少种群的进化时间，提高搜索效率。为此，在遗传算法中引入小生境算法和免疫算法，提出优化的遗传算法。其算法原理如下。

利用疫苗因子引导初始种群的生成，使初始种群内的个体具有优秀个体的某些片段；对个体进行遗传算子(选择、交叉、变异)操作计算之后，比较个体间的适应度距离，如果距离在小生境半径内，则对适应度小的个体进行惩罚，保证了种群的多样性，避免遗传算法中的早熟现象。

优化遗传算法有 3 个核心部分。

1) 小生境半径(L)的选择，利用文献[13]的方法设置 L ，使其跟随种群进化自适应地调整大小。

2) 疫苗因子的生成：在优化遗传算法中，小生境遗传算法搜索到的适应度最大的个体可以认为是参数(C 、 δ)的最优值，用它作为初始疫苗对优化遗传算法的初始种群进行接种，并将得到的最优解作为疫苗因子。

3) 参数寻优范围自动更新：风速数据是按时序变化的，在数据更新时，参数的寻优范围也必须更新，如果参数(C 、 δ)寻优范围设计的较大，会使参数的寻优时间增长，而较小又会使参数收敛于局部最优解。为了在算法的搜索速度与搜索精度间取得平衡，在算法初始阶段对疫苗因子(最优参数(C 、 δ))所处参数范围的位置进行判断：如果在极小值的某个范围内，则相应的减小参数的设定范围；如果在极大值的某个范围内，则增加相应参数的范围；否则参数的寻优范围保持不变。

2.2 算法的实现步骤

步骤 1: 初始化遗传算法的控制参数(种群规模、遗传代数、交叉和变异概率等)和参数(C 、 δ)寻优范围；

步骤 2: 判断疫苗因子是否达到相关参数(C 或 δ)寻优范围的某个极值附近，并对寻优范围作出相应调整；

步骤 3: 利用疫苗因子引导初始种群的生成，计算种群内个体的适应度；

步骤 4: 结束条件判断，如果满足结束条件，将适应度最优个体的编码储存在疫苗因子中，输出寻优结果，否则转向步骤 5；

步骤 5: 进行选择、交叉，变异等遗传操作；

步骤 6: 比较两两个体间的高斯距离 d ，如果 d 在 L 内，则对适应度小者进行惩罚操作；

步骤 7: 根据各个体新的适应度值对个体的适应性进行排序，用前 M 个个体的最小高斯距离更新 L 并转到步骤 4。

3 算例及分析

为了验证优化遗传算法在搜索精度和速度上的有效性，编制程序算法，并设定参数：种群规模 $N=40$ ，遗传代数 $gen=300$ ，交叉概率 $P_x=0.7$ ，二进制编码长度为 2×8 ，根据文献[14]的研究，当 C 固定， δ 在 0 至某一小范围内的增加时，测试精度快速增加而训练精度略有降低；当 δ 固定， C 在 0 至某一范围内增加时，测试和训练精度均随 C 的增加而快速增加，因此，根据多次实验结果，将数据归一化后参数的初始寻优范围设定为 $C[0,10]$ 、 $\delta[0,5]$ 。

采用内蒙古某地区 2009 年 10 月—2010 年 5 月，某 70 m 高测风塔测量的小时风速数据，按时间顺序随机选择 501 个数据，利用前 477 个数据建立预测模型，对后 24 个数据进行预测。针对同样的风速数据样本，采用文献[15]的 GA(ga-svm)寻优、文献[13]的 L 值变化的 GA(nga-svm)寻优、文献[16]直接设定相关参数的 SVM 预测方式，以及 BP 神经网络预测方法进行预测，各方法的预测时间和绝对误差值如表 1 所示。同时，为了验证疫苗因子和参数寻优范围跟随数据变化自动更新后，模型的预测精度情况，对选择训练数据后的 1-30(每次更新 10 个数据，分三次完成)样本进行测试。具体的预测结果如图 1~图 3 所示，预测时间和 e_{MAPE} 如表 2 所示。

表 1 各种方法的风速(m/s)预测时间和误差绝对值

Table 1 Forecast times and errors of wind speed (m/s)

forecasting by different algorithm					
指标	本文方法	nga-svm	ga-svm	svm	BP
T/s	108.37	163.49	154.64	24.08	22.81
e_{MAPE}	4.50%	5.36%	7.23%	8.93%	15.09%

表 2 数据更新后的风速(m/s)预测时间与 e_{MAPE}

Table 2 Wind speed (m/s) predicted times and e_{MAPE} when

data updated			
样本序号	1-10	11-20	21-30
T/s	83.37 s	84.52 s	90.05 s
e_{MAPE}	3.82%	3.91%	5.14%

由表 1 可见，在给定数据样本的情况下，本文的优化遗传算法能够寻找到更适合数据样本的相关参数值，平均绝对百分比预测误差 e_{MAPE} 明显小于 BP 神经网络和直接设定相关参数的 SVM 得到的 e_{MAPE} 、小于 ga 寻优和小生境半径值变化的 nga 寻优的 SVM 预测 e_{MAPE} ，预测精度较高；寻优时间比

BP 神经网络和不经寻优直接设定参数的 SVM 预测要长, 但比 ga 寻优、nga 寻优时间短。

从图 1~图 3 和表 2 可见, 在风速预测模型中, 在风速数据序列更新的情况下, 通过参数寻优范围和疫苗因子的自动更新并引导初始种群的产生, 不仅能减少寻优时间, 而且风速预测的 e_{MAPE} 较小, 预测精度较高。

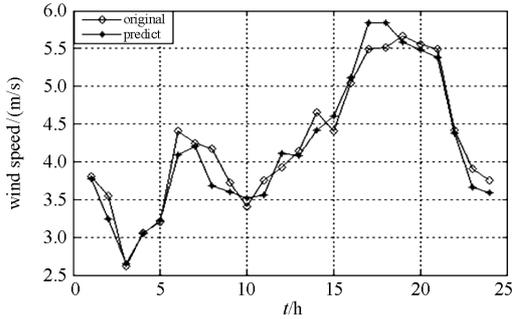


图 1 风速(m/s)预测结果-数据更新 1-10

Fig. 1 Predicted results when data updated 1-10

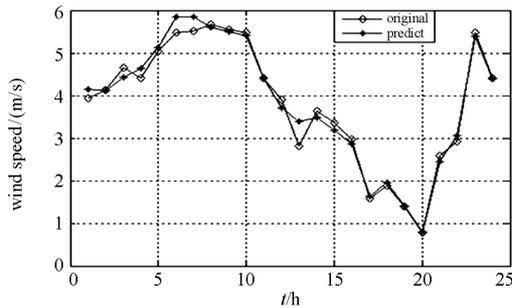


图 2 风速(m/s)预测结果-数据更新 11-20

Fig. 2 Predicted results when data updated 11-20

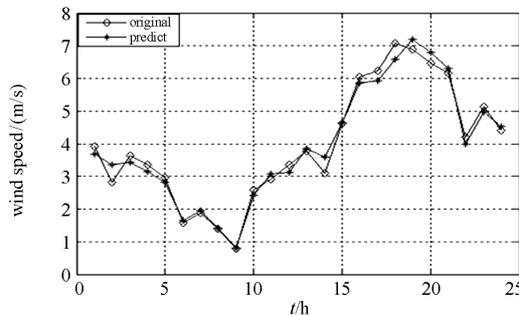


图 3 风速(m/s)预测结果-数据更新 21-30

Fig. 3 Predicted results when data updated 21-30

4 结语

本文将基于优化遗传算法寻优的 SVM 应用于短期风速预测中, 利用小生境算法维护种群多样性

的特点, 结合遗传算法的全局收敛能力, 并引入免疫算法中的疫苗因子引导初始种群的生成, 在减少搜索时间的同时保证种群的多样性, 避免遗传算法的早熟和寻优慢的现象, 提高了寻优速度和精度。通过与其他优化方法的比较, 仿真表明本文的优化遗传算法不仅有良好的收敛速度, 而且能寻找到更适合数据样本的参数值。在数据按时间序列更新时, 疫苗和参数的寻优范围进行自适应的调整更新。因此, 作为一种新的寻优方法, 该优化方法具有一定的理论意义和实际应用价值, 可以应用到其他的时间序列预测模型中。

参考文献

- [1] 廖旋焕, 胡智宏, 马莹莹, 等. 电力系统短期负荷预测方法综述[J]. 电力系统保护与控制, 2011, 39(1): 147-152.
LIAO Nihuan, HU Zhihong, MA Yingying, et al. Review of the short-term load forecasting methods of electric power system[J]. Power System Protection and Control, 2011, 39(1): 147-152.
- [2] 马斌, 张丽艳, 郭成. 一种变权重风电功率最优组合预测模型[J]. 电力系统保护与控制, 2016, 44(5): 117-121.
MA Bin, ZHANG Liyan, GUO Cheng. An optimal combination forecasting model with variable weight for wind power[J]. Power System Protection and Control, 2016, 44(5): 117-121.
- [3] 高阳, 张倩然, 许傲然, 等. 聚类分析在短期风电功率预测模型中的应用[J]. 电器与能效管理技术, 2015(4): 12-15.
GAO Yang, ZHANG Qianran, XU Aoran, et al. Application of cluster analysis in the short-term wind power prediction model[J]. Electrical & Energy Management Technology, 2015(4): 12-15.
- [4] 时锐, 焦然, 陈艳霞. 含大规模风电电网的静态电压仿真分析[J]. 高压电器, 2014, 50(12): 105-110.
SHI Rui, JIAO Ran, CHEN Yanxia. Static voltage simulation analysis of grid with large-scale wind power[J]. High Voltage Apparatus, 2014, 50(12): 105-110.
- [5] 修春波, 任晓, 李艳晴, 等. 基于卡尔曼滤波的风速序列短期预测方法[J]. 电工技术学报, 2014, 29(2): 253-259.
XIU Chunbo, REN Xiao, LI Yanqing, et al. Short-term prediction method of wind speed series based on Kalman filtering fusion[J]. Transactions of China Electrotechnical Society, 2014, 29(2): 253-259.
- [6] 丁明, 张立军, 吴义纯. 基于时间序列分析的风电场

风速预测模型[J]. 电力自动化设备, 2005, 25(8): 32-34.
 DING Ming, ZHANG Lijun, WU Yichun. Wind speed forecast model for wind farms based on time series analysis[J]. Electric Power Automation Equipment, 2005, 25(8): 32-34.

[7] 肖迁, 李文华, 李志刚, 等. 基于改进的小波-BP神经网络的风速和风电功率预测[J]. 电力系统保护与控制, 2015, 42(15): 80-86.
 XIAO Qian, LI Wenhua, LI Zhigang, et al. Wind speed and power prediction based on improved wavelet-BP neural network[J]. Power System Protection and Control, 2015, 42(15): 80-86.

[8] 周建强, 李玉娜, 屈卫东, 等. 基于风速时空信息的BP神经网络超短期风速预测研究[J]. 电网与清洁能源, 2015, 31(1): 109-112.
 ZHOU Jianqiang, LI Yuna, QU Weidong, et al. A study on BP neural ultra-short term wind speed forecast based on historical and spatial wind speed data[J]. Power System and Clean Energy, 2015, 31(1): 109-112.

[9] 张华, 曾杰. 基于支持向量机的风速预测模型研究[J]. 太阳能学报, 2010, 31(7): 928-932.
 ZHANG Hua, ZENG Jie. Wind speed forecasting model study based on Support Vector Machine[J]. Acta Energetica Sinica, 2010, 31(7): 928-932.

[10] 刘苏苏, 孙立民. 支持向量机与RBF神经网络回归性能比较研究[J]. 计算机工程与设计, 2011, 32(12): 4202-4205.
 LIU Susu, SUN Limin. Performance comparison of regression prediction on support vector machine and RBF neural network[J]. Computer Engineering and Design, 2011, 32(12): 4202-4205.

[11] CORTES C, VAPNIK V N. Support vector network[J]. Machine Learning, 1995, 20(3): 273-295.

[12] 刘爱国, 薛云涛, 胡江鹭, 等. 基于GA优化SVM的风电功率的超短期预测[J]. 电力系统保护与控制, 2015, 43(2): 90-95.
 LIU Aiguo, XUE Yuntao, HU Jianglu, et al. Ultra-short-term wind power forecasting based on SVM optimized by GA[J]. Power System Protection and Control, 2015, 43(2): 90-95.

[13] 颜晓娟, 龚仁喜. 基于改进遗传算法寻优的SVM风能短期功率预测[J]. 电测与仪表, 2014, 51(8): 38-41.
 YAN Xiaojuan, GONG Renxi. Short-term wind power prediction based on SVM and improved genetic optimization[J]. Electrical Measurement & Instrumentation, 2014, 51(8): 38-41.

[14] 彭璐. 支持向量机分类算法研究与应用[D]. 长沙: 湖南大学, 2007.
 PENG Lu. The study and application of SVC algorithm[D]. Changsha: Hunan University, 2007.

[15] 杨洪, 古世甫, 崔明东, 等. 基于遗传优化的最小二乘支持向量机风电场风速短期预测[J]. 电力系统保护与控制, 2011, 39(11): 44-48.
 YANG Hong, GU Shifu, CUI Mingdong, et al. Forecast of short-term wind speed in wind farms based on GA optimized LS-SVM[J]. Power System Protection and Control, 2011, 39(11): 44-48.

[16] 戚双斌, 王维庆, 张新燕. 基于SVM的风速风功率预测模型[J]. 可再生能源, 2010, 28(4): 25-29.
 QI Shuangbin, WANG Weiqing, ZHANG Xinyan. Model building for wind speed and wind power prediction based on SVM[J]. Renewable Energy Resources, 2010, 28(4): 25-29.

收稿日期: 2015-05-27; 修回日期: 2016-01-10

作者简介:

颜晓娟(1985-), 女, 硕士, 助教, 研究方向为智能优化控制与智能信息处理; E-mail: 617956088@qq.com

龚仁喜(1962-), 男, 博士, 教授, 研究方向为微电子技术方面的教学和科研; E-mail: rxgong@gxu.edu.cn

张千锋(1982-), 男, 通信作者, 硕士, 副教授, 研究方向为智能信息处理及其在机械设计中的应用。E-mail: 119064088@qq.com

(编辑 姜新丽)