

# 基于 BP 神经网络的智能电表数据分析预测

Analysis and prediction of smart grid data based on BP neural network

郑建柏<sup>1</sup> 王芳<sup>2</sup>

(1 福建省电力有限公司厦门电业局, 福建厦门 361000; 2 福建省电力有限公司检修分公司, 福建厦门 361000)

**[摘要]** 针对智能电表高频率快速采集的实时数据, 采用 BP 神经网络的方法对用户的用电行为进行分析预测。根据预测结果制定分时电价, 将用户的用电行为从用电高峰向低谷时段引导, 平衡用电高峰和低谷的波动幅度, 减少最大用电负荷, 提高电网的运行效率。

**[关键词]** 智能电表; BP 神经网络; 用电行为分析; 分时电价

**[中图分类号]** TM933.4

**[文献标志码]** A

## 引言

智能电表是智能电网数据采集的基本设备之一, 主要作用为采集、计量和传输原始电能数据, 以实现信息集成、分析优化和信息展现, 主要包括高级量测体系(AMI)和自动抄表系统(AMR)两大核心模块<sup>[1]</sup>。国内对智能电表的定义是以微处理器为核心的, 可存储测量信息并能对测量结果进行实时分析、综合和做出各种判断的仪器。智能电表具有自动测量功能, 强大的数据处理能力, 人机交互功能和一定的人工智能<sup>[2]</sup>。另外, 近年来载波、GPRS、ZigBee 等通信功能和用户计量功能等特征也被引入到智能电表的概念中<sup>[2, 3, 4]</sup>。

从应用角度看, 智能电表一方面能够将用电信息实时反馈给用户, 便于用户管理自身的用电量, 达到节能减排的目的; 另一方面, 智能电表每隔 15 分钟即可采集一次用户用电信息, 可以对用户历史用电数据进行挖掘分析, 根据用户在一天之中各个时段的用电信息动态调整电价, 对在不同时段的用电量进行控制, 达到削峰填谷的效果。本文主要研究后者阐述的问题, 利用智能电表强大的数据采集、数据分析和处理能力, 在智能电表中引入 BP 神经网络模型对用户在未来各时段的用电趋势进行预测, 通过预测结果及时调整电价, 以节约电网的运行成本, 减少最大用电负荷并提高电网的运行效率。此外, 从电能计费的

角度出发, 本文中讨论的用户用电量均指智能电表计量的分时有功电能。

## 1 BP 神经网络原理

### 1.1 多层前馈神经网络结构

BP 神经网络是在多层前馈神经网络的基础上进行学习的。多层前馈神经网络的主要特点是, 每个节点的输出只能作为下一节点的输入, 而不允许向其上一层节点进行反馈<sup>[5]</sup>。图 1 是多层前馈神经网络的示意图。

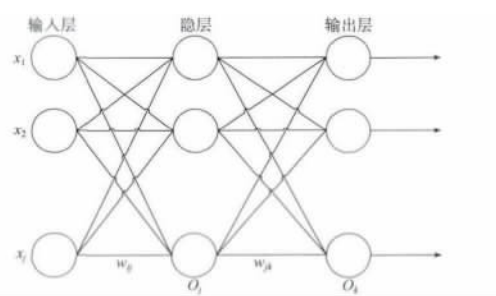


图 1 多层前馈神经网络示意图

多层前馈神经网络包括输入层、隐层和输出层三部分。输入层由一系列代表训练样本的非类别标识属性组成, 输入层的输出形成神经网络的隐层(理论上神经网络的隐层最多两层即可, 实际应用中, 通常只需要一个隐层), 输出层由训练样本的已知类别标识属性组成, 表示对给定样本的网络预测<sup>[5]</sup>。

使用神经网络进行用电量预测的前提是对神

经网络的拓扑结构进行定义, 主要包含两个方面: 首先是确定网络拓扑图, 即确定输入层中输入节点的数量、隐层的数量、各个隐层中节点的数量以及输出层中的单元数量; 其次是对输入节点的属性值进行标准化, 将输入属性值标准化到  $[0, 1]$  区间进行取值。

神经网络拓扑结构的定义是需要不断探索的过程, 结构的好坏直接影响神经网络学习的精确度。实际应用中, 需要用输入数据对不同的网络结构进行反复训练, 寻找最合适的拓扑结构。

## 1.2 BP 算法的流程

BP 算法是使用最广泛的神经网络学习方式, 主要由模式顺传播和误差逆传播两个过程组成<sup>[5]</sup>。模式顺传播的过程指将一组训练样本作为输入信息由输入层经中间层向输出层的传播过程; 误差逆传播的过程指网络的期望输出与实际输出之间的误差信号由输出层经中间层向输入层逐层修正连接权值的过程。重复模式顺传播和误差逆传播的过程, 直到权重最终收敛或达到稳态, 即输出的预测类标识与实际样本的类标识之间的均方差达到指定的误差范围内时, 完成神经网络的训练。

BP 神经网络的具体训练流程包含 4 个步骤。

1) 权重初始化。初始化权重往往在一个较小的范围内随机生成, 比如  $-0.5 \sim 0.5$ 。

2) 模式顺传播。计算输入层、隐含层和输出层的每个单元的输入输出情况。对于输入层而言, 其输入值等于输出值。即对输入节点  $i$ , 有  $O_i = I_i$ 。对于隐层和输出层中的节点, 每个单元的输入为上一层的输出与相应权值的复合, 即对于节点  $i$  的下一层节点  $j$ , 有公式:

$$I_j = \sum_i \omega_{ij} O_i + \theta_j$$

$I_j$  为下一层节点  $j$  的输入,  $O_i$  为上一层节点  $i$  的输出,  $\omega_{ij}$  为上一层节点  $i$  到下一层节点  $j$  的权值,  $\theta_j$  为节点  $j$  的阈值。

3) 误差逆传播。计算输出层误差的公式为:

$$Err_j = O_j(1 - O_j)(T_j - O_j)$$

其中,  $O_j$  是  $j$  单元的实际输出,  $T_j$  是基于给定训练样本类别标识的取值。隐层中节点  $i$  到输出层节点  $j$  的误差计算公式为:

$$Err_j = O_j(1 - O_j) \sum_i Err_i \omega_{ij}$$

权值的修正公式为:  $\Delta \omega_{ij} = (l) Err_j O_i$

新的权值更新为:  $\omega_{ij} = \omega_{ij} + \Delta \omega_{ij}$

阈值的修正公式为:  $\Delta \theta_j = (l) Err_j$

新的阈值更新为:  $\theta_j = \theta_j + \Delta \theta_j$

其中  $l$  表示学习率, 有  $0 < l < 1$ , 其作用是避免 BP 网络陷入局部最优解。

4) BP 算法的终止条件。BP 算法终止主要有三种情况, 首先是前一阶段的权重变化值  $\Delta \omega_{ij}$  已足够小, 低于设定的阈值; 其次是前一阶段的所有样本中出现错误分类的比率低于设定的阈值; 最后是神经网络的训练次数超过了设定的周期数。

## 1.3 神经网络预测原理

构建神经网络的过程实质上是通过神经网络记录神经元之间关联的权值。由于初始权值是在某个范围内随机生成的, 需要借助智能电表在过去各时段采集的历史数据对神经网络进行训练, 不断调整权值, 使神经网络的输出值与实际值的偏差缩小到某个可接受的范围内。经过训练的神经网络提取出了蕴藏于样本中的非线性映射关系, 并以权值分布的形式存储起来<sup>[6]</sup>。在预测阶段, 以新的时间段作为输入数据, 便能完成从输入的预测变量(时间)到输出的预测对象(用电量)的非线性映射, 这种通过神经网络将时间序列的历史数据映射到未来数据的预测方法称为趋势预测。

## 2 策略知识库与系统运行分析

智能电表每隔 15 分钟即可快速采集用户的用电数据, 表 1 给出了某个区域在某一天的用电量数据。图 2 直观地给出了用电量的变化情况, 可以清晰地看出用户每天用电的高峰时段和低谷时段。搜集连续 20 天每个时段的历史数据, 用 BP 神经网络对第 21 天各时段的用电量分别进行预测。

实验中采用单隐层 BP 神经网络, 隐层节点数为 10 个。采用 Matlab 进行实验, 分别用过去连续 20 天中某特定时间点的用电量训练网络, 完成对第 21 天各时刻的用电量预测。以 0:00 时间点为例, 建立连续 20 天中 0:00 时刻的用电量数据矩阵:

$$P = \begin{bmatrix} 8.36 & 8.066 & 8.255 & 7.841 & 7.791 \\ 8.031 & 7.763 & 8.006 & 7.722 & 7.652 & 8.342 \\ 8.046 & 7.776 & 8.264 & 7.642 & 8.35 & 8.058 \\ 7.862 & 8.331 & 7.765 \end{bmatrix}$$

表 1 某区域一天之中的用电量数据

时间	用电量	时间	用电量	时间	用电量	时间	用电量
0:00	8.360	6:00	15.761	12:00	19.053	18:00	17.606
0:15	8.171	6:15	15.532	12:15	18.581	18:15	17.460
0:30	7.930	6:30	15.363	12:30	18.809	18:30	17.691
0:45	8.118	6:45	15.157	12:45	18.649	18:45	17.524
1:00	8.450	7:00	15.886	13:00	18.780	19:00	16.850
1:15	8.396	7:15	16.125	13:15	18.900	19:15	17.336
1:30	8.994	7:30	16.537	13:30	18.758	19:30	16.980
1:45	8.915	7:45	16.528	13:45	18.901	19:45	17.190
2:00	10.186	8:00	18.947	14:00	16.861	20:00	16.004
2:15	10.357	8:15	19.075	14:15	17.001	20:15	16.375
2:30	10.616	8:30	19.026	14:30	16.826	20:30	16.432
2:45	10.855	8:45	19.066	14:45	17.209	20:45	16.337
3:00	12.098	9:00	17.712	15:00	17.070	21:00	14.643
3:15	12.337	9:15	17.776	15:15	16.784	21:15	15.089
3:30	11.775	9:30	17.392	15:30	17.278	21:30	14.763
3:45	12.059	9:45	17.810	15:45	16.640	21:45	15.354
4:00	12.700	10:00	17.095	16:00	17.220	22:00	13.489
4:15	12.369	10:15	17.348	16:15	16.808	22:15	13.140
4:30	12.174	10:30	17.104	16:30	17.347	22:30	13.412
4:45	12.863	10:45	16.835	16:45	16.897	22:45	13.139
5:00	13.710	11:00	16.512	17:00	20.833	23:00	11.136
5:15	14.054	11:15	16.434	17:15	20.320	23:15	11.400
5:30	13.384	11:30	15.882	17:30	20.458	23:30	11.678

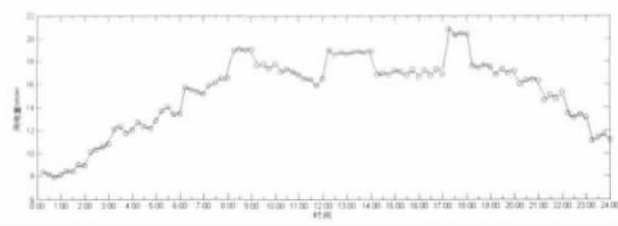


图 2 用户用电量变化图

将最大训练次数设为 1 000，误差精度设为 0.005，训练网络如下：

```
net = newff ( [0 19], [10 1], { 'tansig', 'purelin' } );
net.trainParam.epochs = 1000;
net.trainParam.goal = 0.005;
net = train ( net, X, P );
```

训练网络完成后预测第 21 天 0: 00 时刻的用电量结果为 8.015MWH。

运用同样的方法完成对第 21 天各时段用电量的预测。预测结果见图 3。

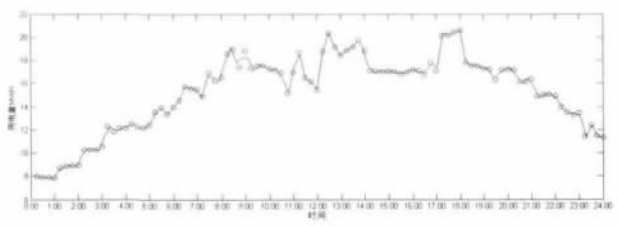


图 3 根据前 20 天的历史用电数据预测的第 21 天用电情况

从预测的用电量结果中可以看出，预测结果大致符合用户历史用电情况的变化规律。用电的高峰时段依然集中在 12: 00 ~ 14: 00 和 17: 00 ~ 18: 00 之间，可以适当抬高这些时段的电价，将用户或企业的用电行为向用电低谷时段引导。直到用电量曲线趋于平稳，使得分时电价趋于稳定。

3 总结

利用智能电表采集数据的实时性和高效性，对其产生的数据进行分析挖掘，可以预测用户的用电行为。可以将预测模块集成到智能电表，对采集的数据进行即时预测，根据预测结果对用电高峰期制定较高的电价，对用电低谷期制定较低的电价，通过分时电价对用户的用电行为进行控制，避免在一天之中的某一时期系统的负载量过大。本文利用 BP 神经网络对用户在一天之中各个时段的用电量进行预测，预测结果与历史数据所反映出的用户用电行为大致相同，具有较高的可信度，可以以此作为制定分时电价的标准。

参考文献：

[1] 王思彤，周晖，袁瑞铭，易忠林. 智能电表的概念及应用 [J]. 电网技术，2010. 34 (4)： 17 ~ 23

[2] 李会容. 基于 DSP 的智能电表的研究 [D]. 成都：电子科技大学，2007

[3] 李剑. 单相智能复费率电能表的设计与开发 [D]. 上海：同济大学，2010

[4] 李大鹏，王祁. 三相复费率网络化智能电表的研制 [J]. 仪器仪表学报，2006. 37 (6)： 1811 ~ 1812

[5] M. 巴斯蒂安. 数据仓库与数据挖掘 [M]. 北京：冶金工业出版社，2003. 271 ~ 275

[6] 张恩亮. 基于 BP 神经网络的道路交通安全预测研究 [D]. 北京：北京交通大学，2006