附件 1：

第十九届全国大学生信息安全竞赛（作品赛）暨第三届“长城杯”网数智安全大赛（作品赛）命题挑战赛赛题

一、面向大模型及其应用的安全性研究

（一）研究方向说明

本题要求从红队视角研究当前热门的大语言模型及智能化应用的典型攻击面（如提示注入、模型越狱、训练数据泄露、滥用风险、工具调用劫持、记忆中毒、环境感知污染等），并设计一套可嵌入或旁路的行为监督机制，对智能化应用的工具调用、代码执行、文件访问进行实时审计与异常判定。最后构造对抗性输入，并设计可落地的防御策略（如输入输出过滤、上下文隔离、模型行为监测等）。

（二）预期成果形态

1. 安全风险分析报告：至少 3 类攻击场景，每类场景要求包括模型对抗样本与越狱测试用例集、智能体攻击脚本。

2. 行为监督原型系统：拦截智能体集群与外部工具的交互，基于安全策略（如允许、拒绝、询问）或异常检测模型进行监控。要求提供一个开源智能化应用（如 OpenClaw）、模拟业务工具（发送邮件、读写文件、调用 API）、模型调用链

路的安全监控插件、基座模型检测或过滤原型、监督端实时展示告警或阻断记录。

二、低空无人机网络安全检测与防护

（一）研究方向说明

本题要求研究无人机及其地面站之间的飞控链路、遥测数据、固件更新接口存在的劫持、篡改、GPS 欺骗等风险，设计针对无人机通信协议(如 MAVLink 等)的异常检测方法，或对地面站控制指令的防护策略。

（二）预期成果形态

1. 无人机攻击复现脚本（仿真环境或真机）。

2. 链路异常检测模块（流量或指令级）。

3. 地面站防护模块：收到恶意指令后的阻断或告警。

三、面向卫星互联网领域的安全检测技术

（一）研究方向说明

本题要求研究卫星互联网（如 Starlink 式星座等）存在的链路劫持、欺骗、协议安全、终端认证等问题和应对方法，提出适用于受限计算环境的加密与检测机制，可在仿真平台中研究卫星的攻击面。

（二）预期成果形态

1. 仿真环境中的检测模块（GNS3 或 OpenSAND）。

2. 完整演示:攻击注入(如非法终端接入、下行链路劫持、下行链路加密或编码方式破解等），检测告警和缓解建议。

四、面向数据要素流转的隐私集合协同计算增强技术

（一）研究方向说明

本题要求在不暴露各方原始数据情况下，利用隐私计算技术，实现集合匹配、求交、求并等安全操作，形成计算开销与安全分析报告，为可信数据流通与协同应用提供技术支撑。

（二）预期成果形态

1.面向隐私集合运算的原型系统(命令行或简单 Web 界面)。

2.针对结构化、半结构化数据的计算开销与安全分析报告。

五、面向海量多模态数据的实时加密与隐私保护技术

（一）研究方向说明

本题要求研究终端算力受限场景（如高清摄像头）下，提出多模态数据（如声音、图片、视频等）的实时数据加密与隐私保护技术（如隐水印、数字签名等）。

（二）预期成果形态

1.原型系统（采集终端、处理电路、接收端），系统要求延时增加不超过 10%，流量不低于 1GB/s。

2.演示要求：系统运行总时长不少于 30 秒，运行期间系统满足时延及流量要求。

六、基于海量日志数据的网络风险识别技术

（一）研究方向说明

本题要求采用无监督或弱监督方法,从TB级安全日志(如Syslog、ETW、WAF、DNS 解析日志等）中提取罕见攻击链或异常模式，并有效降低人工研判成本。

（二）预期成果形态（以下成果任选其一）

1.日志风险识别分析引擎原型系统，可支持常见日志格式，支持典型攻击场景的识别召回率与误报分析，展示风险聚合与溯源线索。

2.恶意域名检测引擎原型系统，可支持轻量推断服务（HTTP API），支持实时域名判定（白、黑、灰），提供真实或公开 DNS 日志验证报告。

七、面向加密通信协议的恶意行为检测技术

（一）研究方向说明

本题要求提出在不解密加密通信（如 SSL、SSH 等）会话或有限解密条件下的异常行为检测技术，能够识别异常命令序列、异常隧道、暴力破解后操作等。

（二）预期成果形态

1.加密通信流量分析工具（元数据、时序行为）。

2.异常行为检测规则或模型。

3.实验环境：正常流量与攻击流量的检测对比。

八、基于大语言模型的应用安全审计技术

（一）研究方向说明（两个方向任选其一）

1.本题要求使用大语言模型分析 API 请求序列、参数分布与访问模式，识别越权、参数遍历、接口滥用等行为，并给出可解释的安全告警。

2.本题要求通过静态分析、动态分析或混合分析手段，使用大语言模型发现源代码中的安全漏洞，重点关注在第三方库、供应链的场景下。鼓励研究传统程序分析方法（如污点追踪、符号执行优化等）与大语言模型的有效结合。

（二）预期成果形态（与研究方向一致）

1.API 安全分析工具（代理或日志分析模式）：提供正常调用与异常调用的评测结果，完整演示通过工具发起的调用过程。

2.代码审计原型：支持 Java、Python、C 等多种语言，提供测试用例集（包含已知漏洞的代码、不存在已知漏洞的代码），提供漏洞检出率、误报率对比分析。

九、APT 供应链攻击检测与溯源系统

（一）研究方向说明

本题要求构建一个面向软件供应链的检测与溯源系统，支持识别组件版本异常、构建环境篡改、依赖混淆等攻击手段，并尝试还原攻击路径。

（二）预期成果形态

1.供应链攻击检测原型：支持如 npm、pypi、rubyGen 等软件包管理工具，完成软件成分及行为特征分析，提供溯源图谱展示（如被污染环节、受影响资产情况等）。

2.针对不少于 2 个典型 APT 供应链攻击案例复现验证。