基于挑战响应的身份鉴别商用密码应用安全性评估实践

刘仁素[1]　张　湛[2]

（1. 重庆若可网络安全测评技术有限公司，重庆 401121；

2. 重庆电子科技职业大学，重庆 401331）

摘要：随着《中华人民共和国密码法》与《商用密码应用安全性评估管理办法》的相继颁布实施，我国从立法层面构建起商用密码合规应用与安全性评估的强制性规范框架。本研究针对身份鉴别领域的合规性测评需求，基于挑战－响应（Challenge-Response）认证机制，构建典型应用系统仿真场景，重点从应用安全与数据安全双维度对身份鉴别测评的实施路径展开系统性研究。通过梳理测评指标体系、验证流程及技术验证方法，本文旨在为商用密码应用安全性评估工作提供可复用的方法论参考，同时揭示当前身份鉴别机制在合规实践中的潜在风险与优化方向。

关键词：商用密码应用安全性评估；身份鉴别；挑战－响应

引言：近年来，随着《中华人民共和国密码法》《商用密码应用安全性评估管理办法》以及《信息安全技术 信息系统密码应用基本要求》（GB/T 39786-2021）等法规标准的相继实施，我国商用密码的合规化应用与安全性评估体系已进入法治化、规范化新阶段。学界对此展开了多维度探索：在行业应用层面，王伟忠等[1]构建了工业互联网商用密码应用体系框架。在安全性评估方法研究方面，杜嵘等[2]从标准符合性角度提出密码应用安全建设范式。

然而，现有研究多聚焦于密码技术横向部署与纵向建设，对身份鉴别这一核心安全机制的评估方法论缺乏深度探讨。本研究基于挑战－响应认证技术构建典型应用系统仿真环境，以《GB/T 39786-2021》为基准，从测评指标可量化、验证流程可回溯、风险场景可复现三个维度，系统解析身份鉴别测评的实施路径。通过揭示测评过程中暴露的协议脆弱性与合规性偏差，本文不仅为商用密码安全性评估提供动态验证框架，更为身份鉴别机制的优化升级提出可落地的技术策略。

1. 测评对象分析

1.1 保护对象及密码应用分析

本论文研究构建基于挑战－响应（Challenge-Response）机制的业务用户登录仿真场景，其密码应用方案严格遵循《GB/T 39786-2021 信息安全技术 信息系统密码应用基本要求》中应用与数据安全层的身份鉴别指标要求。核心保护对象为应用系统业务用户，密码安全需求聚焦于用户身份真实性保障。

技术实现框架如下：

1）认证协议设计：

用户终端通过集成智能密码钥匙（USB Key）及个人数字证书，对挑战值执行签名运算，

生成数字签名响应值；应用系统端调用签名验签服务器完成证书链验证与签名解密，实现双向身份鉴别。

2）密码产品配置：

签名验签服务器：部署 SM2/SM3 算法，承担证书验证与签名解密功能。

智能密码钥匙：作为用户私钥安全载体，实现签名运算的物理隔离保护。

3）密钥全生命周期管理：

非对称密钥对：

用户公钥：在智能密码钥匙中产生，采用证书形式分发存储于签名验签服务器，支持验签调用。

用户私钥：在智能密码钥匙内生成、使用及销毁，杜绝导出风险。

1.2 身份鉴别工作流程建模

本研究通过流程图与网络拓扑双重视角解析身份鉴别协议的交互逻辑，其具体工作流程如图 1 所示。

核心交互时序如下：

1）挑战生成阶段：

用户通过 HTTP 访问应用系统时，应用系统向签名验签服务器发起认证请求，生成符合商用密码标准要求的安全随机挑战值（Nonce），并通过安全通道传输至用户终端。

2）响应生成阶段：

用户终端调用智能密码钥匙，使用私钥对挑战值执行 SM2 签名运算，生成响应值（含签名数据及数字证书）在回传至应用系统。

3）验证决策阶段：

应用系统将响应值提交至签名验签服务器。

服务器端通过 PKI 体系验证证书有效性，并基于公钥解密验证签名合法性。

验证结果（成功 / 失败）返回至应用系统，触发相应的访问控制策略。

安全边界设计如下：

1）网络隔离

签名验签服务器部署于服务器区，终端访问应用系统通过 http 协议进行访问，在网络边界处部署 SSL VPN 和防火墙进行隔离。

协议合规性：全流程符合 GM/T 0036《SM2 密码算法使用规范》与 GM/T 0015《基于 SM2 算法的数字证书格式规范》。

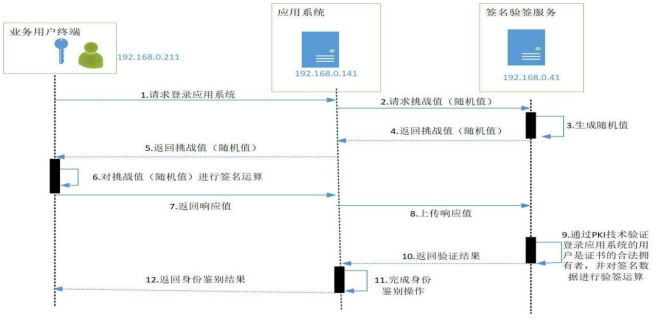

图 1　工作流程图

2. 测评过程及分析

2.1 测评对象及指标映射

基于 GB/T 39786-2021 第 8.4(a) 条款要求，本场景的测评对象为应用系统业务用户，核心测评指标为：“应基于密码技术实

现用户身份鉴别，确保用户身份真实性”。其合规性要求可分解的技术验证维度为测评对象：应用系统业务用户身份鉴别；密码技术应用：SM2/SM3 挑战－响应协议；验证目标：身份真实性、抗重放攻击。

2.2 测评过程与合规性分析

环境和工具准备：

工具：采用 WireShark 4.0.3 协议分析工具，设置过滤条件为 ip.addr == [应用服务器 IP]，抓取终端－应用服务器－签名验签服务器间的全链路通信数据包（持续时长近 10 分钟）。

环境基线：建立无干扰环境基准流量模型，确保数据可比性。

在测评过程中，抓包是商用密码应用安全性评估（以下简称：密评）实践活动中的必要证据获取手段。在本次实践中，采用了 WireShark 工具进行抓包分析。本次抓包设置了应用服务器地址为观察地址，抓取应用服务器与终端、签名验签服务器三者的数据包。

通过对抓取的数据包进行分析，按照梳理出的工作流程对数据包进行分析，以下为分析过程。

1）业务用户终端访问应用系统，终端发送证书到应用服务器并请求随机值，在数据包中找到 POST/getRandomZero，找到 cert；

2）在应用服务器返回给终端的数据包中找到返回的 randnum 值；

3）业务用户前端收到随机值后，对随机值进行签名，并把随机值、签名值、证书信息发给应用服务器，在终端发给应用服务器的 POST/checkcertsing 包中找到对应证书值和签名值；

4）应用服务器将用户证书、随机数、签名值发送到签名验签服务器，找到对应的 DATA 数据；

5）对签名值转换为 16 进制，得出业务应用终端签名值与应用系统传送给签名验签服务器的签名值一致；

6）签名验签服务器验签通过后，把验证结果返回给业务应用终端，完成身份鉴别；

7）从数字证书中提取的公钥（048ef5e13c7210a4957bc632427f4c1a29a84d2b1b46689848 fd2fffff49af49a4193d660b2cef7b80995d46681f333458a185ef3342a8274cb45387ef710f6ab0）、签名值（MEQCIHLYGcWkFgC6JXgdSudIZl8o0lNYEUtSM4ZYhoxsCJXYAiBAyk2wQRFp5ZfFiEZFW6MmwdUQNJ+1TDrncZ3iW85Wqw==）、随机数（AAAAAAAAAAAAAAAAAA==）进行验签，验签通过。

8）查看签名验签服务器和智能密码钥匙的商用密码产品认证证书与实际产品一致。

9）查看密码设备相应配置，确认采用的算法与密码应用分析中的算法一致，数字证书算法为 SM2 和 SM3。

10）查看源代码，发现随机值不是由签名验签服务器发起的随机数，随机数为固定值。

2.3 问题分析与改进建议

经过对数据包、源代码和产品认证证书的分析，得出以下结论：该应用系统调用了签名验签服务器进行验签，密码技术使用有效性 D 符合。通过对密码算法进行分析，采用国密算法 SM2 和 SM3，密码算法 / 技术合规性 A 符合。通过对密码产品认证证书的分析，智能密码钥匙和签名验签服务器秘钥管理安全 K 符合。但是在数据分析过程中，发现在随机数生成上存在问题，该随机数为固定值，随机数产生机制存在安全问题。根据《信息系统密码应用高风险判定指引》得出，该随机数产生机制存在高风险。

通过以上分析提出高风险项和对应改进建议如下：

1）高风险项

·随机数熵值不足：建议集成符合《GM/T 0005-2021》的真随机数发生器（TRNG）。

·传输层保护缺失：需强制启用 TLS 1.3 协议，配置 SM2/SM4 国密套件。

3. 测评实践中的方法论困境与优化路径

密码技术应用的强专业性与业务耦合性，导致密评人员在技术验证、风险研判等环节面临多重挑战。尽管现行标准已解决部分共性问题，但在测评方法学适配性与动态风险评估层面仍存在显著短板。本研究结合案例实证分析，提炼了实践中的三类典型问题并提出结构化改进框架。

1）数据捕获工具部署的实践瓶颈与拓扑适配策略

问题描述：在身份鉴别协议验证中，抓包工具的快速部署直接影响测评效率与数据完整性。现行标准对网络拓扑感知与接入点选择缺乏操作性指引，导致密评人员需消耗大量时间成本进行环境适配。

实证分析：本研究案例采用双模采集策略，仍然花费了大量的部署和抓包时间。

优化建议：

·拓扑感知工具开发：基于《GM/T 0086-2020 网络安全等级保护测评指南》，研发自动化网络拓扑测绘插件，实现测评对象与抓包节点的智能映射。

·轻量化部署协议：制定镜像端口配置模板，通过相关协议自动下发交换设备的镜像策略，减少人工干预。

2）密码技术有效性的多源证据链构建

问题描述：单一数据源（如抓包结果）难以全面验证密码技术实现的合规性与安全性，需构建跨维度的证据交叉验证体系。

实证分析：在本研究的随机数固定值漏洞发现中，通过以下证据链完成技术归因：

·日志溯源：验签服务器日志未记录随机数生成事件。

·代码审计：应用系统源码中随机数函数硬编码固定种子值。

·抓包验证：连续 10 次交互中挑战值检测结果恒定。

优化建议：建立四维验证矩阵（技术实现－协议交互－设备配置－日志审计），并基于标准设计熵值测试用例库可提高准确性。

3）风险叠加效应与动态评估模型缺失

问题描述：现有风险评估多采用孤立漏洞评级法，忽视中低风险项在特定业务场景下的叠加放大效应。

实证分析：以本案例为例，风险叠加路径表现为固定随机数（高风险）+ 明文传输（高风险）导致身份伪造（重大隐患）。

优化建议：引入对应评分体系，尝试通过环境度量量化业务场景权重。构建攻击树模型，比如基于 CNNVD 公布或者 OWASP Top 10 漏洞清单，模拟多漏洞协同攻击风险。

基于以上，本研究认为后续研究中，密评工作的评估工具链集成开发，比如开发支持国密算法的一体化测评平台，集成抓包分析、代码扫描、日志聚合等功能，且逐步进行知识图谱构建，即基于历年密评案例构建风险模式图谱，实现问题智能诊断与修复建议推送可以极大提高密评工作效率和工作准确性。

参考文献：

[1] 王伟忠，闫瑞泽，查奇文，等．工业互联网商用密码应用体系研究 [J]. 信息安全研究 ,2024,10(6):519-525.

[2] 杜嵘，陈洁．基于商用密码应用类标准的密码应用安全建设 [J]. 无线互联科技 ,2024,21(6):9-15.

作者简介：刘仁素（1993 － ），女，重庆，本科，网络安全等级保护，重庆若可网络安全测评技术有限公司（401121），

张湛（1974 － ），男，重庆，研究生（博士），控制科学与工程，重庆市沙坪坝区陈家桥镇。