

中长期负荷预测的模糊竞争学习聚类神经网络算法

岳璐, 张尧

(华南理工大学电力学院, 广东 广州 510640)

摘要: 电力系统中长期负荷预测受大量不确定因素的影响, 聚类方法能够将各种影响因素综合引入预测模型, 提高了预测精度。本文将神经网络引入到模糊聚类分析中, 建立了中长期负荷预测的新方法, 并且对聚类神经网络的学习算法进行了改进, 利用模糊竞争学习完成网络运算, 弥补了网络输出结果二值性的不足, 使得学习规则中权值矩阵的改变速度加快, 因而算法的收敛速度有很大提高。运用文中所述模型及算法综合考虑了历史负荷情况和未来不确定因素等对未来负荷变化的影响。通过与传统方法进行中长期负荷预测比较, 结果表明该方法可以提高负荷预测的精度。

关键词: 中长期负荷预测; 聚类神经网络; 模糊竞争学习; 信息扩充法

The clustering neural network based on fuzzy competitive learning algorithm for middle and long term load forecasting

YUE Lu, ZHANG Yao

(Electric Power College, South China University of Technology, Guangzhou 510640, China)

Abstract: Middle and long term load forecasting of power system is affected by various uncertain factors. Using clustering method, numerous relative factors can be synthesized for the forecasting model so that the accuracy of the load forecasting would be improved significantly. The new method introduces the neural network into the fuzzy clustering and establishes the new model of mid-long term load forecasting. The method also makes improvement in the learning algorithm. It adopts the fuzzy competitive learning to solve the binary results of the network output and makes the change rate of the weight matrix speed up. So the convergence speed is improved effectively. The proposed model considers the influences of both history and future uncertain factors. Compared with the traditional methods, the results show that the new algorithm improves the accuracy of load forecasting considerably.

Key words: mid-long term load forecasting; clustering neural network; fuzzy competitive learning algorithm; information expansion

中图分类号: TM715 文献标识码: A 文章编号: 1003-4897(2008)08-0055-04

0 引言

电力系统中长期负荷预测作为电力系统规划建设的基础和前提, 具有十分重要的意义。随着我国电力事业的发展, 电网管理日趋现代化, 负荷预测问题的研究也越来越引起人们的重视, 成为现代电力系统科学中的一个重要领域。准确的负荷预测对提高电网的经济运行以及电网的优化规划和增容有十分重要的作用^[1]。

中长期负荷预测精度受经济、政策、气候等各种随机因素的影响, 准确进行预测是一项复杂的工作。常规的预测方法如相关分析法、回归分析法、时间序列法等往往直接建立数学表达式对相关因素和负荷之间的关系加以描述^[2], 不能对其中的随机

成分做统计处理, 准确性得不到很好的保证。而灰色系统预测法克服了常规方法中概率统计的弱点, 使原始数列的随机波动性对预测结果精度的影响减小, 有效解决了常规方法中无法考虑随机因素的问题, 但是当历史数据离散度大时, 预测精度也会降低^[3]。近些年兴起的神经网络法的自学习和自适应能力是常规算法和灰色预测法所不具备的, 它能有效提高中长期负荷预测中各种相关变量之间的复杂非线性映射关系, 但是网络学习需要的数据量大, 并且收敛速度慢。相对上述方法而言, 聚类方法可以将各种影响因素综合引入预测模型, 考虑了历史负荷情况和未来不确定因素对未来负荷变化趋势的影响, 避免了由于输入变量过多而导致运算时间过长等不足。因此在中长期负荷预测中有广泛的应

用^[4]。

本文提出了基于模糊竞争学习聚类神经网络的电力系统中长期负荷预测新方法。通过聚类将历年历史数据中的典型特征进行提取归类,得到最优聚类模型,进而由影响未来负荷变化的相关因素来判定未来负荷变化属于哪种模式,建立负荷模型进行预测。该方法将神经网络技术引入到模糊聚类中,发挥了神经网络和模糊聚类理论处理非线性问题的能力,具有训练速度快,学习精度高,数值稳定等优点。并且对学习规则进行了改进,克服了网络输出结果二值性的不足,使得学习算法的收敛速度加快。通过实例验证,本文提出的中长期负荷预测方法具有更高的预测精度。

1 模糊竞争学习聚类神经网络数学模型

1.1 模糊竞争学习聚类神经网络结构

模糊聚类分析是一种重要的数据分析手段。它可以得到样本属于各个类别的不确定性程度,表达样本类属的中性,建立起样本对于类别的不确定性描述,能客观反映现实世界,从而在生产实践中有广泛的应用^[5]。迄今为止提出的各种聚类算法的目的都是为了提高聚类质量和运算速度。随着神经网络技术的飞速发展,具有分布式存储、并行处理和自学习功能的神经网络技术引起人们的广泛重视。C.Bezdek 和 N.R.Pal 提出了一种基于竞争学习算法的聚类神经网络^[6],开辟了用神经网络做聚类分析的新领域。

模糊竞争学习聚类神经网络结构如图 1 所示。设第 m 年的负荷量为 l^m ,影响第 m 年负荷变化的 n 个环境已知量为 x_i^m ($i=1,2,\dots,n$)。 M 为历史数据年份的数目,即有 M 个待分类对象。每年数据由 n 个环境指标和 1 个负荷量指标组成,令 $l^m = x_{n+1}^m$ 。

在对历史数据的聚类中,直接以原始数据对网络进行训练会引起神经元饱和,因此在对网络进行训练之前必须对数据进行归一化处理,以消除原始数据形式不同所带来的不利。不同量纲的数据通过归一化处理映射到 $[0,1]$ 后进行聚类分析。对每年历史数据归一化处理构成如下序列:

$$I^m = (x_1^m, x_2^m, \dots, x_n^m, l^m) = (x_1^m, x_2^m, \dots, x_n^m, x_{n+1}^m) \quad (1)$$

每一年环境变量 x_i^m ($i=1,2,\dots,n$) 和负荷量 l^m 组成输入变量 $I^m = (x_1^m, x_2^m, \dots, x_n^m, x_{n+1}^m)$, 设定聚类数 c 和学习终止条件 ε 。

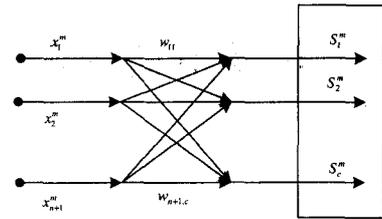


图 1 网络拓扑结构

Fig. 1 Structure of the clustering neural network

图 1 是一个两层前馈神经网络,第一层为输入层,含有 $n+1$ 个神经元 ($x_1^m, x_2^m, \dots, x_n^m, x_{n+1}^m$)。激励函数是线性函数,为 $F(x) = x$ 。第二层为竞争输出层,其神经元个数与聚类类别数 c 相对应。输出值为 $S_1^m, S_2^m, \dots, S_c^m$ 。 w 是第一层和第二层的连接权,形成 $(n+1) \times c$ 阶的矩阵 $w = [w_{ij}]_{(n+1) \times c}$, $i=1,2,\dots,n+1$, $j=1,2,\dots,c$ 。初始化权值矩阵 $w = [w_{ij}]_{(n+1) \times c}$ 。

1.2 模糊竞争学习聚类神经网络算法

步骤 1: 对每年历史数据 (即待分类样本) 分别进行计算。

$$\text{当 } \|w_j - I^m\| = \min_l \|w_l - I^m\|$$

$$(j \in 1, 2, \dots, c, l = 1, 2, \dots, c) \text{ 时,}$$

$$\text{令 } S_j^m = 1, S_l^m = 0, l \neq j.$$

网络中 S_j^m 的作用是判断获胜的类别,当 $S_j^m = 1$ 时,说明第 j 类获胜,输入样本 I^m 属于第 j 类。若 $S_j^m = 0$ 则表明输入样本 I^m 不属于第 j 类。即网络输出为二值函数, $S = 1$ 或 $S = 0$ 。

步骤 2: 用竞争学习规则调整获胜的 w_{ij} , 激励函数为 $F(x_i) = x_i$ 。

步骤 3: 将全部样本学习一遍作为一次循环,计算循环前后 w_{ij} 的变化。当该变化小于学习终止条件 ε 时,学习过程结束,否则转向步骤 1。

整个聚类过程是由这些学习过程反复进行完成的。最终将每一年样本归入到某一类中,形成聚类 U_1, U_2, \dots, U_c 。

2 基于模糊理论的竞争学习算法的改进

2.1 改进的竞争学习规则

竞争学习规则分为有监督和无监督两种,聚类分析问题一般应用无监督竞争学习规则。竞争学习规则的基本观点是只有竞争胜利者才能学习新知识和遗忘旧知识,竞争失败者保持原状态,既不学习新知识也不遗忘旧知识。一个最简单和常用的竞争

学习规则表示为:

$$\dot{w}_{ij} = S_j(F(x_i) - w_{ij}) \quad (2)$$

\dot{w}_{ij} 是连接输入层第 i 个神经元和输出层第 j 个神经元之间权值的修正值。 $S_j = 1$ 或 0 , 为二值函数。

其它学习规则, 如 Hebb 学习规则是改变所有的连接权, 因此基于 Hebb 规则的学习系统在学习新知识的同时要遗忘旧知识。Hebb 学习规则为

$$\dot{w}_{ij} = -w_{ij} + F(x_i)S_j \quad (3)$$

式(2)中, 当第 j 个神经元获胜时, $S_j = 1$, 其余的 $S_i = 0, i \neq j$ 。因此只有当第 j 个神经元获胜时, w_{ij} 才会改变, 遗忘项 $-w_{ij}$ 才会起作用。但在式(3)中遗忘项 $-w_{ij}$ 任何时候都起作用, 它在学习新模式时整个权值矩阵都改变。因此 Hebb 竞争学习算法的收敛速度是比较快的。Kosko 证明了它是以指数级的速度收敛的^[7]。

2.2 模糊竞争学习规则

1.2 节步骤 1 中得出的 S_j^m 是二值函数, 即每次只有一个单元获胜。本文从模糊理论出发, 引入了模糊竞争学习的概念。它的基本思想是竞争中胜负区分不再界限分明, S_j^m 不再是二值函数, 而是一个模糊隶属度函数, S_j^m 定义为

$$S_j^m = \left\{ \sum_{i=1}^c \left(\frac{d_j^m}{d_i^m} \right)^{\frac{2}{t-1}} \right\}^{-1}, \quad (4)$$

$$d_j^m = \|w_j - I^m\|, d_i^m = \|w_i - I^m\|$$

t 为加权指数, 根据经验取值为 5。

由此 1.2 节中步骤 1 改进为, 对每年历史数据 (即待分类样本) 分别进行计算。当 $\|w_j - I^m\| = \min_l \|w_l - I^m\|$ ($j \in 1, 2, \dots, c, l \in 1, 2, \dots, c$) 时,

$$\text{令 } S_j^m = \left\{ \sum_{i=1}^c \left(\frac{d_j^m}{d_i^m} \right)^{\frac{2}{t-1}} \right\}^{-1}, t=5 \quad (5)$$

$$S_l^m = 0, \quad l \neq j \quad (6)$$

说明第 j 类获胜, 输入样本 I^m 属于第 j 类。此时克服了 1.2 节步骤 2 中网络输出的二值性, 使得 S_j 取值范围扩大, 因此在 Hebb 竞争学习规则中,

不仅遗忘项 $-w_{ij}$ 起作用, S_j 的取值也使得权值矩阵改变的速度会更快, 因此竞争学习算法的收敛速度有了显著提高。

3 模糊竞争学习聚类神经网络的中长期负荷预测方法

由历史数据形成聚类 U_1, U_2, \dots, U_c , 其中 U_1 包括 m_1 年历史数据, U_2 包括 m_2 年历史数据, 依次类推, U_i 包括 m_i 年历史数据, U_c 包括 m_c 年历史数据, $m_1 + m_2 + \dots + m_c = M$ 。

将聚类 U_1, U_2, \dots, U_c 向环境因素轴 X 投影, 得到 V_1, V_2, \dots, V_c , 有

$$V_i = \{x^m: I^m \in U_i\}, i=1, 2, \dots, c \quad (7)$$

令

$$V_i = \{x_1^p, x_2^p, \dots, x_n^p, \dots, x_1^q, x_2^q, \dots, x_n^q, \dots\}, \quad (8)$$

$$p, q \in 1, 2, \dots, M$$

计算 V_i 的统计均值与方差, 即

$$\bar{X}_i = \frac{1}{k} (x_1^p + \dots + x_n^p + \dots + x_1^q + \dots + x_n^q + \dots) \quad (9)$$

$$k = m_i \times n$$

$$\bar{X}_{ij} = \frac{1}{m_i} (x_j^p + \dots + x_j^q + \dots), \quad (10)$$

$$i=1, 2, \dots, c, j=1, 2, \dots, n$$

$$\sigma_{ij}^2 = \frac{1}{k} \sum_{q=1}^k (x_q^m - \bar{X}_{ij})^2, \quad (11)$$

$$j=1, 2, \dots, n, k = m_i \times n$$

然后根据 \bar{X}_i 和 σ_{ij}^2 建立 V_i 上的正态模糊集:

$$\forall x^m = (x_1, \dots, x_n) \in X,$$

令

$$A_i(x) = \sum_{j=1}^n \omega_j \exp\left\{-\frac{1}{9\sigma_{ij}^2} (x_j^m - \bar{X}_{ij})^2\right\} \quad (12)$$

其中: $(\omega_1, \omega_2, \dots, \omega_j)$ 是一组取定的权重。

接着将 U_1, U_2, \dots, U_c 向预测轴 Y 投影, 得 W_1, W_2, \dots, W_c , 有

$$W_i = \{I^m: I^m \in U_i\}, i=1, 2, \dots, c \quad (13)$$

令

$$W_i = \{I^p, \dots, I^q, \dots\}, \quad (14)$$

$$p, q \in 1, 2, \dots, M$$

计算

$$\bar{y}_i = \frac{1}{m_i} (l^p + \dots + l^q + \dots) \quad (15)$$

构造以 $(\bar{y}_i, 3\sigma_{ij})$ 为参数的正态模糊数 r_i , 对应于分类 $\{U_1, U_2, \dots, U_c\}$, 建立起特征表

U_1	U_2	$\dots U_c$
A_1	A_2	$\dots A_c$
r_1	r_2	$\dots r_c$

设待预测第 s 年环境因素数据经过标准化后为 $x^s = (x_1^s, x_2^s, \dots, x_n^s) \in X$, 对 x^s 和 $\{A_1, \dots, A_c\}$ 应用最大隶属度原则选出 A_{i_0} , 根据特征表, 取相应的 r_{i_0} 作为第 s 年的负荷预测值, 经过反归一化过程得出最后预测值 l^s .

对未来年份的负荷进行预测时, 本文采用信息扩充法, 即把预测出的未来每一年的信息添加到原有负荷模型中, 使得模型随未来变化情况不断自适应更新。

设需要预测未来 p 年的负荷值, 根据历史数据所建立的最优聚类负荷模型先预测出未来第 1 年的负荷, 接着将该预测年份的信息添加到原有模型中进行信息扩充, 并通过式 (9) (10) (11) 对该年份数据所属类别的均值和方差进行修正, 然后采用新的负荷模型对未来第 2 年的负荷进行预测, 再将预测所得的结果按照预测第 1 年的处理方法对原有模型修正, 依此类推, 最终得出未来 p 年的预测结果。

4 算例及结果分析

以某地区负荷数据为对象进行实例分析, 以重工业产值、轻工业产值、农业产值及时间为相关环境因素变量, 将该地区 1980~2000 年的数据作为历史数据进行聚类分析, 得到聚类结果。输入数据见表 1。

表 1 输入数据

Tab.1 Input data

m	x_1^m	x_2^m	x_3^m	x_4^m	x_5^m
1	0.7236	0.2457	0.2135	0	0.1078
2	0.7365	0.2571	0.2267	0.1034	0.1136
3	0.7399	0.2684	0.2341	0.1135	0.1321
4	0.7485	0.2785	0.2533	0.1157	0.1758
5	0.7586	0.2846	0.2673	0.1201	0.1842
6	0.7641	0.2998	0.2743	0.1217	0.1911
7	0.7677	0.3107	0.2830	0.1219	0.2012
8	0.7802	0.3327	0.3010	0.1281	0.2137
9	0.8342	0.4082	0.3122	0.1352	0.2436
10	0.8452	0.4661	0.3751	0.1421	0.2577
11	0.8534	0.4962	0.3920	0.1578	0.2681

m	x_1^m	x_2^m	x_3^m	x_4^m	x_5^m
12	0.8619	0.5159	0.4083	0.1673	0.2995
13	0.8740	0.5212	0.4242	0.1753	0.3305
14	0.8791	0.5334	0.4966	0.1779	0.3854
15	0.8820	0.5658	0.5141	0.1662	0.4333
16	0.8943	0.5866	0.5237	0.1758	0.4385
17	0.9005	0.5962	0.5353	0.1788	0.4951
18	0.9203	0.6032	0.5655	0.1905	0.5169
19	0.9523	0.6145	0.5911	0.1911	0.5587
20	1	0.6506	0.6031	0.1936	0.5991

然后根据 2001~2005 年的环境因素数据来判定未来负荷变化属于哪种类别, 应用第三节的方法建立负荷模型进行预测, 得出对应年份的负荷值。表 2 还列出了其它几种传统算法的预测误差, 结果如下所示:

表 2 不同预测方法的相对误差 (%)

Tab.2 The relative errors of different methods (%)

年份	综合产值	时间序列	指数	ANN	本文方法
	单耗法	趋势法	平滑法		
2001	1.883	1.985	1.702	1.730	1.632
2002	1.905	2.060	1.889	1.776	1.731
2003	2.356	2.132	2.365	2.207	1.956
2004	2.967	2.855	2.569	2.635	2.139
2005	3.165	2.936	2.987	2.916	2.296

本文采用的模糊竞争学习聚类神经网络方法与其它四种传统预测方法相比, 弥补了它们在计入影响负荷变化因素方面的不足, 将负荷的预测与经济的发展紧密联系起来。由于采用了信息扩充法, 使得模型随未来的变化情况不断自适应更新, 从而使预测更为科学合理。由表 2 的结果可以看出每年预测结果的精度都有明显提高, 负荷值的预测误差基本上都在 2.3% 之内, 能达到实际需要, 适合于中长期负荷预测。

表 3 给出了传统聚类算法和本文采用方法在收敛速度和收敛精度上的比较, 可以看出, 本文算法在收敛速度上明显优于传统的聚类方法。

表 3 不同聚类方法的收敛精度与收敛速度比较结果
Tab.3 Comparison of convergence rate and precision using different clustering methods

方法	硬 C 均值聚类	模糊 C 均值聚类	本文方法
收敛精度	0.0001	0.0001	0.0001
收敛速度/s	0.5983	0.3687	0.1787

5 结论

本文采用模糊竞争学习聚类神经网络方法进行中长期负荷预测, 注重考虑实际环境因素, 将负荷

(下转第 104 页 continued on page 104)

的消化和扬弃。

3 目前国内外采用的母线保护方式对比

由于上述原因,目前采用的变压器后备过流保护方式不能达到速断要求。国内外还采用了微机控制的利用馈线过流闭锁进线速断保护方式,其动作时间为0.3~0.4s,这与短延时速断后备保护时间0.3或0.5s相当,并无多少优越。还是不能满足<0.1s的要求,还要增加线路设施资金投入。因此国外又推广利用电弧光与过流闭锁母线保护方式,并引进到国内大力推崇。虽然短路保护时间能控制在0.06~0.08s<0.1s,这与常规的电流瞬时速断保护时间不相上下,并没有快多少。还要增加大量弧光传感器检测控制系统装置的费用,说明国外也没有解决好这个问题。还有母线差动保护也是在速断保护有死区不能满足的情况下才使用的,是类同变压器内部故障的差动保护。可母线有很多路出线,要想检测总输出电流,它等于各出线回路电流之和,各电流互感器变比与误差都不同。差动保护的原理/计算/检测/整定/装置/接线/维护都很复杂,也很容易引起误动作,投资也大。整定动作电流也是避开尖峰涌流接近 $5 I_n$,保护动作时间也没多少优越。从设计理念讲,简单就是好。只要把目前的过流主保护改成瞬时速断主保护并按上下各

级线路额定电流 I_n 的5~10倍整定消除死区,以短延时速断为后备保护并按线路 I_n 的2~4倍整定避开尖峰涌流,按选择性的电流要求错开上下级的整定值,配合重合闸功能。问题就解决了,不用增加投资,简宜实用。可以相信,仅此一举简单的改动全国每年就可减少损失上亿元。其经济和社会效益是深远的。

以上浅见为引玉之砖,供电力部门有关院校和大企业专家参考借鉴指正。

参考文献

- [1] 车伟扬,等.中低压开关柜母线快速电弧光保护装置的应用[J].电世界,2006,(8).
- [2] ISBN-7-111-06209-4.刘介才/戴绍基.工厂供电(第2版)[M].北京:机械工业出版社,2006.
- [3] 苏文成.工厂供电(第2版)[M].北京:机械工业出版社,1999.
- [4] 华中工学院.电力系统继电保护原理与运行[M].北京:电力工业出版社,1981.
- [5] 卓乐友.电力工程电气设计手册[M].北京:水力电力出版社,1990.

收稿日期:2007-09-03; 修回日期:2007-10-16

作者简介:

王强(1956-),男,高级工程师,从事中压开关设备开发和工厂供电教学研究。E-mail:wangqiang@sdzz.net

(上接第58页 continued from page 58)

预测与经济发展紧密联系起来,通过寻求负荷变化与这些因素相联系的典型模式对负荷的未来走势进行预测,克服了单纯从数学模型角度进行负荷预测的局限性。由于聚类算法运算速度快,神经网络算法易于并行处理,所以使网络训练能快速达到预测精度的要求,并取得了较为理想的预测结果。并且本文采用的模糊竞争学习算法克服了网络输出结果二值性的缺陷,加快了算法的收敛速度。实例分析表明由该方法得到的负荷预测结果精度更高。

参考文献

- [1] 顾洁.电力系统中长期负荷的可变权综合预测模型[J].电力系统及其自动化学报,2003,15(6):56-60.
GU Jie. Study on the Varied Weight Synthesis Model of Mid-long Term Load Forecasting in Power System[J]. Proceedings of the EPSA, 2003, 15(6): 56-60.
- [2] 游仕洪,程浩忠,谢宏.应用模糊线性回归模型预测中长期电力负荷[J].电力自动化设备,2006,26(3):51-53.
YOU Shi-hong, CHENG Hao-zhong, XIE Hong. Mid-and long-term Load Forecast Based on Fuzzy Linear Regression Model[J]. Electric Power Automation Equipment, 2006, 26(3): 51-53.

- [3] 牛东晓,曹树华,赵磊,等.电力负荷预测技术及其应用[M].北京:中国电力出版社,1998.
- [4] 杨争林,唐国庆,宋燕敏,等.改进的基于聚类分析的超短期负荷预测方法[J].电力系统自动化,2005,29(24):83-86.
YANG Zheng-lin, TANG Guo-qing, SONG Yan-min, et al. Improved Cluster Analysis Based Ultra-short Term Load Forecasting Method[J]. Automation of Electric Power Systems, 2005, 29(24): 83-86.
- [5] 高新波.模糊聚类分析及其应用[M].西安:西安电子科技大学出版社,2004.
- [6] Pal N R, Bezdek J C, Tsao E C K. Generalized Clustering Networks and Kohonen's Self Organization Scheme[J]. IEEE Trans on NN, 1993, 4(4): 549-557.
- [7] Kosko B. Fuzzy Engineering[M]. New York: Prentice-Hall Engle-Wood Cliffs, 1996.

收稿日期:2007-09-18; 修回日期:2007-11-29

作者简介:

岳璐(1982-),女,硕士研究生,主要研究方向为电力系统负荷预测;E-mail:yuelu1982@hotmail.com

张尧(1948-),男,教授,博士生导师,主要研究方向为电力系统运行与稳定,电力市场,电网规划。