

# 前 言

随着物联网、5G 网络、人工智能、大数据、边缘计算、工业互联网等新兴技术支撑的人机物互联融合计算模式的快速发展,泛在计算新形态正在兴起,有望成为系统软硬件创新研发和生态构建的研究热点.此外,开源软件作为泛在计算生态中的重要基础,开源软件的使用已涉及到新一代智能信息技术产业的方方面面,包括操作系统、容器技术、人工智能和网络安全等,开源软件供应链的稳定性和开源协议的安全性已经成为新型计算生态下的重中之重,关系到我国计算机产业的健康发展.针对泛在计算带来的海量异构泛在资源、复杂多样应用场景、多模态人机物交互融合和边界开放融合的计算生态,亟需开展相关研究,解决我国在关键领域面临的技术瓶颈和关键挑战.国内学术界和产业界非常重视泛在计算的发展带来的机遇与挑战,并在该领域展开了大量的研究工作,取得了显著的研究成果.

为了进一步推动我国在泛在计算和开源软件领域的研究,及时报道我国学者在泛在计算和开源软件理论与技术方面的最新研究成果,《计算机研究与发展》策划组织了“泛在计算”专题.本专题通过公开征文共收到 3 篇特邀稿件和 46 篇普通投稿,论文从不同的应用场景阐述了泛在计算和开源软件的问题背景、研究意义、算法设计、关键技术和热点应用等,反映了该领域多个方面的最新研究成果.

本专题的审稿严格按照期刊的审稿要求,特邀编委先后邀请了相关领域的专家,历经初审、复审、终审等阶段,最终共精选出 6 篇高质量的论文入选本专题.内容涵盖了泛在计算和开源软件综述、开源软件缺陷检测技术、开源操作系统兼容性分析、开源软件数据分析等研究内容,在一定程度上反映了当前国内工业互联网安全领域的最新研究成果.

## 1. 泛在计算和开源软件综述

机器学习等相关技术在泛在计算和开源软件中得到了广泛发展和应用.陈珍珠等作者的“面向机器学习的安全外包计算研究进展”一文对近 5 年国内外机器学习安全外包研究工作进行了调研,按照任务阶段和云服务器商数量对现有主流的外包模型进行分类和特征归纳,并从逻辑回归、朴素贝叶斯分类、支持向量机、决策树和神经网络等典型机器学习算法角度对机器学习安全外包计算相关研究进展进行了深入阐述和分析,从数据安全与效率均衡、计算完整性保护、迁移学习安全外包三个方面探讨了安全外包计算未来面临的挑战和机遇.田笑等作者的“开源软件缺陷预测方法综述”一文调研分析了 2000 年至 2022 年 12 月软件缺陷预测研究领域的相关文献,以机器学习和深度学习为切入点,梳理了基于软件度量和基于语法语义的两类预测模型.基于这两类模型,分析了软件缺陷预测和漏洞预测之间的区别和联系,并针对数据集来源与处理、代码向量的表征方法、预训练模型的提高、深度学习模型的探索、细粒度预测技术、软件缺陷预测和漏洞预测模型迁移六大前沿热点问题进行了详尽分析,最后指出了软件缺陷预测未来的发展方向.

## 2. 开源软件缺陷检测技术

安全缺陷是泛在计算和开源软件中不可忽视的问题,由于泛在计算涉及到大规模的异构资源和复杂的应用场景,以及开源软件的广泛使用,安全缺陷可能存在于系统设计、开发、部署和运行的各个环节.熊忻等作者的“基于错误路径行为一致性的内核引用计数缺陷检测”一文针对内核中的引用计数缺陷,提出了基于错误路径行为一致性分析的缺陷检测方案,并引入错误路径的语义信息来推断合理的引用计数行为,该方案在 Linux 内核版本 5.6-rc2 和 5.17 上分别发现 21 个和 9 个已经确认的引用计数缺陷,其中,在内核版本 5.6-rc2 中检测出 9 个已

有工作难以覆盖的引用计数缺陷. 余媛萍等作者的“HeapAFL: 基于堆操作行为引导的灰盒模糊测试”一文针对堆内存错误类漏洞, 提出了一种基于堆操作行为引导的灰盒模糊测试方法 HeapAFL, 在不依赖漏洞先验知识的情况下, 通过静态分析插桩基础堆操作函数及其参数监测执行时控制流和数据流的变化, 反馈堆操作行为信息指导模糊测试中种子优先变异阶段, 探索多样化堆操作行为从而更高概率触发堆内存错误类漏洞, 该方案在 6 个真实程序 (binutils、gpac、jpegoptim、mjs、yasm 和 mxml) 中挖掘到了 25 个堆内存错误类漏洞, 包括 19 个已知的漏洞和 6 个未知的漏洞, 且已经获得了 2 个新的 CVE 编号.

### 3. 开源操作系统兼容性分析

开源操作系统兼容性分析对于确保系统稳定性、提高应用程序兼容性、促进软硬件集成和支持版本管理等方面都具有重要的意义. 吴亦泽等作者的“openEuler 中 C 标准库替换的兼容性分析”一文提出兼容性分析算法来研究 openEuler 的 4 种主要软件生态中的 musl libc 兼容性和缺失 API 优先级, 基于应用软件包之间的依赖关系和谷歌 PageRank 算法的思想, 提出了 PackageRank 算法和 APIRank 算法, 分别用于软件包兼容性度量和优先级计算, 该算法为 openEuler 的新 C 库补全工作提供了有效指导.

### 4. 开源软件数据分析

开源软件数据分析对于提供洞察和决策支持、优化软件性能和用户体验、发现和修复软件缺陷, 以及推动开源社区的发展和创新都具有重要意义. 李庚松等作者的“基于多目标混合蚁狮优化的算法选择方法”一文提出了一种算法选择模型和一种多目标混合蚁狮优化算法, 算法选择模型以集成元算法的准确性和多样性作为优化目标, 同时选择元特征和构建选择性集成元算法, 采用多目标混合蚁狮优化算法对模型进行优化, 使用离散型编码选择元特征子集, 通过连续型编码构建集成元算法, 在此基础上应用增强游走策略和偏好精英选择机制提升寻优性能.

承蒙各位作者、审稿专家和编辑部等方面的全力支持, 本专题得以顺利出版. 目前泛在计算技术涉及领域极为活跃, 并且发展迅速, 学科前沿日新月异, 这给特邀编委及审稿人的审稿、选稿工作带来了巨大挑战, 在此我们表示衷心感谢. 由于投稿数量大、主题广泛、时间安排紧张、专题容量有限等原因, 本专题仅选择了部分有代表性的研究工作予以发表, 无法全面体现该领域所有的最新研究工作, 部分优秀稿件无法列入专题发表, 敬请谅解.

我们要特别感谢《计算机研究与发展》编委会和编辑部, 从专题的立项到征稿启事的发布, 从审稿专家的邀请到评审意见的汇总, 以及最后的定稿、修改和出版工作, 都凝聚了他们辛勤的汗水. 本专题的出版期望能给广大相关领域研究人员带来启发和帮助. 在审稿过程中难免出现不尽人意之处, 也希望各位作者和读者包容谅解, 同时请各位同行不吝批评指正. 最后, 再次衷心感谢各位作者、审稿专家、特邀编委和编辑部的辛勤工作.

张玉清 (中国科学院大学)

郭耀 (北京大学)

马华东 (北京邮电大学)

武延军 (中国科学院软件研究所)

刘杨 (新加坡南洋理工大学)

段海新 (清华大学)

马全一 (华为技术有限公司)