

文章编号:1007-6735(2011)01-0039-05

基于复杂系统的短期负荷预测研究及系统设计

孙大帅, 马立新, 王守征
(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: 运用复杂系统理论分析电力系统. 采用聚类方法对用电区域进行子系统划分, 并通过改进的神经网络算法和增加天气因素的预测方法进行短期负荷预测. 通过算例和电力系统应用, 证实了该算法的可行性, 较显著地提高了负荷预测的准确率.

关键词: 复杂系统; 短期负荷预测; 聚类分析; 神经网络

中图分类号: TM 715 **文献标志码:** A

Design of short-term load forecast systems based on the theory of complex systems

SUN Da-shuai, MA Li-xin, WANG Shou-zheng
(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Power network is a complex nonlinear system. The complex system analysis theory was applied to analyze the power network and divide it into some sub-systems according to its cluster analysis. Then back propagation neural network incorporated with weather influence factors was used to forecast the power network short-term load. Experimental results verify that the method is helpful to improve the forecasting accuracy.

Key words: complex systems; short-term load forecast; cluster analysis; neural network

电力系统负荷预测是电力生产管理的重要环节之一. 随着社会经济的发展, 对电力系统负荷预测的精度要求不断提高. 一系列新的智能算法的引入以及计算机技术的发展使得预测精度的提高成为可能. 目前, 主流的负荷预测方法包括回归法、时间序列法、灰色预测法^[1]、支持向量机^[2]及神经网络

等. 这些算法的主要目的是通过对已知负荷数据样本的分析, 提取其中的内在关系, 然后根据这个关系, 预测未来的负荷数据. 例如, 以 BP(back propagation)神经网络算法为例, $y_i = f(x_i)$, 其中, y_i 为要预测的值, x_i 为和预测相关的参数. BP 神经网络算法的目的就是通过分析 x_i 和 y_i 的历史数据, 从而对 f 这个非线性对应关系进行拟合, 然后

收稿日期: 2010-04-06
基金项目: 上海市研究生创新基金资助项目(JWC XSL0902)
作者简介: 孙大帅(1981-), 男, 硕士研究生. 研究方向: 电力系统现代控制技术. E-mail: gisasum@gmail.com
马立新(联系人), 男, 教授. 研究方向: 电力系统稳定性, 电力系统现代控制技术. E-mail: ma.eeepsi@usst.edu.cn

利用得到的 f 来进行预测. 由上述可知, 各种智能算法实际上是对数据样本分析的方法. 尽管有论文中谈到了样本的除错和样本数据的标么化处理, 进行了一些时域上的样本选择^[3], 然而从负荷数据的采集和存储的角度来选择、划分样本数据等并未涉及. 对某个区域的电力负荷进行预测, 现在主流的预测方法是把这个区域的用电负荷作为一个总量来进行分析和预测. 本文将预测区域作为一个复杂系统来对待, 先利用复杂系统的方法将其划分成若干个子系统, 然后分别对各个子系统进行分析和预测, 最后将各个子系统预测值之和作为整个系统的预测值. 对于整个用电系统来说, 分别对其子系统进行考查, 去除了各子系统之间的影响因素, 从而使得负荷预测的精度得到显著提高.

复杂系统是研究系统和复杂性的一门交叉科学, 目前并没有科学规范的统一定义. 对于一个系统来说, 之所以称之为复杂系统, 主要有以下几个定量的考察: 系统规模的复杂性及系统结构的复杂性、系统开放性环境的复杂性等^[4]. 而对应某个区域的电力系统来说显然其符合复杂系统定义的相关要求. 首先, 电力系统是一个巨大的系统, 其结构具有复杂性与时变性, 并且整个电力系统用电负荷受到开放环境的各种因素的影响; 其次, 作为整个用电系统的最根本因素是人, 人具有自主能动性和相互关联性. 从某种程度上来说, 人的能动性和社会性决定了电力系统的复杂性. 所以说负荷的预测是对以人为最终耗电单位的整个复杂系统行为的一种研究.

另外, 随着电网智能化的不断提高, 现在电网中不仅可以得到某个地区的总用电量, 而且可以得到各个分区域节点的负荷值, 这样就可以对各个分区的负荷值分区分类存储和调用, 为从复杂系统角度来分析负荷预测提供了前提条件.

1 预测系统的子系统划分

1.1 划分子系统的理论基础

复杂系统是由许多平行发生作用的子系统组成的网络, 其内部的控制力是相当分散的^[4]. 用电网络即是这样的一个系统, 虽然整个供电网络越来越趋向于智能化, 但是作为每个用电单位来说, 其用电时间和用电量仍具有相当的随机性. 因为, 用电的单位和个人具有自组织性, 所以, 整个用电系统可以理解为一个自组织的系统. 由自组织系统的特性可知自组织系统具有整体协调性. 这种整体协调性在用电

网络中就体现在整个用电网络某个时刻的总的用电量必定是以一个固定值上下浮动. 用电网络整体的协调性带来的是整个用电网络的稳定性, 但是, 这种整体的协调是各个子系统相互作用的结果. 从整个系统来看由于分系统的相互作用, 其各个系统的一些特性就相互交织在一起, 这对于预测算法提取其系统的特性是不利的. 而地区提取系统的各种非线性规则就是预测的重要前提. 所以, 通过对子系统的划分, 可以使得每个子系统之间的干扰减少, 这对于其规则的提取有利, 从而有助于其整个系统预测精度的提高.

1.2 用电系统子系统的划分方案

对于一个供电地区来说, 由于负荷数据采集系统的限制, 可根据用电系统的实际行政区域来进行分区负荷数据采集. 从实际的数据采集系统中可以得到实际各个区的用电数据, 然后分别存储和调用. 但是, 在实际的设计中会遇到以下问题:

a. 当区域划分过多时, 由于区域内用电用户个数的减少, 这个区域内的用电的随机性就会增强, 这样也就增加了对这个区域的预测负荷的难度.

b. 分区域过多时由于要对各个区域分别进行预测, 这样整个预测系统的计算量将大大增加, 预测的时间将会延长.

例如, 以本文设计为例, 共有 n 个分区域 $Z_1, Z_2, \dots, Z_n$, 如果分别对这 n 个区域进行预测, 预测的精度将会显著降低, 且耗时过长不可接受. 为此, 本文通过聚类分析的方法对 n 个区域的历史数据相关特征值进行分组.

本文设计采用系统聚类法(hierarchical clustering methods), 即层次聚类分析法, 这是目前国内外常用的一种方法. 其基本思想是先将参与聚类分析的样品各自作为一类, 然后规定样品之间的距离及类与类之间的距离, 再进行逐级合并, 直至所有的样品都成一类为止. 这里不需要给定分类的标准和类别数, 只是比较客观地从数据自身出发进行分类.

为了进行聚类分析, 首先要提取本文设计中的 n 个区域的特征值. 对于特征值的提取, 为了能最大限度地保留负荷样本数据的特征信息, 同时减少特征值的数量, 以做到减少不必要的计算量, 本文中取待预测日的前 10 d 数据样本, 首先根据样本数据日和待预测日的关联程度加权平均^[5], 取得对应分区的特征历史样本. 得到特征样本数据后, 根据其主成分分析^[6]来确定该分区的特征向量. 本设计中特征向量选为上述特征历史样本的最高负荷、最低负

荷、平均负荷、峰谷差、负荷率和最小负荷率。

通过上述分析可以得到一组聚类分析的样本 $\mathbf{X} = \{x_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, 6\}$, 即 n 组 6 维向量。不同的指标的原始数据具有不同的单位和量纲, 为了避免对分类结果造成影响, 在进行聚类分析之前, 首先对数据进行无量纲化处理。

$$x'_{ik} = \frac{x_{ik} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})}, \quad k = 1, 2, \dots, 6 \quad (1)$$

式中, x'_{ik} 为计算第 i 个样本中的各特征量的标准值。

在标准化之后, 特征值的最大值为 1, 最小值为 0, 其他值都在 0 和 1 之间。对所有的特征向量进行了标准化。

对上述 n 个区域的历史负荷中提取 n 个特征向量, 采用层次聚类方法, 选择欧氏距离来进行区域特征值之间的相似性计算。

$$D_{ij} = \sqrt{\frac{1}{n} \sum_{k=1}^n (x'_{ik} - x'_{jk})^2} \quad (2)$$

式中, D_{ij} 为两个分区的用电相似性系数; x'_{ik} 为 i 区域 k 指标的标准化指标值; x'_{jk} 为 j 区域 k 指标的标准化值; n 为研究分区域总数。

本文的聚类分析法中采用平均法 (between-groups linkage), 即当两个分区之间的用电特征向量之间距离平均值最小时, 就将这两个分区合并, 然后再逐级合并。本文设计中的聚类结果如图 1 所示, 为预测 2009 年 8 月 1 日, 而对其前 10 天数据的聚类分析结果。

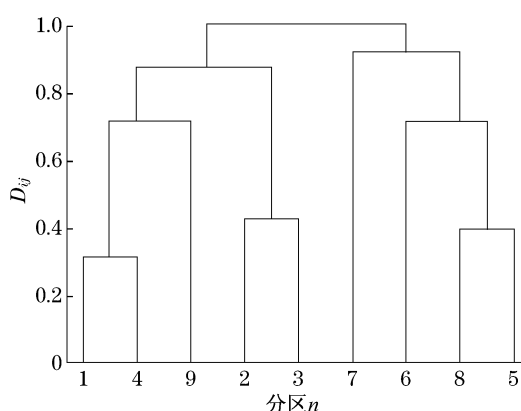


图1 聚类分析的结果

Fig.1 Result of clustering analysis

在分组的过程中主要从聚类分析的角度来考虑各分区域的特征值相似性, 通过设定聚类分析中距离的界限值可以得到特征值的分组。根据特征值的

分组情况, 可以得到负荷分区分组情况。过程如图 2 所示。

将负荷的分区分数据分成若干组后, 再将各组中的分区分负荷值相加作为单独的负荷预测系统来对待。此时即将一个复杂的电力系统分成几个相对简单的子系统。

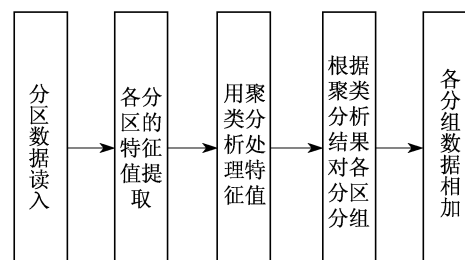


图2 子系统划分步骤框图

Fig.2 Illustration of sub-systems divided

2 预测方法的讨论

当整个地区分成若干子系统后, 将对每个系统的负荷数据分别进行预测。人工神经网络负荷预测方法是近十几年来研究和使用的较多的一种方法。由于神经网络具有并行分布信息、自学习及任意逼近连续函数的能力, 因而能捕获电力负荷的各种变化趋势。现在主流的用于电力系统负荷预测的神经网络主要是基于 BP 的神经元网络, 但这种算法网络训练费时, 并且易趋向局部最小点, 因而在实际应用特别是在线预测时受到一定限制。为此, 本文中采用的是一种高阶的 BP 神经网络 (HBP) 模型^[7]。

2.1 高阶 BP 神经网络介绍

BP 算法是用于前馈多层网络的学习算法, 前馈多层网络的结构如图 3 所示, w 为权值。

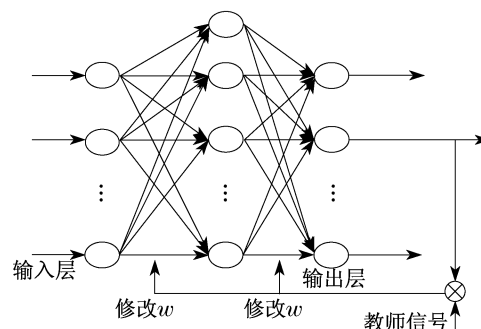


图3 BP神经网络结构图

Fig.3 Structure of the BP-ANN

BP 网络的训练方式采用反向误差传播学习过程, 是按照误差的大小由输出层逐步向输入层调整各神经元的阈值和连接权值的过程。

而本文中的高阶 BP 网络(HBP)主要进行如下改造:由图 4 可以看出,输出层的输入多了一部分,即倒数第二层任意两个神经元的输出相乘再加权到输出层,其余部分和普通 BP 算法相同.在图 4 中,输入层 i 有 n 个节点,隐含层 j 有 m 个节点,输出层 k 有 t 个节点. X 为网络输入, O 为网络输出.对于输出层的每个节点,其输入值和原 BP 网络相比则增加 C_m^2 个输入.在增加这些节点后,前向计算中隐层的输出 b_j 仍为

$$b_j = f\left(\sum_{i=1}^n w_{ij} x_i - \theta_j\right), \quad j = 1, 2, \dots, m \quad (3)$$

式中, w_{ij} 为输入层到隐层的权值; θ_j 为隐层的阈值; f 为激活函数.

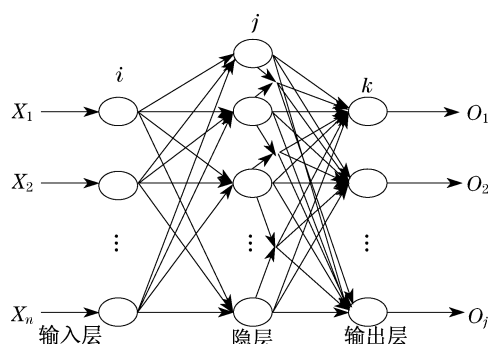


图 4 改进节点结构图

Fig.4 Structure of the changed node

上述计算过程和普通 BP 算法相同,而前向计算中的输出层输出则改为

$$c_k = f\left(\sum_{i=1}^n w_{jk} b_j + x_k - \theta_k\right) \quad (4)$$

$$x_k = \sum_{i=1}^n \sum_{j \in (0, m), i \neq j} b_j b_i w_{ijk} \quad (5)$$

式中, x_k 即为改进 BP 后的新增节点输入值.

而在反向调整的过程中,则有

$$w_{ij}[n+1] = w_{ij}[n] - \beta \frac{\partial E}{\partial w_{ij}} + \alpha(w_{ij}[n] - w_{ij}[n-1]) \quad (6)$$

$$w_{ijk}[n+1] = w_{ijk}[n] - \beta \frac{\partial E}{\partial w_{ijk}} + \alpha(w_{ijk}[n] - w_{ijk}[n-1]) \quad (7)$$

式中, α, β 为调节的步进幅度; E 为 BP 算法中输出层各单元的校正误差值.

2.2 高阶 BP 神经网络用于各分系统的预测建模

采用高阶 BP 神经网络进行短期负荷预测时,主要是利用 BP 神经元网络提取负荷数据在时序上的对应关系.要根据关联性强的原则合理

组织训练样本,即样本的数据所体现出来的规律要和预测的数据有较强的关联性和相似性.本文设计中选择待预测日之前的一段连续时间内的历史数据作为训练样本^[8],同时还要适当地加上待预测日相似日的负荷数据作为训练样本^[9].用上述的样本数据对 HBP 神经网络进行训练.随着用电结构的变化,人们的取暖和降温用电所占的比例不断提高.将天气相关因素从历史负荷中剥离出来,作为独立的输入变量,从而能够给出更加精确的预测值^[8].

3 系统的设计与实现

本文设计中以中原地区某市的负荷为研究样本,通过该市供电局的负荷信息采集系统来分别读取各个区域的负荷数据并进行分区存储.其中,数据库中同时存储与负荷相关的天气信息.然后通过分析各个区域的负荷数据之间的联系将这些分区域分成若干个负荷特性相类似的子系统,这样可以将整个系统的用电数据化分成若干个子系统的用电数据.

整个系统的数据处理流程如图 5 所示,由两部分组成.一部分负责数据的采集和分类存储,此部分和供电部门的信息中心服务器相链接;另一部分为负荷的数据分析和负荷预测部分,其所用的数据均为第一部分所存储的数据.两个部分之间通过数据库交换数据,通过这种负荷预测系统和信息采集系统的隔离,实现了整个预测系统的独立运行,进而保证整个信息网络的安全性.

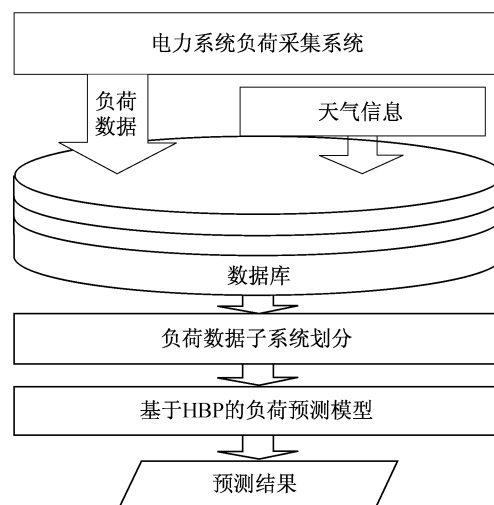


图 5 预测系统数据处理结构图

Fig.5 Data structure of the forecasting system

4 算 例

在实际的负荷预测中,每年夏季由于天气因素变化频繁,用电负荷变化较为复杂,也同时提高了负荷预测的难度.现采用的上述地区 2009 年 8 月的用电负荷为例进行预测验证,模拟真实预测情况,只运用待预测日之前的相关数据.现分别给出分组改进预测方法(G-HBP 法)、分组普通 BP 预测方法(G-BP)、不分组改进 HBP 预测方法和不分组普通 BP 预测方法的验证预测准确率 A.

$$A = \left(1 - \sqrt{\frac{1}{n} \sum_{i=1}^n E_i^2}\right) \times 100\% \tag{8}$$

式中,n 为预测天的预测总点数,本次算例中取为 24 个点;E_i为某一点的相对误差.预测准确率如表 1 所示.

表 1 4 种方法的准确率比较

Tab.1 Correct rate comparison of 4 methods

预测方法	G-HBP	G-BP	HBP	BP
准确率	96.7%	93.9%	92.1%	89.7%

注:均为 2009 年 8 月的月平均准确率

上述准确率计算中需要指出的是:基于 BP 的负荷预测模型相对于改进的 HBP 预测模型,前者在稳定性上落后于后者.在负荷预测中,普通的 BP 预测模型经常出现误差极大的情况,这直接导致了这种预测方法的月平均准确率落后于改进的 HBP 预测模型.

5 结 论

由上述算例可以看出,本文设计从复杂系统的角度来分析电力系统短期负荷预测模型,通过划分用电区域分别进行预测,从而提高了整体的预测精度.同样地,对于其他月份来说,应用本设计中的方法均能在不同程度上提高负荷的预测精度.根据上

述地区的 2009 年的 10 月、11 月和 12 月的时间预测结果来看,其预测的月平均准确率分别为97.4%,97.3%和 96.9%,相对于原算法的预测准确率均有较大程度的提高.

本文提出的预测系统及算法已经应用于电力系统.相对于传统的预测方法,本方法能在整体上提高 1.0%~2.5%的短期负荷预测精度.

参考文献:

[1] TZAFESTAS Spyros, TZAFESTAS Elpida. Computational intelligence techniques for short-term electric load forecasting [J]. Journal of Intelligent and Robotic Systems, 2001,31,(1/2/3):7-68.

[2] 靳忠伟,陈康民,闫伟,等. 基于支持向量机的中长期电力负荷预测研究与应用[J]. 上海理工大学学报, 2008,30(2):129-132.

[3] 康重庆,程旭,夏清,等. 一种规范化的处理相关因素的短期负荷预测新策略[J]. 电力系统自动化. 1999, 23(18),32-35.

[4] 王安麟,马立新,赵群飞. 复杂系统的分析与建模[M]. 上海:上海交通大学出版社,2004.

[5] 清华大学电机系. 电力市场需求预测理论及其应用[M]. 北京:清华大学,1997.

[6] 童其慧. 主成分分析法在指标综合评价中的应用[J]. 北京理工大学学报(社会科学版),2002,4(1):59-61.

[7] MA Lixin,SHIGEI N. Learning characteristics of multi-layered higher order neural networks[C]//International Conference of Computational Methods in Sciences and Engineering,2005.

[8] 张震宇,吕育青,蒋锋. BP 神经网络在计及气温因素的短期日负荷预测中的应用 [J]. 浙江电力,2007,26(1):6-9.

[9] 莫维人,张伯明,孙宏斌,等. 短期负荷预测中选择相似日的探讨 [J]. 清华大学学报(自然科学版),2004,44(1):106-109.

(责任编辑:石 瑛)