

聚类分析和改进贝叶斯算法的短期负荷预测

为了提高短期负荷的预测准确度，提出一种聚类分析和改进贝叶斯算法的短期负荷预测模型。首先收集短期负荷的历史样本，并进行归一化处理，加快建模速度；然后采用模糊均值聚类算法对短期负荷历史样本进行分类，构建贝叶斯算法的学习样本；最后采用贝叶斯算法建立短期负荷预测模型，并针对贝叶斯算法的不足进行相应改进。采用具体短期负荷历史数据序列对模型的有效性进行仿真测试，结果表明，聚类分析和改进贝叶斯算法的短期负荷预测模型提高了短期电力负荷的预测准确度，加快了模型的训练速度，预测结果更加可靠，可以为电力管理部门科学决策提供参考。

邓小亚/四川文理学院计算机学院



邓小亚/讲师

电力负荷预测是电力生产、调度部门的一项重要日常工作，其预测结果有助于电力管理部门进行科学决策，可以提高经济效率，是电力市场化的基础，因此，电力负荷预测具有十分重要的现实意义，尤其是短期电力负荷预测的价值更高^[1]。

影响短期电力负荷值变化的因素众多，主要包括经济、气象以及季节等因素，短期电力负荷值是这些因素综合作用的结果，但不是简单的单一因素叠加的结果，因此短期电力负荷变化非常复杂^[2]。为了提高短期电力负荷的预测准确度，针对负荷预测建模过程中的训练样本选择和预测算法选择问题，提出一种模糊均值（Fuzzy C-Means, FCM）聚类分析算法与改进贝叶斯（Modified Bayesian Learning, MBL）算法^[3]的短期负荷预测模型（FCM-MBL），并采用

具体短期电力负荷数据对FCM-MBL的性能进行测试，以验证其有效性。

短期电力负荷的相关研究工作

多年来，大量的学者和研究机构对短期电力负荷进行了深入研究，当前短期电力负荷的预测模型主要分为三类：传统模型、智能模型以及组合模型。传统模型主要有电力弹性系数模型、回归分析模型和时间序列分析模型，其建模的原理简单、建模速度快，是一类线性预测模型，然而短期电力负荷的影响因素具有多变性，这些传统模型不能反映影响因子与短期负荷之间的非线性变化关系，预测准确度不能满足电力管理部门实际的要求，存在局限性；智能模型是一类非线性预测模型，主要包括专家系统、神经网络等，其预测性能要优于传统模型，但是也存在一定的缺点，如专家系统通用性差、没有学习能力以及适应能力不强，神经网络有逼近任意连续函数的能力，但是泛化能力不强，存在过拟合预测现象。近年来，一些学者根据最优组合规则，即将单一传统模型与单一智能模型进行组合，利用各种模型提供更多的信息，取得了较好的短期负荷预测效果，但是如何确定各单一模型的权值没有比较合适的理论指导，预测结果有待改善。

聚类分析算法和改进贝叶斯算法

1. 聚类分析算法

短期电力负荷变化十分复杂，由于天气、季节变化具有一定的模糊性，因

关键词/Keywords

短期负荷·
改进贝叶斯学习算法·
模糊均值聚类·

四川省教育厅一般项目（14ZB0315），四川文理学院智能计算与物联网工程技术中心资助。

此本文选择FCM聚类分析算法对短期负荷样本的不确定性进行描述,划分各种类型的子样本,从而更加准确地反映短期负荷的变化特性。FCM是一种无监督的学习算法,采用隶属度描述每个元素的类别,假设数据集 $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$; 共有 C 个类别,聚类中心 $\mathbf{V} = \{v_1, v_2, \dots, v_C\}$; $\mathbf{U} = (u_{ij})$ 为每个样本的隶属度矩阵,模糊聚类目标函数为

$$J = \sum_{i=1}^N \sum_{j=1}^C u_{ij} \|x_i - v_j\|^2 \quad (1)$$

式中, u_{ij} 为个体 x_i 属于第 j 类的模糊隶属度,其计算公式为

$$u_{ij} = \begin{cases} \left[\sum_{k=1}^C \frac{\|x_i - v_k\|^{\frac{2}{m-1}}}{\|x_i - v_j\|^{\frac{2}{m-1}}} \right]^{-1} & \|x_i - v_k\| \neq 0 \\ 1 & \|x_i - v_k\| = 0 \text{ 且 } k = j \\ 0 & \|x_i - v_k\| = 0 \text{ 且 } k \neq j \end{cases} \quad (2)$$

式中, m 为权重值; v_j 为第 j 类的聚类中心,其计算公式为

$$v_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (3)$$

FCM聚类分析算法的工作步骤如下:

- 1) 给定模糊权重值 m 。
- 2) 给定初始聚类中心 v 。
- 3) 计算隶属度矩阵 $\mathbf{U}$ 、每一类中心 v 和模糊聚类目标值。
- 4) 如果满足终止条件,迭代结束;否则跳转到步骤3)。

5) 得到模糊隶属度矩阵 $\mathbf{U}$,个体根据矩阵每列最大元素位置判断个体的所属分类。

2. 贝叶斯学习算法

对于短期电力负荷预测问题,设 $\mathbf{x}_i$ 和 $\mathbf{y}_i$ 分别为第 i 组数据的输入与输出向量,那么输入与输出之间可以采用如下方程进行描述

$$\mathbf{y}_i = f(\mathbf{x}_i; \mathbf{w}) + \varepsilon_i \quad (4)$$

式中, $\mathbf{w} = (w_1, w_2, \dots, w_M)^T$ 为待推断的权向量; ε_i 为误差项。

引入核函数, $f(\mathbf{x}_i; \mathbf{w})$ 即可以变为如下表达式

$$f(\mathbf{x}_i; \mathbf{w}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) = \sum_{i=1}^M \omega_i K(\mathbf{x}, \mathbf{x}_i) + \omega_0 \quad (5)$$

在多次独立试验中,目标值数据集的概率函数可以表示为

$$\begin{aligned} p(\mathbf{y} | \mathbf{x}_i, \mathbf{w}, \sigma^2) &= \prod_{i=1}^N N[\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i), \sigma^2] \\ &= \prod_{i=1}^N (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2} [\mathbf{y}_i - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i)]^2\right\} \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{y} - \boldsymbol{\phi}\mathbf{w}\|^2\right\} \end{aligned} \quad (6)$$

式中, $\boldsymbol{\phi}(\mathbf{x}_i) = [1, K(\mathbf{x}_i, \mathbf{x}_1), \dots, K(\mathbf{x}_i, \mathbf{x}_M)]^T$ 。

3. 贝叶斯学习算法的改进

对于标准贝叶斯学习算法,其计算复杂度,建模效率低,影响短期负荷预测的应用范围,为此对标准贝叶斯学习算法进行改进,改进贝叶斯学习算法的 $\mathbf{w}$ 先验概率分布为

$$p(\mathbf{w} | \alpha) = \prod_{i=1}^M N(0, \alpha_i^{-1}) \quad (7)$$

那么, $\mathbf{w}$ 的后验概率分布的计算公式为

$$p(\mathbf{w} | \mathbf{t}, \alpha, \sigma^2) = \frac{p(\mathbf{w}, \alpha, \sigma^2 | \mathbf{t})}{p(\alpha, \sigma^2 | \mathbf{t})} \quad (8)$$

对于新给定输入向量 $\mathbf{x}^*$,可以求得与其对应的目标输出 $\mathbf{t}^*$ 的后验预测分布为

$$p(\mathbf{t}^* | \mathbf{x}^*, \alpha, \beta) = \int p(\mathbf{t}^* | \mathbf{w}, \beta) p(\mathbf{w} | \alpha, \beta) d\mathbf{w} \quad (9)$$

FCM-MBL的短期负荷预测模型

1. FCM-MBL的短期负荷的建模思想

短期负荷是一种复杂的非线性系统,贝叶斯学习算法具有强大的非线性处理能力,因此,本文将其引入到短期负荷建模与预测中,并且采用FCM算法对短期负荷样本进行聚类分析,使样本更具有规律性,因此,FCM-MBL的短期负荷预测模型的工作思想为:首先收集短期负荷历史样本,并进行归一化处理,加快建模速度;然后采用FCM聚类算法对短期负荷历史样本进行分类,构建贝叶斯算法的学习样本;最后采用贝叶斯学习算法建立短期负荷预测模型,并针对贝叶斯算法的不足进行相应改进,具体流程如图1所示。

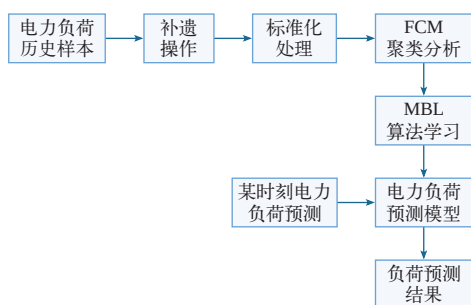


图1 FCM-MBL的短期负荷建模与预测流程

2. FCM-MBL的短期负荷建模步骤

1) 收集短期负荷的历史样本, 由于短期负荷历史样本不可避免会发生丢失或者存在异常点, 因此本文采用该丢失点或异常点的前一个点和后一点平均值进行补遗操作, 具体如下

$$x_i = r_1 x_{i-1} + r_2 x_{i+1} \quad (10)$$

式中, x_{i-1} 和 x_{i+1} 分别为前一个点和后一个点的负荷值; r_1 、 r_2 为权重系数。

2) 短期负荷由于受多种因素影响, 具有时变性, 为此在贝叶斯学习算法建模之前, 对原始短期负荷数据进行标准化处理, 具体如下

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

式中, x'_i 为标准化后的短期负荷值; $x_{\max}$ 、 $x_{\min}$ 分别为短期负荷的最大、最小值。

3) 采用FCM算法对短期负荷数据进行聚类分析, 构建短期负荷样本集类别信息, 计算所有样本与聚类中心之间的距离, 从而得到相应的贝叶斯学习算法的训练样本子集。

4) 采用MBL算法建立短期负荷预测模型。

5) 采用建立的短期负荷预测模型对未来某时刻的短期负荷进行预测。

FCM-MBL在短期负荷预测中的应用

1. 数据来源

为了测试FCM-MBL的短期负荷预测的性能, 采用某地区的2014年5月1日~7月20日的短期负荷数据进行仿真测试, 采样时间为1 h, 这样总共获得了1 200个短期负荷样本, 最后200个样本作为测试样本集, 采用Matlab 2012工具箱实现算

法, 数据具体如图2所示。

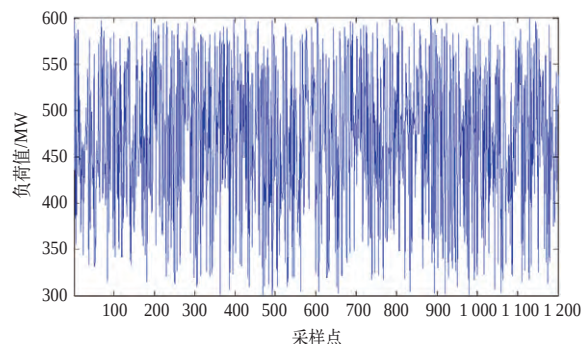


图2 某地区的短期负荷历史样本

2. 预测结果的评价标准

选择没有聚类分析的MBL短期负荷预测模型(MBL)、FCM聚类分析算法和BP神经网络的短期负荷预测模型(FCM-BPNN)作为对比模型, 模型性能评价指标为平均绝对误差(MAPE)和方均根相对误差(RMSE), 即

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^N [\hat{x}(i) - x(i)]^2} \quad (12)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^N \left| \frac{\hat{x}(i) - x(i)}{x(i)} \right| \times 100\% \quad (13)$$

式中, n 为样本数。

3. 结果与分析

(1) 预测准确度分析

MBL、FCM-MBL和FCM-BPNN对测试集的预测结果、预测误差的变化曲线如图3和图4所示, 各模型的MAPE和RMSE误差对比如下表所示。从图3、图4以及下表可知, 所有模型的预测准确度均比较高, 预测误差不大, 但是FCM-

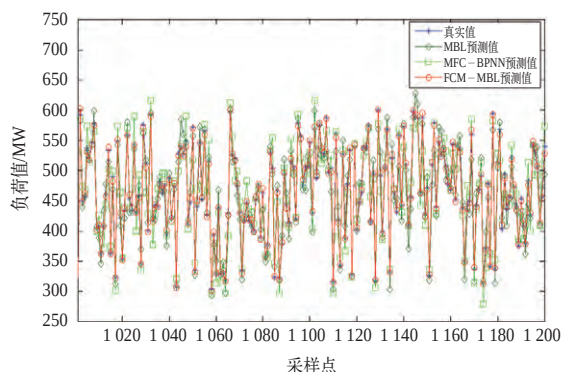


图3 FCM-MBL与对比模型的预测结果变化曲线

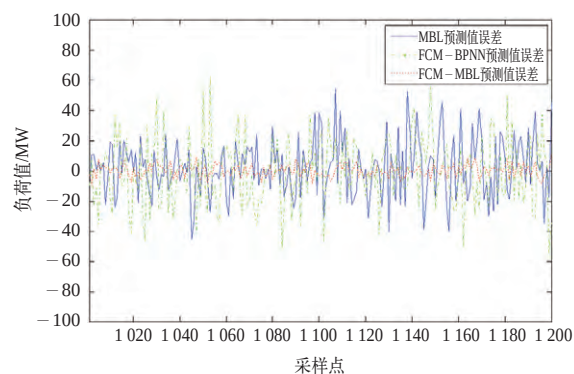


图4 FCM-MBL与其他模型的误差变化曲线

表 FCM-MBL与其他模型的预测误差对比

模型	MAPE (%)	RMSE (%)
MBL	3.34	19.55
FCM-BPNN	4.04	23.47
FCM-MBL	0.67	3.89

MBL的预测性能最优，仔细分析可以得到如下结论：

1) MBL的MAPE和RMSE分别为3.34%和19.55%；FCM-MBL的MAPE和RMSE分别为0.67%和3.89%。相对于MBL模型，FCM-MBL的平均相对误差更小，主要由于FCM算法对短期负荷样本进行聚类分析，获取预测样本点对应的最优训练子集，加强了MBL训练建模的样本数据规律性，减少了MBL对训练集的依赖性，提高了短期负荷预测准确度。

2) FCM-BPNN的MAPE和RMSE分别为4.04%和23.47%；FCM-MBL的MAPE和RMSE分别为0.67%和3.89%，整体预测性能得以改善；且FCM-BPNN的预测结果相对误差波动范围较大，变化比较剧烈，说明模型的预测结果不稳定、不可靠，而FCM-MBL的预测结果更加稳定和可靠，这主要由于采用了泛化能力更优的MBL算法，其克服了BPNN的过拟合、泛化能力差的缺陷，有效提高了短期电力负荷的预测准确度。

(2) 建模效率分析

为了测试FCM-MBL模型与对比模型的建模效率，采用平均训练时间作为评价指标，其平均

训练时间如图5所示。从图5可知，相对于MBL、FCM-BPNN模型，FCM-MBL的平均训练时间最短，计算复杂度最低，建模效率最高，这表明FCM-MBL的应用更广泛，应用价值相对更高。

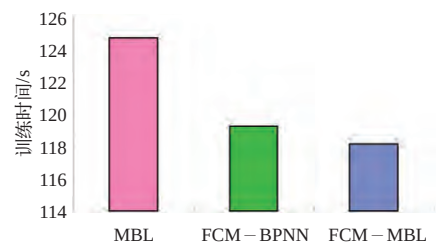


图5 FCM-MBL与对比模型的训练时间对比

结束语

针对短期负荷数据的模糊性、影响因素的多样性以及非线性变化的特点，综合利用FCM算法和MBL算法的优点，提出一种FCM-MBL的短期负荷预测模型。首先采用FCM聚类分析算法的处理不确定性和模糊信息的能力挖掘短期负荷数据的模糊性、多样性，得到短期负荷预测点的相应学习样本集；然后采用MBL算法进行短期负荷非线性建模。仿真实验结果表明，FCM-MBL不仅预测准确度要高于其他短期负荷预测模型，而且建模效率得到相应的改善，短期负荷预测结果更加理想，可以满足短期电力负荷预测的实际要求。

参考文献

[1] 康重庆，夏清，张伯明．电力系统负荷预测研究综述与发展方向的探讨[J]．电力系统自动化，2004，28(17)：1-11.

[2] 杨正瓴，张广涛，林孔元．时间序列法短期负荷预测准确度上限估计[J]．电力系统及其自动化学报，2004，16(2)：36-39.

[3] 王勇，黄国兴，彭道刚．带反馈的多元线性回归法在电力负荷预测中的应用[J]．计算机应用与软件，2008，25(1)：83-84.

(收稿日期：2015-02-10) EA