

基于隶属度的软件可靠性综合预测方法

王二威¹, 侯福均², 郑述招¹

(1.北京理工大学 珠海学院 管理与经济学院, 广东 珠海 519085;

2.北京理工大学 管理与经济学院, 北京 100081)

摘要: 将多模型综合的思想和“变点”思想相结合, 给出了一种简单有效且具有普适性的软件可靠性预测方法——基于模糊隶属度的软件可靠性多模型综合预测方法。首先阐述了多模型综合预测问题的一般描述, 并介绍了模型评价准则, 然后给出了单个模型权重的确定方法和该方法的一般步骤, 最后用实际数据验证了方法的有效性及其普适性。

关键词: 软件可靠性; 模糊隶属度; 多模型综合

中图分类号: TP311

文献标识码: A

文章编号: 1674-7720(2011)15-0004-04

Simple method of multi-model synthesis prediction of software reliability based on fuzzy membership

Wang Erwei¹, Hou Fujun², Zheng Shuzhao¹

(1.Department of Management and Economics, Zhuhai Campus, Beijing Institute of Technology, Zhuhai 519085, China;

2.Department of Management and Economics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Concerned combination concept and change-point idea, a new multi-model prediction method for software reliability based on fuzzy membership was proposed. Firstly, general description of the multi-model prediction problem was expounded, and evaluation criteria of the model was introduced, then the method to determine the weight of single model and the shipunloaders of this multi-model prediction method were proposed. Finally, the method was verified with actual data the effectiveness and universality.

Key words: software reliability; fuzzy membership; combinational models

软件的普遍应用促进人们对其可靠性的关注, 到目前为止软件可靠性的研究取得了很大的成果, 提出了近百种软件可靠性预测模型, 如 JM 模型、GO 模型、NHPP 模型、LV 模型等, 并称此类模型为经典模型。但是单个经典模型存在很大的局限性, 对于一个软件可靠性的评估往往很难用一个模型来处理, 而且很多软件可靠性模型都是建立在概率分布假设的基础上, 这就造成对一个软件有很好的适用性而对其他的软件则效果很差^[1]。20 世纪 90 年代以来, 基于知识的方法被越来越多地应用到软件可靠性预测中, 如神经网络的方法^[2-4]、遗传算法、支持向量机、泛函网络等。然而这些方法算法复杂、计算性能较低、其预测结果的精确性取决于神经网络结构的设计和模型的训练。Cai^[5]在大量对比实验的基础上指出, 大多数情况下, 神经网络方法很难得出满意的量化预测结果。

《微型机与应用》2011 年第 30 卷第 15 期

由于单个经典模型的局限性以及基于新理论方法的复杂性和不成熟性, 软件可靠性预测仍然还有很多问题需要解决: 一是建立能够普遍应用的模型; 二是设计一套行之有效的模型选择方法, 能够让工程人员从众多的软件可靠性模型中选择出最适合实施项目的模型。

Littlewood B 在对大量软件项目研究的基础上提出了“变点”的思想^[6], 表明不能期望用某一个模型来描述软件的失效过程, 而从另一个侧面表明可以用多个模型来描述软件的失效过程。香港中文大学的 Lyu Michael 提出了多个单一经典模型“综合”的思想^[1], 并提出了四种线性“综合”的方法, 试图找到一种能够普遍适用的软件可靠性预测方法。

本文将多模型综合思想和“变点”思想相结合, 提出了一种基于模糊隶属度的软件可靠性多模型综合预测方法, 通过各个经典模型的预测数据和实际数据的模糊

欢迎网上投稿 www.pcachina.com 5

等于程度来动态地改变单个模型的权重,以综合各个模型的优势,进而更准确地预测软件的可靠性。

1 多模型综合预测问题的一般描述

多模型综合预测问题可用如下表达式来描述^[1]:

$$\hat{f}(t) = \sum_{j=1}^n \omega_j \hat{f}_j(t) \quad (1)$$

式中, n 是选定的单个模型的个数, $\hat{f}_j(t)$ 是第 j 个模型的 t 时刻的预测值, ω_j 是第 j 个模型的权重,且 $\sum \omega_j = 1$ 。

这是一个组合预测问题,首先要选定参与预测的单个模型,然后确定各单个模型的权重 ω_j ,即可根据式(1)计算出多模型综合预测值。因此选择参与预测的单个模型和如何计算各模型的权重是解决问题的关键。

2 基于隶属度的软件可靠性多模型综合预测方法

2.1 单个模型权重的确定

2.1.1 二元关系表的建立

为了运用模糊隶属度来描述预测值和实际值的模糊等于程度,进而确定单一模型的权重,首先建立一个二元关系表。为此将选定的几个经典模型作为条件属性,即条件属性集合 $C = \{c_1, c_2, \dots, c_n\}$,则每个时刻这些经典模型预测的失效数为条件属性值;将实际失效数作为决策属性 d ,则决策属性集合 $D = \{d\}$,每个时刻实际的失效数为决策属性值;在 t 时刻各个经典模型预测的失效数和实际失效数看作论域 U 的对象的一条信息,即 $u = (c_1, c_2, \dots, c_n; d_t)$,论域 $U = \{u_1, u_2, \dots, u_m\}$ 。至此,二元信息表建立完成。

2.1.2 隶属函数的确定

本方法用模糊隶属度来描述单个经典模型预测值和实际值的接近程度。将预测值等于实际值模糊化,预测值越接近实际值则其模糊等于实际值的程度就越高;反之就越低,而且会迅速趋近于不等于。采用参照法为单个经典模型预测值和实际值的模糊等于事件确定隶属函数,经过试验比较,本文参照正态分布确定隶属函数,即:

$$\underline{A}(x) = e^{-k(x-d_t)^2} \quad (2)$$

式中, $k = \frac{1}{2\sigma^2}$, σ 为正态分布的标准差; x 为 t 时刻经典模型的预测失效数; d_t 为 t 时刻的实际失效数。则 $\underline{A}(x)$ 表示 x 和 d_t 的模糊等于程度,当 $(x-d_t) \rightarrow \infty$ 时, σ 值越小,趋近于不等于的速度就越快;反之,趋近于不等于的速度就越慢。

2.1.3 单个经典模型权重的计算方法

在进行软件可靠性线性综合模型动态预测时,单个经典模型权重的确定是最重要的。本文给出动态计算单个经典模型权重的方法步骤如下:

(1)在二元信息表的基础上,选择前 h 个时刻作为基

础数据(前 h 个时刻的实际失效数已知)。

(2)计算前 h 个时刻每个模型预测失效数和实际失效数的模糊等于程度 $\underline{A}(c_i)$ 。

(3)计算出各个模型在对 $h+1$ 时刻软件失效数预测的重要度(即权值) W_{c_i} :

$$W_{c_i} = \frac{\sum_t \underline{A}(c_i)}{\sum_i \sum_t \underline{A}(c_i)} \quad (3)$$

本方法采用固定窗口滑动步长的权重确定策略,即对 h 时刻以后每个时刻的失效数进行预测时,都使用该时刻之前的 h 个时刻作为基础数据。这反映了“变点”的思想,综合前 h 时刻各个模型预测的能力,使得每个时间段预测较准的模型总能获得较大的权重,进而获得更好的综合预测效果。窗口 h 的值取决于软件的开发环境、测试计划、操作剖面等要素^[1],但一般取3~5为宜,这样既能够及时感知到单个经典模型预测性能的变化,又不至于平滑太多预测时刻之前的基础数据导致不能反映数据中“变点”的问题。

2.2 本文方法一般步骤

在Lyu提出的软件可靠性多模型综合预测思想的基础上融入“变点”的思想,确定基于模糊隶属度的软件可靠性多模型动态综合预测方法,其步骤如下:

(1)选择经典模型。本文使用Lyu提出的选择原则,即:首先选择那些有很好的预测效果而且应用广泛的经典模型;其次选择那些预测的偏好可以相互抵消的模型,这些模型的预测偏好综合起来以后既不悲观也不乐观。

(2)确定固定窗口长度 h 。用每个经典模型分别进行对前 h 个时刻进行预测,并建立二元信息表。令 $t=h$ 。

(3)计算出各个模型在进行预测 $t+1$ 的重要度,即权值。

(4)预测 $t+1$ 时刻的失效数 d'_{t+1} 。

$$d'_{t+1} = \sum_i c_i \times W_{c_i} \quad (4)$$

(5)当实际测得 $t+1$ 时刻的失效数 d'_{t+1} 后,二元关系表指针下移一位(即 $t=t+1$),返回第三步。

3 综合模型评价

对软件可靠性模型预测性能进行评价最主要的是模型的有效性,它可以用定量的方法来测量。本文采用的评价标准如下:

(1)平均误差AE(Average Error)^[7]:在整个测试阶段模型预测的误差均值,只能针对同一数据进行比较。平均误差越小,模型的预测效果越好。

(2)平均偏差AB(Average Bias error)^[7]:在整个测试阶段模型预测总的偏差趋势。平均偏差越接近于零,模型的预测效果越好。

《微型机与应用》2011年第30卷第15期

(3)均方根误差 RMSE(Root Mean Square Error)^[1]:在整个测试阶段模型预测值和实际值距离。均方根误差值越小,模型的拟合度越高,预测性能越好。

4 实验分析

4.1 实验 1

选取 UDIMM (University of Denmark Informatics and Mathematical Modeling)提供的数据包^[8]中一软件失效数据作为实例进行预测分析。根据 Lyu 所提出的几个经典模型选择原则,选用 JM、GO、LV、YO 四个经典模型。其中, JM 模型是马尔可夫过程模型, GO 模型和 YO 模型都是经典的非齐次泊松过程模型, LV 模型是贝叶斯模型,它们都有很好的预测效果并且应用比较广泛。四个模型预测偏好的综合正好可以相互抵消, GO、LV 模型的预测结果比较乐观,而 JM 模型和 YO 模型的预测结果则有悲观也有乐观。

设固定窗口 $h=4$,即采用前 4 个时刻各个经典模型的预测失效数和实际失效数建立二元关系表。使用基于模糊隶属度的软件可靠性多模型动态综合预测方法步骤进行计算,隶属函数参数 $\sigma=0.25$ 时计算结果如表 1 所示。

表 1 软件失效预测数据表(UDIMM)

数据组	失效发现时间/天	JM	GO	LV	YO	实际失效数/天	预测失效数/天
1	9	1.174	1.726	1.653	1.89	1	
2	21	1.759	2.884	2.462	2.673	2	
3	24	2.849	3.567	2.973	3.953	3	
4	26	5.958	6.976	5.431	7.417	6	
5	31	11.591	12.121	9.531	12.442	11	11.03705
6	43	16.952	18.987	14.263	19.298	16	16.14023
7	50	20.431	21.323	18.692	22.513	19	19.84008
8	58	25.843	25.342	23.215	25.891	24	24.94592
9	63	27.948	27.215	25.723	27.141	25	25.97932
10	70	32.421	31.862	30.431	32.359	29	30.43388
11	71	33.727	32.541	32.294	34.131	32	32.2940
12	77	34.396	33.321	33.119	35.643	33	33.15036
13	78	37.531	34.231	34.391	36.492	36	34.34698
14	85	43.953	40.967	39.532	43.6534	42	40.18896
15	91	49.362	45.543	44.371	48.964	47	44.99298
16	92	52.634	50.452	46.834	51.326	49	48.34846
17	103	53.732	54.343	49.792	53.641	52	53.64195
18	120	57.496	57.653	53.527	56.751	55	57.65300
19	157	60.573	59.634	55.792	58.432	57	58.69413
20	167	61.431	61.239	56.352	60.836	58	58.69413

从表 1 可以看出,该失效数据存在“变点”的问题,即在第 5~第 7 个时刻, JM 和 LV 模型比较适用;从第 7 个时刻开始一直到第 12 个时刻, LV 模型比较适用, GO 模型次之,此后都没有特别合适的模型。基于模糊隶属度的软件可靠性多模型动态综合预测方法能够动态地对各时刻各分模型的权重进行调整,使得整个测试过程

中对失效数的预测都比较准确;而且采用固定窗口滑动步长,能够很好地感知数据的变化,减小了数据噪声。

为了与单个模型的预测性能进行对比,选用平均误差(AE)、平均偏差(AB)、均方根误差(RMSE)作为模型预测性能的评价标准。各个模型评价标准计算结果如表 2 所示。表中 FD 模型即是本文提出的基于模糊隶属度的软件可靠性多模型动态综合预测方法。

表 2 各模型评价标准(UDIMM)

评价标准	FD 模型		JM 模型	GO 模型	LV 模型	YO 模型
	$\sigma=0.25$	$\sigma=0.5$				
AE	0.03000	0.03097	0.063 22	0.063 15	0.044 03	0.0768 0
AB	-0.00846	-0.01769	0.063 22	0.050 06	-0.032 65	0.0768 0
RMSE	0.03410	0.14646	0.067 60	0.077 36	0.055 13	0.0943 5

从表 2 可以看出, FD 模型的绝大多数评价标准值都比单个模型的好,只有 $\sigma=0.5$ 的 FD 模型预测结果的均方根误差较大,说明 FD 模型比单个经典模型有更好的预测性能。 $\sigma=0.25$ 时, FD 模型的三个评价标准取值比 $\sigma=0.5$ 时好很多,也更合理;说明 $\sigma=0.25$ 时隶属函数能够更好地描述模糊等于这个概念,更能反映客观实际情况,即是把更大的权重分配给预测值更贴近实际失效数的模型。当 σ 过小时,综合预测时会过于偏重某一个经典模型,不能够达到综合多个模型优势的效果。

4.2 实验 2

为了与 Lyu 提出的动态权重的线性综合模型(DLC)进行比较,选取 Musa 发表的 DACS 数据集中的一组失效数据(SYS1)^[8]和 Lyu 提供的一组软件项目失效数据(LDATA)^[9]进行试验。同样选用 Lyu 选择的三个经典模型:GO 模型、MO 模型、LV 模型,固定窗口 $h=4$,隶属函数参数取 0.25。分别用两种方法对两组数据进行预测,并对模型预测性能进行评价。两种方法的评价标准计算结果如表 3 所示。

表 3 FD 模型和 DLC 模型评价标准比较(SYS1)

评价标准	FD 模型	DLC 模型	FD 模型	DLC 模型
	(SYS1)	(SYS1)	(LDATA)	(LDATA)
AE	0.02306073	0.03607685	0.03233936	0.06137973
AB	0.01764517	0.01379867	-0.01919051	0.04387112
RMSE	3.10193797	4.17741853	4.38295704	10.27969786

从表 3 可以看出,本文的 FD 模型在各个标准下都优于 DLC 模型。这主要是因为 DLC 模型是根据模型的累积预测精确度来动态确定权重,这样平滑了预测时刻以前所有的基础数据,不能很好地反映“变点”的思想;而 FD 模型采用固定窗口滑动步长来动态确定权重,能够敏锐地感知数据的变化,及时调整权重,又不至于平滑过量的基础数据,从而取得更好的预测效果。

本文提出的基于模糊隶属度的软件可靠性多模型动态综合预测方法结合“变点”思想,采用模糊隶属的方法来动态确定单个模型在综合预测时的权重,以期获得

更准确的预测效果。从实验可以看出,该方法不但比单个经典模型有更高的预测精度和更好的预测效果,而且比 Lyu 提出的线性综合模型更能反映动态反应软件失效数据中“变点”的问题,有更高的预测精度和更好的预测效果。该方法对多个软件项目具有普适性,而且简单有效,在一定程度上解决了一个模型只对某个软件工程项目或者其中某一段时间能够达到较高的预测水平的问题。

参考文献

- [1] LYU M R, NIKORA A. Applying reliability models more effectively[J]. IEEE Software, 1992, 9(4):43-52.
- [2] KARUNANITHI N, WHITLEY D, MALAIYA Y K. Using neural networks in reliability prediction [J]. IEEE Software, 1992, 9(4):53-59.
- [3] SITTE R. Comparison of software -reliability -growth predictions: neural networks vs parametric-recalibration[J]. IEEE Transactions on Reliability, 1999, 48(3):285-291.
- [4] Su Yushen, Huang Chinyu. Neural -network -based approaches for software reliability estimation using dynamic weighted combinational models [J]. Journal of Systems and Software, 2007, 80(4):606-615.
- [5] CAI K Y, CAI L, Wang Weidong, et al. On the neural network approach in software reliability modeling [J]. The Journal of Systems and Software, 2001, 58(1):47-62.
- [6] PHAME H. Reliability handbook [S]. NewYork: Springer Verlag, 2002.
- [7] MALAIYA Y K, KARUNANITHI N. Predictability measures for software reliability models [C]. Fourteenth Annual International Computer Software and Applications Conference, 1990COMPASAC 90. 1990, 1:7-12.
- [8] VLADICESCU F P. Performance evaluation of computers courses technical, University of Denmark Informatics and Mathematical Modeling. <http://www.imm.dut.dk/popentiu/pec/pec.html>.(2009-04-06).
- [9] LYU M R, NIKORA A. CASRE-A computer-aided software reliability estimation tool [C]. Computer -Aided Software Engineering Proceedings. 1992:264-275.

(收稿日期:2011-04-06)

作者简介:

王二威,男,1985年生,硕士,助教,主要研究方向:软件工程,决策理论与方法。

侯福均,男,1967年生,博士,副教授,主要研究方向:决策理论与方法。

郑述招,男,1982年生,硕士,讲师,主要研究方向:系统工程。