

# Short-Term Power Load Forecasting Based on Hybrid ReliefF-PMI-IGWO

Chen Zichun<sup>1,2</sup>, Li Zhen<sup>1</sup>, Fu Hua<sup>1</sup>

<sup>1</sup>School of Electrical and Control Engineering, Liaoning Technical University, Huludao, China

<sup>2</sup>Zhongluan Science and Technology Co., LTD., Tangshan, China

## Email address:

1070547154@qq.com (Chen Zichun), 1937582375@qq.com (Li Zhen), fuhua@163.com (Fu Hua)

## To cite this article:

Chen Zichun, Li Zhen, Fu Hua. Short-Term Power Load Forecasting Based on Hybrid ReliefF-PMI-IGWO. *Science Discovery*. Vol. 10, No. 6, 2022, pp. 414-421. doi: 10.11648/j.sd.20221006.16

**Received:** October 14, 2022; **Accepted:** November 11, 2022; **Published:** November 29, 2022

**Abstract:** With the rapid growth of power grid in China, it is becoming more and more important to forecast short-term power load. Aiming at the problems of complex causes of short-term load forecasting, large amount of data, low efficiency and poor accuracy of traditional model forecasting methods, and the single model forecasting can no longer adapt to the current power load changes, a short-term power load forecasting method based on hybrid ReliefF-PMI-IGWO was proposed. Firstly, two Filter feature selection methods based on hybrid ReliefF and Partial Mutual Information (PMI) were used to screen out irrelevant and redundant input features, and a smaller feature subset was obtained. Then the Wrapper method is used for further screening, among which the Wrapper method selects Improved Grey Wolf Optimization (IGWO) algorithm, and its accuracy is significantly improved. Combined with the case analysis of the hourly load data of China Midea HVAC and Nongfu Spring from May to August 2017, and compared with other five traditional models. The simulation results show that the MAE and MAPE obtained by the proposed model are 49.54, 1.62 and 71.34, 1.98, respectively, which are significantly better than other models, and the prediction accuracy of the proposed model is the highest and the closest to the real value.

**Keywords:** Short Term Power Load Forecasting, Partial Mutual Information, ReliefF, Improved Gray Wolf Optimization Algorithm, Feature Selection

## 基于混合ReliefF-PMI-IGWO的短期电力负荷预测

陈子春<sup>1,2</sup>, 李珍<sup>1</sup>, 付华<sup>1</sup>

<sup>1</sup>辽宁工程技术大学电气与控制工程学院, 葫芦岛, 中国

<sup>2</sup>中滦科技股份有限公司, 唐山, 中国

## 邮箱

1070547154@qq.com (陈子春), 1937582375@qq.com (李珍), fuhua@163.com (付华)

**摘要:** 随着我国电网在快速成长, 对短期电力负荷预测也越来越重要。针对短期负荷预测成因复杂, 数据量大, 以往传统模型预测方法效率低、准确率差, 且单一模型预测已经不能适应现在的电力负荷变化等问题, 本文提出一种基于混合ReliefF-PMI-IGWO的短期电力负荷预测方法。首先使用基于混合ReliefF与偏互信息 (Partial Mutual Information, PMI) 的两种Filter特征选择把无关和冗余的输入特征筛选掉, 得出较小的特征子集; 然后使用Wrapper方法进一步筛选, 其中Wrapper方法选择改进灰狼优化算法 (Improved Grey Wolf Optimization, IGWO), 其准确率明显提升。结合2017年5月-8月中国美的暖通和农夫山泉两个公司的小时级负荷数据的实例分析, 与其他5种传统模型进行对比。仿真结果表明, 所提出的模型得到的两个评价指标MAE和MAPE分别为49.54、1.62和71.34、1.98, 明显优于其他模型, 验证了该模型预测精度最高, 最接近真实值。

**关键字:** 短期电力负荷预测, 偏互信息, ReliefF, 改进灰狼优化算法, 特征选择

## 1. 引言

随着电力行业智能化的快速发展, 提高电力系统智能调节能力, 并且合理安排跨省跨区输电等问题, 科学准确的电力负荷预测才是实现这些决策的依据和保障。电力负荷的特点是非线性和非平稳性, 在复杂多样的社会因素下极易受影响, 导致预测短期负荷的精准度变得艰巨, 所以选择准确、合适的特征对于短期电力负荷预测尤为关键。

早期常采用统计学方法进行短期负荷预测, 但统计学模型对时间序列的平稳性要求较高, 导致预测结果误差大、精度低。因此, 近年来, 常常采用机器学习方法或组合模型预测法去预测短期电力负荷。范海虹先使用自展法生成多个学习器, 然后用卡帕 (Kappa) 测度进行初步筛选, 最后使用萤火虫算法进行集成[1]。王淞瑶等人在GRU模型中加入量子信息处理机制, 对参数进行优选时使用改进粒子群 (PSO) 算法进而完成负荷预测[2]。商立群等人先利用GSO-PCV相结合的特征选择去优先得到最优的特征集, 再利用IGWO算法对ELM的权值和阈值进行优化获得权值和阈值的全局最优解[3]。杨海柱等人先利用FE算法对经过CEEMD处理后的各子序列进行熵值重组, 再用AOA-LSSVM模型对处理后的子序列进行预测, 并将预测叠加输出, 最后通过误差函数对模型进行横、纵向对比[4]。段雪滢等人先利用VMD分解技术获得相对平稳的子序列, 通过提出的 (IPSO) 算法进行参数学习, 同时使用自归一化的GRU结构解决模型训练中梯度消失和爆炸的问题[5]。张淑清等人先建立FPA-VMD-BiLSTM混合预测模型, 完成候选集初步分解预测, 再建立基于BiLSTM神经网络的误差纠正模型[6]。赵婧宇等人提出利用麻雀搜索算法优化长短期记忆神经网络参数的模型 (SSA-LSTM)[7]。Yu Zhuowei等人提出将电力消耗和疫情相关的异构特征编码为负荷图, 再构造残差图卷积网络 (ResGCN) 来拟合负荷图到未来负荷的放线性映射[8]。Prakash N等人提出采用模糊推理系统的蜜蜂优化欧氏范数进行问题识别, 并利用支持向量分类器建立STLF模型[9]。Ghulam Hafeez等人提出将特征工程和改进的萤火虫优化算法与SVM相结合的方法 (FE-SVR-mFFO) [10]。上述文献大部分都是两种方法组合在一起去预测电力负荷, 其中商立群等人、张淑清等人和Ghulam Hafeez等人提出的是用三种方法组合在一起, 但这三个方法都过于繁琐, 以上方法都有一个共同的缺点就是虽然预测模型的精度较高, 但是与真实值还是有较大的差距。

特征选择对提高检测的精度和速度至关重要。特征选择可以分为过滤 (Filter)、包装 (Wrapper) 和嵌入 (Embedded) 三类。其中Filter方法可以快速筛除无关冗余的特征, 具有很高的计算效率, 但是分类效果很差。而Wrapper方法是与机器学习分类器结合, 直接根据预测精度来评价所选特征子集, 其精确度比Filter方法更能获得高标准率的特征子集, 但是每次筛选某一特征子集时都需要训练预测模型, 如果模型复杂度急剧增加, 会导致运行效

率低下。为了提高短期负荷预测的速度和精准度, 本文首先使用基于PMI与ReliefF的Filter方法, 快速砍掉特征候选集中的无关和冗余特征进行预剪枝, 然后对预剪枝后缩小的特征选集使用基于IGWO的Wrapper方法再次修剪。为了验证本文提出的混合ReliefF-PMI-IGWO的特征选择算法和基于此的SVM短期负荷预测模型的性能, 仿真实验通过与其他5种特征选择算法模型进行对比, 再用两个企业案例进行模型验证。

## 2. 混合PMI-FA方法的基本原理

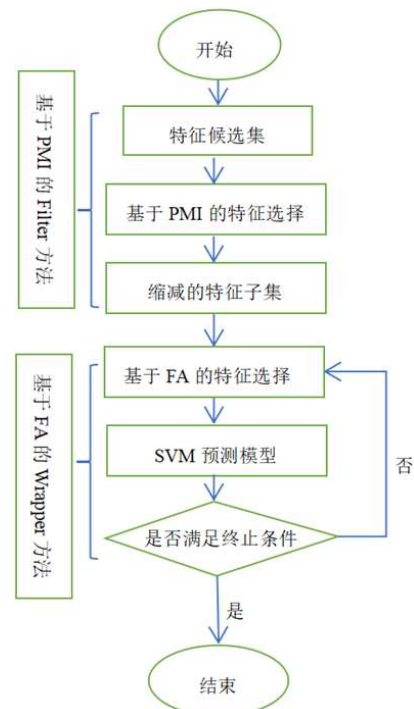


图1 混合PMI-FA方法的流程图。

混合PMI-FA方法的流程图如图1所示。该方法采用的是基于PMI的Filter方法, 通过在已选特征的条件输入和输出变量之间的互信息 (mutual information, MI) 相互衡量, 预先筛选负荷预测模型的输入特征, 去掉无关和冗余的输入特征, 从而缩小特征候选集的范围[11]。再对缩减后的特征候选集使用基于FA的Wrapper方法, 其中预测模型选用支持向量机 (Support Vector Machine, SVM), 输入特征的评价标准选择平均绝对百分比误差 (Mean absolute percentage error, MAPE), 由于Wrapper方法的计算量非常大, 所以选择萤火虫搜索算法 (Firefly algorithm, FA) 可以比较快速的搜索出最优特征子集, 这样可以把Filter的高效率和Wrapper的精度高优势综合在一起。然后通过两个实验North-American电力系统在1991年1月的预测模型结果和使用GEFCom2012竞赛中的20个区域的负荷序列, 去进一步验证该方法的性能, 与其他预测模型相比, 该方法性能比较优异。

### 3. 本文方法

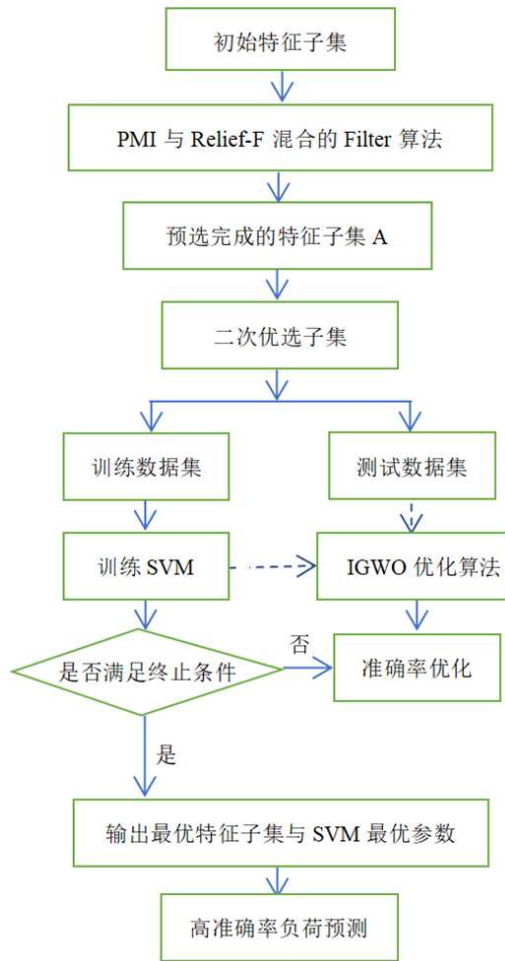


图2 混合ReliefF-PMI-IGWO的短期电力负荷预测的流程图。

基于混合ReliefF-PMI-IGWO的短期电力负荷预测方法流程图如图2所示。为了提高短期负荷预测的精准度，针对其中存在的无关和冗余特征不利于智能预测的准确性的问题，本文提出了一种基于混合ReliefF-PMI-IGWO的短期电力负荷预测方法。首先利用ReliefF和PMI这两种

Filter的混合方法，快速细致的对特征候选集中的无关和冗余特征进行预剪枝，为后续提高搜索算法的运行效率打下基础；然后使用基于IGWO的Wrapper方法对最优特征子集进行全局搜索二次修剪。实验部分使用了两家国内企业的数据集进行预测，结果与其他5种常见的特征选择算法进行对比，去验证本文提出的基于混合ReliefF-PMI-IGWO特征选择算法和基于此的SVM短期负荷预测模型的性能。具体步骤如下：

(1) 在相同情况下，分别采集两组短期电力负荷数据，一组作为训练样本，一组作为测试样本。

(2) 计算出初始特征集中各个特征的综合权重，利用1节中所提的混合Filter方法，并进行降序排列，设置一定的阈值，如PMI<sub>97%</sub>，剔除小于阈值的，构成特征子集A。

(3) 采用基于IGWO的Wrapper方法对预测特征子集A进行二次特征选择，灰狼位置利用IGWO算法的改进机制进行更新。

(4) 在进行二次预测特征优选时，利用IGWO算法同步进行SVM参数的优化，当算法达到最大迭代次数时，输出最优的特征子集与SVM的最佳参数组合。

(5) 利用最佳参数组合与最优特征子集对SVM进行训练与测试，实现高准确率的短期电力负荷预测。

#### 3.1. 混合ReliefF-PMI的Filter方法

##### 3.1.1. ReliefF算法

ReliefF算法是Kononenko在Relief算法的基础上发展而成的一种特征排序方法[12]。Relief算法能处理二分类特征选择问题，其扩展变体ReliefF能处理多分类特征。

ReliefF算法其运行效率高，数据类型也没有限制。如果想综合评估特征的可分性可以通过考察特征在不同近邻样本之间的差异，在同类样本间特征差异小，但在异类样本间特征差异大，则该特征的可分性就越强。

在使用该算法时，对需要评估的某一特征 $x_i$ ，首先随机抽取一个样本点R在训练集中，再从同类和不同类数据集中各自寻找出k个近邻样本，分别记住H与M，最后不断的更新 $x_i$ 的权值根据式(1)，直到达到最大迭代次数m后，得出各个特征的权值向量W。

$$W(x_i) = W(x_i) - \sum_{j=1}^k \text{diff}(x_i, R, H_j) / (mk) + \sum_{c \neq \text{class}(R)} \left[ \frac{P(c)}{1 - P(\text{class}(R))} \sum_{j=1}^k \text{diff}(x_i, R, M_j(c)) \right] / (mk) \quad (1)$$

$$\text{diff}(x_i, R, H_i) = \frac{|R_i - H_i|}{\max(x_i) - \min(x_i)} \quad (2)$$

##### 3.1.2. 偏互信息(PMI)

在变量间的线性与非线性的关联耦合程度上互信息可以较好的衡量，能有效辨别模型的单变量输入。使用MI在特征选择时常常会导致输入变量中冗余性含量很高。因为它辨别不出来。其表达式如下所示：

$$MI = \iint \mu^{X,Y}(x,y) \ln \frac{\mu^{X,Y}(x,y)}{\mu^X(x)\mu^Y(y)} dx dy \quad (3)$$

其中， $\mu^X$ 、 $\mu^Y$ 和 $\mu^{X,Y}$ 分别代表了X、Y的单变量概率密度函数和联合概率密度函数。E Hancer等人提出的差分进

化模糊互信息也会导致输入变量中冗余性含量很高[13]。而胡忠义提出偏互信息则是利用条件期望先筛除掉变量间的相关关系再计算MI，即解决了输入变量间冗余性高的问题，又考虑了输入和输出变量间的信息耦合程度，变量选择的精确度得到了有效的提高[14]。但其文中显性度量值选择的是PMI<sub>95%</sub>，因为候选集过于庞大，因而选择显性度量值PMI<sub>97%</sub>，筛选更加快速精准。在已选择的特征子集Z的条件下输入特征X与输出变量Y之间的PMI估计如下：

$$PMI(X,Y) = \iint \mu^{X',Y'}(x',y') \ln \frac{\mu^{X',Y'}(x',y')}{\mu^{X'}(x')\mu^{Y'}(y')} dx' dy' \quad (4)$$



$$x' = x - E[x|z] \quad y' = y - E[y|z] \quad (5)$$

而 $E[\cdot]$ 代表期望值。 $\mu^{x',y'}$ 和 $\mu^{x'}$ 、 $\mu^{y'}$ 分别代表为样本的联合概率密度估计函数和单变量概率密度估计。变量 $x'$ 和 $y'$ 是 $x$ 和 $y$ 考虑已有的特征子集 $z$ 条件下的残余信息。使用基于PMI的特征选择时，输入特征优先选择拥有最高PMI值的变量。

基于PMI的特征选择方法的简要步骤如下：

(1) 设筛选出的输入特征集合为 $Z$ （在算法初级阶段 $Z$ 为空集），而还没有被选入的输入特征子集为 $Z_{\text{not}}$ 。

(2) 在已选择的 $Z$ 条件下估计 $Z_{\text{not}}$ 与输出变量 $Y$ 中每个变量的PMI值。

(3) 在 $Z_{\text{not}}$ 中识别出 $z_{\text{max}}$ （拥有最大PMI值的变量）。

(4) 想要一个满足 $Y$ 与 $z_{\text{max}}$ 相互独立的原假设的新样本，可以对变量 $z_{\text{max}}$ 中的所有样本使用Bootstrap方法进行抽样，然后再计算新样本的PMI值。对以上计算PMI值的过程进行重复并产生一个PMI数据分布，计算 $z_{\text{max}}$ 的显著性度量则是PMI分布的97%分位值 $\text{PMI}_{97\%}$ 。它可以用来衡量 $z_{\text{max}}$ 是否对预测模型有显著性影响[15]。

(5) 如果上步计算的显著性度量 $\text{PMI}_{97\%}$ 小于变量 $z_{\text{max}}$ 的PMI值时，则表示在预测模型时 $z_{\text{max}}$ 有显著性影响，可以把 $z_{\text{max}}$ 从 $Z_{\text{not}}$ 中剔除并添加在已选的特征子集 $Z$ 中。如果变量 $z_{\text{max}}$ 的PMI值小于 $\text{PMI}_{97\%}$ 或 $z_{\text{max}}$ 中没有变量了，则转到第七步。

(6) 返回并继续执行第二步。

(7) 选完所有显著性的变量后停止执行，结束步骤。

### 3.1.3. 混合ReliefF-PMI方法

徐宇颂等人提出的将特征向量按特点分为2类，分别使用方法确认最优特征向量维度，但这个过程太过繁琐，且耗时很长[16]。胡欣等人提出的基于特征选择的LSTM-Attention预测模型能有效去除冗余因素，但是不能去除无关因素[17]。周传华等人提出的从信息相关性和分类能力对特征进行综合度量，再采用SFS策略对特征进行选择，去除了无关特征，但无法去除冗余特征[18]。杨国华等人提出的基于Holt-Winters指数平滑的特征组合（FCHW），这个方法只能适合少的特征集，数据一旦变多，处理时间就会变长[19]。Yu Binbin等人提出通过两级生命

周期分析消除了冗余和无关特征，但分类时间过长，导致消除特征时间过长[20]。Zhengyi Liu等人提出了交互式滤波包装多目标进化算法GR-MOEA，用来引导和修复策略来选择高质量的特征子集，但引导和修复时间长，不如筛选特征快[21]。Wang Jianzhou等人用模糊造粒技术简化了大规模数据集，提取了原始数据的有效特征，但有效特征存在冗余[22]。针对以上无法去除冗余或无关因素，以及特征选择时间长，且单一的Filter评价准则普遍无法准确的去评价各个特征对诊断精度综合影响的问题，提出了ReliefF与PMI进行混合Filter特征选择，去除了无关和冗余因素，还提升了筛选重要特征的高速性和准确性，具体步骤如下：

(1) 求出初始特征集合中各个特征的ReliefF权重值和PMI值。

(2) 做标准化处理对于所求出的两组权重值，使其均值为0，标准差为1。

(3) 经过第二步处理后，设第 $x$ 个特征的ReliefF权重为 $W_x$ ，PMI值为 $P_x$ ，则该特征的综合权重为 $CW_x = W_x - P_x$ ， $CW_x$ 越大，则该特征越重要。

### 3.2. 基于IGWO的Wrapper方法

由于FA算法很容易因参数设置不当而导致算法无法收敛，还会导致算法收敛过早的问题，曾林俊等人提出的基于GWO-ALO优化ELM的预测方法，两种算法结合导致运算过程复杂，预测时间加长的的问题[23]。因此本文使用了IGWO算法，该算法有较强的收敛性能，而且结构比较简单、参数调节少、运算快，在实验中容易实现，可以自适应调整的收敛因子以及信息反馈机制，还可以实现局部与全局搜索之间的平衡。

#### 3.2.1. GWO算法

灰狼优化算法（Gray Wolf Optimization, GWO）是受到灰狼群体社会等级制度与狩猎行为启发而提出的一种智能优化算法，由澳大利亚格里菲斯大学学者Mirjalili等人在2014年提出的[24]。灰狼的社会等级如图3所示，在群体狩猎过程中发挥着重要的作用，捕食的过程在灰狼 $\alpha$ 的带领下完成。灰狼的狩猎和算法执行包括3个主要部分：包围、追捕和攻击。

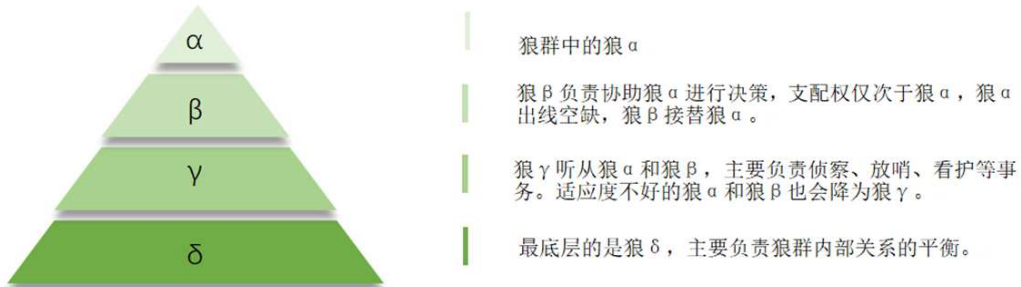


图3 灰狼的社会等级制度。

#### (1) 包围猎物阶段

在狩猎过程中，灰狼群对猎物进行包围，这一行为的数学模型用式(6)和(7)来建立：

$$\vec{D} = |\vec{C} \cdot \vec{X_p}(t) - \vec{X}(t)| \quad (6)$$

$$\vec{X}(t+1) = \vec{X_p}(t) - \vec{A} \cdot \vec{D} \quad (7)$$

式(6)代表灰狼与猎物间的距离向量；式(7)代表灰狼的位置向量更新公式； $t$ 代表算法当前的迭代次数； $\vec{X}$ 和 $\vec{X}_p$ 分别代表一头灰狼和猎物的位置向量， $\vec{A}$ 与 $\vec{C}$ 均为系数向量，公式如下：

$$\vec{A} = 2 \cdot a \cdot \vec{r}_1 - a \quad (8)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (9)$$

其中 $\vec{r}_1$ 与 $\vec{r}_2$ 是[0,1]区间的随机向量； $a$ 是收敛因子，随着迭代次数从2线性减小到0。

### (2) 追捕猎物阶段

种群中其他灰狼的位置由 $\alpha$ 、 $\beta$ 和 $\gamma$ 的位置共同决定：

$$\begin{cases} \vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \\ \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \\ \vec{D}_\gamma = |\vec{C}_3 \cdot \vec{X}_\gamma - \vec{X}|, \vec{X}_3 = \vec{X}_\gamma - \vec{A}_3 \cdot \vec{D}_\gamma \end{cases} \quad (10)$$

当识别出猎物的位置时，灰狼群的位置则根据公式(11)进行更新：

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (11)$$

式中， $\vec{X}(t+1)$ 表示灰狼 $\vec{X}$ 在算法迭代到 $t+1$ 次时的位置。

### (3) 攻击猎物阶段

当猎物停止移动时，灰狼群体攻击猎物以完成最终的狩猎。为了模拟逼近猎物的数学模型，要通过逐渐减少收敛因子 $a$ 来实现，由公式(8)可知，当 $a$ 的减小， $\vec{A}$ 的波动范围也随之减小，表示如(12)下：

$$a = 2 - 2 \left( \frac{t}{t_{\max}} \right) \quad (12)$$

其中， $t_{\max}$ 代表最大迭代次数，在迭代过程中，当 $a$ 的值从2线性下降到0时，与之对应的 $\vec{A}$ 的值在区间 $[a, a]$ 内变化。

## 3.2.2. 改进灰狼优化算法(IGWO)

### (1) 基于余弦规律的收敛因子

当 $|\vec{A}| > 1$ 时，灰狼群体是全局搜索，搜索范围广且收敛速度快；当 $|\vec{A}| < 1$ 时，灰狼群体是局部搜索，缩小搜索范围对猎物进行攻击，其收敛速度慢。因此，GWO算法的全局和局部搜索能力与 $\vec{A}$ 的大小至有至关重要的关系。因为 $a$ 是线性迭代的，但算法在收敛过程中并不是线性的，所以，在实际的优化搜索过程中线性递减的收敛因子 $a$ 不能准确反映出其复杂性。因此，这里提出一种基于余弦规律变化的收敛因子，改进公式如(13)所示。

$$\begin{cases} a = a_{\text{final}} + (a_{\text{initial}} - a_{\text{final}}) \frac{1 + [\cos((t-1)\pi/(t_{\max}-1))]^n}{2}, & t \leq \frac{1}{2} t_{\max} \\ a = a_{\text{final}} + (a_{\text{initial}} - a_{\text{final}}) \frac{1 + [\cos((t-1)\pi/(t_{\max}-1))]^n}{2}, & \frac{1}{2} t_{\max} \leq t \leq t_{\max} \end{cases} \quad (13)$$

其中， $a_{\text{initial}}$ 和 $a_{\text{final}}$ 分别为收敛因子 $a$ 的初始值和最终值，本文取 $a_{\text{initial}}=0$ ， $a_{\text{final}}=2$ ， $n$ 为递减指数， $0 < n < 1$ 。

在迭代初期改进后的收敛因子 $a$ 可以长时间保持较大值，使 $A$ 也保持较大值的时间长些，提高搜索效率，因此减小的缓慢；而到了迭代后期，为了使 $a$ 长时间保持较小值， $A$ 也长时间保持较小值，提高搜索精度，看上去是一条基于余弦规律变化的曲线，这样能使算法很好的平衡全局和局部搜索能力。

### (2) 引入动态权重策略

常见的比例权重有很多，本文采用了一种基于步长欧式距离的比例权重[25]，其表达式如式(15)所示：

$$W_1 = \frac{|X_1|}{|X_1|+|X_2|+|X_3|}, W_2 = \frac{|X_2|}{|X_1|+|X_2|+|X_3|}, W_3 = \frac{|X_3|}{|X_1|+|X_2|+|X_3|} \quad (14)$$

$$X(t+1) = \frac{X_1 \cdot W_1 + X_2 \cdot W_2 + X_3 \cdot W_3}{3} \quad (15)$$

其中， $W_1$ 、 $W_2$ 、 $W_3$ 分别表示 $\delta$ 狼对 $\alpha$ 、 $\beta$ 、 $\gamma$ 狼的学习率。

## 3.2.3. IGWO算法步骤

综合以上描述的改进策略，给出本文所提出的改进灰狼优化算法(IGWO)的具体步骤：

(1) 设置种群规模 $N$ ，随机生成 $a$ 、 $A$ 、 $C$ 等参数和最大迭代次数 $t_{\max}$ 。

(2) 随机初始化灰狼种群。

(3) 计算灰狼种群中所有个体的适应度值，并按照适应度值进行排序，选择最好的前三个灰狼，并记录其位置 $X_\alpha$ 、 $X_\beta$ 和 $X_\gamma$ 。

(4) 利用公式(10)、(14)和(15)更新种群中其他灰狼的位置。

(5) 利用公式(13)计算 $a$ ，然后利用公式(8)和(9)更新 $A$ 、 $C$ 的值。

(6) 判断算法是否满足结束条件，若达到预定的最大迭代次数 $t_{\max}$ ，则停止计算，并输出最优位置 $X_\alpha$ ；否则，则重复执行第(3)-(5)，直至达到满足条件结束。

## 3.2.4. IGWO算法性能测试

为了测试IGWO算法的性能，本文采取了国际上通用的4个经典基准测试函数进行仿真实验，其中 $f_1$ - $f_2$ 为单峰测试函数， $f_3$ - $f_4$ 为多峰测试函数，并与GA算法、PSO算法和基本GWO算法进行对比。测试函数表达式如(16)所示：

$$\begin{cases} f_1(x) = \sum_{i=1}^n x_i^2 \\ f_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n |x_i| \\ f_3(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \\ f_4(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \end{cases} \quad (16)$$

各测试函数的取值范围和最小值等信息如表1所示，其他算法设置如表2所示。

表1 测试函数信息。

| 函数    | 类型             | 维度 | 取值范围         | 最小值 |
|-------|----------------|----|--------------|-----|
| $f_1$ | Sphere         | 30 | [-100,100]   | 0   |
| $f_2$ | Schweffel 2.22 | 30 | [-10,10]     | 0   |
| $f_3$ | Rastrigin      | 30 | [-5.12,5.12] | 0   |
| $f_4$ | Griewank       | 30 | [-50,50]     | 0   |

表2 其他算法的参数设置。

| 算法      | 参数设置                              |
|---------|-----------------------------------|
| PSO     | 粒子速度 $V_{\max}=5$ , $c_1=c_2=0.5$ |
| GWO     | $M_{\max}=500$                    |
| GWO-ALO | $S_{N_{\max}}=500$                |

表3 4种进化算法对4个测试函数的结果比较。

| 函数    | PSO      |         | GWO    |        | GWO-ALO  |          | IGWO      |     |
|-------|----------|---------|--------|--------|----------|----------|-----------|-----|
|       | 平均值      | 标准值     | 平均值    | 标准值    | 平均值      | 标准值      | 平均值       | 标准值 |
| $f_1$ | 8.20E-14 | 5.90E-4 | 0.0399 | 0.2665 | 1.67E-28 | 9.12E-28 | 0         | 0   |
| $f_2$ | 1.22E+1  | 6.69E+0 | 0.1440 | 0.3163 | 4.67E-17 | 2.56E-16 | 1.51E-198 | 0   |
| $f_3$ | 7.23E+1  | 3.91E+1 | 3.0492 | 2.5106 | 0.3087   | 1.6906   | 0         | 0   |
| $f_4$ | 1.37E+1  | 1.06E+1 | 0.0053 | 0.0230 | 0        | 0        | 0         | 0   |

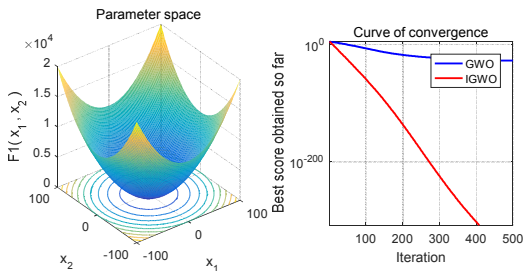


图4 GWO和IGWO的收敛曲线图。

在种群规模和迭代次数相同的情况下，可以从表3的比较结果看出，本文所提出的IGWO算法在4个测试函数上都表现出很明显的优越性，大大提高了收敛精度，IGWO算法收敛到了理论最优值0。从图4的GWO算法与IGWO算法收敛曲线图可以看出，IGWO算法的收敛速度更快，收敛精度更高。

4. 实验结果与分析

本文采用美的暖通和农夫山泉提供的从2017年5月1日到8月15日的电力负荷及气象数据的实例数据集，采样周期为1小时，数据集共有5376条负荷数据。用这两组数据之前，首先要对数据进行预处理，就是把各输入变量分别被归一化至[0, 1]，这样可以避免因变量间的数量级不同而产生误差。

本文的所有实验均使用Matlab R2016a实现，实验平台为Windows10、2.10GHz、16.0GB的计算机。为了验证本文提出的混合Relief-PMI-IGWO的特征选择算法和基于此的支持向量机（SVM）预测模型，选用SVM及作为预测模型，平均绝对误差（mean absolute error, MAE）和平均绝对百分比误差（Mean absolute percentage error, MAPE）为输入特征的评价指标，定义如下：

$$MAE = \frac{1}{N} \sum_{i=1}^N |\tilde{y}_i - y_i| \tag{20}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|\tilde{y}_i - y_i|}{y_i} \times 100\% \tag{21}$$

式中，N为预测步长，这里N为24（小时）； $y_i$ 为待预测样本的真实值； $\tilde{y}_i$ 为模型预测值。其MAPE越小，预测值越接近于真实值。

表4列出了不同特征选择算法在该数据集上的预测模型中的结果。其表中Nu表示筛选出的特征子集及个数、

将算法的种群规模N设置为30，最大迭代次数参数 $T_{\max}$ 为500。实验分别为4个测试函数独立运行30次，避免了单次运行结果所带来的偏差，求得实验的平均值用来反映算法在一定迭代次数下所能达到的寻优精度，而标准差用来反映算法的稳定性。测试函数结果比较见表3。

Tfs表示特征选择所消耗的时间，Ttf表示预测模型的训练和预测时间。

表4 不同特征选择方法的结果分析。

| 方法             | Nu | Tfs(min) | Ttf(min) | MAPE |
|----------------|----|----------|----------|------|
| LSTM-Attention | 19 | 4.13     | 66.52    | 5.43 |
| FCHW           | 25 | 3.49     | 65.47    | 6.17 |
| GR-MOEA        | 7  | 24.36    | 29.36    | 9.63 |
| GWO-ALO-ELM    | 5  | 35.77    | 29.23    | 4.38 |
| PMI-FA         | 10 | 50.81    | 31.98    | 4.96 |
| 本文             | 4  | 23.03    | 24.12    | 4.03 |

由表4可知，本文提出的特征选择方法优于其他5类，能够较好地筛选出与预测时刻相关度高的特征，但本文方法与其中2个方法相比消耗时间稍长，由于负荷决策是电力部门在前一天进行的，所以结果还是可以接受。

为了验证本文模型的优越性，本文在相同最优特征子集条件下，选择不同模型进行对比实验。将2017年5月1日至5月14日两周的负荷数据作为训练集，共672个样本。将2017年5月15至5月17日3天的数据作为预测集，共有144个样本。为了保证所有结果的可靠性和有效性，算法的最大迭代均取200，算法的种群数量均取30。还为了避免单次运行结果容易产生误差，对不同预测模型分别单独运行4次，对其负荷预测模型结果求和并取均值，从而得到6种不同预测模型的预测结果以及评价指标。图5和图6为案例一美的暖通和案例二农夫山泉在不同模型的预测结果对比，表5和表6分别是在不同模型下的预测精度对比。

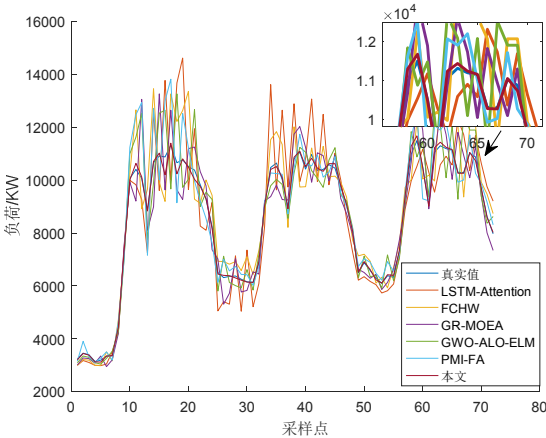


图5 案例一美的暖通不同模型的预测结果对比。

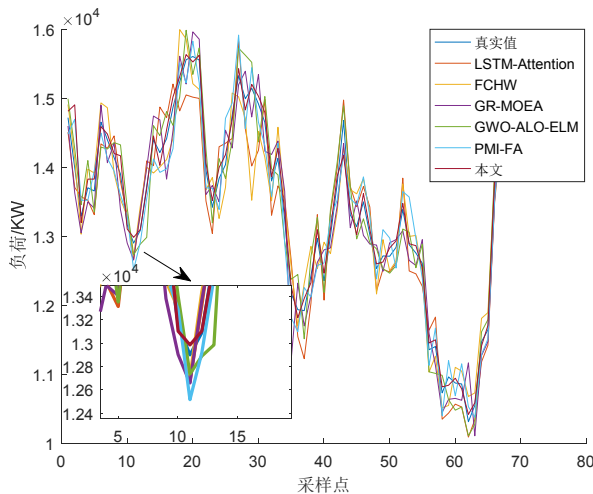


图6 案例二农夫山泉不同模型的预测结果对比。

表5 案例一不同负荷预测模型精度对比。

| 预测模型           | MAE/kw | MAPE/kw |
|----------------|--------|---------|
| LSTM-Attention | 311.34 | 4.99    |
| FCHW           | 235.76 | 4.47    |
| GR-MOEA        | 174.36 | 3.84    |
| GWO-ALO-ELM    | 147.33 | 3.03    |
| PMI-FA         | 104.85 | 2.57    |
| 本文方法           | 49.54  | 1.62    |

表6 案例二不同负荷预测模型精度对比。

| 预测模型           | MAE/kw | MAPE/kw |
|----------------|--------|---------|
| LSTM-Attention | 427.05 | 5.76    |
| FCHW           | 364.66 | 4.49    |
| GR-MOEA        | 231.90 | 4.01    |
| GWO-ALO-ELM    | 176.25 | 3.43    |
| PMI-FA         | 108.55 | 2.58    |
| 本文方法           | 71.34  | 1.98    |

由图5、6的局部放大图显示，本文预测模型所得结果与真实值更加贴合。从平均绝对误差分析，本文所提的预测模型与其它5种相比，在案例一中，MAE分别是261.8kw、186.22kw、124.82kw、97.79kw和55.31kw；在案例二中MAE分别是355.71kw、293.32kw、160.56kw、104.91kw和37.21kw。从平均绝对百分比误差分析，本文所提的预测模型与其它5种相比，在案例一中，MAPE分别是3.37%、2.85%、2.22%、1.41%和0.95%；在案例二中，MAPE分别是3.96%、2.51%、2.03%、1.45%和0.6%。综合分析，本文提出的基于混合ReliefF-PMI-IGWO特征选择算法的短期电力负荷预测模型稳定性有明显提高，负荷预测精度也更好。

## 5. 结论

针对以往电力负荷预测模型预测方法效率低、准确率差，且单一模型预测已经不能适应现在的电力负荷变化等问题，本文提出了基于混合ReliefF-PMI-IGWO的短期电力负荷预测。首先混合两种Filter特征选择方法—Relief-F方法和PMI方法，既提高了重要特征筛选的速度和精度，又避免了一些无关冗余信息的出现。然后选择IGWO算法的Wrapper方法，IGWO算法克服了GWO算法易陷入局部最

优解的缺点，具有更高的稳定性和寻优能力。通过两个实例分析及对比实验可知，基于ReliefF-PMI-IGWO方法的SVR预测模型所获得的2个评价指标MAE和MAPE均是最优，具有更高的预测精度和更好的泛化能力，且与真实值非常接近，在短期电力负荷预测中有很好的应用价值，为后续预测提供了新思路、新方法。但本文模型比其中几个模型的预测时间稍长一点，由于负荷决策是电力部门在前一天进行的，所以结果还是能接受。针对本文模型耗时稍长的问题，在今后的研究中，将研究方向重点集中在数据处理和模型压缩上。

## 致谢

本文为国家自然科学基金项目编号（51974151），辽宁省高等学校创新团队项目（LT2019007）的阶段性成果之一。

## 参考文献

- [1] 范海虹. 基于萤火虫算法的短期电力负荷预测方法 [J]. 中国电力, 2021, 54 (3): 144-148.
- [2] 王淞瑶, 张智晟. 基于量子加权GRU神经网络的电力系统短期负荷预测 [J]. 电力系统及其自动化学报, 2022, 34 (1): 1-7.
- [3] 商立群, 李洪波, 侯亚东等. 基于特征选择和优化极限学习机的短期电力负荷预测 [J]. 西安交通大学学报, 2022, 56 (4): 165-175.
- [4] 杨海柱, 田馥铭, 张鹏等. 基于CEEMD-FE和AOA-LSSVM的短期电力负荷预测 [J]. 电力系统保护与控制. 2022 (13): 126-133.
- [5] 段雪滢, 李小腾, 陈文洁. 基于改进粒子群优化算法的VMD-GRU短期电力负荷预测 [J]. 电工电能新技术, 2022, 41 (5): 8-17.
- [6] 张淑清, 李君, 姜安琪等. 基于FPA-VMD和BiLSTM神经网络的新型两阶段短期电力负荷预测 [J]. 电网技术, 2022, 46 (08): 3269-3279.
- [7] 赵婧宇, 池越, 周亚通. 基于SSA-LSTM模型的短期电力负荷预测研究 [J]. 电工电能新技术, 2022, 41 (6): 71-79.
- [8] Yu Zhuowei; Yang Jiajun; Wu Yufeng; Huang Yi. Short-Term Power Load Forecasting Under COVID-19 Based on Graph Representation Learning With Heterogeneous Features [J]. Frontiers in Energy Research, 2021.
- [9] Prakash N.; Vaikundaselvan B.; Sivaraju S. S. Short-Term Load Forecasting for Smart Power Systems Using Swarm Intelligence Algorithm [J]. Journal of Circuits, Systems and Computers, 2022, 31 (11).
- [10] Ghulam Hafeez, Imran Khan, Sadaqat Jan. Anovel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in smart grid [J]. 2021, Applied Energy, 299 (9): 1171778.



- [11] 胡忠义. 基于文化基因算法的电力负荷预测模型及应用 [D]. 华中科技大学, 2015 (07).
- [12] KONONENKO I. Estimating attributes: Analysis and extensions of RELIEF [C]//European conference on machine learning. Springer, Berlin, Heidelberg, 1994: 171-182.
- [13] E Hancer, An improved evolutionary wrapper-filter feature selection approach with a new initialisation scheme [J]. Special Issue on Feature Engineering, 2021. 11.
- [14] 胡忠义. 基于计算智能技术的电力负荷预测理论及应用 [M]. 武汉: 武汉大学出版社, 2019. 1.
- [15] Sharma A. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1-A strategy for system predictor [J]. Journal of Hydrology, 2000, 239: 232-239.
- [16] 徐宇颂, 邹山花, 卢先领. 基于特征选择和组合模型的短期电力负荷预测 [J]. 中国电力, 2022, 55 (7): 121-127.
- [17] 胡欣, 冯杰, 徐先峰等. 基于特征选择实现多因素电力负荷预测 [J]. 自动化仪表, 2022, 43 (3): 69-74.
- [18] 周传华, 柳智才, 丁敬安等. 基于filter+wrapper模的特征选择算法 [J]. 计算机应用研究, 2019, 36 (7): 1975-2010.
- [19] 杨国华, 郑豪丰, 张鸿皓等. 基于Holt-Winters指数平滑和时间卷积网络的短期负荷预测 [J]. 电力系统自动化, 2022 (06): 73-82.
- [20] Yu Binbin; Li Jianjing; Liu Che; Sun Bo. A novel short-term electrical load forecasting framework with intelligent feature engineering [J]. Applied Energy, 2022, 327.
- [21] Zhengyi Liu, Bo Chang, Fan Cheng. An interactive filter-wrapper multi-objective evolutionary algorithm for feature selection [J]. Swarm and Evolutionary Computation, 2021, 65 (4): 100925.
- [22] Wang Jianzhou; Xing Qianyi; Zeng Bo; Zhao Weigang. An ensemble forecasting system for short-term power load based on multi-objective optimizer and fuzzy granulation [J]. Applied Energy, 2022, 327.
- [23] 曾林俊, 许加柱, 王家禹等. 考虑区间构造的改进极限学习机短期电力负荷区间预测 [J]. 电网技术, 2022 (07): 2555-2563.
- [24] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer [J]. Advances in Engineering Software, 2014, 69 (3): 46-61.
- [25] 王秋萍, 王梦娜, 王晓峰. 改进收敛因子和比例权重的灰狼优化算法 [J]. 计算机工程与应用, 2019, 55 (21): 60-65.