航天嵌入式软件可信性度量方法及应用研究

王婧^{①*}, 陈仪香^{②*}, 顾斌^①, 郭向英^①, 王保华^②, 金晟毅^③, 徐建^①, 张居阳^②

① 北京控制工程研究所, 北京 100190;

② 华东师范大学软件学院, 上海 200062;

③ 中国空间技术研究院总体设计部, 北京 100094

* E-mail: yxchen@sei.ecnu.edu.cn

收稿日期: 2014-12-03; 接受日期: 2014-12-27

国家中长期科技发展规划重大专项和国家自然科学基金(批准号: 91118007)资助项目

摘要 航天嵌入式软件的可信性是确保型号任务成功的重要因素。目前对航天嵌入式软件可信性评价均采用定性的方法, 还没有针对航天嵌入式软件可信性的定量度量方法和模型。本文采用主观和客观相结合的方法, 提出了一种面向航天型号软件的可信性度量模型及分级评价方法, 通过航天软件的可信属性来研究软件的可信度和定量分析评价。基于该度量模型和方法对月地高速再入返回飞行器软件的可信性进行了度量评估, 应用情况表明, 航天嵌入式软件可信性度量模型及分级评价方法是合理的, 科学的, 能够有效评价软件的可信性并发现软件产品研制过程中需要加强的部分, 对于航天型号软件研制水平的提高有重要的意义。

关键词

航天嵌入式软件
软件可信性
度量模型
分级模型
可信属性

1 引言

随着我国航天事业的快速发展, 软件在航天器中的作用和地位越来越突出, 航天器上很多关键, 复杂的功能都逐渐由软件来实现^[1], 软件已成为空间飞行器系统的核心。而软件复杂度和规模的增长导致软件缺陷增多, 国内外由于软件缺陷所引发的灾难事故频发。1996 年欧洲航天局阿丽亚娜 5 型火箭发射后 40 s 即爆炸, 原因是惯性导航系统软件设计的失误^[2]。2003 年发生的北美大停电事件也是由于电力控制系统软件失效引起。因此, 航天嵌入式软件的可信性已成为确保型号任务成功的重要因素。

软件的可信性是指软件系统的动态行为及结果总是符合预期目标, 能适应一定环境和需求因素的

变化, 具有抵御异常情况的能力, 不受偶然的或恶意的原因而破坏, 以至遭遇外界干扰时系统仍然能够保持安全可靠地运行^[3]。

软件可信性度量与评估是对软件的可信性进行量化评价与评估, 通过评估对增强软件可信性的设计与开发提供指导。目前型号发射前主要依靠专家经验对软件的质量进行定性评价, 还没有针对航天嵌入式软件的可信性定量度量方法和模型。航天嵌入式软件在软件可信性理论, 技术的掌握和工程化应用等方面比较薄弱, 定性评价的方法难以为航天嵌入式软件可信开发提供规范性参考和指导。

本文采用主观和客观相结合的方法, 提出了一种航天嵌入式软件可信性定量度量方法, 其中包括基于多维属性的软件可信性度量计算模型和软件可

引用格式: 王婧, 陈仪香, 顾斌, 等. 航天嵌入式软件可信性度量方法及应用研究. 中国科学: 技术科学, 2015, 45: 221-228
Wang J, Chen Y X, Gu B, et al. An approach to measuring and grading software trust for spacecraft software (in Chinese). Sci Sin Tech, 2015, 45: 221-228, doi: 10.1360/N092014-00479

信性分级评价方法. 基于该度量模型和方法对月地高速再入返回飞行器软件的可信性进行了量化评估, 应用情况表明, 本文提出的航天嵌入式软件可信性度量模型及分级评价方法是合理的、科学的, 能够有效评价航天嵌入式软件的可信性并发现软件产品研发过程中需要加强的部分, 对提高航天型号软件质量有重要的指导意义.

2 基于多维属性的软件可信性度量模型

对于软件可信性度量研究有各类不同的方法, 有的通过软件进化来研究软件可信度^[4], 有的通过共生学说研究^[5], 有的则通过数据挖掘角度来研究^[6], 本文通过软件可信属性来研究软件的可信度^[7~23]. 通常有许多因素影响软件的可信度, 我们把这些因素称为软件的可信属性^[7,8], 不同的软件着重于不同的可信属性, 要得到软件的可信度, 可以通过先得到这个软件在其可信属性上的可信程度, 然后综合得到软件可信度^[9~15]. 当一个可信属性还可以细分为更详细的子属性时, 可以通过综合这些子属性的可信度来得到这个可信属性的可信度; 而当一个属性不可细分时, 我们称其为度量元, 度量元只能通过分析软件的可信证据来得到. 所以软件可信度量研究最重要的 2 个环节就是求出度量元的可信度和将度量元和子属性评价综合成可信属性以及软件可信度, 后者我们称其为软件可信度量计算模型. 我们可以如下定义通用的基于属性的软件可信度量计算模型:

$$T = f(y_1, \dots, y_n, \alpha_1, \dots, \alpha_n), \quad (1)$$

其中, T 为软件可信度; f 为度量计算模型; y_i 为可信属性的可信度; α_i 为可信属性 y_i 的权重.

对于多个度量元评价结果的综合, 最简单的办法就是通过一定的规则直接判断和求它们的算术平均值. 文献[16,17]将软件划分为若干等级, 然后依据对软件可信性指标度量结果判断是否满足给定的条件, 最后根据判断结果确定软件的可信等级. 针对软件可信性度量计算模型的选取问题, 文献[12~15]提出软件可信度量模型应该满足的 4 条基本准则, 分别如下.

1) 单调性. 单调性用来保证当软件的可信属性提升时, 软件可信度也提升. 即度量函数 T 对于每一个可信属性 y_i 单调递增. 因此有

$$\frac{\partial T}{\partial y_i} > 0. \quad (2)$$

2) 凝聚性. 凝聚性用来保证上述的提升速度是递减的, 即保持其他 $y_i (i \neq j)$ 不变, 只改变 y_i 的值, 则 y_i 的提升速度对软件可信度的影响在减少. 因此

$$\frac{\partial^2 T}{\partial^2 y_i} < 0. \quad (3)$$

3) 灵敏性. 改变 y_i 值引起的总度量值的变化与其属性权重有关, 其中灵敏性函数定义为

$$\frac{\partial T}{\partial y_i} \frac{y_i}{T} = \lim_{\Delta y_i \rightarrow 0} \frac{\frac{\Delta T}{T}}{\frac{\Delta y_i}{y_i}}. \quad (4)$$

4) 替代性. 一个可信属性 y_i 的值改变可以通过其他属性 $y_j (i \neq j)$ 的值替代来实现, 以求保持总度量值不变. 但这种替代是与属性重要性相关的, 重要属性(关键属性)与非重要属性(非关键属性)相互替代是相对较难的. 即

$$\delta_{ij} = \frac{d(\frac{y_i}{y_j})}{dh_{ij}} \times \frac{h_{ij}}{\frac{y_i}{y_j}}, \quad (5)$$

其中 $h_{ij} = \frac{dy_i}{dy_j}$. δ_{ij} 值的大小决定了相互替代性难易程度, 值越大替代越容易.

很明显, 加权求和模型不具备准则 2) 凝聚性.

基于这 4 条构造准则, 文献[23~25]设计了软件可信性度量模型:

$$T = \left\{ \alpha \left[\min_{1 \leq i \leq m} \left\{ \left(\frac{y_i}{10} \right)^\varepsilon \right\} y_1^{\alpha_1} y_2^{\alpha_2} \dots y_m^{\alpha_m} \right]^\rho + \beta \left[y_{m+1}^{\beta_{m+1}} y_{m+2}^{\beta_{m+2}} \dots y_{m+s}^{\beta_{m+s}} \right]^\rho \right\}^{-\frac{1}{\rho}}, \quad (6)$$

其中, α 表示关键属性的权重, β 表示非关键属性的权重; α_i 表示各关键属性在关键属性中的权重, β_i 表示各非关键属性在非关键属性中的权重; ε 表示调控参数, 调控最小关键属性对软件可信性的影响; ρ 是替代系数, ρ 越大关键属性和非关键属性之间的替代性就越难.

运用此模型可以有效地通过可信子属性或者度量元计算出软件总体可信度. 文献[18]在以上 4 条构

造准则基础上, 提出了第 5 条准则: 稳定性.

5) 稳定性. 稳定性用来描述软件可信评估值在所有属性可信评估值的范围之内, 即, 如果 $\forall i, y_i \in [a, b]$, 那么 $T \in [a, b]$.

3 航天型号软件可信性度量及评估方法

航天型号软件可信性度量及评估主要包括基于等级的子属性评分, 软件可信性度量计算和软件可信性分级评定 3 部分. 子属性的评分采用主观和客观结合的方法, 由领域专家根据软件客观数据通过主观判断给出评分. 图 1 为航天软件可信度量评估体系结构, 其中 Z_{ij} 为第 i 个属性中第 j 个子属性的可信度, Z_{ij}^k 为第 k 个专家对第 i 个属性中的第 j 个子属性的评分, y_i 为第 i 个属性的可信度, T 为软件的可信度, L 为软件的可信级别. 首先由专家采用基于等级的评分方式给出每个子属性的评分 Z_{ij}^k , 随后根据软件可信性度量计算模型计算出每个子属性的可信度 Z_{ij} , 每个属性的可信度 y_i 及软件的可信度 T , 最后综合给出软件的可信级别 L .

3.1 基于等级的子属性评分

为了对航天型号软件进行全面的可信性度量, 从产品特性, 开发过程 2 个维度建立了包含 9 个属性, 28 个子属性, 103 个度量元的 3 层可信属性结构, 如图 2 所示.

然而, 由于度量元的数目较多, 如果采用对每个

度量元评分的方式, 会产生过大的时间成本, 在对实际软件产品的可信性度量过程中, 难以对每个度量元的可信度打分.

另一方面, 在提取的度量元中, 各度量元的可信度对子属性可信度的影响方式不同, 各度量元的评分方式也不同. 按照度量元的可信度对子属性可信度的影响方式, 可将全部的度量元划分基础型度量元, 增强型度量元和混合型度量元. 按照度量元的评分方式, 可将全部度量元划分为二值型度量元和程度型度量元. 假设对于子属性的可信度 Z_{ij} , $0 \leq Z_{ij} \leq 100$ 且 S 表示可信和不可信的分界. 基础型度量元的评分影响 Z_{ij} 在 $[0, S]$ 的变化, 增强型度量元的评分影响 Z_{ij} 在 $[S, 100]$ 的变化, 混合型度量元的评分影响 Z_{ij} 在 $[0, 100]$ 的变化. 对于二值型度量元, 度量元的分值只有 0 和 100, 对于程度型度量元, 度量元的分值可以在多个分值之间变化. 因此, 如何将多种类型的度量元的分值计算综合为子属性的可信度是软件可信度量过程中难题.

基于此, 本文提出基于等级的子属性评分方式: 将子属性的可信度划分为 4 个等级: A, B, C, D, 其中 A 为最高级别, 依次递减, 每个等级规定了该子属性包含的全部度量元必须达到的可信度. 其中, 子属性可信度等级划分的原则为

- 1) 高等级中各度量元的可信度均高于或等于低等级中各度量元的可信度;
- 2) 最高级中各度量元的可信度均为最高;
- 3) 所有不满足次低级别度量元可信度要求的情况均划为最低级别.

3.2 软件可信性度量计算

航天嵌入式软件主要关注软件功能的正确性, 性能指标符合性, 可靠性以及应对故障情况的能力, 所有度量属性均为关键属性. 因此, 为了面向航天型号软件的可信性度量, 对第 2 节中的度量计算模型进行了简化, 从中去除了关键属性和非关键属性的区分, 设定所有可信属性都为关键属性, 同时设置 ε 调控参数为 0, 即所有可信属性对软件可信性的影响都取决于属性的权重. 简化后的可信性度量计算模型为

$$\begin{cases} \alpha_1 + \alpha_2 + \cdots + \alpha_n = \sum_{i=1}^n \alpha_i = 1 (0 \leq \alpha_i \leq 1), \\ T = y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_n^{\alpha_n} = \prod_{i=1}^n y_i^{\alpha_i} (0 \leq y_i \leq 1), \end{cases} \quad (7)$$

图 1 软件可信度量评估体系结构图

图 2 软件可信属性结构图

其中, T 为软件可信度, $0 \leq T \leq 1$; y_i 为第 i 个属性的可信度; α_i 为第 i 个属性的权重值。

简化后的模型满足第 2 节中的单调性, 凝聚性, 灵敏性, 替代性和稳定性这 5 条基本准则。

直观上分析, 此模型满足“木桶原理”, 即某软件若要达到某一可信度值, 那么必须每项可信属性的可信度都需要达到一定限度的可信度值, 否则就没法最终达到总体可信度。可以将模型中各属性形象地理解为电路设计中构成串联关系的元器件。

基于 3.1 节中的航天软件可信属性结构以及基于等级的子属性评分方式, 属性的可信度 y_i 需要由各专家给出的子属性可信度级别通过以下计算得出。

令 m 为专家人数, k 表示第 k 位专家 ($1 \leq k \leq m$); n 为软件的可信属性个数, i 表示第 i 个可信属性 ($1 \leq i \leq n$); ω_i 为第 i 个可信属性的子属性个数, j 表示第 j 个子属性 ($1 \leq j \leq \omega_i$); Z_{ij} 为第 i 个属性中第 j 个子属性的可信度, Z_{ij}^k 为第 k 个专家对第 i 个属性中的第 j 个子属性的评分; α_i 为第 i 个属性的权重; β_{ij} 为第 i 个属性中第 j 个子属性所占权重; y_i 为第 i 个属性的可信度值; T 为软件的可信度值。

由于专家对每个子属性按 4 个等级进行评分, 因此在计算某个子属性的可信度时, 首先将子属性可信度等级转换为具体的数值 Z_{ij}^k 。

定义子属性可信度等级转换为具体的数值的转换方式为

- A: $Z_{ij}^k = 100$,
- B: $Z_{ij}^k = 90$,
- C: $Z_{ij}^k = 70$,
- D: $Z_{ij}^k = 20$.

第 j 个子属性的可信度为

$$Z_{ij} = \sum_{k=1}^m Z_{ij}^k / m, \quad (8)$$

第 i 个属性的可信度为

$$y_i = Z_{i1}^{\beta_{i1}} \cdot Z_{i2}^{\beta_{i2}} \cdots Z_{i\omega_i}^{\beta_{i\omega_i}}, \quad (9)$$

软件的可信度值为

$$T = y_1^{\alpha_1} \cdot y_2^{\alpha_2} \cdots y_n^{\alpha_n}. \quad (10)$$

3.3 软件可信性分级评定

软件的可信级别不仅仅取决于软件的可信度值, 它还与每个可信属性的可信度值相关。本文通过可信属性的可信度以及软件的可信度值综合评定软件的可信级别。

基于可信值越大提高的难度越大, 可信级别高对属性的可信性值要求也高的原则, 在可信分级模型中, 分级值域区间不是等分的。从最低级向上增加可信级别的值区间近似黄金分割的比例递减, 有

$$\begin{aligned} 45 \cdot (\sqrt{5} - 1) / 2 &\approx 25, \\ 25 \cdot (\sqrt{5} - 1) / 2 &\approx 15, \\ 15 \cdot (\sqrt{5} - 1) / 2 &\approx 10, \\ 10 \cdot (\sqrt{5} - 1) / 2 &\approx 5. \end{aligned} \quad (11)$$

根据(11)式中的近似黄金分割的比例递减区间划分方式, 提出了软件可信分级模型, 如表 1 所示, 软件可信级别的评定需要软件的可信度值 T 和属性的可信度 y_i 分别满足表中软件可信度值要求和属性可信度值要求。其中分别使用 I, II, III, IV, V 表示五个不同的可信等级, I 级可信等级最低, 依次递增, V 级可信等级最高。

在表 1 中, 可信性级别评定的值完全覆盖软件可信性度量值的可能分布区间。 T 在每个区间值的每一种情况都可以且只可以评定一个级别, 保证了其可信等级的唯一性。

此外, 级别的评定需要软件既要满足软件可信度量值的要求, 同时每个属性也要达到相应的单个属性要求。对于相同度量值的 2 个软件, 若其中 1 个软件的某一属性未达到单个属性要求, 那么它的可信等级就会降低。

表 1 软件可信分级评定表

软件可信度值要求	属性可信度值要求	级别评定
$95 \leq T$	低于 95 分的属性个数不超过 3 个且没有低于 85 分的属性	V
$85 \leq T < 95$ 或者 $T > 95$ 且不能评为“非常可信”级别	低于 85 分的属性个数不超过 3 个且没有低于 70 分的属性	IV
$70 \leq T < 85$ 或者 $T > 85$ 且不能评为“比较可信”级别及以上者	低于 70 分的属性个数不超过 3 个且没有低于 45 分的属性	III
$45 \leq T < 70$ 或者 $T > 70$ 且不能评为“弱可信”级别及以上者	低于 45 分的属性个数不超过 3 个	II
$T < 45$ 或者 $T > 45$ 且不能评为“不可信”级别及以上者	无要求	I

4 应用及分析

2014 年 4 月开展了针对月地高速再入返回飞行器 23 个软件的专项审查, 在审查中采用本文提出的航天器软件可信性度量与评价方法对以上软件的可信度进行度量. 这是首次采用定量分析方法对航天嵌入式软件进行评价.

由 10 名专家组成软件专项审查专家组对软件进行评分, 软件的代码总规模约 30 万行. 评分的方式是各专家采用主观与客观结合的方式对 3.1 节提到的 28 个可信子属性按照 A, B, C, D 四个等级进行基于等级的评分. 在参与审查的 23 个软件中, 本文选取具有代表性的 11 个软件的评分结果进行以下分析和统计.

4.1 子属性的评分分布情况

图 3 为各子属性的评分分布情况, 即对于各个子属性统计评分分别为 A, B, C, D 等级的软件个数.

可以看出, 个别子属性的评分分布不太符合正

态分布. 产生这个现象的原因有以下 2 个.

1) 部分子属性的分级存在不足. 有些子属性 A 级和 B 级的要求较高, 有些要求较低, 对于这些子属性, 每个等级规定的度量元可信度和度量元可信度的组合方式还存在需要进一步完善.

2) 航天型号软件研制存在共性. 航天型号软件在一些方面普遍完成较好, 例如, 由于航天型号软件均按照 GJB5000A 标准进行软件开发, 因此软件变更控制情况普遍情况良好; 同时也存在一些共性问题, 例如, 由于动态时序, 空间使用, 数据使用, 控制行为等方面的设计, 测试验证技术不先进, 因此普遍在专项测试方面存在不足.

因此, 虽然部分子属性的分级还存在不足, 但是子属性的评分与软件产品研制的实际情况是符合的, 真实反应了软件在各属性上的可信度.

4.2 软件可信性度量计算及分级情况

图 4 所示为运用本文提出的航天软件可信性度

图 3 各子属性的评分分布情况

图4 属性可信度及软件可信度分值分布情况

量得出的各属性的可信度和软件可信度。

分析图形曲线可以看出,度量计算模型不会因为个别子属性得分过高而过大提高整体软件的可信度值,而得分较低的子属性会拉低整体软件可信度值。即软件若要达到某一可信度值,那么必须每项可信属性的可信度都需要达到一定限度的可信度值,否则就没法最终达到总体可信度,符合“木桶原理”。

表2为11个软件的可信等级。其中3个软件的可信等级为IV级,其余软件的可信等级为III级,全部软件均达到较高的可信等级。从中也可以看出,软件想要达到最高可信级别V级是很难的,这符合可信值越大提高的难度越大,可信级别高对属性的可信性值要求也高的原则。

对于航天型号软件,不仅需要软件整体可信度达到要求,还需要每个属性的可信值均达到对应的要求。因此,级别的评定需要软件既要满足软件可信

表2 各软件可信等级分布情况

软件编号	软件可信度值	可信等级
2	82.45634	III
4	80.23482	III
6	78.89698	III
7	80.28511	III
9	80.69644	III
18	90.25450	IV
19	88.26221	IV
20	82.00907	III
21	82.21979	III
22	81.46237	III
23	90.23548	IV

度量值的要求,同时每个属性也要达到相应的单个属性要求。对于相同度量值的2个软件,若其中1个软件的某一属性未达到单个属性要求,那么它的可信等级就会降低。这种设计是符合航天软件的可信度要求需要和特点的,适用于航天型号软件的可信分级。在本次度量的软件中没有软件因为单个属性的可信度不满足可信级别要求而降级。

4.3 改进建议

应用情况及分析数据表明,本文提出的航天嵌入式软件可信性度量计算模型及分级评价方法是科学,合理,有效的,能够反映出软件之间存在的相对差异。在本次软件评价中,由于选取的度量属性还不够全面,导致个别软件得分偏低。如何更全面地度量评价软件本身的可信水平,还需要进一步研究和思考。

另一方面,航天型号软件在动态时序,空间使用,数据使用,控制行为等专项测试方面还需加强。目前,相关技术及学术成果还没有完全转化为工具或工具仍存在不完善。因此,尤其需要加强相关工具的研究力度,并在航天型号软件的研制过程中大规模推广应用。

5 总结

本文针对航天软件可信性度量与评估问题,运用基于多维属性的软件可信性度量模型,设计了航天软件可信性度量及评估分级方法,并在型号软件研制中进行了应用。应用情况表明,面向航天软件的可信性度量模型及可信性评估分级方法适用于航天型号软件的可信性量化分析。度量评估模型和方法,以及属性和子属性设计是合理,科学的,模型能够有效识别软件产品研制过程中的需要加强的部分,对于航天型号软件研制水平的提高有重要的意义。为了进一步提升航天嵌入式软件可信性,下一步工作将依据本文提出的可信度量模型与评估方法,制定航天嵌入式软件可信开发规范,标准和指南。

参考文献

- 1 陈建新,张志,王磊,等.嫦娥三号巡视器综合电子系统的设计与实现.中国科学:技术科学,2014,44:450-460

- 2 Sekding P B. Faulty software caused Ariane 5 failure. *Space News*, 1996, 25: 24–30
- 3 Liu K, Shan Z G, Wang J, et al. The trusted software basic research summary of major research plan. NSFC, 2008, 145–151
- 4 Yang J. A classification evaluation model for software trustworthiness based on trustworthiness evolution. In: *Proceedings of BMEI 2011*. Guangzhou, 2011. 222–227
- 5 Tang Y X, Liu S X, Liu Z L. The metric model of software trustworthiness based on symbiosis theory. In: *Proceedings of EMEIT 2011*. Harbin, 2011. 3465–3469
- 6 Yuan Y Y, Han Q. Data Mining Based Measurement Method for Software Trustworthiness. In: *Proceedings of IPTC 2010*. Huanggang, 2010. 293–296
- 7 Avizienis A, Laprie J C, Randell B. Fundamental concepts of dependability. *ISW-2000*. 2000
- 8 Avizienis A, Laprie J C, Randell B, et al. Basic concepts and taxonomy of dependable and secure computing. *IEEE T Depend Secure*, 2004, 1: 11–33
- 9 Cai S B, Zhou Y Z. A framework to support a credible assessment of the software resources. *J Software*, 2010, 21: 359–372.
- 10 Zhen Z M. The software credibility kinetic characteristics and evolution of complexity. *Sci China Inform Sci*, 2009, 39: 946–950
- 11 Zhen Z M. The complexity of its dynamics software trustworthiness statistical analysis method. *Sci China Inform Sci*, 2009, 39: 1050–1054
- 12 陶红伟. 基于属性的软件可信性度量模型研究. 博士论文. 上海: 华东师范大学, 2011
- 13 Tao H W, Chen Y X. A New Metric Model for Trustworthiness of Softwares. In: *Proceedings of ICISA 2010*. Seoul, 2010. 69–72
- 14 Tao H W, Chen Y X. A Metric Model for Trustworthiness of Softwares. In: *Proceedings of WI-IAT 2009*. Milan, 2009. 69–72
- 15 Tao H W, Chen Y X. A New Metric Model for Trustworthiness of Softwares. *Telecommun Syst*, 2012, 51: 95–105
- 16 Lang B, Liu X D. A software trustworthiness classification model. *Comput Sci Expl*, 2010, 4: 231–239
- 17 Amoroso E, Watson J, Nguyen T, et al. Toward an approach to measuring software trust. In: *Proceeding of IEEE Symposium on Research in Security and Privacy*. Oakland, 1991. 198–218
- 18 Zhang L W, Zhou Y, Chen Y X, et al. Stability of software trustworthiness measurements models. In: *Proceeding of SSCPS 2013*. Gaithersbur, 2013. 219–224
- 19 Chen L, Ping C, Liu W. The model and method of trustworthiness level evaluation for software product. In: *Proceedings of ICNC 2010*. Yantai, 2010. 709–715
- 20 Rong M. A model for CPS software system trustworthiness evaluation based on attributes classifying. In: *Proceeding of ICCSE 2013*. Colombo, 2013. 1309–1314
- 21 Yang S L, Ding S, Zhu W. Trusted software evaluation method based on utility and evidence theory. *Computer Res Develop*, 2009, 46: 1152–1159
- 22 Shi L, Yang S L. The evaluation of software trustworthiness with FAHP and FTOPSIS methods. In: *Proceeding of CiSE 2009*. Wuhan, 2009. 1–4
- 23 Bao T, Liu S F, Han L. Research on an analysis method for software trustworthiness based on rules. In: *Proceedings of CSCWD 2010*. Shanghai, 2010. 43–47

An approach to measuring and grading software trust for spacecraft software

WANG Jing¹, CHEN YiXiang², GU Bin¹, GUO XiangYing¹, WANG BaoHua², JIN ShengYi³, XU Jian¹ & ZHANG JuYang²

¹ *Beijing Institute of Control Engineering, Beijing 100190, China;*

² *Software Engineering Institute, East China Normal University, Shanghai 200062, China;*

³ *Institute of Spacecraft System Engineering, China Academy of Space Technology, Beijing 100094, China*

The trustworthiness of spacecraft software is one of the key factors to ensuring space mission's success. However, the evaluation of spacecraft software is only qualitative heretofore. The software trustworthiness measurement is the quantification of software trustworthiness and is conducive to the improvement of software quality. In this paper, an attribute-based approach which combines subjective and objective to measuring and grading software trust for spacecraft software is proposed. Based on the proposed method, 23 spacecraft software is measured. Analytical results show that the proposed metric model and measurement approach is reasonable and effective. Through the measurement, weak link in the progress of software development can be easily found, which indicates that the software trustworthiness measurement plays significant role in improving Spacecraft software's quality.

spacecraft software, software trustworthiness, measurement model, grading criteria, hierarchy attributes

doi: 10.1360/N092014-00479