军用软件测试过程的质量管理改进

付鹏程
中国软件评测中心　北京　102206

摘　要：随着信息化、智能化战争形态的加速演进，军用软件作为现代武器装备系统的“神经中枢”与“智能大脑”，其质量可靠性直接关乎作战效能、战场生存乃至国家安全。软件测试作为保障军用软件质量的核心环节，其过程管理的有效性至关重要。然而，当前军用软件测试过程仍面临需求模糊、测试体系割裂、技术手段滞后、人员能力不足等多重挑战，难以完全满足高可靠、高安全、强对抗环境下的严苛要求。本文深入剖析了军用软件的独特属性及其对测试质量管理提出的特殊挑战，系统梳理了当前测试过程中存在的主要问题，并在此基础上，从构建全生命周期一体化测试体系、深化基于风险的测试策略、推动智能化与自动化测试技术应用、强化测试数据治理与知识沉淀、以及打造专业化复合型测试人才队伍等五个维度，提出了系统性的质量管理改进路径。旨在为提升我国军用软件测试效能与质量保障能力提供理论参考与实践指导。

关键词：军用软件；软件测试；质量管理；全生命周期；智能化测试；测试体系

引言

在21世纪的军事变革浪潮中，软件定义战争已成为不争的事实。从精确制导武器的火控系统，到复杂战场环境下的指挥控制（C4ISR）网络，再到无人作战平台的自主决策算法，军用软件无处不在，其复杂度与重要性呈指数级增长。与民用软件不同，军用软件运行于极端、动态且充满敌意的环境中，任何微小的缺陷都可能被对手利用，导致灾难性后果，轻则任务失败，重则造成重大人员伤亡和战略被动。因此，“零缺陷”或“近零缺陷”是军用软件质量追求的终极目标。软件测试，作为发现软件缺陷、验证软件是否满足规定需求的关键活动，是实现这一目标不可或缺的“守门人”。然而，传统的、以事后验证为主的测试模式，在面对现代军用软件的高复杂度、高实时性、高安全性以及快速迭代的需求时，已显捉襟见肘。测试过程本身的质量管理——即如何规划、执行、监控和改进测试活动——成为制约军用软件整体质量提升的瓶颈。

1 军用软件的特殊属性及其对测试质量管理的挑战

1.1 极致的可靠性与安全性要求

军用软件失效的代价是无法承受的。例如，一枚导弹的制导软件若在关键时刻出现逻辑错误，可能导致误伤友军或平民；一个通信加密模块的漏洞，可能使整个作战网络暴露于敌方监听之下。这种“一次失败即全局崩溃”的特性，要求测试必须覆盖所有可能的边界条件、异常场景和故障注入情况，确保软件在任何可预见甚至部分不可预见的情况下都能保持稳定或安全降级。

1.2 复杂的运行环境与强对抗性

军用软件运行的物理和电磁环境极其恶劣，包括高温、低温、高湿、强振动、强电磁干扰等。更重要的是，它时刻处于“红蓝对抗”的博弈之中，必须能够抵御来自敌方的网络攻击、电子欺骗、数据篡改等恶意行为[1]。这要求测试不仅要验证功能正确性，更要进行高强度的鲁棒性测试、安全性测试（如渗透测试、模糊测试）和抗毁性测试，模拟真实战场上的极限压力。

1.3 高度的保密性与封闭性

出于国家安全考虑，军用软件的开发、测试全过程都必须在高度保密的环境中进行。这限制了开源工具和外部资源的使用，也使得测试数据的生成、管理和共享变得异常困难。同时，封闭的生态也阻碍了先进测试理念和技术的快速引入与融合。

1.4 快速迭代与持续交付的压力

现代战争节奏加快，作战需求瞬息万变，要求武器装备及其软件系统能够快速响应、敏捷升级。传统领域强调“一次成功、长期稳定”的开发模式形成张力。测试过程必须在保证质量的前提下，大幅缩短周期，支持DevSecOps等敏捷开发模式，这对测试的自动化、并行化和效率提出了极高要求。

2 当前军用软件测试过程中的主要问题分析

2.1 测试活动与研发过程脱节，“两张皮”现象严重

测试往往被视为研发流程末端的一个独立阶段，而非贯穿始终的质量保障活动。需求分析阶段缺乏测试人员的早期介入，导致需求本身可能存在模糊、不可测或矛盾之处，为后续测试埋下隐患。开发与测试团队沟通不畅，信息壁垒森严，使得测试用例设计滞后，难以有

效指导开发自测。

2.2 测试体系碎片化，缺乏统一标准与协同机制

在大型武器装备项目中，通常涉及多个分系统、多个承研单位。各单位可能采用不同的测试方法、工具和标准，导致测试资产（如用例、脚本、数据）无法共享复用，测试结果难以横向对比和集成验证。整个项目的测试视图是割裂的，无法形成对系统整体质量的全面、一致评估。

2.3 测试技术手段相对滞后，自动化与智能化水平不足

许多测试工作仍依赖大量手工执行，效率低下且易出错。对于复杂的嵌入式系统、分布式系统和AI驱动的智能软件，传统的黑盒/白盒测试方法显得力不从心。特别是在面对海量状态空间和非线性行为时，如何高效生成高覆盖度的测试用例，如何对AI模型的决策逻辑进行有效验证，仍是技术难题。智能化测试（如基于AI的测试用例生成、缺陷预测）的应用尚处于探索阶段。

2.4 测试数据管理薄弱，知识资产流失严重

测试数据是测试活动的核心燃料，但其生成、标注、存储、版本控制和安全防护往往缺乏规范管理。高质量的对抗性测试数据（如模拟的网络攻击流量、电子干扰信号）尤为稀缺。此外，测试过程中积累的经验、教训、典型缺陷模式等宝贵知识未能有效沉淀和结构化，随着人员流动而流失，导致同类错误反复发生。

2.5 专业化测试人才队伍建设滞后

军用软件测试不仅需要扎实的软件工程基础，还需具备深厚的军事领域知识、系统工程思维以及对特定硬件平台的理解[2]。然而，当前测试队伍中，兼具“懂软件、懂军事、懂系统”的复合型人才严重短缺。测试岗位的职业发展通道不清晰，也影响了高水平人才的吸引与保留。

3 军用软件测试过程质量管理的系统性改进路径

针对上述问题，必须采取系统工程的思维，从体系、策略、技术、数据和人才五个层面协同推进测试过程的质量管理改进。

3.1 构建全生命周期一体化测试体系

打破测试与研发的壁垒，将测试活动深度融入软件全生命周期（SDLC）。（1）左移（Shift-Left）： 在需求工程阶段，测试专家即应参与需求评审，运用可测试性设计（Design for Testability, DFT）原则，确保需求的明确性、完备性和可验证性。同步开展测试需求分析，初步规划测试策略和重点。（2）右移（Shift-Right）：将测试延伸至部署后阶段，通过在真实或近似真实的作战环境中收集运行数据（Telemetry），进行持续监控和反馈，用于改进后续版本的测试。建立从战场到实验室的闭环反馈机制。（3）一体化协同平台： 建立统一的、支持高安全等级的测试协同管理平台，集成需求管理、测试设计、用例管理、缺陷跟踪、自动化执行和报告生成等功能。确保所有干系人（开发、测试、系统工程师、用户代表）在同一平台上协同工作，实现信息无缝流转和过程透明化。

3.2 深化基于风险的测试（Risk-Based Testing, RBT）策略

面对资源有限的现实，必须将宝贵的测试资源精准投向风险最高的区域。（1）系统化风险评估： 建立科学的风险评估模型，综合考虑功能失效的严重性（Safety, Mission Criticality）、发生的可能性（Complexity, New Technology Adoption）以及探测难度等因素，对软件模块、功能点进行风险评级。（2）动态调整测试强度： 根据风险评级，动态分配测试资源。对高风险模块，采用更严格的测试方法（如形式化验证、模型检测）、更高的覆盖准则（如MC/DC）和更充分的回归测试。对低风险模块，则可适当简化测试，提高整体效率[3]。（3）量化决策支持： 利用历史项目数据和测试过程数据，不断校准风险评估模型，使其更加精准。通过量化的风险暴露度指标，为项目管理者提供清晰的决策依据。

3.3 推动智能化与自动化测试技术的深度融合

以技术创新驱动测试效能的革命性提升。（1）分层自动化测试架构： 构建涵盖单元测试、接口测试、系统测试和验收测试的多层次自动化测试金字塔。大力推广开发人员的单元测试和接口测试自动化，为上层测试奠定坚实基础。（2）AI赋能的智能测试：一是智能测试生成： 利用机器学习、符号执行、搜索算法等技术，自动分析代码或模型，生成能有效触发潜在缺陷的高价值测试用例，突破人工设计的局限。二是智能测试预言（Oracle）： 对于AI/ML驱动的软件，其输出往往是概率性的，传统“预期结果vs实际结果”的断言方式失效。需研究基于 metamorphic relations（蜕变关系）或统计学方法的智能预言机，以判断AI行为的合理性。三是缺陷智能分析与预测： 通过对历史代码变更、测试结果和缺陷数据的挖掘，构建预测模型，提前识别出代码中可能存在缺陷的“热点”区域，指导测试资源聚焦。（3）数字孪生与虚拟仿真测试： 为复杂的武器平台构建高保真的数字孪生体，可以在虚拟环境中进行大规模、高并发、高危险的测试，极大地降低实装测试的成本和风险，并能模拟现实中难以复现的极端场景。

3.4 强化测试数据治理与知识资产沉淀

将数据和知识视为核心战略资产进行管理。（1）构建测试数据湖/仓库：建立集中、安全、标准化的测试数据管理平台。对数据进行分类分级（如正常数据、异常数据、对抗数据），并实施严格的访问控制和生命周期管理。利用数据合成、变异等技术，低成本地扩充高质量测试数据集。（2）知识图谱驱动的知识管理：将分散的测试经验、缺陷模式、解决方案、最佳实践等非结构化知识，通过自然语言处理（NLP）等技术，构建成结构化的测试知识图谱。这不仅能方便查询和复用，还能为智能测试提供上下文支持，例如，当发现一个新缺陷时，系统能自动推荐历史上相似的案例和修复方案[4]。（3）建立组织级测试资产库：鼓励跨项目、跨团队的测试资产（通用测试用例、自动化脚本框架、测试工具插件）共享与复用，避免重复造轮子，提升组织整体的测试成熟度。

3.5 打造专业化、复合型的测试人才队伍

人才是所有改进措施落地的根本保障。（1）明确职业发展通道：设立清晰的技术专家序列（如初级测试工程师 -> 高级测试工程师 -> 测试架构师/测试科学家）和管理序列，让测试人员看到长远的职业前景。（2）构建多元化培养体系：开展定制化的培训，内容不仅包括前沿测试技术，还应涵盖军事理论、作战条令、特定武器系统原理等。鼓励测试人员参与实战化演习，加深对真实作战场景的理解。（3）营造质量文化：在组织内部倡导“质量是每个人的责任”的文化，打破“测试就是找茬”的旧观念。建立正向激励机制，表彰在质量保障中做出突出贡献的个人和团队。

4 实施保障与未来展望

上述改进路径的落地，离不开强有力的组织保障和持续的投入。首先，高层领导的支持与承诺至关重要。必须将测试能力建设提升到战略高度，在资源配置、政策制定上给予倾斜。其次，需要建立健全的标准规范体系，为测试活动的规范化、标准化提供依据。再次，应加强军民融合与国际合作，在确保安全的前提下，积极吸收借鉴民用领域（如金融科技、自动驾驶）和国际先进军队在软件测试方面的成功经验。展望未来，随着人工智能、量子计算等颠覆性技术的发展，军用软件的形态和复杂度将再次跃升。未来的测试将更加自主化、认知化和自适应。测试系统将不仅能执行预设的测试任务，更能像一个“智能陪练”一样，主动学习软件的行为模式，动态生成对抗策略，并根据测试反馈实时调整自身行为，以最高效的方式逼近软件的质量极限。这要求我们今天的改进工作必须具有前瞻性，为迎接下一代军用软件的质量挑战做好充分准备。

5 结语

军用软件的质量是国家安全的基石，而高效的测试过程是保障这一质量的生命线。面对新时代军事斗争的严峻挑战，我们必须摒弃陈旧的、被动的测试观念，以系统性思维推动测试过程质量管理的全面革新。通过构建全生命周期一体化的测试体系，实施精准的基于风险的测试策略，深度融合智能化与自动化技术，强化数据与知识资产管理，并最终打造出一支世界一流的军用软件测试专业队伍，我们才能真正筑牢军用软件的质量防线，为打赢未来智能化战争提供坚实可靠的软件支撑。这是一项复杂而艰巨的系统工程，但也是我们必须完成的时代使命。

参考文献

[1]朱翠云.基于过程分析的X军用软件测试质量管理改进研究[D].中国矿业大学,2024.

[2]徐豪.基于GJB5000B军用软件测试过程研究[J].中国战略新兴产业,2025,(17):175-177.

[3]杨赛,赵贺.GB/T 25000.51—2016在军用软件质量特性测试中的应用研究[J].电子质量,2024,(08):104-109.

[4]赵严,宋文爱,张日飞,等.基于平台支撑流程驱动的军用软件测试过程实施[J].火力与指挥控制,2021,46(08):155-161.