铁路安全软件测试评估

TIELU ANQUAN RUANJIAN CESHI PINGGU

主编　吴芳美
主审　赵志熙

中国铁道出版社

铁道科技图书出版基金资助

铁路安全软件测试评估

主编　吴芳美
主审　赵志熙

中　国　铁　道　出　版　社
2001年·北京

(京)新登字 063 号

内 容 简 介

本书介绍了安全软件测试的基本思想、评估基础和实现途径，测试评估平台构成的技术。并以铁路信号安全软件的测试为实例，阐述了安全软件测试评估平台的结构以及软件测试的具体方法。本书还介绍了实际使用的铁路车站信号计算机联锁软件测试评估平台、计算机联锁软件出厂检测装置和能进行现场测试的便携式联锁软件测试仪的概况和操作方法。这些思路和方法对铁路其他采用计算机控制的系统，以及铁路以外其他安全控制系统来说都是适用的，对于希望深入了解这些技术的人员会有启发作用。

本书可以作为相关专业的研究生教材，也可供从事相关工作的科研和工程技术人员参考。

图书在版编目(CIP)数据

铁路安全软件测试评估/吴芳美著 .—北京：中国铁道出版社，2001.8
ISBN 7-113-04241-4

Ⅰ.铁… Ⅱ.吴… Ⅲ.铁路运输－交通运输安全－软件－测试 Ⅳ.U298

中国版本图书馆 CIP 数据核字(2001)第 042027 号

书　　名：铁路安全软件测试评估
主　　编：吴芳美
出版发行：中国铁道出版社(100054，北京市宣武区右安门西街 8 号)
责任编辑：傅立谚　　编辑部电话：路电(021)73115
封面设计：李艳阳　　市电(010)63549465
印　　刷：中国铁道出版社印刷厂
开　　本：787×1092　1/16　印张：9.75　字数：240 千
版　　本：2001 年 10 月第 1 版　　2001 年 10 月第 1 次印刷
印　　数：1～2 000 册
书　　号：ISBN 7-113-04241-4/TP·574
定　　价：38.00 元

前　言

安全控制(Safety Critical Control)问题和许许多多技术领域有关。凡是涉及可能危及人的生命或大宗财产损失的系统,如航空、航天、铁路、核能、化工、医疗仪器中都存在安全控制的问题。在这些系统中安全问题不光是需要保证系统在正常情况下的安全运行,还必须考虑在系统出现某些不正常的情况下,如何避免出现危及人身安全或大宗财物损失的事故或减少事故的损失。铁路是大容量和大众化的交通运输工具,安全运输是铁路追求的重要目标。通常所说"安全第一"表明安全在铁路的地位。"故障—安全(Fail-Safe)"理念的技术化也源于铁路。

近些年来,由于计算机被大量引入控制系统,控制系统的技术要求提高了,控制的内容被大大地丰富了。相应地,系统的复杂性也随之增加。系统中可能导致重大损失的不安全因素,如果没有作必要的特殊考虑,也会随之增多。

在采用计算机控制的系统里,各种复杂的功能主要依靠软件来实现。因此嵌入在安全控制系统中的软件,除了保证完整地实现系统的控制功能外,当发生意外事件时的安全防护也是考察软件性能的一个重要的指标。

软件的编制主要依靠人的脑力劳动。由于人的因素造成软件差错而引起严重后果的事例不胜枚举。人因差错(Human Initiated Error)可能来自开发人员对系统需求或者对前一道工序理解上的偏差,可能来自对任务缺乏周密的考虑,也可能来自开发过程中的某种疏忽。对于复杂系统来说开发过程中的软件差错几乎是不可避免的。

软件测试正是发现软件差错,纠正差错的主要手段。按照系统的需求对软件实施尽可能完备的测试是保证软件可靠性的关键。但是由于用来验证软件所需要的测试案例量非常大,对一个比较复杂的软件来说常常是一个难以承受的开销。但是开销是否值得支出,就要看一个可靠软件会带来的得益,和一个不可靠软件可能带来的损失,应该通过审慎的权衡来做决定。对安全系统来说,首先要做一个完整的功能测试是必不可少的。同时还必须考虑尽可能多的安全性测试,所谓安全性测试就是用包括可能发生的意外事件在内的案例进行测试。意外事件包含一个庞大的集合,一般是不可能穷尽的。这就需要通过广泛地采集意外事件的信息,归纳专家的意见,挑选最有效的测试案例,并通过精心的测试脚本设计,实现最低的测试支出达到最佳的测试效果。要实现高的测试效率,还有重要的一点是尽可能实现自动测试,以减少人工投入,降低测试成本。

黑箱测试是开发方及用户以外的第三方测试和软件开发中后一个工序对前一工序结果进行验收中常用的测试办法。黑箱测试基本上是对系统行为(Behavior)的检验和确认而不涉及软件内部的结构,这就可以绕过知识产权问题以及回避对许多复杂细节的考察。实现自动黑箱测试需要一套能自动生成测试案例、演绎测试脚本、判定系统行为和记录测试结果的测试平台。当被测件和平台通过网络连通后,平台即能无人介入自动运转。测试人员只须监视测试

结果并对测出的问题进行人工或自动复核。

软件的功能、可靠性，特别是安全性的定性、定量评估是一项需求迫切和实现难度大的课题。一般来说，定性评估较容易，定量评估较难。至今尚没有一套令大家满意和非常实用的软件安全性定量评估理论和方法。在本书中，提出基于软件黑箱测试和风险分析的安全性分级定量评估理论，并结合铁路安全控制软件的实际，给出一种分级定量评估方法及软件安全性投放准则和软件安全性比较准则，试图通过实践验证理论和方法的正确性。

本书介绍了安全软件(Safety Critical Software)测试的基本思想、评估基础和实现途径，测试评估平台构成的技术。并以铁路信号安全软件的测试为实例，阐述了安全软件测试评估平台的结构以及软件测试的具体方法。本书还介绍了实际使用的铁路车站信号计算机联锁软件测试评估平台、计算机联锁软件出厂检测装置和能进行现场测试的便携式联锁软件测试仪的概况和操作方法。这些思路和方法对铁路其他采用计算机控制的系统，以及铁路以外其他安全控制系统来说都是适用的。

本书虽然是介绍铁路信号安全软件的测试评估技术，但是也从另一个方面介绍了铁路信号，特别是车站信号采用计算机控制时的软件设计思想和设计方法。对于希望深入了解这些技术的人员会有启发作用。

阅读本书时，可能需要一点近世代数、概率论和随机过程方面的知识，读者可以参阅相关的书籍，但不必过于深入，以能理解本书的一些公式为限。

本书可以作为相关专业的研究生教材，也可供从事相关工作的科研和工程技术人员参考。

本书由同济大学（原上海铁道大学）吴芳美教授主编，虞翊、徐中伟、屠海滢博士，朱程荣副教授，李巍巍、荆剑、覃崇乾硕士参加了编写工作。书中第一章由吴芳美编写；第二章由吴芳美、朱程荣共同编写；第三章由李巍巍、徐中伟共同编写；第四章由虞翊、覃崇乾共同编写；第五章由徐中伟编写；第六章由屠海滢编写；第七章由荆剑编写；第八章由李巍巍编写；第九章由虞翊编写；第十章由吴芳美编写。

全书由北方交通大学赵志熙教授主审。

本书在编写过程中同济大学郦萌教授对全书提出了十分具体的意见并作了修改和审订。屠海滢博士对全书作了大量的编辑排版和插图的设计工作。在此一并表示感谢。

编　者

2000 年 12 月

目　录

第一章 概 论

软件作为有别于电气、电子元器件等硬件的新元素被引进了计算机系统，有人认为，软件寿命是无限的，因为软件没有物理的磨损或耗散。它的可靠性问题与硬件相比较具有较特殊的属性，因此，软件的可靠性和安全性的研究开展得比较晚，与硬件相比研究的难度也更大。研究软件可靠性的领域比较广，而研究软件安全性的领域则比较窄，主要在航空、航天、核电站、铁路、金融、网络等部门处于重要地位。而且在不同部门中的安全性含义又不尽相同。金融、网络系统通常指的是保密意义上的安全性(Security)；而铁路、航空、核电站等安全控制系统中的安全性(Safety)的精髓则是故障—安全(Fail-Safe)，即要求满足故障导向安全原则。

在铁路信号计算机控制安全软件中，安全侧和危险侧不像在继电器控制时有明显的物理对比，如继电器接点闭合与断开的位置、通电断电等。而在软件中，问题可能就在一条指令或一项数据的某些差错或一次不当的修改就埋下危险的隐患。因此对在铁路计算机安全控制和防护系统中的软件有更高的安全性要求。

安全软件一般是指对那些安全性要求很高的软件，例如控制航天飞机、控制武器火力、控制管理核电站、铁路车站计算机联锁控制及列车超速防护等的软件。在开发这些软件时，一般应遵循如下原则：

1. 安全系统软件必须保证系统遵循故障安全原则。发生故障时，必须使系统维持在安全状态或转向安全状态并锁定在安全状态。

2. 安全系统软件的功能，特别是安全保障功能，只能比继电控制系统功能高而不能低。

3. 应能防止人为非正常操作带来的事故，以提高整个人—机系统的安全性。

4. 安全软件设计还应着眼于提高安全系统的包括可靠性、可用性、可维护性、可测试性、健壮性等在内的可信性。

5. 每大类设备安全软件应用程序的基本模块应具备通用性。

6. 对安全关键软件，必须列出可能出现的不期望事件，例如，铁路的不期望事件是指将导致行车事故的事件，并给出防止不期望事件出现的软件处理要求。

具体在铁路安全控制系统，上列原则又进一步细化成如下原则：

1. 铁路信号计算机安全控制和防护系统(以下简称铁路安全系统或称铁路信号系统)除了保证行、调车作业安全外，还要为提高运输效率、改善劳动条件和文明生产创造条件。

2. 安全系统中的安全软件系统整体上必须具有可靠性的要求，并对故障采取符合故障安全原则的尽可能的对策，即满足铁路安全系统的安全性(亦称铁路信号的安全性)要求。

3. 按安全性要求划分软件安全性完善度等级，并采用与确定等级相适应的技术和措施。

4. 根据软件安全性完善度等级，遵照软件质量保证体系和软件生命周期来设计、开发和测试软件。即按软件工程化开发原则开发安全系统的安全软件。

5. 遵循国际电信联盟 ITU(International Telecommunication Union)建议的通信协议确定系统的通信方式，以便和其他信息系统联网。

6. 在铁路安全控制系统中，在安全功能和系统功能不可分割的情况下，可在系统需求规

格说明中包含与安全有关的需求说明，即在基本需求说明中反映出安全性需求。

7. 在编制软件需求规格说明时，应同时提出一个合适的软件安全性体系结构。

8. 集成软件和硬件后，要验证是否在整体上构成了要求的系统安全性完善度等级上的安全软件。

9. 每一个安全软件产品需要单独加以确认和进行安全性评估，并将结果作为软件文档的一部分交给用户。

10. 如果在运行过程中要求维护软件，可按有关的规范、准则进行。

11. 应对软件开发人员、测试评估人员的能力有相应的要求。

上列许多活动都是同软件开发过程交叉进行的，其中包括软件验证、评估和质量管理。

在硬件可靠性和安全性技术取得众多成果的今天，对软件可靠性和安全性的深入研究就显得十分重要，并引起国内外相关领域专家的极大重视。近 30 年来计算机技术迅速发展的历程表明，在 20 世纪 60 年代末提出的“软件危机”并非是耸人听闻的，只需举几个有名的例子便可说明。60 年代中期，美国飞船的一次失败仅仅是由于没有发现用 FORTRAN 语言编写的程序中在 DO 语句漏掉了一个标点符号。直至 80 年代，由软件造成的危害并没有减少，首次发射的航天飞机也因在预定发射时间前 20 分钟发现一个软件错误而被迫推迟了发射。至于由于软件错误使放射治疗机射线过量释放导致 5 位病人死亡的事故更使人心有余悸。这说明，当软件复杂到一定程度后，软件中的错误能够躲过一系列检验而在系统中隐藏下来，并将在某些特定的环境下影响运行，带来危害。Dijestra 对软件的测试作用就有“只能证明它有错，而不能证明它没有错”的名言。国外 70 年代以来十分关注对软件可靠性的研究，我国在该领域的研究起步较晚，始于 80 年代。

自 20 世纪 70 年代以来，国内外致力于将计算机应用于铁路信号系统。近十余年来，计算机控制的信号设备或系统在国外已得到较为广泛的应用。国内也正在积极开发，有越来越多的该类设备投入使用。由计算机取代传统的继电器实现信号设备安全控制和进行行、调车作业的指挥，已成为铁路信号控制和防护系统发展的主流。特别是近年来，在我国铁路以提速、重载为目标进行技术改造的环境下，计算机控制的信号设备得到广泛采用。例如，用于车站安全控制和防护的信号设备，即车站计算机联锁设备在我国得到始料未及的迅猛发展。出现在整条线电气化改造工程中，全线车站装备计算机联锁的局面。在站场扩容、继电联锁到大修期限等情况下，也选择计算机联锁作为更替设备。加上全国几千个联锁车站的大市场的吸引力，不仅国外多家厂商把目光投了过来，国内也有许多单位开发生产计算机联锁装备。对于这样一个有着极高安全性完善度等级要求的设备或系统来说，如何判定其是否满足功能和安全性需求，如何科学地加强质量管理已到了迫在眉睫的时刻。

为了使我国开发的计算机控制信号系统，特别是安全控制软件达到国际先进水平，保证铁路信号软件的质量，需要研究制定计算机信号系统的安全软件一整套技术条件、设计规范和管理方式。编写设计规范，确定各开发阶段统一的文档格式和内容要求。制定一套软件验收中必须遵循的程序和规定。在多方面积累经验和已取得技术成果的基础上，有必要根据最新技术制定为专家认定，用户和开发者能共同接受和遵循，并能推动技术发展的铁路计算机控制信号安全软件技术条件。以便能更好地利用软件手段进一步提高系统安全性。安全软件技术条件应详细规定软件安全性完善度等级的划分原则、人员及职责、软件生存周期模型、软件各开发阶段技术要求、软件确认和评估、软件质量保证及系统配置、技术措施及文档内容和格式等具体要求等。其中提出的要求分为强制执行的、建议执行的和指南性质的。其内容将有助于

提高铁路信号系统安全软件的安全性，并为我国在此领域的安全性技术得以按国际、国内标准进行客观公正的评价。

软件工程化包括了软件产品本身和软件开发过程两个方面。它们都必须符合软件工程学所规定的标准。一般将软件开发标准分为3类：

1. 强制性标准——用于所有软件开发而无一例外地被强制执行；

2. 建议实施的标准——虽不是强制性的，但强烈地建议实施，不执行时要作出特别说明和报请有关技术主管部门批准；

3. 指导性的标准——不执行可以不说明理由。

我国铁路信号部门需要根据自己的特点按不同类别标准一一作出规定。

铁路安全系统开发原则中有关安全性完善度及其等级部分将在第二章讨论，测试评估将是贯穿全书的核心内容。本章仅就软件体系结构、软件工程化开发及软件工程管理展开讨论。

一、软件体系结构

软件体系结构对满足软件功能需求和安全性需求是至关重要的。应依据软件需求规格说明、安全需求规格说明、软件质量保证计划编制软件系统结构说明和软件安全性计划。由软件开发者建立软件体系结构，该结构应考虑与安全性完善度等级相适应，并在软件体系结构说明书中详细叙述。软件体系结构应该：

1. 考虑按软件安全性完善度等级的要求实现软件需求规格的可行性。

2. 识别、评价和详细说明系统硬件、软件相互作用的有效性。

3. 识别所有软件成分是否是新开发的、已存在的或专利保护的；它们是否经过安全成分还是非安全成分的确认及确认条件。

4. 对应不同的安全性完善度等级，使用标准软件时，应作如下限制：等级0可直接使用；等级1或2，要经测试确认后使用；等级3或4，使用前除进行测试确认外，还应进行失效分析，确立失效影响的防护系统，宜使用标准软件中有稳定可靠性记录而且较简单的功能块。

5. 在设计中可使用已经验证的根据标准开发的软件模块。

6. 软件体系结构规格说明应阐明在软件安全性完善度等级要求范围内的软件开发、集成、验证、确认和维护策略。该规格说明除了要求清楚、精确、不含糊、可验证测试等外，还应能返回跟踪软件需求规格说明。

7. 软件体系结构说明应包括软件组成及采用的故障检测技术、容错和避错技术、安全性技术、人工智能技术等。软件组成主要包括：操作系统、程序结构、数据结构、输入输出、接口、通信等。

二、软件工程化开发和软件工程管理

这里特别要强调软件的工程化开发和质量管理。实现软件工程化开发及软件质量管理是改变软件手工作坊式生产和缺少科学管理现状的重要途径。

（一）软件工程化开发

软件开发管理是以保证软件质量为中心目标。开发中在综合考虑进度、经费、质量等因素时，应把软件的可靠性和安全性放在首要地位。

对于铁路信号这种计算机应用新领域，由于软件手工“作坊”生产方式，潜伏着很多安全上的隐患。解决问题的出路是要把软件当作产品来“生产”，按软件工程学的方法（即软件开发和

生产过程的工程化)生产软件。当按软件工程的要求来开发软件时,应当全面分析软件开发中的各道“工序”,按“工序”将软件的开发和生存过程划分为计划、需求分析和说明、系统设计、编码(编程)、测试及投入使用后的运行维护等几个阶段。规定各阶段分别完成的不同任务。诚然,不同阶段之间又是密切相关的,前面阶段的结果直接影响后面阶段的开发,后面阶段发现的问题需要一直追溯到前面阶段的工作。并在开发过程中,实行分阶段的工程管理和质量管理。这样的关系绘在框图上,看似逐次泻落的瀑布,故称瀑布模型,如图 1-1-1 所示。应当指出,在软件开发管理过程中文档具有重要作用,所有开发及维护活动都必须反映在文档中。

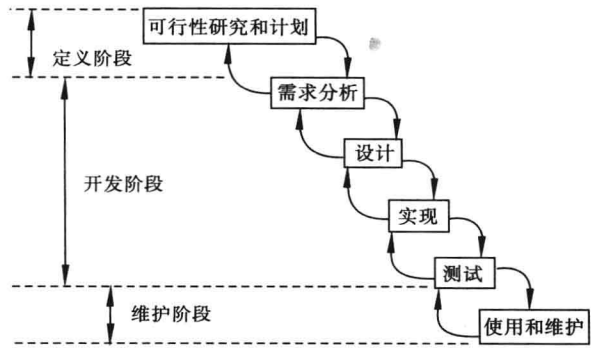

图 1-1-1 生存周期的瀑布模型

在软件生存周期的定义阶段,即将软件开发分成几个确定的阶段和规定每一阶段应展开的活动。全体开发人员必须经过系统的技术培训,使他们对阶段的划分和各阶段的任务有统一的认识。

在按工程化开发软件的过程中应首先确定软件生存周期模型,例如选用瀑布模型。注意在各开发阶段实施时有可能同步运作的软件质量保证计划,在某阶段开始前即对该阶段活动的起止加以分析界定,最主要的是完成软件开发计划。在需求分析阶段,通过软件需求分析,编写详尽、正确、无二义性等的软件需求规格说明书。它是其后各阶段的依据,主要包括开发应具备的条件、硬软件资源、约束条件、人员分工及职责等,力求其完整。阶段成果是软件需求规格说明书。

设计阶段又可细分为二个阶段:概要设计阶段和详细设计阶段。对较为简单的系统可省略概要设计阶段。这个阶段是依据需求规格说明进行软件设计直到模块级和模块的内部细节,为编码提供必要的说明。阶段成果是概要设计和详细设计说明书。

实现阶段为编写程序和程序单元测试。阶段成果是可供集成和集成测试的程序单元。

测试阶段主要工作是根据需求规格说明的要求制定测试计划,进行软件集成测试和系统(软、硬件)集成测试。阶段成果是提供安装验收的系统软件和软件测试报告及安全性评价报告。

运行维护阶段从软件通过验收测试投入运行开始,运行中应进行正确性、完善性、适应性等维护。软件生存期的最后一个阶段是报废阶段。

这里需要再特别强调一下在整个软件开发过程中建立软件开发文档的重要性。文档对于安全软件的质量管理和保证软件的开发、运行和维护质量极为重要。在软件生存周期模型确定以后,为了达到安全性管理效果,应该在生存周期的每一个阶段按要求如期地建立规定的文档。文档中应记录软件整个生存周期的所有信息。开发过程中的各种活动都必须记录在文档中。每个阶段文档是以后各阶段工作的依据,又是对前面阶段工作的复审。软件工程管理在很大程度上是通过文档来实现的。而文档标准化是文档管理的重要技术手段。通过文档的建立,应该保证在任何一位软件开发人员因故不能继续工作时,其他水平相当的人员能够不困难地接替他的工作,使开发工作得以继续进行。这样,不仅确保工程项目能按时按质地完成,也

保证承担软件生产的法人利益不因个人的原因而受到损害。同时软件文档增加了软件对用户的透明度,这将为软件的运行、维护提供了保障手段。

软件设计应和硬件相适应。应能证明软件和硬件集成后的相互作用能够完成系统功能和安全性需求。

为了完成软件硬件集成,应根据系统需求规格说明书(含软件、硬件)、系统安全性需求规格说明书(含软件、硬件)、系统体系结构说明书(含软件、硬件)及设计、测试说明书等文档,制定软硬件集成计划并在集成后完成软硬件集成报告。集成计划中含有集成测试案例生成和测试步骤及方法。集成除将软件纳入计算机硬件外,还包含有与传感器及其他对象的集成。

在制作集成计划时应考虑:硬件选用、总线检测、加电检测、电源失效防护、主控计算机失效防护、传感器(仪表及其他器具)失效防护、电磁干扰防护、人为误操作防护、维护开销等。

(二) 软件工程管理

软件工程管理是提高软件质量的重要保证,为此,必须加强软件开发工程管理。应在现有技术领导体系和质量管理体系的基础上形成适合软件开发特点的软件工程管理体系。

在软件质量保证工作中必须借鉴专家的重要见解和质量保证过程中的工作经验。在软件开发任务确定之后,任务承担一方应该制定质量保证计划。该计划的制订和实施是软件开发管理的一个重要组成部分。计划中主要应包括:计划的实施机构、机构的职责、软件配置管理、文档及其管理、各开发阶段的评审和审查、设施和工具的要求、问题报告和修改活动、记录所有软件管理的一切活动及其他管理项目等。

在安全性质量管理计划中要确定在软件安全性生存周期中的所有管理及活动的明确定义。如,为满足《技术条件》的手段和技术、评价过程、有效性检验、配置管理的变更控制、人员培训及再培训、监测安全功能的频度、检测评估机构(人员)的独立性等。

软件管理的主要职能包括:

1. 人员管理

按软件项目需要,在各开发阶段配置适量的不同层次的技术人员和管理人员。

2. 计划管理

按项目需要,规定任务、目标、资源分配及工作进度等。

3. 标准化管理

指对软件开发期的各阶段技术性的和管理性的活动作出合理的、统一的规定,包括设计标准化、软件文档标准化及软件项目管理标准化等,例如,用标准化的格式编写文档。

4. 软件配置管理

在软件工程化开发的软件生存周期的各阶段活动中产生的各种软件配置项(如,文件、数据、表格、报告等),进行标识、控制、审查等的管理活动。

当今软件开发、运行和维护已被提到管理技术的高度而得到相当的重视。一般地说,要完成大的复杂的软件系统开发和运行维护任务,不采取有效的管理技术是不行的。这需要一套井井有条的管理思路和严格的管理手段。在软件的整个生存周期,各阶段的成功不但是建立在技术质量和设计质量的基础之上,也是以管理质量为基础的。必须重视软件项目的管理还与当前软件开发背景有关,原因之一就是绝大部分的有经验的软件工程师并不同时具备管理才能,而具备管理才能的人又不一定懂许多软件开发知识,致使软件质量得不到保证。由于软件工程的成败与管理有着密切的关系,因而它是涉及软件整个生存周期的工作。软件工程管理已开始得到计算机软件工作者的重视,特别是领导和决策人的重视。

虽然我国铁路尚未形成信号软件生产的专门企业，但从现在开始，把软件作为“产品”来认识不算为时尚早。首先要求供、求双方及技术监督的主管部门都要掌握软件产品的特殊性，然后按其特性对产品生产的全过程进行有效的管理。

软件质量保证目的是通过必要的技术和管理活动使软件达到要求的质量性能。并证明这些技术和活动在软件生存周期（安全性生存周期）的全过程中已被实现。

在现行技术条件下，无论是质量保证方法（所谓预防措施）的应用还是软件容错方法的应用均不能保证系统的“绝对”安全性。现有的方法和技术尚无法证明在相当复杂的安全软件中不存在错误，尤其是规格说明和设计中的错误。为此，安全软件必须设计成能在任何条件下检测出整个软件或系统潜在的不安全条件或状态，当检测出一个潜在的危险状态时，除了即刻提醒操作人员注意并采取行动，同时能够通过恢复到某个安全状态的方法来防止危险的发生和发展，应保证不回避检测出来的不安全状态，软件对不安全状态应有停机措施。这是对故障安全的必要补充。

软件是非实物性的产品，它：可见性差，具有抽象性；具有严密的逻辑性，不允许存在差误；软件的批量生产不应该看作是简单的复制；软件工程活动包含了大量的高强度脑力劳动，致使软件质量难于用简单的手段加以度量；运行中还会不断出现错误及必须进行维护等特点。软件的这些特点给软件质量管理带来了许多实际困难。

目前我国安全软件开发、测试、鉴定和验收手段与预期的软件工程化管理尚有不小距离，只有少数部门投入一些人力和物力进行本领域研究和开发。例如航空航天部的有关研究所曾成功开发了专用的软件测试评估平台。在铁路信号领域此项工作已经迈出了坚实的一步并取得了阶段性的成果，它们是用以验证标准和规范的实施，以及设备投产前的功能验证、安全性和可靠性定性及定量评估所必须的。对安全软件测试评估方法进行深入的研究，在此基础上开发一套测试评估工具，并把它作为对信号安全软件规定的工程标准和规范来加以实施。

需要说明的是，不论是那一种评价方法，由于系统软件的开发者、测试者和用户对待软件的测试评估的主观意愿是不同的，开发者主观上往往希望在测试时软件没有错误；用户则希望找出尽可能多的错误和那些被认为是隐患的问题以排除后患。为此，信号软件测试和评价需要由独立于开发者和用户的专门人员使用规定的标准、方法及工具来进行第三方测试，并在测试的基础上对安全软件给出公正客观的评估。

第三方测试不仅能按我国产品质量监督检验部门要求的做到“公正性、科学性和权威性”，能对研制方送检的软件作出公正客观的评估，而且能促使研制者不断改进软件，在完善功能的同时，提高软件的可靠性和安全性。第三方测试者从测试中获取各研制方的技术信息及专家知识，在不断提高测试评估水平的同时，还能推荐各家之所长，促使我国计算机联锁软件水平的进一步提高。

软件的安全性评价是所有涉及安全的计算机控制系统面临的重要课题。通过安全性测试实施对软件安全性验证和评价的探索性方案并加以实现。同时也把铁路信号设备过去以手工方式进行开通试验作为安全验证手段上升到一个以计算机和人工智能为基础的测试评估的新台阶。它不仅可以模拟手工试验无法考虑的一些意外环境，而且可以快速而有效地探测到被测系统在安全防护上可能达到的极限。这个极限就是进行安全性评价的依据。实践表明，用已经开发成功的铁路车站计算机联锁软件测试评估平台，对若干个车站联锁软件进行了测试（以自动测试为主，手工测试为辅），取得了很好的社会效益和经济效益。从得到的测试结果分析，由于安全软件的缺陷，存在必须认真对待的危及安全的问题和影响运输效率的问题。这一

测试评估工具除了适用铁路车站计算机联锁软件外，完全可以用于其他计算机控制的信号设备，特别是比较复杂的系统。扩展功能的测试评估平台同样也可以用来实施对系统的软硬件可靠性、安全性的测试和综合评价。还可以进行包括温度及其他物理环境变化条件下的测试评估。

测试评估平台提供的数据为建立铁路安全性定量指标模型创造了条件，使我们有可能在这些数据基础上，应用现代数学工具推断比较符合实际的安全性指标。

在下面的章节中，在提出安全软件测试评估的有关理论、方法、实现途径等问题时，全书主要以铁路车站计算机联锁软件测试评估为例，深入阐述独立于开发方和用户方的第三方测试评估的方法及工具，并扩展到其他铁路安全软件的测试评估。对联锁软件的测试评估分为3个层次：

1. 联锁软件制式测试评估

指计算机联锁系统开发商所开发的一种联锁软件制式是否能满足我国车站联锁的基本要求的测试评估。测试的对象是一个为由铁路业务主管部门指定的标准车站设计的联锁软件。这一层次的测试具有考察开发商对计算机联锁技术需求的理解水平和解决联锁技术问题及完善安全性能力的性质。这一测试由铁道部产品质量监督检验中心铁路车站计算机联锁检验站进行。在这一层次对联锁功能要作到完备(穷举)测试，并要加载尽可能多的安全性测试案例，并对是否达到需求规格(含功能及安全性需求)作出评价。

2. 联锁软件出厂检测

指对每一个即将投入使用的车站联锁系统产品，在出厂前对其联锁软件进行测试。由厂方在规定的出厂检验装置上进行，通过测试的产品方可出厂。

3. 联锁软件开通验收及日常维护测试

指新设备到达现场或站场改造、软件更新后，开通前的验收测试和日常维护中的测试(如类似继电集中联锁进行每年一次的联锁功能测试)。这类测试可以采用便携式检测装置，通过按 TCP/IP 标准协议的网络互连，由冗余模块转接，相继对备用或冗余模块及主模块中的软件进行测试。

车站计算机联锁系统是一个强实时性的逻辑运算系统，它的安全性完善度等级为4级，是最高的等级。以它为例展开下面章节的讨论将具有典型性和不失一般性。

第二章　铁路安全控制软件评价基础及实现途径

铁路安全控制软件的评价依据是软件的安全性完善度等级。对已确立等级的软件的评价要按对应等级的要求进行。安全性完善度等级和安全性之间的差别在于安全性完善度等级本来并不直接评价安全性，而是根据开发人员对被开发系统安全性问题的估计、防范设计的完善性出发的一种安全性评价，因此，它的用途更多地用来作为开发阶段的指导。按不同等级提出一系列指导性的意见，比如对应某一个等级，应该采用什么技术或不应该采用什么技术。如果仅从黑箱测试而言，测试方无法了解软件系统内究竟采用什么技术。但是铁路安全软件验证制度要求被测方提供必要的文档作为检验的一个部分。因此根据安全性完善度的要求来考察被测软件仍然是软件检验的重要部分。另一方面，根据欧洲铁路信号软件标准，安全性完善度等级实际上还是蕴涵了每个等级对安全性的定量要求，软件中每一个被检测出的差错都可以根据它造成危害的可能性和危害带来的损失，来评价被测软件的安全性。在黑箱测试基础上的评价也需将测试案例分成几个对应等级的子集，为评价作必要的数据准备。这些案例集可以按安全性完善度等级划分，采用对比环境或动态判定（表）判定软件是否存在缺陷，然后根据缺陷属于哪一个等级，就可以判定被测软件的安全性完善度。

第一节　铁路安全控制软件评价基础

一、关于软件安全性完善度等级

安全性完善度及其等级是一个安全系统设计、开发的准绳，是对系统评价的依据。因此在国内外的相关技术标准都把它作为关键概念贯穿其中。对软件也按安全需求将其划分成不同的安全性完善度等级。这里首先明确给出安全性完善度及等级的定义。

安全性完善度（Safety Integrity）——在安全系统中体现确保安全的能力。其定量指标可以采用在给定时刻系统维持安全功能完善的概率来表示。

软件安全性完善度等级（Software Safety Integrity Level）——对软件所要求的安全性完善水平的一种定量指标。是将安全性完善度根据软件关键功能失效的频率和产生的危险严重程度划分成的等级。

首先应该对软件安全性要求进行分级，然后按安全性要求将软件划分成不同的安全性完善度等级。而安全性要求通常是按软件失效后果的严重性确定的。对不同等级的软件提出不同的技术要求，以便适当地选用软件可靠性及安全性技术措施来达到相应的技术条件。采用国内、外推荐的方法一般将软件分为 4 个软件安全性完善度等级：

A 级（4 级）——软件失效将造成人员伤亡和（或）大宗财产损失的严重后果，一般称之为灾难性的或致命的后果；

B 级（3 级）——软件失效将造成人员伤残、引起严重职业病和（或）一般财产损失，通常称之为关键性的或严重的后果；