

国家 863 重点项目

“高可信软件生产工具与集成环境”技术文档



软件可信分级规范

Software Trustworthiness Classification Specification

(TRUSTIE-STC V 2.0)

2009 年 5 月 30 日

发布声明

牵头单位：北京航空航天大学计算机学院

参研单位：国防科技大学计算机学院

北京大学信息科学技术学院

中国科学院软件研究所

中创软件公司

执笔人：刘旭东、郎波、谢冰、毛晓光

版本号：2.0

发布时间：2009年5月30日

审核人：王怀民

版权声明

本技术规范得到国家 863 重点项目“高可信软件生产工具及集成环境”第一课题“可信的国家软件资源共享与协同生产环境”的资助，版权归“可信的国家软件资源共享与协同生产环境”课题组所有。

本规范在以下条件下可以自由传播：

- (1) 保持本规范的完整性（包括发布声明）；
- (2) 未经课题组许可，任何人不得以本规范盈利。

Trustie 技术规范

变更记录

变更版本	日期	A/M/D	原因与修改情况描述	修订人	审核人
V1.0	2009.5.4	A	1.1 背景	郎波	刘旭东
V1.0	2009.5.4	M	软件可信级别定义	郎波	刘旭东
V1.0	2009.5.4	M	1.2 概述	郎波	刘旭东
V1.1	2009.5.4	M	第 3, 4, 6, 7, 8, 9 章	郎波	刘旭东
V2.0	2009.5.30	M	根据 09 年 5 月长沙讨论修改	刘旭东	

注：A - 增加 M - 修改 D - 删节

目 录

第 1 章	引言	1
1.1	背景.....	1
1.2	概述.....	2
第 2 章	范围与引用	5
2.1	范围.....	5
2.2	引用.....	5
第 3 章	术语和定义	6
3.1	软件.....	6
3.2	软件产品.....	6
3.3	软件资产.....	6
3.4	软件质量.....	6
3.5	软件可信.....	6
3.6	软件可信性.....	6
3.7	软件可信属性.....	6
3.8	软件可信等级.....	6
3.9	软件可信证据.....	7
3.10	软件可信分级指标体系.....	7
3.11	软件可信分级评估.....	7
第 4 章	软件可信分级框架	8
第 5 章	软件可信属性模型	10
5.1	可用性.....	11
5.1.1	功能符合性.....	11
5.1.2	功能准确性.....	11
5.1.3	易理解性.....	11
5.1.4	易操作性.....	11
5.1.5	适应性.....	11
5.1.6	易安装性.....	11
5.2	可靠性.....	11
5.2.1	成熟性.....	11
5.2.2	容错性.....	11
5.3	安全性.....	12
5.3.1	机密性.....	12
5.3.2	完整性.....	12
5.4	实时性.....	12
5.5	可维护性.....	12
5.5.1	易诊断性.....	12
5.5.2	可修改性.....	12
5.5.3	稳定性.....	12
5.5.4	易测试性.....	12

5.6	可生存性.....	12
5.6.1	抗攻击性.....	13
5.6.2	攻击识别能力.....	13
5.6.3	恢复性.....	13
5.6.4	自我完善性.....	13
第 6 章	软件可信等级定义.....	14
6.1	第 0 级—未知级.....	14
6.2	第 1 级—可用级.....	14
6.3	第 2 级—证实级.....	14
6.4	第 3 级—实用级.....	14
6.5	第 4 级—评估级.....	15
6.6	第 5 级—证明级.....	15
第 7 章	软件可信证据参考模型.....	16
7.1	软件开发阶段的可信证据.....	17
7.2	软件提交阶段可信证据.....	17
7.3	软件应用阶段的可信证据.....	18
第 8 章	软件可信证据度量与获取.....	19
8.1	软件提交阶段证据的度量.....	19
8.2	软件提交阶段可信证据的获取方法.....	22
第 9 章	软件可信分级评估.....	23
9.1	软件可信评估的基本方法与评估的动态性.....	23
9.2	一个软件可信评估机制示例.....	23
9.2.1	示例中的可信证据模型.....	23
9.2.2	示例中的可信分级定义.....	24

第1章 引言

1.1 背景

软件的可信性是用户所关心的重要软件特性，而高可信软件的设计与生产已经成为当前软件产业发展的重要目标。为了对软件可信的概念、软件可信等级定义、可信证据以及软件可信分级评估方法等进行规范性约定，为“可信的国家软件资源共享与协同生产环境”中软件资源的可信分级评估、以及基于该环境可信软件开发提供规范性参考和指导，“可信的国家软件资源共享与协同生产环境”课题组特开展了软件可信分级规范的研究和制定工作。该项工作由北京航空航天大学牵头负责，由国防科技大学、北京大学、中国科学院软件所、中创软件中间件有限公司等单位共同参与。

可信分级规范工作组从 2008 年 1 月份开始工作，一年来，工作组成员进行了大量的研究和广泛的交流讨论工作，最终形成了本参考规范 1.0 版，该规范包含附件《软件可信证据框架参考规范 1.0 版》。

本规范形成过程中的关键事件如下：

2008 年 3 月 29 日，在长沙召开的课题组技术会议上，课题负责人国防科技大学王怀民教授提出了一种可信等级定义，将软件可信等级定义为 6 级，即：未知级、可用级、实用级、自主验证级、权威评估级、理论证明级。在 2008 年 4 月至 5 月，课题组在软件开发协同平台的 Wiki 上，对可信的含义、可信的证据模型以及可信分级问题进行了广泛讨论。2008 年 5 月 29 日工作组在北京航空航天大学召开了第二次可信分级讨论会。会上对北京航空航天大学提出的基于实体可信属性的证据模型和国防科技大学提出的过程/实体/行为/应用/信誉模型，以及面向可信评估的分级机制等关键问题进行了研讨。

在 2008 年 6 月 29 日北京西郊宾馆召开的项目组交流会上，工作组做了可信分级规范草案的工作汇报。经过与会的专家与项目组成员研讨，确定了软件可信的含义，肯定了基于软件生命周期确定的过程/实体/应用的三阶段可信证据参考模型。

2008 年 10 月 26 日至 27 日，在长沙中创软件园召开的课题组技术会议上，

对可信分级规范进行了进一步的研讨。北京航空航天大学针对如何保证可信分级规范的科学性与实用性问题，提出在软件可信分级规范中，将与软件类型无关的概念与模型抽象为规范的理论基础，而进一步在理论上建立多种面向不同类型软件的可信分级指标体系，建立层次化的可信规范框架。这次会议上，确定了层次化的可信规范框架的思路，并基本确定可信分级规范将由可信分级总论、可信证据模型以及可信评估等三部分组成。

2008年11月14日，王怀民教授在无锡召开的中科院技术科学论坛第34次学术报告会议上做了题为“构建可信的国家软件资源共享与协同生产环境”的报告，该报告对软件可信等级进行了调整，即：未知级、可用级、验证级、实用级、评估级、证明级。2009年3月27日，在北京召开的课题组技术会议上，将“验证级”的名称改为“证实级”。

2008年11月27日在北京大学召开了由南京大学李宣东教授主持由北大、北航、中科院等多家单位代表参加的软件可信证据框架研讨会，确定了涵盖软件开发、软件实体提交和软件应用等三阶段证据的软件可信证据框架，并在项目各课题组范围内进行了意见征询和问卷调查。2009年3月27日，在北京召开的课题组技术会议上，对该证据框架进行进一步讨论，并从证据与软件可信的相关性等方面考虑，对证据框架中的某些证据项进行了修改。2009年5月11日在长沙召开的课题组技术会议上，再一次对可信证据框架进行了讨论和修改。

本规范起草的主要成员包括：国防科技大学王怀民、毛晓光、卢刚，北京航空航天大学刘旭东、郎波、刘超，北京大学谢冰、郝丹，中国科学院软件所魏峻、杨叶，南京大学李宣东。

1.2 概述

当前，以通信、存储和计算为核心的信息基础设施已经渗透到政治、经济、军事、文化和社会生活的各个层面，软件作为信息基础设施的灵魂，在信息社会中发挥着至关重要的作用。日趋庞大的软件需求愈来愈多，复杂度愈来愈高，可用性要求愈来愈强，软件系统却愈来愈脆弱，常常发生各种各样的问题，并对人们的工作生活带来不利的影响，甚至造成巨大的损失。例如：1996年6月，软件失效导致欧洲 Ariane 五型火箭首发失败；2003年8月，美国电力检测与控制管理系统的软件失效造成了美国东北部大面积停电；2006年，我国中航信离港

系统发生三次软件故障，造成近百个机场值机系统瘫痪，等等。人们发现软件并不总是完全可以让人信任的，其行为和结果并不完全符合人们的预期，因此，人们对软件的正确性、可靠性、安全性、可生存性等“可信”性质给予了高度关注，这就是所谓的“软件可信性”问题。高可信软件的设计与生产技术已成为软件领域重要研究方向。

从可信软件的生产和应用的需求出发，软件开发过程中所集成的服务、构件和架构等软件资源以及所开发完成的软件系统是否可信、可信的程度如何都将成为人们关注的重要问题，而如何确定一个软件是否可信、以及如何度量软件的可信程度则是软件可信分级的主要研究内容。软件可信分级通过软件可信等级评估来进行。

软件可信分级评估模型由三个部分组成，如图 1 所示。软件可信等级定义即对软件的可信级别进行划分，并对每一可信级别的含义进行定义，用第 0 级、第 1 级、第 2 级等表示，不同的级别表示软件具有不同程度的可信性。软件可信证据模型则依据软件可信等级的定义，规定了达到某个可信等级的软件需要具有的证据集合。软件可信等级评定则依据可信等级定义，并根据所获得的可信证据和特定的评估准则确定软件的可信等级。

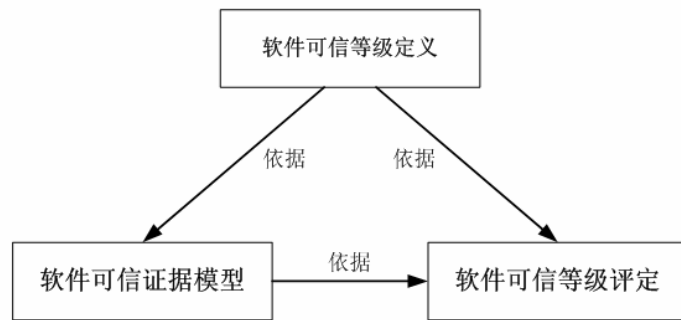


图 1 软件可信分级评估模型

因此，要进行科学、合理的软件可信分级，就要给出恰当的软件可信等级定义，确定能充分支持软件可信等级的可信证据集合，并建立有效的软件可信等级评估机制。

本规范旨在对软件可信的概念、软件可信属性、软件可信等级定义、软件可信属性的度量方法以及软件可信分级评估方法等进行明确的规定，为“可信的国家软件资源共享与协同生产环境”资源库中的软件资源可信分级提供标准规范。

本规范围绕软件可信分级评估这个核心问题，从软件可信等级定义、软件可

信证据框架以及软件可信等级评定三个方面，对软件可信的内涵（定义）、可信属性模型、可信证据模型、软件可信等级定义和软件可信等级评定框架等进行规范性描述，为软件可信评估奠定基础，为软件可信评估机制的建立提供指导。

在软件可信内涵方面，给出软件可信的定义，研究了软件可信与可信属性之间的关系，提出软件可信属性模型，为软件可信等级评定确立概念基础。

依据软件生命周期的阶段划分，建立了由软件开发阶段可信证据、软件提交阶段可信证据和软件应用阶段可信证据等三个阶段证据组成的软件可信证据模型，同时给出了一种可信证据的度量方法。

在软件可信等级定义方面，针对软件满足用户所期望可信属性的程度和所提交可信证据的类型，提出了软件可信等级的划分和每一等级的定义，为软件可信等级评定提供依据。

在软件可信等级评定方面，提出一种由软件可信分级依据、软件可信证据以及软件可信分级评估三部分组成的层次化可信分级框架，以此确立软件可信分级评估的基本方法体系。

第2章 范围与引用

2.1 范围

本规范对软件可信的概念、软件可信属性、软件可信等级定义、软件可信属性的度量方法以及软件可信分级评估方法等进行明确的规定，本规范还给出了一种软件可信分级评估的参考模型。

本规范适用于软件制品的可信分级，也适用于有需要的组织和个体对目标软件进行可信评估和改进。软件的可信性随着可信等级的增高而增强。

2.2 引用

下列文件中的有关条款通过引用而成为本标准的条款。凡注日期或版次的引用文件，其后的任何修改（不包括勘误的内容）或修订版都不适用于本标准，但提倡使用本标准的各方探讨使用其最新版本的可能性。凡不注明日期或版次的引用文件，其最新版本适用于本标准。

ISO/IEC 14598-1:1999 Information technology - Software product evaluation-Part 1: Grneral overview

ISO/IEC TR 9126-1: 2003 Software engineering – Product quality

第3章 术语和定义

3.1 软件

信息处理系统的全部或部分程序、规程、规则和相关文档。

3.2 软件产品

一组计算机程序、规程以及可能有的相关文档和数据。又称为软件制品。

3.3 软件资产

是软件生存周期中具有复用价值的软件制品，它包括软件需求、软件设计、软件代码、测试案例和文档等。软件资产可以用于创造新的软件产品以及相关的制品。又称软件资源。

3.4 软件质量

与表征软件能够满足规定或隐含需求能力有关的软件的所有特性。

3.5 软件可信

如果一个软件系统的行为总是与用户预期的行为和结果相一致，则称该软件可信（trustworthy）。

3.6 软件可信性

软件按照用户期望提供安全可靠服务的能力。

3.7 软件可信属性

用以描述和评价软件系统可信的一组属性。软件可信属性可被细化成多级子属性。

3.8 软件可信等级

尺度中某一范围的值可用来按照明确或隐含的需求对软件可信进行分级评定，这些可信级别称为可信等级，又称为软件可信级别。软件的可信等级的评定与软件的应用领域、软件类别等密切相关。

3.9 软件可信证据

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。

3.10 软件可信分级指标体系

软件可信等级定义中的各个可信等级应该具有的软件可信证据及其度量值的集合，称为软件可信分级指标体系。

3.11 软件可信分级评估

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，称为软件可信分级评估，简称为软件可信评估。

Trustie 技术知识

第4章 软件可信分级框架

软件可信分级的目的是通过分级的方式对软件制品的可信性进行标度，以方便用户根据需求选用恰当的产品。不同应用领域和不同类别软件的用户对于软件制品在可信属性的要求存在很大的差别。国家标准《信息技术 软件产品评价 质量特性及其使用指南》（GB/T 16260-1996）中对软件质量等级评估有过这样一段描述：“质量与给定需求有关，不可能有通用的等级，每一次具体的评价都必须对等级进行定义”。软件可信等级也是一样，一方面是因为软件可信等级与软件质量等级之间具有十分密切的关系，另一方面，衡量一个软件系统可信程度的高低，与应用领域的用户对该类型软件的相关可信属性的期望相关，不同应用领域、不同类别的软件，其用户关心的可信属性以及对这些可信属性的期望值都可能不同，因此，不存在一种适用于所有软件制品的统一的可信分级标准。

虽然用户对于各种类型软件可信属性的期望有所差异，但软件可信的内涵对于各种类型软件应该是基本一致的，基于这种可信的内涵，可以定义一种标度软件可信程度的可信等级。而某一特定类别软件的可信等级的评定，需要根据该类软件用户的期望，确定度量软件可信的证据并建立特定的分级机制。因此，软件可信的内涵与可信等级的定义，可以作为软件可信分级的基本依据。图 2 给出了分层式软件可信分级框架，表示了本规范的基本思想。

本规范将首先通过建立软件可信的概念模型（即软件可信的定义）和软件可信属性模型来确定软件可信内涵，并明确软件可信等级的定义；在此基础上，进一步从软件开发、软件提交和软件应用等软件生命周期的三个阶段，分析影响软件可信属性的重要因素，构建软件可信证据的参考模型。该模型对各类软件的可信评估体系的建立都具有指导意义。不同应用领域、不同类别的软件可信评估具有其特定性，即不同类型软件由于用户的期望不同，因此需要建立特定的可信证据模型和相应的可信分级指标体系。

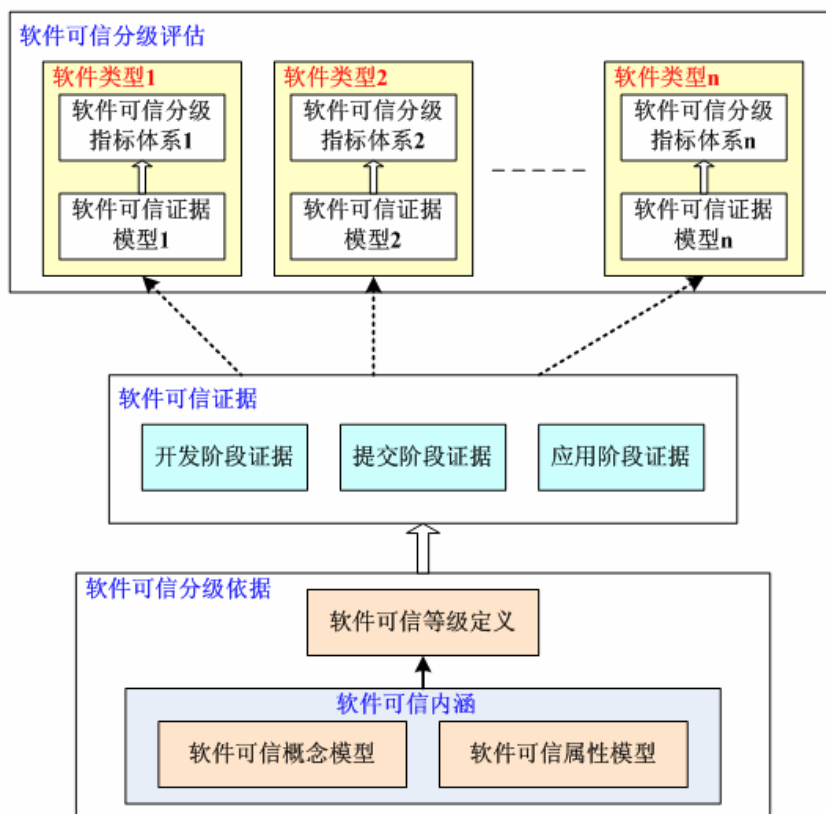


图 2 软件可信分级框架

图 2 所示，本规范对于软件可信分级的定义与描述分为三个层次，分别是软件可信分级依据、软件可信证据以及软件可信分级评估部分。

软件可信分级依据是本规范的理论基础，主要包括两个部分。一是通过软件可信概念模型与软件可信属性模型展现的软件可信内涵；另一部分是基于软件可信内涵的软件可信等级定义，可信等级表达软件对可信属性的满足程度，是软件可信程度的一种标度。

软件可信证据从软件生命周期的开发阶段、提交阶段、应用阶段提取软件过程保障、软件实体、软件应用信誉等三个维度的可信信息，定义了支持软件可信分级评估的证据模型。该模型不特定于任何类型软件，但对于建立特定的可信证据模型具有指导作用。

软件可信分级评估面向软件可信分级的具体实施。评估的基础是软件可信证据模型以及一种基于可信分级证据模型建立的软件可信分级指标体系或评估机制。软件可信分级指标体系由证据对软件可信级别的支持关系组成。在对某类软件的软件可信评估时，需要面向该类软件建立特定的证据模型与分级指标体系。

第5章 软件可信属性模型

软件可信属性是软件（客体）获得用户（主体）对其行为实现预期目标的能力的信任程度的客观依据。主体通过客体所具有一组表达其可信属性的客观能力事实，从而信任客体的行为能够实现其设定的目标。因此，若软件可信，则意味着软件拥有了一系列与软件可信属性相关的能力；反过来，若软件具有了一系列与软件可信属性相关的能力，则可以相信该软件能达到其预设目标。进一步，有关软件可信证据模型的构建、软件可信分级的评估都将以软件可信属性为基础。

根据目前关于软件可信性的几种典型概念框架和软件质量模型，本规范规定软件可信属性是软件按用户的期望提供正确、安全、可靠等特性服务的能力，不但应涵盖功能性、可靠性、易用性、效率、可维护性和可移植性等软件质量特性，还应包括安全性、实时性、可生存性等其他软件特性。

考虑到软件可信属性测量和度量的可操作性，软件可信分级评估可只关注主要的软件可信特性。因此，本规范定义软件可信属性包括：可用性（Availability），可靠性（Reliability），安全性（Security），实时性（Real Time），可维护性（Maintainability）和可生存性（Survivability）。上述每个特性又由若干子特性构成，这些属性构成了软件可信属性模型，如图3所示。

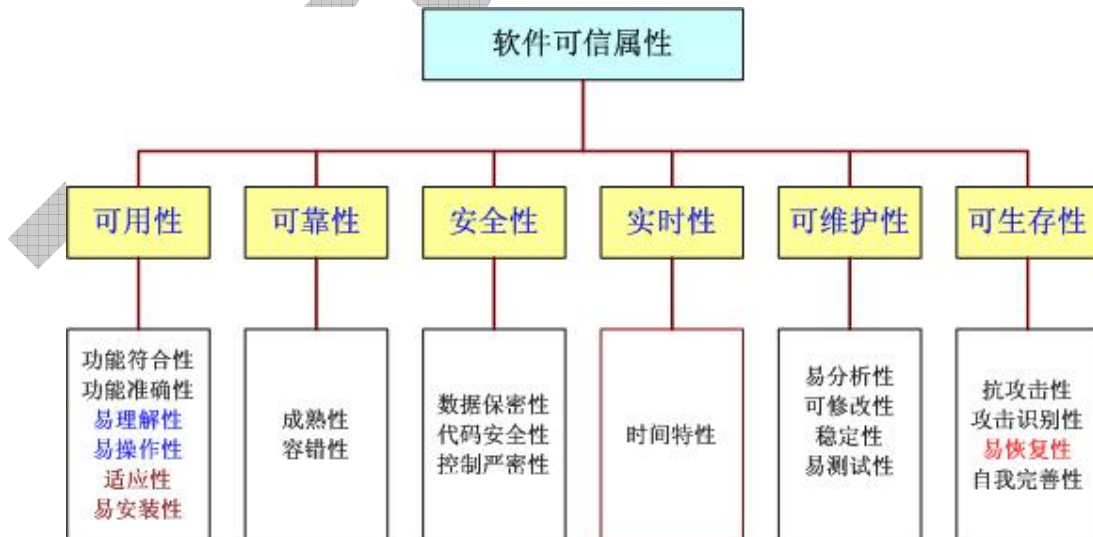


图3 软件可信属性模型

5.1 可用性

当软件在指定条件下使用时，软件产品持续提供满足明确和隐含需求的功能的能力，以及软件产品被理解、学习、使用和移植的能力。

5.1.1 功能符合性

软件产品为指定的任务和用户目标提供一组合适的功能的能力。

5.1.2 功能准确性

软件产品提供具有所需精确度的正确或相符的结果及效果的能力。

5.1.3 易理解性

软件产品能够使用户理解软件是否满足要求，使用户知道在特定背景下如何使用软件，以及使用的条件。

5.1.4 易操作性

软件产品使用户能操作和控制它的能力。

5.1.5 适应性

软件产品无需采用特殊手段就可能适应不同的指定环境的能力。

5.1.6 易安装性

软件产品在指定环境中被安装的容易程度。

5.2 可靠性

在规定的环境下、规定的时间内软件无失效运行的能力。

5.2.1 成熟性

软件本身存在的故障而导致的软件失效的可能程度。

5.2.2 容错性

在软件出现故障或者违反指定接口的情况下，软件产品维持规定的性能级别

的能力。

5.3 安全性

软件系统对数据和信息提供保密性、完整性、可用性、真实性保障的能力。

5.3.1 机密性

软件系统中的信息不被非法用户所获取。

5.3.2 完整性

软件系统中的信息不被非法篡改。

5.4 实时性

软件在指定的时间内完成反应或提交输出的能力。

5.5 可维护性

软件产品可被修改的能力。修改可能包括修正、改进或软件适应环境、需求和功能规格说明中的变化。

5.5.1 易诊断性

软件产品诊断软件中的缺陷或失效原因以及标识待修改部分的能力。

5.5.2 可修改性

软件产品使指定的修改可以被实现的能力。

5.5.3 稳定性

软件产品避免由于软件修改而造成意外结果的能力。

5.5.4 易测试性

软件产品使已修改部分能被确认的能力。

5.6 可生存性

软件在受到攻击或失效出现时连续提供服务并在规定时间内恢复所有服务的能力。

5.6.1 抗攻击性

软件抵抗攻击的能力。

5.6.2 攻击识别能力

软件探测已经发生的入侵并评估其危害程度的能力。

5.6.3 恢复性

软件在被攻击后，恢复服务的能力。

5.6.4 自我完善性

针对从干扰及攻击中获得的信息来改进系统生存性的策略，从整体上增强系统的可生存性的能力。

Trustie 技术园地

第6章 软件可信等级定义

用户对某一类别软件所期望的可信属性是进行软件可信等级划分的基础，而软件对用户所期望的可信属性的满足程度是软件可信等级划分的重要依据。用户对软件所期望的可信属性的满足程度的高低实际上反映了软件在可信性上的各种能力，这些能力可以通过软件测试、分析与验证工具进行评估，或者在软件的应用和维护过程中得以体现。

下面基于用户对软件所期望的可信属性的满意程度给出软件可信等级的定义。

6.1 第 0 级—未知级

未获得关于软件可信性的任何证据，不能判定软件是否能满足用户对该类别软件可信属性的期望，软件的可信等级定义为未知级。

6.2 第 1 级—可用级

软件实体可访问，并且能按照软件提供者指定的模式正常运行，隐含表明该软件能满足用户对该类别软件可信属性的基本期望。软件的可信等级定义为可用级。

6.3 第 2 级—证实级

在可用级的基础上，软件提供者依据特定的已成文的软件可信属性发布规范发布软件可信属性声明，该声明可通过软件可信性分析、测试或验证工具以及其它可信评估机制进行确认，表明该软件能满足用户对该类别软件可信属性的普遍期望，且用户期望的可信属性均得到了确认。软件的可信等级定义为证实级。

6.4 第 3 级—实用级

在证实级的基础上，软件已在相关应用领域得到应用，并且有可证实的成功应用案例，隐含表明该软件能满足用户对该类别软件可信属性的普遍期望，且得到实际应用的证实。软件的可信等级定义为实用级。

6.5 第 4 级—评估级

在实用级的基础上，软件的可信性通过了权威软件可信分级评估机构依据特定的已成文的可信分级评估规范进行的评估，表明该软件能满足用户对给类别软件可信属性的较高期望，且用户期望的可信属性均得到了权威机构的评估保证。软件的可信等级定义为评估级。

6.6 第 5 级—证明级

在评估级的基础上，所提交的软件可信性属性都是可被严格证明的，软件的可信等级定义为证明级。证明级是最高的可信级别。

Trustie 技术规范

第7章 软件可信证据参考模型

软件可信性是主体的主观感受，软件可信评估是对这种主观感受以分级的方式进行客观表达，是依据软件对可信属性的满足程度进行分级的。而对某种可信属性满足程度的确定，需要依据特定的证据进行。软件所具有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。一种可信属性可能通过多个可信证据从不同的角度反映出来。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。

从软件的全生命周期角度来看，软件可信性受到软件研制过程的影响，并通过软件自身的特性与软件的应用情况以及软件的信誉等方面表现出来。因此软件可信证据包括如下三个方面：

- 软件开发阶段的证据：软件设计与生产的规范化（可信的设计与生产）为软件的可信性提供重要保障，因此软件开发阶段的证据成为软件可信证据的重要组成部分。这部分证据关注软件过程对软件可信性的影响。
- 软件提交阶段的证据：生产完成并被提交的软件称为软件实体。软件实体的自身可信特性是用户判断“实体是否符合期望”的主要依据，相关证据称为软件提交阶段的证据。这些证据集中体现在软件可信属性模型中的各个特性上，主要通过自动化的分析、测试和验证工具，以及人工分析和评估等手段获得。
- 软件应用阶段的证据：软件应用案例、应用的广泛程度、用户的满意程度和第三方的评价等也是建立用户对软件实体信心的重要依据。相关证据称为软件应用阶段的证据。

因此，本规范依据软件可信属性模型，从软件生命周期的开发阶段、提交阶段以及应用阶段，归纳可信证据，定义了一种软件可信证据参考模型，如图 4 所示。

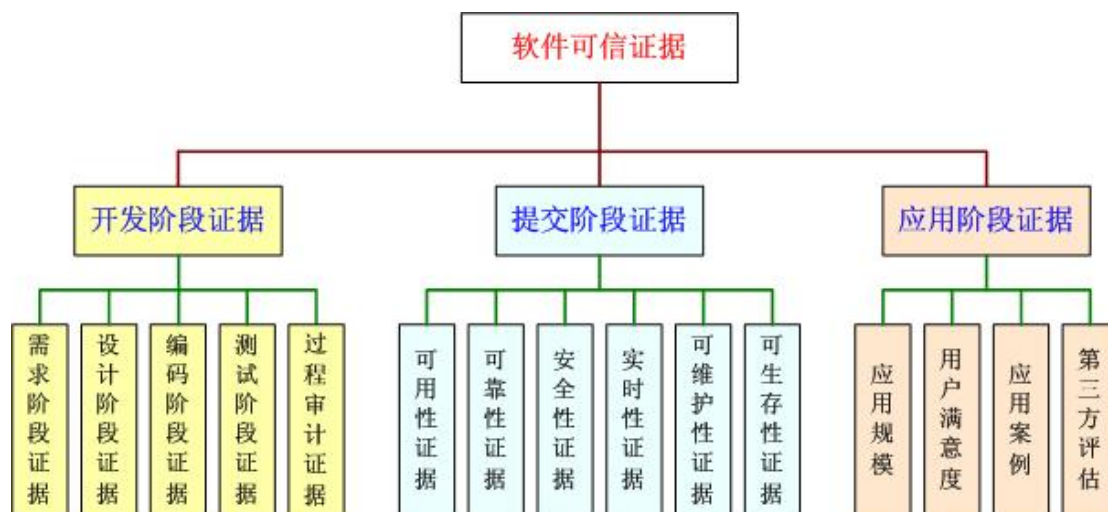


图 4 软件可信证据模型

7.1 软件开发阶段的可信证据

软件开发阶段的证据是在软件开发中的需求分析、设计以及编码与测试等过程中，能够提供的与保障软件可信性相关的一系列证据，主要包括：

- 需求阶段证据：需求规范描述方法，是形式化、半形式化还是非形式化，是否引用标准或引用模型；需求分析人员的能力等级；需求变更频率，即需求变更次数与需求总数之比；需求变更比率，即总需求变更数与总需求数之比；需求评审结论，需求阶段的评审缺陷密度和缺陷清除率等都是需求分析阶段的有效可信证据。
- 设计阶段证据：设计人员的能力等级，设计阶段的需求变更数，设计评审结论，设计阶段的评审缺陷密度和缺陷清除率等都是该阶段有效的可信证据。
- 编码阶段证据：编码人员的能力等级，编码阶段的需求变更数，单元测试强度、代码规模、代码可维护性等都是该阶段有效的可信证据。
- 测试阶段证据：测试人员的能力等级、测试工具支持的有效性以及测试缺陷趋势等都是该阶段有效的可信证据。
- 过程审计证据：过程不合格项产生的趋势是该阶段有效的可信证据。

7.2 软件提交阶段可信证据

软件提交阶段的可信证据指从软件实体可以获取的可信证据，也称为软件实体证据，这些证据集中体现在软件可信属性模型中的各个特性上，主要通过自动

化的分析、测试和验证工具，以及人工分析和评估等手段获得。主要包括：

- 软件可用性证据，包括功能符合性、功能准确性、资源利用率、易理解性、易操作性以及易安装性等证据。
- 软件安全性证据，主要包括数据保密性、代码安全性以及控制严密性等证据。
- 软件可靠性证据，主要包括软件成熟性以及容错性证据。
- 软件可生存性证据，主要包括软件抗攻击和识别攻击的能力、软件在被攻击后的恢复自恢复能力以及自我完善能力。
- 软件实时性证据，主要指软件的时间响应特性。
- 软件可维护性证据，主要包括软件故障的易诊断性，错误易于修改的程度，错误修改后软件运行的稳定性，以及软件维护的易测试性等。

7.3 软件应用阶段的可信证据

应用阶段证据是在软件实际使用过程中形成的与软件可信性相关的证据，包括：

- 应用方面证据：主要指软件的应用案例、应用规模以及用户对软件的满意程度等。
- 第三方评价：独立第三方对该软件的综合评价也将作为评估该软件可信性的重要证据。

软件可信证据模型是面向各类软件的一种概念模型。由于不同类型的软件，用户对可信性质的期望有所不同，因此不同类型软件的可信证据模型也是各不相同。特定类型软件的可信证据模型应该在本规范提出的可信证据参考模型的基础上形成。

第8章 软件可信证据度量与获取

三类软件可信证据中，软件开发阶段证据的载体是软件开发过程中的各类文档、规范、评审和测试数据等，软件应用信誉证据主要是软件提供者的基本属性信息以及通过合适途径获取的软件应用与信誉等信息，而软件提交阶段证据的获取与度量是最为复杂的，也是对软件可信性影响最直接的。本规范将重点对软件提交阶段的证据即软件实体证据的度量与获取方法进行说明。

8.1 软件提交阶段证据的度量

软件可信属性度量是使用特定的度量方法和度量尺度，给一个软件可信属性赋予一个值。

在本规范中，软件可信证据度量分为原始度量与分级度量两个层次。原始度量是指按照证据特性所确定的证据初始度量。分级度量是在原始度量的基础上，对原始度量的值域按照一定规则进行完整分割，每个分割称为一个等级，由此将证据的原始度量划分为若干个级别，并可根据属性原始度量所属的级别确定其分级度量的值。

本规范中给出软件可信证据分级度量的定义，如表 1 所示。关于软件可信证据原始度量以及分级度量中每一级的含义（即对原始度量的分割），本规范中不做具体定义，可参见软件测试与评估的相关标准。

表 1 软件可信属性分级度量

可信证据名称	值域	原始度量定义	分级度量定义
可用性	$0 \leq X \leq 1$	$X = cf * (1 - A/B)$ ($0 \leq X \leq 1$) A=在评价中检测有问题的功能数 B=被评价的功能数 cf = 评测机构的影响因子(信任度)	5级：功能完全正确 ($X=1$) 4级：功能略有问题 ($0.85 \leq X < 1$) 3级：功能有问题 ($0.7 \leq X < 0.85$) 2级：功能有较多问题 ($0.5 \leq X < 0.7$) 1级：功能有很多问题 ($0 \leq X < 0.5$)
	$0 \leq X \leq 1$	$X = 1 - A$ ($0 \leq X \leq 1$) A=软件产品没有提供所需精确度的功能数的比例	5级：准确 ($X=1$) 4级：比较准确 ($0.85 \leq X < 1$) 3级：基本准确 ($0.7 \leq X < 0.85$) 2级：部分准确 ($0.5 \leq X < 0.7$) 1级：不准确 ($0 \leq X < 0.5$)

	资源利用率	$0 \leq X$, 单位:	内存使用大小, CPU 使用率等	5级: 好 4级: 较好 3级: 中 2级: 较差 1级: 差
	易理解性	$0 \leq X \leq 1$	$X = cf * (1 - A/B)$ ($0 \leq X \leq 1$) A=不能被理解的功能(或功能的类型)数 B=功能(或功能的类型)总数 cf= 用户的影响因子(信任度)	5级: 好 ($X=1$) 4级: 较好 ($0.85 \leq X < 1$) 3级: 中 ($0.7 \leq X < 0.85$) 2级: 较差 ($0.5 \leq X < 0.7$) 1级: 差 ($0 \leq X < 0.5$)
	易操作性	$0 \leq X \leq 1$	$X = 1 - A/B$ ($0 \leq X \leq 1$) A=用户发现不易操作的功能数 B=功能的总数	5级: 极易操作 ($X=1$) 4级: 较易操作 ($0.7 \leq X < 1$) 3级: 可操作 ($0.5 \leq X < 0.7$) 2级: 较难操作 ($0.3 \leq X < 0.5$) 1级: 难操作 ($0 \leq X < 0.3$)
	适应性	$0 \leq X \leq 1$	$X = 1 - A/B$ ($0 \leq X \leq 1$) A=在结合环境硬件进行测试时未能完成任务的运行功能数 B=进行测试的功能总数	5级: 好 ($X=1$) 4级: 较好 ($0.85 \leq X < 1$) 3级: 中 ($0.7 \leq X < 0.85$) 2级: 较差 ($0.5 \leq X < 0.7$) 1级: 差 ($0 \leq X < 0.5$)
	易安装性	$0 \leq X \leq 1$	$X = A/B$ ($0 \leq X \leq 1$) A=用户为自己的方便成功地改变安装操作的次数 B=用户为自己的方便企图改变安装操作的总次数	5级: 好 ($X=1$) 4级: 较好 ($0.85 \leq X < 1$) 3级: 中 ($0.7 \leq X < 0.85$) 2级: 较差 ($0.5 \leq X < 0.7$) 1级: 差 ($0 \leq X < 0.5$)
可靠性	成熟性	$0 \leq X \leq 1$	$X = cf * A$ ($0 \leq X \leq 1$) A=测试过程对软件产品测试覆盖的充分度	5级: 非常充分 ($X=1$) 4级: 较充分 ($0.85 \leq X < 1$) 3级: 中 ($0.7 \leq X < 0.85$) 2级: 不充分 ($0.5 \leq X < 0.7$) 1级: 差 ($0 \leq X < 0.5$)
	容错性	$0 \leq X \leq 1$	$X = 1 - A/B$ ($0 \leq X \leq 1$) A=损坏发生的次数 B=软件失效的数目	5级: 好 ($X=1$) 4级: 较好 ($0.85 \leq X < 1$) 3级: 中 ($0.7 \leq X < 0.85$) 2级: 较差 ($0.5 \leq X < 0.7$) 1级: 差 ($0 \leq X < 0.5$)
安全性	机密性	{5, 4, 3, 2, 1}		5级: 安全级别为验证保护级(A)($X=5$) 数据加密存储传输
	完整性			4级: 安全级别为强制保护级(B)($X=4$) 数据加密存储传输 3级: 安全级别为可控存取保护级(C2)($X=3$)

				2级:安全级别为自主安全保护级(C1)(X=2) 1级:安全级别为无保护级(D)(X=1)
可 生 存 性	抗攻击性	$0 \leq X \leq 1$	$X = \{Tr/Ta\}$ Ta =被攻击次数 Tr =成功抵抗攻击次数	5级:好($0.9 \leq X \leq 1$) 4级:较好($0.75 \leq X < 0.9$) 3级:中($0.6 \leq X < 0.75$) 2级:较差($0.4 \leq X < 0.6$) 1级:差($0 \leq X < 0.4$)
	攻击识别能力	$0 \leq X \leq 1$	$X = \{Tr/Ta\}$ Ta =被攻击次数 Tr =成功识别攻击次数	5级:好($0.9 \leq X \leq 1$) 4级:较好($0.75 \leq X < 0.9$) 3级:中($0.6 \leq X < 0.75$) 2级:较差($0.4 \leq X < 0.6$) 1级:差($0 \leq X < 0.4$)
	恢复性	$0 \leq X \leq 1$	$X = \{Tr/Ta\}$ Ta =遭受攻击次数 Tr =成功恢复次数	5级:好($0.9 \leq X \leq 1$) 4级:较好($0.75 \leq X < 0.9$) 3级:中($0.6 \leq X < 0.75$) 2级:较差($0.4 \leq X < 0.6$) 1级:差($0 \leq X < 0.4$)
	自我完善能力	$0 \leq X \leq 1$	$X = \{Tr/Ta\}$ Ta =遭受攻击种类数 Tr =通过自我完善成功抵抗攻击的种类数	5级:好($0.9 \leq X \leq 1$) 4级:较好($0.75 \leq X < 0.9$) 3级:中($0.6 \leq X < 0.75$) 2级:较差($0.4 \leq X < 0.6$) 1级:差($0 \leq X < 0.4$)
	实时性	$0 \leq X$, 单位: 秒	$T = cf * ((\text{获得结果的时间}) - (\text{完成命令输入的时间}))$	5级:快 4级:较快 3级:中 2级:较慢 1级:慢
可 维 护 性	易诊断性	$0 \leq X \leq 1$	$X = 1 - \sum Ai$ ($0 \leq X \leq 1$) Ai=不利于软件产品分析的因素	5级:极易分析(X=1) 4级:较易分析($0.85 \leq X < 1$) 3级:可分析($0.7 \leq X < 0.85$) 2级:较难分析($0.5 \leq X < 0.7$) 1级:难分析($0 \leq X < 0.5$)
	易改变性	$0 \leq X \leq 1$	$X = 1 - \sum Ai$ ($0 \leq X \leq 1$) Ai=不利于软件产品修改的因素	5级:极易修改(X=1) 4级:较易修改($0.85 \leq X < 1$) 3级:可修改($0.7 \leq X < 0.85$) 2级:较难修改($0.5 \leq X < 0.7$) 1级:难修改($0 \leq X < 0.5$)
	稳定性	{5, 4, 3, 2, 1}		5级:很稳定 4级:较稳定 3级:稳定 2级:较不稳定 1级:不稳定

	易测试性	$0 \leq X \leq 1$	$X = A/B$ ($0 \leq X \leq 1$) A = 实际测试中使用了内 置式测试方法或数据的测 试用例数 B = 实际测试使用的测试 用例数	5级: 很有效 ($X=1$) 4级: 较有效 ($0.85 \leq X < 1$) 3级: 有效 ($0.7 \leq X < 0.85$) 2级: 效果较差 ($0.5 \leq X < 0.7$) 1级: 无效 ($0 \leq X < 0.5$)
--	------	-------------------	---	--

8.2 软件提交阶段可信证据的获取方法

软件提交阶段可信证据的获取方法主要有软件测试与分析、软件的服务质量 (QoS) 监测等, 如表 2 所示。软件提交阶段可信证据的获取操作主要在软件试运行阶段进行。

表 2 软件提交阶段可信证据的获取方法

可信证据名称		软件可信证据获取方法	
		软件测试与分析	QoS 监测
可用性	功能适合性	√	
	准确性	√	
	资源利用率	√	√
	易理解性	√	
	易操作性	√	
	适应性	√	
	易安装性	√	
可靠性	成熟性	√	
	容错性	√	√
安全性	机密性	√	√
	完整性	√	√
可生存性	抗攻击性	√	√
	攻击识别能力	√	√
	恢复性	√	√
	自我完善能力	√	√
实时性		√	√
可维护性	易诊断性	√	√
	可改变性	√	
	稳定性	√	
	易测试性	√	

第9章 软件可信分级评估

9.1 软件可信评估的基本方法与评估的动态性

在不同应用领域中，用户对于软件的期望可能包括不同的内涵，具体表现在所关注的软件可信属性的种类和关注程度不同。因此，在对一个软件进行可信等级评估时，应首先分析确定用户所关注的软件可信属性集，并根据软件对该可信属性集的满足程度，对其进行可信分级。

软件可信评估是确定软件可信级别的过程。软件可信评估中，需要针对特定类型的软件，基于用户对该类软件可信性的期望，并依据可信等级的定义，确定该类软件可信证据模型，利用一定的方法和技术获取软件可信证据并对可信证据进行度量，最后依据可信分级定义并依据可信证据确定软件可信级别。

在软件可信评估中，由于评估者能够获取的证据类型与准确程度受限于环境因素和技术上的可操作性，一些实际可获取的证据数据可能不是直接对应到软件可信证据参考模型中的某项证据，而是一种间接对应关系，因此可信证据可能有多种形态。

另外，软件可信分级应该是具有动态性，即随着软件自身可信特性的提高/降低、软件可信证据的增加/减少或增强/减弱，软件的可信级别也随着动态变化。

9.2 一个软件可信评估机制示例

9.2.1 示例中的可信证据模型

图 5 所示给出了一种可操作的软件可信证据模型。

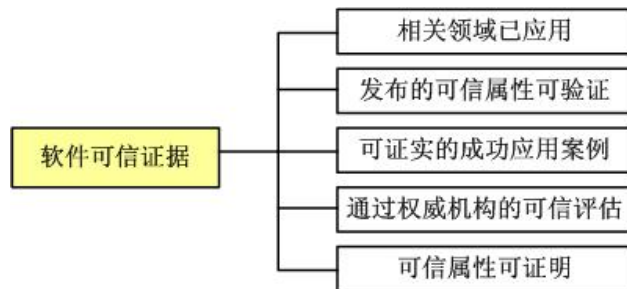


图 5 一种软件可信证据模型

- 相关领域已使用：软件在其所属的应用领域得到应用，这可通过软件投

入使用的时间、使用的频度以及发行量等进行标度。

- 自发布的可信属性可验证：软件生产者发布的有关软件可信属性声明可得到验证。
- 可证实的成功应用案例：能够提供证据，如用户使用报告等，证明软件在相关领域的成功应用案例。
- 通过权威机构可信评估：软件通过所属专业领域中的权威评估机构进行的软件可信评估。相关评估机构所采用的软件可信评估标准应符合本领域软件的用户期望可信属性集。
- 可信属性可证明：软件的用户期望可信属性可证明。

9.2.2 示例中的可信分级定义

基于上述简单的软件可信证据模型，本节中给出一种软件可信分级的评估机制示例，如表 3 所示。

表 3 基于软件可信证据模型的软件可信分级机制示例

等级 证据	未知级	可用级	证实级	实用级	评估级	证明级
相关领域已使用		√	√	√	√	√
可信属性可证实			√	√	√	√
成功应用案例				√	√	√
通过权威机构可信评估					√	√
可信属性可证明						√