

Weibull 分布下基于遗传算法的设备寿命预测

邓炳杰, 陈晓慧

(重庆大学机械传动国家重点实验室, 重庆 400030)

摘要: 本文研究了 Weibull 分布下设备可靠性寿命预测的问题. 利用改进后的遗传算法, 主要是对遗传算法编码、目标函数和遗传操作的改进, 实现对 Weibull 参数的估计. 获得了 Weibull 分布模型和设备的可靠性寿命模型.

关键词: Weibull 分布; 遗传算法; 寿命预测

MR(2010) 主题分类号: 92D10

中图分类号: O119

文献标识码: A

文章编号: 0255-7797(2016)02-0385-08

1 引言

正确预测设备的剩余寿命^[1]对于保证设备安全运行、提高经济效益有着十分重要的意义. 潘冬等^[2]应用 Achard 磨损模型, 充分考虑了齿轮负载及转速对齿轮副齿面磨损的综合影响, 建立了齿轮磨损寿命预测模型, 但模型的实际应用价值没有体现. Xia 等^[3]建立了四参数的 Weibull 分布模型进行失效研究. Dragan Juki 等^[4]运用加权误差拟合方法实现了对 Weibull 参数的估计, 但缺少估计方法的对比说明. Tan 等^[5]使用极大似然估计进行二参数的 Weibull 分布的参数估计, 并用实例验证方法的可行性. 但对于常用的三参数的 Weibull 分布没有进行相应的研究. Zhou 等^[6]对几种常用的 Weibull 分布参数估计方法进行了说明和比较, 相比之下, 得出了极大似然估计方法具有比较优越的估计精度, 但是没有将智能算法与估计方法相结合, 导致计算难度大.

综上所述, Weibull 分布是在设备寿命的可靠性分析中最常用的概率分布之一, 对 Weibull 分布的参数估计可得出设备的可靠性寿命, 从而进行寿命预测^[2]. 但目前对 Weibull 分布的参数估计大多使用极大似然估计或最小二乘估计等, 运用这两种方法不仅存在计算难度大, 而且很难保证计算过程的收敛性等问题. 因此, 寻找新的智能方法进行 Weibull 参数估计是势在必行的.

遗传算法是一种模拟自然界的選擇与遗传的机理寻找最优解的优化方法. 吴俊杰等^[8]运用遗传算法寻找最优路线. 李伟等^[9]等对遗传算法的选擇算子进行改进, 实现了对 BP 神经网络的优化. 任子武等^[11]运用粒子群优化算法来优化遗传算法, 取得了不错的效果. 所以, 由于遗传算法不要求函数连续、具有可扩展性、具有潜在的并行性等优点, 已得到了广泛的运用. 并且, 可以根据实际情况, 对遗传算法进行相应的改进并取得非常有效的效果. 本文就是在 Weibull 分布的基础上, 运用改进的遗传算法和极大似然估计相结合, 进行 Weibull 分布

*收稿日期: 2013-09-11 接收日期: 2014-04-23

基金项目: 机械传动国家重点实验室自主项目资助 (SKLMT-ZZKT-2012 MS 02); 国家自然科学基金重点项目 (51035008).

作者简介: 邓炳杰 (1988-), 男, 重庆, 硕士, 主要研究方向: 主要从事寿命预测模型研究.

的参数估计, 进而预测设备的寿命. 由于遗传算法本身所具有的特点, 可以方便、简单的估计出 Weibull 分布的参数.

2 设备寿命分布模型的选择

2.1 Weibull 分布的优势和特点

在设备寿命的可靠性分析中, 常用的分布模型有 Weibull 分布、正态分布、对数正态分布、指数分布、瑞利分布等. 其中, Weibull 分布是可靠性分析中最常用的分布. 它是由瑞典物理学家 wallodi Weibull 于 1939 年引进, 是可靠性分析及寿命检验的理论基础. Weibull 分布^[3] 有很多种形式, 包括一参数、二参数、三参数或混合 Weibull. 三参数的 Weibull 分布是一种比较完善的分布, 在拟合随机数据方面十分灵活, 适应性很强. 因此, 三参数 Weibull 分布能更准确地表达出设备寿命的概率分布, 而且, 在可靠性研究领域中的几种常用分布, 如指数分布、瑞利分布等都可以看作是三参数 Weibull 分布的特例. 三参数的 Weibull 分布由形状、尺度 (范围) 和位置三个参数决定. 其中形状参数是最重要的参数, 决定分布密度曲线的基本形状; 尺度参数起放大或缩小曲线的作用, 但不影响分布的形状. 三参数 weibull 分布的分布函数为 (假设设备在时间 γ 前没有故障发生):

$$F(X) = 1 - \exp\left[-\left(\frac{x-\gamma}{a}\right)^\beta\right], \quad x \geq \gamma. \quad (2.1)$$

概率密度函数和失效率函数分别为

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{\alpha}\right)^\beta\right], \quad x \geq \gamma, \quad (2.2)$$

$$\lambda(x) = \left(\frac{\beta}{\alpha}\right) \left[\frac{x-\gamma}{\alpha}\right]^{\beta-1}, \quad x \geq \gamma, \quad (2.3)$$

其可靠度函数为

$$R(x) = \exp\left\{-\left[\frac{x-\gamma}{\alpha}\right]^\beta\right\}, \quad x \geq \gamma, \quad (2.4)$$

其中 $\alpha, \beta > 0, \gamma \geq 0$, α 为尺度参数, β 为形状参数, γ 为位置参数^[4].

Weibull 分布必须具备以下两个条件: 第一, 满足客观事物的随机性; 第二, 变量必须大于零. 设备的寿命是一个连续的随机变量, 其值大于零. 所以, 设备寿命分布符合 Weibull 分布的条件, 可以建立 Weibull 分布模型进行设备的寿命预测. 即可以得出设备寿命的 Weibull 分布模型, 其设备寿命损耗分布函数、概率密度函数和失效函数分别为 (2.1)-(2.3) 式所示^[5].

2.2 设备失效率与 Weibull 参数之间的联系

在失效率函数 (如公式 (2.3) 所示) 中, 由于 α 为尺度参数, 其大小反应为曲线图横纵坐标的放大或缩小, 它的变化不影响图形的形状, γ 是设备的初始损耗值. 所以, 在失效率公式中, 起非常重要作用的是 β 的值. 令 $\alpha = 1, \gamma = 0$ 可以模拟出设备寿命的浴盆曲线图, 如图 2.1 所示. 由图可知, 当 $\beta < 1$ 时, $\lambda(x)$ 随着时间的增大而减少, 设备运行在早期失效期 I, 设备的寿命损耗程度低. 当 $\beta = 1$ 时, $\lambda(x)$ 随着时间的增大基本不变, 设备运行在偶发失效期 II, 设备进入正常稳定工作. 当 $\beta > 1$ 时, $\lambda(x)$ 曲线呈递增分布, 当 $\beta = 3 \sim 4$ 时, $\lambda(x)$ 曲线与正态分布形状近似, 设备的寿命损耗程度上升, 设备运行在损耗失效期 III^[7].

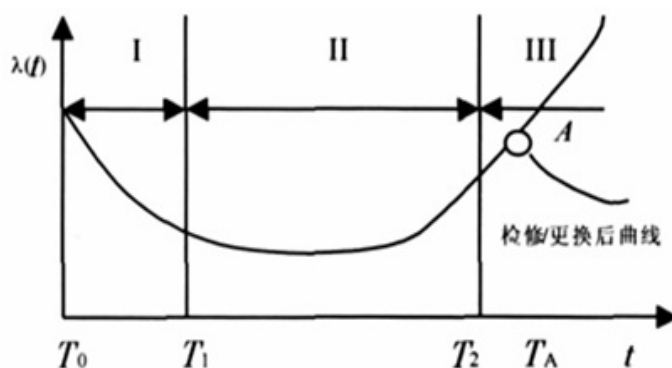


图 2.1: 设备寿命曲线图

3 应用遗传算法对 Weibull 参数进行估计

Weibull 分布的参数估计方法有: 极大似然估计、矩估计、最小二乘估计、线性回归估计等. 对二参数的 Weibull 分布模型, 以上方法都能取得比较满意的结果, 并且计算难度也不是很大. 但对于三参数的 Weibull 分布模型, 由于模型本身的复杂性, 运用上述方法时, 存在计算难度大, 估计精度不够理想等问题^[6]. 因此希望对它进行改进, 以达到减少计算和提高精度的双重效果. 遗传算法具有求解速度快、精度高等优点, 运用遗传算法和极大似然估计相结合, 可以达到满意的结果.

根据同种设备的历史故障失效数据以及 Weibull 分布的参数特点, 运用遗传算法与极大似然估计相结合, 就可确定 α, β, γ 的具体值, 从而得出设备的 Weibull 分布模型, 并且可计算出设备的可靠性寿命及各种相关指标. 根据具体的 Weibull 分布模型, 可评价设备的工况及获得对设计、工作和检修有实际意义的可靠性参数.

3.1 基本的遗传算法

遗传算法是一种采用生物进化理论进行选择问题的近似最优解的方法^[8]. 它具有不要求函数连续、可扩展性、潜在的并行性等优点, 已得到了广泛的运用^[9]. 它的实施过程包括编码、产生初始种群、计算适应度、选择、复制、交换、突变、反复迭代、终止等操作^[10]. 遗传算法的工作步骤是^[11]:

- (1) 确定个体的字符串的组成及长度.
- (2) 随机建立初始种群.
- (3) 计算各个体的适应度并选择优秀个体.
- (4) 根据遗传概率, 用下述操作产生新个体:
 - ① 复制. 将已有的优良个体复制后添入新群体中, 删除劣质个体.
 - ② 交换. 将选出的两个个体进行交换, 产生的新个体进入新群体.
 - ③ 突变. 随机地改变某个体的某一字符后, 将新个体添入新群体.
- (5) 反复执行 (3) 及 (4), 直到达到终止条件, 选择最佳个体作为遗传算法的结果.

群体的编码是遗传算法要解决的首要问题, 它直接决定着算法的迭代次数、是否收敛等问题. 传统的二进制编码字符串过长降低了算法的效率, 而且二进制编码本身也不直观、精度

也不高. 群体的适应度函数的确定是一个关键问题, 一般是以所要进行求解的目标函数作为遗传算法的适应度函数. 遗传操作主要包括选择、交叉和变异. 传统的选择是根据概率值来确定, 这种方法有可能破坏群里新个体的产生. 传统的交叉和变异都是事先设定好的, 不具有个体的交叉和变异概率随适应度函数值的变化而自动调整的能力.

3.2 改进的遗传算法

考虑到传统遗传算法存在一些不足, 在其基础上进行如下改进:

3.2.1 编码方案的改进

将染色体设计成二层递阶结构. 在生物医学领域中, 染色体是由控制基因和序列基因组成, 控制基因表明此染色体的作用和功能, 而序列基因用来实现它的作用和功能. 控制基因采用二进制编码, 1 表示下层基因被激活, 其中的序列基因参加遗传操作. 0 表示下层基因未被激活, 其中的序列基因不参与任何遗传操作. 序列基因采用实数编码. 染色体的结构如图 3.1 所示. 图 3.1 表示的染色体由 5 个控制基因和 5 个整数表示的序列基因组成, 其表示为 $X = (4, 5, 7, 8)$. 这样的染色体结构, 既可以缩短染色体结构的长度, 也可以通过控制基因来确定变异的概率, 保护优秀的基因.

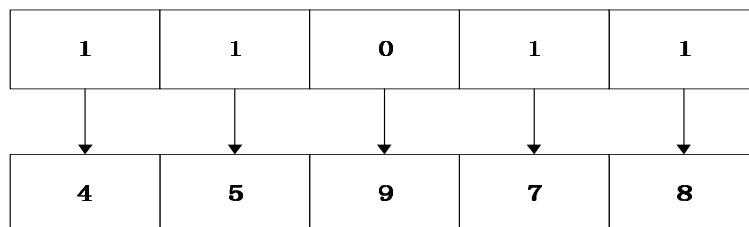


图 3.1: 双层染色体结构图

3.2.2 遗传算法目标函数的建立

为了实现用遗传算法进行参数估计, 用极大似然估计与遗传算法相结合来构造遗传算法的目标函数如下:

$$\max F(\theta) = \ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta), \quad (3.1)$$

等价于

$$\min F(\theta) = -\ln L(\theta) = -\sum_{i=1}^n \ln f(x_i; \theta), \quad (3.2)$$

其中 $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ 为待估参数.

3.2.3 遗传操作的改进

(1) 采取保留最优个体的选择方法^[12]. 由于遗传算法本身选择具有随机性, 即使最优良个体也可能在选择算法中被淘汰, 而最差的个体也可能被选择. 所以, 在遗传算法中采用保留

最优个体的选择方法^[13], 即每代都保留几个最优秀的个体不参与选择和变异操作, 但参与交叉操作, 这样能保证优秀基因传递给下一代, 从而群体整体性能得到提升.

(2) 交叉概率 P_C 和变异概率 P_M 随适应度函数值自动改变. 当个体适应度值低于整体平均适应度值时, 表明该个体性能相对较差, 需采用较大的 P_C 和 P_M . 适应度值越大, P_C 和 P_M 就越小. 具体取值如下所示:

$$\begin{cases} P_C = P_{C1} - (P_{C1} - P_{C2})(f - \bar{f}) / (f_{\max} - \bar{f}) \cdots f \geq \bar{f}, \\ P_C = P_{C1} \cdots f < \bar{f}, \\ P_M = P_{M1} - (P_{M1} - P_{M2})(f_{\max} - \bar{f}) / (f_{\max} - \bar{f}) \cdots f \geq \bar{f}, \\ P_M = P_{M1} \cdots f < \bar{f}, \end{cases} \quad (3.3)$$

其中 \bar{f} 为群体得平均适应度值, f_{\max} 为群体的最大适应度值, f 为个体的适应度值, $P_{C1} = 0.9, P_{C2} = 0.6, P_{M1} = 0.1, P_{M2} = 0.001$. 这样 P_C 和 P_M 的值随适应度函数值的变化而变化, 再加上最优个体保存策略, 即可以帮助遗传算法快速收敛, 又可以避免陷入局部最优解.

(3) 终止条件的改进. 传统的遗传算法一般采用遗传迭代次数来终止算法. 这样, 即有可能陷入局部最优、也不能确保误差的大小, 所以对遗传终止条件进行改进 - 即当两轮遗传操作的最优适应度相对误差小于 5%, 遗传操作结束. 为尽可能避免陷入局部最优, 还加入了遗传次数的限制, 即同时满足遗传次数和误差条件才终止遗传算法^[14,15].

3.3 遗传算法估计 Weibull 参数的具体过程

运用遗传算法进行 Weibull 参数估计的具体流程如图 3.2 所示.

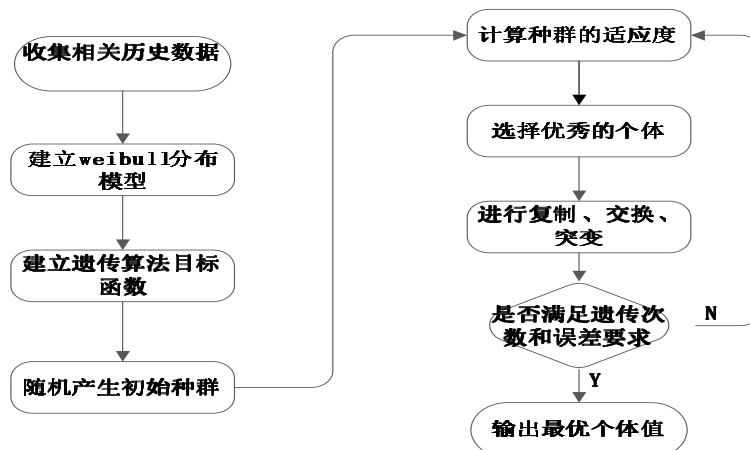


图 3.2: 遗传算法估计 weibull 参数的具体流程图

4 设备寿命预测实例

Weibull 分布下基于遗传算法的设备寿命预测的具体步骤如下:

- (1) 建立符合三参数的设备 Weibull 分布寿命模型, 确定其概率密度函数.
- (2) 采集同种设备的历史失效数据.
- (3) 运用改进后的遗传算法和历史失效数据对 Weibull 参数进行估计.

(4) 根据参数的估计值, 判断此种设备的寿命损耗程度及可靠性等信息.

表 4.1 是管道疲劳寿命失效数据统计表 (数据来源于文献 [16]), 管道失效数据 (从管道疲劳裂纹出现到不满足使用要求为止) 以某院研制的埋地油气漏磁检测装置的检测数据为准. 文献 [16] 采用极大似然估计的方法, 通过一定的计算, 建立了比较满意的模型, 得出了比较精确的结果. 为进一步减少计算量和提高精度, 本文运用改进后的遗传算法建立 Weibull 分布模型. 具体过程如下:

表 4.1: 管道疲劳寿命失效数据统计表 (单位: 天)

序号	1	2	3	4	5	6	7	8	9
时间	160	178	196	218	231	246	255	286	326
序号	10	11	12	13	14	15	16	17	18
时间	358	408	446	512	643	690	700	712	740

采用极大似然估计的理念, 通过 Matlab 语言编程, 运用传统的遗传算法和改进后的遗传算法分别求出参数值, 其优化迭代如图 4.1 和图 4.2 所示:

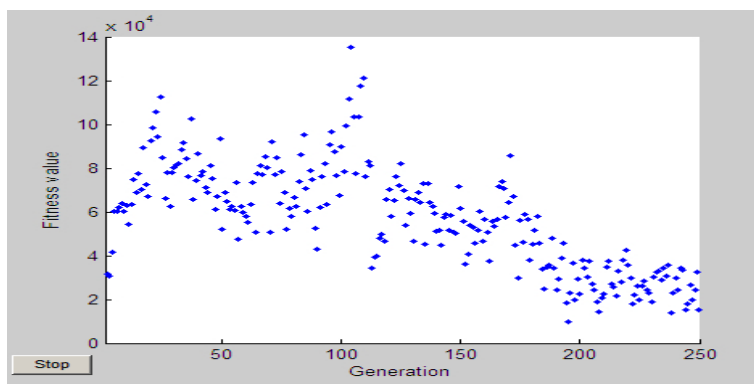


图 4.1: 传统遗传算法优化迭代图

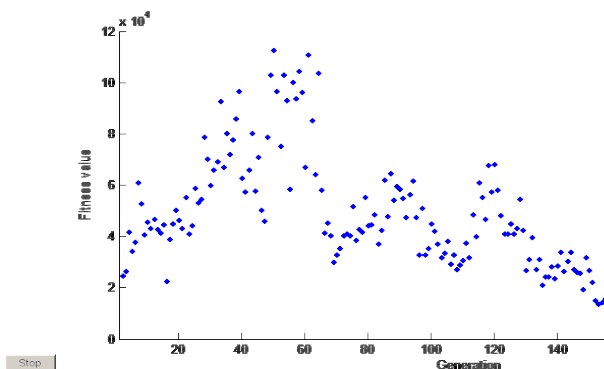


图 4.2: 改进后的遗传算法优化迭代图

两者求出的结果及文献 [16] 的结果如表 4.2 所示:

表 4.2: 结果显示表

	传统遗传算法	改进的遗传算法	文献 [16]
参数具体值	$\alpha = 215, \beta = 3.12,$ $\gamma = 149$	$\alpha = 290, \beta = 3.24,$ $\gamma = 138$	$\alpha = 274, \beta = 3.08,$ $\gamma = 157$
迭代次数	250	223	-
目标函数的最大值	-192.23	-147.22	-152.83

由表 4.2 可知, 传统的遗传算法所需要的迭代次数较多, 并且求出的结果精度不高, 而运用改进后的遗传算法, 不仅目标值的波动范围减少、算法的迭代次数有所减少, 而且还获得了更好的最优个体, 并且求得的精度高于文献 [16], 计算量也得以减小.

将运用改进后的遗传算法所求得的参数值代入公式 (2.1)–(2.4), 可求得 Weibull 分布模型为

$$F(t) = 1 - \exp \left[- \left(\frac{t - 138}{290} \right)^{3.24} \right], t \geq 138.$$

概率密度函数和失效率函数分别为

$$f(t) = 0.011 \left(\frac{t - 138}{290} \right)^{2.24} \exp \left[- \left(\frac{t - 138}{290} \right)^{3.24} \right], t \geq 138,$$

$$\lambda(t) = 0.011 \left[\frac{t - 138}{290} \right]^{2.24}, t \geq 138.$$

可靠度函数为

$$R(t) = \exp \left\{ - \left[\frac{t - 138}{290} \right]^{3.24} \right\}, t \geq 138.$$

至此, Weibull 分布模型已经成功建立, 根据此模型, 可以方便的计算出管道运行的失效概率和可靠性寿命. 假设管道在疲劳裂纹出现后工作了 300 天, 则其失效概率为 $\lambda(300) = 3 \times 10^{-3}$, 可靠度函数 $R(300) = 85.9\%$. 若要保证管道疲劳裂纹缺陷引起的疲劳失效的可靠度达到 85.9%, 则疲劳裂纹出现后, 缺陷管道运行的时间为 300 天左右. 即疲劳裂纹出现后, 工作 300 天左右, 就需要对管道进行挖掘检修或更新, 以保证油气输送的安全性. 有了定量分析, 极大地方便了工作人员检修或更换设备, 节约了成本、提高了效率, 为企业带来了相当可观的利润.

5 结论

本文运用改进后的遗传算法, 估计 Weibull 分布的相关参数, 从而获得设备的可靠性寿命的相关指标. 根据极大似然函数的理念构建遗传算法的目标函数, 合理利用遗传算法的特点, 解决了极大似然估计的计算难度大、不易收敛等缺点. 在此基础上, 对遗传算法本身的选择和终止条件的改进, 可以在一定程度上解决遗传算法本身的随机性选择和陷入局部最优的缺陷. 但是, 本文只对单一的 Weibull 分布进行相应的研究, 没有对多重 Weibull 混合模型进行相应的参数估计等相关研究, 这是本文存在的不足之处.

参 考 文 献

- [1] 张小丽, 陈雪峰等. 机械重大装备寿命预测综述 [J]. 机械工程学报, 2011, 47(9): 100–116.
- [2] 潘冬, 赵阳等. 齿轮磨损寿命预测方法 [J]. 哈尔滨工业大学学报, 2012, 44(9): 29–33.
- [3] Xia Zhaopeng, Yu Jianyong. Studying the mechanical properties of jute/cotton blended yarns using the Weibull mode[J]. J. Donghua Univ., 2009, 4: 393–396.
- [4] Dragan Juki, Darija Markovi. On nonlinear weighted errors-in-variables parameter estimation problem in the three-parameter Weibull model[J]. Appl. Math. Comput., 2010, 215: 3599–3609.
- [5] Tan Zhibin. A new approach to MLE of Weibull distribution with interval data[J]. Reliab. Engin. Sys. Saf., 2009, 94: 394–403.
- [6] Zhou Dan. Comparison of parameter estimation methods for transformer Weibull lifetime modelling[J]. 高电压技术, 2013, 39(5): 1170–1177.
- [7] 刘树新, 刘长武等. 基于损伤多重分形特征的岩石强度 Weibull 参数研究 [J]. 岩土工程学报, 2011, 33(11): 1786–1791.
- [8] 吴俊杰. 船体装配线划线优化遗传算法研究 [J]. 大连理工大学学报, 2012, 52(3): 381–386.
- [9] 李伟. 基于粗糙集和改进遗传算法优化 BP 神经网络的算法研究 [J]. 西北工业大学学报, 2012, 30(4): 601–606.
- [10] 刘大莲, 徐尚文. 求解约束优化问题的内外交叉遗传算法 [J]. 系统工程理论与实践, 2012, 32(1): 189–195.
- [11] 任子武, 伞治. 实数遗传算法的改进及性能研究 [J]. 电子学报, 2007, 35(2): 269–274.
- [12] 赵亮, 吕剑虹. 基于改进遗传算法的风电场多目标无功优化 [J]. 电力自动化设备, 2010, 30(10): 84–88.
- [13] 宋以胜. 基于收缩精度的遗传算法逐级进化策略 [J]. 解放军理工大学学报, 2010, 11(3): 343–347.
- [14] 范青武, 王普. 遗传算法交叉算子的实质分析 [J]. 北京工业大学学报, 2010, 36(10): 1328–1336.
- [15] 崔建国, 赵云龙. 基于遗传算法和 ARMA 模型的航空发电机寿命预测 [J]. 航空学报, 2011, 32(8): 1506–1511.
- [16] 宋志强, 李著信. 基于 Weibull 分布的埋地油气管道疲劳寿命分析与评估 [J]. 机床与液压, 2011, 39(7): 130–133.

THE EQUIPMENT LIFE PREDICTION BASED ON WEIBULL DISTRIBUTION AND GENETIC ALGORITHM

DENG Bing-jie, CHEN Xiao-hui

(State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400030, China)

Abstract: In this paper, we predict the equipment life reliability which is based on Weibull distribution. By using the improved GA which includes the improvement of the coding scheme, the establishment of the objective function, the improvement of the genetic operation, the problem of parameter estimation of Weibull distribution is resolved, then the Weibull distribution model and the equipment reliability life model are established.

Keywords: Weibull distribution; genetic algorithm; the prediction of life

2010 MR Subject Classification: 92D10