

一种考虑软件补丁的软件可靠性模型研究

付方智, 沈元隆

(南京邮电大学 电子科学与工程学院, 江苏 南京 210003)

摘要: 将调整因子引入到关于软件产品的可靠性预计模型中, 提出了改进的关于软件产品的软件可靠性预计模型。通过一组数据仿真证明, 改进模型比原模型有着更好的性能, 更加符合现实的软件可靠性活动过程。

关键词: 软件可靠性模型; 软件补丁; 参数估计

中图分类号: TP311

文献标识码: A

文章编号: 1674-7720(2010)11-0046-02

The research of software reliability prediction model based on software patch

FU Fang Zhi, SHEN Yuan Long

(Electronic Science and Engineering College, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Improved software reliability prediction model is drawn through introducing the adjustment factor into software reliability prediction model. The paper proves that the improved model has a better performance compared to the original model by simulating a set of data. It is more suitable for realistic course of activities of software reliability.

Key words: software reliability model; software patch; parameter estimation

从 1972 年至今, 为了研究软件可靠性, 已经提出一百多种模型, 但这其中大部分模型主要是从消除软件中的错误数的角度来研究软件的可靠性模型。而另一种现象也逐渐被重视起来, 就是一种软件在发布后, 软件产品的可靠性伴随着时间增长, 而错误数并没有纠正的情况下, 软件的可靠性也得到了改善。为了描述这种软件产品可靠性的增长, 2008 年 Pankaj Jalote 和 Brendan Murphy 从用户角度, 针对软件产品销售后可靠性增长的现象提出一种关于软件产品的可靠性预计模型^[1]。但是, 在 Pankaj Jalote 和 Brendan Murph 提出的模型中, 其假设在某些方面过于严格, 特别是所提出的模型中忽视了软件补丁的作用。而现在软件补丁的作用越来越大, 因而软件的补丁不能被忽视了。本文全面分析了实际情况, 提出了考虑软件补丁的关于软件产品的软件可靠性增长模型。

1 原始的关于软件产品的软件可靠性增长模型

1.1 基本假设与公式

2008 年 Pankaj Jalote 和 Brendan Murphy 提出的一种关于软件产品的可靠性增长模型^[1], 假定如下:

(1) 假定软件销售出去以后, 没有新的版本发布出来, 不考虑软件的补丁的问题, 认为补丁的影响微乎其微。

(2) 假定用户初始的失效率为 $(\lambda_0 + \lambda_f)$, 稳定时的失效率为 λ_f 。

(3) 假定可以通过软件产品技术支持服务, 获得每个月的产品销售量与每个月的总的失效总数。

(4) 针对用户, 软件的失效率以常数每个月来进行衰减, 并且 α_{i+1} 为 α_i 的 k 倍 ($k > 1$), 直到它达到稳定的状态。

失效率的定义: 假定用户安装了 N 套某种软件产品, 从安装时刻开始, 经历了 T 的时间间隔, 总的失效数为 F , 则:

$$\lambda = \frac{F}{N \times T} \quad (1)$$

从用户的角度, 给一个失效率公式:

$$\lambda(i) = \lambda_0 \times \alpha_i + \lambda_f \quad (2)$$

基于以上的假定, 可以得出如下模型:

假定: 第一月卖出 N_1 套, 第二个月卖出 N_2 套, ..., 第 i 个月卖出 N_i 套; 令这种软件的暂态的失效率是 λ_0 ,

每个月软件衰退率为 α ，稳态的失效态是 λ_f ，由此可以得到每个月的总的失效数 F_i ：

$$F_i = \lambda_0 N_i + \lambda_0 N_{i-1} \alpha + \lambda_0 N_{i-2} \alpha^2 + \dots + \lambda_0 N_1 (\alpha_1 \alpha_2 \dots \alpha_{i-1}) + \lambda_f (N_i + N_{i-1} + \dots + N_1) \quad (3)$$

由于 $\alpha_i = \alpha$ ，则可以得到：

$$\alpha_i = k^{i-1} \alpha \quad (4)$$

将式(4)代入式(3)得：

$$F_i = \lambda_0 N_i + \lambda_0 N_{i-1} \alpha + \lambda_0 N_{i-2} k \alpha^2 + \dots + \lambda_0 N_1 k^{(i-1)(i-2)/2} \alpha^{(i-1)} + \lambda_f (N_i + N_{i-1} + \dots + N_1) \quad (5)$$

由于任何一个月的故障率都是不可能增加的，故 k 的限定条件为：

$$0 < k^{(i-1)(i-2)/2} \alpha^{(i-1)} < 1, \quad i > 1 \quad (6)$$

1.2 模型的参数估计

利用收集到每月的失效总数和每月的产品销售数，并采用最大似然法或者最小二乘法来对 λ_0 、 λ_f 、 α 、 β 、 γ 、 k 进行估计，以改进模型。然后，软件产品公司再利用估计的参数和模型，对某种软件产品进行可靠性预测，从而根据预测进行对产品的改进。利用最小二乘法对于模型参数的估计公式如下：

$$X' = \sum_{i=1}^n [F_i - (\lambda_0 N_i + \lambda_0 N_{i-1} \alpha + \lambda_0 N_{i-2} k \alpha^2 + \dots + \lambda_0 N_1 k^{(i-1)(i-2)/2} \alpha^{(i-1)} + \lambda_f \sum_{m=1}^i N_m)]^2 \quad (7)$$

将式(7)代入 $\frac{\partial Y}{\partial \lambda_0} = 0$ ， $\frac{\partial Y}{\partial \lambda_f} = 0$ ， $\frac{\partial Y}{\partial \alpha} = 0$ 和 $\frac{\partial Y}{\partial k} = 0$ ，利用数值求解可以求得 λ_0 、 λ_f 、 α 、 k 的值，从而获得软件可靠性模型。

2 考虑软件补丁的关于软件产品的软件可靠性增长模型

2.1 新模型的假设与基本公式

基于前面的分析，本文重新修正 2008 年 Pankaj Jalote 和 Brendan Murphy 提出的关于软件产品的软件可靠性预计模型的假设，并提出改进的模型，并称其为 FP 模型。

FP 模型是对改进的模型考虑了补丁之后的模型，并认为补丁虽然可以纠正软件中的错误，但是同样会带来新的错误，因此引进 2 个新的参数 β 、 γ ，其假设如下：

(1) 假定软件销售出去以后，没有新的版本发布出来，补丁及时发布，用户能及时安装和使用，而新的补丁安装后会带来新的错误。

(2) 假定用户初始的失效率为 $\lambda_0 + \lambda_f$ ，稳定时的失效率为 λ_f 。

(3) 假定可以通过软件产品技术支持服务获得每个月的产品销售量与每个月的总的失效率。

(4) 针对用户，由于考虑到了补丁的影响以及会带来新的问题，所以软件的每个月失效率为 α_i ，直到它达到稳定的状态。

根据上述假设可以得到此模型的失效公式：

$$\lambda(i) = \lambda_0 \alpha_i + \lambda_f \quad (8)$$

其中 λ 为时间。

$$\alpha_i = k^{i-1} \alpha - \beta + \gamma^i \quad (9)$$

将原模型的公式修正为：

$$F'_1 = \lambda_0 N_1 + \lambda_f N_1$$

$$F'_2 = \lambda_0 N_2 + \lambda_0 N_1 (\alpha - \beta + \gamma) + \lambda_f (N_1 + N_2)$$

...

$$F'_i = \lambda_0 N_i + \lambda_0 N_{i-1} (\alpha - \beta + \gamma) + \lambda_0 N_{i-2} (\alpha - \beta + \gamma) (k\alpha - \beta^2 + \gamma^2) + \dots + \lambda_0 N_1 (\alpha - \beta + \gamma) (k\alpha - \beta^2 + \gamma^2) \dots (k^{(i-2)} \alpha - \beta^{(i-1)} + \gamma^{(i-1)}) + \lambda_f \sum_{m=1}^i N_m \quad (10)$$

2.2 FP 模型的参数估计

利用收集的每月的失效总数和每月的产品销售数，并采用最大似然法或者最小二乘法来对 λ_0 、 λ_f 、 α 、 β 、 γ 、 k 进行估计，FP 改进模型，软件产品公司再利用估计的参数和模型，对某种软件产品进行可靠性预测，从而根据预测进行对产品的改进。利用最小二乘法对模型参数的估计公式如下：

$$Y = \sum_{i=1}^n [F_i - (\lambda_0 N_i + \lambda_0 N_{i-1} (\alpha - \beta + \gamma) + \lambda_0 N_{i-2} (\alpha - \beta + \gamma) (k\alpha - \beta^2 + \gamma^2) + \dots + \lambda_0 N_1 (\alpha - \beta + \gamma) (k\alpha - \beta^2 + \gamma^2) \dots (k^{(i-2)} \alpha - \beta^{(i-1)} + \gamma^{(i-1)}) + \lambda_f \sum_{m=1}^i N_m)]^2 \quad (11)$$

将式(11)代入 $\frac{\partial Y}{\partial \lambda_0} = 0$ ， $\frac{\partial Y}{\partial \lambda_f} = 0$ ， $\frac{\partial Y}{\partial \alpha} = 0$ ， $\frac{\partial Y}{\partial \beta} = 0$ ， $\frac{\partial Y}{\partial \gamma} = 0$ 和 $\frac{\partial Y}{\partial k} = 0$ ，利用数值求解可以求得 λ_0 、 λ_f 、 α 、 β 、 γ 、 k 的值，从而获得软件可靠性模型。

3 实例仿真

本实例仿真数据是来源于 2008 年 Pankaj Jalote 和 Brendan Murphy 在 ACM Trans 发表的文献^[1]。仿真数据如表 1 所示。

表 1 某公司每月销售情况和每月失效总数

月份	总的失效数	每月的销售数量
1	367	4 618
2	853	14 385
3	835	5 608
4	791	6 186
5	956	9 829
6	805	5 584
7	967	8 240
8	1 218	7 656
9	1 031	4 914
10	1 144	5 295
11	1 058	7 418

通过数值方法求得 FP 模型和原模型的参数值如表 2 所示。

表 2 FP 模型和原模型的参数值

参数	λ_0	λ_f	α	k	β	γ
原模型	0.047 0	0.006 8	0.437 9	2.103 6		
FP 模型	0.046 9	0.006 9	0.424 5	2.157 6	0.996 7	1.004 3

3.1 图形仿真

用 Matlab 7.1 仿真 FP 模型和原模型, 仿真结果如图 1 所示。

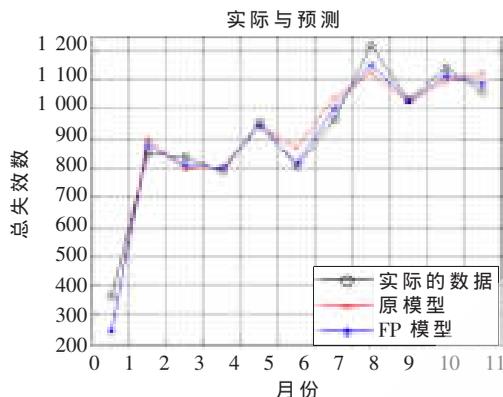


图 1 Matlab 7.1 仿真图

从图 1 可以看出, 改进的模型比原来的模型更接近实际的数据。

3.2 模型的常用预测能力比较

误差平方和 $SSE^{[2]}$ 用来描述累计失效数的观测值与预测值之间的距离, 定义如下:

$$SSE = \sum_{i=1}^n (y_i - \hat{m}(t_i))^2 \quad (12)$$

在式(12)中, n 表示实效数据集中失效样本的数量, y_i 表示到时间 t_i 为止累计故障数的实测值, $\hat{m}(t_i)$ 表示到时间 t_i 为止累计故障数的估计值。

SSE 越小, 则表示曲线的拟合能力越好^[3]。由表 3 可知, 改进的 FP 模型在预测能力上优于原来的模型。

表 3 两种模型的常用预测能力参数 AE 比较

	原始模型	FP 模型
SSE 值	86 572	18 679

通过分析和研究原始的关于软件产品的软件可靠性预计模型, 考虑到软件补丁的作用不能忽视, 提出了软件发布后的软件可靠性模型——一种考虑软件补丁的软件可靠性模型。本文首先分析了原模型的假设和软件补丁对软件可靠性的影响, 提出修订的假设条件, 并由假设条件推出改进的 FP 模型, 然后对改进的 FP 模型的参数估计展开讨论, 并得到相应的函数表达式。利用最小二乘法求得模型参数, 再通过仿真实验对原始的关于软件产品的软件可靠性预计模型和改进的 FP 模型进行分析比较, 可以得出改进的 FP 模型的预测精度比原始模型要高, 并且其拟合效果都要优于原始模型。

参考文献

- [1] JALOTE P, MURPHY B. Post-release reliability growth in software products [J]. ACM Transactions on Software Engineering and Methodology (TOSEM), 2008, 17(4).
- [2] ZHANG Xue Mei, TENG Xiao Lin, PHAM H. Considering fault removal efficiency in software reliability assessment[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems And Humans, 2003, 33(1):114-119.
- [3] 赵靖, 刘宏伟, 崔刚, 等. 考虑测试与运行差别的软件可靠性增长模型[J]. 计算机研究与发展, 2006, 43(3):503-508.

(收稿日期: 2009-11-14)

作者简介:

付方智, 男, 1985 年生, 硕士研究生, 主要研究方向: 通信系统的可靠性技术。

沈元隆, 男, 1948 年生, 硕士, 教授, 主要研究方向: 通信系统的可靠性技术。