

基于神经网络的距离保护训练算法的研究

席晓丽, 张举

(华北电力大学电气工程学院, 河北 保定 071003)

摘要: 对神经网络式距离保护做了简要概述, 并且为了使该新型保护更加完善, 对其存在的训练算法问题做了深入的研究。对 8 种改进的 BP(Back - Propagation) 算法训练神经网络进行了比较, 通过反复训练, 得出 8 种算法在权值调整和误差收敛过程中所遇到的共同点和差别, 并得出初始权值、激励函数和神经元个数等因素的选取对神经网络学习的影响。实验证明: LM 算法和尺度共轭梯度反向传播算法训练网络收敛速度快, 受权值影响小, 但占用较大的内存; 贝叶斯正则化算法有较好的泛化能力。

关键词: 神经网络; 距离保护; BP 算法; 电力系统

中图分类号: TM773 **文献标识码:** A **文章编号:** 1003-4897(2005)04-0026-05

0 引言

距离保护一直是中高压电网中应用最广泛的保护方案。在实际应用中, 距离保护通常被用作中压电网的主保护和高压电网的后备保护。距离保护有很多优点, 但是也存在严重的不足之处: 由于过渡电阻的存在, 系统振荡的影响, 系统运行方式的变化, 均可能导致保护性能的下降等。目前, 作为传统距离保护的改进, 基于神经网络的距离保护已经成为一个热点话题。

神经网络 ANN(Artificial Neural Network) 在电力系统中的应用越来越多。神经网络式距离保护的基本原理就是从系统的各种参数的组合模式来识别故障状态, 从这个意义上说, 神经网络式距离保护的基本原理已与传统保护原理完全不同。它既不用测量阻抗, 也不用规定阻抗继电器的特性。之所以仍称之为距离保护, 是因为它可以实现距离保护的各种功能: 振荡识别、故障检测等。并且为了与传统保护概念相一致, 将这种按阶梯时限工作的保护仍称之为距离保护。神经网络式距离保护的特点是: 原理简单; 反应迅速、准确; 即使在不同的故障条件和网络条件下也能正常工作; 能避免常规保护存在的问题, 如由于过渡电阻或者振荡引起的拒动误动等等。将神经网络应用于保护中, 一个比较关键的问题就是训练网络算法的选择。本文对采用 ANN 实现电力系统输电线的距离保护算法的选择进行了详细说明, 并分析了其各自的优缺点。

1 神经网络式距离保护简介

神经网络式距离保护可由 3 个独立的子网络来

组成^[1]: 振荡识别子网络 ANN1、故障检测和选相子网络 ANN2、故障定位子网络 ANN3。ANN1 用以识别系统的振荡状态; ANN2 用以检测系统中是否发生内部故障及识别故障类型; ANN3 用以确定故障是否发生在保护动作范围内。对于距离保护的每一个保护段, 需要 3 个子网络一起作出判断, 给出系统所处状态。图 1 为神经网络式距离保护结构示意图。来自保护安装处的 CT、PT 的二次侧的电流、电压量经过同步数据采集及处理后, 各子网络分别对其进行特征量提取, 加于网络的输入层节点, 经过逐层计算后, 各个子网络分别给出计算结果。由各个子网络输出的结果经过综合判断后, 确定系统所处的状态。当保护启动, 确定为区内故障时, 按故障相别发出相应的保护跳闸命令, 并给出保护启动、故障相别等信号, 如果 ANN3 判断为无故障, 而 ANN1 判断为有振荡, 则将保护段退出。段是否应退出视具体情况而定。如无振荡信号, 则可不退出保护。在振荡过程中出现保护启动并判有区内故障时, 仍可动作切除故障。神经网络式距离保护由几个相互独立、并行处理的网络组成, 根据各个子网络的功能, 从而确定网络的结构。对于功能简单的网络, 选择常规的 BP 算法就足以在时间、精度上取得很好的结果; 而对于复杂的、学习样本数多的子网络, 其需要进行分类的模式种类很多, 因此仍采用常规的 BP 算法将会浪费大量的训练时间, 同时由于误差曲面的复杂性, 可能会陷入局部极小值。对此, 本文就多种改进的算法加以讨论, 并通过反复训练, 得出这几种算法在权值调整和误差收敛过程中所遇到的共同点和差别, 得出初始权值、激励函数和神经元个数等因素的选取对神经网络学习的影响。

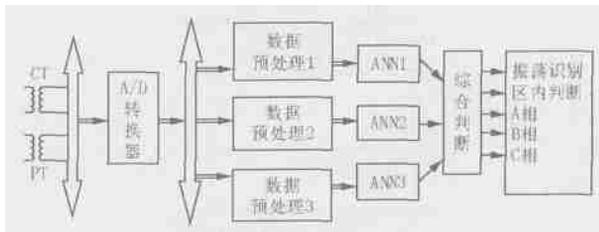


图1 神经网络式距离保护结构示意图

Fig. 1 Configuration of distance protection based on ANN

2 BP 算法概述

BP 算法已成为目前应用最为广泛的神经网络学习算法,据统计有近 90% 的神经网络应用是基于 BP 算法的^[3]。BP 神经网络采用的传递函数通常是 Sigmoid 型可微函数,所以可以实现输入和输出间的任意非线性映射。对数 Sigmoid 函数在 Matlab 中为 logsig 函数,正切 Sigmoid 函数在 Matlab 中为 tansig 函数。

2.1 BP 网络拓扑结构及训练算法

BP(Back Propagation)神经网络的拓扑结构如图 2 所示, w_{ij} 为输入层第 i 个神经元与隐层第 j 个神经元之间的连接权值; v_{jk} 为隐层第 j 个神经元与输出层第 k 个神经元之间的连接权值; x_{pi} 为输入层第 i 个神经元的输入; c_{pk} 为输出层第 k 个神经元的输出; $i=1,2, \dots, n; j=1,2, \dots, l; k=1,2, \dots, M; n, l, M$ 分别为输入层、隐含层和输出层的神经元个数; $p=1,2, \dots, P, P$ 为样本总数。神经元的工作特性为:

$$S_{pj} = \sum_{i=1}^n w_{ij} \times x_{pi} - \theta_j \quad b_{pj} = f(S_{pj})$$

$$t_{pk} = \sum_{j=1}^l v_{jk} \times b_{pj} - \theta_k \quad c_{pk} = f(t_{pk})$$

式中: θ_j 为隐含层第 j 个神经元的阈值; θ_k 为输出层第 k 个神经元的阈值; s_{pj}, b_{pj}, t_{pk} 分别为输入层神经元的输出、隐含层神经元的输出和输出层神经元的输出。期望输出信号,即教师信号为 y_{pk} ,则误差函数和权值调节公式为:

$$E_p = \sum_{k=1}^M (y_{pk} - c_{pk})^2 / 2$$

$$w_{ij}(t+1) = w_{ij}(t) - \frac{\partial E_p}{\partial w_{ij}}(t)$$

2.2 BP 算法的改进

BP 算法的主要缺点: 收敛速度慢; 局部极值; 难以确定隐层和隐节点的个数。

标准 BP 算法实质上是一种简单的最速下降静态寻优法,在修正 $w_{ij}(k)$ 时,只是按照 k 时刻的负梯

度方向进行修正,而没有考虑到以前积累的经验,即以前时刻的梯度方向,从而常常使学习过程发生振荡,收敛缓慢。

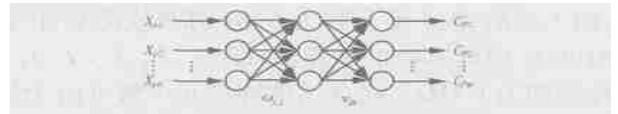


图2 BP神经网络的拓扑结构

Fig. 2 Topology structure of BP ANN

1) 动量 - 自适应学习算法

$w_{ij}(k+1) = w_{ij}(k) + [(1 - \alpha) D(k) + \alpha D(k-1)]$
其中: $w_{ij}(k)$ 既可表示单个的权值,也可表示权值向量; $D(k) = -\frac{\partial E}{\partial w_{ij}(k)}$ 为 k 时刻的负梯度; $D(k-1)$ 为 $k-1$ 时刻的负梯度。 α 为学习率, $\alpha > 0$; α 为动量因子, $0 < \alpha < 1$ 。这种方法所加入的动量项实质上相当于阻尼项,它可以减少学习过程的振荡趋势,从而改善收敛性。

2) 自适应调整学习率的改进算法:

$$w_{ij}(k+1) = w_{ij}(k) + \eta(k) \times D(k)$$

$$\eta(k) = 2^{-1} \times \eta(k-1)$$

$$\eta(k) = \text{sign}[D(k) \times D(k-1)]$$

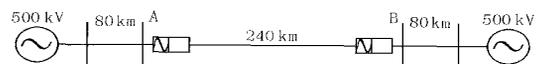
当连续两次迭代其梯度方向相同时,表示下降太慢,这时可以使步长加倍;当连续两次迭代其梯度方向相反时,表明下降过头,这时可以使步长减半。

3) LM(Levenberg - Marquardt) 优化方法的权值调整率选为 $\eta = (J^T J + \mu I)^{-1} J^T e$ 。其中 J 为误差对权值微分的 Jacobian 矩阵; e 为误差向量; μ 为一个标量,当 μ 很大时,上式就接近于梯度法。 μ 是自适应调整的。

2.3 各种方法训练网络的比较

2.3.1 用于仿真的电力系统模型

模型如图 3 所示。



$$Z_{m1} = 1.2857 + j55.6377 \quad Z_1 = 0.01808 + j0.27747$$

$$Z_{m0} = 2.87644 + j37.1267 \quad c_1 = 0.01297 \mu\text{F/km}$$

$$Z_{m1} = 1.6714 + j35.5391 \quad Z_0 = 0.23084 + j0.97208$$

$$Z_{m0} = 4.7000 + j37.9784 \quad c_0 = 0.008116 \mu\text{F/km}$$

图3 电力系统仿真模型

Fig. 3 Model of power system simulation

如图 3 一个 500 kV 的双端电源的系统。其中元件参数均已给出。使用 ATP 对该系统进行各种故障

仿真,仿真数据作为训练和测试的样本。以单向接地故障作为典型情况(因为 80% 以上的故障均为单相接地故障)。我们研究的保护是安装在 A 处的保护。构建一个 ANN,用于判断区内、外故障。网络选为 3 层,输入层的神经元个数为 6,分别取故障后半周期内电压、电流基波的采样值,即 U_a 、 U_b 、 U_c 、 I_a 、 I_b 、 I_c 。采样周期为 1 kHz。隐含层的神经元个数可由大量的实验确定,这里分别对隐含层神经元个数为 32、34、36、38、22、24、26、28 时作出实验,经过大量实验证明,当个数为 32 个时,收敛效果最好,因此下图仅列出了此时的误差曲线。输出层神经元个数为 1 个(1 表示区内故障,0 表示区外故障)。激励函数分别选用 $\tan\text{sig}$ (隐含层) - tansig (输出层)型和 tansig (隐含层) - purelin (输出层)型。在训练网络之前,须对网络的输入样本进行归一化处理。

2.3.2 各种训练方法的列举

1) 采用梯度下降反向传播算法对网络进行训练

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{traingd}(\text{net}, \text{Pd}, \text{Ti}, \text{Ai}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

该方法实现了最速梯度下降算法,即基本 BP 算法,其它梯度下降算法都是该算法的变形。如图 4 所示(最大训练次数为 1 500 次,横坐标为训练次数,纵坐标为误差大小,goal 为均方误差指标,performance 为最终的误差值)。

2) 根据已设定的权值和阈值学习函数对网络随机进行训练

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{trainr}(\text{net}, \text{Pd}, \text{Ti}, \text{Ai}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

此方法采用随机训练,即网络每接收一个输入数据,都要对权值和阈值进行调整,而且输入数据是以随机顺序排列的。如图 5 所示。

3) 采用 LM 反向传播算法对网络进行训练

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{trainlm}(\text{net}, \text{Pd}, \text{Ti}, \text{Ai}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

该方法依据 Levenquardt - Marquardt 优化理论对网络的权值和阈值进行调整。如图 6 所示。

4) 采用贝叶斯正则化对网络进行训练

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{trainbr}(\text{net}, \text{Pd}, \text{Ti}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

该方法也是依据 Levenquardt - Marquardt 优化理论对网络的权值和阈值进行调整的。如图 7 所示。

5) 采用尺度化共轭梯度反向传播算法对网络进行训练

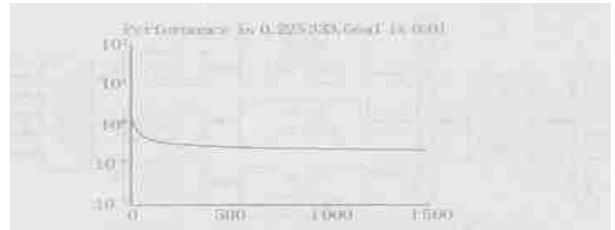


图 4 traingd 法误差曲线

Fig. 4 Error curve of traingd method

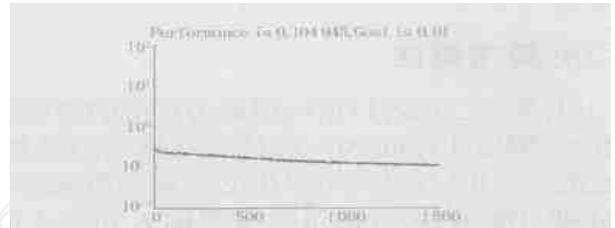


图 5 trainr 法误差曲线

Fig. 5 Error curve of trainr method

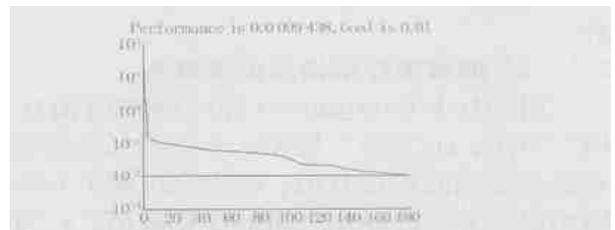


图 6 LM 法误差曲线

Fig. 6 Error curve of LM method

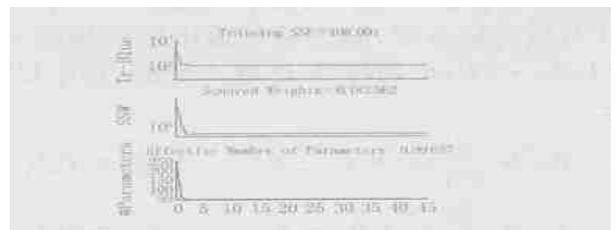


图 7 trainbr 法误差曲线

Fig. 7 Error curve of trainbr method

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{trainscg}(\text{net}, \text{Pd}, \text{Ti}, \text{Ai}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

该方法不仅是共轭梯度算法的一种变形,而且结合了 Levenberg - Marquardt 算法中的模型置信区间方法和共轭梯度算法,避免了耗时巨大的线搜索过程。如图 8 所示。

6) 采用自适应学习速率梯度下降反向传播算法对网络进行训练

格式: $[\text{net}, \text{TR}, \text{Ac}, \text{El}] = \text{traingda}(\text{net}, \text{Pd}, \text{Ti}, \text{Ai}, \text{Q}, \text{TS}, \text{VV}, \text{TV})$

该方法是在最速梯度下降算法的基础上将学习率进行自适应调整。如图 9 所示。

7) 采用动量梯度下降反向传播算法对网络进行训练

格式: [net, TR, Ac, El] = traingdm(net, Pd, Tl, Ai, Q, TS, VV, TV)

该方法在最速梯度下降算法的基础上,对权值和阈值更新时,不仅考虑当前的梯度方向,而且还考虑了前一时刻的梯度方向,这种调整方式是通过设置动量因子来实现的。如图 10 所示。

8) 采用自适应学习速率动量梯度下降反方向传播算法对网络进行训练

格式: [net, TR, Ac, El] = traingdx(net, Pd, Tl, Ai, Q, TS, VV, TV)

该方法结合了方法 6 和方法 7。如图 11 所示。

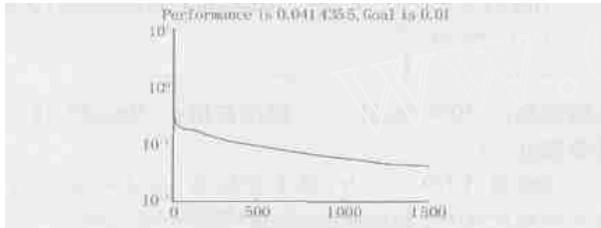


图 8 traingcgm 法误差曲线

Fig. 8 Error curve of traingcgm method

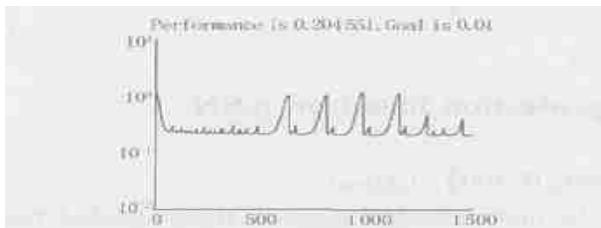


图 9 traingda 法误差曲线

Fig. 9 Error curve of traingda method

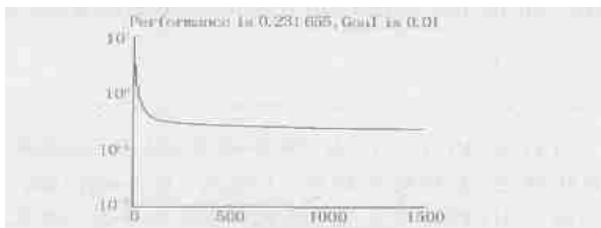


图 10 traingdm 法误差曲线

Fig. 10 Error curve of traingdm method

2.3.3 各种方法训练效果比较

上面的误差曲线图是在相同的条件下(相同的均方误差性能指标,相同的最大训练次数,相同的层数以及神经元个数,传递函数均采用 tansig 和 purelin) 得出的,误差公式为:

$$E_p = \sum_{k=1}^M (y_{pk} -$$

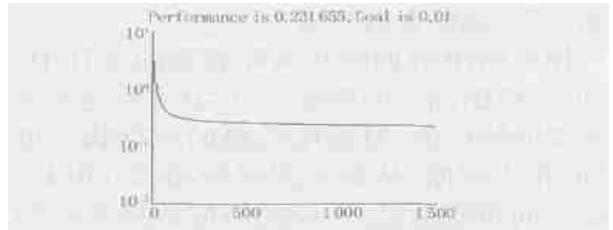


图 11 traingdx 法误差曲线

Fig. 11 Error curve of traingdx method

$c_{pk})^2/2$, 可以得出以下结论:

1) LM 方法比其它方法的误差均减小的快得多,学习次数少。但其占用内存较大,占用内存量与隐含层个数相关,如果神经元个数选得太多,训练时必须要有足够的内存,否则程序无法运行。

2) Trainr 方法由于不是成批处理数据,因此训练速度很慢,并且收敛效果不佳。其训练结果与初始权值的选取有较大关系,与隐含层神经元个数的多少没有太大的关系。因此此种方法只适合于小型网络的训练。

3) Trainbr 方法的显著优点在于其训练出来的网络有很强的推广能力,这一点在用测试样本检验网络的实验中已被证实,而且经过大量的实验,证明这种方法不会出现“训练过度”的情况。该方法训练速度较快,占用内存不大,收敛效果较好。

4) Traingda 方法极容易出现局部极小值,其收敛效果与初始权值有较大的关系。占用内存少。

5) Traingdx、traingdm 方法的收敛速度较慢,与初始权值的设定有很大关系。

6) Traingcgm 方法的收敛效果较 LM 法差,但要比其它的方法都好,训练速度快,收敛效果与神经元个数有很大关系。

7) Traingd 训练速度快,但收敛效果不好,占用内存较小。

8) 适当增加神经元的个数会使各种算法的收敛性变好,神经元的激励函数用 tansig 和 purelin 结合使用比单一使用效果好得多。

3 结论

LM 算法和尺度化共轭梯度反向传播算法 traingcgm 要比其它的方法效果好得多,尤其是 LM 算法,收敛速度非常快,受初始权值的影响小,但是占用的内存比较大。贝叶斯正则化 Trainbr 方法有很强的泛化能力。因此在训练较大型网络时,推荐使用以上 3 种网络。另外,网络的初始权值和隐含层的神经元个数的选取也很重要。对于具体的网络,经过

权衡之后,选择合适算法。

神经网络的训练在实际装置的设计中,应该离线训练,得出固定的权值,然后从实际运行系统中提取需要的特征值,对输电线路的运行状态进行分析识别,作出判断,从而实现保护功能。但是,如果从自适应的角度来讲,神经网络应该将实时的运行状况作为知识,重新存储到权值里,这样就要求在线训练网络。可是以上8种方法中,耗时最短的是LM方法,用了25s完成网络的训练,达到了最小误差。可是这个时间绝对满足不了保护对快速性的要求。因此,我们期待着更好的算法出现。

参考文献:

- [1] 段玉倩,贺家李.基于人工神经网络的距离保护[J].中国电机工程学报,1999,19(5):67-70.
DUAN Yu-qian, HE Jia-li. Distance Protection Based on ANN[J]. Proceedings of the CSEE,1999,19(5):67-70.
- [2] 褚法玉,陈晓科,王宁.基于人工神经网络输电线的速断保护[J].电力自动化设备,2003,23(3):67-69.
CHU Fa-yu, CHEN Xiao-ke, WANG Ning. Transmission Line Protection Based on ANN[J]. Electric Power Automation Equipment,2003,23(3):67-69.
- [3] 许东.基于MATLAB6.X的系统分析与设计——神经网络(第二版)[M].西安:西安电子科技大学出版社,2003.
XU Dong. System Analysis and Design Based on MATLAB6.X——ANN,Second Edition[M]. Xi'an: Xidian University of Press,2003.
- [4] 都洪基,邓烽,苏炜宏.基于人工神经网络的自适应距离保护[J].继电器,2002,30(2):38-41.
DU Hong-ji, DENG Feng, SU Wei-hong. Adapting Distance Protection Based on ANN[J]. Relay,2002,30(2):38-41.
- [5] 朱洪波,许兆凤,马燕凤,等.两种改进BP算法神经网络滤波的比较[J].电力自动化设备,2003,23(6):35-36.
ZHU Hong-bo, XU Zhao-feng, MA Yan-feng, et al. Comparison of Two Improved BP Algorithms ANN Filtering[J]. Electric Power Automation Equipment,2003,23(6):35-36.

收稿日期: 2004-06-21; 修回日期: 2004-07-18

作者简介:

席晓丽(1979-),女,硕士研究生,主要研究方向为电力系统继电保护;E-mail: xixiaoli875@sohu.com

张举(1946-),男,教授,从事微机继电保护的教学与研究工作。

Research of training algorithm for distance protection based on ANN

XI Xiao-li, ZHANG Ju

(North China Electric Power University, Baoding 071003, China)

Abstract: Distance protection based on ANN is introduced. In order to make the new protection perfect, the method of training network is profoundly studied. Eight improved BP algorithms, which are used in ANN(Artificial Neural Network) training network, are compared. By repeating training, their common ground and difference in the process of weight adjusting and error convergency are drawn, and the effect of neural number, exciting function and initial weight on the training of ANN are summarized. It is proved that LM arithmetic and trainscg method have high convergence rates and are less affected by the weights in training networks while using large memory; and trainbr method has a better ability of generalization.

Key words: neural network; distance protection; BP algorithm; power system

(上接第25页 continued from page 25)

收稿日期: 2004-06-02; 修回日期: 2004-06-29

作者简介:

胡彩娥(1971-),女,博士研究生,研究领域为电力系统电压稳定、无功优化规划;E-mail: hucaie@cau.edu.cn

杨仁刚(1953-),男,教授,博士生导师,研究领域为无功补偿装置、配电网自动化。

Optimal planning for reactive power with voltage stability

HU Cai-e, YANG Ren-gang

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: The mathematical model of fuzzy multi-objective optimal planning for reactive power is established. The objectives consist of four important terms including network loss, investment of reactive compensation equipment, static voltage stability margin, and voltage deviation of PQ bus. First, voltage stability sensitivity factor index method is applied to identify weak bus as candidate nodes for installing new reactive compensation equipment, then capability and location are optimized with genetic algorithm (GA). The present method is applied to IEEE30 bus systems with satisfactory result and has improved the voltage stability of the whole power system.

Key words: voltage stability sensitivity; fuzzy multi-object; static voltage stability margin