课题类型：□ 重大　□ 重点　■ 一般
课题周期：■ 一年　□ 两年　□ 三年

拟解决的关键问题：
研究背景：
明年生效的《保护关键基础设施（电脑系统）条例》将激发香港本地关键基础设施提供者的网安需求，但目前从业人员不足，现拟研发安全领域相关算法，紧扣[AI+]主题，以提升香港移动网信安全服务产品及效能、品牌科技实力
拟解决的关键问题：
（1） 安全运营中 SEIM/HIDS 日志误报排查需要联合多维度主机日志等，人工排查耗时长
（2） 拟解决渗透测试中的信息搜集耗时长、搜集后严重漏洞脆弱性流程较复杂，需要大量的人力匹配资产组件和漏洞对应关系

总体研究内容与目标：
- 基于安全运营测试中产生的多模态数据和检索增强技术，研究一种日志智能检索与推理模型算法
- 构建渗透测试多模态知识库，拟研发一套基于大型语言模型（LLM）的自动化渗透测试算法

预期成效和价值：
（1）3 项技术创新专利（包括一种检索增强技术方法、一种模型微调方法以及一种测试方法）；
（2）1 种基于多模态数据的自动化渗透大模型原型系统；
（3）1 种基于检索增强的安全运营优化原型系统；
（4）2 篇相关高水平论文。

成果转化目标及预期转化方式：
1. 完成基于 AI 的自动化渗透测试系统。
2. 与研究机构共享算法与源码，后续共同推动相关成果落地。

2025 年 7 月——2025 年 12 月 年度课题指南
年度人员投入：高校方 16 人月；移动方 8 人月
阶段研究要点：
1. 完成多模态安全数据采集与向量化编码原型。
2. 制定大模型＋RAG 总体技术方案，并提交 1 份专利交底书。
3. 启动第一版算法 / 模型微调，完成核心代码框架。

2025 年 7 月——2026 年 12 月 年度课题指南
年度人员投入：高校方 15 人月；移动方 4 人月
阶段研究要点
1. 完善大模型与 RAG 融合能力，提升告警识别与渗透推演性能。
2. 优化推理效率，验证模型在典型安全任务下的实用性。
3. 撰写高水平论文，持续推进专利公开与落地。

年度成果形式及考核指标：

序号	成果名称	交付时间	关键技术/成效指标	知识产权（共有/独有）	标志性成果
1	项目可行研究和技术方案	2025.09	完成《可行性研究报告》和《项目技术方案》	共有	《可行性研究报告》/《项目技术方案》
2	自动化渗透测试输出报告；RAG 知识库构建	2025.11	渗透测试投入人力时间缩短 30%；在特定情景下 RAG 检索的准确率达到 60%	共有	提交相关专利申请
3	RAG 检索策略优化安全运营的相关专利；大模型微调	2025.11	获得微调后的算法模型	共有	提交相关专利申请
4	自动化渗透测试的相关论文和 RAG 检索策略优化安全运营的相关论文	2026.05	高水平学术会议包括不局限于 CCF	共有	撰写高水平学术论文两篇
5	自动化渗透测试输出报告；RAG 方法输出	2026.06	自动化程度达 70%，其中从信息收集到脚本调用实现完全自动化；渗透测试投入人力时间节省 60%；渗透成功率（标准靶场环境下）提升 10%；在特定情景下的准确率达到 70%;基于RAG的模型部署时间减少 50%；基于 RAG 的模型误报率降低 ，**安全报警误报率从 50%降至 20%（数据来源：CMHK 安全告警工单处理统计）**。	独有	模型、检索方法