

# 基于数据挖掘的 FNN 短期电力负荷预测方法研究

崔和瑞<sup>1</sup>, 宋秀莉<sup>1</sup>, 葛曼倩<sup>2</sup>

(1. 华北电力大学工商管理学院, 河北 保定 071003; 2. 青岛电力设计院有限公司, 山东 青岛 266002)

**摘要:** 针对人工神经网络模型在进行负荷预测时, 大多不考虑气象等因素的影响, 提出了一种基于数据挖掘预处理的改进短期电力负荷预测的方法, 应用数据挖掘的聚类功能, 寻找与预测日同等气象类型的多个历史短期负荷数据序列进行预测, 从而提高预测的精度。鉴于 ANN 模型对不确定性和模糊信息学习处理能力较差的缺点, 引用模糊系统的理论, 构建模糊神经网络 (FNN) 模型。通过实例预测和预测结果比较分析表明, 提出的方法具有较高的预测精度。

**关键词:** 数据挖掘; 聚类分析; 短期电力负荷预测; 人工神经网络; 模糊神经网络

## Research on FNN short-term electric load forecasting based on data mining technology

CUI He-rui<sup>1</sup>, SONG Xiu-li<sup>1</sup>, GE Man-qian<sup>2</sup>

(1. School of Business Administration, North China Electric Power University, Baoding 071003, China;  
2. Qingdao Electric Power Design Institute, Qingdao 266002, China)

**Abstract:** The artificial neural network model does not consider the weather and other factors in load forecasting. Based on data mining (DM), this paper proposes the improved short-term load forecasting method, by the clustering feature of DM, we search the historical daily load with the same meteorological category as the forecasting day and to compose data sequence with highly similar meteorological features, so as to enhance the accuracy of the forecasting. In view of the poor ability of ANN model to deal with the uncertainty and ambiguity information, this paper cites the theory of fuzzy to build fuzzy neural network (FNN) model. Comparing with ANN and FNN, the approach in this paper has achieved greater forecasting accuracy.

This work is supported by National Natural Science Foundation of China (No. 70671039).

**Key words:** data mining; cluster analysis; short-term electric load forecasting; artificial neural network; fuzzy neural network

中图分类号: TM715 文献标识码: A 文章编号: 1674-3415(2009)22-0054-04

## 0 引言

电力负荷预测是电力调度、用电、计划、规划等管理部门的重要工作之一, 其中, 短期负荷预测是电力系统调度管理部门制订开停机计划及在线安全分析的基础, 也是电力市场中实现电能计划管理的基础。

多年来, 许多学者对短期负荷预测方法进行了深入的研究<sup>[1,2]</sup>, 这些方法大致可以分为3类: 传统预测方法<sup>[2-4]</sup>(如时间序列模型、回归模型等)、智能预测方法<sup>[5-7]</sup>(如专家系统、人工神经网络模型和模糊逻辑等)以及一些新的短期负荷预测方法(如混沌预测法<sup>[5]</sup>等)。但由于负荷变化的影响因素多且过程复杂, 这些方法各自都存在一定的缺陷和局限性, 时间序列法计算量小, 速度较快, 但预测精度不能

保证满足工程需要, 且不具备自适应学习能力, 预测系统的鲁棒性没有保障<sup>[8]</sup>。专家系统可以避开复杂的数值计算, 但通用性较差, 缺乏学习能力<sup>[9,10]</sup>。人工神经网络具有很强的鲁棒性、记忆能力、非线性映射能力及强大的自学习能力, 但普通神经网络模型<sup>[11]</sup>大多不考虑气象因素影响, 并且 BP 网络学习算法实际上是利用梯度下降法调节权值使目标函数达到极小, 导致了 BP 网络过分强调克服学习错误而泛化性能不强; 对不确定性和模糊信息学习处理能力较差等缺点。

基于以上分析, 本文引入数据挖掘强大的数据处理技术, 首先对负荷预测的历史数据进行聚类分析, 寻找与预测日相似气象类型的多个历史负荷数据, 组成具有高度相似气象特征的数据序列, 进而再运用模糊神经网络 FNN (fuzzy neural network) 模型进行负荷预测, FNN 是将模糊系统和神经网络结合。预测结果显示, 本文的方法 (DMFNN) 比

模糊神经网络(FNN)以及单纯使用人工神经网络(ANN)具有更高的预测精度。

### 1 聚类分析介绍

数据挖掘是一个从大量数据中抽取挖掘出未知的、有价值的模式或规律等知识的复杂过程, 它的主要功能有: 分类、聚类、关联和预测。聚类是指将数据对象分组成为多个类, 在同一类中的对象之间具有较高的相似性, 而不同类中的对象差别较大。聚类做为数据挖掘的一种方法被广泛地应用。它既可以作一个独立的方法对数据进行分析, 也可以作为其它数据挖掘方法的预处理步骤。聚类的目的在于确定原始数据在其度量空间中是否存在集群分布的现象。如果集群现象存在, 则将相应的数据集分成相应数量的数据子集。

将原始数据的度量空间用  $n$  维的实数空间  $R_n$  表示, 设给定样本  $X = \{x_1, x_2, \dots, x_n\}$ , 其中,  $x_i \in R_n$ , 且  $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\} \quad i=1, 2, \dots, n$ 。

在短期负荷预测中, 运用聚类分析技术进行样本数据的选择, 还要解决样本数据用怎样的一个  $n$  维的特征向量表示的问题, 对负荷影响较大的天气因素有日最高温度、日最低温度、日平均温度和日降雨量、湿度、气压等。由于不同的物理量量纲, 难以进行比较。为了消除指标特征值不同物理量量纲的影响, 使这些因素的取值具有可比性, 就要根据  $x_i$  的各项指标, 将数据进行标准化处理,

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

其中:  $s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ 。

将标准化数据压缩到  $[0, 1]$  区间,

$$x''_{ij} = \frac{x'_{ij} - x'_{j\min}}{x'_{j\max} - x'_{j\min}} \quad (2)$$

其中:  $x'_{j\max}$  和  $x'_{j\min}$  分别是标准化数据中的最大和最小值。

计算  $x_i$  和  $x_j$  的相似系数  $r_{ij}$ ,

$$r_{ij} = \frac{\left| \sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \right|}{\left[ \sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \times \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2 \right]^{1/2}} \quad (3)$$

其中:  $\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ik}, \bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{jk}, i, j=1, 2, \dots, n$ 。

对于相似关系矩阵  $R = (r_{ij})_{n \times n}$ , 一般只满足反射性和对称性, 因此, 还需将  $R$  转化成等价矩阵  $R^*$ 。为加速收敛, 通常采用求传递闭包的方法, 直到计算出  $R^{2^k} = R^k = R^*$ , 此时,  $R^*$  即为等价矩阵。再根据不同的聚类水平进行聚类分析。在适当的阈值  $\lambda (\lambda \in [0, 1])$  上进行截取, 便可得到所需要的分类。

### 2 模糊神经网络模型介绍

BP 神经网络作为人工神经网络的一个主要模型, 有着广泛的应用。它是一种单向传播的多层前向网络, 其输入信号从输入节点依次传过各隐含层, 然后传到输出节点, 每一层节点的输出只影响下一层节点的输入。它通过将神经网络的实际输出与期望输出之间的误差反向传播, 整个训练的过程是网络全局误差趋向极小值的过程。

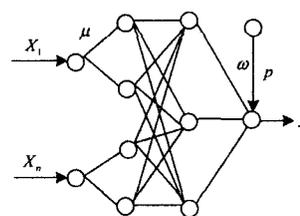


图1 模糊神经网络结构

Fig.1 Structure of fuzzy neural network

为了适合短期负荷预测的研究, 图1所示模糊神经网络模型由输入层、隶属函数生成层、模糊推理层、反模糊化层以及“干预项”组成<sup>[13]</sup>。层与层之间依据模糊逻辑系统的语言变量、模糊 if-then 规则、模糊推理方法、反模糊化函数所构建。第一层的各个结点直接与输入向量相连接; 第二层的每一个结点代表一个语言变量值; 其作用是计算各输入向量的隶属函数; 第三层的每一个结点代表一条模糊规则, 其作用是匹配模糊规则, 并计算出每条规则的适用度。

隶属函数采用高斯型:

$$\mu_i^j = \exp\left[-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}\right] \quad (4)$$

$$1 \leq i \leq n, \quad 1 \leq j \leq m$$

其中:  $\mu_i^j$  为隶属度函数生成层的数值,  $x_i$  为实际输入,  $n$  为输入变量的个数,  $m_i$  为  $x_i$  的模糊分割数,  $c_{ij}$  和  $\sigma_{ij}$  分别为隶属函数的中心和宽度。

反模糊化层的输出函数为:

$$Y = (\pi_1 \omega_1 + \pi_2 \omega_2 + \dots + \pi_m \omega_m)(1 + p) \quad (5)$$

其中： $m = \prod_{i=1}^n m_i$ ， $Y$  为网络输出的归一化数据， $\pi_k$  为采用乘积推理规则的结果， $p$  为干预项隶属函数值， $\omega_k$  为模糊推理层与反模糊化层的连接权。

此模糊神经网络采用 BP 算法，平方误差为：

$$E = (y - Y)^2 / 2 \quad (6)$$

其中： $Y$  为实际输出， $y$  为期望输出。

$$c_{ij}(t+1) = c_{ij} - \eta \frac{\partial E}{\partial c_{ij}} \quad (7)$$

$$\sigma_{ij}(t+1) = \sigma_{ij} - \eta \frac{\partial E}{\partial \sigma_{ij}} \quad (8)$$

$$\omega_i(t+1) = \omega_i - \eta \frac{\partial E}{\partial \omega_i} \quad (9)$$

其中： $\eta$  为学习率， $t$  为迭代次数。

温度变化  $\Delta T$  和修正因子  $Z$  的语言值均为： $NV$  (负大)、 $N$  (负)、 $ZE$  (零)、 $P$  (正)、 $PV$  (负大)。假设有两条规则：

- 规划 1: IF  $\Delta T$  is  $NV$ , then  $Z$  is  $N$
- 规划 2: IF  $\Delta T$  is  $ZE$ , then  $Z$  is  $ZE$

每条规则的使用度为：

$$\begin{cases} \alpha_1 = \mu_{NV}(\Delta T) \\ \alpha_2 = \mu_{ZE}(\Delta T) \end{cases} \quad (10)$$

根据误差的大小来修正最终的预测结果，对训练时误差较大的点进行修正，修正的计算为：

$$E = \frac{\alpha_1 z_1 + \alpha_2 z_2}{\alpha_1 + \alpha_2} \quad (11)$$

本文采用预测日的归一化曲线的方法预测和模糊推理修正法得到最终的归一化曲线。具体算法流程如图 2 所示。

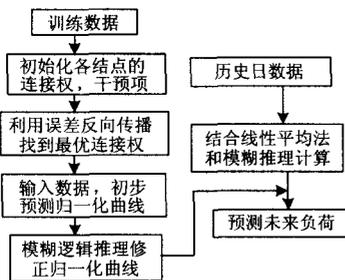


图 2 预测算法流程

Fig.2 Flow of forecasting algorithm

### 3 实证分析

#### 3.1 样本选择

运用本文所提新方法，以佛山某电网的电力负荷数据为对象进行短期负荷预测，该地区电网的电力负荷变化规律受气象影响非常显著，按照通常经典预测方法和普通人工神经网络方法预测有较大误差。选取该电网自 2004 年 1 月至 2005 年 1 月期间的负荷数据库作为训练样本，2005 年 2 月的负荷数据作为测试样本。取对负荷影响较大的天气因素：日最高温度、日最低温度、日平均温度和日降雨量，通过本文提出的数据挖掘技术找出和预测点在气象特性和预测时段都相近的数据建立 FNN 模型。多次取样预测显示， $p$  取 0.036 时的预测结果与实际负荷非常接近。

#### 3.2 误差测量与结果分析

本文选取平均相对误差  $e$  作为预测效果判断的根据：

$$e = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - y_i}{Y_i} \right| \times 100\% \quad (12)$$

其中： $Y_i$  和  $y_i$  分别表示实际值和预测值，以 3% 为评价准则，若  $e > 3\%$ ，则该预测结果合格。用  $r$  表示负荷预测结果的不合格率：

$$r = [\text{num}(e > 3\%) / n] \times 100\% \quad (13)$$

其中： $\text{num}(e > 3\%)$  表示日预测结果平均相对误差超过 3% 的个数， $n$  为总负荷点数。

用本文介绍的模型 (DMFNN) 和单纯的 FNN 方法、普通人工神经网络模型 (ANN) 对 2005 年 2 月 13 日至 26 日连续 14 日的负荷进行预测，并进行比较。从表 1 可以看出，本文方法的相对误差值为 2.427%，比 FNN 的 3.037% 和 ANN 的 3.547% 小。本方法的不合格率平均值为 4.337%，比 FNN 的 5.520% 和 ANN 的 6.5453% 小。

表 1 预测误差比较

Tab. 1 Comparison of forecasting error

日期	DMFNN		FNN		ANN	
	$e$ (%)	$r$ (%)	$e$ (%)	$r$ (%)	$e$ (%)	$r$ (%)
2-13	2.58	4.218	2.71	5.593	3.01	6.426
2-14	2.63	4.352	2.78	5.028	3.10	6.102
2-15	2.64	5.016	5.13	7.226	7.84	7.638
2-16	1.22	3.260	1.57	4.196	1.61	5.017
2-17	1.68	4.217	1.79	6.382	2.95	7.142
2-18	3.02	3.749	4.93	6.784	4.32	7.210
2-19	1.74	5.348	1.65	3.602	1.80	6.438
2-20	3.39	4.663	3.19	5.386	4.06	6.321
2-21	1.98	4.321	2.10	5.129	2.41	6.359
2-22	3.30	3.615	5.73	6.217	6.23	7.001
2-23	1.72	5.390	1.86	6.079	2.25	7.204
2-24	3.26	4.512	3.75	5.050	3.94	6.702
2-25	2.65	4.058	2.79	5.751	3.16	6.821
2-26	2.17	3.992	2.54	4.852	2.98	5.363

由此可以看出,本文方法预测效果好、精度高,是一种新的实用型方法,按此方法处理气象影响因素,能够有效地提高短期负荷预测的精度,能相对简化负荷预测模型,有效地克服普通网络的缺点,使预测误差进一步减小,达到了实用的要求。

#### 4 结论

(1) 本文重点考虑了气象中的主要影响因素对负荷预测的影响,提出了基于聚类分析的数据挖掘技术,以减少数据规模,减少噪声数据的干扰,提高训练速度,并将聚类后的数据作建模预测的样本数据。

(2) 本文运用模糊神经网络的预测方法,引入了“干预项”和模糊推理补偿了其他因素的影响,该方法具有处理不确定性和模糊信息的能力,能够有效地提高预测精度。

(3) 通过与FNN模型和单纯的ANN模型的实际预测比较,证明本文所建模型具有较高的精度和实用性。

#### 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.  
KANG Chong-qing,XIA Qing,ZHANG Bo-ming. Review of Power System Load Forecasting and Its Development[J].Automation of Electric Power Systems, 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.  
YANG Zheng-ling,ZHANG Guang-tao,LIN Kong-yuan. Upper Limit Estimating of Short Term Load Forecasting Precision by Time Series Analysis[J].Proceedings of the CSU-EPSSA, 2004,16(2):36-39.
- [3] 蒋刚.基于模糊支持向量核回归方法的短期峰值负荷预测[J].控制理论与应用,2007,24(6):986-990.  
JIANG Gang. Short-term Peak Load Forecasting Based on Fuzzy Support Vector Kernel Regression Method[J].Control Theory & Applications, 2007,24(6):986-990.
- [4] 林辉,郝志峰,蔡瑞初.基于双时间序列神经网络的短期电网负荷预测[J].计算机工程与应用,2007,43(32):218-220.  
LIN Hui,HAO Zhi-feng,CAI Rui-chu. Short-term Load Forecasting Based on Two-time-series Neural Network Model[J]. Computer Engineering and Applications, 2007,43(32):218-220.
- [5] 李广,邹德忠,谈顺涛.基于混沌神经网络理论的小电网短期电力负荷预测[J].电力自动化设备,2006,2:50-52.  
LI Guang,ZOU De-zhong, TAN Shun-tao. Short-term Load Forecast for Small Power Net Based on Chaos-artificial Neural Network Theory[J].Electric Power Automation Equipment,2006,26(2):50-52.
- [6] 赵宇红,肖金凤,陈忠泽.混合模糊神经网络在短期负荷预测中的应用[J].电力系统及其自动化学报,2006,2:99-104.  
ZHAO Yu-hong, XIAO Jin-feng,CHEN Zhong-ze. Application of Hybrid Fuzzy Neural Network in Short-Term Load Forecasting[J]. Proceedings of the CSU-EPSSA, 2006,2:99-104.
- [7] 王楠,刘晨光,马立肖,等.基于模糊逻辑和神经网络的电力负荷预测[J].微计算机信息,2007,23(2):313-315.  
WANG Nan,LIU Chen-guang, MA Li-xiao, et al. ANN and Fuzzy Logic Based Short-term Load Forecasting[J].Micro Computer Information, 2007,23(2):313-315.
- [8] 罗滇生,姚建刚,何洪英.基于自适应滚动优化的电力负荷多模型组合预测系统的研究与开发[J].中国电机工程学报,2003,23(5):58-61.  
LUO Dian-sheng,YAO Jian-gang,HE Hong-ying. Research and Development of Multi-model Combining Load Forecasting System Based on SELF-adaptive Rolling Optimization[J].Proceedings of the CSEE, 2003,23(5):58-61.
- [9] 莫维仁,张伯明,孙宏斌.扩展短期负荷预测的原理和方法[J].中国电机工程学报,2003,23(3):1-4.  
MO Wei-ren,ZHANG Bo-ming,SUN Hong-bin. Extended Short-termload Forecasting Principle and Method[J]. Proceedings of the CSEE, 2003,23(3):1-4.
- [10] 陈耀武,汪乐宇,龙洪玉.基于组合式神经网络的短期电力负荷预测模型[J].中国电机工程学报,2001,21(4):79-82.  
CHEN Yao-wu,WANG Le-yu, LONG Hong-yu. Short-term Load Forecasting with Modular Neural Networks[J]. Proceedings of the CSEE,2001,21(4):79-82.
- [11] XIA You-shen, WANG Jun. A General Projection Neural Network for Solving Monotone Variational Inequalities and Related Optimization Problems[J]. IEEE Trans on Neural Networks, 2004,15(2):318-328.
- [12] 陈安,陈宁,周龙骧.数据挖掘技术及应用[M].北京:科学出版社,2006  
CHEN An,CHEN Ning,ZHOU Long-xiang. Data Mining Technology and Application[M]. Beijing: Science Press,2006.
- [13] 耿伟华,孙衢,张翠霞,等.基于改进的模糊神经网络的短期负荷预测[J].电力系统及其自动化学报,2007,19(5):29-32.  
GENG Wei-hua,SUN Qu,ZHANG Cui-xia,et al. Short-Term Load Forecasting Based on Improved Fuzzy Neural Network[J].Proceedings of the CSU-EPSSA,2007,19,5:29-32.

收稿日期:2008-11-26

作者简介:

崔和瑞(1967-),男,博士,教授,硕士研究生导师,主要研究方向为区域经济可持续发展、数量经济分析;E-mail: cuiherui1967@126.com

宋秀莉(1981-),女,硕士研究生,主要研究方向为区域经济可持续发展;

葛曼倩(1975-),女,工程师,主要从事电力工程设计与开发工作。