智能网络与内生安全实验室

（ 吴春明教授 wuchunming@zju.edu.cn ）

实验室主要从事互联网体系架构与演进、软件定义网络、可编程网络、程序分析与漏洞发现、网络空间内生安全、AI 大模型技术领域的研究与应用实践。近年来承担、参与国家重点研发计划包括宽带通信和新型网络、网络空间安全、多模态网络与通信等专项课题，工信部工业互联网创新工程及省级重点研发计划项目二十余项。授权和申请国家发明专利九十余件，发表学术论文百余篇，其中 CCF-A 四十余篇，高影响因子期刊二十余篇。研究成果在安全顶会 ACM CCS 2021 和 IEEE/ACM IWQoS 2021 上荣获 Best Paper Awards；在 IEEE INFOCOM 2021 上荣获 Best Paper Candidates; 2022 年荣获了全球 Pwnie Awards 最具创新研究奖提名；在全球软件工程领域顶会 ISSTA 2025 上荣获 Distinguished Paper Award；多项安全成果入选世界黑帽大会议题项，发现云上系统漏洞 56 项，获得包括 Redhat、Oracle 在内的全球知名厂商 CVE 致谢 22 项。曾参与获得国家科技进步二等奖，中国自动化学会科技进步一、二等奖，中国电力科技创新一等奖，中国通信学会科学技术二等奖，公安部颁发的网络安全保障工作突出贡献一等奖章，浙江省自动化学会高等教育教学成果二等奖，以及中国网络安全产业联盟网络安全优秀创新成果大赛二等奖、未来网络科技创新大赛全国二等奖。实验室成立的摇光安全战队 FSL 多次参加了国内外网络安全赛事，队员曾荣获包括 XCTF、DEFCON CTF、DataCon 等大赛的冠亚军战绩。实验室培养的研究生荣获多项校级及国家级荣誉与项目，包括竺可桢奖学金、首批国家自然科学基金博士生项目、中国科协青年人才托举工程博士生专项、中国电子学会优秀博士学位论文等。

研究动态与成果简介

1. 网络安全技术

实验室提出了基于语义感知的虚拟机管理程序的 Fuzzing 框架 V-SHUTTLE，通过解耦化嵌套结构并启用类型感知来实现全自动化的 Fuzzing。V-SHUTTLE 被发表在网络系统安全领域四大顶会之一的 ACM CCS 2021 上，并获得 Best Paper Award，当年是中国研究团队第二次以第一作者身份获得安全四大会议的 Best Paper 奖项。 V-SHUTTLE 迄今累计发现了 QEMU 和 VirtualBox 两款虚拟机管理程序中百余个未知漏洞，多个被授予了 CVE 编号，得到了厂商的公开致谢。其核心模块已在蚂蚁集团进行了实际部署与应用，成果的最大贡献在于有效保障了绝大多数互联网云厂商基础支撑软件的安全。

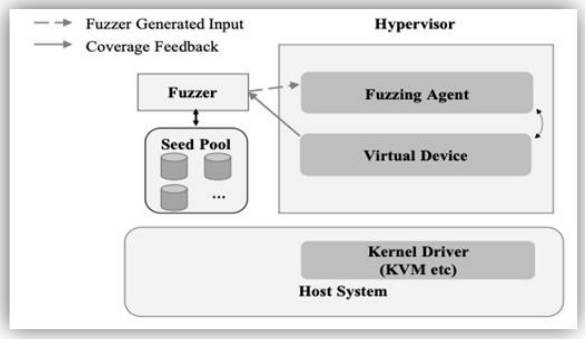

图 1 V-SHUTTLE 的工作流程

图 2 ACM CCS 2021 最佳论文奖

进而深入研究虚拟机管理程序中的异步时钟管理模块，在虚拟机管理程序中有许多设备使用异步时钟来处理它们的任务，例如网络、USB、磁盘和加解密设备等等，其目的是避免调用线程的阻塞，从而提高软件的响应能力。通过研究发现，攻击者可以利用异步时钟进行一些竞争条件攻击，并益于漏洞的利用。基于此，实验室提出了一种新的攻击利用技巧 Timekiller，利用异步时钟机制，使得原本难以利用的单个堆溢出写漏洞变得可利用。结合堆溢出写漏洞和 Timekiller 攻击技巧，成功实现一套针对 QEMU/KVM 的虚拟机逃逸方法。Timekiller 入围了 HITBSecConf2023-Phuket 黑客大会。

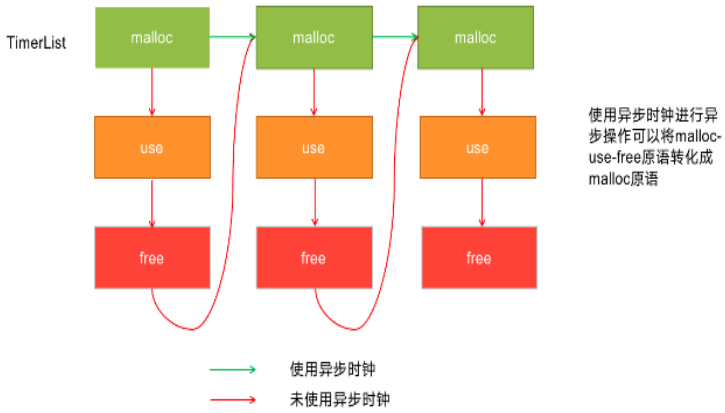

图 3 Timekiller 核心思想

2. 可编程网络测量技术

网络智能化成为未来网络的一大发展趋势，该趋势将面临优化互联网路由控制技术，解决传统网络控制调度的弊端，降低管理成本的同时提升用户体验。如何利用有限的网络资源快速部署测量系统，迅捷精准地捕捉网络流量的变化、获得高精度的测量结果对于智能化网络的研究与发展意义重大。实验室提出的 LightNF 框架在可编程网络中提供了自动化的网络测量任务部署方案：首先提供了一套基于高层次用户意图的编程原语，允许网络管理员使用这套原语来构建任意的网络测量任务；随后，LightNF 调用其自动化网络测量任务分析及部署优化框架，计算由用户构建的网络测量任务在底层网络中的最佳部署方案，并自动部署至底层网络可编程交换机，满足部署需求的同时屏蔽底层细节，使整体网络测量任务部署过程保持高效、高性能、高精度的优势。LightNF 发表于计算机网络领域国际会议 IEEE/ACM IWQoS 2021 会议上，并获得大会唯一最佳论文奖（图 4、图 5）。

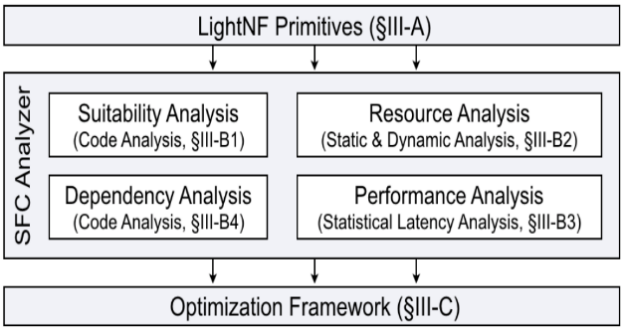

图 4 LightNF 的架构设计

图 5 IEEE/ACM IWQoS 2021 最佳论文奖

在可编程网络测量场景中，Sketch 具备低资源开销与理论精度保证，是一类至关重要的网络测量算子。然而，将 Sketch 高效、准确地部署到底层网络中仍面临一系列挑战，实验室围绕这一问题展开了系统性的研究探索。在全网 Sketch 部署场景中，最优性与可扩展性之间长期难以兼顾。为此，实验室提出可扩展且近似最优的 Sketch 部署框架 Eagle，将全网 Sketch 部署分解为多个子问题，并引入一系列加速求解技术，在复杂网络拓扑下将 Sketch 部署速度提升最高达 255 倍，同时最小化最优性损失，研究成果发表于计算机通信网络领域的顶级会议 ACM SIGCOMM 2024（图 6）。完成 Sketch 的全网部署后，单台交换机上的寄存器资源短缺仍是制约 Sketch 测量精度的关键瓶颈。对此，实验室提出寄存器虚拟化框架 Phantom，创新性地利用交换机内部数据循环通道作为虚拟寄存器的存储介质，并通过交换机与监控服务器的协同设计完成 Sketch 操作的状态化处理。实测结果表明，Phantom 可虚拟化 10 的 6 次方量级寄存器，将网络监控应用的测量精度提升最高达 86%，研究成果发表于计算机系统领域的顶级会议 ACM EuroSys 2025（图 7）。随着新一代多流水线交换机的广泛部署，Sketch 测量又面临阵列分割带来的冗余测量与测量失衡等新挑战。面向这一新型硬件架构，实验室提出高精度 Sketch 测量框架 SketchPipe，通过无分割部署策略与异步状态报文测量机制，从根本上消除上述问题。实测结果表明，SketchPipe 在多种网络监控应用中将测量精度提升最高两个数量级，研究成果发表于计算机网络领域的顶级会议 USENIX NSDI 2026（图 8）。

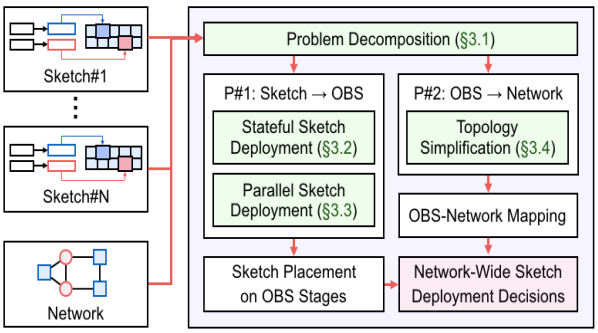

图 6 ACM SIGCOMM 2024 论文的 Eagle 框架

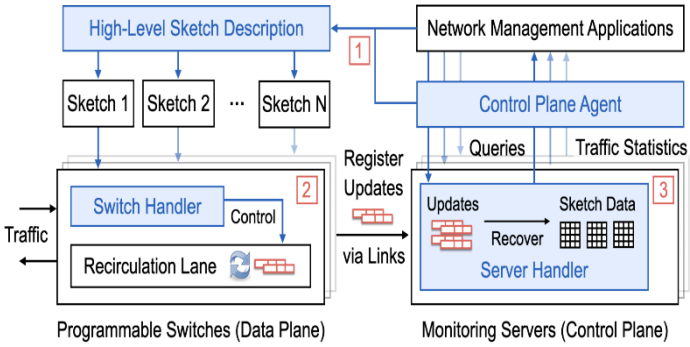

图 7 ACM EuroSys 2025 论文的 Phantom 框架

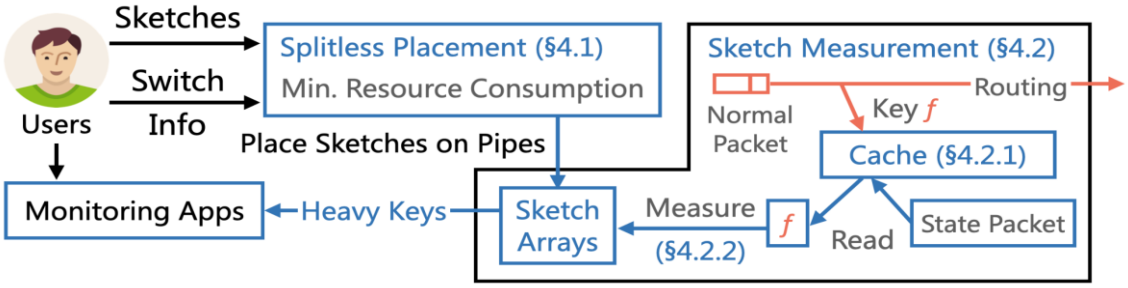

图 8 USENIX NSDI 2026 论文的 SketchPipe 框架

3. AI 大模型技术与应用

实验室在 AI 大模型研究与应用方面，围绕高效、可信、可扩展 AI 系统中的资源供给、模型优

化、数据安全与应用落地，形成了覆盖“基础设施、模型系统、数据安全、算法应用”的全栈式研究体系：

- 在基础设施层，提出超大规模边缘云带宽编排平台（图 9），通过带宽弹性编排、主动异常感知与细粒度限速机制，实现弹性、精准、可预期的带宽供给，部署于全球 3200 余个边缘集群中稳定运行近 3 年，实现 99.9%的带宽分配精度，降低全网带宽开销 10%。
- 在模型系统层，设计了大模型自动训推配置框架，融合 AI 理论与分布式系统协同优化，构建从训推信号采集到参数配置的自动闭环，已在蚂蚁集团及某上市企业部署，降低 92%的模型开发时间，推理速度提升 4–16 倍不等。
- 在数据安全层，提出跨域隐私计算自动部署框架，对不同机构的异构环境进行统一抽象，显著降低安全计算部署成本，已在百余家企业和机构落地应用，部署周期从数月缩短至数天，平均额外执行开销仅为 8.4%。
- 在算法应用层，构建图谱驱动的大模型知识服务，对复杂系统中海量、多样、异构的数据语义进行统一关联，驱动大模型实现边缘云性能异常诊断，在超大规模边缘云的 3200 余个集群中达成分钟级更新与诊断闭环，异常定位准确率达 91%（图 10）。

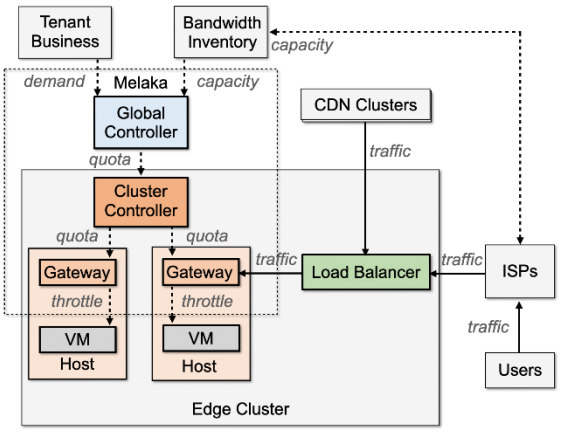

图 9 超大规模边缘云带宽编排架构

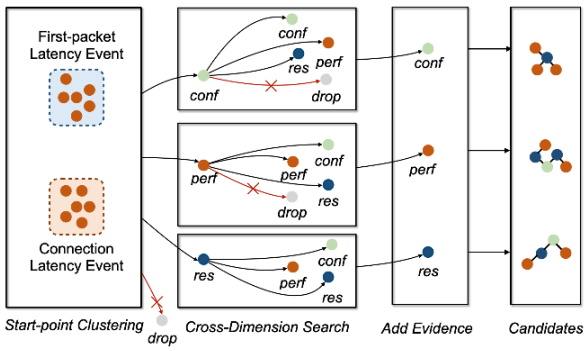

图 10 大规模边缘云性能诊断

4. 智能化程序分析与漏洞挖掘

实验室长期深耕程序分析与漏洞挖掘领域，形成了从静态分析、动态模糊测试到 AI 驱动安全的完整研究体系。在静态程序分析方面，面向混合语言应用与跨语言交互场景，提出了精确的跨语言静态分析方法，有效解决了多语言软件的安全审查难题。在模糊测试方面，除 V-SHUTTLE 在虚拟化管理程序 fuzzing 上取得的突破性成果外，实验室还提出了面向闭源 SDK 库的自动化 Fuzz Driver 生成框架，实现了对闭源软件的高效漏洞挖掘；同时提出了面向深度学习编译器 TVM 的调度优化模糊测试框架 Scuzer，将模糊测试技术拓展至 AI 基础设施安全领域。在系统攻击面缩减方面，通过系统调用限制等手段有效收敛运行时攻击面，相关成果在 Black Hat Europe 2023 上进行了报告。在 AI 赋能安全方面，实验室深入探索了深度学习与程序分析的融合路径：发现了"算法一

致性"是决定深度模型能否形成可控、可验证推理能力的关键因素，系统证明了所提出方法可显著提升图神经网络的可靠性与跨任务的泛化能力。这一成果产生的深远影响在于：将这一思路迁移到大模型与智能体场景中，把程序分析的形式化约束与工具链编排进大模型的推理与决策闭环中，可获得更稳定、更可解释、可复现的自动化程序分析能力。该成果在全球软件工程顶会 ISSTA 2025 上荣获了杰出论文奖（图 11）；进一步将大语言模型与程序分析工具链深度耦合，利用 LLM 引导符号执行揭示智能合约中的不一致性漏洞，实现了基于 LLM 的原型链污染自动化检测与利用生成。上述研究已累计发现 90 余项 CVE 安全漏洞，获得 Apple、Microsoft、Oracle、QEMU 等全球知名厂商的公开致谢。

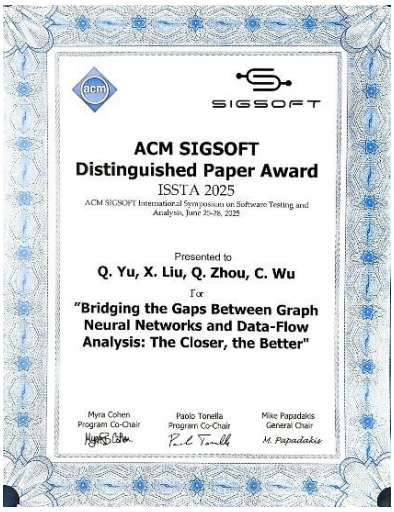

图 11 ISSTA 2025 杰出论文奖

5. 智能攻防与 AI 安全

（1）实验室研究并评估了云环境 Web 应用防火墙（Web Application Firewall，WAF）的安全性，提出了基于语义解析树及上下文无关文法的攻击载荷变异方法，研发了能对云 WAF 进行自动化攻击的框架 AutoSpear。AutoSpear 在真实云上环境对 Amazon AWS、F5、Fortinet、Cloudflare、ModSecurity 等国内外主流厂商的 WAF 实施了安全评估与漏洞发现。在实现最高 99%攻击成功率的情况下，成功发现了若干主流厂商 WAF 存在的高危安全漏洞。AutoSpear 的相关成果入围 Black Hat Asia 2022（图 12）。此外，将内生安全、拟态防御技术应用于 WAF 系统之中（图 13），成功部署于国家电网省级公司内网、教育部大科学装置平台内网等系统中，并防护省级多家三甲医院官网、千岛湖旅游门户网站等公网系统，年均拦截内外网恶意请求近亿条。

图 12 AutoSpear 在世界黑帽大会上亮相

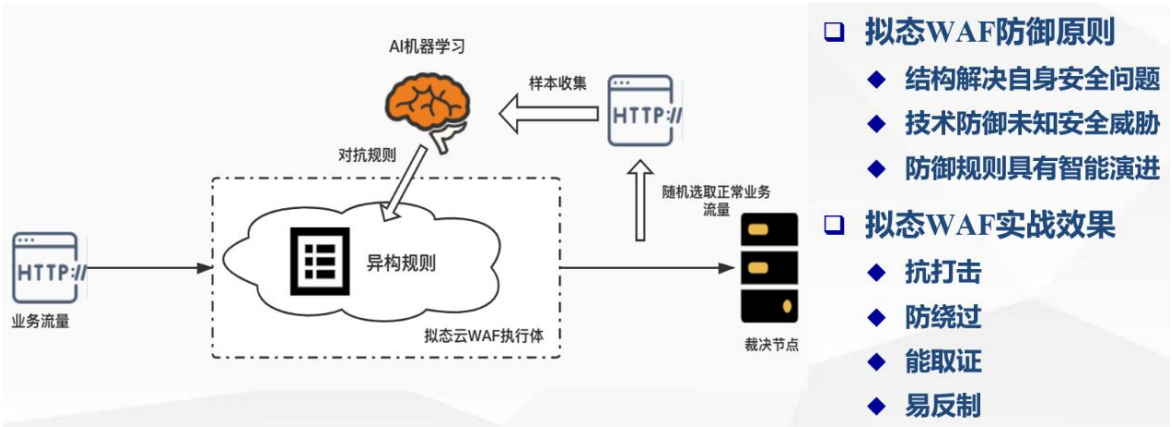

图 13 拟态云 WAF 的架构设计

（2）随着大语言模型的快速部署与广泛应用，其安全性已成为亟待解决的关键挑战。实验室围绕大模型安全评估与攻防技术开展了系统性研究。实验室提出了 面向大语言模型越狱攻击的规模化自动化评估框架，通过自动化的变异策略生成多样化的越狱提示，系统性地评估 LLM 在面对各类攻击时的鲁棒性，为大模型安全评估提供了可扩展的方法论与工具支撑。在此基础上，实验室进一步揭示了多轮对话场景下大模型的安全薄弱环节，提出了 ICON 框架——基于"意图-上下文耦合"（Intent-Context Coupling）的高效多轮越狱攻击方法。ICON 突破了单轮攻击的局限，通过建模攻击意图与对话上下文的耦合关系，在多轮交互中逐步诱导模型产生有害输出，显著提升了对具备安全对齐的大语言模型的攻击成功率，揭示了大模型在复杂交互场景下仍存在的深层安全隐患，为大模型安全对齐与防御机制的研究提供了重要的攻防视角与评估基准。

6. 国内外竞赛

实验室成立的摇光安全战队 FSL 多次参加了国内外网络安全赛事，组员荣获包括 XCTF、DEFCON CTF、DataCon 等大赛的冠亚军战绩。在中国网络安全产业联盟网络安全优秀创新成果大赛、未来网络科技创新大赛上也曾分别荣获全国二等奖。

图 14 实验室摇光安全战队斩获佳绩