

## 项目度量中对基本 COCOMO 进行校准的研究

周海玲, 孙涌

(苏州大学 计算机科学与技术学院, 江苏 苏州 215006)

**摘要:**所有成功的软件组织都将度量作为保证自己管理和技术质量的重要手段, 软件成本估计则是软件度量<sup>[1,2]</sup>的核心任务。为了提高成本估算的准确性, 文中根据特定软件企业中的历史项目数据对基本 COCOMO 模型进行校准, 在具体的参数修正方法上利用对数数据相关算法进行校正, 并与其它方法进行了比较, 得到了满意的结果。校准后的模型对项目开发成本的预测将会更加准确, 从而切实体现 COCOMO 成本度量工作对于软件项目的指导价值。因此, 文中所做的成本估算模型的校准工作, 对软件开发企业非常具有实用价值。

**关键词:**软件成本估算; 基本 COCOMO; 对数数据相关; 估算精度

**中图分类号:** TP311.5

**文献标识码:** A

**文章编号:** 1005-3751(2006)02-0023-03

## Research on Calibrating Basic COCOMO in Project Metrics

ZHOU Hai-ling, SUN Yong

(School of Computer Science and Technology, Soochow University, Suzhou 215006, China)

**Abstract:** All of successful software organizations regard metrics as an important means for their management and technological quality. In order to improve the accuracy of the basic COCOMO in process of project metrics, this paper will calibrate some parameters in the model according to software developing environment. So the model will yield more accurate effort predictions. Above all, provide a calibration methodology log data related to yield superior predicted estimates and contrast it with other approaches. The prediction of the development cost of the project of model after regulating will be more accurate, reflect COCOMO cost measure between work and guidance value, software of project conscientiously. So cost calibration to estimate model this text do, have practical value to software development enterprise.

**Key words:** software cost estimation; basic COCOMO; log data related; prediction accuracy

## 0 概述

COCOMO(Constructive Cost Model)是Boehm利用加利福尼亚的一个咨询公司的大量项目数据推导出的一个成本模型<sup>[3]</sup>。该模型于1981年首次发表了原始COCOMO模型。为适应软件工程领域的快速变化, COCOMO经过多次的更新, 如1987年的Ada版本, 并于1994年发展演变为COCOMO II模型<sup>[4,5]</sup>。

软件项目成本估算不是一劳永逸的活动, 它是随项目的进行而进行的一个逐步求精过程。早期的估算因未知因素很多, 虽然估算精度相对较低, 但意义较大, 对制定项目计划起到至关重要的作用。原始COCOMO中的基本模型以程序规模为基础, 在开发初期只要估算出程序规模即可估算出项目工作量及开发成本。但是, 该模型中的参数是从产生该模型的软件配置环境中依据经验推导出来

的, 在不同的软件开发环境下其成本估算精度不理想。鉴于此, 文中提出一种方法, 对基本COCOMO中相关参数根据各软件企业历史项目数据进行动态调整优化, 从而使估算出的参数切合本企业实际情况, 使之在项目开发中的估算精度大大提高, 很具实用价值。在校正过程中, 利用对数数据相关算法<sup>[6]</sup>进行参数估计, 并将结果与线性回归和Bayesian算法的结果进行对比, 得到了较理想结果。

## 1 基本 COCOMO 模型简介

1981年发表的原始COCOMO模型是一个分层次的系列软件成本估算模型, 包括基本模型、中级模型和详细模型3个子模型。

原始COCOMO模型中的基本模型的工作量估算公式为:

$$E = a \times S^b \quad (1)$$

其中,  $E$  表示以人月为单位的工作量,  $S$  表示以千代码行(KLOC)计数的程序规模,  $a$  和  $b$  是两个随开发模式而变化的因子。

## 2 基本 COCOMO 的参数校准方法

为了改善基本COCOMO的校正精度, 文中设计了利

收稿日期: 2005-05-05

基金项目: 国家“八六三”计划资助项目(2003AA430190)

作者简介: 周海玲(1973—), 女, 江苏徐州人, 硕士研究生, 研究方向为软件工程领域的项目管理与度量; 孙涌, 副教授, 硕士研究生导师, 主要从事软件工程、数据库及面向对象程序设计等方面的教学和研究工作。

用软件企业中历史项目数据进行参数校准优化的方法,首先对众多的历史项目数据进行整理分类,利用对数数据相关算法求出参数  $a$  和  $b$ 。求出的  $a$  和  $b$  应使各例子中实际工作量和使用该参数估算出的工作量误差为最小。其主要的工作流程如图 1 所示。

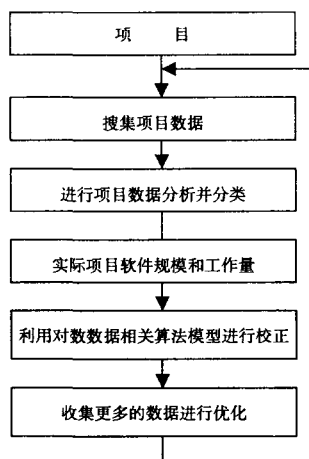


图 1 基本 COCOMO 模型进行参数校正的方法

第一环节:搜集项目数据。对于 COCOMO 的校正过程是一个基于数学建模和历史项目数据的处理过程,历史数据搜集的全面准确与否,严重影响着本模型校正的质量和精度。所以,若想进行 COCOMO 的校正处理,一定要有意识地全面搜集和整理以往完成的项目数据,项目数据越多,则进行校正后的 COCOMO 精度越高。

第二环节:进行项目数据分析与分类。在收集了足够多的项目数据之后,就要进行项目数据分析,分析的第一个步骤就是确定项目软件的类别,即确定软件项目的规模和性质。这是 COCOMO 校正的重要一步,有了这一步将极大地提高校正的准确性。项目软件的规模分类见表 1。

表 1 项目软件的规模分类

类型	小型	中型	大型	特大型
尺寸(KLOC)	≤50	50 < SIZE ≤ 300	300 < SIZE ≤ 1000	>1000

按软件的性质分类有:有机式、半分离式和嵌入式。有机式指项目相对简单,一小组有经验的程序员在极为熟悉的环境中开发软件;半分离式指介于有机式和嵌入式之间的中间方式,项目为中等规模,开发小组可能由经验不同的混合人员组成;嵌入式指项目必须在严格的约束条件下开发,要解决的问题很少见,因而无法借助于经验。

第三个环节:搜集项目软件的实际规模和工作量。

第四个环节:利用对数数据相关模型进行校正,求出常数项  $a$  和  $b$ 。为了保持一定的精度,校正工作中每种分类采集的数据点至少应为 10 个。

第五个环节:收集更多的数据进行动态的调整优化。参与校正的历史项目数据越多越及时,则该模型的校正精度就越高,专家们对新项目的成本估计的精确也越高,如此便可帮助企业有效地规避风险,保证项目软件的正常进度和软件产品的质量。

### 3 基本 COCOMO 模型校准算法的选择

对于基本 COCOMO 的参数校正,就是根据历史项目数据确定合适的  $a$  和  $b$ 。文中对基本 COCOMO 参数校正方法进行研究,经过大量计算与历史项目数据验证后认为,利用数据处理中的对数数据相关校正算法将会得到一个更为理想的结果。

文中以 Sunita Devnani - Chulani 的论文“Bayesian analysis of software cost and quality models”第 49 页中提供的 5 个历史项目数据及用 Bayesian 方法<sup>[7~10]</sup>处理的结果为例,分别用一元回归算法<sup>[6]</sup>和对数数据相关算法进行了处理,分析的情况如下:

例子项目数据:

工作量(PM)	代码行数(LOC)
4.8	3200
6.1	4500
7	5200
8	6000
9	7000

假设对于各个例子,它们都是相互独立的。对于第一个软件项目数据,开发一个 3200SLOC 的软件需要 4.8PM (PM=152 小时)。

3 种方法的校正结果见表 2。

表 2 3 种方法得到的校正参数

参数名称	线性回归方法	Bayesian 方法	对数数据相关方法
A	2.26	1.39	1.84
B	0.607	0.987	0.815

利用线性回归方法得到模型公式如下:

$$\text{Effort} = 2.26 \times \text{Size}^{0.607} \quad (2)$$

利用 Bayesian 方法得到模型公式如下:

$$\text{Effort} = 1.39 \times \text{Size}^{0.987} \quad (3)$$

利用对数数据相关方法得到模型公式如下:

$$\text{Effort} = 1.84 \times \text{Size}^{0.815} \quad (4)$$

为了验证 3 种校正结果,现将例子数据代入,并利用相对错误等级进行验证。相对错误等级 MRE (Magnitude of relative errors) 的定义为:

$$\text{MRE} = \left| \frac{\text{估计值} - \text{实际值}}{\text{实际值}} \right|$$

从表 3 中可以明确看出,在对基本 COCOMO 进行参数校正的方法中,线性回归法的估算误差是最大的,在本例中平均相对错误等级 MRE 达到了 11.766%。Bayesian 方法则比线性回归校正法精确度高的多,它的相对错误等级 MRE 只有 3.524%,比线性回归法的精度提高了 2 倍多。而对数数据相关校正法的平均相对错误等级 MRE 只有 1.142%,它比 Bayesian 校正法的估算精确度提高了 2 倍多,比线性回归法的估算精度提高了一个数量级。由此文中可以得出一个相当谨慎的结论:在基本 COCOMO 参数校准的算法中,对数数据相关算法是比 Bayesian 方法更为合适的算法。

表3 校正方法对比表

实际工作量	线性回归方法			Bayesian 方法			对数数据相关方法		
	估算工 作量	工作量 误差	MRE	估算工 作量	工作量 误差	MRE	估算工 作量	工作量 误差	MRE
4.8	4.579	0.221	4.60	4.381	0.419	8.73	4.748	0.052	1.08
6.1	5.631	0.469	7.69	6.134	0.034	0.56	6.269	0.169	2.77
7	6.148	0.852	12.17	7.075	0.075	1.07	7.053	0.053	0.76
8	6.706	1.294	16.18	8.148	0.148	1.85	7.925	0.075	0.94
9	7.364	1.637	18.19	9.487	0.487	5.41	8.896	0.014	0.16
平均 MRE	11.766			3.524			1.142		

#### 4 总结

进入 20 世纪 90 年代以来,软件工程领域一直在经历着巨大的变化,各种软件开发的模型、技术及方法不断出现,但是经过上述方法校正后的基本 COCOMO,利用软件公司中不断出现的历史项目数据校正优化后,可以不断随时间和软件工程界的变化而调整自己,使自己一直能保持在一个非常好的状态,从而保证了对新项目成本预测的精确度,更好地服务于项目度量。在国家“八六三”项目《装载机远程服务系统与智能挖掘机》及项目《工程机械远程故障诊断智能数据库及其应用》中,文中所研究的 COCOMO 成本估算及校正系统已在标准数据库无故障信号处理软件系统的开发中用于项目前期工作量的度量,有关实际数据正在进行统计、协调之中。

(上接第 22 页)

由于 Web Services 中的数据是以 XML 格式利用 SOAP 协议传输的,故下面就用 Web Services 方法的 SOAP 请求:

```
<? xml version="1.0" encoding="gb2312" ? >
< soap: Envelope xmlns: xsi = http://www. w3. org/2001/
XMLSchema - instance "xmlns: xsd = " http:// www. w3. org/
2001/XMLSchema "xmlns: soap = " http://schemas. xmlsoap. org/
soap/envelope />
< soap: Body >
< AddNewTicket xmlns = "http://tempuri. org/" >
< name >String< /name >
< number >int< /number >
< ordertimer >dateTime< /ordertimer >
< /AddNewTicket >
< /soap: Body >
< /soap: Envelope >
```

可以看到,整个 SOAP 消息包含在 SOAP Envelope 中。而 SOAP Body 中的请求消息的主要内容,包括调用的方法 AddNewTicket 和要传递给方法的参数 name, number, ordertime 和一些可选项。

该方法通过 HTTP 发送 SOAP 消息到平台后,可以利用 WSDL 信息进行 UDDI 注册。当用户需要该服务时 UDDI 又可以为其查找和绑定,让该方法对用户服务。

由于采用了 ASP.NET Mobile Controls 进行开发,代

#### 参考文献:

- [1] Jones C. Applied Software Measurement[M]. [s.l.]: McGraw - Hill, 1997.
- [2] 覃征,杨利英,高勇民,等. 软件项目管理[M]. 北京:清华大学出版社,2004.
- [3] Boehm B W. 软件工程经济学[M]. 李师贤等译. 北京:机械工业出版社,2004.
- [4] CSE Center for Software Engineering. COCOMO II Model [EB/OL]. <http://sunset.usc.edu/research/COCOMOII/>, 2005.
- [5] NASA National Aeronautics and Space Administration. Cost Estimating Web Site[EB/OL]. <http://www1.jsc.nasa.gov/bu2/COCOMO.html>. 2005.
- [6] McGarry J, Card D. 实用软件度量[M]. 吴超英,廖彬山译. 北京:机械工业出版社,2003.
- [7] Box G, Tiao G. Bayesian Inference in Statistical Analysis[M]. [s.l.]: Addison Wesley, 1973.
- [8] Devnani - Chulani S, Boehm B W, Steece B. Bayesian Analysis of Empirical Software Engineering Cost Models[J]. IEEE Transactions on Software Engineering, 1999, 25(4): 573 - 583.
- [9] 何渝. 计算机常用数值算法与程序(C++版)[M]. 北京:人民邮电出版社,2003.
- [10] Gelman A, Garlin J, Stern H, et al. Bayesian Data Analysis [M]. [s.l.]: Chapman & Hall, 1995.

码可以针对不同的移动设备解释成不同的无线标记语言,以适应特定设备的 Web 浏览器的方式呈现。在 ASP.NET 中可以用 DeviceSpecific/Choice 结构为不同的设备配置不同的控件属性,这样可以避免一些空间在某些设备上无法显示的错误。

#### 3 结束语

Web Services 和 ASP.NET Mobile Controls 的结合扩大了企业的服务范围,可以将企业集成方便地向移动设备延伸。用户能得到更好的服务,也方便了开发人员对系统的开发和维护。

#### 参考文献:

- [1] 许源,谢康林. 基于 Web Services 的电子商务移动扩展[J]. 微型电脑应用, 2004, 20(9): 19 - 21.
- [2] 马凌,陈洪亮. Web Services 技术在企业信息集成中的应用[J]. 微型电脑应用, 2004, 20(10): 19 - 22.
- [3] 梁俊斌,苏德富. 开发基于 Web Services 的多层分布式数据库[J]. 微机发展, 2004, 14(6): 45 - 47.
- [4] 基于 Windows Mobile 的 Pocket PC 和 Smartphone 的开发工具简介[EB/OL]. [www.microsoft.com/china/MSDN/library/archives/library/dnppcgen/html/devtoolsMobileapps.asp](http://www.microsoft.com/china/MSDN/library/archives/library/dnppcgen/html/devtoolsMobileapps.asp), 2003 - 08.
- [5] 张亚鹏. 基于 ASP.NET 的电子商务系统用户权限设计与实现[J]. 计算机安全, 2004(9): 33 - 35.