

doi: 10.7690/bgzdh.2016.04.023

一种改进粒子群算法优化 BP 神经网络实现核素识别方法

刘议聪, 朱泓光, 宋永强

(绵阳市维博电子有限责任公司, 四川 绵阳 621000)

摘要: 为获得全局最优点, 提出一种改进粒子群算法优化 BP 神经网络实现核素识别方法。该算法用一种动态改变惯性权重与学习因子的自适应方法, 优化 BP 神经网络的阈值与权值, 通过训练 BP 神经网络识别模型得到粒子群的全局最优解, 利用最优权值与阈值实现核素识别。分析结果表明: 该方法不仅能更快地收敛于最优解, 同时能更好地平衡全局搜索和局部搜索能力, 有效地改善算法的收敛速度和识别精度。

关键词: 粒子群算法; 核素识别; 全局最优点; 惯性权重; 学习因子

中国分类号: TP301.6 **文献标志码:** A

Optimize BP Neural Network by an Improved Particle Swarm Optimization to Implement Nuclide Identification

Liu Yicong, Zhu Hongguang, Song Yongqiang

(Mianyang Weibo Electronic. Ltd, Mianyang 621000, China)

Abstract: To get the global optimal point, propose an optimize BP neural network by an improved particle swarm optimization (PSO) to implement nuclide identification. It changes inertia weight and learning factor dynamically with self-adaption to optimize the weight value and threshold value of BP neural network. It gets the global optimal value of the particle swarm by training BP neural network to identify models. Finally, it implements nuclide identification by using the optimal weight and threshold value. The experiment shows our proposed method can not only converge to the optimal value faster but also do a good balance between local search and global search. Therefore, it significantly improves the convergence speed and the accuracy of nuclide identification.

Keywords: PSO; nuclide identification; global optimal point; inertia weight; learning factor

0 引言

粒子群优化算法 (particle swarm optimization, PSO) 最早是由美国电气工程师 Eberhart 和社会心理学家 Kennedy 于 1995 年发明的一种基于鸟类捕食原理的优化进化算法。该算法根据自定的适应函数决定的适应度不断调整粒子的速度和位置, 最终所有粒子趋于全局最优解, 具有结构简单、参数调整简单易行以及适合计算机编程处理等优点^[1]。

BP 神经网络具有良好的自学习能力和强大的非线性映射能力, 采用了基于梯度下降的 BP 算法, 具有良好的非线性映射能力、泛化能力和容错能力, 已在众多领域得到广泛应用。但 BP 神经网络存在学习效率低、收敛速度慢和易陷入局部极小等局限, 尤其是 BP 神经网络结构的选取对网络的学习能力和泛化能力有很大影响^[2]。粒子群算法作为最有潜力的神经网络训练方法, 具有收敛速度快、全局搜索能力强等优点^[3]。

为改善上述问题, 笔者引入一种改进粒子群优化算法 (improved particle swarm optimization, IPSO)

对 BP 神经网络进行优化, 并通过实验进行验证。

1 理论基础

1.1 基本粒子群算法

群鸟在觅食的过程中, 每只鸟的初始状态及飞翔的方向均为随机。随着时间的推移, 这些初始处于随机位置的鸟类通过群内相互学习、信息共享和个体不断积累寻觅食物的经验, 自组织积聚成一个群落, 并逐渐朝食物靠近。每只鸟能够通过一定经验和信息估计目前所处的位置对于能寻找到食物的价值, 即适应值; 每只鸟能够记住自己所找到的最好位置, 称之为局部最优 (Pbest)。此外, 还能记住群鸟中所有个体所能找到的最好位置, 称为全局最优 (Gbest), 整个鸟群的觅食中心都趋向全局最优移动, 通过鸟群觅食的位置不断移动, 即不断迭代, 可以使鸟群朝食物步步进逼。最终鸟群聚集在一个全局最优的位置。

在上述鸟类觅食的模型中, 每只鸟可被视为一个粒子, 所有的群体可被视为一个粒子群。若在一

收稿日期: 2015-12-27; 修回日期: 2016-02-02

作者简介: 刘议聪 (1991—), 男, 四川人, 硕士, 助理工程师, 从事智能信号处理与控制研究。

个 D 维的目标搜索空间中, 由 m 个粒子组成一个群体, 其中第 i 个粒子 ($i=1,2,\dots,m$) 位置表示为 $X_i=(x_i^1,x_i^2,\dots,x_i^D)$, 即第 i 个粒子在 D 维搜索空间中的位置是 X_i 。将代入 X_i 适应度函数即可得到其对应的适应度值, 算法中可以自定义粒子的筛选规则。选出规则中的最优粒子个体, 记其经历过的最好位置为 $P_i=(p_i^1,p_i^2,\dots,p_i^D)$, 整个粒子群体中的最优位置即为 $P_g=(p_g^1,p_g^2,\dots,p_g^D)$ 。粒子个体的速度即为 $v_i=(v_i^1,v_i^2,\dots,v_i^D)$ 。PSO 算法采用如下公式对粒子的位置和速度进行更新:

$$v_i^d = wv_i^d + c_1r_1(p_i^d - x_i^d) + c_2r_2(p_g^d - x_i^d); \quad (1)$$

$$x_i^d = x_i^d + \alpha v_i^d. \quad (2)$$

其中: $i=1,2,\dots,m$; $d=1,2,\dots,D$; w 是非负数, 称惯性权重; 学习因子 c_1 和 c_2 为非负常数; r_1 和 r_2 是 $[0,1]$ 范围内的随机数; α 为约束因子, 用于控制速度的权重。 $v_i^d \in [-v_{\max}^d, v_{\max}^d]$, 每个粒子 i 的飞翔速度 v_i 均被一个最大速度 $v_{\max}=(v_{\max}^1, v_{\max}^2, \dots, v_{\max}^D)$ 限制。如果粒子速度超过 v_{\max} , 则将其速度限制在 v_{\max}^d 。算法可自定义迭代终止条件, 一般达到预设最大迭代次数或粒子群到目前为止搜索到的最优位置满足目标函数的最小容许误差即可终止^[4]。

1.2 改进粒子群算法

由于 PSO 算法局部搜索能力较差, 搜索精度不高。同时算法存在容易早熟和粒子在算法后期易在全局最优解附近“振荡”的现象, 所以算法不能保证搜索到全局最优解。PSO 算法的惯性权重 w 的选择直接影响算法的收敛性, 是影响 PSO 算法行为和性能的关键所在。其中, 较大的 w 有利于提高算法的全局搜索能力, 粒子跨越幅度较大, 容易丧失局部寻优能力, 而较小的 w 会增强算法的局部搜索能力, 但全局搜索能力会减弱。在众多的改进算法中, 线性调整策略相对简单且收敛速度快; 因此, 被广泛采用。笔者引入一种非线性递减策略调整 w 的粒子群更新算法, 提出一种开口向上的 w 值递减策略, 具体表现形式如下式所示:

$$w^t = 2(w_{\text{start}} - w_{\text{end}})(t/t_{\text{max}})^2 + 4(w_{\text{end}} - w_{\text{start}})(t/t_{\text{max}}) + w_{\text{start}} \quad (3)$$

为了在全局搜索和局部搜索之间找到一个折中点, 并同时算法早期通过加快 w 的递减速度, 故在式 (3) 中加入一个非线性递减权重策略, 即平方项。 w 的变化轨迹如图 1。

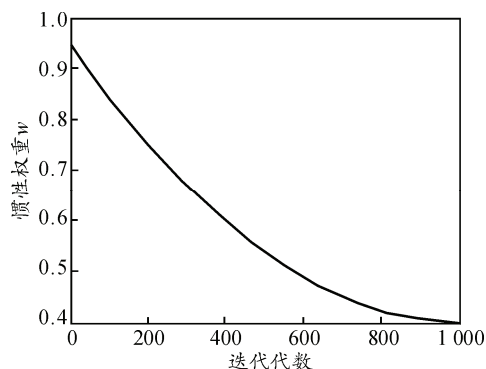


图1 惯性权重 w 的变化轨迹

开始阶段, w 较大, 目的是在迭代的初期加大 w 的递减速度以使算法保持其较强的探索能力, 并更快地进入局部搜索以均衡全局和局部搜寻能力。随着优化过程的进行, w 逐渐减小, 在达到最大迭代次数时迭代至设定的最小值。

学习因子 c_1 、 c_2 分别控制“认知”部分和“社会”部分对粒子速度的影响。合适的 c_1 和 c_2 可以加快收敛且不易陷入局部最优。在算法迭代初期, c_1 较大、 c_2 较小, 粒子保持自我认知能力较强, 即各粒子首先自我寻找最优位置, 故此时“自我”部分较大, 而群体的“社会信息”较小。随着优化过程的进行, c_1 逐渐减小, c_2 逐渐加大, 群体中的粒子之间进行信息交互意识逐渐增强。“社会信息”加大, 这样设置学习因子有利于算法收敛于全局最优解。其对应的曲线方程如下:

$$c_1^t = 2(c_{1\text{start}} - c_{1\text{end}})(t/t_{\text{max}})^2 + 4(c_{1\text{end}} - c_{2\text{start}})(t/t_{\text{max}}) + c_{1\text{start}} \quad (4)$$

$$c_2^t = 2(c_{2\text{start}} - c_{2\text{end}})(t/t_{\text{max}})^2 + 4(c_{2\text{end}} - c_{2\text{start}})(t/t_{\text{max}}) + c_{2\text{start}} \quad (5)$$

其中: t_{max} 为最大允许迭代次数; t 为当前的迭代次数; $c_{1\text{start}}$ 和 $c_{1\text{end}}$ 分别是 c_1 的初始取值和迭代到最大允许迭代次数时的取值。 c_1 、 c_2 的变化轨迹如图 2~图 3。

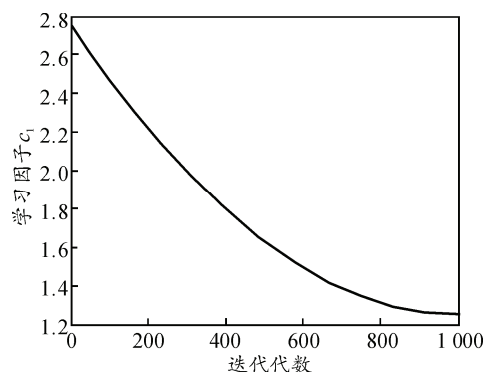


图2 学习因子 c_1 的变化轨迹

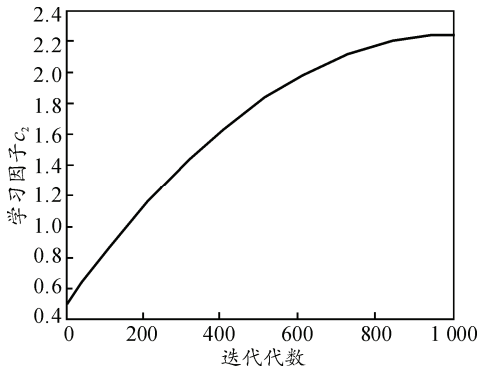


图 3 学习因子 c_2 的变化轨迹

2 IPSO 优化 BP 神经网络实现核素识别

利用 IPSO 算法优化 BP 神经网络, 关键是得出一组使 BP 神经网络能达到全局最优的网络权值 ω 和阈值 b , 寻找最优权值与阈值的流程如下:

1) BP 神经网络初始化、数据预处理;

2) 粒子群初始化, 包括种群粒子数、最大迭代次数、惯性权重和学习因子的初始取值、位置和速度的取值范围以及随机生成每个粒子的位置;

3) 根据具体情况设定适应度函数。将种群中的每个粒子位置作为待优化的 BP 神经网络权值和阈值, 输入训练样本, 对网络进行训练并测试, 并求出网络输出。种群中每个粒子的适应度值被定义为:

$$fit = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^m (y_i - t_k)^2 \quad (6)$$

其中: y_i 表示神经网络的实际输出; t_k 表示神经网络的目标输出; m 表示神经网络输出结点数; n 表示训练样本数。测试样本的测试输出与期望输出的均方误差, 适应度函数即为均方误差函数。故在文中算法适应度值与最优位置是成反相关关系;

4) 由式 3) 中得出的适应度值确定新的各粒子最优位置 P_{best} 和当前全局最优位置 G_{best} ;

5) 按照式 (1) 和式 (2) 更新粒子速度和当前位置。并判断其是否在 $[x_{min}, x_{max}]$ 及 $[v_{min}, v_{max}]$ 范围内, 若超出此范围, 则取其边界值, 同时更新惯性权重和学习因子;

6) 对较优位置按一定概率执行变异操作;

7) 判断是否到达终止条件, 是则结束, 否则返回第 3) 步进行下一次迭代;

8) 当算法达到最大迭代次数后, 将最终得到的群体最优位置 G_{best} 赋给 BP 神经网络的连接权值和阈值;

9) 输入测试数据, 对 BP 神经网络进行测试, 输出网络核素识别结果, 并计算正确识别率。

由于 Robert Hecht-Nielson 证明了具有一个隐含层的 3 层 BP 神经网络, 可以有效地逼近任意连续函数^[5], 故文中实验借助 Matlab 平台建立一个 3 层神经网络, 包含一个隐藏层。在实验中, 笔者对实验室测得的 20 组核素数据样本进行测试分析, 由于输入神经网络的样本中共有 7 种核素, 故输出共用 7 个节点, 输出量用 7×1 的列向量表示。文中约定核素排列的顺序为 ^{241}Am 、 ^{109}Cd 、 ^{238}Pu 、 ^{226}Rn 、 ^{133}Ba 、 ^{60}Co 、 ^{137}Cs 。在样本输出中, 若某个样本中存在某种核素, 则在输出矩阵中该核素对应的位置上元素为 1。例如, 若某个样本中存在 ^{137}Cs , 则其输出矩阵为 $[0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1]^T$; 若某个样本中存在 ^{241}Am 与 ^{60}Co , 则样本输出矩阵为 $[1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0]^T$ 。

3 实验结果对比与分析

3.1 实验结果

为了算法结果的可对比性, 本实验设置了 3 个模型分别实现核素识别: 传统 BP 神经网络识别模型、基本 PSO-BP 识别模型和 IPSO-BP 识别模型。传统的 BP 神经网络模型设置为 3 层, 其中有两层隐含层, 一层输入层, BP 神经网络结构确定为 20-10-7, 即输入层、隐含层、输出层的节点个数分别为有 20、10、7。基本 PSO-BP 的原输入空间数据为 1 024 维, 经归一化并进行 PCA 降维后, 提取出新的输入空间中 15 维主成分, 组成输入数据。BP 神经网络结构确定为 15-32-7, 学习因子 $c_1=c_2=1.494 \ 5$, 最大迭代代数为 1 000, 种群个数为 40。IPSO-BP 的原输入空间数据为 1 024 维, 经归一化并进行 PCA 降维后, 提取出新的输入空间中 15 维主成分, 组成输入数据。BP 神经网络结构确定为 15-32-7, 学习因子 $c_{1start}=2.75, c_{1end}=1.25, c_{2start}=0.5, c_{2end}=2.25, w_{start}=0.95, w_{end}=0.4$, 最大迭代代数为 1 000, 种群个数为 40。其预测结果数据如表 1。

传统 BP 神经网络实现核素识别的结果如图 4。

从图 4 中可以很直观地看出, 传统的 BP 神经网络在 20 组样本数据中出现了第 12 组样本识别错误的情况, 识别的准确率为 95%。

基本 PSO-BP 与 IPSO-BP 在寻优迭代过程中的最佳适应度表现如图 5 所示。

表 1 IPSO-BP 算法用于核素识别测试结果数据

样本序号	待识别核素	训练核素						
		²⁴¹ Am	¹⁰⁹ Cd	²³⁸ Pu	²²⁶ Rn	¹³³ Ba	⁶⁰ Co	¹³⁷ Cs
1	²⁴¹ Am	1.000 0	-0.000 2	0.000 8	0.000 1	0.000 7	0.000 9	0.000 2
2	²⁴¹ Am	1.000 1	0.000 4	-0.000 2	0.000 0	-0.000 3	-0.000 5	0.000 0
3	²⁴¹ Am	1.093 9	-0.146 3	0.105 6	0.070 7	-0.111 8	-0.128 0	0.035 8
4	²⁴¹ Am	1.001 3	-0.001 5	0.002 1	0.000 3	0.001 1	0.001 0	0.000 0
5	⁶⁰ Co	-0.246 6	-0.024 6	-0.028 7	0.020 0	-0.026 3	1.082 7	0.050 0
6	⁶⁰ Co	0.034 7	0.045 9	0.075 8	-0.025 9	0.069 9	0.977 1	-0.042 1
7	⁶⁰ Co	0.211 1	-0.049 9	-0.138 1	0.108 4	-0.078 3	0.914 4	-0.109 7
8	⁶⁰ Co	-0.132 1	0.038 8	-0.001 0	0.069 0	0.096 1	1.059 3	-0.023 2
9	⁶⁰ Co	0.149 1	0.064 3	-0.117 6	-0.024 3	-0.043 2	0.823 3	0.057 3
10	⁶⁰ Co	0.027 4	0.018 4	-0.043 8	-0.028 2	-0.051 8	0.951 3	-0.001 4
11	²⁴¹ Am	0.895 3	-0.123 0	-0.003 8	0.092 6	0.003 1	0.082 7	0.200 7
12	¹³⁷ Cs	0.101 3	-0.029 3	-0.026 8	-0.173 9	0.064 5	0.133 4	0.943 2
13	¹³⁷ Cs	-0.049 3	0.057 5	-0.055 6	0.009 5	-0.001 7	0.034 3	1.013 0
14	¹⁰⁹ Cd	0.000 1	0.999 7	0.000 5	-0.000 1	0.000 7	0.000 8	0.000 1
15	²³⁸ Pu	-0.000 2	-0.000 1	1.000 1	0.000 2	0.000 0	0.000 0	-0.000 2
16	²²⁶ Rn	-0.000 2	-0.000 5	0.000 9	0.999 9	0.000 9	0.000 8	0.000 4
17	¹³³ Ba	0.000 0	-0.000 1	-0.000 1	-0.000 1	0.999 7	0.000 0	0.000 2
18	²⁴¹ Am+ ¹⁰⁹ Cd	1.000 2	1.000 5	0.001 3	0.000 5	0.000 9	0.000 5	-0.000 5
19	¹⁰⁹ Cd+ ²³⁸ Pu	0.000 5	1.000 4	0.999 5	0.000 0	-0.000 9	-0.000 4	-0.000 3
20	²³⁸ Pu+ ¹³³ Ba+ ¹⁰⁹ Cd	0.000 4	0.999 7	1.000 1	0.000 4	1.000 2	0.001 0	0.000 2

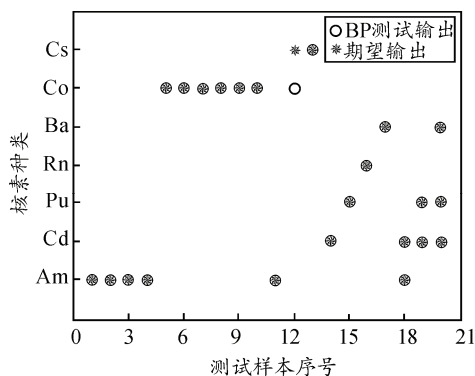


图 4 BP 神经网络预测结果

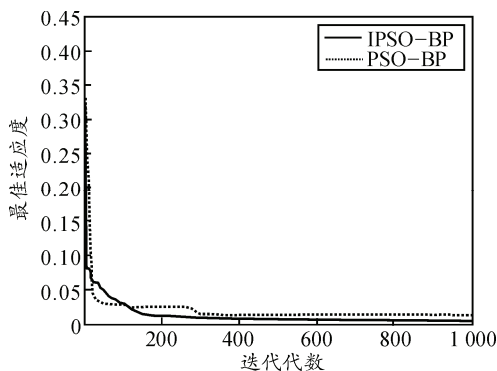


图 5 PSO-BP 与 IPSO-BP 预测结果精度对比

从图 5 中可以看出：PSO-BP 算法在迭代初期迅速收敛，早熟现象较为严重，在第 300 代时逐渐趋近于全局最优解，最佳适应度值为 0.013 9。相比于基本 PSO-BP 算法，IPSO-BP 算法能够较为平稳地收敛至全局最优解，不会出现早熟现象，在迭代至第 700 代以后，算法逐渐趋近于全局最优解，至第 1 000 代时，找到的最优适应度值不再变化，故

此最佳适应度值所对应的权值与阈值即可以确定为全局最优解，最佳适应度值为 0.004 7。

基本 PSO-BP 算法最终的核素识别结果如图 6 所示，识别准确率为 100%。

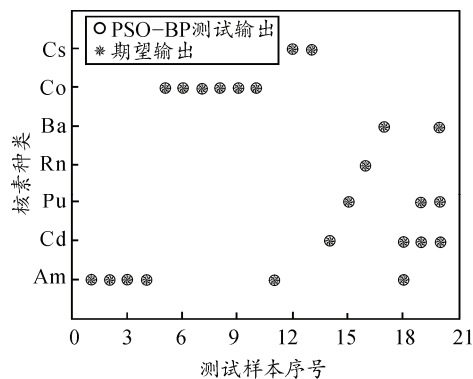


图 6 PSO-BP 算法的核素识别结果

IPSO-BP 算法最终的核素识别结果如图 7 所示，识别准确率为 100%。

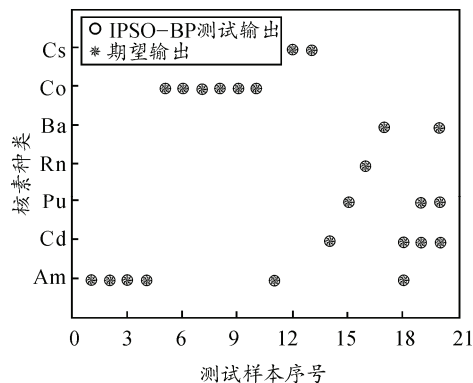


图 7 IPSO-BP 算法的核素识别结果

3.2 结果分析

为了比较文中方法(IPSO-BP 模型)、经过标准粒子群算法优化的 BP 神经网络模型和标准的 BP 模型之间的优点,笔者主要采用了均方误差(mean square error, MSE)作为评判指标。文中设置的最佳适应度值即为每代预测值的最小均方误差,从上述最佳适应度值表现图中可以看出:IPSO-BP 算法的最小均方误差明显小于 PSO-BP 算法,即 IPSO-BP 算法的预测精度明显高于 PSO-BP 算法。

4 结束语

笔者针对基本 PSO 优化 BP 神经网络容易存在早熟、不易达到全局最优等缺点,改进了其惯性权重 w 和学习因子 c_1 、 c_2 ,将其从经典的固定值改进为一种自适应的动态非线性递减策略调整 w 以及动态调整的学习因子。实验结果表明,IPSO 算法可以有效提高 BP 神经网络的识别精度。比传统的 PSO 在收敛速度、计算精度、全局最优解的搜索能力、

算法稳定性等方面有明显的优势,并且 IPSO-BP 算法与传统 BP 算法、标准 PSO-BP 算法的训练结果相比,具有更小的训练误差和检验误差,不仅能更快地收敛于全局最优解,而且提高了结果精度的数量级,在一定程度上改善了学习能力与泛化能力。

参考文献:

[1] 谢晓锋,张文俊,杨之廉. 微粒群算法综述[J]. 控制与决策, 2003, 18(2): 129-134.
 [2] 李祚泳,彭荔红. BP 网络学习能力与泛化能力满足的不确定关系式[J]. 中国科学(E 辑), 2003, 33(10): 887-895.
 [3] SHI Y, EBERHART R C. A modified particle swarm optimizer[C]//IEEE International Conference on Evolutionary Computation. Alaska: Anchorage, 1998: 69-73.
 [4] 卓金武. Matlab 在数学建模中的应用[M]. 2 版. 北京: 北京航空航天大学出版社, 2014: 151-152.
 [5] Robert Hecht-Nielsen. Theory of the Back Propagation Neural Network[C]. Orlando, Florida, USA: Proceedings of the International Joint Conference on Neural Networks, 1989: 593-605.

(上接第 87 页)

5 结束语

在综合考虑主客观赋权法的优缺点及评估对象特点的前提下,笔者选取了德尔菲法和熵值法,从主、客观 2 方面给出飞行机务保障能力评估指标权重,然后采用基于最小二乘法和熵值法的综合集成赋权法将主客观的权值进行适当综合,能够避免单一赋权方法在赋权中存在的局限性,给出更合理的评估指标权重。但是综合集成赋权法所采用的集成方法又有多种,并且很难确定哪种方法更可靠,目前大都根据不同实际而采用不同的集成方法,这都有待于进一步研究。

参考文献:

[1] 舒正平,周盛华. 装备维修保障效能评估指标权重组合赋权法的研究[J]. 兵工自动化, 2010, 29(12): 9-11.
 [2] 贾梦杰,孙荣平,董华,等. 机载设备技术保障能力评

估指标体系[J]. 四川兵工学报, 2013, 34(5): 62-64.
 [3] 韩书稳,任风云,李永辉. 装备保障能力评估指标体系研究[J]. 科技信息(学术版), 2006(5): 32-34.
 [4] 齐建,王新阁,郑刚. 航空兵飞行团机务保障能力评估指标体系[J]. 装备制造技术, 2010(4): 139-141.
 [5] 中国人民解放军海军装备部. 海军航空兵部队修理工作细则[S]. 2005.
 [6] 中国人民解放军海军装备部. 海军航空维修一线管理细则[S]. 2005.
 [7] 王秀华,李洲,杜汪洋. 提高部队装备战备水平的思考[J]. 装备指挥技术学院学报 2007, 18(2): 38-41.
 [8] 赵希锦,叶宏,邓新华,等. 基于熵值法的环境科技能力综合评价研究[J]. 四川环境, 2013(32): 138-142.
 [9] 周宇峰,魏法杰. 组合赋权法在企业绩效评价中的应用[J]. 工业与工程管理, 2007(4): 51-74.
 [10] 郝海,踪家峰. 系统分析与评价方法[M]. 北京: 经济科学出版社, 2007: 29-33, 75.
 [11] 王长青,张一农,许万里. 运用最小二乘法确定后评估指标权重的方法[J]. 吉林大学学报, 2010, 28(5): 513-518.