

D-S 证据理论融合改进 BP 网络的短期负荷预测研究

孙奇, 杨伟

(南京理工大学动力工程学院, 江苏 南京 210094)

摘要: 针对传统 BP 网络收敛速度较慢且实时性较差的缺点, 在考虑了气候因素的情况下, 分别用三种 BP 网络算法对某地区的负荷进行预测, 通过对预测误差的分析, 用证据理论的 Dempster 合成法则对算法进行融合, 通过选取待预测日之前几天的数据作为融合样本, 规定相应的基本信度函数, 得到融合后的信度分配, 从而决定相应时刻的预测模型, 仿真结果表明, 经过证据理论融合后选择的负荷预测算法具有更高的预测精度。

关键词: 证据理论; Dempster 合成法则; 负荷预测; 改进 BP 网络

Research on load forecasting of fusion between D-S evidential theory improved BP network

SUN Qi, YANG Wei

(College of Power Engineering, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract: Aiming at the shortcomings of the traditional BP network, three kinds of changed BP network calculating ways are used in load forecasting at an area, according to the situation of weather factor. By analyzing the error of load forecasting, calculating ways have fusion through Dempster fusion rule of evidential theories. Through choosing the data of several days before estimated day as fusion samples and ruling the homologous basic belief degree function, belief degree allotment is fetched after the fusion. The load forecasting model is decided according to the corresponding time. The results of simulation demonstrate that the load forecasting calculating ways after fusion of evidential theories have higher accuracy.

Key words: evidential theory; Dempster synthesize rule; load forecasting; improved BP network

中图分类号: TM715

文献标识码: A

文章编号: 1003-4897(2007)07-0061-05

0 引言

负荷预测是电力系统领域的一项重要工作, 随着负荷预测工作的不断开展, 各种预测模型也层出不穷, 许多专家提出了神经网络预测、灰色预测、组合预测等智能化预测方法。其中 BP 网络是一种有效的预测手段, 但是其收敛速度较慢且控制的实时性较差。针对其缺点, 许多专家对其做出了相应的改进, 也取得了良好的预测效果^[1~3]。

然而, 负荷特性随着预测时刻的变化而不断改变, 单一的预测模型不可能始终保持较高的预测精度, 如何随时间变化选择精度更高的预测模型, 是迫切需要解决的问题。本文针对三种 BP 算法, 通过证据理论的 Dempster 合成法则对待预测日之前的部分预测数据进行融合, 从而根据融合后的误差情况决定选择何种算法, 以减小预测结果的误差^[4, 5]。结果表明, 融合后的预测模型能达到更高的预测精度。

1 D-S 证据理论的相关概念

1.1 基本定义

证据理论是目前信息融合中最常用的一种方法, 它建立在集合论的基础之上, 主要用来解决不确定性问题, 下面首先给出证据理论的几个相关定义^[6, 7]。

定义 1: 对于一个集合 A , 我们称集合 $B = \{C | C \subseteq A\}$ 是集合 A 的幂集, 表示为 2^A 。对于一个判决问题, 用集合 Θ 表示我们所能认识到的所有可能结果, 该集合 Θ 就称为识别框架。通常, 我们感兴趣的命题都对应着 Θ 的一个子集, 由定义 1 可知, Θ 的所有子集构成的集合称为 Θ 的幂, 表示为 2^Θ 。

定义 2: 设 Θ 为识别框架, 如果集函数 $m: 2^\Theta \mapsto [0, 1]$ (其中 2^Θ 为 Θ 的幂集) 满足以下两个条件: a) $m(\phi) = 0$, b) $\sum_{A \subseteq \Theta} m(A) = 1$, 则

称 m 为框架 Θ 上的基本信度分配；如果 $\forall A \subseteq \Theta, m(A)$ 称为 A 的基本信度值；若 $m(A) > 0$ ，则称所有满足该条件的集合 A 为 m 的焦点元素，并将与基本信度分配 m 相对应的信度函数记为 Bel 。

1.2 Dempster 合成法则

对于同一识别框架 Θ 上基于不同的证据所得到的信度函数 $Bel_i (i=1,2,\dots,n)$ ，若该 n 批证据不完全冲突，则可用 Dempster 合成法则计算出一个新的信度函数 Bel ，这个信度函数可以作为在这 n 批证据联合作用下产生的信度函数。 Bel 称为合成前信度函数 $Bel_i (i=1,2,\dots,n)$ 的直和，用符号表示为如下形式： $Bel = Bel_1 \oplus Bel_2 \oplus \dots \oplus Bel_n$ 。

假设 Bel_1 和 Bel_2 是同一识别框架 Θ 上的基于两个彼此独立证据的信度函数， m_1 和 m_2 分别是其对应的基本信度分配，设 m_1 对应的焦点元素为 A_1, A_2, \dots, A_p ， m_2 对应的焦点元素为 B_1, B_2, \dots, B_q ，且有 $\sum_{A_i \cap B_j = \phi} m_1(A_i)m_2(B_j) < 1$ ，其中 $i=1,2,\dots,p, j=1,2,\dots,q$ ，则合成的基本信度分配函数 $m: 2^\Theta \mapsto [0,1]$ 可以定义如下：

$$m(A) = \begin{cases} 0, A = \phi \\ \frac{\sum_{A_i \cap B_j = A} m_1(A_i)m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i)m_2(B_j)}, A \neq \phi \end{cases}$$

以上是两个信度函数 Bel_1 和 Bel_2 的 Dempster 合成法则，记为 $Bel_1 \oplus Bel_2$ 。若待合成的信度函数多于两个，则可以用同样的算法将前一步的合成结果与下一个信度函数进行合成，直到所有信度函数合成完毕。

2 网络模型和算法

2.1 惯性修正法

为了使收敛速度更快，可在加权系数修正公式中增加一个惯性项，使得加权系数变化得更加平稳。以典型的三层式网络为例，输出层的任意神经元 k 在样本 p 作用时的加权系数改进公式为：

$$w_{ki}(k+1) = w_{ki}(k) + \eta \delta_k^p o_i^p + \alpha [w_{ki}(k) - w_{ki}(k-1)]$$

隐含层的任意神经元 i 在样本 p 作用时的加权系数改进公式为：

$$w_{ij}(k+1) = w_{ij}(k) + \eta \delta_i^p o_j^p + \alpha [w_{ij}(k) - w_{ij}(k-1)]$$

式中： α 为惯性系数， $0 < \alpha < 1$ ， η 为学习速率， o_i^p 为隐含节点 i 在样本 p 作用时的输出， o_j^p 为输入节点 j 在样本 p 作用时的输出， w_{ki} 表示输出层节点 k 与隐含层节点 i 之间的权值， w_{ij} 表示隐含层节点 i 与输入层节点 j 之间的权值。

2.2 步长变化法

传统 BP 网络的一阶梯度寻优算法收敛速度较慢，其重要原因是学习速率 η 较难选择。如果 η 选的太小，则收敛太慢；如果太大，则可能修正过头，导致振荡甚至发散，为了解决这个问题，可以将步长设置为变化量，根据收敛情况不断改变，具体算法如下：

$$w(k+1) = w(k) + \eta(k)D(k)$$

其中： $\eta(k) = 2^\lambda \eta(k-1)$ ， $\lambda = \text{sgn}[D(k)D(k-1)]$ ，且 $D(k) = -\partial J / \partial w(k)$ 为 k 时刻的负梯度， J 为每次训练的总误差函数。以上公式说明，当连续两次迭代的梯度方向相同时，表明下降太慢，这时可使步长加倍；当连续两次迭代的梯度方向相反时，表明下降过头，这时可以使步长减半。

2.3 尺度变化法

尺度变化法建立在导数的基础之上，试图通过迭代次数逐渐建立对海森矩阵的近似，它具有二阶梯度算法收敛较快的特点，但又不用直接计算二阶梯度，节省了大量计算时间，具体算法如下

$$M_{k+1} = M_k + \left(1 + \frac{\Delta g_k^T M_k \Delta g_k}{\Delta X_k^T \Delta g_k}\right) \frac{\Delta X_k \Delta X_k^T}{\Delta X_k^T \Delta g_k} - \frac{\Delta X_k \Delta g_k^T M_k + \Delta M_k \Delta g_k \Delta X_k^T}{\Delta X_k^T \Delta g_k}$$

$$X_{k+1} = X_k + H^{-1} g_k$$

式中： M 是海森矩阵 H 的近似值， X 为网络的权重和阈值， g 为梯度， k 为训练次数。

3 三种算法用于负荷预测

下面用 BP 网络对某地区的负荷进行预测。在待预测日的前一天，每间隔两小时对电力负荷进行一次测量，共得到 12 组数据，另外，还将当天最高气温、最低气温、天气（晴天为 0，阴天为 0.5，雨天为 1）三个典型的气候因素作为神经网络的三个输入，这样确定了网络共有 15 个输入神经元，网络的输出神经元共有 12 个，为待预测日每隔两小时的负荷值，隐含层的神经元数为 31 个。预测前首先用公式 (1) 将数值归一化，处理成 $[0, 1]$ 上的数据。

$$\hat{x} = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (1)$$

其中: x 为当天负荷的实际值, x_{\max} 和 x_{\min} 为当天负荷最大和最小值。归一化的结果如表 1 所示, 表 1 中最后三列为气候因子各自的影响系数。

结合表 1 中数据的特点, 可以看出, 由于每天同一时刻的负荷值变化较小, 另外考虑到网络的训练速度, 因此选用较少的训练样本进行网络训练。分别将 1~10 日的电力负荷与 2~11 日的气候因素作为 10 组训练样本(每组 15 个数据), 对网络进行

训练。当然, 若每天各时刻的负荷差别较大, 为了保证较高的训练精度, 也可以选用更多组数据进行网络训练。训练的收敛结果如图 1 所示。

训练完毕后, 将 11 日的电力负荷与 12 日的气候因素作为检测样本, 预测 12 日的电力负荷情况。并分别将三种算法的预测结果与真实负荷值进行比较。三种算法在各时刻的绝对误差和平均绝对误差如表 2 所示。

表 1 原始数据归一化结果

Tab.1 Return result of original data

日期	各时刻的电力负荷												气候因子		
	0 点	2 点	4 点	6 点	8 点	10 点	12 点	14 点	16 点	18 点	20 点	22 点	高温	低温	天气
1	0.25	0.15	0.13	0.22	0.55	0.66	0.70	0.69	0.68	0.69	0.75	0.82			
2	0.22	0.16	0.14	0.23	0.51	0.53	0.68	0.71	0.73	0.68	0.78	0.83	0.242	0.303	0
3	0.25	0.16	0.15	0.24	0.55	0.56	0.70	0.73	0.75	0.70	0.81	0.82	0.239	0.313	0
4	0.20	0.11	0.12	0.20	0.50	0.52	0.68	0.70	0.70	0.68	0.78	0.79	0.222	0.270	1
5	0.21	0.12	0.13	0.20	0.55	0.57	0.70	0.70	0.72	0.70	0.80	0.81	0.235	0.251	0.5
6	0.23	0.13	0.15	0.22	0.56	0.58	0.72	0.73	0.74	0.75	0.81	0.82	0.254	0.313	0
7	0.24	0.14	0.17	0.22	0.58	0.60	0.71	0.71	0.73	0.72	0.81	0.82	0.260	0.320	0
8	0.23	0.14	0.16	0.21	0.57	0.58	0.71	0.71	0.71	0.71	0.80	0.81	0.258	0.301	0
9	0.21	0.12	0.13	0.18	0.50	0.53	0.69	0.69	0.70	0.73	0.77	0.80	0.230	0.287	0.5
10	0.20	0.11	0.12	0.16	0.46	0.50	0.66	0.67	0.68	0.70	0.75	0.78	0.223	0.280	1
11	0.21	0.13	0.13	0.21	0.56	0.57	0.71	0.71	0.72	0.74	0.80	0.81	0.231	0.298	0
12	0.21	0.12	0.16	0.22	0.62	0.62	0.72	0.72	0.72	0.73	0.82	0.82	0.232	0.294	0

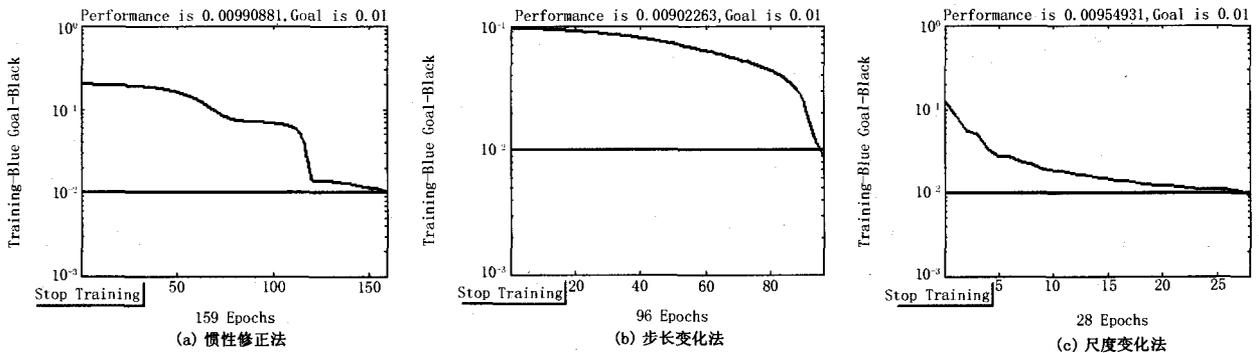


图 1 训练过程的收敛曲线

Fig.1 Concentration curve of training process

表 2 三种算法预测 12 日负荷的绝对误差

Tab.2 Absolute error of three kinds of calculating ways in prediction on 12th day

预测算法	0 点	2 点	4 点	6 点	8 点	10 点	12 点	14 点	16 点	18 点	20 点	22 点
惯性修正法	0.16	0.11	0.14	0.04	-0.20	-0.10	0.07	0.02	0.03	0.01	0.03	0.02
步长变化法	0.01	0.05	0.05	0.01	0.01	0.12	0.08	0.06	0.03	0.05	-0.17	0.02
尺度变化法	0.03	0.02	0.10	0.20	0.08	0.02	0.003	0.02	0.05	0.04	0.07	-0.12
平均绝对误差	0.067	0.06	0.097	0.083	0.097	0.080	0.051	0.033	0.037	0.033	0.090	0.053

4 证据理论融合负荷预测模型

4.1 融合样本的选取

针对表 2 的预测误差可以看出, 每一种预测算

法在某些时刻的预测误差较小, 而在某些时刻预测误差却较大, 为了更加精确地得到预测结果, 可以针对每一时刻分别选择相应的预测算法。显然, 针对某一时刻, 预测误差最小的模型可以被当作预测

模型,例如在时刻 8 点,由于误差 $0.01 < 0.08 < 0.20$, 则可以选择最小值 0.01 对应的步长变化法作为预测算法。然而,当三种算法的误差非常接近甚至相等时,该办法却无能为力,例如时刻 16 点惯性修正法与步长变化法的绝对误差均为 0.03。为了克服该传统方法带来的缺陷,将证据理论的 Dempster 合成法则运用于预测算法的择优,具有较高的预测精度。

下面首先选取融合样本,用三种算法训练好的神经网络分别对待预测日(12日)之前 5 天(7~11日)的负荷进行预测,网络的 5 组输入分别由 6~10 日的电力负荷以及 7~11 日的气候因子所组成,以时刻 0 点为例,预测结果误差即为时刻 0 点的原始融合样本,如表 3 所示。

表 3 三种算法 7~11 日时刻 0 点的负荷预测误差

Tab.3 Load forecasting error of three kinds of calculating ways at 0:00 from day 7 to 11

预测算法	7日	8日	9日	10日	11日
惯性修正法	0.20	0.02	0.21	0.12	0.12
步长变化法	0.08	0.001	0.01	0.07	0.008
尺度变化法	0.08	0.21	0.11	0.10	0.12

4.2 证据理论合成

下面用证据理论对时刻 0 点的三种预测算法进行融合。根据证据理论的相关概念,可以将误差值分为“优等”、“一般”和“较差”三个等级,分别记为 A、B、C,则预测精度识别框架的最小基本元素为 {A, B, C}。当误差值 e 越小,预测精度越高时,应该给预测精度识别框架的元素 A 赋予较高的基本信度值;反之,当误差值 e 越大,预测精度越低时,则应该给预测精度识别框架的元素 C 赋予较高的基本信度值。同时,考虑到前面定义 2 规定的条件 $m(A) + m(B) + m(C) = 1$,可以将基本信度分配规定为如下的形式:

当 $e < 0.05$ 时:

$$m(A) = 0.8, m(B) = 0.15, m(C) = 0.05$$

当 $0.05 \leq e \leq 0.1$ 时:

$$m(A) = 0.1, m(B) = 0.8, m(C) = 0.1$$

当 $e > 0.1$ 时:

$$m(A) = 0.05, m(B) = 0.15, m(C) = 0.8$$

根据上述的基本信度分配规定,综合表 3 的结果,可以得到时刻 0 点满足基本信度分配定义的融合样本,如表 4 所示。

根据表 4 的基本信度分配,设 7~11 日的基本信度函数分别为 $Bel_i (i = 7, 8, 9, 10, 11)$, 根据前述的 Dempster 合成法则,将这五个信度函数进行融合。即首先将 Bel_7 与 Bel_8 融合,然后将结果与 Bel_9

融合,逐次推进,共分成四步融合完毕。将融合后的最终结果记为: $Bel = Bel_7 \oplus Bel_8 \oplus Bel_9 \oplus Bel_{10} \oplus Bel_{11}$, 每一步的融合结果如表 5 所示。

表 4 时刻 0 点的基本信度分配

Tab.4 Basic belief degree allotment at 0:00

预测算法	基本信度分配	7日	8日	9日	10日	11日
惯性修正法	$m(A)$	0.05	0.8	0.05	0.05	0.05
	$m(B)$	0.15	0.15	0.15	0.15	0.15
	$m(C)$	0.8	0.05	0.8	0.8	0.8
步长变化法	$m(A)$	0.1	0.8	0.8	0.1	0.8
	$m(B)$	0.8	0.15	0.15	0.8	0.15
	$m(C)$	0.1	0.05	0.05	0.1	0.05
尺度变化法	$m(A)$	0.1	0.05	0.05	0.1	0.05
	$m(B)$	0.8	0.15	0.15	0.8	0.15
	$m(C)$	0.1	0.8	0.8	0.1	0.8

表 5 信度函数的融合结果

Tab.5 Fusion result of belief degree function

预测算法	合成信度分配	融合结果			
		1步融合	2步融合	3步融合	4步融合
惯性修正法	$m(A)$	0.782 23	0.763 32	0.743 26	0.722 07
	$m(B)$	0.171 92	0.194 72	0.218 41	0.242 99
	$m(C)$	0.045 85	0.041 96	0.038 33	0.034 941
步长变化法	$m(A)$	0.478 26	0.800 28	0.471 24	0.791 26
	$m(B)$	0.507 25	0.198 34	0.528 36	0.208 7
	$m(C)$	0.014 43	0.001 387	0.000 408	4.027e-5
尺度变化法	$m(A)$	0.095 238	0.090 171	0.029 86	0.027 237
	$m(B)$	0.820 11	0.838 14	0.961 65	0.965 59
	$m(C)$	0.084 656	0.071 689	0.008 493 4	0.007 171 7

通过表 5 中时刻 0 点的最终融合结果可以看出,各种预测算法对应的信度值都按一定的规律进行变化,为了选取 0 点时刻的最佳预测模型,将经过四步融合后的最终结果进行对比,如图 2 所示。

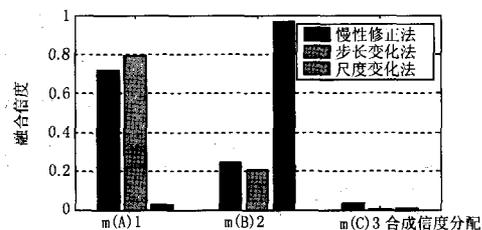


图 2 时刻 0 点基本信度分配的最终融合结果

Fig.2 Last fusion result of basic belief degree allotment at 0:00

从图 2 中可以看出,与误差等级“优等”相对应的合成信度分配 $m(A)$, 其步长变化法的融合信度最高,说明步长变化法的预测误差较小;而与误差等级“较差”相对应的合成信度分配 $m(C)$, 其

步长变化法的融合信度最低,这同样说明步长变化法的预测误差较小。因此,根据上述的分析可以断定,在三种 BP 算法中,步长变化法是时刻 0 点预测误差最小且精度最高的预测方法,于是选用步长变化法对时刻 0 点的负荷进行预测。

同理,可以按照与上面相同的融合方法对 12 日其余时刻(2 点~22 点)选取最佳的预测算法,对负荷进行预测,融合前后的预测曲线比较如图 3 所示,各种预测算法的平均误差如图 4 所示。

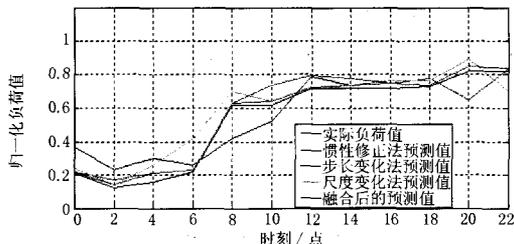
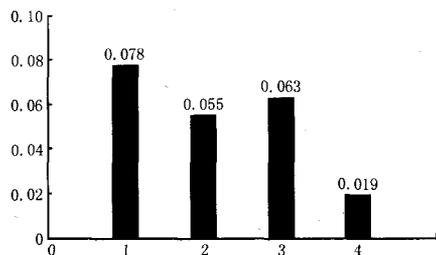


图 3 各时刻负荷预测曲线对比

Fig.3 Comparison of estimating curve at all time



1: 惯性修正法; 2: 步长修正法; 3: 尺度修正法; 4: 融合后的算法

图 4 各种预测平均误差的对比

Fig.4 Comparison of average error of all kinds of predict

从图 3 和图 4 可以看出,算法融合后的预测曲线比融合前的预测曲线更接近真实负荷曲线,且融合后的预测平均误差远低于融合前的平均误差,这进一步说明了证据理论融合算法用于负荷预测的优越性。

5 结论

1) 本文考虑到传统 BP 算法所存在的缺陷,用三种改进的 BP 网络算法对某地区的负荷进行预测,同时还考虑到了最高气温、最低气温等影响负荷的重要因素,具有较强的实用性。

2) 强调了由于各时刻负荷变化规律的不同,从而导致在负荷预测过程中要根据时刻的变化不断改变预测模型和算法,以达到更高的预测精度。

3) 提出了将证据理论用于负荷预测算法的融合,通过分析待预测日之前几天同一时刻负荷的变化规律,用信度函数量化了预测误差。通过规定基本信度分配,用 Dempster 合成法则将信度进行四重

融合,利用融合的结果选择与该时刻相对应的误差最小的预测模型。仿真结果表明,用融合后的算法对负荷进行预测时,预测的精度更高。

参考文献

- [1] 朱斌,刘晓军.应用神经网络算法进行短期负荷预测[J].江苏电机工程,2006,25(1):57-59.
ZHU Bin, LIU Xiao-jun. The Discussion of Short term Load Forecasting Applying Manual Neural Network Arithmetic[J]. Jiangsu Electrical Engineering, 2006, 25(1): 57-59.
- [2] 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,2006,18(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, 18(2):99-104.
- [3] 刘刚,黄剑华,刘学仁.基于改进型BP神经网络的短期电力负荷预测[J].电气应用,2005,24(12):46-49.
LIU Gang, HUANG Jian-hua, LIU Xue-reng. Short-term Load Forecasting Based on Improved BP Neural Network[J]. Electricity Application, 2005, 24(12): 46-49.
- [4] 倪明,单渊达.证据理论及其在应用[J].电力系统自动化,1996,20(3):76-80.
NI Ming, SHAN Yuan-da. Evidential Theory and Its Applications[J]. Automation of Electric Power Systems, 1996, 20(3): 76-80.
- [5] 陈静,吴杰,陈昊,等.多种负荷预测曲线的优化综合方法[J].江苏电机工程,2006,25(1):38-41.
CHEN Jing, WU Jie, CHEN Hao, et al. Optimal Synthesis Manifold Load Prediction Curve[J]. Jiangsu Electrical Engineering, 2006, 25(1): 38-41.
- [6] 倪明,高晓萍,单渊达.证据理论在中期负荷预测中的应用[J].中国电机工程学报,1997,17(3):199-203.
NI Ming, GAO Xiao-ping, SHAN Yuan-da. Application of Evidential Theory in Middle-term Load Forecasting [J]. Proceedings of the CSEE, 1997, 17(3): 199-203.
- [7] 蔡兴国,马平.基于证据理论的并发故障诊断方法[J].哈尔滨工业大学学报,2003,35(4):447-450.
CAI Xing-guo, MA Ping. Simultaneous fault diagnosis based evidence theory [J]. Journal of Harbin Institute of Technology, 2003, 35(4): 447-450.

收稿日期:2006-09-12; 修回日期:2006-11-30

作者简介:

孙奇(1982-),男,硕士研究生,研究方向为电力系统负荷预测、电力系统继电保护和配电网自动化; E-mail: laoqiang_ren_200@163.com

杨伟(1965-),男,副教授,主要从事电力系统运行、控制以及电力市场方面的研究和教学工作。