人工智能硬件安全白皮书

DUEROS × 百度安全

联合出品

目录

前 言.....3

1．范围.....4

2．术语定义与缩略语.....4

3．概述.....5

4．总体安全框架.....6

5．DuerOS 生态安全规范.....7

6．DuerOS 生态的安全问题.....16

前 言

为适应人工智能硬件快速发展的安全需要，由百度安全事业部AI安全团队组织制定“人工智能硬件生态安全白皮书 “，推荐相关开发者参考采用。有关对此白皮书的建议和意见，向百度安全事业部AI安全团队反映。

此白皮书将持续进行迭代，请及时关注白皮书更新信息。

此白皮书为精简版本，完整版本将面向百度合作伙伴开放。

此白皮书由百度安全事业部AI安全团队提出并归口。

人工智能硬件安全白皮书

1. 范围

本标准分析了DuerOS生态存在的安全威胁，并以此为基础提出了DuerOS应用设备层、网络传输层、数据处理层以及数据安全和隐私层等方面的系统安全需求。

本标准适用于DuerOS系统安全技术领域。

2. 术语定义与缩略语

2.1 术语与定义

2.1.1 应用设备层

应用设备层主要由各种RFID标签、有线或无线传感器、智能硬件设备组成，本文也包括运行于其上的操作系统；

2.1.2 网络传输层

通过各种电信/传感网络与互联网融合，与应用设备进行实时通信；

2.1.3 数据处理层

对所收集信息进行处理，实现智能化识别、定位、跟踪、监控、管理等实际应用；

2.2 缩略语

下列缩略语适用于本文件。

DoS 拒绝服务 Denial of Service

DDoS 分布式拒绝服务 Distributed Denial of Service

IoT 物联网 Internet of Things

MITM 中间人 Man In The Middle

PKI 公钥基础设施 Public Key Infrastructure

RFID 无线射频识别 Radio Frequency IDentification

SIM 客户识别模块 Subscriber Identity Module

UICC 通用集成电路卡 Universal Integrated Circuit Card

VPN 虚拟专用网络 Virtual Private Network

KARMA 自适应系统热修复 Kinetic Adaptive Repair for Many AI-systems

CSR 证书注册请求 Certificate Signing Requests

3. 概述

DuerOS是百度度秘事业部研发的对话式人工智能系统，为厂商全新打造的分层解决方案。能够满足各个类型厂商不同层次的需求。

本标准按照DuerOS的网络层次划分方式，以设备应用层、网络传输层、数据处理层，数据安全和隐私层为出发点，制定DuerOS生态所面临的安全威胁和制定DuerOS生态安全标准需求，主要包括：从DuerOS的DuerOS终端、感知延伸网络、接入、核心网络、

应用层、控制管理以及DuerOS层间方面等不同的角度分析DuerOS面临的安全威胁和存在的安全隐患，并在此基础上，提出DuerOS设备应用层、网络传输层、数据处理层以及相关业务应用的安全需求。

4. 总体安全框架

根据通用分层模型，DuerOS在逻辑功能上划分为四层，即：应用设备层、网络传输层、数据处理层和数据安全和隐私层。

基于通用分层模型, DuerOS生态安全需求主要划分为：应用设备层的硬件的安全需

求、软件安全需求；网络层的网络层安全需求；数据处理层的安全需求；数据安全和隐私层的安全需求，分别包括：

1. 应用设备层安全需求，包括末端节点的安全需求，包括硬件及固件所带来的安全漏洞以及 DuerOS 基于操作系统带来的传统安全问题；
2. 网络传输层安全需求，包括数据传输的信息泄露和网络攻击等；
3. 数据处理层安全需求，包括云端安全问题以及移动控制端、web 控制端的安全问题；
4. 数据安全和隐私层安全需求，包括敏感数据及凭证信息的安全问题；

5. DuerOS 生态安全规范

针对可能遇到的攻击，制定对应标准如下：

5.1 应用设备层安全规范

5.1.1 硬件安全规范

5.1.1.1 物理接口相关规范

级别：推荐执行

说明：包含显示端口与隐式端口规范方案，降低硬件层面的安全风险；

5.1.1.2 硬件芯片安全规范

级别：推荐执行

说明：采用符合安全标准的芯片，降低硬件层面安全风险；

5.1.1.3 开发板规范

级别：推荐执行

说明：针对开发板布局与引脚进行规范，增加硬件破解的难度；

5.1.2 软件安全规范

5.1.2.1 软件通用安全规范

a） 登录密码规范

级别：强制执行

说明：提升密码强度，增加暴力破解的难度；

b） 加密强度规范

级别：强制执行

说明：按照目前的标准，在 2020 年前，AES128 级别加密等级是安全的；目前推荐的 RSA 加密位数为 2048，ECC 加密位数为 163；

c） 更新功能与过程规范

级别：强制执行

说明：当系统、固件出现重大漏洞，可及时进行修复；防止安装包被伪造篡改；

d） 系统安全补丁规范

级别：强制执行

说明：更新官方补丁，关注 CVE 等漏洞平台，降低漏洞对系统安全的影响；

e） 系统版本规范

级别：推荐执行

说明：产品外发时，确保两年及以上的系统安全支持时间；

f） 固件编译规范

级别：推荐执行

说明：固件编译时去除符号表信息，增加调试及破解难度；

g） 登录密码补充规范

级别：推荐执行

说明：推荐硬件 ID 绑定方案，安全性优于口令登录；

h） 系统第三方应用规范

级别：推荐执行

说明：防止通过系统第三方应用漏洞攻击系统；

j) 开源库及第三方组件规范

级别：强制执行

说明：防止通过开源库及第三方组件漏洞攻击系统；

5.1.2.2 应用层安全规范

a） Android 系统应用层安全规范

级别：强制执行

说明：包含系统配置、系统应用相关安全要求；

b）Linux 系统应用层安全规范

级别：推荐执行

说明：包括服务、用户及用户组相关安全规范；

5.1.2.3 内核层安全规范

a） Linux 内核安全规范

i. 热补丁修复支持

级别：推荐执行

说明：支持热补丁修复的系统能快速对新增漏洞进行修复，将漏洞对系统的影响降至最小，推荐使用百度安全的相关热补丁修复技术；

ii. Linux 安全漏洞修复规范

级别：强制执行

说明：针对漏洞及相关安全问题有紧急应对方案，降低内核漏洞对系统安全的影响；

5.2 网络传输层安全规范

5.2.1 网络协议选择规范

级别：强制执行

说明：针对通信方案，选择主流的协议，确保安全性；

5.2.2 网络通信规范

级别：强制执行

说明：包含 TLS 私钥与证书规范、TLS 配置规范及应用设计规范；

5.2.3 网络通信补充规范

级别：推荐执行

说明：选择适当的加密算法，保护密钥，防止中间人攻击；

5.2.4 Wi-Fi 规范

级别：推荐执行

说明：针对伪热点等钓鱼攻击进行防范；

5.2.5 运营商网络规范

级别：推荐执行

说明：防止通过运营商网络及相关技术漏洞来发动攻击；

5.3 数据处理层安全规范

5.3.1 云端安全规范

5.3.1.1 云端服务器规范

级别：推荐执行

说明：降低渗透攻击，DDoS 攻击对业务的影响；

5.3.1.2 云端服务器需对用户的网络请求处理规范：

级别：强制执行

说明：需进行身份认证及授权验证；对用户请求数据进行有效过滤；防止非法访问，SQL 攻击，XSS 攻击，CSRF 等网络攻击；

5.3.1.3 事件与响应规范

级别：强制执行

说明：包含事件响应准备、事件检测与分析、事件遏制、事件根除与恢复、事件响应时间及事件报告与持续跟踪等相关规范；

5.3.1.4 身份授权和管理规范

级别：强制执行

说明：针对场景和约束条件，采用正确的协议；

5.3.1.5 权限隔离规范

级别：强制执行

说明：包含管理平面、虚拟化与容器权限隔离及应用程序权限隔离相关规范；

5.3.1.6 安全服务规范

级别：推荐执行

说明：选择安全服务提供商，加强安全防护能力；

5.3.1.7 云服务安全检测补充

级别：推荐执行

说明：通过白盒与黑盒多方面检测，推荐使用百度安全事业部的检测工具；

5.3.2 移动控制终端安全规范

5.3.2.1 通用规范

a) 外发版本规范

级别：强制执行

说明：外发版本相关配置修改，防止软件信息泄露；

b) 安全规范与响应机制

级别：推荐执行

说明：针对新漏洞有相关安全处理方案，有效降低漏洞对系统的影响；

c) 反破解，反劫持能力规范

级别：推荐执行

说明：推荐使用百度提供的应用加固能力，防止攻击者通过逆向，调试，劫持等手段对应用进行攻击；

a） 证书验证

级别：强制执行

说明：移动设备面临网络窃听、网络劫持等中间人攻击行为，因此对于敏感信息需要加密传输，并且对接收到的重要数据也需要进行完整性校验。

5.3.2.2 Android 控制终端安全规范

a) 软件更新规范

级别：强制执行

说明：防止安装包遭篡改或被伪造，除应用商店更新外，推荐使用百度安全事业部提供的安全更新方案；

b) 模块加载规范

级别：强制执行

说明：防止加载执行漏洞，推荐使用百度安全事业部提供的模块安全检测方案；

c) 系统配置规范

级别：强制执行

说明：包含组件服务、端口、进程权限、文件权限等相关规范；

d) WebView 使用规范

级别：强制执行

定义：针对用户输入数据进行充分校验；

说明：推荐使用百度安全事业部提供的安全 WebView 方案；

5.3.2.3 iOS 控制终端安全规范

a) UIWebView 使用规范

级别：强制执行

说明：基于传统 Web 安全做延伸，针对用户输入数据进行充分校验；

b) URLScheme 安全规范

级别：强制执行；

说明：确定应用注册的所有 URLScheme，针对 URLScheme 的实现逻辑进行审计，确保对输入的数据充分校验；

5.3.3 Web 控制端安全规范

5.3.3.1 网络的请求处理规范

级别：强制执行

说明：防止非法访问，SQL 攻击，XSS 攻击，CSRF 等网络攻击；

5.3.3.2 路径参数处理规范

级别：强制执行

说明：防止攻击者能够访问未授权的目录；

5.4 数据安全与隐私层安全规范

5.4.1 数据安全

5.4.1.1 敏感信息存储规范

级别：强制执行

说明：降低隐私泄露的风险；

5.4.1.2 身份凭证存储规范

级别：强制执行

说明：防止因密钥，身份凭证泄露导致的非法访问；

5.4.1.3 敏感数据权限管理规范

级别：强制执行

说明：防止隐私信息泄露；

5.4.1.4 信息加密与传输规范

级别：强制执行

说明：防止通过明文传输获得信息或进行 XSS 攻击；

5.4.1.5 云端数据安全规范

级别：强制执行

说明：数据安全保护本身相关的控制，加密是最重要的控制手段，也需要其他的方案配合。

6. DuerOS 生态的安全问题

6.1 应用设备层安全问题

6.1.1 硬件安全问题

通过与承载 DuerOS 的物理设备进行物理接触，包括且不限于接口、存储芯片，从而获取控制权、系统信息与数据，具有严重的安全风险。本安全准入标准从物理层面进行防范，尽量从底层开始降低安全风险。

6.1.2 软件安全问题

基于 DuerOS 所运行的操作系统，无论是 Linux、Android 或嵌入式操作系统如 FreeRTOS 等，本身存在漏洞等安全问题。在其上运行 DuerOS 时，需要保证底层操作系统的安全性，从而保证 DuerOS 的权限与安全性。本安全准入标准主要从系统漏洞、安全更新等方面进行防护，及时更新设备软件，降低被攻击的风险。

6.2 网络传输层安全问题

DuerOS 体系架构中，涉及到多种体系设备间的连接与配对，其中包含不同的连接协议，如蓝牙、ZigBee 和 WI-FI 等，使用每种协议通信时，均需要注意数据传输中的安全性问题等，落实到具体协议时，需要注意数据加密与解密安全问题，其中在互联网上进行数据传输时，更加要关注数据安全传输的相关配置与校验。

6.3 数据处理层安全问题

数据处理与应用层包含云端服务器与控制终端两个层面。其中云端服务器负责存储、传输大量用户数据与信息，针对服务器安全，除了针对权限、授权、认证等方面的严格规定外，必须有完整的处理攻击时的备案；控制终端中，主要考虑运行于 iOS 与 Android 操作系统上的的应用安全，防止出现信息泄露等安全问题。

6.4 数据安全与隐私层安全问题

除了传统安全外，由于搭载 DuerOS 设备可能涉及摄像头、语音输入等相关隐私数据，所以在信息存储与加密上，必须保证权限隔离与足够的加密强度，才能保证隐私数据的安全。