

# 基于优化的 BP 神经网络地层可钻性预测模型

董青青, 梁小丛

(中国地质大学(武汉)工程学院, 湖北 武汉 430074)

**摘要:**提出了一种粒子群算法(PSO)优化的 BP 网络模型预测地层可钻性的新方法。利用粒子群算法优化 BP 网络模型的参数,避免了 BP 网络陷入局部极小值的缺点,提高了模型的预测速度和精度。结合钻探实例,利用测井资料和地层可钻性级别的关系建立了可钻性级别实时预测模型,并将该模型与传统的 BP 网络进行对比,结果表明,该模型优于 BP 网络,具有较高的精度和较快的收敛速度,有一定的适用性。

**关键词:**地层可钻性;BP 网络模型;粒子群算法;预测模型

中图分类号:P634.1 文献标识码:A 文章编号:1672-7428(2012)11-0026-03

**A Model for Predicting Formation Drillability Based on Optimized BP Neural Network/DONG Qing-qing, LIANG Xiao-cong** (Engineering College, China University of Geosciences, Wuhan Hubei 430074, China)

**Abstract:** A new method for predicting formation drillability was proposed according to the theory of BP networks based on PSO. The use of PSO optimizing the parameters of BP networks is to improve the convergence speed and precision of BP neural networks. Combining with the examples of drilling and based on the relationship of log information formation drillability grade, a real-time formation drillability grade model was established. The results show that the model is superior to BP network with higher accuracy and faster convergence rate and it is an effective way to predict formation drillability.

**Key words:** formation drillability; BP network model; PSO; prediction model

## 0 研究背景

在岩土钻掘工程设计与实践中,我们常常希望能事先预知地层岩石的破碎难易程度,以便正确选择钻进工艺、方法,制定生产定额。因此探讨地层岩石可钻性的方法显得极为重要。目前国内外普遍采用的岩石可钻性实验方法主要有:室内岩心实验法,和参数分析法 2 大类。都是钻后评价方法,实验用的岩心脱离了地下的真实环境,因此实验数据具有一定的偏差性,并不能真实反映岩石的可钻性级数。

而 BP 网络模型是目前广大研究者用于地层岩石可钻性的钻前预测模型。利用测井资料作为网络模型的输入参数,岩层可钻性级值作为输出,此模型结构简单,具有良好的自适应性、自学习性和极强的非线性逼近、容错能力等特点。但该模型容易陷入局部极小值,且泛化能力较弱、网络收敛速度较慢。在本文中提出一种基于粒子群算法优化的 BP 网络预测模型,在该模型中用粒子群算法代替了梯度下降法对 BP 网络的权值和偏差进行优化,提高了模型的收敛速度和预测的精度。因此利用此模型建立地层可钻性级值的预测具有更高的精度和更广泛的适用性。

## 1 粒子群算法(PSO)基本原理

粒子群优化算法<sup>[1,2]</sup>把群体中的每个个体视为

搜索空间的“鸟”,也即‘粒子’。在粒子群优化问题中,每个解就是空间中的粒子,所有的粒子都有一个被优化的函数所决定的适应值。空中的粒子以一定的速度飞行,并根据粒子本身最优适应值  $pbest$  以及全局的最优适应值  $gbest$  对速度进行动态调整,从而形成群体寻优的正反馈机制,直到寻找到群体最优适应值为止。设在一个  $D$  维的目标搜索空间中,有  $N$  个粒子组成一个群落,将每个粒子代入一个目标函数即可算出其适应值,根据适应值的大小衡量是否为所要求解的最优解。设第  $i$  个粒子的位置为  $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$ ,速度也是一个  $M$  维向量,记为  $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$ 。记第  $i$  个粒子迄今为止搜索到的最优位置为  $pbst_i = (pb_{i1}, pb_{i2}, \dots, pb_{id})$ ,而整个粒子群搜索到的最优位置为  $gbst_i = (gb_{i1}, gb_{i2}, \dots, gb_{id})$ 。粒子位置和速度改变的 2 个公式:

$$V_{ij}(k+1) = wV_{ij}(k) + c_1 rand_1 [pb_{ij}(k) - X_{ij}(k)] + c_2 rand_2 [gb_{ij}(k) - X_{ij}(k)] \quad (1)$$

$$X_{ij}(k+1) = X_{ij}(k) + V_{ij}(k+1) \quad (2)$$

式中: $c_1, c_2$ ——加速系数,即分别调节向全局最好粒子和个体最好粒子方向飞行的最大步长; $rand_1, rand_2$ —— $[0, 1]$ 之间的随机数; $V_{ij}$ ——第  $i$  个粒子在第  $k$  次迭代中第  $j$  维的速度; $X_{ij}(k)$ ——粒子  $i$  在第  $k$  次迭代中第  $j$  维的当前位置; $pb_{ij}(k)$ ——粒子  $i$

收稿日期:2012-07-07

作者简介:董青青(1987-),女(汉族),陕西人,中国地质大学(武汉)博士在读,地质工程专业,从事地质灾害研究工作,湖北省武汉市鲁磨路 388 号,dongqingqing2012@sina.com。

在第 $j$ 维第 $k$ 此迭代个体极值点的位置; $gb_j(k)$ ——整个粒子群在第 $j$ 维第 $k$ 次迭代全局极值点的位置; $w$ ——惯性权重。

惯性权重是最重要的参数,较大的权重有利于提高算法的全局搜索能力,而较小的权重会增强算法的局部搜索能力,因此在本文中采用线性递减权重。

$$w = \frac{(w_{\text{star}} - w_{\text{end}})(t_{\text{max}} - t)}{t_{\text{max}}} + w_{\text{end}} \quad (3)$$

式中: $t_{\text{max}}$ ——最大允许迭代次数; $t$ ——当期迭代次数; $w_{\text{star}}$ 、 $w_{\text{end}}$ ——分别是初始权重和进化到最大迭代次数的权重。

在更新过程中,粒子飞行速度有最大值 $V_{\text{max}}$ 的限制,若 $V(k+1) > V_{\text{max}}$ 则 $V(k+1) = V_{\text{max}}$ ;若 $V(k+1) < -V_{\text{max}}$ 则 $V(k+1) = -V_{\text{max}}$ 。较大的 $V_{\text{max}}$ 可以保证粒子群的全局搜索能力, $V_{\text{max}}$ 较小则种群的局部搜索能力加强。同样粒子的位置也限制在 $[-X_{\text{max}}, X_{\text{max}}]$ 区间。

## 2 优化网络模型基本原理

神经网络模型具有较强的非线性映射能力。而BP网络是目前应用最广泛、最成熟的神经网络之一。在一般情况下BP网络取一层隐含层就可以逼近绝大部分非线性的问题。BP网络权值和偏差一般采用误差梯度下降法逼近,有时并不能保证所求为误差超平面的全局最小解,很可能是一个局部极小解。所以在一些复杂工程状况采用梯度下降法训练权值和偏差并不能达到预期的值。而用粒子群算法优化权值和偏差避免了梯度算法中要求函数可微、对函数求导的过程,缩短了神经网络的训练时间,在一定程度上改善了BP算法性能,使其跳出局部极小值的范畴。

用PSO算法优化BP网络模型参数,定义粒子群初始位置向量 $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$ ,向量中的维数的大小 $D$ ,即为BP网络中权值和偏差的总和,向量中每一元素分别表示网络中权值或偏差。PSO训练BP网络时使用如下适应值函数:

$$E = \frac{1}{n} \sum_{i=1}^n N_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^M (Y_{i,j} - y_{i,j})^2 \quad (4)$$

式中: $n$ ——训练的样本总数; $M$ ——网络输出神经元的个数; $Y_{i,j}$ ——第 $i$ 个样本的第 $j$ 个网络输出节点的理想输出值; $y_{i,j}$ ——第 $i$ 个样本的第 $j$ 个网络输出节点的实际输出值。

粒子群算法优化BP网络的详细步骤如下:

(1)对各参数进行初始化:①初始化神经网络输

入层个数 $N$ ,隐含神经元个数 $L$ 和输出层个数 $M$ ;②初始化粒子群算法种群维数 $D = NL + ML + L + M$ ,粒子数 $n = 25$ ,设定初始和终止惯性权重 $w_{\text{star}} = 0.9$ 、 $w_{\text{end}} = 0.2$ ,加速因子 $c_1 = c_2 = 2$ ,误差精度 $\varepsilon = 0.001$ ,最大代数 $T = 2000$ ;③初始化粒子群位置和速度矩阵;④初始化每一个粒子的个体极值和全局最优值。

(2)逐个输入粒子并根据BP算法计算粒子的适应度。

(3)由计算所得粒子的适应值与它经历最好位置的适应值进行比较,若现在的位置适应值比它经历最好位置更好,则更新个体极值。

(4)比较每个粒子的适应值与全局所经历最好位置的适应值,如果更好,更新全局极值点。

(5)根据上述方程(1)和方程(2)更新粒子的速度和位置。

(6)如未达到足够好的适应值或一个预设的最大迭代数则结束,否则返回步骤2。

## 3 优化的BP网络模型的建立

由文献[3~5]可知岩层的可钻性级别与声波时差 $\Delta t$ 、地层密度 $\rho$ 、电阻率 $R_t$ 以及泥质质量分数 $V_{\text{sh}}$ 关系十分密切。而这些参数都可以由测井资料所获得,因此在进行深部钻进时可以利用已经获得的上层井段或者同一地区已钻井的测井资料作为模型的分析参数,并基于粒子群优化的BP神经网络建立未钻岩层可钻性级别预测模型。将与地层可钻性密切相关的参数(声波时差、地层密度、泥质质量分数、地层深度)作为网络的输入量,以地层可钻性级别 $K_d$ 作为网络的输出,构成的岩层可钻性模型如图1所示。

## 4 地层可钻性预测模型实例分析

根据上述所建立粒子群算法优化的BP网络模型,基于MATLAB建立编程软件,利用文献[6]数据建立如表1所示地层参数。

将与地层可钻性密切相关的测井参数( $\Delta t$ 、 $\rho$ 、 $R_t$ 、 $V_{\text{sh}}$ )作为网络输入层的输入量,以地层可钻性极值 $K_d$ 作为输出层的期望输出量。对于隐含层神经元个数并没有确定的规定值,隐含层神经元个数越多训练所需时间越长,网络性能越差。因此可以利用for循环及网络性能评价函数(5)和误差评价函数(6)选取最优隐含层神经元个数:

$$\text{error} = \text{mse}(e, \text{net}) \quad (5)$$

$$\text{res}(i) = \text{norm}(\text{error}) \quad (6)$$

其中 $e$ 为真实值和预测值误差,根据所有的 $\text{res}$

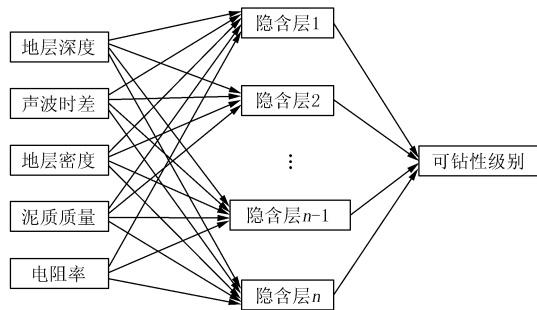


图1 岩层可钻性预测模型

表1 地层测井参数和岩层可钻性级数数据

岩样号	深度 /m	$\Delta t / (\mu s \cdot ft^{-1})$	$\rho / (g \cdot cm^{-3})$	$R_t / (\Omega \cdot m)$	$V_{sh} / \%$	$K_d$
19	4194.000	45.8	2.74	2496.0	1.7	4.452
37	4195.875	50.5	2.65	778.4	1.2	6.848
65	4198.250	53.2	2.63	1246.3	0	7.769
72	4198.875	50.8	2.76	1095.5	0.7	7.309
83	4200.000	43.9	2.81	1955.4	1.3	7.966
125	4205.500	47.5	2.77	2395.4	0.2	6.548
185	4212.125	45.0	2.80	3981.2	1.3	7.003
193	4213.125	47.4	2.78	3724.8	0.9	8.355
234	4217.875	48.7	2.74	3031.4	0.2	5.676
262	4221.250	44.1	2.61	14561.6	2.4	6.385
356	4231.250	54.2	2.67	3975.4	0	3.070
387	4235.000	52.5	2.61	1482.7	0.4	4.669
403	4237.250	57.9	2.59	1513.8	0	4.693
423	4239.250	58.5	2.57	1716.2	0	3.845
518	4249.250	50.3	2.72	927.7	23.7	5.128
591	4257.125	54.2	2.67	1792.0	0.2	5.805
605	4258.625	52.0	2.68	2063.6	0.5	6.425
704	4269.250	47.6	2.67	40984.6	0.8	6.007
708	4269.750	48.7	2.69	40982.5	0.3	6.057
733	4272.750	47.8	2.72	4449.7	1.2	5.716
750	4275.500	51.7	2.65	1955.3	0.4	5.942
777	4278.375	47.9	2.78	4597.9	4.5	7.079

值选取最小值所对应的隐含神经元个数作为最优隐含层。同时粒子群算法中取粒子数 25, 最大迭代次数 2000。粒子位置的上界为  $[1 \times 10^2, 1 \times 10^6]$ , 下界为  $[1 \times 10^{-6}, 1 \times 10^{-1}]$ , 粒子最大速度为  $[5 \times 10^{-2}, 5 \times 10^4]$ 。

利用表 1 中数据对所设计好的模型进行训练, 并利用后 10 组数据对地层可钻性极值进行回判, 所得岩层可钻性级别回判结果如表 2 所示。将所得结果绘制成曲线如图 2 所示。由表 2 和图 2 可以知基于粒子群优化 BP 网络模型其相对误差一般不超过 3%, 比未经优化过的 BP 网络精度高, 而且波动也较稳定, 与真实值拟合的程度高。因此基于粒子群优化 BP 网络模型能够更好预测地层可钻性, 反应地层可钻性的变化趋势。

表2 PSO-BP与BP预测值对比

岩样号	实际值	PSO-BP 预测	相对误差 /%	BP 预测	相对误差 /%
403	4.693	4.681	0.25	4.511	3.88
423	3.845	3.744	2.62	3.594	6.52
518	5.128	5.226	1.91	5.479	6.84
591	5.805	5.703	1.75	5.980	3.01
605	6.425	6.447	0.34	6.174	0.04
704	6.007	6.057	0.83	6.189	0.03
708	6.057	6.158	1.66	5.852	3.38
733	5.716	5.617	1.73	5.927	3.69
750	5.786	5.796	0.17	5.824	0.65
777	6.008	6.108	1.66	6.354	4.56

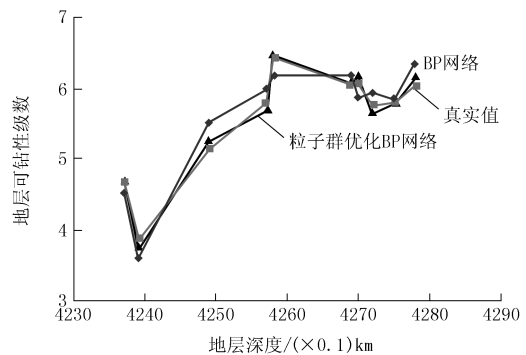


图2 两种预测地层可钻性模型的曲线对比

## 5 结论

(1) 由以上工程实例分析可知, 用粒子群算法优化的神经网络模型具有更高的精度, 误差较小, 而且避免了 BP 神经网络模型容易陷入局部极小值的局限, 具有较大的适应性, 为地层可钻性提供了一种较高精度的预测模型。

(2) 在实际钻进中利用上部地层或者相邻已钻地层的测井参数和对应的地层可钻性级数对网络进行训练。利用训练好的网络和钻头底下未钻地层的相关测井参数即可预测出地层可钻性级数, 因此该模型对未钻地层的可钻性实时预测有一定的推广性。

## 参考文献:

- [1] 谭明交, 张红梅. 粒子群优化算法及其在神经网络中的应用[J]. 计算机与信息技术, 2008, (7).
- [2] 徐乐华, 凌卫新, 熊丽琼. 基于面向对象自适应粒子群算法的神经网络训练[J]. 计算机应用研究, 2009, (1).
- [3] 高成军, 湛海云, 卢青兰, 等. 测井资料在预测地层参数剖面方面的应用[J]. 西南石油大学学报(自然科学版), 2008, (3).
- [4] 刘向君, 宴建军, 罗平亚, 等. 利用测井资料评价岩石可钻性研究[J]. 天然气工业, 2005, (7).
- [5] 王克雄, 魏凤奇. 测井资料在地层抗钻特性参数预测中的应用研究[J]. 石油钻探技术, 2003, (5).
- [6] 夏宏泉, 刘之的, 陈平, 等. 基于 BP 神经网络的岩石可钻性测井计算研究[J]. 测井技术, 2004, (2).