

改进的支持向量回归机在电力负荷预测中的应用

唐承娥¹ 韦 军²

1 广西职业技术学院机电与信息工程学院 南宁 530226

2 广西壮族自治区招生考试院 南宁 530021

摘要 电力预测是一项重要的工程应用。为了解决多层次粒度支持向量回归机(Dynamical Granular Support Vector Regression Machine, DGSVRM)预测电力负荷精度的问题,提出一种基于萤火虫群优化(Glowworm Swarm Optimization, GSO)算法与模式搜索算法(Pattern Search, PS)的混合算法来优化 DGSVRM 预测模型的关键参数。仿真实验表明,通过优化参数之后,预测模型的预测精度得到很大提高。

关键词:多层次粒度支持向量回归机;萤火虫群优化;模式搜索算法

中图分类号 TP391

Application of Power Load Prediction Based on Improved Support Vector Regression Machine

TANG Cheng-e¹ and WEI Jun²

1 College of Electromechanical and Information Engineering, Guangxi Vocational and Technical College, Nanning 530226, China

2 Guangxi Zhuang Autonomous Region Admission Examination Institute, Nanning 530021, China

Abstract Electricity forecasting is an important engineering application. In order to solve the accuracy problem of dynamical granular support vector regression machine for power load forecasting (DGSVRM), this paper proposes a hybrid algorithm of glowworm swarm optimization (GSO) and pattern search (PS) to optimize the key parameters of DGSVRM forecasting model. Simulation results show that the prediction accuracy is greatly improved by optimizing the parameters of the prediction model.

Keywords Dynamical granular support vector regression machine, Glowworm swarm optimization, Pattern search algorithm

1 引言

国家电力理性发展规划及电力企业高效运行的基础之一就是全社会用电量的合理预测。社会用电量的预测不仅可以正确引导电力系统的发电计划,还可以调整社会各种能源的配置,这就需要我们加强对社会电力电量负荷的预测^[1]。

从 20 世纪 50 年代开始,世界各国学者和各大电力公司的研究人员投入到理论模型研究、算法设计、软件开发等方面展开研究,且取得了一些优异的模型、算法。无论是利用传统方法预测^[2-3],还是结合新的理论技术(如人工智能技术、非微分几何代数等),都不同程度地提高了预测精度。但基于概率统计的一些预测方法无法解决非线性、复杂多变的电力负荷问题。新的理论技术虽然可以解决非线性、复杂多变的电力负荷预测问题,但也带来计算量大、参数设计困难及收敛慢等问题。研究学者将神经网络及相应的改进措施用于电力负荷预测^[4-6],而神经网络模型面临优化参数较多的问题,在工程领域的应用受到影响。另外,神经网络模型也存在过学习和预测扩展能力不足的缺点。支持向量机^[7](Support Vector Machine, SVM)可以解决学习过程中的维数不确定和过度学习等问题^[8],因此比神经网络模型有更好的扩展预测能力^[9]。SVM 被应用在模式识别、回归分析、函数估计等领域。研究学者提出将孪生支持向量机、模糊支持向量机、粒度支持向量

机及结合粒子群算法等模型应用在各种电力负荷预测场景^[10-12],提高了预测能力。

根据过去大部分研究结果^[13-15],SVM 预测模型的预测任务是找出正确设置或者优化惩罚因子、核函数及其参数、损失函数等关键参数。因此,在使用 SVM 进行预测时,研究人员都注重对 SVM 参数优化进行理论研究,利用各种工具或者算法进行优化。研究人员采用模拟退火算法、差分进化算法、果蝇算法、遗传算法、混沌粒子群算法、人工蚁群算法、免疫算法或上述算法的混合等优化 SVM 模型参数^[16-18],提高了模型的预测能力和精度。从已有研究中发现,有些智能算法善于全局搜索,而局部搜索能力较弱,对提高解的精度不起作用;而另一些智能算法善于局部搜索,全局搜索能力则较弱,容易陷入局部最优值收敛,导致不容易找到 SVM 预测模型的最优参数^[19]。

Krishnanand 和 Ghose 等^[20-23]提出了一种新的群体智能算法——萤火虫群优化(Glowworm Swarm Optimization, GSO)算法。该算法是根据模拟自然界萤火虫在晚上群聚活动的自然现象而提出的,萤火虫的群聚主要是通过萤火虫发出荧光素与同伴交流进行各项活动。GSO 算法拥有找出全局极值并在多极值问题中找到多个极值的能力,已被应用到许多工程领域。多层次粒度支持向量回归机^[24](DGSVRM)能够以较快的速度完成动态粒划的过程并收敛,在保持较高

基金项目:2018 年广西高校中青年教师基础能力提升项目(2018KY0956)

This work was supported by the 2018 Guangxi College Young and Middle-aged Teacher's Basic Ability Improvement Project(2018KY0956).

通信作者:唐承娥(735438514@qq.com)

训练效率的同时可有效提高传统粒度支持向量回归机的泛化性能。

为了充分发挥萤火虫群优化算法的全局搜索能力和模式算法的局部搜索能力,在 DGSVRM 训练效率和泛化能力有保证的基础上,本文提出一种新的混合算法 GSOPS——萤火虫群优化算法与模式搜索算法的混合,来优化 DGSVRM 的 5 个关键参数。对 5 个重要参数进行最优化,就变成考虑 5 个变量的组合优化问题。在构建 GSOPS 的过程中,GSO 找出全局解空间,并负责寻找最优解的潜在区域;而 PS 则负责精细化搜寻潜在区域,从而提高电力负荷预测的精度。为验证混合优化算法优化 DGSVRM 的效果,利用两个实际负荷案例进行了验证,选择 4 种算法与 GSOPS 算法进行比较来说明 GSOPS 优化 DGSVRM 的效果,选择 3 种模型与 DGSVRM 进行比较来说明 DGSVRM 的预测性能。

2 多层次粒度的支持向量回归机

2.1 基本的支持向量机

设给定原始训练集为 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset R^n \times R$ 。 n 表示训练集数据数量, x_i 为输入特征向量, y_i 为输出值。 SVM 找到最优的回归超平面, 设定目标函数:

$$f(x) = \langle w, \phi(x) \rangle + b \quad (1)$$

$$R = R_{\text{emp}} + \frac{1}{2} \|w\|^2 = \frac{C}{n} \sum_{i=1}^n L(y_i, f(x)) + \frac{1}{2} \|w\|^2 \epsilon \quad (2)$$

$$L(y_i, f(x)) = \begin{cases} |y - f(x)| - \epsilon, & \text{if } |y - f(x)| \geq \epsilon \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中, \langle, \rangle 是内积函数的符号, w 代表权重向量, b 表示偏移向量, $\phi(x)$ 是映射函数, 用来表示输入空间 x 的高维空间。式(2)中的第一项 R_{emp} 或 $\frac{C}{n} \sum_{i=1}^n L(y_i, f(x))$ 代表训练误差。在 SVRM 模型中, Vapnik^[25] 在式(3)中用 ϵ 表示不敏感损失度, 通常用式(2)评估模型的误差程度。式(2)中的 $\frac{1}{2} \|w\|^2$ 表示正则项, 代表模型的复杂度; C 表示惩罚因子, 用来平衡训练误差和正则项。为了充分发挥 SVRM 的预测能力, 采用智能算法来寻优参数 C 和 ϵ 等。

把两个正数的松弛变量 ξ 和 ξ^* 代入式(2), 则目标函数和它的约束条件如下:

$$\min R = C \sum_{i=1}^n (\xi + \xi^*) + \frac{1}{2} \|w\|^2$$

$$\text{s. t. } y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi \quad (4)$$

$$\langle w, x_i \rangle + b - y_i \leq \epsilon + \xi^*$$

$$\xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n$$

对式(4)使用沃尔夫对偶定理及根据鞍点的条件, 有:

$$\max_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle \phi(x_i), \phi(x_j) \rangle - \epsilon \sum_{i,j=1}^n (\alpha_i + \alpha_i^*) + \sum_{i,j=1}^n y_i (\alpha_i - \alpha_i^*) \quad (5)$$

$$\text{s. t. } \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) = 0 \text{ and } \alpha_i, \alpha_i^* \in [0, c] \quad (6)$$

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \phi(x_i)$$

通过解二次线性规划方程求得 α_i 和 α_i^* , 两个常数表示非负的拉格朗日系数。

为了解决在高维空间中内积运算难的问题, 利用核函数把问题从低维度转到高维度进行求解, 则式(1)就可以转为:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \quad (7)$$

其中, $K(x_i, x_j)$ 代表核函数。当核函数满足 Mercer 条件并且满足 $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ 时, 采用核函数就可以降低计算复杂度。通常, 核函数选择线性核函数、高斯函数和径向基函数、S 型核函数等^[25]。根据许多文献选取核函数的经验, 本文选取径向基核函数^[26], 如式(8)所示, 其中 δ 为核参数, 代表了映射高维空间的结构, 影响模型的复杂度。为了提高 SVRM 的运算效率, 采用算法对 δ 进行优化, 选出合适的值。

$$K(x_i, x_j) = \exp\left[-\frac{\|x_i - x_j\|^2}{2\delta^2}\right] \quad (8)$$

2.2 基于多层次粒度的 SVRM

为了提升 SVRM 的训练效率, 许多研究人员通过粒度计算理论构建了粒度支持向量回归机模型^[27-28]。这些研究大多运用静态粒度(一次单层粒划)的方式, 很少运用多层次粒度(多次多层粒度)。为此, 本文利用粗粒度划分样本的方式改进 DGSVRM 模型, 克服静态粒划分的缺点。

假设原始训练集为 $T = (X, D) = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset R^n \times R$, 通过 ϕ 作用后, 在多维空间 R^N 记为 $T = \{(\phi(x_i), y_i), i = 1, \dots, l\}$, 样本分有 k 个粒 G_1, \dots, G_k , 其中 $G_i = \{(\phi(x_{ij}), y_{ij})\}, i = 1, \dots, k, j = 1, \dots, n_i$ (n_i 为第 i 个粒中样本的个数), 定义粒的中心和半径概念如下。

定义 1 在空间 N 维中, 粒 X_i 被看成单个粒超球, 则 X_i 的中心 μ_i , 半径 r_i 如下:

$$\mu_i = \frac{1}{n_i} \sum_{p=1}^{n_i} \phi(x_p)$$

$$= \sqrt{\frac{1}{n_i^2} \left(\sum_{p=1}^{n_i} \phi(x_p) \right)^2}$$

$$= \frac{1}{n_i} \sqrt{\sum_{p=1}^{n_i} \sum_{q=1}^{n_i} K(x_p, x_q)} \quad (9)$$

$$r_i = \max_{x_s \in G_i} (\phi(x_s) - \mu_i)$$

$$= \max_{x_s \in G_i} \sqrt{(\phi(x_s))^2 - 2\phi(x_s) \cdot \mu_i + \mu_i^2}$$

$$= \max_{x_s \in G_i} \sqrt{K(x_s, x_s) - \frac{2}{n_i} \sum_{p=1}^{n_i} K(x_s, x_p) + \frac{1}{n_i^2} \sum_{p=1}^{n_i} \sum_{q=1}^{n_i} K(x_p, x_q)} \quad (10)$$

根据定义 1, N 维空间中任一样本 $\phi(x_j)$ 到第 i 个粒超球 G_i 的距离为:

$$d(\phi(x_j), G_i) = \sqrt{K(x_j, x_j) - \frac{2}{n_i} \sum_{p=1}^{n_i} K(x_j, x_p) + \frac{1}{n_i^2} \sum_{p=1}^{n_i} \sum_{q=1}^{n_i} K(x_p, x_q)} \quad (11)$$

粒划分算法的具体步骤如算法 1 所示。

算法 1 依照核函数的粒划分

- Step1 输入开始样本集 T 、粒划参数 k 和参数 θ ;
- Step2 选取 k 个样本点, 定义初始化的粒心;
- Step3 根据距离公式(11)计算所有样本点的距离, 采用核空间进行粒划分;
- Step4 利用式(9)调整粒心, 判断粒心是否在给定 θ 的范围内, 若不在则返回 Step3, 否则转到 Step5;
- Step5 停止运算, 得出最优粒集 $X \rightarrow \{G_1, \dots, G_k\}$ 。

把样本粒到初始近似超平面 f_1 的距离与粒的密度进行

比较,再决定进行多层次粒划,进而继续获得近似最优回归的超平面。

定义 2 N 维空间中的单个粒 G_i 与 $f: y = \omega \cdot \phi(x) + b$ 的距离如下:

$$d(G_i, f) = \frac{(\omega - 1) \cdot \mu_i^T + b}{\sqrt{\omega^2 + 1}} - r_i$$

$$= \frac{\frac{1}{n_i} \sum_{k=1}^{n_i} \sum_{j=1}^{|\text{SV}_i|} \alpha_j \cdot y_j \cdot K(x_j, x_k) + b}{\sqrt{\sum_{j=1}^{|\text{SV}_i|} \alpha_j \cdot \alpha_k \cdot y_j \cdot y_k \cdot K(x_j, x_k)}} - r_i \quad (12)$$

其中, SV_i 表示 SVM 的集合。式(12)为粒到超平面的距离。为了更好地刻画这个状态,分别定义信息粒和粒密度。

在 CGSVRM 模型中,有超平面 $f: y = \omega \cdot \phi(x) + b$ 及两者粒 G_i 和 G_j , 两个相距为 2η , 则有以下判断:

假如 $d(G_i, f) > \eta - 2r_i$, 则高概率拥有支持向量的样本信息; 反之, 就低概率拥有支持向量的样本信息。

定义 3 有粒 G_i (用 μ_i 表示粒心, r_i 表示半径), 有 $f: y = \omega \cdot \phi(x) + b$, 如果粒 G_i 到 f 的距离 $d(G_i, f)$ 不超过 $\eta - 2r_i$ (其中 η 表示 SVRM 的回归间隔), 则称粒 G_i 是包含支持向量信息的信息粒。

定义 4 (粒密度) 给定粒 $G_i = \{x_{ij}\}, j = 1, 2, \dots, n_i$, 其中心和半径记为 μ_i 和 r_i , n_i 为样本数量, 那么粒 G_i 的密度 ρ_i 定义为:

$$\rho_i = \frac{n_i}{\sum_{j=1}^{n_i} d(\mu_i, x_{ij})}$$

$$= \frac{n_i}{\sum_{j=1}^{n_i} \sqrt{\frac{1}{n_i} \sum_{p=1}^{n_i} \sum_{q=1}^{n_i} K(x_p, x_q) - \frac{2}{n_i} \sum_{p=1}^{n_i} K(x_p, x_q) + K(x_j, x_j)}} \quad (13)$$

如果粒密度比较大, 需要进行深层次粒划分, 在保证留下包含支持向量信息的同时降低训练运算量; 如果粒密度值小, 表明运算量降低, 就可以对样本进行学习训练。

定义 5 对于第 Le 层的信息粒 $G'_{lev,j}$, 有:

$$k'_{lev,j} = \lceil \frac{r'_{lev,j} - \rho'_{lev,j}}{d_para} \rceil \quad (14)$$

称式(14)为多层次粒划。式(14)中 d_para 表示多层次粒划参数, 是 DGSVRM 进行粒划的关键参数, 直接影响该模型的泛化能力。通过 GSOPS 对 DGSVRM 参数进行优化。在多层次粒划中, 每一层次的支持向量回归机都使用本层全部粒(包含信息粒与非信息粒)的粒心, 计算得出本层的近似回归超平面。在多层次粒划中, 每层的粒数量为:

$$k'_{lev+1} = k_{lev} - k'_{lev} + \sum_{j=1}^{k'_{lev}} k'_{lev+1,j} \quad (15)$$

$T = (X, D) = \{x_i, d_i\}$ 为样本集, $i = 1, \dots, l, x_i \in R^n, d_i \in R$ 。样本划分有 k 个粒。新建的训练样本集合 $T' \in T, T'$ 的规模为 l' ($l' \ll l$)。若在 T' 上获得超平面 f', f' 的误差计算如下:

$$R_{\text{emp}}[f'] = \frac{1}{l'} \sum_{i=1}^{l'} c(x_i, d_i, f'(x), \epsilon) \quad (16)$$

其中, ϵ 表示回归损失函数。

定义 6 对数据集 T 进行处理之后得出新训练集为 T' , f 和 f' 分别表示在 T 和 T' 上获得超平面, 模型误差定义如下:

$$E_M = \lim_{l, l' \rightarrow \infty} |R[f'] - R[f]| \quad (17)$$

根据文献[25]的推论, 可任意逼近原始支持向量机的模型。

算法 2 基于多层次粒划的 SVRM

初始化给定的训练集为 $T_0 = (X_0, D) = \{x_i, d_i\}, i = 1, \dots, l_0$ 且 $x_i \in R^n, d_i \in R$ 。初始化 3 个参数: $k_0, \text{Lev}_0 = 0, d_para$ 。

Step1 初始粒划并进行 SVRM 训练。

采用算法 1 对 T_0 中的样本子集 X_0 进行粒划, 得到 $X_0 \rightarrow \{G_{1,1}, \dots, G_{1,k}\}$ 的粒划集, 其中 $G_{1,i} = \{\phi(x_i)\} (i = 1, \dots, l_{1,i})$, $l_{1,i}$ 表示第一层第 i 个粒中的样本大小, 设 $\text{Lev} = 1$ 。

Step2 多层次粒划。

Step2.1 训练样本集导出近似回归超平面 f_{Lev}

Step2.2 通过式(12)计算距离, 根据定义 3 找出粒 $\{G'_{lev,j}\} (j = 1, \dots, k'_{lev})$;

Step2.3 根据式(13)计算密度 $\rho'_{lev,j}$, 根据式(14)对 Lev 计算 $k'_{lev,j}$;

Step2.4 对上一步算出"数量大于 2 的信息粒再次进行粒划分";

Step2.5 计算粒中心和半径, 若不满足条件则转 Step2.1; 否则执行 Step3。

Step3 SVRM 训练, 得出最优超平面 l_{last} 。

Step4 算法结束。

为了验证 DGSVRM 模型的性能, 将其用于电力负荷预测中。首先, 选择径向基函数作为核函数, DGSVRM 模型需要调整 4 个关键参数, 即多层次粒划参数 d_para 、核参数 δ 、惩罚因子 C 及不损失函数带 ϵ 。4 个参数的设置直接影响 DGSVRM 的预测性能^[29-30]。为了提升该模型的预测精度, 须对上述 4 个参数寻找最优值, 下一节将叙述使用 GSO 和 PS 构造混合算法 GSOPS 来优化参数。

3 基于 GSOPS 算法优化 DGSVRM 的关键参数

本节主要介绍 GSOPS 混合算法的构建过程及如何优化 DGSVRM 参数。混合算法中, 萤火虫算法作为全局搜索工具, 模式搜索为局部搜索工具。

3.1 初始化

DGSVRM 的参数优化问题是一个维度为 4 的连续变量组合优化问题。在用 GSOPS 算法求解这个问题时, 设萤火虫(或个体)表示 4 个参数组合解, 并记为 $X_i = \langle d_para, C, \delta, \epsilon \rangle$ 。一组萤火虫种群由多个萤火虫组成, 为了提高算法的运算效率, 初始种群使用拉丁超立方体抽样法(Latin Hypercube Sampling, LHS)产生初始种群。具体过程是先把待解的空间划分为多个子空间, 再在每个子空间中随机抽取样本点。这样组成的初始种群就可以确保初始种群中的萤火虫个体能相对均匀地分布在全局解空间中, 比随机初始化更能求得最优解。

3.2 适应值函数

DGSVRM 负荷预测模型由于拥有比较强的泛化能力, 为了能准确预测未来的电力负荷, 在使用 GSOPS 优化 DGSVRM 的参数时选取到最优的参数是保证泛化能力的关键。为了找到最优参数, 选取能代表泛化能力的适应值函数来评价非常重要。在验证集上获得定义 5 的模型误差最小值的参数组合解, 即是最优解, 即验证集上的模型误差为适应值函数。

3.3 采用萤火虫优化算法建立全局搜索策略

萤火虫优化算法^[20]主要是通过模拟萤火虫利用发光特性吸引同伴的机制而建立起来的算法。在萤火虫算法中,萤火虫更新位置由式(18)确定:

$$v_i = v_i + \beta_0 e^{-r_{ij}^{\gamma}} (v_j - v_i) + \alpha(\text{rand} - \frac{1}{2}) \quad (18)$$

设 $\gamma=1, \beta_0=1$, 用 αS_k 代替 α , 其中 S_k 为待优化问题各维度的比例参数。式(18)中第二项为萤火虫 j 对 i 的吸引力, 当 $r_{ij}=0$ 时, 吸引力等于 β_0 。式(18)中第三项即随机扰动项, 用于防止算法陷入早熟收敛, 实现全局最优值区域的搜索。

3.4 模式搜索算法负责局部搜索的策略

在 GSOPS 中, 利用模式搜索算法在最优值区域进行进一步的精细化搜索。模式搜索算法是一类采用梯度下降法的常用优化算法, 算法参数简单且计算高效, 在许多参数优化问题中被成功应用^[31]。该算法既可以检验当前解的邻域, 也可以非常高效地在局部区域搜索并收敛至局部最优解。算法的极强收敛性已在有或无约束问题中从理论角度得到证明^[32]。根据已有文献关于该算法的证明, 在局部搜索区域使用模式搜索算法进行搜索, 能提高粒子个体的精细化搜索能力。

算法3 模式搜索算法

Step1 设定搜索半径 Δ_0 , 把待求解问题编码成萤火虫个体 p_0 ; 设初始化搜索半径 $\Delta = \Delta_0$;

Step2 计算搜索 $f(x) = f(p_0) + f(\Delta_0 p_k)$;

Step3 如果当前搜索的解比已有解优, 更新当前个体的位置; 按 $\Delta = \Delta_0/3$ 缩小半径继续搜索;

Step4 若终止条件满足, 结束运行, 否则转到 Step2。

正如算法3所描述, 模式搜索算法先以当前个体 p_0 为初始点, 以初始邻居半径 Δ_0 为大小搜寻 p_0 的邻域, 找到邻域内的最优点, 同时替换当前个体; 如果可以找到比当前个体更优的个体, 就缩小邻域半径 Δ_0 , 用新的邻域半径继续搜寻邻域。不断迭代, 直到达到最大搜索次数或邻域半径达到预先指定的临界值。邻域内的邻居用式(19)求得。

$$P_k = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \quad (19)$$

算法4给出了混合算法 GSOPS 的具体步骤。

算法4 GSOPS

Step1 初始化算法的参数: 种群大小 n , 搜索范围(半径) Δ , 最大迭代数值 Max , 参数组合解的比例参数 S_k 等。

Step2 用拉丁超立方体抽样法产生初始群体 $x_i, i=1, 2, \dots, n$;

Step3 使用式(17)计算种群各个体的适应值, 选定群体中最好适应度函数值的个体的邻域为全局最优解所在的邻域;

Step4 根据式(18)求解粒子函数值, 更新全局最优解;

Step5 使用算法3对全局最优解所在的区域进行更精细的搜索;

Step6 如果满足终止条件, 算法停止; 若不满足终止条件, 转到 Step4。

3.5 GSOPS 算法的构建过程

GSOPS 算法的主要思想是用 GSO 算法在全局解空间中对最优解的可能区域进行搜索, 并利用 PS 算法在局部区域进行个体的开采与学习。为了不过多增加计算量, 在算法中以概率 $p_l(x_k)$ 对萤火虫个体进行局部搜索, 而不是随机进行局部搜索。概率 $p_l(x_k)$ 使用轮盘赌局的方法定义^[33]:

$$p_l(x_k) = \frac{f_{\max}(P) - f(x)}{\sum_{y \in P} (f_{\max}(P) - f(y))} \quad (20)$$

个体适应值越大, 获得局部精细搜索的概率就高。

4 实验的相关设置

4.1 实验数据及预处理

为了验证 DGSVRM-GSOPS 的负荷预测模型的性能, 使用两个实际电力系统的负荷数据进行实验。第一个实际电力数据集是 PJM 公司 Pennsylvania-New Jersey-Maryland¹⁾ 采集监测的电力负荷数据(单位为), 从2010年1月1日至2011年6月31日, 合计18个月, 总计13104h的数据; 第二个数据集是我国东北一城市的月度电力负荷数据^[34], 数据采集时间段为2004年1月-2009年4月, 共64个月。

为了弥补实际电力负荷数据的缺失, 先对数据进行预处理, 采用邻近数据的均值进行补充。使用式(21)对原始电力负荷数据进行归一化处理:

$$x' = \frac{x - \min_A}{\max_A - \min_A} \quad (21)$$

其中, 原始负荷数据用 x 表示, 归一化后的数据用 x' 表示, 原始数据序列 A 的最大值和最小值分别用 \max_A 和 \min_A 表示。式(21)处理后的结果都在 $[0, 1]$ 之间。归一化之后, 能减小误差, 节省存储空间, 并提高训练效率^[35-36]。通过训练, 得到预测结果, 再反向归一化到原始量纲, 在原始数据上对预测误差进行计算。

4.2 评价指标

预测模型性能的评价指标有很多^[37], 具体需要从模型本身及预测实际问题来考虑^[38]。为了检验 DGSVRM 模型的预测性能, 使用平均绝对标准化误差(MASE)、平均绝对百分比误差(MAPE)及方向变化统计量(DS)3个预测指标。三者的定义如式(22)~式(24)所示。此外, 为了更准确地评价 DGSVRM 的泛化能力, 使用模型误差式(17)作为评价指标。MAPE 用于比较真实值与预测值两者之间的百分比误差, 是常用的预测误差评价指标。MAPE 越小, 预测值越逼近于真实值。MASE 是通过 naive 预测模型缩尺的标准化误差, 其值在 0 与 1 之间。MASE 值越小, 表明该预测模型越优于 naive 模型。MASE 利于处理异常值且不敏感、易解释, 得到了许多研究学者的积极推荐^[37]。模型的方向性误差用 DS 来评价, 其值越大, 证明模型的性能越优。根据定义6误差越小时, 表明利用压缩处理得到训练集训练得到的近似回归超平面越接近原始超平面, 即模型泛化能力越接近标准的 SVRM 模型。

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{t+i} - \hat{y}_{t+i}}{y_{t+i}} \right| \times 100 \quad (22)$$

$$\text{MASE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{t+i} - \hat{y}_{t+i}}{\frac{1}{t-1} \sum_{j=2}^t |y_j - y_{j-1}|} \right| \times 100 \quad (23)$$

$$\text{DS} = \frac{\sum_{i=2}^N d_i}{N-1} \times 100, \quad d_i = \begin{cases} 1, & \text{if } (y_{t+i} - y_{t+i-1})(y_{t+i} - \hat{y}_{t+i-1}) \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (24)$$

单一预测样本等于模型所有预测样本的平均误差。

¹⁾ <http://www.pjm.com>

为了多角度比较不同模型,另外使用 Wilcoxon 符号秩检验^[39-40]检测两个电力负荷预测模型的预测误差是否存在显著性差异。

5 实验结果与分析

将第一个 PJM 公司的电力系统的每小时负荷数据序列分为训练集、验证集和测试集 3 部分,具体情况如表 1 所列。先使用训练集和验证集对模型进行学习,确定最优参数;然后使用最优参数建立预测模型,并在训练集和验证集上继续训练预测模型;最后通过测试集校正模型误差。

表 1 PJM 公司的数据划分情况

Table 1 Data division of PJM

数据的类型	各时间段	样本的大小
训练集	2010/1/1—2010/12/31	365 * 24
验证集	2011/1/1—2011/3/31	365 * 24
测试集	2011/4/1—2011/6/31	365 * 24

为了充分判断使用 GSOPS 算法优化电力负荷预测模型中的参数($d_para, C, \delta, \epsilon$)是否达到最优,将其与其他算法的预测模型进行精度的对比。选用的对比算法包括遗传算法(GA)、粒子群优化算法(PSO)、模拟退火算法(SA)和萤火虫算法(GSO)。

初始化 SVRM 参数分别为 $\log_2 C \in [-6, 6]$, $\log_2 \gamma \in [-6, 6]$, $\log_2 \epsilon \in [-6, 6]$, $d_para \in [1, 2.5]$ 。先行实验确定每个算法中的相关参数,各进化算法的种群大小为 50;算法的终止判断条件是:算法迭代次数达到 150 次或者最优适应值连续 50 次无改进。GSO 与 GSOPS 中的参数 S_k 都赋值 1,代表各参数上界为 $\frac{1}{6}$ 。

为了验证 GSOPS 优化 DGSVRM 的性能,采用 4 种常见的算法优化 DGSVRM 负荷预测模型参数之后确定模型,不同评价指标的结果如表 2—表 5 所列。需要指出的是,4 个表中的计算结果都是 30 次独立实验的结果均值,避免了算法的偶然性发生。首先,整体上使用 GSOPS 优化 DGSVRM 预测模型获得了比其他算法更好的结果;其次,采用 GSOPS 寻优 DGSVRM 的参数时优于单独的 GSO,因为 GSOPS 利用模式搜索算法提高了 GSO 的局部搜索能力;最后,基于 GSOPS 的 DGSVRM 负荷预测模型获得了最小的 MAPE、最小的 MASE、最小模型误差值 E_m 和最大的 DS,这充分验证了使用 GSOPS 寻优 DGSVRM 参数及提高负荷预测模型性能取得的优势。

表 2 MAPE 的对比

Table 2 Comparison of MAPE

(单位:%)

月份	GSOPS	GSO	GA	PSO	SA
4 月	0.95	1.11	1.27	1.19	1.35
5 月	0.94	1.23	1.37	1.43	1.57
6 月	1.18	1.23	1.34	1.45	1.67
4 月至 6 月	1.02	1.21	1.36	1.45	1.63

表 3 MASE 的对比

Table 3 Comparison of MASE

月份	GSOPS	GSO	GA	PSO	SA
4 月	0.29	0.35	0.42	0.46	0.49
5 月	0.35	0.42	0.48	0.51	0.48
6 月	0.39	0.44	0.43	0.49	0.57
4 月至 6 月	0.35	0.43	0.45	0.48	0.52

表 4 DS 的对比

Table 4 Comparison of DS

(单位:%)

月份	GSOPS	GSO	GA	PSO	SA
4 月	97.19	95.74	93.13	93.09	92.89
5 月	93.76	92.32	93.59	93.28	92.28
6 月	96.35	93.35	93.45	92.46	91.56
4 月至 6 月	95.74	94.32	93.68	92.99	92.34

表 5 模型误差的对比

Table 5 Comparison of model errors

月份	GSOPS	GSO	GA	PSO	SA
4 月	0.82	0.92	1.1	0.99	1.1
5 月	0.81	0.93	1.03	1.15	1.07
6 月	0.82	0.95	1.09	0.92	1.17
4 月至 6 月	0.83	0.92	1.06	0.99	1.12

5 种常见算法优化 DGSVRM 模型之后求得近似回归超平面与原始回归超平面的误差直方图如图 1 所示。

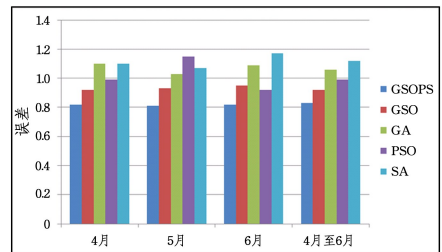


图 1 5 种算法模型的模型误差

Fig. 1 Model errors of the five algorithm models

表 6 列出了 5 种算法在优化模型参数时所使用的时间。从表中可以看出,GSOPS 算法比其他 4 个算法消耗了更多的时间,这是因为 GSOPS 算法加入了模式搜索算法,在个体的精细学习中增加了耗时。从应用在实际电力系统角度看,提前一天进行预测决策对于 GSOPS 多消耗一点时间是能接受的。上述结果表明,使用 GSOPS 寻优 DGSVRM 的参数能力强,可以使用 GSOPS 优化 DGSVRM 模型。

表 6 5 种算法寻优参数所需时间

Table 6 Need time to optimize parameters of five algorithm

算法	CPU/m
GSOPS	33
GSO	24
GA	27
PSO	28
SA	26

为了更充分地检验基于 GSOPS 的 DGSVRM 负荷预测模型的预测能力,选择了常用径向基神经网络(RBFNN)、基于 LM 的 MLP 神经网络(MLP-LM)、多目标模糊时间序列预测模型(TV-ROMP)^[41]这 3 个常用的电力负荷预测模型来进行对比。RBFNN 和 MLP-LM 由 MATLAB 相应工具箱来实现。为了在统一标准做对比,RBFNN,MLP-LM 和 TV-ROMP 均采用一样的数据预处理、输入选择处理及参数做出相应选择。表 7—表 9 列出了不同预测模型在不同月份和整体测试集上的 3 个误差指标结果。由这些表可以看出,不论是单独的月份还是整个测试集,MLP-LM 模型的预测结果最

不理想,这主要是由于 MLP-LM 模型属于非线性模型,在处理非线性特征的负荷数据时,无法找到参数权值最优状态值,因此效果就差。另外,使用 GSOPS 优化 DGSVRM 负荷预测模型时各评价指标都优于其他模型。通过使用 Wilcoxon 符号秩检验对结果做统计检验,表明本文提出的模型在 0.03 的显著水平下显著优于其他对比模型。

表 7 MAPE 的对比

Table 7 Comparison of MAPE

(单位:%)				
月份	DGSVRM	RBF	MLP-LM	TV-ROMP
4月	1.04	2.37	2.35	1.21
5月	1.14	2.38	2.39	1.34
6月	1.18	2.4	2.42	1.45
4月至6月	1.5	2.4	2.42	1.24

表 8 MASE 的对比

Table 8 Comparison of MASE

月份	DGSVRM	RBF	MLP-LM	TV-ROMP
4月	0.31	0.6	0.59	0.52
5月	0.34	0.6	0.63	0.53
6月	0.39	0.64	0.66	0.58
4月至6月	0.34	0.61	0.63	0.55

表 9 DS 的对比

Table 9 Comparison of DS

(单位:%)				
月份	DGSVRM	RBF	MLP-LM	TV-ROMP
4	96.49	90.12	90.34	93.51
5	94.73	89.15	89.01	93.81
6	94.84	89.68	89.51	94.07
4月至6月	94.84	89.68	89.51	93.97

为了比较 GSOPS 与常用混合进化算法搜寻 DGSVRM 模型中参数的效果,选用文献[34]中使用的电力负荷模型在第二个案例上进行预测。为了统一标准进行对比,用 DGSVRM 模型与已有文献进行一样的季节调整策略^[34],并用于预测最后 6 个月的负荷数据值。表 10 列出了电力负荷数据的真实值和各个模型的预测结果。选取文献[42]的 GSO- ϵ -SVM-SA(简记为 GSOSA),使用 SA 对 SVRM 的参数进行优化。CGASA 和 S-CGASA 分别为不使用与使用季节调整策略的 SVRM 预测模型,两个模型均使用混沌遗传算法模拟退火算法优化 SVRM 预测模型中的参数。GSOPS 使用本文提出的基于 GSOPS 的 DGSVRM 预测模型,而 SGSOPS 则在 GSOPS 的基础使用了季节调整策略。如表 10 所列,GSOSP 比 GSO- ϵ -SVM-SA 和 CGASA 有更小的 MAPE 和 MASE,但微差于 S-CGASA,原因在于 S-CGASA 使用了季节调整策略。而基于 GSOPS 的 DGSVRM 模型也使用季节调整方式,与 S-CGASA 相比,获得了更好的 MAPE, MASE。因此,在本案例中,与已有文献中的 SA 和 CGASA 算法相比,本文提出的 GSOPS 算法获得了比 SVRM 负荷预测模型更优的参数。获得优势有两个方面的原因:1)GSOPS 提升了全局搜索能力和局部精细能力;2)通过轮盘赌局的方式,进一步搜索局部精细学习个体,使全局与局部的能力得到了较好的平衡。为了直观地检验本文提出的基于 GSOPS 的 DGSVRM 模型的预测效果,图 2—图 6 给出了 5 种算法的评价指标曲线图,可以明显看出预测曲线非常接近实际曲线,且误差值很小。

表 10 5 种算法优化模型的预测结果与真实值的比较

Table 10 Comparison of actual and predicted values of five algorithms optimized model

Time	真实值	GSOSA	CGASA	S-CGASA	GSOPS	SGSOPS
10月8日	181.07	183.50	177.30	174.64	179.90	178.25
11月8日	180.56	189.34	177.44	184.21	181.55	184.26
12月8日	189.03	202.98	177.58	189.91	190.45	188.97
1月9日	182.07	194.75	177.73	181.97	182.58	181.80
2月9日	167.35	167.58	177.87	163.28	165.45	161.94
3月9日	189.30	184.94	178.01	182.17	187.82	181.92
4月9日	174.84	180.16	178.68	177.63	174.25	176.11
MAPE/%	3.80	3.73	3.73	1.90	1.33	1.58
MASE	0.58	0.55	0.55	0.44	0.22	0.34
DS/%	82.33	33.33	33.33	83.33	87.33	83.33
E_m	0.75	0.90	0.89	0.92	0.68	0.86

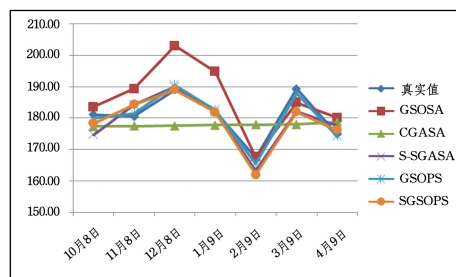


图 2 5 种算法优化模型所得实际值与预测值的比较

Fig. 2 Comparison of the actual and predicted values of five algorithms optimized model

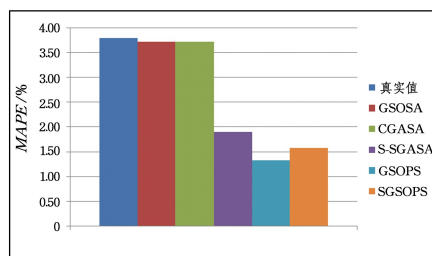


图 3 不同算法优化模型的 MAPE 对比

Fig. 3 MAPE comparison of different algorithms optimized model

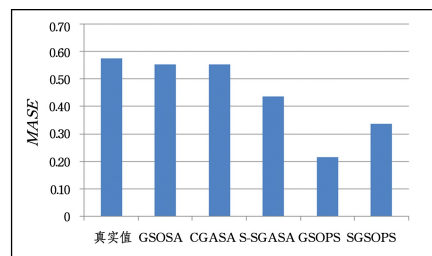


图 4 5 种算法寻优模型的 MASE 对比

Fig. 4 MASE comparison of five algorithms optimized model

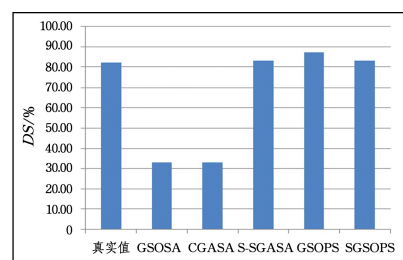


图 5 不同算法优化模型的 DS 对比

Fig. 5 DS comparison of different algorithms optimized model

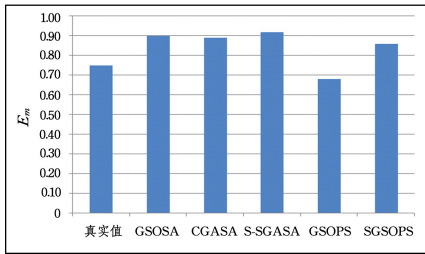


图 6 5 种算法优化模型的 E_m 对比

Fig. 6 E_m comparison of five algorithms optimized model

结束语 为了提高预测电力负荷精度,本文提出了萤火虫群优化算法与模式搜索算法的混合算法 GSOPS,并将其用于寻优 DGSVRM 负荷预测模型的关键参数,获得了最好的预测结果。在混合算法中,GSO 负责搜索全局解空间并搜索最优解潜在区域,PS 算法则被个体用于在局部空间执行进一步搜索学习,进而提高 GSO 算法的局部搜索能力。为了验证模型的预测能力,用两实例电力负荷的预测进行了实验。首先,第一个预测实例中,利用 4 种常用的智能算法(GSO,GA,PSO,SA)优化的模型确立最优模型,再用 3 种常用的电力预测模型(RBFNN,MLP,ARIMA)进行预测比较。从预测结果得出本文提出的 GSOPS 算法强于其他 4 种智能算法,有效地搜寻了构建 DGSVRM 模型的最优参数。从验证中得出,采用 GSOPS 算法优化的 DGSVRM 负荷预测模型优于其他 3 种模型。其次,使用 DGSVRM 模型和已有文献中的预测模型在第二个实例中进行预测以检测各种模型的预测能力,预测结果表明使用 GSOPS 优化 DGSVRM 负荷预测模型参数得到的 DGSVRM 模型比已有文献中提出的预测模型预测更准确。

参考文献

[1] SAAB S, BADR E, NASR G. Univariate modeling and forecasting of energy consumption; the case of electricity in Lebanon [J]. *Energy*, 2001, 26(1): 1-14.

[2] GOIA A, MAY C, FUSAI G. Functional clustering and linear regression for peak load forecasting [J]. *International Journal of Forecasting*, 2010, 26(4): 700-711.

[3] AMARA WICKRAMA H A, HUNT L C. Electricity demand for Sri Lanka; a time series analysis [J]. *Energy*, 2008, 33(5): 724-739.

[4] HIPPERT H S, PEDREIRA C E, SOUZA R C. Neural networks for short-term load forecasting: A review and evaluation [J]. *IEEE Transactions on Power Systems*, 2001, 16(1): 44-55.

[5] MANDAL P, SENJYU T, NAOMITSU U, et al. A neural network based several-hour-ahead electric load forecasting using similar days approach [J]. *International Journal of Electrical Power & Energy Systems*, 2006, 28(6): 367-373.

[6] LAURET P, FOCK E, RANDRIANARIVONY R N, et al. Bayesian neural network approach to short time load forecasting [J]. *Energy Conversion and Management*, 2008, 49(5): 1156-1166.

[7] CRISTIANINI N, TAYLOR J S. 支持向量机导论 [M]. 李国正, 王猛, 曾华军, 译. 北京: 电子工业出版社, 2004.

[8] VAPNIK V N. The nature of statistical learning theory [M]. Beijing: Tsinghua University Press, 2000.

[9] SAPANKEVYCH N L, SANKAR R. Time Series Prediction

Using Support Vector Machines: A Survey [J]. *IEEE Computational Intelligence Magazine*, 2009, 4(2): 24-38.

[10] ELATTAR E E, GOULERMAS J, WU Q H. Electric Load Forecasting Based on Locally Weighted Support Vector Regression [J]. *IEEE Transactions on Systems Man and Cybernetics Part C-Applications and Reviews*, 2010, 40(4): 438-447.

[11] WANG J, ZHU W, ZHANG W, et al. A trend fixed model combined with the E-SVR for short-tenor on firstly and seasonal adjustment forecasting of electricity demand [J]. *Policy*, 2009, 37(11): 4901-4909.

[12] DUAN P, XIE K C, GUO T T, et al. Short-Term Load Forecasting for Electric Power Systems Using the PSO-SVR and FCM Clustering Techniques [J]. *Energies*, 2011, 4(1): 173-184.

[13] BAO Y, HU Z, XIONG T. A PSO and pattern search based memetic algorithm for SVMs parameters optimization [J]. *Neurocomputing*, 2013, 117: 98-106.

[14] CHAPELLE, VAPNIK V, BOUSQUE T, et al. Choosing multiple parameters for support vector machines [J]. *Machine Learning*, 2002, 46(1/2/3): 131-159.

[15] GOMES T A F, PRUDENCIO R B C, SOARES C, et al. Combining meta-learning and search techniques to select parameters for support vector machines [J]. *Neurocomputing*, 2012, 75(1): 3-13.

[16] HONG W C, DONG Y, ZHANG W Y, et al. Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm [J]. *International Journal of Electrical Power & Energy Systems*, 2013, 44(1): 604-614.

[17] HONG W C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm [J]. *Energy*, 2011, 36(9): 5568-5578.

[18] AREIBI S, YANG Z. Effective Memetic Algorithms for VLSI design genetical gorithms local search mufti level clustering [J]. *Evolutionary Computation*, 2004, 12(3): 327-353.

[19] BAO Y, HU Z, XIONG T. A PSO and pattern search based memetic algorithm for SVMs parameters optimization [J]. *Neurocomputing*, 2013, 117: 98-106.

[20] KRISHNANAND K N, GHOSE D. Glowworm swarm based optimization algorithm for multimodal functions with collective robotics applications [J]. *Multiagent and Grid Systems—An International Journal*, 2006, 2: 209-222.

[21] YARINEZHAD R, SARABI A. A New Routing Algorithm for Vehicular Ad-hoc Networks based on Glowworm Swarm Optimization Algorithm [J]. *Journal of AI and Data Mining*, 2019, 7(1): 69-76.

[22] KALAISELVI T, NAGARAJA P, ABDUL B Z. A Comprehensive Study on Glowworm Swarm Optimization [J]. *Computational Methods, Communication Techniques and Informatics*, 2017, 8(1): 332-337.

[23] QIONG, PENG L. A Self-Adaptive Step Glowworm Swarm Optimization Approach [J]. *International Journal of Computational Intelligence and Applications*, 2019, 18(1): 1950004-1-195004-11.

[24] GUO H S, WANG W J. Dynamical Granular Support Vector Regression Machine [J]. *Journal of Software*, 2013, 24(11): 2535-2547.

- [25] VAPNIK V N. Statistical Learning Theory [M]. New York: John Wiley & Sons, 1998.
- [26] BERMEJO P, GAMEZ J A, PUERTA J M. AGRASP algorithm for fast hybrid (filter-wrapper) feature subset selection in high-dimensional datasets. Pattern [J]. Recognition Letters, 2011, 32(5): 701-711.
- [27] VAPNIK V, GOLOWICH SE, SMOLA A. Support vector method for function approximation, regression estimation, and signal processing[M] // Proc. of the Neural Information Processing Systems. Cambridge: MIT Press, 1997.
- [28] COLLOBERT R, BENGIO S. SVM Torch; Support vector machines for large-scale regression problems[J]. Journal of Machine Learning Research, 2001, 1: 143-160.
- [29] TAY FEH, CAO L J. Application of support vector machines in financial time series for forecasting[J]. Omega, 2001, 29(4): 309-317.
- [30] HERNANDEZ L, BALADRON C, AGUIAR J M, et al. A Survey on Electric Power Demand Forecasting; Future Trends in Smart Grids, Microgrids and Smart Buildings[J]. IEEE Communications Surveys & Tutorials, 2014, 16(3): 1460-1495.
- [31] MOMMA M, BENNETT K P. A pattern search method for model selection of support vector regression[C] // Proceedings of the Second Siam International Conference on Data Mining. 2002: 261-274.
- [32] DOLAN E D J, LEWIS R M, TORCZON V. On the local convergence of pattern search[J]. SIAM Journal of Optimization, 2003, 14(2): 567-583.
- [33] KENEDY J R, EBERHART T. Optimization[C] // Proceedings Particle swarm of IEEE International Conference on Neural Networks. 1995: 1942-948.
- [34] ZHANG W Y, HONG W C, DONG Y C, et al. Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting[J]. Energy, 2012, 45(1): 850-858.
- [35] HU Z Y. Research on Parameter Optimization of Support Vector Machine Based on SVM[D]. Wuhan: Huazhong University of Science and Technology, 2011.
- [36] HSU C W, CHANG C C, LIN C J. A practical guide to support vector classification, Department of Computer Science[D]. National Taiwan University, 2008.
- [37] HYNDMAN R J, KOEHLER A B. Another look at measures of forecast accuracy[J]. International Journal of Forecasting, 2006, 22(4): 679-688.
- [38] ARMSTRONG J S, COLLOPY F. Error measures for generalizing about forecasting methods: empirical comparisons[J]. International Journal of Forecasting, 1992, 8(1): 69-80.
- [39] DIEBOLD F X, MARIANO R. Comparing predictive accuracy [J]. Journal of Business and Economic Statistics, 1995, 13: 253-263.
- [40] CONOVER W. Practical nonparametric statistics(2nd ed)[M]. New York: Wiley & Sons, 1980.
- [41] LIU X J, FANG J A. Long Term Load Forecasting Based on a Time-Variant Ratio Multiobjective Optimization Fuzzy Time Series Model[J]. Mathematical Problems in Engineering, 2013, 2013: 1-7.
- [42] WANG J, ZHU W, ZHANG W, et al. A trend fixed model combined with the E-SVR for short-tenor on firstly and seasonal adjustment forecasting of electricity demand [J]. Policy, 2009, 37(11): 4901-4909.



TANG Cheng-e, born in 1983, lecturer. Her main research interests include neural networks and automation of electric power systems.

(上接第 44 页)

- [15] CHO K, VAN M B, GULCEHRE C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation [J]. arXiv:1406.1078, 2014.
- [16] XU Y, MOU L, LI G, et al. Classifying relations via long short term memory networks along shortest dependency path [J]. arXiv:1508.03720, 2015.
- [17] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient Estimation of Word Representations in Vector Space [J]. arXiv:1301.3781, 2013.
- [18] CHE W, LI Z, LIU T. LTP: A Chinese Language Technology Platform [C] // Proceedings of the 23rd International Conference on Computational Linguistics Demonstrations Volume. Stroudsburg, PA: ACL, 2010: 13-16.
- [19] KINGMA D P, BA J. Adam: A method for stochastic optimization [J]. arXiv:1412.6980, 2014.



LYU Yi-lin, born in 1997, undergraduate student. His main research interest is information extraction.



WAN Huai-yu, born in 1981, Ph.D, associate professor, Ph.D supervisor, is a member of China Computer Federation. His main research interests include social network mining, text mining, user behavior analysis, and traffic data mining.