

航空机载软件测试质量评价方法研究

刘涛, 李娜

(西南电子技术研究所 天奥软件测评中心, 成都 610036)

摘要: 软件测试是提高软件产品质量和降低软件维护成本的重要手段; 针对高安全关键航空机载软件测试质量评价难题, 结合航空机载设备研制生存周期、航空机载软件研制生存周期和航空机载软件测试阶段, 提出一种全生存周期航空机载软件测试质量评价方法, 并建立了基于软件研制过程测试、三方测评、定型/鉴定测评和用户使用 4 个阶段的质量评价模型, 包括测试需求分析和策划质量、测试设计和实现质量、测试执行质量和测试总结质量 4 项活动和 18 种度量元; 通过工程实践证明, 该方法技术上可行, 具有评价要素更完整、评价模型更合理、评价结果更客观的特点, 为解决航空机载软件测试质量评价提供了一种新方法, 最终能达到降低软件维护成本和提高软件测试质量的目的。

关键词: 机载软件; 软件测试; 测试质量; 评价模型; 评价方法

Research on Quality Evaluation Method of Airborne Software Testing

Liu Tao, Li Na

(Dept. of software testing, Southwest China Institute of Electronic Technology, Chengdu 610036, China)

Abstract: Software testing is an important means to improve the quality of software products and reduce the cost of software maintenance. Aiming at the quality evaluation problem of high safety key airborne software testing, a whole life cycle is put forward in combination with the life cycle of airborne equipment development, the life cycle of airborne software development and the testing stage of aero airborne software. The quality evaluation model based on four stages of software development process testing, independent testing, finalization testing and user use evaluation is established, which includes 4 activities and 18 measure elements, such as test requirement analysis and planning quality, test design and implementation quality, test execution quality and test summary quality. Through engineering practice, it is proved that the method is feasible and has the characteristics of more complete evaluation elements, more reasonable evaluation model and more objective evaluation results. It provides a new method for solving the quality evaluation of airborne software testing, and can finally achieve the goal of reducing the cost of software maintenance and improving the quality of software testing.

Keywords: airborne software; software testing; testing quality; evaluation model; evaluation method

0 引言

航空电子系统是飞机上所有电子设备的总和, 涵盖通信、导航、雷达、电子战、飞行控制和管理等各个功能子系统, 同时也包括这些系统间用于信息交换和资源共享的信息综合系统, 通过机载网络和软硬件技术组成一个有机的整体。随着航空电子技术不断发展, 现代飞机的体系架构已由分离式、联合式逐渐向模块化、高度综合化方向发展, 系统功能越来越复杂, 软件实现的功能在航空电子系统中所占的比重越来越大, 航空机载软件具有嵌入式、实时性、结构复杂、高安全关键等级等特点^[1], 航空机载软件已成为决定航空电子系统乃至整个航空器质量的关键^[2]。

由于软件是航空机载装备的重要组成部分, 机载软件质量问题愈显突出, 甚至成为机载计算机软件能否发挥其优越性能的一个制约因素^[1], 软件测试是保障航空机载软件质量可靠性的重要手段, 也是机载软件全生存周期缺陷管理和分析的关键环节。目前航空机载软件测试是由国家

授权的软件测评机构依据国标、国军标、中国民用航空规章以及国际通行标准 (DO-178B/C 等) 考察航空机载产品中的软件是否符合研制总要求、软件研制任务书、软件需求规格说明、软件设计说明等技术文件中明确的要求和实现方式。软件测试质量的好坏直接影响软件测试结果的有效性。但由于航空机载软件的复杂性和特殊性, 虽然解决了如何测试的问题, 但软件测试质量如何评价是软件测试今后发展过程中需要面对的问题^[3]。建立软件测试质量模型和选择科学、合理的评价方法是软件测试质量评价的两个基本问题, 也是当前软件测试质量评价研究中的难点。国内外学者从测试过程、可靠性增长、缺陷残留风险、测试效率和测试成本等多个角度提出了许多有重要意义的软件测试质量评价方法^[4-6], 但缺乏从软件研制全生存周期进行量化评价的方法。

本文针对高安全关键航空机载软件测试质量评价难题, 提供一种能够降低软件维护成本, 提高软件测试质量, 全生存周期软件测试质量评价度量方法, 以帮助软件研发人员和项目管理人员客观、公正的量化软件测试质量, 从而持续提高软件质量^[7-8]。

收稿日期: 2018-04-23; 修回日期: 2018-05-23。

作者简介: 刘涛(1981-), 男, 四川通江人, 高级工程师, 硕士研究生, 主要从事软件工程化、软件测试技术方向的研究。

1 航空机载设备研制和软件测试阶段

根据航空机载设备研制管理要求，航空机载设备研制生存周期分为 F 方案研制阶段、C 型件研制阶段、S 型件研制阶段、D 定型阶段和 P 使用阶段^[9]，如图 1 所示。航空机载软件研制生存周期包括系统分析与设计、软件需求、软件设计、软件编码、软件集成、软件验收、系统集成、软件鉴定/定型、软件运行与维护。航空机载软件测试阶段包括研制过程测试阶段、三方测评阶段^[10]、鉴定/定型测评阶段、用户使用阶段。根据航空机载设备研制要求和型号软件工程化要求，航空机载软件在 C 型件研制阶段需开展研制过程测试质量度量，在 S 型件研制阶段需开展三方测评质量度量，在 D 定型阶段需开展鉴定/定型测评质量度量，三方测评和鉴定/定型测评质量度量是从缺陷的探测程度来衡量研制过程测试工作的质量，在 P 使用阶段需将软件产品交付用户使用后开展的用户使用质量度量，在一定时间内暴露出来缺陷数量的多少也是评价软件测试质量的一个重要维度。

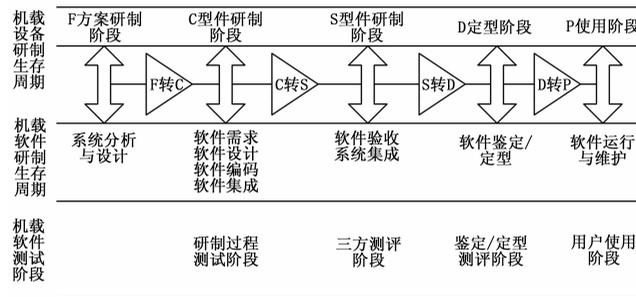


图 1 航空机载设备研制阶段和软件测试阶段对应关系图

2 航空机载软件测试质量评价模型

航空机载软件测试质量评价模型包括阶段层、活动层和度量元层（见图 2）。阶段层包括研制过程测试质量、三方测评质量、鉴定/定型测评质量和用户使用质量 4 个阶段，活动层包括测试需求分析和策划质量、测试设计和实现质量、测试执行质量和测试总结质量 4 项活动，度量元层包含 18 种度量元（见表 1），度量元的含义及计算方法见指标描述，度量元的取值范围采用归一化处理。针对机载软件测试评价开展的时机、技术和管理要求，度量元的种类和数量可以不同，因项目需要而设置。

航空机载软件研制过程测试是由研制单位按照项目总体测试要求开展的软件内部测试工作，包括测试需求分析和策划、测试设计和实现、测试执行和测试总结 4 项活动，研制过程测试质量根据测试策划质量、测试设计质量、测试执行质量和测试总结质量进行质量度量。测试策划质量包括显性需求覆盖率、隐性需求覆盖率、测试策略充分性和测试环境差异性 4 种度量元，测试设计质量包括测试需求覆盖率、测试用例颗粒度 2 种度量元，测试执行质量包括测试用例执行率、回归测试用例比例和软件缺陷接受率 3 种度量元，测试总结质量包括组织资产贡献度、测试工具分析度和测试人员能力分析度 3 种度量元。

表 1 软件测试质量度量元

序号	度量元	取值范围	指标描述	指标意义
1	显性需求覆盖率	[0,1]	测试项对需求文档中明确描述的需求点覆盖的百分比	越大越好
2	隐性需求覆盖率	[0,1]	测试项对需求文档中未明确描述但需软件实现需求点覆盖的百分比	越大越好
3	测试策略充分性	[0,1]	测试项中采用的测试策略占所有策略的百分比	越大越好
4	测试环境差异性	[0,1]	测试环境中使用真实设备占所有设备的百分比	越大越好
5	测试需求覆盖率	[0,1]	测试用例对测试项覆盖的百分比	越大越好
6	测试用例颗粒度	[0,1]	测试用例总数占机载软件有效代码执行的百分比	越大越好
7	测试用例执行率	[0,1]	测试用例执行数量占测试用例总数的百分比	越大越好
8	回归测试用例比例	[0,1]	回归测试用例数占测试用例总数的百分比	越大越好
9	软件缺陷接受率	[0,1]	开发人员接受缺陷数占测试人员提交软件缺陷总数的百分比	越大越好
10	组织资产贡献度	[0,1]	形成组织资产的典型缺陷数和可复用的用例数占有效软件缺陷和测试用例总数的百分比	越大越好
11	测试工具分析度	[0,1]	已分析的测试工具数占使用测试工具总数的百分比	越大越好
12	测试人员能力分析度	[0,1]	测试人员技能分析数占参与测试人员总数的百分比	越大越好
13	三方测评缺陷发现率	[0,1]	三方测评发现的缺陷数/(研制过程测试发现的缺陷数+三方测评发现的缺陷数)	越小越好
14	三方测评严重缺陷发现率	[0,1]	三方测评发现的严重和关键缺陷数/(研制过程测试发现的严重和关键缺陷数+三方测评发现的严重和关键缺陷数)	越小越好
15	鉴定/定型测评缺陷发现率	[0,1]	鉴定/定型测评发现的缺陷数/(研制过程测试发现的缺陷数+三方测评发现的缺陷数+鉴定/定型测评发现的缺陷数)	越小越好
16	鉴定/定型测评严重缺陷发现率	[0,1]	鉴定/定型测评发现的严重和关键缺陷数/(研制过程测试发现的严重和关键缺陷数+三方测评发现的严重和关键缺陷数+鉴定/定型测评发现的严重和关键缺陷数)	越小越好
17	用户使用缺陷发现率	[0,1]	用户使用发现的缺陷数/(研制过程测试发现的缺陷数+三方测评发现的缺陷数+鉴定/定型测评发现的缺陷数+用户使用发现的缺陷数)	越小越好
18	用户使用严重缺陷发现率	[0,1]	用户使用发现的严重和关键缺陷数/(研制过程测试发现的严重和关键缺陷数+三方测评发现的严重和关键缺陷数+鉴定/定型测评发现的严重和关键缺陷数+用户使用发现的严重和关键缺陷数)	越小越好

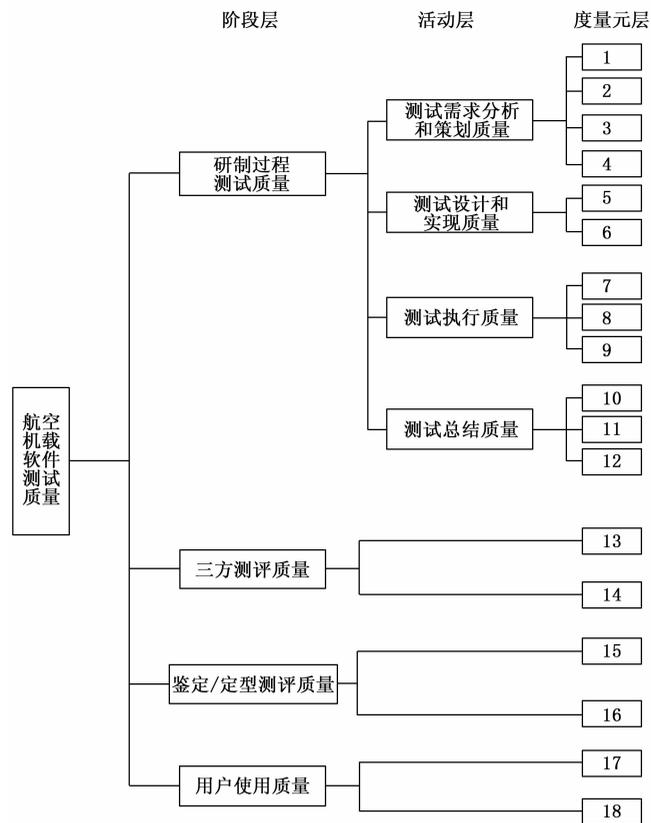


图 2 航空机载软件测试质量评价模型

在研制过程测试后,项目总体单位委托具有软件测评资质的测评机构依据三方测评方案和实验室质量管理体系开展第三方测评。三方测评质量是从缺陷的探测程度来衡量软件测试工作的质量,由于测试不可能实现穷举,需在一定的时间和人力成本内进行,因此所有的测试工作理论上均不可能探测到所有缺陷,即不存在完美的软件或系统。三方测评质量采用三方测评结果与研制过程测试结果进行对比的方式进行质量评价,三方测评质量包含三方测评缺陷发现率和三方测评严重缺陷发现率 2 个度量元,属于减分项。

在三方测评后,项目主管机关委托具有软件测评资质的测评机构依据鉴定/定型测评方案和实验室质量管理体系开展鉴定/定型测评,鉴定/定型测评质量采用鉴定/定型测评结果与研制过程测试结果和三方测评结果进行对比的方式进行质量评价,鉴定/定型测评质量包括鉴定/定型测评缺陷发现率和鉴定/定型测评严重缺陷发现率 2 个度量元,属于减分项。

在鉴定/定型测评后,航空机载软件灌装于机载装备并交付用户,在一定时间内用户使用过程中暴露出来的问题数量也是评价软件测试质量的一个重要维度,用户使用质量采用用户使用结果与研制过程测试结果、三方测评结果和鉴定/定型测评结果进行对比的方式进行质量评价,用户使用质量包括用户使用缺陷发现率和用户使用严重缺陷发现率 2 个度量元,属于减分项。

3 航空机载软件测试质量评价实施过程

航空机载软件测试质量采用下面 6 个步骤对被测软件测试质量进行评价。

第一步:确定研制过程测试阶段中测试需求分析和策划、测试设计和实现、测试执行、测试总结 4 项活动的权重系数 a ,权重系数按顺序定义为 $\{a_0, a_1, a_2, a_3, \dots, a_i\}$,且满足式(1)。每个活动 a_i 包含若干度量元,确定每个度量元的权重系数 b ,按顺序定义为 $\{b_0, b_1, b_2, \dots, b_m\}$,满足式(2);收集研制过程测试阶段各度量元数值 $M(k)$, a_i 活动的质量评价数值为 t_i ,由式(3)得出;按照式(4)计算出研制过程测试质量评价价值 T_1 。

$$\sum_{i=0}^3 a_i = 1 \tag{1}$$

$$\sum_{j=0}^m b_j = 1 \tag{2}$$

$$t_i = \sum_{k=0}^m (b_k * M(k)) \tag{3}$$

$$T1 = \sum_{i=0}^3 (a_i * t_i) \tag{4}$$

第二步:确定三方测评阶段中三方测评缺陷发现率和三方测评严重缺陷发现率两个度量元的权重系数,收集三方测评阶段两个度量元数值,计算出三方测评质量评价价值 T_2 。

第三步:确定鉴定/定型测评阶段中鉴定/定型测评缺陷发现率和鉴定/定型测评严重缺陷发现率两个度量元的权重系数,收集鉴定/定型测评阶段两个度量元数值,计算出鉴定/定型测评质量评价价值 T_3 。

第四步:确定用户使用阶段中用户使用缺陷发现率和用户使用严重缺陷发现率两个度量元的权重系数,明确一段时间内收集用户使用阶段两度量元数值,计算出用户使用质量评价价值 T_4 。

第五步:按照式(5)计算航空机载软件测试质量评价价值 T 。

$$T = T1 - T2 - T3 - T4 \tag{5}$$

第六步:航空机载软件测试质量评价价值 T 与航空机载软件研制单位组织测量库的质量评价基准值进行比较,最终确定航空机载软件测试的质量优劣。

4 实验结果和分析

某航空机载软件研制单位采用图 1 的质量评价模型拟对 3 个被测对象(A、B、C)开展软件测试质量定量评价。假设研制过程测试质量总分为 100,三方测评质量总分为 10、鉴定/定型测评质量总分为 20,用户使用质量总分为 30,其中属于减分项的质量总分为 50。研制单位组织测量库的质量评价基准值为优秀(大于等于 80 分)、合格(大于等于 60 分且小于 80 分)和不合格(小于 60 分)。

步骤 1:设置航空机载软件研制过程测试 4 项活动的权重分别为 $\{0.4, 0.3, 0.2, 0.1\}$;研制过程测试、三方测评、鉴定/定型测评和用户使用的度量元采取默认平均权重方式确定权重。