软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2016,27(4):955−968 [doi: 10.13328/j.cnki.jos.005024]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

软件可信评估研究综述:标准、模型与工具*

沈国华[1], 黄志球[1], 谢 冰[2], 朱羿全[1], 廖莉莉[1], 王 飞[1], 刘银陵[1]

[1](南京航空航天大学 计算机科学与技术学院,江苏 南京 211106)

[2](北京大学 计算机科学技术系,北京 100871)

通信作者: 沈国华, E-mail: ghshen@nuaa.edu.cn

摘 要: 安全攸关软件的可信性关乎生命安全和财产保全,因此,分析评价软件可信性是否符合用户的预期(即软件可信评估)至关重要.软件可信评估从主观和客观两个方面度量软件的质量,对软件生产和应用有着重要的意义.综述了可信评估管理中涉及到的标准、模型和工具,而非关注软件度量本身.首先分析对比了软件可信性、可信评估的定义,并在研究了与可信性密切相关的软件质量的联系与区别之后,从相关国际标准、评估涉及的模型(包括质量属性模型、证据模型、分级规范等)以及软件工具支持等方面综述了软件可信评估研究工作.并且区分了这些方面中领域相关、领域无关的不同之处.目前软件可信评估已取得了一定的理论成果,并开发了若干工具辅助进行可信评估,但仍需在通用性、可伸缩性等方面有所加强.

关键词: 软件可信性;可信评估;软件质量;软件度量;安全攸关软件

中图法分类号: TP311

中文引用格式: 沈国华,黄志球,谢冰,朱羿全,廖莉莉,王飞,刘银陵.软件可信评估研究综述:标准、模型与工具.软件学报,2016,27(4):955−968. http://www.jos.org.cn/1000-9825/5024.htm

英文引用格式: Shen GH, Huang ZQ, Xie B, Zhu YQ, Liao LL, Wang F, Liu YL. Survey on software trustworthiness evaluation: Standards, models and tools. Ruan Jian Xue Bao/Journal of Software, 2016,27(4):955−968 (in Chinese). http://www.jos.org.cn/1000-9825/5024.htm

Survey on Software Trustworthiness Evaluation: Standards, Models and Tools

SHEN Guo-Hua[1], HUANG Zhi-Qiu[1], XIE Bing[2], ZHU Yi-Quan[1], LIAO Li-Li[1], WANG Fei[1], LIU Yin-Ling[1]

[1](College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

[2](Department of Computer Science and Technology, Peking University, Beijing 100871, China)

Abstract: The failure of safety-critical software could result in death, injury and damage to people or loss of equipment or property. Therefore, it is important to evaluate whether software trustworthiness fulfills the user needs (i.e., trustworthiness evaluation). This paper first compares the definition of software trustworthiness and its evaluation. Then, it surveys the software trustworthiness evaluation from three different aspects: Standards, models, and CASE tools. This work studies these aspects from the view of domain-independent as well as domain-dependent. In summary, there is great progress being made for software trustworthiness evaluation theoretically and practically while its universality and scalability are still need to be improved.

Key words: software trustworthiness; trustworthiness evaluation; software quality; software metrics; safety-critical software

* 基金项目: 国家自然科学基金(61272083); 国家高技术研究发展计划(863)(2015AA015303); 中央高校基础科研业务费专项资金(NS2015093)

Foundation item: National Natural Science Foundation of China (61272083); National High-Tech Research and Development Plan of China (863) (2015AA015303); Fundamental Research Funds for the Central Universities (NS2015093)

收稿时间: 2014-07-02; 修改时间: 2014-12-22; 采用时间: 2015-12-28; jos 在线出版时间: 2016-01-11

CNKI 网络优先出版: 2016-01-12 11:22:22, http://www.cnki.net/kcms/detail/11.2560.TP.20160112.1122.003.html

1 软件可信与可信评估

1.1 软件可信性

当前软件已普遍运用于航空航天、武器装备、交通、核能等安全攸关领域.软件的大量应用使系统性能有了很大飞跃,有效提高了整体系统的精确性、灵活性和快速反应能力.以航空为例,机载软件是一种典型的嵌入式软件,在目前的第三代和第四代飞机中,机载软件已经成为飞行控制、通信导航、火力控制以及维修保障的核心(例如,军机 F-22 飞机上嵌入式系统代码高达 300 万行,F-35 机载和地面嵌入式系统代码高达 1 500 万行[1]).

在开放、动态的环境中,软件(特别是嵌入式软件)行为的不可控性和不确定性都使得软件面临可信性问题的重大挑战.用于安全攸关领域的软件,由于各种故障、错误的发生或是受到外部危险源的侵害、人为攻击,导致出现软件失效事件,直接或间接带来不利影响,甚至造成巨大财产损失和人员伤亡.因此,这类软件被称为安全攸关软件(safety-critical software).

由于软件失效引发的事故甚至灾难不胜枚举,如:1996 年 Ariane5 型火箭首次发射中软件数据转换错误,导致发射 40s 后爆炸;2009 年,法国航空公司 AF447 航班的 A330-200 型飞机由于测速仪结冰,飞行控制软件给出了错误的攀升指示,而软件中未设置高度值上限,最终导致飞机在大西洋上坠毁[2].分析这些事故发生的原因可以发现:软件失效是由自身系统质量的缺陷或者使用质量的保障不完善等原因造成的,比如系统遇到某外部危险源(如结冰、雷击等)或故障时如何通过约束软件的行为使其不会出现不可接受的违反系统安全的行为[3],凸显了安全性、可靠性等问题.

软件是否安全稳定、能否成功运行并给用户提供预期的服务成为人们最为关心的问题,即软件可信性(software trustworthiness)问题.关于软件可信的定义至今仍没有达成共识,形成统一的规范.多年来,各国的学者、研究组织立足于不同的研究领域,从不同角度出发给出了软件可信定义的表述.本文从提出的组织、术语、可信的定义、视角、关注的可信属性等方面对主要的几种定义进行了对比,见表 1.

Table 1 Definition of software trustworthiness

表 1 软件可信的定义

提出者	术语	可信定义	视角	关注的属性
美国科学与技术委员会(NSTC)	High confidence	对系统行为符合设定期望的可预测性的一种度量[4]	从用户(主体)的角度出发,强调行为的可预测性	功能正确性、安全性、容错性、实时性和保密性等
微软公司	Trustworthy	一种可用、可靠和安全的计算,如电力系统、自来水服务和电话[5]	从用户体验的角度,强调用户对软件行为的可信赖	可靠性、安全性、保密性、业务完整性等
可信计算组织(TCG)	Trusted	一个实体在实现给定目标时,若其行为能够完全遵循期望的方式,则该实体是可信的[6]	从用户(主体)的角度出发,强调行为与设定目标的符合性	保密性等
王怀民等人	Trustworthy	如果一个软件系统的行为总是与预期的相一致,则称为可信[7]	从用户(主体)的角度出发	可靠性、可用性、安全性等
ISO/IEC15408	Security	一个可信的组件、操作或者过程的行为在任意操作条件下是行为可预测的,并能很好地抵抗应用程序软件、病毒以及一定的物理干扰造成的破坏[8]	从软件(客体)的角度出发,强调客体的可预测性和抗毒、干扰的能力	可靠性、可用性、安全性等
ISO/IEC 25010	Dependability	提供指定服务给最终用户的能力,使其能够依赖并信任系统提供的服务[9]	从用户(主体)的角度出发	可靠性、可维护性、保密性等
Avizienis 等人	Dependability	提供可信赖服务、避免频繁出现不可接受的、严重的服务故障的能力[10]	从软件(客体)的角度出发	可用性、可靠性、安全性、机密性、完整性、可维护性
林闯等人	Trustworthy	网络系统的行为及其结果是可以预期的,能够做到行为状态可监测,行为结果可评估,异常行为可控制[11]	从网络行为的角度出发	安全性、可生存性、可靠性等

从以上对比可以看出,可信定义主要从用户(主体)和软件(客体)两个不同的角度对可信进行定义.从用户角度出发的软件可信定义侧重用户主观感受,强调用户对软件行为的信任、是否符合用户的期望;从软件角度出

发的软件可信定义侧重软件的客观能力,强调软件本身应该具备哪些能力才能获得用户的信任.前者表达的是主观“可信”,后者表达的是客观“可信性”.即“可信”是用户对软件客观质量的主观认同,而“可信性”是软件客观具有的质量[12],是软件的行为符合人们期望的能力,它是软件系统的可用性、可靠性、安全性、正确性、完整性等诸多属性的综合反映.综合以上文献,具备了用户期望的可信性的软件对用户而言才是可信的,用户觉得“可信”的软件一定程度上具备了满足用户预期的“可信性”,二者有一些差别并能通过可信评估联系起来.

与可信性密切相关的术语还有很多,如安全性(safety)、保密性(security)、可靠性(reliability)、可生存性(survivability)、完整性(integrity)等.总体上可信性覆盖以上这些术语的内涵[13],相互关系如图 1 所示.

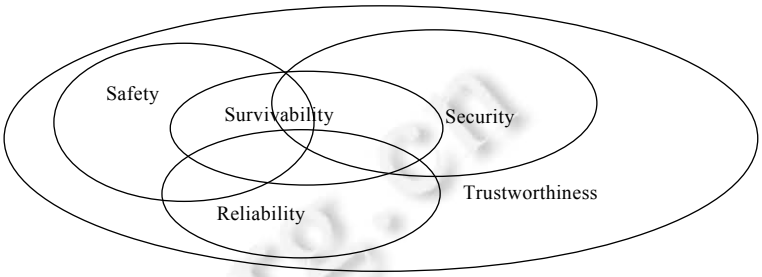


Fig.1 Software trustworthiness and related terms

图 1 软件可信性及相关术语

随着可信性问题的日益凸显,软件可信性受到了高度关注,围绕“软件可信性”形成了新的研究热点.美国计算研究协会(Computing Research Association,简称 CRA)和美国国防部高级研究计划署(Defense Advanced Research Projects Agency,简称 DARPA)都将高可信软件系统视为目前计算机研究领域必须应对的五大挑战之一.很多国家的政府组织、科研机构等意识到软件可信研究的重大价值和美好前景,纷纷提出了各自的研究计划.例如:国际组织 TCPA(Trusted Computing Platform Alliance)和 TCG(Trusted Computing Group)制定了关于可信计算平台、可信存储、可信网络连接、可信计算框架等一系列技术规范,致力于促进新一代具有安全、信任能力的硬件计算平台的发展;美国国家科学基金会(National Science Foundation,简称 NSF)制定了一个可信计算计划,重点关注如何管理网络世界的安全和隐私;欧洲于 2006 年启动了由 23 个科研机构和工业组织参与的“开放式可信计算(Open Trusted Computing)”研究计划.我国也十分重视软件系统的可信性问题,国家自然科学基金委员会于 2007 年开始批准立项了“可信软件基础研究”重大研究计划[14];国家高技术研究发展计划(863)中设立了专门的项目来研究高可信软件生产工具和集成环境[15];国家重点基础研究发展计划(973)将可信软件的研究置于重要地位,研究基于网络的复杂软件的可信度和服务质量[15],并提出将“安全攸关软件系统的共性理论和构造方法”作为 2014 年度重要支持方向.安全攸关软件的可信性是软件与硬件、系统、物理世界、人等深度融合过程中必须面临的一个挑战.

1.2 软件可信评估及其标准

可信软件的研究范畴非常广泛,主要包括以下 4 个方面[14]:(1) 软件可信性度量与建模;(2) 可信软件的构造与验证;(3) 可信软件的演化与控制;(4) 可信环境的构造与评估.近年来,软件可信性受到了持续关注,围绕“软件可信性”形成了新的研究热点.人们针对软件生命周期的各个阶段,研究提高和保障软件可信性的理论和方法.判定一个软件资源是否可信、度量软件资源的可信程度是软件可信性研究中的一个基本问题,也是本文关注的重点.

在软件开发和应用中,为评判软件资源的安全性和可靠性等是否符合用户的预期,需要对软件资源的可信性进行分析与评价,即可信评估(trustworthiness evaluation)[16].量化是一种工程科学成熟的重要标志,可信评估给出软件可信性一个量化的评估结果,不仅有利于软件使用者对软件制品进行选择,而且有利于软件研发者开发出高可信的软件.可信评估的研究为软件制品的可信评估标准的制定以及相关可信评估工具的开发奠定了理论基础.目前,可信评估仍是软件可信研究中的一个发展方向,相关的理论和方法还处在研究阶段[2],目前尚未有成熟的评估方法可以评估出多数人认可的软件可信标度[17].

关于软件质量及软件评估的一个重要国际标准是 ISO/IEC 9126,于 1991 年颁布,并于 2001 年在原标准的基础上修改,形成了两个相关新标准 ISO/IEC 9126:2001(software engineering-product quality)和 ISO/IEC

14598(software engineering-software product evaluation).ISO/IEC 9126:2001 分成 4 个部分:ISO/IEC 9126-1、ISO/IEC TR 9126-2、ISO/IEC TR 9126-3、ISO/IEC TR 9126-4,分别描述质量模型(quality model)、外部度量(external metrics)、内部度量(internal metrics)和使用质量度量(quality in use metrics).之后,ISO/IEC 9126:2001 和 ISO/IEC 14598 被 ISO/IEC25000(ISO/IEC software engineering-software product quality requirements and evaluation,简称 SQuaRE)标准替代,SQuaRE 是一个系列标准,包括以下部分:质量管理(ISO/IEC 2500n)、质量模型(ISO/IEC 2501n)、质量测量(ISO/IEC 2502n)、质量需求(ISO/IEC 2503n)、质量评估(ISO/IEC 2504n)以及待扩展部分.其中,质量模型定义了 1 个软件产品质量模型和 1 个系统使用质量模型.质量测量中 ISO/IEC 25023 定义了外部属性测量、ISO/IEC 25022 定义了内部属性测量、ISO/IEC 25024 使用质量的测量,分别替代了 ISO/IEC 9126-2、ISO/IEC 9126-3 和 ISO/IEC 9126-4,而 ISO/IEC 2504n 则对应 ISO/IEC 14598.ISO/IEC 9126 及其相关标准的演化关系如图 2 所示.

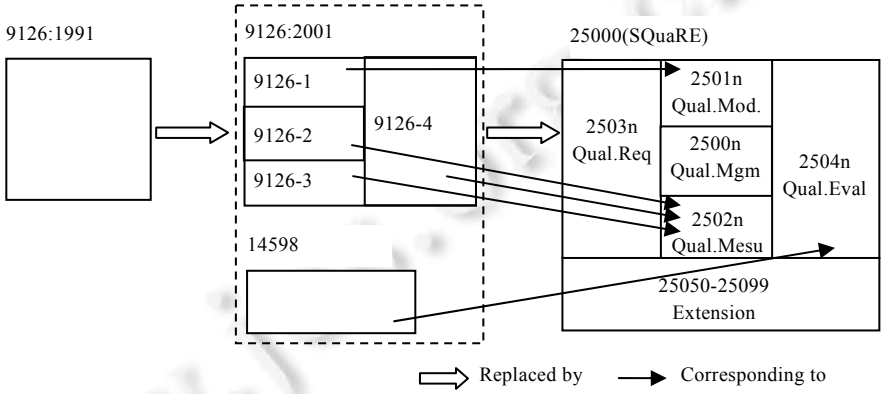


Fig.2 International standards for software quality and evaluation

图 2 软件质量及软件评估的国际标准

然而,目前对可信评估的定义国内外尚未有一致的定论,且视角不同带来了方法上的大相径庭.学者们一般从评估方法的角度对可信评估进行定义,例如,文献[2]提出采用分级模型对软件可信性进行评估,认为可信评估就是依据可信等级定义,根据所获得的可信证据和可信评估指标体系确定软件的可信等级;软件可信分级规范[18]将可信评估定义为依据特定的已成文的软件可信评估准则,确定特定的软件产品是否达到某一特定可信等级的活动;文献[19]认为,“评估可信性对命题 P 为真的支持程度”就是可信评估.本文基于 ISO/IEC25000,并综合以上各文献,定义“可信评估”是针对特定的软件制品 Target(被评估对象),面向可信质量模型 Quality model(评估目标),由评估人员 Evaluator(评估人)采集相关的可信证据 Evidence(评估依据),应用可信分级规范 Classification(评估准则)对软件可信性进行测量,并得到量化评估结果 Result(评估结果)的过程,如图 3 所示,其中,与可信评估直接关联的元素用灰色表示,是本文主要分析的内容(因评估对象和评估人这两个元素易于理解,则不详细展开介绍).即可信评估可表示如下:

Result=eval(Target,Evaluator,QualityModel,Evidence,Classification).

软件可信评估对软件的生产和应用都有着重要的意义.

从可信软件生产的需求出发,对软件开发过程中集成的构件、服务和设计模型等软件资源的可信度进行评估,有利于软件研发者开发出高可信的软件.因为可信性来源于对过程、产品或资源与软件目标之间符合程度的把握,而只有对软件的目标有了清晰的定义、对软件的现状有了实际的调查,才有可能找出需要改进的地方.此外,软件产品的可信评估结果也可以提供给客户,作为其了解软件产品可信性的有力依据.

从可信软件应用的需求出发,软件可信度的评估结果有利于保障软件使用者选择可信的软件制品.随着 Internet 技术的普及,用户可以自主地选择构件、服务等资源进行复用,然而软件资源质量难以预测和控制,使得软件复用者常常下载或集成不符合需求的软件.根据不同领域的应用需求对软件可信性进行评估的结果为软

件复用者的选择提供了依据.

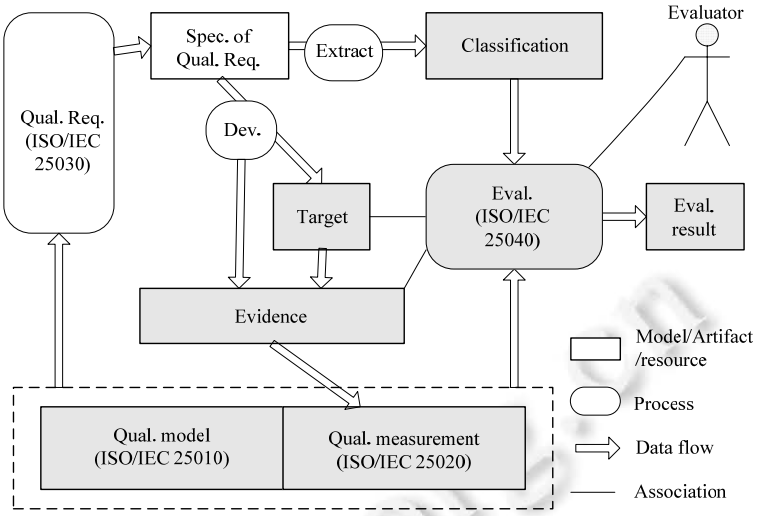


Fig.3 Definition of software trustworthiness evaluation

图 3 软件可信评估的定义

2 软件可信评估与传统软件测量

2.1 软件质量及其测量

根据国际标准化组织(International Organization for Standardization,简称 ISO)的定义,软件质量(software quality)是软件产品满足规定和隐含需求能力有关的所有特征和所有特性的总和.软件测量(software measurement)是依照清晰定义的规则将这些特性映射为数字或符号的整个过程.该标准提出从内部质量(internal quality)、外部质量(external quality)、使用质量(quality in use)和过程质量(process quality)这 4 个角度来理解和保障软件质量的属性,如图 4 所示.

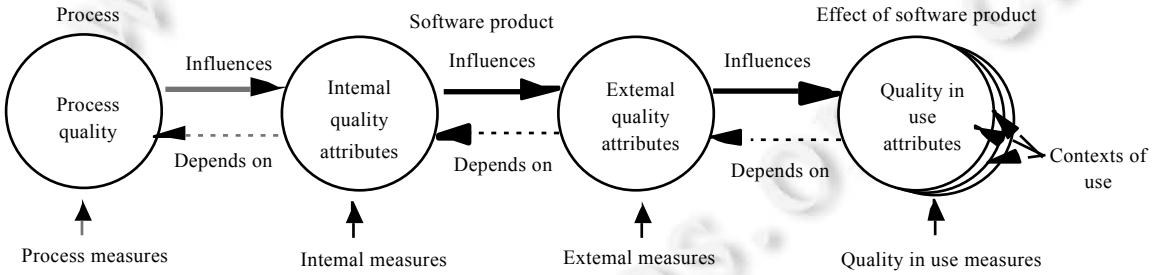
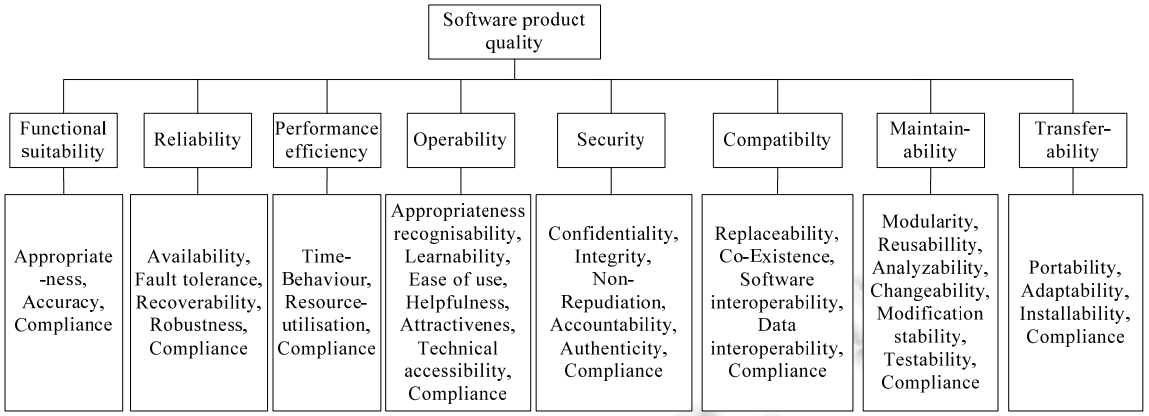


Fig.4 ISO software quality model

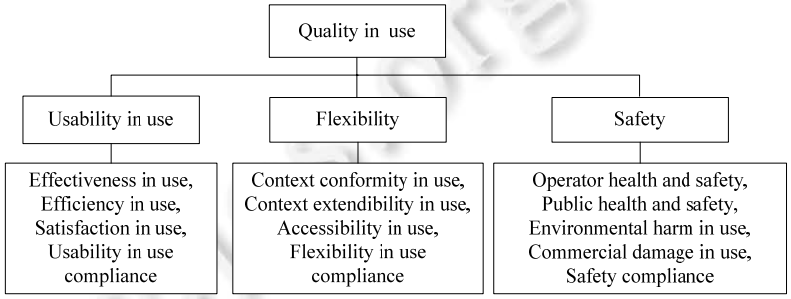
图 4 ISO 软件质量模型

过程质量的测量(measurement)和评估通常作用于软件开发过程中,目的在于预测过程的未来性能,减少过程结果的偏差,对软件过程的行为进行目标管理,帮助发现软件开发过程中的瓶颈或问题所在.内部质量的测量和评估通常作用于开发阶段的非执行的软件产品(比如设计模型、数据库表、程序代码等),通过对中间制品的质量测量预估最终产品的质量.外部质量的测量和评估通常在测试阶段和运行阶段使用,通过对软件行为能力的评价实现对软件产品质量的评估.使用质量的测量和评估通常在软件应用阶段进行,当软件用于指定的使用环境和条件时,从用户体验的角度实现对产品质量的评估.内部质量和外部质量统称为系统质量,通过功能性(functionality)、可靠性(reliability)、易用性(usability)、效率(efficiency)、可维护性(maintainability)、可移植性(portability)等若干特性及其子特性进行表达,如图 5(a)所示.使用质量一般在真实系统环境下获得,通过有效性

(effectiveness)、生产率(productivity)、安全性(safety)、满意度(satisfaction)等特性来表达,如图 5(b)所示.



(a) ISO 软件质量模型(内部/外部质量)



(b) ISO 软件质量模型(使用质量)

Fig.5 ISO software quality model

图 5 ISO 软件质量模型

2.2 软件度量与软件可信评估

文献[15]认为,软件可信性是软件质量的一种特殊表现形式,是传统软件质量概念在互联网时代的延伸.因此,可信评估本质上也是软件质量的测量,但是它与传统的软件质量测量又有所不同.

首先,传统的质量测量偏重于软件系统本身,遵循“以系统质量为核心、兼顾使用质量”的思路进行质量保障.可信评估过程中将重心逐步转移到了使用质量上,遵循“以使用质量为核心,系统质量为基础”的思路关注使用层面的综合化质量属性及其保障.随着商用成品(commercial off-the-shelf,简称 COTS)、网构软件、服务计算等新型软件形态的发展与普及,软件的运行环境从封闭、静止、稳定转向开放、动态、不确定,软件在运行时可能会受到木马、病毒、窃听等外界的恶意攻击,传统的仅考虑自身系统质量的质量测量已难以适用,需要考虑软件实际运行时的使用质量.

其次,传统的质量测量通常针对具体的质量属性,如正确性、容错性、易安装性等,较少考虑不同质量属性的综合.而可信性是软件系统的可用性、可靠性、安全性、正确性、可预测性等诸多属性在使用层面的综合反映,因此可信评估关注的是不同质量属性的综合.

再次,传统软件质量测量的客观性较高,而可信评估则是主观与客观的结合[15].可信本身是一个复杂的概念,不同的研究学者对其有不同的认识.软件可信性既有客观的因素,又有主观的成分[20],不同应用领域(甚至同一领域)、不同任务需求、不同人员在不同时间段对可信性的定义和标准都可能不一样.而且使用质量的度量依赖于进行测量的环境,随着评估人的不同而发生变化,相对而言,可信评估在较大程度上涉及到用户个性化的体验和评价.

基于以上几点,可信评估与传统质量测量在关注的质量重心、考虑的质量属性以及评估的主客观方式上的

不同,使得传统的软件质量测量和保障技术难以满足可信评估的需求,需要建立新型的质量保障体系.因此, 本文重点关注可信评估管理中涉及到的标准、模型和工具,而非关注软件度量本身.例如,本文不太关注通过某种特定方法(如测试或验证)获得了一些程序质量度量,但关注这些度量指标与可信属性的支持关系,它们与何种证据有内在的追踪关系,评价指标的标准如何定义,以何种过程来进行管理,是否有工具支持.

传统软件质量测量分为产品、过程、资源等方面,同理,可信评估也可从这些视角展开.例如,面向软件产品的可信评估可采用 McCall 模型、ISO 的内部/外部质量模型.此外,文献[21,22]等也提出了相应模型;面向软件过程的可信评估,可采用 ISO 的过程质量模型以及文献[23,24]等提出的相关模型.本质上,不同视角的可信评估能采用质量模型(见第 3.1 节)统一来加以描述,即根据不同的需求定义所需的质量模型,因此,本文不再根据这种分类视角分别加以阐述.

软件从传统形态发展到 COTS、构件、服务等不同形态,新的形态普遍呈现出相对开放、商用化、标准化等新特点.尽管如此,可信性评估的内涵仍保持相对恒定,而且其外延的新变化也有类似之处.由此,对它的评估主要关注新特点的部分.文献[25]综述了从 1995 年以来对于 COTS 的选择方法,其中包括对它的评估.COTS 的评估涉及 7 个开放性研究问题(R1~R7),诸如(问题 R1/R2)如何让评估适用不同领域和项目;(问题 R3/R4)采用何种判定规则能作更有效的选择;(问题 R5)提供协商构件用于解决冲突;(问题 R6)设计知识库用于评估;(问题 R7)如何基于 COTS 供应商的信誉度调整评估值.由于相似性,对服务等的评估也可参照使用这些研究问题.

2.3 可信评估面临的挑战

如上所述,软件可信评估是可信性研究中的一个基本问题,在软件可信性保障活动中占据重要的地位.虽然,目前国内外采用各种软件分析、验证方法进行软件质量保证的研究众多;然而,专注于可信评估的研究相对有限,尚面临着一系列的问题有待解决:如可信评估模型的结构和语义、可信评估的工具支持等.

(1) 可信评估模型的结构和语义问题:对软件资源进行可信评估首先要明确评估的需求,其次,根据需求选择或定制评估模型,模型中对评估的内容、度量公式、决策标准等做出明确的规划,然后评估人员遵循评估模型收集相应的证据实例,按照模型中定义的算法进行量化计算,得到度量的结果.目前可信评估方法的研究大多从数学建模的角度探讨评估的算法,对于评估过程中涉及的概念、评估模型的结构却没有一致的定义和明确的说明.我们分析发现,在不同的方法描述中普遍存在着一词多义、异词同义的现象.由于缺乏明确的结构和语义,可信评估执行人员在实践中很难有效地建立评估模型,从而造成对可信评估需求描述的不精确性.

(2) 可信评估的工具支持:可信评估是一项复杂的活动,仅依靠人工手段,没有任何工具支持进行评估必然是低效的,既容易出现错误也不利于评估过程的管理和复用.然而,目前还没有出现成熟、得到广泛应用的可信评估管理工具.由于可信评估对象的形态(如构件、服务、工具)多样,且涉及领域(如商业、军事)各自不同,这给工具的实现带来难度.针对于此,出现了若干针对较一般软件形态、非特定领域的可信评估的研究及其工具实现;同时也出现了一批适用于某种软件形态、特定领域的研究及工具.面向前一种情况,北京航空航天大学在国家 863 计划的支持下设计、实现了“软件结构及代码的审查和综合评估工具”[26];北京大学在软件资源库中建立了可信评估的原型系统[16];南京航空航天大学在国家 863 计划的支持下设计并实现了“软件可信评估管理工具”[27],这些工具一般具有较普遍的适用性.面向后一种情况,文献[28]给出了一个支持仿真系统可信评估的工具 HIT-CET;文献[29−31]开发了适用于 BPMS 软件可信评估的管理工具.这些工具往往支持相对特定的领域或形态,对于目前各类不同应用环境下的软件制品的支持度还有所欠缺.

(3) 新的软件形态导致的适用性调整问题:如第 2.2 节所述,新的 COTS 和服务等软件形态使得传统评估需要调整以适用于新的可信评估.以 COTS 评估[25]涉及的问题为例,不同领域和项目适用性问题(R1/R2)本质上需要借助(元)模型的通用性和针对目标领域定制实现;判定规则(R3/R4)和信誉度选择(R7)分别对应评估模型中的度量公式、决策标准;基于知识库的评估(R6)本质上需要引入一些领域知识(如构件与服务的行业分类法),对应评估模型的语义问题,因此这些问题可归结为挑战(1),即可信评估的模型.而解决冲突的协商机制(R5)则采用一个过程模型解决,可归结到挑战(2),即利用可信评估工具解决.

因此,下文主要从可信评估的相关模型(包括质量模型、证据模型和可信分级规范等)与工具支持等方面进

行分析.

3 软件可信评估的模型与工具

3.1 可信评估的质量模型与证据模型

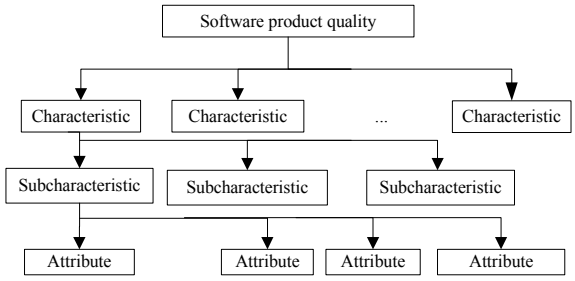


Fig.6 Structure used for the quality model (quality meta-model)

图 6 软件质量模型结构(软件质量元模型)

软件可信涉及的模型主要包括质量模型和证据模型.质量模型关注具备哪些特性的软件是可信的,即软件可信性与哪些因素相关,从哪些角度来保障软件可信性.而证据模型关注用于支撑可信属性的素材的来源、收集和分类方式.下面分别进行描述.

(1) 软件可信的质量模型

总体上看,软件质量模型分为两个层次:质量元模型和质量模型.ISO/IEC 2501n 定义了质量模型结构(structure used for the quality model)以及软件质量模型(software quality model).质量模型结构即为软件质量元模型,总体上将质量模型划分为若干特性(characteristics),特性可进一步细化为子特性(sub-characteristics)和质量属性(attributes),特性与子特性是属性的类别划分,如图 6 所示.

《软件可信分级规范》[18]依据目前几种典型软件质量模型(如 ISO/IEC 2501n,McCall 模型),定义软件可信属性为“软件按用户的期望提供正确、安全、可靠等特性服务的能力”,涵盖功能性、可靠性、易用性等软件质量特性,还包括安全性、可生存性、实时性等其他软件特性.由此可见,该规范中的“可信属性”可以指代 ISO/IEC 2501n 质量模型中的特性、子特性和属性,即通过可信属性可不断迭代细分为下一级的可信子属性.这种质量模型不但保留了 ISO 质量模型结构的本质含义,而且更简单灵活,易于扩展.

软件质量模型是质量模型结构(即质量元模型)的一个实例,如 ISO 定义了一个软件产品质量模型(software product quality model)和一个使用质量模型(model for quality in use),分别如图 5(a)和图 5(b)所示.软件产品质量模型包括若干个特性,可进一步划分子特性;系统使用质量模型亦包括若干特性,也可进一步划分子特性.

文献[32]提出了一种对软件体系结构的可信属性模型,该模型利用最大熵原理(POME)和灰色决策方法(GDMM)对体系结构进行可信评估;文献[33]综合考虑了软件在身份、能力和行为等方面的可信属性,提出了一个 Internet 环境下软件的可信概念模型(concept model for trustworthiness)及可信保障框架(trustworthiness assurance framework),这里的可信概念模型即为可信评估的质量模型;文献[7]给出了一种面向互联网虚拟计算环境(internet-based virtual computing environment,简称 iVCE)的互联网软件可信概念模型,从身份可信、能力可信和行为可信这 3 个方面来保障网络软件的可信性;文献[34]分析了 13 个不同的可信模型(trust model,即为可信评估的质量模型),提出用上下文环境、第三方信息、可用性、自信力、行为等因素诠释软件可信,并给出软件可信概念模型的本体描述.文献[27]依据软件可信分级规范,定义了可信属性元模型,其中包含属性、属性间分解关系,能够依据评估目标定制可信属性模型,具有通用性.

(2) 可信证据模型

与软件相关的能够反映其某种可信属性的数据、文档或其他信息,称为软件可信证据[18].传统的软件度量主要从软件客体的视角出发,关注于客观的软件属性,所以较少提及证据管理;而可信评估加入了软件用户对开发过程、使用质量的关注,因此需要搜集并管理证据.

一个软件所有可信证据的集合以某种结构进行组织后,就构成了软件可信证据模型.证据的组织结构因分类视角的不同而差异很大.例如,软件可信分级规范从软件生命周期的角度将可信证据划分为以下 3 个方面:开发阶段证据、提交阶段证据和应用阶段证据[18],每个阶段的证据可进一步细分为下一级的证据.此外,还可以根据证据的来源、存在的方式进行证据的组织.文献[27]依据软件可信分级规范,并经过适当抽象,定义了可信证据

元模型,其中包括证据类、证据两种基本元素,证据类可迭代划分为子证据类,最终证据类可包含证据.可信证据模型是多层次树状结构,证据类是其中非叶子节点,表示划分依据,证据是叶子节点,不可再分.基于可信证据元模型,能够依据评估实际要求定制可信证据模型,具有灵活性.

文献[20]分析了传统的可信证据收集与分类方法的不足,提出了一种基于验证的可信证据模型和可信评估方法.该模型由人从不同角度对软件进行思考形成,反映了人对软件认识的深入程度,与传统证据模型相比,该模型便于多维评估,有利于软件的可信演化.文献[35]综合考虑了软件环境、使用体验和生产过程中的可信要素,从声誉可信、交互可信、机理可信这 3 个剖面对可信证据进行管理,并给出了模型的应用方法.文献[36]通过对开源社区 SoureForge 中的项目进行分析,发现项目角色的配置与软件可信性存在紧密的关联,据此提出了将项目角色配置视为软件可信证据的想法.

3.2 可信评估的分级规范

这部分主要阐述可信的分级规范.分级规范是可信评估的准则,包括可信分级的定义,每个分级应满足的决策标准等.下面从领域无关和领域相关两个方面进行分析.

(1) 领域无关的可信评估分级规范

在可信评估分级规范中,需要定义软件可信分级(例如从第 0 级~第 5 级)和软件可信分级指标体系.可信分级指标体系是可信证据及其指标度量值的集合,由证据对软件可信级别的支持关系组成[18].文献[15]采用了度量和提高软件过程可信性的方法来保障软件可信,构建了可信过程管理框架 TPMF(trustworthy process management framework);文献[2]描述了一种采用软件分级模型对软件的可信性进行评估的方法,给出了一种软件可信分级模型,为软件可信评估机制的建立提供了一种理论和方法;文献[16]提出了一个支持软件资源可信评估的框架,并给出了该框架在北京大学软件资源库中设计的实现方案;文献[35]提出了需求导向、多目标融合的可信测量模型(trusted measurement model,简称 TMM),与可信计算组织(TCG)采用的完整性度量机制相比,TMM 灵活、易扩展,更贴合实际应用的需求.文献[17]针对可信评估中评估需求的动态性、多变性、专家决策的主观性等问题,提出了基于效用和证据理论的评估方法,该方法支持动态构建评估指标树,一定程度上满足了评估需求动态、多变的需要.

文献[27]提出了一种通用的可信评估元模型:包括质量模型、证据模型和评估指标体系.其中,评估指标体系是对可信评估分级规范中的分级指标体系进行了细化定义,以便于评估工具的实现.分级指标体系定义为指标(indicators)的集合,而指标通过定义分析模型(analysis model)和标度(scale)来描述,其中,分析模型又包括公式(formula)和决策标准(decision criteria).通过指标将原本相对独立的属性模型和证据模型关联起来.可信评估元模型支持依据不同领域、不同目标定制不同的质量模型、证据模型和指标体系,从而满足了可信评估需求动态、多变的需要.文献[37]提出了一套评估指标集,并且使用模糊综合评估模型对软件可信属性进行评估.

关于可信度量值的分析模型与标度,文献[38]提出了一种模型,将可信度量值定义为各子属性度量值的函数,形如 $T=f(y_1,y_2,\ldots,y_n)$,该函数满足单调性、增长性、敏感性和替代性这 4 种性质.此外,将子属性分为关键与非关键两类,并分别赋予权重.文献[39]在此基础上,针对最小度量值的关键子属性的影响因子作了改进,使其更符合实际情况.文献[40]则增加了第 5 种性质:稳定性,使得模型的使用效果有所提高.这组研究成果提供了较通用的由子属性度量值计算父属性度量值的计算模型,并设定标度范围为[0,10],对属性度量具有明确的指导意义.

(2) 特定领域软件的可信评估分级规范

在实际工作中实施的可信评估往往需要面对某个特定领域,通过分析针对特定领域内软件的可信需求,详细设计评估过程中每一阶段的任务,包括评估的内容、度量的机制、决策标准等.

例如,面向业务流程管理系统(business process management system,简称 BPMS)的软件可信评估:文献[29,30]给出了 BPMS 软件的可信指标体系和分级规范,建立了评估过程模型,设计、实现了基于上述模型的可信评估管理工具.文献[41]制定了 20 个功能评估标准和 10 个模块评估标准,并将 BPM 产品划分为支持 BPM、BPM 引擎、通用型 BPM 产品、专业级 BPM 应用这 4 个等级.文献[42]从概念、方法和工具支持这 3 个维度实现了一个 BPM 管理和评估框架.文献[31]从耦合度和内聚度这两个方向上对 BPM 软件的流程设计进行评估,

给出了具体的实现方案和评估工具 CoCoFlow.

面向商用成品 COTS 的选择:文献[43]提出了一种基于产品领域的间接选取方法,并在一个在线贸易系统中加以应用,结果表明,该方法比传统的直接选取方法更有效.文献[44]提出了一种软件功能需求驱动的评估和选择 COTS 构件的方法,首先将功能需求分解到每一个模块,再基于每一个模块识别出候选的构件,根据给定的评估模板对构件的可用性、易用性进行评估,最后求解构件给定成本约束下的最优组合.文献[15]提出一种基于差异分析的构件评估方法,通过分析构件与需求在功能性上的差异度选择构件.

对于开源软件的评估有若干主要的评估模型和方法,如 OSMM[45](open source maturity model)、QSOS[46](qualification and selection of open source software)、OpenBRR[47](open business readiness rating)等.这些评估模型和方法总体思路比较类似,首先定义质量属性模型(包含若干属性/子属性),其次收集原始评估数据并获得度量,依据一个计算模型(如乘权求和)由度量计算出属性.这里,以 OpenBRR 为例作一说明:OpenBRR 是由 Carnegie Mellon Silicon Valley Center 与 Intel 等公司联合发起的,采用了 12 个开源软件的属性分类(categories),如功能、质量、性能、支持度、社区、文档等,通过采集规范化的度量,计算获得每个分类,最终将所有分类乘权求和得到最终的 BRR 分级.总体上,这些评估模型和方法兼顾了通用性和开源软件的部分特点(如支持度、社区).但由于其属性模型扁平化,且计算模型相对简单,标度多采用离散的分级(如 1~5 级),使其对于其他较为复杂的软件和领域的适用性有所不足.

文献[48]通过可信证据的成熟度计算可信属性的可满足性,并根据可满足性裁剪软件可信属性,然后,基于该可信属性对 CPS 系统进行建模与评估.文献[49]研究了作战仿真系统的可信性,结合实际建立了一套可信性评估指标体系,提出了专家权重定量计算方法.文献[50]在综合分析电力系统的特点和可信需求后,从系统构建、实施过程管理、可信评估和可信证明这 4 个方面对如何构建可信的电力系统进行了研究.文献[51]针对作战飞机的任务效能评估问题,提出了自己的解决方案.文献[52]研究 Web 服务的可信性问题,从可用性、可靠性和安全性这3个方面对Web服务的可信性进行评估,给出了具体评估标准和算法.文献[53]针对开源构件可信性进行了评估.文献[54]在分布式数据库服务器系统 DDSS(distributed database server system)中引入可信机制,通过建立多层次信任链结构,改进存取控制方式,加强客户端角色管理、认证机制等技术,一定程度上提高了 DDSS 系统的可信性.文献[55]提出了适用于无线传感网络的可信评估模型.文献[56]给出了评估工作流型软件产品可信性的关键指标.文献[57]引入相关度的概念来评估面向服务的工作流的性能.文献[58]提出了使用形式化验证技术来验证时序安全属性,以提高医疗设备软件的可信性.

可见,针对特定领域的可信评估呈现出明显多样性的特征.即添加了特定领域的知识,并采纳了定制的度量和决策标准等,使之更专有、更具体,但不易迁移使用.

3.3 可信评估工具

软件可信评估工具对于可信评估的实施至关重要,然而,由于可信评估面临的应用领域众多,关注的质量属性侧重点不同,评价的标准变化,证据收集困难等,使得工具的实现具有一定的难度.

文献[40]设计并实现了一个基于 Java 的软件可信评估工具 SPATRUME,采用针对文献[38]改进的属性度量值计算模型,具有通用性,可能的不足之处在于缺乏证据信息以及属性度量值与证据的关联关系.北京航空航天大学设计并实现了“可信软件结构及代码的审查和综合评估及支持工具”[26],具有比较完整的软件质量(在代码及结构方面)分析与度量功能以及初始的可信评估能力,但并未深入关注评估.北京大学开发了软件资源库[16],可针对软件系统、构件、服务进行发布、检索,并管理可信证据,另外还包括可信分级的描述以及软件可信评估功能.软件可信评估过程包括发布者提供证据信息,交由专家人工评定可信级别.软件资源库系统可以辅助软件可信评估,但并不提供评估定制功能,且专家评定具有一定的主观性,对可信分级与证据间的关联关系表达有一定的欠缺.南京航空航天大学设计并实现了“软件可信评估管理工具”[27],该工具提供了可信评估元建模能力,因此是领域无关的,可提供完整的评估定制能力,并可根据领域特性定制属性模型和证据模型,并依据分级体系自动生成评估结果.另外,该工具还实现了与北京大学资源库系统的连接,实现了两个系统中资源信息、证据包的数据交换.上述工具主要适用于通用领域或应用场景,具有较为普遍的通用性,并在一定程度上具备经过定制

用于特定领域的能力.

文献[28]在提出可信度层次评估过程模型的基础上,设计实现了一个针对仿真系统的可信评估工具HIT-CET.文献[29−31]均针对业务流程管理系统(BPMS),实现了软件可信评估的管理工具.其中,文献[29]提出了一个完整的领域构件可信评估体系,并在此基础上实现了领域构件可信评估系统的开发;文献[30]提出了可信指标体系、评估过程模型及算法模型,并设计实现了适用于 BPMS 软件可信评估的管理工具;文献[31]提出了一种在管理设置阶段对过程设计的创建和评价提供指导的启发式方法.上述工具主要适用于特定的领域或应用场景,具有较强的针对性.

4 结束语

安全攸关软件的可信性关乎生命安全和财产保全,因此,相应的软件可信评估至关重要.软件可信评估从主观和客观两个方面度量软件的质量,对软件生产和应用有着重要的意义.本文重点关注软件可信评估的管理,综合分析、对比了目前可信评估的研究现状,从相关标准、评估涉及的模型(包括质量属性模型、证据模型、分级规范等)以及软件工具支持等方面综述了软件可信评估研究工作.

总体上看,软件可信评估取得了大量研究成果,在理论(模型和框架)上分两个层面展开,第 1 层是具有通用性、领域无关的模型/元模型,第 2 层是在前一层的基础上,对通用模型/元模型进行定制或实例化,以满足不同领域、不同软件形态、不同使用环境的实际需求.此外,针对理论成果开发出相应的工具,用于辅助实际的软件可信评估过程管理,为研制出高可信软件提供技术保障和支持.

未来主要关注的发展方向包括:如何将可信评估向更专有的领域、更特定的应用去发展,并适用于不同软件形态和不同的运行环境;如何表示可信证据/可信属性在软件工程的各阶段(从需求到设计、编码、测试、维护)之间的追踪关系,以及它们对于可信标准的依从性.

致谢 在此,我们向对本文工作给予支持和提出宝贵建议的同行表示感谢.尤其要感谢在国家 863 计划“高可信软件生产工具及集成环境”专题中有课题合作的北京大学计算机科学技术系赵俊峰、张伟副教授、复旦大学计算机科学技术学院的赵文耘教授、彭鑫副教授.感谢南京航空航天大学计算机学院钱巨副教授、刘春勇、洪宏、彭焕峰等同志对可信评估工具开发付出的努力.另外,还要感谢神州数码信息系统有限公司的徐拥军执行总监,他们对于评估工具的应用案例提出了很多宝贵的意见.

References:

[1] Athalye P, Maksimovic D, Erickson R. High-Performance front-end converter for avionics applications. IEEE Trans. on Aerospace and Electronic Systems, 2003,39(2):462−470. [doi: 10.1109/TAES.2003.1207258]

[2] Final Report on the accident flight AF 447 Rio de Janeiro-Paris. BEA, 2012. http://www.bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf

[3] Wu WH, Kelly T. Safety tactics for software architecture design. In: Proc. of the 28th Annual Int’l Computer Software and Applications Conf. 2004. [doi: 10.1109/CMPSAC.2004.1342860]

[4] NSTC. Research challenges in high confidence systems. In: Proc. of the Committee on Computing, Information and Communications Workshop. 1997. http://www.hpcc.gov/pubs/hcs-Aug97/intro.html

[5] Gates B. Trustworthy computing. 2002. http://www.wired.com/2002/01/bill-gates-trustworthy-computing/

[6] TCG. Specification architecture overview specification. Revision 1.4.2nd, 2007.

[7] Wang HM, Tang YB, Yin G, Li L. Credible mechanism of Internet software. Science in China-Series E: Information Sciences, 2006,36(10):1156−1169 (in Chinese with English abstract).

[8] ISO/IEC 15408-1:2009. Information technology-security techniques-evaluation criteria for IT security. Part1: Introduction and General Model, 2009.

[9] ISO/IEC 25010:2011: Systems and software engineering—Systems and software quality requirements and evaluation (SQuaRE)—System and software quality models. 2011.

[10] Avizienis A, Laprie JC, Randell B, Landwehr C. Basic concepts and taxonomy of dependable and secure computing. IEEE Trans. on Dependable and Secure Computing, 2004,1(1):11−33. [doi: 10.1109/TDSC.2004.2]

[11] Lin C, Peng XH. Research on trustworthy networks. Chinese Journal of Computers, 2005,28(5):751−758 (in Chinese with English abstract).

[12] Wang HM, Yin G. Evolution of software trustworthiness in the network age. Communication of China Computer Federation, 2010,6(2):28−34 (in Chinese with English abstract).

[13] Fan XG, Chu WK, Zhang FM. Surveys of software safety. Computer Science, 2011,38(5):8−13 (in Chinese with English abstract).

[14] Liu K, Shan ZG, Wang J, He JF, Zhang ZT, Qin YW. Overview on major research plan of trustworthy software. Bulletin of National Natural Science Foundation of China, 2008,22(3):145−151 (in Chinese with English abstract).

[15] Mei H. Software credibility: Internet brings challenges. China Computer Federation, 2010,6(2):20−27 (in Chinese with English abstract).

[16] Cai SB, Zou YZ, Shao LS, Xie B, Shao WZ. Framework supporting software assets evaluation on trustworthiness. Ruan Jian Xue Bao/Journal of Software, 2010,21(2):359−372 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3786.htm [doi: 10.3724/SP.J.1001.2010.03786]

[17] Yang SL, Ding S, Chu W. Trustworthy software evaluation using utility based evidence theory. Journal of Computer Research and Development, 2009,46(7):1152−1159 (in Chinese with English abstract).

[18] Liu XD, Lang B, Xie B, Wang HM. Software trustworthiness classification specification (TRUSTIE-STC V 2.0). In: High Confidence Software Production Tools and Integrated Environment Technical Documentation. 2009. http://www.doc88.com/p-3008711993507.html

[19] Lu G, Wang HM, Mao XG. A cognitive-based evidence model for software trustworthiness evalution. Journal of Nanjing University (Natural Sciences), 2010,46(4):456−463 (in Chinese with English abstract).

[20] Ding XL, Wang HM, Wang YY, Lu G. Verification oriented trustworthiness evidence and trustworthiness evaluation of software. Journal of Frontiers of Computer Science and Technology, 2010,4(1):46−53 (in Chinese with English abstract).

[21] Voas J. Trusted software's holy Grail. Software Quality Journal, 2003,11(1):9−17. [doi: 10.1023/A:1023679926998]

[22] Ding S, Yang SL. Research on evaluation index system of trusted software. In: Proc. of the 4th Int'l Conf. on WiCOM. 2008. 1−4. [doi: 10.1109/WiCom.2008.1869]

[23] Amoroso E, Taylor C, Watson J, Weiss J. A process-oriented methodology for assessing and improving software trustworthiness. In: Proc. of the 2nd ACM Conf. on Computer and Communications Security. New York: ACM, 1994. 39−50. [doi: 10.1145/191177.191188]

[24] Qian HB, Yan HH, Zhang ML,Yang HY, He ZT, Zhu XJ. A test-oriented software measurement and evaluation method. China, CN200910082587.4, 2009 (in Chinese).

[25] Mohamed A, Ruhe G, Eberlein A. COTS selection: Past, present, and future. In: Proc. of the 14th Annual IEEE Int'l Conf. and Workshops on the Engineering of Computer-Based Systems. 2007. 103−114. [doi: 10.1109/ECBS.2007.28]

[26] Liu C. Comprehensive review and assessment of the structure and code of trusted software and support tools. China Science and Technology Achievements, 2010,11(16):21−22 (in Chinese with English abstract).

[27] Shen GH, Huang ZQ, Qian J, Xu YJ, Hao J, Zhao WY, Peng X. Research on software trustworthiness evaluation model and its implementation. Journal of Frontiers of Computer Science & Technology, 211,5(6):553−561 (in Chinese with English abstract).

[28] Qin LG, Yang M, Fang K. Research on the simulation credibility evaluation assistant tool based on hierarchical evaluation. Computer Simulation, 2010,27(6):118−121 (in Chinese with English abstract).

[29] He JS. Research and implementation of BPM field component credible evaluation system [MS. Thesis]. Xi'an: Northwestern University, 2010 (in Chinese with English abstract).

[30] Yang J. Research and implementation of software credibility assessment tools [MS. Thesis]. Xi'an: Northwestern University, 2010 (in Chinese with English abstract).

[31] Van der Feesten I, Reijers HA, van der Aalst WMP. Evaluating workflow process designs using cohesion and coupling metrics. Computers in Industry, 2008,V01.59:420−437. [doi: 10.1016/j.compind.2007.12.007]

[32] Jiang R. A trustworthiness evaluation method for software architectures based on the principle of maximum entropy (POME) and the grey decision-making method (GDMM). Entropy, 2014,16(9):4818−4838. [doi: 10.3390/ e16094818]

[33] Wang HM, Tang YB, Yin G. Trustworthiness of internet-based software. Science in China-Series F: Information Sciences, 2006,49(6):759−773. [doi: 10.1007/s11432-006-2024-4]

[34] Viljanen L. Towards an ontology of trust. In: Proc. of the 2nd Int'l Conf. on Trust, Privacy and Security in Digital Business (TrustBus 2005). 2005,3592:175−184. [doi: 10.1007/11537878_18]

[35] Zhang LG, Zhang HG, Zhang F. The amount of credibility mechanism in turstede computing. Journal of Beijing University of Technology, 2010,36(5):586−591 (in Chinese with English abstract).

[36] Yuan L, Wang HM, Yin G, Shi DX, Mi HB. A role-based software credible assessment techniques. Journal of Beijing University of Technology, 2010,36(5):611−615 (in Chinese with English abstract).

[37] Zhang YJ, Zhang YM, Hai M. An evaluation model of software trustworthiness based on fuzzy comprehensive evaluation method. American Journal of Engineering and Technology Research, 2011,11(9):1145−1149.

[38] Tao HW, Chen YX. A metric model for trustworthiness of softwares. In: Proc. of the 2009 IEEE/WIC/ACM Int'l Conf. on Web Intelligence and Intelligent Agent Technology. 2009. 69−72. [doi: 10.1109/WI-IAT.2009.233]

[39] Liu YZ, Luo X, Xue K, Luo P. A metric model research based on attributes for trustworthiness of software. Computer Science and Application, 2012,2:121−125 (in Chinese with English abstract).

[40] Zhang LW, Zhou Y, Chen YX, Zhang M, Zhang JY. Stability of software trustworthiness measurements models. In: Proc. of the 7th Int'l Conf. on Software Security and Reliability Companion. 2013. 219−224. [doi: 10.1109/SERE-C.2013.23]

[41] Pedraza-Garcia G, Astudillo H, Correal D. Modeling software architecture process with a decision-making approach. The Jornadas Chilenas de Computación (JCC2014), 2014. http://www.jcc2014.ucm.cl/jornadas/WORKSHOP/WBPM%202014/WBPM-6.pdf

[42] Delgado A, Ruiz F, García-Rodríguez de Guzmán I, Piattini M. MINERVA: Model drIveN and sErvice oRiented framework for the continuous business process improVement and relAted tools. In: Dan A, Gittler F, Toumani F, eds. Proc. of the ICSOC/Service Wave 2009. LNCS 6275, 2010. 456−466. [doi: 10.1007/978-3-642-16132-2_43]

[43] Leilng KRPH, Leung HKN. On the efficiency of domain based COTS product selection method. Information and Software Technology, 2002,44:703−715. [doi: 10.1016/S0950-5849(02)00118-0]

[44] Sheng JF, Chen SQ, Wang B. Software requirements-driven COTS evaluation. Computer Engineering, 2005,(24):99−101 (in Chinese with English abstract).

[45] Golden B. Succeeding with Open Source. Reading: Addison-Wesley Professional, 2004.

[46] Semeteys R, Pilot O, Baudrillard L, Le Bouder G, Pinkhardt W. Qualification and selection of open source software (QSOS), Version 2.0. Technical Report, Atos Origin, 2013.

[47] OpenBRR. Business Readiness Rating for Open Source. A Proposed Open Standard to Facilitate Assessment and Adoption of Open Source Software, Request for Comments, 2005.

[48] Rong M. A model for CPS software system trustworthiness evaluation based on attributes classifying. In: Proc. of the 8th Int'l Conf. on Computer Science & Education (ICCSE 2013). 2013. 1309−1314. [doi: 10.1109/ICCSE.2013.6554124]

[49] Tang JB. Research on credibility of warfare simulation system [Ph.D. Thesis]. Changsha: University of Defense Technology, 2009 (in Chinese with English abstract).

[50] Bao T, Liu SF, Wang XY. Research on a trusted construction method for electric power production management system. Acta Electronica Sinica, 2010,38(9):2166−2171 (in Chinese with English abstract).

[51] Zhang JK, Cheng L, Huang J, Wu Z. Mission-Based operational effectiveness evaluation model of combat aircraft. Journal of Beijing University of Aeronautics and Astronautics, 2005,31(12):1279−1283 (in Chinese with English abstract).

[52] Wang XL, Wang HW. Requirements for trust evaluation of Web services. Computer Systems & Applications, 2009,(4):36−39 (in Chinese with English abstract).

[53] Palviainen IM. Trustworthiness evaluation and testing of open source components. In: Proc. of the 7th Int'l Conf. on Quality Software (QSIC 2007). 2007. [doi: 10.1109/QSIC.2007.4385514]

[54] Tian JF, Xiao B, Ma XX, Wang ZX. The trust model and its analysis in TDDSS. Journal of Computer Research and Development, 2007,44(4):598−605 (in Chinese with English abstract). [doi: 10.1360/crad20070408]

[55] Hur J, Lee Y, Yoon H, Choi D, Jin S. Trust evaluation model for wireless sensor networks. Advanced Communication Technology, 2005,491−496. [doi: 10.1109/ICACT.2005.245914]

[56] Perez M, Rojas T. Evaluation of workflow-type software products: A case study. Information and Software Technology, 2000,V01.42:489−502. [doi: 10.1016/S0950-5849(00)00093-8]

[57] Liu B, Fan YS. Service-Oriented workflow performance evaluation and correlation analysis for key performance indicators. Computer Integrated Manufacturing Systems, 2008,14(1):160−165 (in Chinese with English abstract).

[58] Li CX, Raghunathan A, Jha NK. Improving the trustworthiness of medical device software with formal verification methods. Embedded Systems Letters, 2013,5(3):50−53. [doi: 10.1109/LES.2013.2276434]

附中文参考文献:

[7] 王怀民,唐扬斌,尹刚,李磊.互联网软件的可信机理.中国科学(E 辑),2006,36(10):1156−1169.

[11] 林闯.可信网络研究.计算机学报,2005,28(5):751−758.

[12] 王怀民,尹刚.网络时代的软件可信演化.中国计算机学会通讯,2010,6(2):28−34.

[13] 樊晓光,褚文奎,张凤鸣.软件安全性研究综述.计算机科学,2011,38(5):8−13.

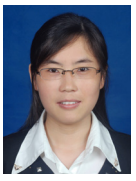
[14] 刘克,单志广,王戟,何积丰,张兆田,秦玉文.“可信软件基础研究”重大研究计划综述.中国科学基金,2008,22(3):145−151.
[15] 梅宏.软件可信性:互联网带来的挑战.中国计算机学会通讯,2010,6(2):20−27.
[16] 蔡斯博,邹艳珍,邵凌霜,谢冰,邵维忠.一种支持软件资源可信评估的框架.软件学报,2010,21(2):359−372. http://www.jos.org.cn/1000-9825/3786.htm [doi: 10.3724/SP.J.1001.2010.03786]
[17] 杨善林,丁帅,褚伟.一种基于效用和证据理论的可信软件评估方法.计算机研究与发展,2009,46(7):1152−1159.
[18] 刘旭东,郎波,谢冰,毛晓光,王怀民.软件可信分级规范.版本 2.0,国家高技术研究发展计划(863)重点项目“高可信软件生产工具与集成环境”技术文档,TRUSTIE-STC V2.0,2009.
[19] 卢刚,王怀民,毛晓光.基于认知的软件可信评估证据模型.南京大学学报(自然科学),2010,46(4):456−463.
[20] 丁学雷,王怀民,王元元,卢刚.面向验证的软件可信证据与可信评估.计算机科学与探索,2010,4(1):46−53.
[24] 钱红兵,晏海华,张茂林,杨海燕,何智涛,朱小杰.一种面向测试过程的软件可信性度量与评估方法:中国,CN200910082587.4,2009.
[26] 刘超.可信软件结构及代码的审查和综合评估及支持工具.中国科技成果,2010,11(16):21−22.
[27] 沈国华,黄志球,钱巨,徐拥军,郝进,赵文耘,彭鑫.软件可信评估模型及其工具实现.计算机科学与探索,2011,5(6):553−561.
[28] 秦立格,杨明,方可.仿真可信度评估辅助工具研究.计算机仿真,2010,27(6):118−121.
[29] 贺久松.BPM 领域构件可信评估系统的研究与实现[硕士学位论文].西安:西北大学,2010.
[30] 杨静.软件可信性评估工具的研究与实现[硕士学位论文].西安:西北大学,2010.
[35] 张立强,张焕国,张帆.可信计算中的可信度量机制.北京工业大学学报,2010,36(5):586−591.
[36] 袁霖,王怀民,尹刚,史殿习,米海波.基于角色的软件可信评估技术.北京工业大学学报,2010,36(5):611−615.
[39] 刘彦钊,罗峋,薛凯,罗平.一种基于属性划分的软件可信性度量模型研究.计算机科学与应用,2012,2:121−125.
[44] 盛津芳,陈松乔,王斌.软件功能需求驱动的商业构件评估.计算机工程,2005,(24):99−101.
[49] 唐见兵.作战仿真系统可信性研究[博士学位论文].长沙:国防科学技术大学,2009.
[50] 包铁,刘淑芬,王晓燕.电力生产管理系统的可信构造方法研究.电子学报,2010,38(9):2166−2171.
[51] 张建康,程龙,黄俊,武哲.基于任务的作战飞机效能评估模型.北京航空航天大学学报,2005,31(12):1279−1283.
[52] 王秀利,王宏伟.Web 服务可信评估要求.计算机系统应用,2009,(4):36−39.
[54] 田俊峰,肖冰,马晓雪,王子贤.TDDSS 中可信模型及其分析.计算机研究与发展,2007,44(4):598−605. [doi: 10.1360/crad20070408]
[57] 刘博,范玉顺.面向服务的工作流性能评价及指标相关度分析.计算机集成制造系统,2008,14(1):160−165.

沈国华(1976－),男,江苏丹阳人,博士,副教授,CCF 高级会员,主要研究领域为需求工程,软件可信评估,软件安全性.

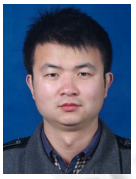
黄志球(1965－),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为软件工程,形式化方法,隐私保护.

谢冰(1970－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为软件工程,形式化方法,分布式系统.

朱羿全(1990－),男,硕士,主要研究领域为软件安全性,Web 服务.



廖莉莉(1989－),女,硕士,主要研究领域为本体度量,语义 Web,描述逻辑.



王飞(1990－),男,硕士生,CCF 学生会员,主要研究领域为软件安全性,软件可追踪性.

刘银陵(1989－),男,硕士生,主要研究领域为软件安全性.