

DOI:10.12158/j.2096-3203.2021.06.024

大数据背景下考虑删失特点的继保设备运行状态评估

张雷¹, 王光华¹, 李金铄², 耿宏贤², 戴志辉²(1. 国网河北省电力有限公司保定供电分公司, 河北 保定 071000;
2. 华北电力大学电气与电子工程学院, 河北 保定 071003)

摘要:未能记录到设备失效准确时间的失效数据称为删失数据。针对大数据背景下继保设备失效数据存在删失的特点,提出了大数据背景下考虑删失数据的继保设备运行状态评估方法。首先,通过对继保设备失效数据进行特点分析,对失效数据进行预处理,利用期望-最大化(EM)算法并结合指数分布和威布尔分布模型,估计继保设备失效模型的参数。其次,将参数估计值代入失效模型中,得到设备的可靠度、故障概率密度、失效率、平均无故障时间等可靠性指标。然后,通过仿真验证对比分析偶然失效期和老化失效期内不同估计方法得到的模型参数精度,验证了文中方法处理删失数据的有效性。最后,以某型号继保设备为例分析可靠性指标,验证了利用文中方法规划设备检修周期的可行性。

关键词:删失数据;期望-最大化(EM)算法;运行状态评估;大数据;继保设备;检修周期

中图分类号:TM77

文献标志码:A

文章编号:2096-3203(2021)06-0185-08

0 引言

继电保护是电力系统的第一道防线,有效利用继电保护运行数据对继保设备状态进行评估,是保持电网安全稳定运行的有效途径^[1-2]。

针对继保设备的状态评估,国内外做了大量研究。从继保设备的原理、硬件、配置等方面进行分析、建模、仿真试验,然后对其可靠性进行分析,为设备运行状态评估提供理论依据,也为运维人员提供了提高继保设备可靠性的方法和策略参考^[3-4]。例如文献[5]详细介绍了继保设备可靠性研究现状和发展趋势,并指出设备可靠运行是保证电网稳定的重要因素。文献[6]指出设备本身、信息系统和人为因素是保护系统风险的主要来源,并进一步分析了继保设备故障风险对一次系统的影响。文献[7]提出了一种基于变权重模糊综合评价的继保设备状态评价方法,其指标更加完善,但需要足够多的失效样本,而继保设备本身可靠性较高,失效样本数量较少,因此难以满足精度要求。对此,文献[8]在三参数威布尔分布模型的基础上结合灰色模型,在小样本条件下对继保设备进行寿命评估,取得了精度较高的计算结果。文献[9]从扩充样本的角度出发,使用蒙特卡洛法进行抽样,提高了小样本情况下评估结果的稳定性。文献[10]考虑了继保设备缺陷数据的随机截尾特征,基于两参数威布

尔分布,采用极大似然估计(maximum likelihood estimation, MLE)实现参数估计,但未考虑区间删失数据的影响。整体上,目前继保设备运行状态评估方法很少计及删失数据的特点,且对包含删失数据的保护运行“大数据”未充分挖掘,造成了信息损失、可靠性评估结果偏颇和分析模型假设过多的问题。围绕电力大数据的分析与利用,国内外学者已进行了大量研究^[11-12]。但在大数据背景下,对继保设备的失效数据进行挖掘、分析,从而为判别继保设备状态、评估其运行水平提供可行途径的相关研究,还较少见到。研究基于大数据的继保设备可靠性分析与运行状态评估方法具有重要意义。

文中结合继保设备运行数据特点,首先基于期望-最大化(expectation-maximization, EM)算法对失效模型的参数进行估计,并计算得到设备的可靠性指标;其次,对比各失效期内不同估计方法得到的模型参数精度,验证了方法处理删失数据的有效性;最后,通过算例验证了利用文中方法规划设备检修周期的可行性。

1 大数据背景下失效数据特点分析

长期实践经验表明,设备失效率与累计运行时间之间的函数关系可用浴盆曲线表示。设备的全寿命周期由早期失效期、偶然失效期与老化失效期组成^[13]。对于继保设备,出厂前会经过充分的测试以排除隐患,现场安装后,运维人员也会对设备进行严密的调试。因此,可认为继保设备投运后其运行状态便处于偶然失效期。继保设备的失效数据

收稿日期:2021-06-05;修回日期:2021-08-02

基金项目:国家自然科学基金资助项目(51877084);国家电网有限公司科技项目(SCHEBD00KZJS1800301)

一般是在其投入运行之后收集的现场数据。

受观测条件与监测手段的限制,现场收集到的失效数据可能包含 4 种类型。

(1) 精确数据。继保设备的自检功能日益完善,自检周期非常短,因此自检发现缺陷的时间即可认为是缺陷发生的准确时间。此外,在周期较短的巡视过程中的缺陷发现时间也可认定是缺陷发生时间。这类失效数据称为精确数据,如图 1 中编号为 1、2 的失效数据,线段的左端点表示设备的投运时间,右端点表示缺陷发生时间。

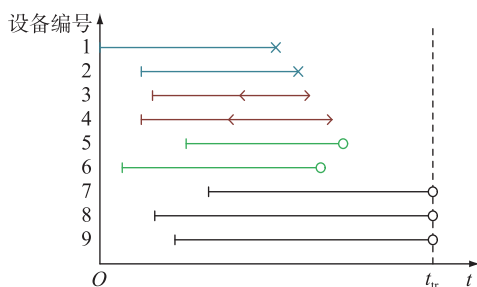


图 1 原始失效数据
Fig.1 Raw failure data

(2) 区间删失数据。部分隐蔽性较强的缺陷无法自检查出,无法确定具体发生时间,如专业巡视时发现的“复归按钮脱落”“备用线圈断线”等缺陷,只可推断出发生在本次巡视前、最近一次巡视之后。文中称此类无法确定缺陷发生的确切时间,只能推断时间区间的失效数据为区间删失数据。如图 1 中编号为 3、4 的失效数据,线段的左端点表示设备的投运时间,尖括号之间的线段表示缺陷发生时间所在的区间。

(3) 右删失数据。在实际工作情况下收集继保设备的失效数据,往往伴随着各种偶然事件。在观察结束之前,有可能出现样本设备中途退出运行的情形。设备在退出运行之前始终保持正常工作,只能得到设备保持正常工作的时长,此类数据称为右删失数据。如图 1 中编号为 5、6 的失效数据,线段的左端点表示设备的投运时间,右端点表示设备退出运行的时间。

(4) 随机截尾数据。继保设备的可靠性水平较高,因此直到观察结束,大部分设备仍能保持正常工作。此类直至观察结束仍未出现失效的样本数据称为截尾数据^[10],如图 1 中编号为 7、8、9 的失效数据,线段的左端点表示设备的投运时间,右端点表示观察结束的时间,即截尾时间 t_{tr} 。

因各继保设备的投运时间不同,所以当到达 t_{tr} 时,各继保设备的运行时间也不尽相同。如果把各样本设备的投运时刻移动到同一计时起点,这些设

备的失效数据将表现出随机截尾的特点,见图 2。

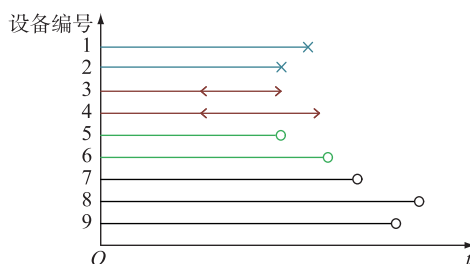


图 2 计时起点对齐后的失效数据
Fig.2 Failure data after starting time is aligned

针对继保设备失效数据的上述特点,常用的最小二乘法等算法已无法满足设备失效模型参数估计的需求。文中结合 EM 算法,利用带有删失特性的失效数据,估计继保设备失效模型的参数。

2 继保设备失效分布模型

失效分布模型能反映设备所处运行阶段及设备的运行状态。采用指数分布模型和威布尔分布模型,通过计算分布模型的参数,估计设备的可靠度 R 、故障率 λ 、平均无故障时间 t_{MTBF} 等指标。其中,指数分布反映偶然失效期的失效特征,常用于拟合具有恒定失效率的设备寿命分布,运行的继保设备多处于该失效期。威布尔分布则可通过不同的形状和尺度参数反映早期失效期(失效率递减)、偶然失效期和损耗失效期(失效率呈加速趋势,如继保设备的老化期)特征,具备较强的灵活性。

2.1 指数分布

指数分布模型^[14]的可靠性指标如下。

可靠度函数:

$$R(t) = e^{-\lambda t} \quad (1)$$

式中: t 为设备失效时间。

故障概率密度函数:

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (2)$$

失效率:

$$\lambda(t) = \lambda \quad (3)$$

平均无故障时间:

$$t_{MTBF} = 1/\lambda \quad (4)$$

2.2 威布尔分布

威布尔分布模型^[15-17]可分成三参数和两参数 2 种形式。两参数模型的函数表达式如下。

可靠度函数:

$$R(t) = e^{-(t/\eta)^k} \quad (5)$$

式中: $\eta > 0$ 为尺度参数; $k > 0$ 为形状参数。

故障概率密度函数:

$$f(t) = \begin{cases} (k/\eta) (t/\eta)^{k-1} e^{-(t/\eta)^k} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (6)$$

失效率:

$$\lambda(t) = (k/\eta) (t/\eta)^{k-1} \quad (7)$$

平均无故障时间:

$$t_{\text{MTBF}} = \eta \Gamma(1 + 1/k) \quad (8)$$

式中: $\Gamma(\cdot)$ 为伽马函数。

3 继电设备运行状态评估

3.1 EM 算法

EM 算法的基本思想是:当失效数据样本中存在删失数据时,若已知分布模型的参数,则可以结合已知样本数据估计删失数据的数学期望。反之,若已知删失数据的值,则可利用 MLE 估计分布模型的参数^[18]。

EM 算法流程简述如下。(1) 输入包括删失数据的样本数据;(2) 对分布模型的参数赋初值,并设定可接受的最大参数估计误差;(3) 执行 E 步:将上次迭代得到的参数估计结果代入分布模型,估计删失数据的概率分布,并计算删失数据服从此概率分布下对数似然函数的数学期望;(4) 执行 M 步:对参数寻优,使上述期望达到最大,将寻优结果作为本次参数估计的结果并计算估计误差;(5) 若参数估计误差小于可接受的最大误差,计算结束,否则返回步骤(3)。

3.2 基于 EM 算法的指数分布(EM-exp)模型参数估计

指数分布的故障概率密度函数如式(2)所示,其对数似然函数的数学期望为:

$$Q(\theta | \theta') = n \ln \lambda - \lambda \sum_{i=1}^n E(t_i | z, \lambda') \quad (9)$$

式中: n 为样本总数; t_i 为第 i 个样本的真实失效时间; z 为发生数据删失的样本集; θ, θ' 分别为不同分布模型中本次和上次迭代得到的参数;对指数分布, λ', λ 分别为上次和本次估计的故障率参数; $Q(\cdot)$ 为似然函数的数学期望; $E(\cdot)$ 为 t_i 在服从上次估计的故障率参数下对数似然函数的数学期望。

当 $Q(\theta | \theta')$ 取得极大值时,有:

$$\frac{\partial Q}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n E(t_i | z, \lambda') = 0 \quad (10)$$

根据 EM 算法的收敛性质,当迭代次数 $\omega \rightarrow +\infty$ 时, λ 收敛于定值,即:

$$\lambda = \lambda' \quad \omega \rightarrow +\infty \quad (11)$$

迭代次数足够多时,将式(10)中的各数学期望项展开,并将式(11)代入式(10),得:

$$\frac{n}{\lambda} = \sum_{i=1}^j t_i + \sum_{i=j+1}^{j+l} \frac{[-(t+1/\lambda)e^{-\lambda t}]_{Z_i^-}^{Z_i^+}}{[1 - e^{-\lambda t}]_{Z_i^-}^{Z_i^+}} + \sum_{i=j+l+1}^n \left(Z_i + \frac{1}{\lambda} \right) \quad (12)$$

式中: j 为失效时间是精确信息的样本数; l 为失效时间是区间删失数据的样本数,删失区间为 $[Z_i^-, Z_i^+]$; Z_i 为发生数据删失的样本集中的第 i 个元素。失效时间是右删失数据或随机截尾数据的样本数为 $n-j-l$ 。

采用数值方法求解上述方程即可得到指数分布模型的参数估计值。

3.3 基于 EM 算法的威布尔分布(EM-wbl)模型参数估计

威布尔分布的故障概率密度函数如式(6)所示。为简化计算,令 $\mu = 1/\eta^k$, 则式(6)可改写为:

$$f(t) = \begin{cases} k\mu t^{k-1} e^{-\mu t^k} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (13)$$

威布尔分布模型对数似然函数的数学期望为:

$$Q(\theta | \theta') = n \ln k + n \ln \mu + (k-1) \sum_{i=1}^n E(\ln t_i | z, k', \mu') - \mu \sum_{i=1}^n E(t_i^k | z, k', \mu') \quad (14)$$

式中: k', μ' 分别为上次迭代所得威布尔分布的形状参数和尺度参数估计结果; k, μ 为本次迭代估计的参数。

当 $Q(\theta | \theta')$ 取得极大值时,有:

$$\frac{\partial Q}{\partial \mu} = \frac{n}{\mu} - \sum_{i=1}^n E(t_i^k | z, k', \mu') = 0 \quad (15)$$

$$\frac{\partial Q}{\partial k} = \frac{n}{k} + \sum_{i=1}^n E(\ln t_i | z, k', \mu') - \mu \frac{\partial}{\partial k} \sum_{i=1}^n E(t_i^k | z, k', \mu') = 0 \quad (16)$$

当迭代次数 $\omega \rightarrow +\infty$ 时,参数 k, μ 将收敛于定值,即:

$$\begin{cases} k = k' \\ \mu = \mu' \end{cases} \quad \omega \rightarrow +\infty \quad (17)$$

迭代次数足够多时,将式(15)、式(16)中的各数学期望项展开,并将式(17)代入式(15)、式(16),得:

$$\frac{n}{k} = \sum_{i=1}^j [(\mu t_i^k - 1) \ln t_i] + \sum_{i=j+1}^{j+l} \frac{[-(\mu t^k \ln t + 1/k) e^{-\mu t^k}]_{Z_i^-}^{Z_i^+}}{[1 - e^{-\mu t^k}]_{Z_i^-}^{Z_i^+}} + \sum_{i=j+l+1}^n (\mu z_i^k \ln z_i + 1/k) \quad (18)$$

$$\frac{n}{\mu} = \sum_{i=1}^j t_i^k + \sum_{i=j+1}^{j+l} \frac{[-(t_i^k + 1/\mu)e^{-\mu t_i^k}]_{z_i^+}}{[1 - e^{-\mu t_i^k}]_{z_i^+}} + \sum_{i=j+l+1}^n (z_i^k + 1/\mu) \quad (19)$$

求解上述隐函数方程组,即可求得威布尔分布参数 k, μ 的数值解。

传统 MLE 在处理删失数据时只计及了随机截尾数据,不能处理区间删失数据,会影响继电设备运行状态评估的准确性。而基于 EM 算法的失效分布模型,在考虑区间删失数据的基础上,可通过对删失数据的还原处理,得到理想的参数估计值并据此计算设备的可靠度、故障率、平均无故障时间等指标,从而实现继电设备运行状态的评估。

4 算例分析

为检验模型的正确性,利用计算机生成带有删失特性的失效数据进行分析。首先,使用蒙特卡罗法生成服从指数分布和威布尔分布的精确失效数据^[19-21],并以此为基础生成带有删失特性的样本集。为模拟区间删失数据,随机选取部分样本进行删失处理。在失效数据的左右两侧分别扩充出删失区间,两侧的删失区间长度服从参数为 ζ_2, σ_2 的正态分布,且相互独立。

为模拟右删失数据,在非区间删失样本中随机选取部分样本数据,令其删失时间 t_c 服从 (t_0, t_r) 上的均匀分布,若删失发生在失效之后,则重新分配小于失效时间的 t_c 。其中, t_0 为投运时间。

为模拟失效数据的随机截尾特征,令计时起点为 $t_0 = 0$,各样本设备的投运时间服从参数为 ζ_1, σ_1 的正态分布。若设备的投运时间在计时起点之前,则重新分配大于 t_0 的投运时间。失效时间大于 t_r 的样本数据即为截尾数据。

4.1 偶然失效区

继电设备工作于偶然失效期时,其失效率基本保持不变。使用计算机生成服从指数分布的失效数据,样本容量为 100。生成失效数据时所用参数如表 1 所示。

表 1 参数表(偶然失效期)

Table 1 Parameters table (random failure)

参数	数值
$1/\lambda$	60 000
ζ_1, σ_1	500, 500
ζ_2, σ_2	1 000, 1 000
t_r	30 000

分别使用 EM-exp 模型、EM-wbl 模型和传统基

于 MLE 的指数分布模型(MLE-exp)进行参数估计。由于 MLE-exp 方法不能处理区间删失数据,所以在使用此类数据时,选取缺陷发现时间,即删失区间的右端点作为设备的失效时间。参数估计结果如表 2 所示,表中 p_1 为使用计算机生成数据时,精确数据在所有非截尾数据中所占比例; p_2 为区间删失数据所占比例; p_3 为右删失数据所占比例; t_{cal} 为计算耗时。

表 2 偶然失效区参数拟合结果

Table 2 Parameter fitting results of random failures

失效模型	不同 (p_1, p_2, p_3) 下的拟合结果				
	(1,0,0)	(0.8,0.1,0.1)	(0.7,0.2,0.1)	(0.6,0.3,0.1)	
EM-exp	$\lambda/10^{-5}$	1.669 5	1.659 1	1.646 5	1.574 3
	t_{cal}/s	0.002 21	0.013 70	0.013 00	0.029 00
	t_{MTBF}/h	59 897	60 273	60 734	63 521
EM-wbl	k	1.000 7	1.016 7	1.090 0	1.126 6
	$(1/\eta)/10^{-5}$	1.670 5	1.716 1	1.778 6	1.740 2
	t_{cal}/s	0.156 0	0.312 5	0.594 0	0.156 0
MLE-exp	t_{MTBF}/h	59 846	57 326	54 415	55 027
	$\lambda/10^{-5}$	1.669 5	1.633 8	1.610 8	1.443 9
	$t_{cal}/10^{-3} s$	0.652	0.868	0.753	0.635
	t_{MTBF}/h	59 897	61 206	62 080	65 061

由表 2 可得:

(1) 删失数据所占比例较小时,与 EM-exp 模型相比,MLE-exp 模型的参数估计耗时最少,且 2 种模型得到的 λ 的估计值与真值之间的差距较小。因此,对于失效数据为精确数据的继电设备,宜采用 MLE-exp 模型进行参数估计。

(2) 随着删失数据增多,3 种模型得到的参数估计值与真值之间的差距不断增大,此时 EM-exp 模型准确性最好。因此,对于失效数据中含删失数据的继电设备,宜采用 EM-exp 模型进行参数估计。

(3) 对处于偶然失效期的设备,EM-wbl 模型的参数估计结果精度较差。因此,对于具有恒定失效率或定期维护的继电设备,宜采用 EM-exp 模型进行参数估计。

4.2 损耗失效区

使用计算机生成服从威布尔分布的失效数据,样本容量为 100。生成数据时所用参数,如表 3 所示。

指数分布模型不适用于失效率随时间不断变化的情形,所以本节只讨论威布尔分布模型的参数估计精度。分别使用 EM-wbl 模型和传统的基于 MLE 的威布尔分布模型(MLE-wbl)进行参数估计,估计结果如表 4 所示。

表3 参数表(损耗失效区)

Table 3 Parameters table (wearout failure)

参数	数值
k, η	1.5, 60 000
ζ_1, σ_1	500, 500
ζ_2, σ_2	1 000, 1 000
t_{tr}	30 000

表4 损耗失效区参数拟合结果

Table 4 Parameter fitting results of wearout failures

失效模型		不同 (p_1, p_2, p_3) 下的拟合结果			
		(1,0,0)	(0.8,0.1,0.1)	(0.7,0.2,0.1)	(0.6,0.3,0.1)
EM-wbl	k	1.490 5	1.593 6	1.615 0	1.691 9
	$1/\eta$	59 732	59 677	57 660	55 822
	t_{MTBF}/h	53 965	53 525	51 653	49 824
MLE-wbl	k	1.490 5	1.675 2	1.771 6	1.972 6
	$1/\eta$	59 732	57 658	54 399	51 114
	t_{MTBF}/h	53 965	51 499	48 146	45 311

由表4可知,当样本数据全为精确数据时,2种方法计算得到的参数 k, η 以及 t_{MTBF} 的值均相同。随着删失数据所占比例不断上升,2种方法的估计误差均有所增加,但EM-wbl模型的估计结果更接近准确值,准确度更高。

5 实例分析

5.1 数据预处理

选取某市同一型号的继保设备50台,记设备投运时间为 t_0 ,停止观察的时间为截尾时间 t_{tr} ^[22-23]。对于直到观察结束仍未发生缺陷的设备,其累计运行时长 t_{acc} 为:

$$t_{acc} = t_{tr} - t_0 \quad (20)$$

对于没有发生故障,但在观察期间退出运行的设备,记其停运时间为 t_c ,其累计运行时长为:

$$t_{acc} = t_c - t_0 \quad (21)$$

对于发生缺陷的设备,若缺陷发现方式为自检,则发现缺陷的时间即为缺陷发生时间 t_f ,设备的失效时间记为:

$$t = t_f - t_0 \quad (22)$$

若发现缺陷的方式为巡视过程,则设备的失效时间区间 (t^-, t^+) 的计算方法为:

$$\begin{cases} t^- = \max T - t_0 \\ t^+ = t_f - t_0 \end{cases} \quad (23)$$

式中: T 为设备历次巡视时间,取其中距缺陷发现时间最近的一次巡视时间作为设备失效区间的左端点,缺陷发现时间为失效区间的右端点。

整理后的部分失效数据如表5所示,观察期间,

50台设备中共有9台发生失效,6台设备的失效时间为精确数据,3台设备的失效时间为区间删失数据,无设备在正常运行情况下中途退出运行。

表5 继保设备失效数据

Table 5 Failure data of relay protection equipment

数据编号	数据值	数据类型
1	7 283	精确数据
2	7 479	精确数据
3	10 533	精确数据
4	11 129	精确数据
5	14 957	精确数据
6	21 519	精确数据
7	(12 156, 15 928)	区间删失
8	(15 015, 19 597)	区间删失
9	(16 717, 21 831)	区间删失
⋮	⋮	⋮
49	24 329	随机截尾
50	24 977	随机截尾

5.2 计算分析

将经预处理的失效数据代入EM-wbl模型中,得到的参数估计值为:

$$\begin{cases} k = 1.717 2 \\ \eta = 60 641 \end{cases} \quad (24)$$

将参数值代入失效模型中,可得出此型号继保设备的可靠性指标如下。

可靠性:

$$R(t) = e^{-(t/60 641)^{1.717 2}} \quad (25)$$

故障概率密度:

$$f(t) = 1.050 6 \times 10^{-8} t^{0.717 2} e^{-(t/60 641)^{1.717 2}} \quad (26)$$

失效率:

$$\lambda(t) = 1.050 6 \times 10^{-8} t^{0.717 2} \quad (27)$$

平均无故障时间:

$$t_{MTBF} = 54 071 \quad (28)$$

可靠度、故障概率密度、失效率随时间变化的曲线如图3—图5所示。

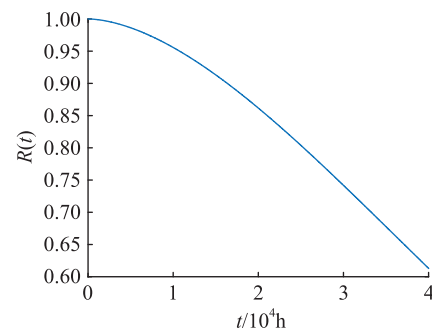


图3 可靠度曲线

Fig.3 Reliability function curve

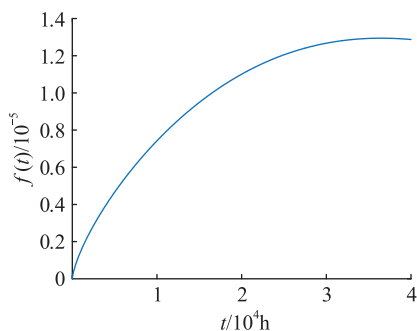


图4 故障概率密度曲线

Fig.4 Fault probability density curve

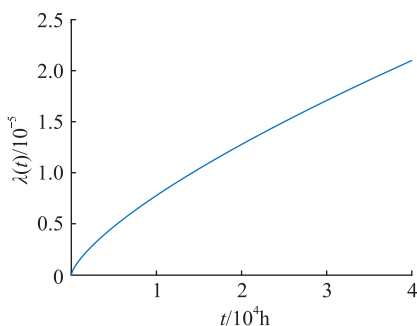


图5 失效率曲线

Fig.5 Failure rate curve

如图3所示,当 $R(t) = 0.98$ 时,得到继保设备运行时间为6 250 h,实际运行情况中设备首次失效时间为7 283 h,即通过EM算法得到的参数估计值可以很好的拟合设备可靠度曲线。形状参数 $k > 1$,由图4、图5可知,故障概率密度及失效率随运行时间的增加逐渐增大,反映了继保设备处于损耗期的特征。平均无故障时间 t_{MTBF} 可用于检修周期制定,通常情况下状态检修周期 $T_m = 0.1t_{MTBF}$,即建议状态检修周期为5 407 h。

综上所述,当存在删失数据时,采用EM-wbl得到的参数估计值可更精确地拟合设备实际的失效情况,再根据继保设备可靠度、失效率、故障概率密度等指标随时间变化的曲线预判其寿命分布,达到评价继保设备运行水平,预测设备运行年限的目的。

此外,上述分析的失效数据来自同类型设备,因此得出的结论适用于该类型设备。但所提模型对不同型号设备具有普适性。

6 结语

首先分析了继保设备失效数据的特点,总结了现场收集到的失效数据类型:精确数据、区间删失数据、右删失数据和随机截尾数据。然后在失效数据样本包含删失数据的情况下,对比分析了偶然失效期和老化失效期内不同估计方法得到的模型参

数精度。证明在偶然失效期内,EM-exp模型得到的参数估计精度更高;在老化失效期内,EM-wbl模型能以更高的精度拟合继保设备的失效特性曲线。

最后通过对某型号设备数据处理,结合失效分布模型,分析了该继保设备可靠度、失效率、故障概率密度等可靠性指标的时变特征。利用设备可靠度曲线,可提前判断设备可靠度,降低到阈值所需的时间,并及时发出告警提醒工作人员加强巡视。利用失效率曲线,可得设备的故障风险随时间变化的规律,用于安排检修计划。利用故障概率密度曲线,可计算设备的平均无故障时间,为合理安排服役年限提供参考。下一步将针对不同类型删失数据,对比研究不同算法的参数估计精度。

参考文献:

- [1] 高翔. 继电保护状态检修应用技术[M]. 北京:中国电力出版社,2008.
GAO Xiang. Application technology of state maintenance of relay protection[M]. Beijing:China Electric Power Press,2008.
- [2] 陈星田,熊小伏,齐晓光,等. 一种用于继电保护状态评价的大数据精简方法[J]. 中国电机工程学报,2015,35(3):538-548.
CHEN Xingtian, XIONG Xiaofu, QI Xiaoguang, et al. A big data simplification method for evaluation of relay protection operation state[J]. Proceedings of the CSEE,2015,35(3):538-548.
- [3] 桂海涛,陈宏山,李任明. 基于继电保护数据分析的配电网供电可靠性提升策略研究[J]. 供用电,2020,37(6):53-59.
GUI Haitao, CHEN Hongshan, LI Renmin. Research on strategy of distribution network power supply reliability improvement based on relay protection data analysis[J]. Distribution & Utilization,2020,37(6):53-59.
- [4] 张琳波,李本瑜,石恒初,等. 基于信息融合的继电保护智能管控系统研究[J]. 供用电,2020,37(3):52-57.
ZHANG Linbo, LI Benyu, SHI Hengchu, et al. Study of relay protection and intelligent management and control system based on information fusion[J]. Distribution & Utilization,2020,37(3):52-57.
- [5] 黄烽,邓勇,任晓辉,等. 继电保护装置运行年限评价的研究[J]. 电力系统保护与控制,2017,45(22):158-163.
HUANG Feng, DENG Yong, REN Xiaohui, et al. Assessment for service life of relay protection device[J]. Power System Protection and Control,2017,45(22):158-163.
- [6] 郭创新,陆海波,俞斌,等. 电力二次系统安全风险评估研究综述[J]. 电网技术,2013,37(1):112-118.
GUO Chuangxin, LU Haibo, YU Bin, et al. A survey of research on security risk assessment of secondary system[J]. Power System Technology,2013,37(1):112-118.
- [7] 戴志辉,刘媛,邱小强,等. 基于变权重模糊综合评判法的保护装置状态评价[J]. 电测与仪表,2021,58(4):150-157.
DAI Zhihui, LIU Yuan, QIU Xiaoqiang, et al. Condition evaluation approach of protective devices based on fuzzy com-

- hensive evaluation with variable weight[J]. *Electrical Measurement & Instrumentation*, 2021, 58(4):150-157.
- [8] 王嘉琦,徐岩,彭雅楠,等. 基于灰色-三参数威布尔分布模型的继电保护装置可靠性参数估计[J]. *电网技术*, 2019, 43(4):1354-1360.
WANG Jiaqi, XU Yan, PENG Ya'nan, et al. Estimation of reliability parameters of protective relays based on grey-three-parameter weibull distribution model[J]. *Power System Technology*, 2019, 43(4):1354-1360.
- [9] 李芷筠,戴志辉,焦彦军. 小样本失效数据下保护可靠性的贝叶斯-蒙特卡罗评估方法[J]. *电力系统及其自动化学报*, 2016, 28(5):9-14.
LI Zhijun, DAI Zhihui, JIAO Yanjun. Bayes Monte-Carlo reliability assessment method of protection systems with small failure samples[J]. *Proceedings of the CSU-EPSA*, 2016, 28(5):9-14.
- [10] 王文焕,杨国生,周泽昕,等. 基于随机截尾数据及极大似然估计的继电保护可靠性分布[J]. *电力系统保护与控制*, 2019, 47(12):125-131.
WANG Wenhuan, YANG Guosheng, ZHOU Zexin, et al. Reliability distribution of relay protection based on random truncation data and maximum likelihood estimation[J]. *Power System Protection and Control*, 2019, 47(12):125-131.
- [11] DIAMANTOULAKIS P D, KAPINAS V M, KARAGIANNIDIS G K. Big data analytics for dynamic energy management in smart grids[J]. *Big Data Research*, 2015, 2(3):94-101.
- [12] 程向向,秦川,张功林,等. 基于营配贯通电力大数据的配电网画像建模技术与运行状态评估研究[J]. *供用电*, 2020, 37(11):21-27.
CHENG Xiangxiang, QIN Chuan, ZHANG Gonglin, et al. Research on distribution network portrait modeling technology and operation status evaluation based on big data of the marketing and distribution fusion [J]. *Distribution & Utilization*, 2020, 37(11):21-27.
- [13] 王丽敏,杨国生,詹荣荣,等. 基于 MIF 的继电保护状态评价及其在设备全寿命周期管理中的应用[J]. *发电技术*, 2019, 40(3):300-306.
WANG Limin, YANG Guosheng, ZHAN Rongrong, et al. State evaluation of relay protection based on multi-source information fusion and its application in the life cycle management of equipment[J]. *Power Generation Technology*, 2019, 40(3):300-306.
- [14] 何旭,姜宪国,张沛超,等. 基于 SVM 的小样本条件下继电保护可靠性参数估计[J]. *电网技术*, 2015, 39(5):1432-1437.
HE Xu, JIANG Xiangguo, ZHANG Peichao, et al. SVM based parameter estimation of relay protection reliability with small samples[J]. *Power System Technology*, 2015, 39(5):1432-1437.
- [15] 苏傲雪,范明天,张祖平,等. 配电系统元件故障率的估算方法研究[J]. *电力系统保护与控制*, 2013, 41(19):61-66.
SU Aoxue, FAN Mingtian, ZHANG Zuping, et al. A new approach for the evaluation of failure rates on distribution system components[J]. *Power System Protection and Control*, 2013, 41(19):61-66.
- [16] ZHU T F. Statistical inference of weibull distribution based on generalized progressively hybrid censored data[J]. *Journal of Computational and Applied Mathematics*, 2020, 371:112705.
- [17] 陈勇,李胜男,张丽,等. 基于改进平均秩次法的保护装置可靠性评估[J]. *电力系统保护与控制*, 2020, 48(2):165-171.
CHEN Yong, LI Shengnan, ZHANG Li, et al. Reliability evaluation of protective device based on improved average rank order method[J]. *Power System Protection and Control*, 2020, 48(2):165-171.
- [18] 孔祥芬,张利寒,蔡峻青. EM 算法下飞机 IDG 删失数据的可靠性分析[J]. *机械科学与技术*, 2019, 38(7):1138-1142.
KONG Xiangfen, ZHANG Lihan, CAI Junqing. Reliability analysis of aircraft IDG failure data censored randomly by EM algorithm[J]. *Mechanical Science and Technology for Aerospace Engineering*, 2019, 38(7):1138-1142.
- [19] 徐岩,王士铭,刘婧妍,等. 继电保护装置的一种新型评估方法[J]. *电测与仪表*, 2018, 55(5):119-124, 136.
XU Yan, WANG Shiming, LIU Jingyan, et al. A new evaluation method for relay protection device[J]. *Electrical Measurement & Instrumentation*, 2018, 55(5):119-124, 136.
- [20] 薛安成,王睿琛,刘蔚,等. 继电保护装置恒定失效率估算方法[J]. *电力系统自动化*, 2012, 36(4):6-10.
XUE Ancheng, WANG Ruichen, LIU Wei, et al. Estimation methods for constant failure rate of protection equipments[J]. *Automation of Electric Power Systems*, 2012, 36(4):6-10.
- [21] ANDRÉS CHRISTEN J, RUGGERI F, VILLA E. Utility based maintenance analysis using a random sign censoring model[J]. *Reliability Engineering & System Safety*, 2011, 96(3):425-431.
- [22] 徐长宝,赵立进,高吉普,等. 基于灰色马尔科夫链的继电保护装置寿命研究[J]. *电力科学与技术学报*, 2019, 34(3):114-119.
XU Changbao, ZHAO Lijin, GAO Jipu, et al. Research of relay protection device life cycle based on gray Markov model[J]. *Journal of Electric Power Science and Technology*, 2019, 34(3):114-119.
- [23] 黄涛,周鹏,赖秋频,等. 一种智能变电站保护装置寿命预测方法[J]. *电测与仪表*, 2018, 55(12):39-44.
HUANG Tao, ZHOU Peng, LAI Qiupin, et al. A life prediction method for intelligent substation protection device[J]. *Electrical Measurement & Instrumentation*, 2018, 55(12):39-44.

作者简介:



张雷

张雷(1986),男,学士,工程师,从事电力系统继电保护与调度工作(E-mail:2112805760@qq.com);

王光华(1978),男,学士,高级工程师,从事电力系统继电保护与调度工作;

李金铄(1997),女,硕士在读,研究方向为电力二次运行评估与智能运维。

Operating status assessment of protective equipment considering censored characteristics under background of big data

ZHANG Lei¹, WANG Guanghua¹, LI Jinshuo², GENG Hongxian², DAI Zhihui²

(1. State Grid Baoding Power Supply Company of Hebei Electric Power Co., Ltd., Baoding 071000, China;

2. School of Electrical and Electronic Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: Failure data without exact observed failure time of equipment is referred to as censored data. Aiming at the characteristics of censored failure data of protective equipment under the background of big data, a method for evaluating the operating life of protective equipment considering censored data under the background of big data is proposed. Firstly, based on analyzing failure-data features of protective equipment, the failure data were preprocessed. By combining the expectation-maximization (EM) algorithm with the exponential distribution model and the weibull distribution model, failure model parameters of protective equipment were estimated. Secondly, the estimated parameters are substituted into the failure model to obtain reliability indices, such as the reliability, failure probability density, failure rate, and meantime between failures. Subsequently, through simulation study, the estimation accuracy of model parameters in accidental-failure and aging-failure periods obtained by different methods was comparatively analyzed, and the effectiveness of the proposed method for processing censored data was verified. Finally, taking a certain type of protective equipment as an example to investigate reliability indices, the feasibility of applying this method to plan equipment maintenance cycle is verified.

Keywords: censored data; expectation-maximization (EM) algorithm; operating status assessment; big data; protective equipment; maintenance cycle

(编辑 陆海霞)

(上接第 156 页)

Non-intrusive load detailed disaggregation algorithm for electrothermal load based on three dimensional characteristics vector

LIU Xi'ang, ZHOU Gan, XU Xin, LI Zhi

(School of Electrical Engineering, Southeast University, Nanjing 210096, China)

Abstract: Non-intrusive load monitoring and disaggregation (NILMD) technology is an important data acquisition method for deep improvement of residents' energy service and the interaction between power supply and demand. However, it is unable to disaggregate accurately the electrothermal load by the NILMD algorithm of edge detection which is widely used in current engineering. In order to solve this problem, a NILMD algorithm for typical electrothermal load based on three dimensional characteristics vector is proposed in this paper. Firstly, the edge detection algorithm is used to extract the electrothermal events through active power, reactive power and current harmonic, and the three dimensional characteristics vector model is constructed together with the non-electric characteristics such as running duration, start and stop times based on active power. Then, the learning rules and algorithm for typical electrothermal load detailed disaggregation are designed by sequential covering. Finally, the disaggregation accuracy of the electrothermal load is more than 85% based on experimental verification. Experimental results show that the NILMD algorithm for typical electrothermal load proposed in this paper effectively improve the disaggregation accuracy of four typical electrothermal load.

Keywords: non-intrusive load monitoring and disaggregation; typical electrothermal load disaggregation; three dimensional characteristics vector; edge detection; edge extraction; sequential covering

(编辑 陆海霞)