我们在未来！

并行與分布式計算通訊

地 址：武汉市华中科技大学东五楼二楼
邮 编：430074
电 话：15387112483
E-mail：yany@hust.edu.cn
Homepage：http://grid.hust.edu.cn

BDTS 大数据技术与系统国家地方联合工程研究中心
SCTS 服务计算技术与系统教育部重点实验室
CGCL 集群与网格计算湖北省重点实验室

并行與分布式計算通訊

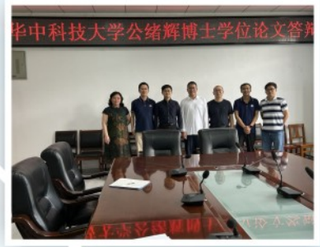
BING XING YU FEN BU SHI JI SUAN TONG XUN

2023年第2期 总第53期 2023年06月

封面人物：赵进——努力造就成就，态度决定高度
2023届毕业生成果展示
计算机科学与技术专业建设50周年学术讲座在武汉举行
科研之路在于发现问题解决问题

http://grid.hust.edu.cn

2023届博士及硕士毕业生顺利完成答辩



为庆祝计算机科学与技术专业建设50周年实验室举办多期高端学术讲座

风采

致实验室 2023 届毕业生

在日复一日的打卡音中，在三点一线的脚步里，四季轮转，时过三载。小小的东五楼，承载着你们努力的汗水，见证了你们拼搏的岁月。而今，夏蝉奏响离别的乐章，实验室即将与你们挥手告别，全体老师将看着你们踏上新的征程，走向更广阔的舞台。

本期通讯将展示实验室2023届全体毕业生的学位毕业论文摘要，记录和总结同学们专心科研取得的丰硕成果。同时，实验室全体老师将为你们送上衷心的祝福和殷切的期望，希望你们一路有鲜花为伴，有清风相随，在各自的道路上谱写出更加绚丽的新华章：

陈姝彤：从数学小能手到计算机学者，你的未来潜力无限！

公绪辉：希望在今后的工作岗位上站得更高，看得更远，做出更好的成绩。

罗瑞坤：研究无边缘，志向上云天！

牛轶佩：六边形战士，前途无量！

石建：重整待发，扬帆起航，期待你在未来工作中勇挑重担，脱颖而出！

王坤：希望坚持内核方向，将来做出能落地的高水平成果！

王庆刚：沉着冷静，踏实前行，进步巨大，未来可期！

吴文超：努力成为大模型时代的系统架构师！

于水英：祝未来有更多坦途，事业顺利，家庭幸福！

柏东山：愿研究生学业之路奠定坚实基础，毕业之际展翅高飞，实现自己的人生理想。

陈佳杰：希望养成早起早睡的良好习惯，继续保持强大的工程能力！

陈丽君：入则进德修业，出则任重道远。

陈铭涛：天道酬勤，所有的努力都有收获，祝你爱情事业双丰收，从此人生皆坦途。

陈子睿：祝子睿宝宝前途似锦，开心每一天！

但镇武：祝镇武勇于担当，步步高升！

董硕华：保持心中热爱，奋楫笃行，奔赴光明山海，祝硕华顺顺利利！

窦昊辰：祝昊辰博士学业顺利，前程似锦！

方一霖：求学顺利，学业有成！

冯豆：面朝大海春暖花开，加油！

贺康：达摩院的大佬，求带兄弟们一起飞啊！

贾亚茹：你曾面对挑战，努力后收获满满；服务器里留有你的口罩图片，更有你的奋斗痕迹，记得回来看看；愿我们爱笑的甜美女孩事业顺利，生活幸福美满！

姜新宇：百折不挠，继续前行，将自己所学为国家经济发展贡献自己力量！

蒋志国：祝志国前程似锦，人生路上以奋斗为邻，收获成功。

李可欣：祝学业顺利，早日拿到博士学位！

李铭：希望研究生三年的学习生活能让你做事更加细心，祝你工作后保持这种状态，相信以后你一定能干出一番事业。

李星儒：祝组里唯一的小美女工作和生活都开心！

李妍：聪颖、善良、内秀，新的旅程即将开启，希望你更加自信、坚强、乐观！

刘烬凯：基本上把项目的事情交给你就完全不用操心，你在将来的工作中一定可以乘风破浪。

刘晓耕：为国而生，师夷长技以制夷！

刘旭东：付出总有回报，希望你用积极踏实的态度把好人生之舵，拥有光辉灿烂的未来。

陆建荣：不惧风雨，苟利国家生死以！

陆弈帆：对你的技术深信不疑，一定能成长为独当一面的网络安全技术大牛！

泥俊沛：祝俊沛在筑梦历程中破浪杨帆、前程似锦！

潘婷婷：毕业不是终点，而是新的起点。祝你在新的领域里继续发光发热！

潘婉宁：研究生毕业是新的起点，愿你不

断追求知识和突破，创造出令人瞩目的成就，成为引领行业发展的杰出人才！

齐蒙：期待你能把技术和社交完美结合，成长为一名复合型人才。

秦轶群：愿自信独立的你逢山开路，遇水搭桥，有前程可奔赴，亦有岁月可回首！

沙岩：祝你在未来的道路上勇往直前，无所畏惧！

史瑞泽：继续加油！祝早日拿到博士学位！

侍言：脚踏实地，善于探索，渴求新知，希望你在新的工作岗位上早日成为技术大拿。

司冰茹：愿冰茹载着美好信念、美丽智慧和抗挫勇气勇敢启航！

宋钊：努力，奋进，向上。

宋振：从青涩到成熟，成长的收获伴你振翅高飞！

孙茂林：不忘初心，砥砺前行，追求卓越，未来一定不可限量！

孙元伟：愿你未来的征程一切顺利！

唐存志：好好干，你一定能为国防做出重要贡献！

万俊：继续加油，前程似锦！

王宏恩：勇于迎接挑战，不断突破自我，为职业发展和人生幸福努力奋斗！

王洪磊：祝在新的平台上创业成功！

王淼：祝王淼不忘初心，砥砺奋进！

王千千：作为第一个走出实验室的千禧娃娃，希望你能在外面打出实验室的品牌，无论是争先还是躺平，都要秉承初心。加油吧孩子！

王勇聪：毕业快乐！未来更主动，更自信，前途会更加光明！

魏昊辉：以梦为马，以汗为泉，不忘初心，不负韶华！祝前途似锦！

吴敏康：祝前程似锦，心想事成！

吴锐：学习态度端正，沟通能力强，做好时间的主人，你将无往而不利！

徐钫洲：沉稳踏实、聪明勤奋、温和平静，希望你永远孜孜以求，知难而进，经得起风雨，方能见彩虹。

徐润泽：祝润泽事业一帆风顺，生活幸福！

许鸿皓：祝鸿皓小美女事事顺心，阖家幸福！

杨静：祝前程似锦，心想事成！

杨茂根：相信自己，事业有成！

杨艺欣：读博期间一定要恪守规矩，好好听话，替华科争光！

杨赟：乐于追求卓越，祝一切顺利！

叶翔宇：新的征程即将开始，祝在新的岗位一如既往地创造价值。

叶贻良：小伙子遇事别太焦虑，祝事业有成！

于亮亮：做项目总是充满着热情，踏实认真，在以后的工作岗位上一定能游刃有余。

于泽霖：插上梦想的翅膀，你可以飞得更高！

岳镜涛：脚踏实地、行稳致远，祝镜涛在与时代同频共振中绽放光芒！

张帆：新的征程，扬帆起航！

张亮：开朗阳光是你的第一印象，勤奋好学，敢于迎难而上是你最大的优点。天道酬勤，相信未来的你定能闯出自己的一片天地。

张雨婷：毕业快乐！愿你的未来能一如既往的美丽，灿烂，夺目！

张哲畅：早日学成归国，期待你能以海外优秀人才引进回母校。

赵瑞欣：你用毅力取得今天的收获，又将以果敢接受明天的挑战。愿我们坚韧的文静女孩坚持努力，遇见幸运，前程似锦，美梦成真！毕业快乐！

赵洧：毕业快乐！你是最优秀的！

朱春雪：愿你愿望成真，巾帼不让须眉，为五彩人生添加永恒的绿色！

朱占洋：百炼成钢，愿在新的岗位宏图大展！

同学们，去日不可追，来日尤可期。愿你们将遗憾留在过去，带着笑容与希望，踏着朝阳，在这个热烈的季节扬帆起航，奔向远方。

海阔凭鱼跃，天高任鸟飞，我们相信优秀的同学们将在五湖四海自由施展才能，绽放光芒。实验室将始终以你们为荣、以你们为傲！

实验室全体老师

二〇二三年六月

主　编：金　海
本期执行主编：张　腾
编　委：（按姓氏拼音排序）陈汉华、代炜琦、戴小海、丁晓锋、段卓辉、樊　浩、耿　聪、顾　琳、何　强、胡胜山、华强胜、黄　宏、黄晨明、蒋文斌、李　珍、李　志、廖小飞、刘海坤、刘英书、陆　枫、马晓静、毛伏兵、羌卫中、邵志远、石宣化、陶　莉、万　瑶、王多强、王虹飞、王　雄、文　明、吴　松、肖　江、徐　鹏、姚德中、叶晨成、余　辰、袁　斌、袁平鹏、张　腾、张　宇、郑　龙、郑　然、邹德清
责任编辑：燕　燕

地　址：武汉市华中科技大学东五楼二楼
邮　编：430074
电　话：15387112483
E-mail：yany@hust.edu.cn
Homepage：http://grid.hust.edu.cn
（此刊仅供内部交流学习）

卷首语
……1

热点
……4

封面人物
努力造就成就，态度决定高度……赵　进　9

专栏
2023 届毕业生成果展示……11

声音
论人工智能的低碳转型……储朝阳　48
解释型语言生态中的供应链攻击与防御……郭啸辰　51

动态
为庆祝计算机科学与技术专业建设 50 周年实验室举办多期高端学术讲座……燕　燕　54

推荐
Combining Optimization Objectives: New Modeling Attacks on Strong PUFs……万才学　推荐　55
GraphSPD：Graph-Based Security Patch Detection with Enriched Code Semantics…谢子凡　推荐　57
On Sharding Across Heterogeneous Blockchains……刘存扬　推荐　61

交流
科研之路在于发现问题解决问题……陈　铭　64

微信首推刷掌支付，便捷之外如何保障安全

（孙一飞 整理）

没带公交卡、手机没电，也能坐地铁了。

近日，微信支付联合北京轨道交通推出了大兴机场线刷掌乘车服务。乘客注册开通刷掌乘车服务后，即可在该条地铁线路刷掌过闸乘车。这也标志着微信支付正式发布“微信刷掌”产品。

乘车前，乘客需在大兴机场线各站自动售票处的刷掌注册机录入掌纹，并在微信小程序“北京地铁刷掌乘车”完成授权后，可开通刷掌乘车功能。乘坐地铁时，用户在带有圆环标志的刷掌闸机上扫描手掌，即可进出站。

据微信方面透露，“微信刷掌”采用了“掌纹+掌静脉”识别技术，针对“无介质”“非接触”“高便捷”“高安全”等用户需求提供了补充支付方式。除交通场景外，该产品正逐步应用于办公、校园、健身、零售、餐饮等领域。

“相对于刷脸支付，刷掌支付更容易在便捷性和安全隐私方面找到一定的平衡。”多位分析人士对《财经》记者表示，“刷掌”不受口罩影响，相对人脸信息来说掌纹信息的敏感度也相对较低。

不过，对于刷掌支付的前景，多位业内人士表示不太乐观。“近几年大众对隐私保护重视程度日渐提高，大家是否愿意继续将自身生物识别特征与支付绑定需要打个问号。”博通咨询金融业资深分析师王蓬博对《财经》记者表示，从此前各支付公司推广“刷脸支付”的情况来看，线下机具改造、渠道铺设中存在的一些问题和争议，在“刷掌支付”的推广中可能依然存在。

刷掌支付采集的生物识别数据安全也颇受关注。华东政法大学数据法律研究中心主任高富平对《财经》记者表示，作为一种生物识别信息，掌纹与指纹、人脸等信息一样，都有可能被用于跟踪用户个人轨迹。在他看来，支付机构使用个人生物识别数据进行身份验证无可厚非，但应该严格限制除身份验证之外的二次开发应用。“这需要监管部门加强规制。”高富平表示。

此外，招联金融首席研究员董希淼表示，此前有诈骗人员骗过银行“人脸识别”系统盗取储户资金，刷掌技术作为一种验证手段的安全性亦需要进一步验证。

Wind（万得）数据显示，5月22日午后刷掌概念走高，熵基科技拉升涨超10%，汉王科技、远方信息、百胜智能等盘中纷纷冲高。截至当日收盘，汉王科技涨停，熵基科技、远方信息、百胜智能均涨超7%。

（参考链接：https://www.huxiu.com/article/1591255.html）

无线体域网络（WBAN）

（岳 航 整理）

无线体域网络（Wireless Body Area Network，WBAN）是一种基于无线通信技术的网络，用于监测人体内或周围的生理参数和环境参数。WBAN可以通过传感器和无线通信技术实现对人体健康状况的实时监测和远程管理，具有广泛的应用前景，如医疗保健、运动健身、智能家居等领域。

WBAN的传感器通常包括生理参数传感器和环境参数传感器。生理参数传感器可以监测人体内的生理参数，如心率、血压、体温、血氧饱和度等；环境参数传感器可以监测周围环

境的参数，如温度、湿度、气压等。这些传感器可以通过蓝牙、ZigBee、WiFi等无线通信技术与移动设或云端服务器进行数据传输和处理。

WBAN的应用场景非常广泛。在医疗保健领域，WBAN可以用于远程监测患者的生理参数，提高医疗服务的效率和质量；在运动健身领域，WBAN可以用于监测运动员的生理参数，帮助他们更好地掌握自己的身体状况；在智能家居领域，WBAN可以用于监测家庭环境的参数，实现智能化的环境控制。

除了以上提到的应用场景，WBAN还可以在其他领域发挥作用。例如，在军事领域，WBAN可以用于监测士兵的生理参数和环境参数，以提高作战效率和保障士兵的安全；在工业领域，WBAN可以用于监测工人的生理参数和环境参数，以确保工作场所的安全和健康；在智能交通领域，WBAN可以用于监测驾驶员的生理参数，以提高驾驶安全性和减少交通事故的发生率。

为了实现这些应用场景，工程师需要不断地研究和改进WBAN的技术。例如，可以采用更高效的传感器和通信技术，以提高数据传输的速度和准确性。同时，可以采用更智能化的算法和模型以实现更准确的数据分析和预测。

此外，工程师还需要关注WBAN的可扩展性和互操作性。由于WBAN涉及到多种传感器和通信技术，因此必须确保这些技术之间的互操作性和可扩展性，以便将来可以轻松地添加新的传感器和技术。

（参考链接：Singh S, Prasad D. Wireless body area network (WBAN): A review of schemes and protocols[J]. Materials Today: Proceedings, 2022, 49: 3488-3496.）

大语言模型在软件自动修复领域的应用

（江宗泽 整理）

随着软件程序和系统在日常生活中变得越来越普遍，软件错误也变得越来越常见。由于软件系统在医疗保健、运输等领域得到广泛应用，这些错误可能会导致危险的安全问题和财务损失。因此，开发人员通常需要花费大量时间和精力来修复软件错误。为了帮助开发人员减少这种手动工作，已经建立了自动程序修复（APR）工具，可以根据原始有错误的程序自动生成潜在的补丁。在传统的APR技术中，基于模板的APR被广泛认为是最先进的技术。这些技术利用由人类专家设计的修复模板来修复源代码中特定类型的错误。因此，这些APR工具受到底层修复模板的限制，因为它们只能修复特定类型的错误。为了解决这个问题，研究人员提出了基于学习的APR工具，通常将程序修复建模为神经机器翻译（NMT）问题，其目标是将有错误的程序翻译成修复的程序。这些基于学习的APR工具需要包含有错误代码和已修复代码对的受监督训练数据集，通常是通过从开源代码库中挖掘历史错误修复获得的。尽管基于学习的APR工具已经在修复可修复的错误数量和类型方面取得了改进，但它们仍然受限于它们的训练数据，这些数据可能包含无关的提交并且仅包含有限的错误修复类型，可能无法推广到未见过的错误类型。

最近，针对构建大型预训练语言模型（LLM）的最新发现为程序修复提供了一种替代解决方案，而无需依赖于历史错误修复。虽然LLMs通常是用于自然语言处理（NLP）任务的通用工具（例如GPT-3），但它们也已经被用于通过代码微调（例如Codex和ChatGPT）进行编程语言。与专门设计的基于学习的APR模型

不同，LLMs是以无监督的方式训练的，使用了多达数十亿个文本/代码标记，并且可以用于各种代码任务。AlphaRepair[3]提出利用CodeBERT，这是一个在数百万个代码片段上预训练的大型代码模型，直接用于APR。AlphaRepair的关键见解是，我们可以直接使用模型根据其周围的上下文（包括前缀和后缀）预测正确的代码样子，而不是学习从有错误的代码到修复代码的转换，即填充式APR。利用这个想法，AlphaRepair展示了最先进的修复结果，而无需在错误修复数据集上进行微调。尽管AlphaRepair已经显示出比以前的基于学习的APR更好的结果，但它使用的模型（125M参数）比当前最先进的LLMs（Codex：12B参数和GPT-3：175B参数）要小得多。

（参考链接：http://lingming.cs.illinois.edu/publications/icse2023a.pdf）

人工智能赋能威胁情报分析

（朱嘉豪 整理）

随着网络威胁的不断演化和复杂化，传统的安全防御手段往往无法及时发现和应对新型的攻击和威胁。基于人工智能的方法通过利用机器学习、数据挖掘和自然语言处理等技术，能够自动化地分析和利用大量的威胁情报数据，从中发现模式、识别异常和预测潜在的网络威胁。

该领域的研究聚焦于多个方面，包括威胁情报的收集、整合和分析。研究人员致力于开发自动化的机制和工具，从不同的数据源中收集威胁情报，并将其整合到一个统一的平台中。通过智能算法和模型，威胁情报分析和挖掘的研究旨在揭示数据中的隐藏模式和关联性，以便识别新的威胁、预测攻击趋势和生成实时的威胁情报报告。人工智能和机器学习的技术可以帮助自动化和加速这个过程，从海量的数据中发现模式、识别异常和预测威胁。以下是该领域的几个重要研究方向：

1. 威胁情报收集与整合：这个方向关注如何有效地收集、整合和管理威胁情报数据。研究人员致力于发展高效的数据收集方法，包括主动和被动方式，从各种来源获取威胁情报数据。同时，他们还致力于构建威胁情报平台和工具，以整合和管理大规模的威胁情报数据，使其更易于访问和分析。人工智能技术可以用于自动化和智能化地收集威胁情报数据，减轻研究人员的负担。例如，机器学习算法可以分析网络流量和日志数据，自动发现异常活动和潜在的威胁源。同时，自然语言处理和信息抽取技术可以从各种来源的文本数据中提取关键的威胁情报信息。

2. 威胁情报分析与挖掘：这个方向关注如何从海量的威胁情报数据中提取有用的信息和洞察，并进行深入的分析和挖掘。研究人员利用机器学习、数据挖掘和统计分析等技术，发展智能化的威胁情报分析方法，用于识别威胁模式、发现新的攻击技术和漏洞，并预测潜在的威胁行为。

3. 威胁情报可视化与交互：这个方向关注如何有效地呈现和传达威胁情报的结果和洞察。研究人员致力于开发可视化工具和技术，将复杂的威胁情报数据以直观和易于理解的方式展示出来，帮助安全专业人员更好地理解和分析威胁情报，并做出相应的决策和响应。人工智能技术可以提供交互式的威胁情报可视化工具，使安全专业人员能够直观地理解和分析威胁情报。这些工具可以将复杂的数据转化为可视化图表、图形和动画，帮助用户发现威胁

模式和趋势，以及进行针对性的决策和响应。

4. 威胁情报共享与合作：这个方向关注如何促进不同组织之间的威胁情报共享和合作。研究人员致力于设计安全的共享平台和机制，确保威胁情报的安全传输和保护，并建立可信任的合作模型和激励机制来促进组织之间的信息共享。人工智能技术可以在安全信息共享和合作中发挥关键作用。例如，机器学习算法可以用于自动化标记和分类威胁情报数据，使其更易于共享和理解。

基于人工智能的威胁情报分析为网络安全提供了新的思路和方法，可以有效地提升威胁检测和响应的能力。它不仅可以帮助组织及时识别和防御各种威胁，还能够提供更准确的威胁情报，指导安全决策和策略制定。这一领域的研究和应用具有重要的意义，对于保护网络安全和应对日益复杂的网络威胁具有重要的价值。

（参考链接：https://zhuanlan.zhihu.com/p/398338693）

Meta推出的“分割一切（SAM）”模型已经迎来二创

（孙一飞 整理）

Meta 近期发布的「分割一切」AI 模型（Segment Anything Model，简称SAM）。Meta 表示，「SAM 已经学会了关于物体的一般概念，可以为任何图像或视频中的任何物体生成 mask，甚至包括在训练过程中没有遇到过的物体和图像类型。SAM 足够通用，可以涵盖广泛的用例，并且可以在新的图像『领域』即开即用，无需额外的训练。」

这一模型的发布在计算机视觉领域引发轰动，预示着 CV 也将走向「一个全能基础模型统一某个（某些？全部？）任务」的道路。当然，大家对此早有预感，但没想到这一天来得如此之快。

比基础模型迭代更快的是研究社区「二创」的速度。论文才刚刚发布两天，几位国内工程师就基于此想出了新的点子并将其付诸实践，组建出了一个不仅可以「分割一切」，还能「检测一切」、「生成一切」的视觉工作流模型。

具体来说，他们使用一个 SOTA 的 zero-shot 目标检测器（Grounding DINO）提取物体 box 和类别，然后输入给SAM模型出 mask，使得模型可以根据文本输入检测和分割任意物体。另外，他们还将其和 Stable Diffusion 结合做可控的图像编辑。

这个三合一模型项目名叫 Grounded Segment Anything，三种类型的模型既可以分开使用，也可以组合使用。

对于 Grounded Segment Anything 未来的用途，项目作者构想了几种可能：

可控的、自动的图像生成，用于构建新的数据集；提供更强的基础模型与分割预训练；引入 GPT-4，进一步激发视觉大模型的潜力；一条自动标记图像（带 box 和 mask）并生成新图像的完整 pipeline等。

（参考链接：https://www.thepaper.cn/newsDetail_forward_22624391）

数据异构的联邦学习

（石月鑫 整理）

联邦学习（Federated Learning）是一种新兴的机器学习范式，旨在让多个客户端共同训练一个模型，而不必共享隐私数据。每个客户端将其本地数据用于本地模型的训练，然后将更新后的本地模型参数上传到中央服务器进行

聚合和更新，以进一步提高模型的准确性。与传统的分布式机器学习相比，联邦学习具有更好的隐私保护和安全性。

数据异构是联邦学习研究的热点方向。它指的是在参与方的数据具有不同类型、格式或分布的情况下进行模型训练，也被称为非独立同分布（Non-IID）数据。例如，不同医院拥有的病人数据可能会有所不同，如年龄、地域以及病史等，导致数据分布不同。再比如，当用户通过不同的手机操作系统使用应用程序时，手机设备上的传感器数据有可能存在发生规律改变，从而导致数据收集环境的差异。总之，当各个客户端的本地数据分布相互异构时，每个客户端的本地模型朝向各自的最优解优化，最终导致聚合后的全局模型偏移，精度严重下降。

解决联邦学习中的数据异构通常有以下几种策略：局部更新全局模型：在每个本地设备上，只更新本地设备上的数据样本对全局模型的影响，而不是直接更新全局模型。加权聚合：在全局模型上进行加权平均，其中每个本地模型的权重与其训练数据的相关性有关。迁移学习和元学习：将从其他领域中获得的知识应用于联邦学习，或者使用元学习方法来针对非独立同分布的数据集进行建模。截止到发稿，CVPR2023公布了3篇关于数据异构的联邦学习的文章，下面进行逐一介绍。

Dengsheng Chen等提出了一种弹性联邦学习聚合方案，弹性聚合根据参数敏感度自适应地对聚合各个客户端的本地模型。其中，参数敏感度通过计算每个参数变化时整体预测函数输出的变化量来衡量。弹性聚合减少了对敏感参数的更新幅度，以防止全局模型的模型参数向任何一个客户端模型的分布偏移，同时，弹性聚合提高了对不太敏感参数的更新的机会，使全局模型能更好地探索不同的客户端分布。大量实验结果表明，弹性聚合可以在完全不可知客户端异质性的前提下，在凸和非凸设置中均实现高效训练，并且对大量客户端、部分参与和不平衡数据具有鲁棒性。此外，弹性聚合与其他联合优化器配合良好，并在各方面实现了显着改进。

Jian-hui Duan等提出了FedFusion，FedFusion是一种基于数据不可知分布融合的模型聚合方法，以优化本地数据集异构（非独立同分布）的联邦学习。在此基础上，异构客户端的数据分布可以用具有不同参数和权重的多个虚拟融合组件的全局分布来表示。作者开发了一种变分自动编码器（VAE）方法，以基于从本地模型中提取的有限统计信息来学习分布融合组件的最佳参数，并应用派生的分布融合模型来优化具有非IID数据的联邦模型聚合。大量的实验结果表明，与其他基线模型相比，FedFusion实现了显着的性能提升。

Bo Li等通过实验观察和理论证明，对在数据异构条件下联邦学习的全局模型参数偏移进行了研究。作者对不同客户端的本地模型进行了对比，发现各个客户端的本地模型的特征提取层差异不大，而分类层（全连接层）差异最大，也就是说，分类层的差异导致了聚合后的全局模型的参数偏移。此外。作者还通过理论证明了结论的正确性。基于这个结论，作者提出了FedPVR，FedPVR在模型聚合时，特征提取层直接使用FedAvg进行参数平均聚合，分类层在聚合时使用全局模型参数进行正则化校正。大量的实验结果表明，FedPVR在数据异构条件下显著优于其他基线模型。

（参考链接：https://openaccess.thecvf.com/CVPR2023?day=all）

努力造就成就，态度决定高度

时光如白驹过隙，回首在博士五年的时光，从初入实验室的青涩，到论文投稿时的紧张，再到论文录用时的喜悦，历历在目，恍如昨日。借此机会，谈一谈博士期间的生活和经历，期望能够对学弟学妹有所帮助。

与实验室结缘

记得在2016年大三结束，在中国地质大学（武汉）计算机学院进入保研阶段，从学院老师了解到了金老师团队的招生广告。随即我也在实验室官网上了解了实验室的整体研究方向，对实验室的相关研究内容十分感兴趣，印象深刻，后来也有听其他保研到实验室的师兄说金老师的实验室做研究很厉害。因此，在看到这个招生相关信息后，我就果断联系了实验室系统组的廖小飞老师。廖老师也给我了很多做系统软件相关研究的建议，这对我后续的发展有着极大的帮助，最后也顺利地拿到保研名额并进入了实验室。从与实验室结缘到毕业至今，已经过去了六个年头，每当想起实验室这个大家庭以及在实验室里度过的时光，心中满是感激。

在实验室科研

在确定保研名额后，我就加入到了东五楼实验室的相关研究中，当时带我的老师是实验室的张宇老师。张宇老师在我大四的时候就给我安排了相关学习和研究工作，包括本科的毕业论文也是在张宇老师指导下完成的。在大四的时候我就在接触图计算相关的研究课题，张宇老师给我发了大量相关的学术论文供我学习。在看论文学习的过程中，廖老师和张宇老师也提醒我要学会总结论文的核心思想，即发现了什么本质问题？采用了什么创新方法？达到什么效果？在此期间，我养成了在每篇论文阅读时做标记和总结的习惯，通过阅读大量的文献使我在往后的工作中受益匪浅，使我能够快速地接触新的研究领域，找寻最本质的问题和痛点。

在正式加入实验室之后，我主要开始做图计算相关的研究工作。刚开始的一段时间里比较迷茫，虽然阅读了较多的论文，然而并不知道自己要做的问题是什么，也觉得很多方向是可以做的，但不知道该怎么下手。在这个时候，通过和张宇老师进行沟通和交流，确定了后续的研究工作。针对并发图计算任务来做相关的系统软件，这在当时的研究是比较少的。相信很多学弟学妹们在科研的过程中也会遇到类似的问题，在这个时候往往需要和自己的指导老师进行沟通交流，因为他们往往有丰富的经验来判断哪个方向是有可以挖掘的问题，同时也要结合自己的兴趣点来做判断。然而，一旦自己做了决定，那么就要付出持之以恒的努力，要将这个问题死磕到底。

于是，在确定具体的研究方向之后，便潜心研究并发图计算相关的问题。这个过程无疑是漫长的，需要对现有的大量的图计算系统进行测试分析，以发现他们在支持并发图计算时面领的瓶颈问题，同时也要在方法论层面来分析他们的问题，随后将针对这些问题来设计方法并进行验证。在论文撰写的时候，由于是第一次写这种英文小论文，张宇老师就建议我找几篇相关的顶会论文，依葫芦画瓢的进行撰写。因为学术论文的写作是有固定模式的，需要多学习和借鉴顶会论文的框架，并积累别人论文中好的语句。

同时，在自己将论文写好后，离投稿截稿还有时间的话，可以将自己的论文多给其他老

师看一下，提提意见。因为自己在写论文时可能会陷入自己的思维逻辑中，导致很多地方写的不清楚，或者逻辑不清晰，从而使得审稿人很难看懂你的论文。因此，我们在写论文的时候要多考虑和揣测读者的感受，要让自己的思想清晰且有逻辑地展示给读者。随后将论文投稿到SC 2019上，并在当时申请了新加坡国立大学访问交流的机会。幸运的是，在访问交流期间，新加坡国立大学的何丙胜老师以及英国华威大学何黎刚对论文rebuttal提供了很多宝贵的意见，在当年8月份，这篇论文被SC 2019录用了。直至现在我都记得当时看到论文录用消息时的喜悦和激动，第一次投稿CCF A类会议就能够直接录用。同时，这也给我更大的鞭策，在后续的工作中要持续努力，而不是昙花一现。

为此，在后续的科研生活中，基于第一篇录用论文中所采用的方法，我们进一步分析了新的问题和挑战，并再进一步做了相关工作。这是因为廖小飞老师和张宇老师多次强调，博士期间的工作要聚焦，不要做的太分散，这样后续的博士论文才能很好地将这些工作串在一起。随着写论文数量的增加，我也慢慢领悟到一篇好的论文，不仅需要一个好的想法，同时也需要用精妙的表达把想法呈现给读者。所谓精妙的表达，就是论文要逻辑精密、用词准确、表达流畅、重点突出。写作时要时常站在评审人的角度来审视自己的论文。这些能力不仅仅是在科研上有用，在日常生活和工作中让别人能够很轻松地理解你的想法也是十分重要的。

科研感悟

科研的道路往往不是一帆风顺，而是充满了坎坷和艰辛。一颗坚持不懈、死磕到底的决心是突破这些困难的必备条件。当我们在选择了一个研究领域后，一定要坚持不懈地对其深挖，切记不能不停地更换研究方向，这样很难做出有影响力的工作。同时，对投稿的失利也不要气馁。我在第一篇论文录用后，后续的投稿也连续被拒了三四次。唯一能做的就是根据审稿人的意见对自己的工作查漏补缺，持续的优化。相信只要自己一直努力，总会是收获到好的结果。

同时，要充分利用实验室和课题组的资源，加强合作交流。例如通过实验室和课题组的平台与其他老师建立联系，进行访学交流等。当你在对一个课题感到迷茫时，这些老师往往能够为你理清思路，让你少走很多弯路。同时，他们也会从不同的视角来看待问题，这也可能会为你带来新的研究思路。

最后，一篇科研论文所解决的问题以及内在机理的发现往往是其精髓所在，决定了这篇论文的高度。因此，我们在构思一个工作时需要在这方面多费心思。通常，需要从所研究领域最本质的特征或特点中来寻找新的问题。这是因为往往这样的问题是一个共性且重要的问题。同时，也要多阅读其他领域的顶级论文，这些论文也会给我们带来不一样的视角和想法，对我们所困扰的问题带来全新的解决思路。

致谢

回顾博士五年，首先要感谢廖小飞老师和张宇老师的悉心指导，若不是两位老师在科研过程中授业解惑，我很难在博士期间发表多篇论文并顺利毕业。同时也十分感谢金老师对我的帮助与鼓励，以及CGCL大家庭其他老师和同门们对我的关心和帮助，在此向您们致以最真挚的感谢。同时祝福实验室越来越好，人才辈出！师弟师妹们都有更好的发展，前程似锦！

赵　进

2022届博士毕业生、之江实验室/华中科技大学博士后

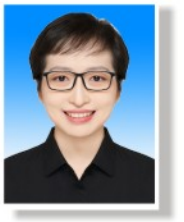
研究方向：图计算、体系结构、系统软件

Email：zjin@hust.edu.cn

2023届毕业生成果展示

博士毕业生

边缘数据中心能效的协同管理方法



姓　　名：陈姝彤

研究方向：数据中心能效管理、边缘计算

导　　师：刘方明

指导老师：刘方明

E-mail：shutongcs@qq.com

QQ：873742144

联系电话：13545240274

毕业去向：广西大学

“万物互联”趋势使计算模式向分布式的边缘计算拓展，也促进了其底层基础设施——边缘数据中心在人口密集区域的海量部署。庞大的规模与尚未成熟的能效管理方案不仅增加了边缘数据中心的用电成本，更加大了在“双碳”目标下的环保压力，如何提高边缘数据中心能源效率已受到工业界和学术界的高度重视。

边缘数据中心的负载与部署特性赋予了边缘数据中心独特的能效优化能力，即边缘架构自主优化与外部设施辅助优化。一方面，边缘计算负载具有动态性、移动性等特点，使边缘数据中心可利用负载管理等方式调整自身能耗，独立实现能效优化目标。另一方面，边缘数据中心大规模部署在城市区域中，可充分借助智能电网与区域供暖系统等城市基础设施带来的交叉生态机遇，协助边缘数据中心进行能效优化。然而，边缘数据中心能效管理也极具挑战。首先，边缘计算负载具有复杂的性能–能耗关系及严格的服务质量要求，需要在性能与能效之间谨慎权衡。其次，边缘数据中心能效管理是一个在线问题，必须在未知未来信息全貌的情况下做出合理决策，避免长期性能与能效损失。最后，边缘数据中心与外部设施间的协同机制尚不完善，难以调动边缘数据中心与边缘计算服务提供商的积极性，影响了边缘数据中心能效协同优化的潜力。

针对上述问题，深入挖掘了边缘数据中心的能耗调节能力，结合在线优化算法与市场激励机制，设计了边缘数据中心能效的协同管理方法：在边缘架构自主优化方面，实现边缘负载性能与能效的协同优化，降低边缘数据中心能源消耗；在外部设施辅助优化方面，从边缘数据中心能源的“需”与“供”两方面切入，分别利用电力需求响应与废热回收供暖降低用电成本与提高能源利用率。边缘架构自主优化为外部设施辅助优化打下基础，外部设施辅助优化为边缘架构自主优化扩宽空间。具体包括以下三个方面的研究内容。

在边缘架构自主优化方面，针对依赖于高功率硬件加速器的边缘计算代表性应用——实时视频流分析，提出了一种负载与能效协同优化机制DynamiE。设计了一种基于深度强化学习的自适应实时视频流分析框架，深入挖掘了该应用复杂的用户体验–能耗关系，在满足能耗限制要求的同时优化了用户体验。此外，通过用户研究分析了用户所处视频内容场景对用户体验倾向的影响，设计了一种多目标深度强化学习解决方案以提升自适应策略训练效率，并提出了一种视频内容场景对自适应策略的映射算法，保证了动态场景下稳定优秀的用户体验与能耗控制表现。通过实际实时视频流分析系统验证了DynamiE在降低能耗、优化用户体验

等方面的高效性。

在外部设施辅助优化方面，提出一种边缘数据中心的计算与电力需求响应机制EdgeDR，通过负载管理与基础设施启停状态管理维护计算与用电高峰期的边缘数据中心可用资源与智能电网稳定性。针对边缘计算服务提供商缺乏参与动机的难题，设计了公平的经济补偿方案，提高服务提供商积极性，辅以高效的负载管理策略实现了服务质量保证。此外，设计了一种动态支付与召回机制，平衡了边缘数据中心运营商短期利润与长期收益。针对成本控制问题，设计了一种在线算法，在无法预知未来信息的情况下决策边缘数据中心启停时机。通过理论分析证明了EdgeDR出色的竞争比与竞标诚实性等性质。通过真实需求响应事件驱动的仿真实验，验证了EdgeDR在实现边缘需求响应、保证服务质量与控制全局成本方面的有效性与优越性。

在外部设施辅助优化方面，针对边缘数据中心另一能源形式——热能，提出一种边缘数据中心废热回收供暖的在线交易与双赢激励机制EdgeHeat，与区域供暖系统协同实现能源再利用。通过实证分析，探讨了利用区域供暖系统实现废热回收供暖的可行性、市场规模与利润等关键问题。针对废热回收供暖场景动态变化问题，设计了废热回收供暖反向拍卖机制与在线成本控制算法，提高了边缘数据中心废热回收供暖积极性，并优化了废热回收供暖的全局运营成本。通过理论证明验证了所提出机制的竞争比与公平性等性能，并通过真实区域供暖系统数据驱动的仿真实验，验证了EdgeHeat在实现废热回收供暖、全局运营成本优化等方面的高效性。

边缘计算环境下存储成本效益优化的数据管理机制研究

姓　　名：罗瑞坤
研究方向：边缘计算、数据存储管理
导　　师：金　海
指导老师：何　强
E-mail：rkluo@hust.edu.cn
QQ：1156532057
联系电话：18940927605
毕业去向：华中科技大学博士后

目前，边缘计算作为云计算的延伸，已经成为分布式计算领域的研究热点。边缘计算带来的低延迟、高带宽和上下文感知等优势使得越来越多的数据与计算任务开始依赖在边缘服务器上进行存储处理。然而，相较于云数据中心，边缘服务器的存储资源是极度受限的。如何在边缘计算环境下高效地存储和管理海量边缘数据以提升存储资源利用率，实现存储成本效益优化 是一个亟待解决的难题。论文分析了传统云计算环境下的数据管理机制在边缘计算环境下所面临的诸多挑战，着眼于探索边缘计算环境下具有存储成本效益优化的数据管理机制，这包括了数据存储前的纠删码分块数据放置策略、数据存储后的冗余数据去重策略和存储资源的负载均衡策略。旨在充分挖掘边缘计算环境的特点，在保障边缘计算环境下数据的低访问延迟的前提下，提高系统的资源利用率，降低数据存储成本。本文主要的研究内容和创新点可以总结为：（1）针对边缘计算环境下数据存储放置主要采用多副本机制所带来的存储资源浪费问题 提出了基于纠删码的边缘数据存储策略；（2）针对边缘计算环境下边缘服务器的高度地理分布式部署所带来的数据冗余度高，以至于边缘服务器存储资源不能被充分利用和释放的问题，建立了以边缘服务器网络拓扑为出发点的边缘数据去重策略。通过删除部分边缘服务器上的冗余数据副本以实现降低数据存储成本。（3）针对边缘计算环境下边缘服务器存储资源的异构性导致数据存储

负载不均衡，以至于降低了存储资源利用率的问题，提出了基于服务器协同的边缘数据负载均衡策略。实验表明，所提出的方法从数据存储管理的三个方面均能够有效较低边缘计算环境下的数据存储成本，方法具有可行性和高效性。

云原生微服务应用的性能优化研究

姓　　名：牛获佩

研究方向：云计算

导　　师：刘方明

指导老师：刘方明

E-mail：newypei@qq.com

QQ：303157334

联系电话：15727008641

毕业去向：华为技术有限公司

云原生微服务应用的运行平台会在不同层面引发应用性能的下降。对于基于PaaS的微服务应用而言，PaaS的资源规模调节和负载均衡机制不够精细，无法适应高度动态的用户请求规模，导致微服务应用性能下降。此外，如何调节云容器实例规模、均衡应用负载是云用户亟需解决的问题。对于基于FaaS的微服务应用，由于FaaS平台能够为每一个用户请求自动化触发和销毁云函数，云用户无需关心云资源规模的调节和负载均衡。然而，BaaS提供的云服务无法实现基于网络的微服务通信，导致微服务应用的性能严重下降。

针对以上问题，围绕两种主要的云原生微服务应用部署平台，解决基于PaaS微服务应用的资源规模调节和负载均衡问题，解决基于FaaS微服务应用的网络通信问题，具体包括以下三个方面的研究内容：

提出了一种性能软保证机制应对突发流量，设计了一种竞争比为1-5ε的在线算法动态调整公有云容器规模适应短时间内波动的用户请求，在降低请求响应时延的同时控制云资源费用，通过真实请求数据驱动的仿真实验，验证了所设计的算法能够高效地分配负载，动态地适应波动的用户请求规模，最终为用户提供高性价比的资源规模调节方案；分析了现有负载均衡算法无法处理微服务链之间竞争和性能需求差异的问题，提出了一种面向微服务链的负载均衡方法，利用消息队列技术极大地简化了应用运维，在降低请求响应时延的同时满足不同微服务链的性能需求，基于真实请求数据的仿真实验表明，所设计的负载均衡策略能够感知不同微服务链的性能需求，显著降低微服务应用的请求响应时间；提出了一种面向网络的云函数直接通信系统，支持复杂工作流中的常见通信模式，支持自主式函数协调，降低了微服务应用复杂功能的端到端时延，大量的系统性能实测数据表明，系统能够显著降低微服务应用中复杂功能的端到端时延。

面向容器的内核资源隔离方法研究

姓　　名：王　坤

研究方向：容器、操作系统内核

导　　师：吴　松

指导老师：吴　松

E-mail：wkhust@hust.edu.cn

QQ：1063894981

联系电话：13317190606

毕业去向：暂未确定

容器因其启动速度快、运行效率高、资源占用少等优点，但是容器的隔离性较弱，限制了容器的进一步发展。目前针对容器隔离性的研究大部分集中于硬件资源隔离性优化，对于操作系统内核资源隔离性的研究较少。现有内核资源管理方法没有充分考虑容器的高部署密度等特点，导致容器场景下内核资源的竞争较为激烈，隔离性严重不足，影响了容器的运行。

现有内核管理机制难以保证容器有序、合理、安全的使用内核资源，使得容器系统面临着内核锁资源竞争较激烈、消耗型文件系统资源分配不公平、系统日志隐私保护不严格等问题。

针对内核锁资源竞争较激烈的问题，提出了基于系统调用频率限制的内核锁资源隔离方法。该方法可以准确地控制容器的系统调用频率，防止容器对于内核锁资源的过度竞争，从而解决内核锁资源隔离不足的问题。针对消耗型文件系统资源分配不公平的问题，提出了基于最大配额的消耗型文件系统资源隔离方法。该方法可以为容器设定资源可用量的最大配额，确保容器不会使用超出其最大配额的资源，从而解决消耗型文件系统资源隔离不足的问题。针对容器共享系统日志造成隐私保护不严格的问题，提出了基于私有化的系统日志资源隔离方法。该方法可以为每个容器提供独立的系统日志服务，使得系统日志不在容器之间共享，从而解决系统日志资源隔离不足的问题。

基于FPGA的图计算加速器访存优化机制研究

姓　　名：王庆刚
研究方向：图计算、体系结构
导　　师：廖小飞
指导老师：郑　龙
E-mail：1245231954@qq.com
QQ：1245231954
联系电话：15207108504
毕业去向：之江实验室博士后

图计算是基于图结构进行内在因果关系挖掘和分析的使能技术，广泛渗透服务于经济发展和社会治理等重要领域。由于FPGA的可重构，高并行和高能效的特性极大地满足了图数据高效处理所需的技术要求，基于FPGA的图计算专用加速器设计成为了研究热点。然而，当前的FPGA图计算加速器研究在访存优化方面存在诸多的问题。首先，图拓扑关联关系不规则性与现有高并发计算流水线存在不适配，片上访存冲突频次高。其次，动态图边数据乱序访问频繁性与仅优化点数据乱序访问的现有片上缓存架构存在不适配，片外访存时间开销大。最后，高阶图数据随机访存模式与现有索引驱动的访存行为调度机制存在不适配，片上缓存命中比例低。

针对上述挑战，论文围绕着“无冲突、低延迟和高命中”的核心目标展开了基于FPGA的图计算加速器访存优化机制研究，具体包括以下三个方面：（1）为了解决访存冲突频次高的问题，从稀疏矩阵向量乘的角度重新审视了图计算中的数据冲突问题，观察到图邻接矩阵对角线上的边数据往往彼此独立、互不依赖。基于这一发现，提出了拓扑依赖性感知的访存冲突消解机制，通过调度无依赖关系的边数据参与并行计算实现无冲突图计算。（2）为了解决片外访存时间开销大的问题，对真实动态图更新行为进行剖析，观察到动态图更新的空间局部性。基于这一发现，提出了空间局部性感知的高效片上缓存机制，通过在片上驻留被高频访问的少量边数据，将动态图更新中出现的大多数片外访存转换为高速的片上缓存访问，从而降低片外访存开销。（3）为了解决片上缓存命中比例低的问题，对高阶图数据处理的访存模式进行剖析，观察到访存时间局部性与高阶图结构重叠特征紧相关。基于这一观察，提出了时间局部性感知的访存行为调度机制，通过充分地利用被识别的数据重用机遇来提高缓存命中率。

通信高效的隐私保护数据聚合协议研究

姓　　名：公绪辉
研究方向：并行与分布式计算、信息隐私与安全
导　　师：华强胜

指导老师：华强胜

E-mail：772738308@qq.com

QQ：772738308

联系电话：15927370301

毕业去向：暂时未定

隐私保护数据聚合（Privacy-Preserving Data Aggregation, PPDA）问题是近年来随着大数据技术的迅速发展而兴起的一个研究热点。在实际应用中，大量的数据需要被收集和聚合，然而这些数据往往包含着用户的隐私信息，例如个人身份、财产信息等。因此，如何在保护隐私信息的前提下对这些数据进行聚合成为了研究的重点。

在隐私保护数据聚合中，隐私保护数据采集、隐私保护集合并集和隐私保护无重复计数是其中三个基础且研究广泛的问题。这些问题的解决对于确保数据安全和数据精确性具有重要意义。虽然目前大部分已有协议通过使用同态加密(Homomorphic Encryption, HE)、秘密分享(Secret Sharing, SS)和不经意传输(Oblivious Transfer, OT)等密码学技术保护隐私，但是它们未能充分利用各种聚合运算的特性和相应密码学技术的优点，从而导致较高通信和计算代价。针对上述基础聚合运算问题，围绕“高效（低通信复杂度）与安全”核心目标提出了一系列隐私保护数据聚合协议。

（1）隐私保护数据采集：收集多个参与方所持数据集合的隐私数据，且相同元素不删除。隐私保护的目标是保护各个元素的源头（元素具体来源于哪方）。已有研究工作大都通信与计算开销高，或需要可信任第三方，导致安全性受损。针对该问题，基于抽样与划分技术，提出了不需要可信任第三方的(n-k)-源匿名协议，其中 n 和 k 分别表示参与方和同谋者的数量。通过理论证明，所提协议与采用可信任第三方的工作相比，仅仅增加了一个关于 n 的对数多项式的通信代价。

（2）隐私保护集合并集：收集多个参与方所持数据集合的隐私数据，且相同的元素只保留一个，即他们所持数据集合的并集。隐私保护的目标是保护各个元素的源头及其重数（在所有参与方集合所组成的多重集合中出现的次数）信息。已有研究工作大都把集合元素编码为多项式，且各集合中相同元素重复参与通信与计算过程，从而造成较高通信与计算代价。针对该问题，通过设计或（OR）扰动加密算法和使用空间高效的“布隆滤波”（Bloom Filter，BF）数据结构，提出了一个通信高效的协议。该协议的通信复杂度近乎匹配通信复杂度下界，同时该协议具有常数轮数复杂度。针对两方情况，基于 OT 协议与布隆滤波，设计了一个通信最优且避免使用计算昂贵的公钥运算的两方协议。

（3）隐私保护无重复计数：计算在多个参与方所持数据集合中不同元素的数量，即并集势。隐私保护的目标是保护各个元素具体信息（包括源头，元素本身及其重数）。现有隐私保护无重复近似计数的工作主要是基于空间高效的数据结构“Flajolet-Martin（FM） Sketches”。这些工作或存在隐私泄露风险，或其通信复杂度与通信下界仍有差距。现有隐私保护无重复精确计数的工作主要是基于集合并集协议，导致较高通信与计算代价。针对以上问题，首先设计了混合随机抽样技术，并基于该技术与 FM Sketches 提出了一个隐私保护无重复近似计数协议。之后，通过结合多重散列和混合随机抽样技术，提出了一个隐私保护无重复精确计数协议。与现有工作相比，所提两个协议的通信复杂度分别近乎匹配无重复近似和精确计数问题的通信下界。

综上所述，本文所提隐私保护数据聚合协议不仅可以保护参与方隐私，并且具有较高的

实用性和可行性，对于促进隐私保护数据聚合技术的发展和推广具有积极的参考意义。

分布式流连接系统数据分发技术研究

姓　　名：于水英
研究方向：流处理、流连接
导　　师：陈汉华
指导老师：陈汉华
E-mail：yushuiying2006@126.com
QQ：176878136
联系电话：18907141586
毕业去向：暂时未定

流连接广泛应用于提取多源流数据之间的关联信息，作为流处理系统中重要的基础性操作，成为领域的研究焦点。在分布式流连接处理系统中，数据分发策略是保证系统低处理延迟、高吞吐率和可扩展性的关键。然而，流数据的倾斜性和多源性特征给分布式流连接系统带来负载不均和可扩展性差等严重问题，进而提高处理延迟，降低吞吐率。

首先，针对流数据范围分布的倾斜性导致负载不均问题，提出了一种面向倾斜分布的流连接自适应分发策略。设计了基于收益整合的迁移评估模型，将迁移导致的分区数量变化和负载不均程度统一为逗留时间的变化，进而衡量迁移带来的收益。收益最高的迁移方案即为低迁移开销和可控制分区数量的迁移方案。

其次，针对流数据键值分布的倾斜性导致负载不均问题，提出了负载感知的流连接动态分发策略。该策略选择部分处理单元作为迁移键值的迁入处理单元，设计了迁入处理单元数量递增的负载感知探测选择方法，并通过基于排队论测量的性能评估模型来评估迁出键值对系统性能的影响。

最后，针对流数据的多源连接过程复杂导致可扩展性差问题，提出了可扩展多源流连接对称分发策略。该策略使得每个元组到达系统后可以立即与中间结果连接生成最终结果。进一步，设计了基于动态属性索引图的中间数据分发结构，以减少中间结果的数量，从而减少系统计算开销。大规模真实数据的实验结果表明，相比现有设计，以上三种技术提高了系统吞吐率，降低了处理延迟。

硕士毕业生

分布式组

基于知识蒸馏的个性化联邦学习研究

姓　　名：柏东山
研究方向：联邦学习，分布式机器学习
导　　师：金　海
指导老师：姚德中
E-mail：964270816@qq.com
QQ：964270816
联系电话：18868105895
毕业去向：阿里云计算有限公司

个性化场景中的联邦学习面临着数据异构和模型异构的挑战。现有方法未深入分析客户端初始化策略对个性化性能的影响，也未深入挖掘异构客户端之间的相似性。

为了解决上述挑战，提出了基于知识蒸馏的个性化联邦学习算法。针对个性化场景中的数据异构问题，提出了客户端侧自蒸馏算法pFedSD。该算法首次探究了联邦学习初始化策略在个性化场景中引起的客户端知识遗忘现象，并保留历史个性化模型来指导客户端更新。pFedSD算法有效缓解了遗忘现象，实现了客户端个性化模型的隐性集成。该算法虽然一定程度上缓解了数据异构问题，但依然受限于模型同构。

更进一步地，针对个性化场景中的模型异构问题，提出了服务器侧注意力集成蒸馏算法pFedAD。该算法在服务器的模型融合阶段使用

集成蒸馏来代替传统联邦学习中的参数平均，从而突破模型同构的限制。为了更细粒度地传递全局知识，客户端给予其他客户端不同的注意力，从而加强相似客户端之间的协作。为了个性化算法的快速验证和公平比较，实现了一个通用的个性化联邦学习系统来集成上述算法。

在多个公开数据集上的测试结果表明，pFedSD算法比目前最优的方法准确率最高提升4.46%，收敛速度提升5倍，pFedAD算法比主流的模型异构算法准确率最高提升2.41%。

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Master Degree in Engineering（基于容器的层共享微服务放置策略研究）

姓　　名：陈子睿
研究方向：云原生优化、边缘计算
导　　师：陈汉华
指导老师：顾　琳
E-mail：751005427@qq.com
QQ：751005427
联系电话：18652638359
毕业去向：交通银行股份有限公司

在边缘平台中提供基于轻量化容器的微服务成为了边缘计算的发展趋势。然而，边缘平台中有限的资源成为了普及边缘应用与增长服务规模的瓶颈。因此，如何最有效地利用边缘资源、提升服务效率成为近年来的主要研究热点之一。

针对上述挑战，考虑边缘资源的分散性与有限性，设计利用容器层共享的微服务放置策略，提升资源利用率和边缘吞吐率。首先分析在边缘服务器间实行层共享的可行性与必要性，进一步挖掘微服务在边缘放置过程中利用层共享优化存储资源使用率的可能性。其次，分析基于容器的微服务放置问题的理论场景，实现联合微服务放置与层存储问题数学模型的建立。而后，证明联合微服务放置与层存储问题的NP难特性，将模型由整数线性规划问题转化为松弛后的线性规划问题，设计基于层共享的联合微服务放置与层存储算法，最终获得性能优异且具有良好近似比的微服务放置决策结果。

最终通过实验，在现实应用场景下，分别验证基于层共享的联合微服务放置与层存储算法在小规模实验与大规模实验中的有效性与稳定性。该算法和当前最先进的两种算法相比，可以分别多放置6倍和12倍微服务，并且边缘请求吞吐率分别提高了27.74%和38.46%。

X-Stream: 面向视频流分析任务隐私保护的灵活自适应视频转换器

姓　　名：冯　豆
研究方向：边缘计算、视频流分析隐私
导　　师：刘方明
指导老师：刘方明
E-mail：821878167@qq.com
QQ：17786520904
联系电话：17786520904
毕业去向：奥可纳软件技术(上海)有限公司

视频流分析存在巨大应用前景的同时也具有严重的隐私暴露风险。为了缓解隐私安全问题，提出了X-Stream，为海量的视频流分析任务提供便捷的隐私保护接口。X-Stream主要提出了三点创新性设计：（1）声明性查询接口，其为用户提供简单但表达力强的接口来描述隐私保护和视频流分析任务需要保留的视觉信息；（2）视频上下文感知机制，一种基于当前视频上下文动态选择最合适的隐私保护技术及其参数的自适应机制；（3）高效的执行引擎，结合了多任务重复数据消除和帧间推理的优化。实现了X-Stream隐私保护系统，并利用真

实的数据集进行了测试，主要测试了其在两大主流视频流分析任务下的隐私保护表现，分别是异常行为检测和背景提取。结果表明，与最先进的技术相比，X-Stream具有高易用性的同时，显著提高了隐私保护质量和性能。

面向容器的CPU调度虚拟化研究

姓　　名：李星儒
研究方向：云计算与虚拟化、容器、调度系统
导　　师：吴　松
指导老师：吴　松
E-mail：1021551729@qq.com
QQ：1021551729
联系电话：15629109360
毕业去向：阿里云计算有限公司

应用通常可以通过对调度系统中的策略和参数进行个性化配置来提高其执行效率，而容器为了保证运行过程中的安全性和隔离性，在默认情况下仅允许应用使用单一策略且不可配置参数。容器环境下应用面临着安全和性能的权衡问题。

为解决该问题，面向容器的CPU调度虚拟化方案在宿主机CPU调度器上为每个容器虚拟化一个独立的、可配置的CPU调度器，其中用全局调度器和局部调度器来区分描述宿主机上和容器内的CPU调度环境。它基于容器的两层调度模式实现，在全局调度器中使用默认的完全公平调度器来保证容器间运行的安全性和隔离性，同时为局部调度器提供独立的任务调度方式和参数视图来实现容器内调度的灵活性。因此面向容器的CPU调度虚拟化方案实现了容器内应用可以通过对其局部调度器上策略和参数的配置来提升执行效率，同时保证了这种配置不影响全局调度器及其它容器的局部调度器，减少了容器环境下的性能损失。

基于hackbench的性能评估表明，面向容器的CPU调度虚拟化实现了容器间良好的隔离性，并满足了容器内应用的定制化需求。其中测试结果指出相较于默认环境，面向容器的CPU调度虚拟化能使得容器分别从策略和参数的优化配置中获得54%和17%的性能提升。

图式区块链连续型事务处理方法研究

姓　　名：泥俊沛
研究方向：区块链
导　　师：胡　侃
指导老师：肖　江
E-mail：junpeini@qq.com
QQ：553166577
联系电话：13125133600
毕业去向：深圳嘉易科技创新有限公司

图式区块链在连续型事务处理中存在两个关键问题：1）无需的事务执行和存储，导致连续型事务无法正确执行。2）事务一致性难以保证，导致连续型事务之间可能产生冲突。

针对以上问题，提出了一种面向图式区块链的高效事务处理方法FLUID。为支持连续型事务的按序执行，FLUID设计了基于事务依赖跟踪结构的存储模型，强制连续型事务按照依赖关系进行提交和处理，从而保障连续型事务处理的顺序执行和存储。为保证连续型事务处理的一致性，FLUID提出一种基于可信共识组的冲突解决机制，通过投票获取轮换的可信共识节点集合，在事务提交前监听并校验事务有效性，在提交前解决潜在冲突，有效降低并发事务的确认延迟。同时，FLUID提出一种校验和机制，并建立了一个基于检查点的连续型事务一致性验证机制，在任何节点可通过校验和高效地校验账本数据之间的一致性，矿工节点将账本中的最新校验和打包成检查点事务，并基于检查点在可信共识组范围内实现对账本中事务处理结果的确定性共识。

为了证明FLUID处理连续型事务的有效

性，在广域网真实环境中对FLUID系统进行了实验评估。结果表明，FLUID相比于目前最先进的图式区块链存储模型OHIE，可提升66%的吞吐量并将延迟降低两个数量级。

面向异构环境的高效联邦学习算法研究

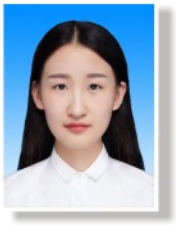
姓　　名：潘婉宁
研究方向：联邦学习
导　　师：丁晓锋
指导老师：姚德中
E-mail：834234014@qq.com
QQ：834234014
联系电话：17371266472
毕业去向：抖音视界有限公司

联邦学习是一种基于边缘计算的新型分布式深度学习范式，相较于传统的分布式深度学习，联邦学习能更好地保护用户隐私。然而，联邦学习的发展面临两个关键的挑战：（1）系统异构性挑战，即各客户端设备具有不同的计算和网络资源，而联邦学习通常假设各客户端训练相同的模型，导致服务提供商需要在模型复杂度和可访问的数据量之间权衡。（2）数据异构性挑战，即各客户端的数据分布不一致，导致模型在本地训练时产生偏移，降低了模型的收敛速度和准确度。

为了应对系统异构性挑战，提出了一种基于低秩压缩的异构联邦学习框架。该框架在服务器端将满秩的全局模型预先分解，得到不同规模的混合网络结构低秩模型，分发给不同客户端进行低秩优化训练。训练后的模型传输回服务器，被还原为统一大小的稀疏大模型，最终通过参数平均实现异构模型聚合。这一框架对资源异构的设备灵活配置不同复杂度的模型，可以降低对计算和网络资源的严格要求，从而提升模型的训练效率和准确度。为了应对数据异构性挑战，提出了一种基于知识蒸馏的联邦学习优化算法，通过将全局模型的历史知识蒸馏给本地模型，在特征空间缓解本地模型训练的偏移问题，提升了模型收敛速度和准确度。

通过在多个数据集和多种异构设置下的算法性能评测表明，所提出的压缩框架支持算力和网络资源异构的设备加入训练，在小模型的表现上，相较于现有最先进的异构联邦学习框架最高提升了6.93%的准确度。在数据异构场景下，相较于经典联邦学习算法，所提出的优化算法最高提升了5.62%的准确度。最后，通过集成异构压缩框架和优化算法实现了异构联邦学习系统。在数据异构和系统异构同时存在的场景中，相较于同构训练小模型，所提出的异构联邦学习系统最高降低了67.6%的通信开销。

基于自适应权重的图式区块链网络传输方法研究



姓　　名：司冰茹
研究方向：区块链网络
导　　师：陈汉华
指导老师：肖　江
E-mail：bingru1023@163.com
QQ：921667635
联系电话：18234637116
毕业去向：中国船舶集团有限公司第七0九研究所

由于传统链式区块链在高并发应用场景中存在性能不足的局限性，基于有向无环图的图式区块链受到学术界和工业界的广泛关注。图式区块链网络传输存在两个关键问题：第一，现有传输方法未对单个区块的传输进行优化；第二，现有传输方法未考虑由于区块并发传输造成的账本不一致问题。

针对以上问题，提出一种基于自适应权重的图式区块链网络传输方法SplitTrans。为了加快区块传输速度，SplitTrans将区块头和区块体

的广播过程解耦，设计了基于区块头优先传播的区块传播机制，节点无需等待完整区块即可立即进行转发。为了优化邻居选择，节点根据区块头的传输延迟和父区块的上链情况评估邻居节点的性能，并利用响应数据计算邻居节点权重。此外，为了实现高区块写入率和低延迟，SplitTrans设计了基于节点自适应权重的伪随机邻居选择方法，通过选择一定数量的高权重邻居优化区块体广播，同时随机选择少量邻居提高网络的安全性。

区块链网络节点规模庞大，难以在实际网络环境中验证传输方法的有效性和可行性，因此设计并实现了一个图式区块链网络仿真工具BDAGsim。BDAGsim采用多层架构，用于SplitTrans和Gossip的模块化实现。仿真实验结果表明，与现有的Gossip算法相比，SplitTrans将区块传输时间降低了60.2%，具有良好的网络可扩展性和稳定性。

面向可移植字节码模块的Serverless函数调度系统

姓　　名：宋　振
研究方向：云计算
导　　师：刘方明
指导老师：刘方明
E-mail：galaxy0305@qq.com
QQ：1355594513
联系电话：15827417738
毕业去向：支付宝（杭州）信息技术有限公司

近年来，大量研究工作将计算和数据密集型应用迁移到Serverless平台上进行并行计算。然而，现有的函数执行机制难以提供足够的函数并行度以及启动和扩展能力。此外，函数之间依赖于远程存储服务的状态交互模式将会产生过量的数据移动。

针对上述问题，引入了可移植字节码模块技术。为了提高函数启动性能的同时减少其资源占用，提出了一种新型的函数执行机制WaFunc，在原生的沙箱环境下进行高效的资源隔离和共享，并在此基础上实现了函数快照和线程调度等特性。同时以WaFunc为基础，基于数据本地性优先的调度策略和迁移优化方法设计并实现了一个高效的Serverless函数调度系统WASH，该系统提供了：（1）简洁易用的上层编程接口，提高用户的开发和使用效率；（2）细粒度的函数隔离和执行，增加函数的并行度和扩展度；（3）高效合理的函数调度，减少应用工作流执行期间的数据移动。

在对系统进行了充分的实验评估之后，结果表明，WASH提供了简洁易用的用户编程接口。同时，相对于现有的主流Serverless方案，WASH在占用更低内存的情况下具有19.3%~32.9%的函数启动延迟性能提升，提高了23.9%~48.1%的系统整体吞吐率，并降低了28.2%~54.8%的数据移动开销。

基于虚拟内核的容器内存管理策略定制化方法

姓　　名：王洪磊
研究方向：容器、内存管理
导　　师：吴　松
指导老师：吴　松
E-mail：321582424 @qq.com
QQ：321582424
联系电话：17633908764
毕业去向：北京志凌海纳科技有限公司

在云计算多租户环境下，同一宿主机上容器应用的内存使用模式不尽相同，共享的宿主机内存管理策略无法满足不同容器应用在不同场景下对透明大页和虚拟内存管理策略的定制化需求，从而严重影响容器应用性能。

针对容器应用共享内核与定制化需求的矛盾，提出了基于虚拟内核的容器内存管理策略定制化方法，在保证容器轻量化特性的前提

下，支持容器内存管理策略的定制化。首先，针对容器共享内核造成无法定制内存管理策略的根本问题，设计了基于Rust内核模块和轻量级重定向机制的虚拟内核框架，支持为容器构建包含私有内核代码与数据的虚拟内核，以低开销的方式实现容器和宿主机内核的解耦。同时，针对不同场景下容器应用内存管理策略定制化需求，在虚拟内核框架中基于完全隔离原则和投票机制设计了可定制的容器透明大页和虚拟内存管理策略。最后，设计了内存管理策略自适应组件，支持为容器动态调整内存管理策略，从而进一步提高容器应用的性能和稳定性。

实验测试通过为容器应用定制透明大页管理策略，容器密集型内存访问性能提升230%；通过为容器应用定制虚拟内存管理策略，容器应用吞吐率最高提升24%。

边缘场景下基于模型拆分和批量推理的端边协同推理方法研究

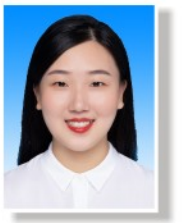
姓　　名：王千千
研究方向：边缘智能、端边协同
导　　师：余　辰
指导老师：余　辰
E-mail：540436756@qq.com
QQ：540436756
联系电话：17671788544
毕业去向：蚂蚁科技集团股份有限公司

深度学习与边缘计算技术的发展孕育出了海量的边缘智能应用。面向用户的任务关键型边缘智能应用需要低延迟推理来保障高质量服务。

为了解决边缘服务器负担大量模型计算任务时的过载与带宽不足的问题，基于延迟反馈机制和上下文相关的置信区间上界算法设计了两种模型拆分方法，对推理任务各拆分方案的端到端执行延迟进行实时预估和最优决策。实现了动态运行环境中端边之间计算负载的高效分配，降低了推理任务的端到端执行延迟。为了缓解密集任务场景下边缘服务器的过载问题，在协同推理的基础上提出了在边缘服务器上对拆分任务序列进行批量推理的方法。该方法以松弛预估的推理任务执行延迟为基础，设计了一种服务质量感知的任务合并调度算法，通过对拆分任务序列的并行批处理缩减了任务的平均计算开销。

对比实验结果表明，在各种边缘环境下，协同推理方法的系统吞吐率分别较本地推理和计算卸载方法平均提升了2.37倍和1.12倍；在三种级别的服务水平协议设置下，协同推理方法较计算卸载方法平均提升了25.31%；在动态变化的网络带宽和边缘服务器负载环境下，协同推理方法表现出了良好的鲁棒性。验证了在端–边计算结构下，协同推理方法有效提升了深度学习模型的推理速度与系统任务吞吐率。

面向分布式仓库的容器镜像拉取策略研究



姓　　名：许鸿皓
研究方向：边缘计算、云原生优化
导　　师：顾　琳
指导老师：顾　琳
E-mail：812486914@qq.com
QQ：812486914
联系电话：13007160762
毕业去向：国家电网天津电力分公司

面向分布式仓库的容器镜像拉取策略需要考虑仓库的地理分散性、存储内容差异、拉取负载平衡等多个复杂因素。因此，综合分布式仓库的特性和容器固有的分层结构，设计了以层为单位的新型容器镜像拉取策略。同时，充分利用容器间层共享特性，进行分布式仓库镜像的去重与并行拉取，显著减少拉取流量，提

高了存储资源利用率，增加了边缘吞吐量。首先详细分析了本地层共享和邻居层共享两种层共享机制，验证其在容器拉取与微服务部署问题中的可行性。然后，将研究问题构建为整数规划形式的容器拉取与微服务部署数学模型，以最大化边缘吞吐量为目标，建立完善的资源约束和逻辑约束。接着，通过推理证明了该问题的NP难复杂度以及近似子模特性，由此提出了一种基于迭代贪婪的镜像拉取算法，在满足所有约束的前提下，每轮作出可以获得当前最大吞吐量的决策，并提供了具有近似比的理论性能保证。通过大量基于真实容器公开数据集的实验，验证了不同场景与不同数据规模下算法所具有的优良性能。实验结果表明，与现有的两种先进技术相比，该算法将边缘吞吐量分别提高了38.47%和44.95%，同时在边缘多部署了58.98%和66.68%的微服务。

固态硬盘公平性事务处理方法研究

姓　　名：叶贻良
研究方向：云计算与虚拟化、I/O、存储
导　　师：石宣化
指导老师：吴　松
E-mail：2948278083@qq.com
QQ：2948278083
联系电话：15827082733
毕业去向：阿里云计算有限公司

在共享环境下，由于现代高性能固态硬盘（Solid-State Drives, SSD）可以并行处理多个负载，并发负载间公平地使用SSD资源成为I/O系统的重要保障。如果I/O系统缺乏公平性控制机制，那么负载间将由于相互干扰而引发公平性问题。然而，现代SSD为了提高性能而绕过I/O软件栈，导致传统的公平性控制机制失效。这给I/O系统的公平性带来巨大的挑战。

目前，有众多研究工作在SSD内部提供了公平性事务调度，但是这些工作受到当前SSD处理事务框架的限制，仅能提供有限的公平性保障，并产生了不可忽略的性能损失。一方面，芯片队列间的调度结果相差较大，难以保障总体公平性；另一方面，忽略了事务并行执行条件以及后端存储的有限并行能力，导致调度决策无法有效执行。为此，设计了公平性保障增强的事务处理方法。该方法包含二阶段地址映射和并行能力感知的事务派发控制。二阶段地址映射的第一阶段旨在分别对负载进行轮询分配，尽可能在芯片队列间维持相同的调度结果；第二阶段进行全局轮询分配，并根据公平性选择地开启晶粒级隔离，以满足事务并行执行条件；并行能力感知的事务派发控制基于负载的特征进行调度，并与第二阶段地址映射协作，共同实现SSD并行能力的高效利用。该方法包含三个子步骤，事务先由第一阶段地址映射确定通道和芯片地址，之后进行并行能力感知的事务派发控制，最后再由第二阶段地址映射确定晶粒和面地址。

实验结果表明，相比于现有的公平性控制机制，该方法平均提升了96%的公平性；相比于面向并发的事务调度，该方法平均仅降低了0.4%的资源利用率。

面向图式区块链的高性能共识机制研究

姓　　名：岳镜涛
研究方向：图式区块链、共识协议
导　　师：陈汉华
指导老师：肖　江
E-mail：824502870@qq.com
QQ：824502870
联系电话：15922788037
毕业去向：北京抖音信息服务有限公司

基于有向无环图的图式区块链可支持并行区块，有望突破传统链式区块链仅支持串行区

块造成的吞吐量瓶颈，因此受到学术界和工业界的广泛关注。但现有图式区块链共识机制仍存在以下问题。首先，现有共识机制的各个组件高度耦合，导致图式区块链系统的调整和优化困难，难以灵活适配多元化需求。其次，现有共识机制基于概率性共识协议，在保证安全性的同时增加了延迟，难以有效支持延迟敏感型应用。

针对以上问题，提出了基于模块化设计思想和混合共识协议的高性能图式区块链共识机制ElasticDAG。为了灵活适配多元化应用需求，ElasticDAG将图式区块链共识机制解耦为相互独立的存储机制和确认机制，并为两者设计了通用的交互模式，从而实现一个模块化共识机制。为了降低图式区块链的确认延迟和客户端的存储开销，ElasticDAG提出了一个可支持轻节点的图式区块链快速确认机制，包括适配图式区块链的低延迟混合共识协议和交易合法性证明机制。此外，为了在实现顺序一致性的同时提升图式区块链的吞吐量，ElasticDAG采用了可避免抛弃区块的高活性图式区块链存储机制。

为了验证ElasticDAG设计的有效性，在真实的广域网环境下对ElasticDAG共识机制进行了测试。实验结果表明，ElasticDAG的确认延迟为32秒，每秒可确认2874笔交易。相比于目前最先进的图式区块链共识机制OHIE，ElasticDAG 可在不损失吞吐量的情况下降低95.6% 的确认延迟。

系　统　组

面向高速网络的IP查找算法加速器设计方法研究

姓　　名：陈佳杰
研究方向：高性能计算、网络
导　　师：廖小飞
指导老师：廖小飞
E-mail：983411610@qq.com
QQ：983411610
联系电话：15827135430
毕业去向：阿里云计算有限公司

近年来，随着网络链路速度呈现爆炸式增长，面向高速网络的报文处理加速算法 受到广泛的关注。而IP查找算法作为报文处理流程中最典型的查表算法之一，如何支持400Gb/s高速网络的性能需求成为重要问题。与此同时，核心网和数据中心的网络路由转发规模也在飞速增长，IP查找的路由表存储容量占用问题也成为IP查找算法落地实现需要考虑的重要因素。

Ip查找加速器在现场可编程逻辑门阵列(FPGA)平台通过定制计算来加速算法的处理速度。首先在算法层面上采用多叉前缀树和位图的方法，设计高效的适用于硬件实现的算法。通过优化算法架构，减少单次IP查找的计算和存储开销，高吞吐量以支持高速网络查表需求。并通过解耦合的计算逻辑和访存逻辑，以及高效握手协议和多级队列的方法来实现一个高查找效率的硬件流水线结构。此外，为了应对未来路由表规模高速增长的趋势，在基于片上存储器的部署方案的基础上开发了基于片外存储器的加速器，并对两种方案在IP查找场景下的存储容量和访存带宽进行了分析，讨论未来IP查找算法加速器存储器部署的性能瓶颈和发展趋势。

基于依赖感知的并发图处理加速方法研究

姓　　名：姜新宇
研究方向：图计算、并行分布式计算
导　　师：王多强
指导老师：张　宇
E-mail：270336962@qq.com
QQ：270336962
联系电话：13762769230

毕业去向：华为技术有限公司武汉研究所

为了提高分布式图处理系统的并发图处理性能，现有的并发图处理解决方案采用数据共享的方法来减少系统的数据访问开销。但是，由于并发图分析任务的数据访问行为具有强烈的不规则性，现有方法往往难以高效实现并发任务间的数据共享，从而导致系统的并发图处理性能依旧较差。为解决上述问题，设计实现了一个依赖感知的并发图处理系统TGraph。该系统能够与现有的分布式图处理系统进行集成，减少系统执行并发图处理时产生的数据访问与通信开销，从而有效提高系统的并发图处理性能。与现有方案不同，TGraph设计了一种依赖感知的并发图分析任务同步执行模型。该模型根据图数据间固有的依赖关系来调度并发图分析任务同步执行，使其数据访问行为更加规则，从而高效实现并发图分析任务间的数据共享。此外，TGraph还提出了一种面向并发图处理的负载均衡策略，通过预测每个图分析任务的具体负载来更加精准地实现计算节点间的负载均衡。最后，TGraph还实现了一种基于联合通信的消息管理机制，通过联合并发图分析任务间的状态同步通信来有效减少系统的通信成本。

基于FPGA-HBM的BFS算法加速器优化方法研究

姓　　名：李可欣
研究方向：高性能计算
导　　师：郑　然
指导老师：邵志远
E-mail：kexin.li_cecilia@qq.com
QQ：309073331
联系电话：18361335632
毕业去向：华中科技大学攻读博士

广度优先搜索（Breadth First Search，BFS）算法是一种基础的图计算算法，随着数据规模和业务需求的不断扩大，BFS算法的效率成为一个重要研究课题。配置了高带宽内存（High Bandwidth Memory，HBM）的现场可编程门阵列（Field Programmable Gate Array，FPGA）拥有可编程、低功耗、高带宽特性，能够为BFS算法提供加速。

设计BFS算法加速器Hi-BFS，在ScalaBFS的基础上从HBM访问、计算模块处理和布局布线三个方面提出加速器性能提升优化方法。具体来说，设计填充压缩稀疏矩阵行/列的高密度图数据存储结构，匹配HBM提供的大位宽；同时构建独立的高频HBM读取模块，提高访问HBM的效率。对片上存储设备进行细粒度访问与粗粒度访问的隔离，优化算法执行过程中的数据通路。此外，还对电路实现阶段的布局布线方案进行优化，提高电路频率，以此提高加速器的整体性能。

测试结果表明，基于FPGA-HBM平台的Hi-BFS加速器具有良好的可扩展性，能够充分利用HBM所有的内存通道，在赛灵思U280加速器卡上每秒遍历的边数最高可达75.34千兆，与ScalaBFS相比提升了2.53~3.82倍。与运行在英伟达A100 GPU上的经典图计算加速器Gunrock相比，Hi-BFS取得了1.34~3.17倍吞吐量和7.38~17.30倍能源效率的加速。

动态图神经网络推理的硬件加速技术研究

姓　　名：潘婷婷
研究方向：嵌入式系统、硬件加速
导　　师：金　海
指导老师：叶晨成
E-mail：1741947285@qq.com
QQ：1741947285
联系电话：15211049045
毕业去向：之江实验室

针对动态图的图神经网络推理过程通常需

要处理连续时间点的多个动态图，会带来巨大的计算复杂度和高昂的时间成本，限制了动态图神经网络在实际应用中的可行性和效率。

为了解决这一问题，提出动态增量更新计算模型。捕获不同时间点的动态图之间的变化信息，然后将其快速地传播到受影响的节点上，实现局部的信息更新和特征提取，从而更加精细地控制动态图的更新和信息传递过程。通过增量更新计算模型，可以有效地减少动态图神经网络的计算复杂度和计算时间。为了适应动态图数据集的海量规模以及单个时间点极稀疏变化的性质，在动态图神经网络的推理过程中，实现高效的适应动态图的压缩矩阵行存储格式CRR，在节省图数据集存储空间的同时还支持不同计算方法的切换。此外，还根据不同动态图的变化规模以及动态图中不同数据区域的数据分布预分析结果，设置适应性的分块大小，以减少DRAM访问，进一步提高计算效率。

在五个实际应用场景中的数据集上的性能模拟实验数据表明，动态增量更新计算模型与静态计算模型相比，可大幅减少计算操作数量。动态图神经网络加速器DyGA与图神经网络加速器GCNAX相比，性能得到了有效提升，并减少了DRAM访问次数。

FPGA-Based Acceleration Architecture for Apache Spark Operators（基于分布式FPGA架构的Spark算子优化方法研究）

姓　　名：孙元伟
研究方向：大数据、分布式
导　　师：张　宇
指导老师：刘海坤
E-mail：1027414152@qq.com
QQ：1027414152
联系电话：15245459320

毕业去向：之江实验室

传统以 CPU 为核心的内存计算平台 Spark，通过并行扩展计算节点的方式来提高系统的处理速度，在现如今海量数据环境下无法满足应用对计算响应时间的需求。目前的解决方案是采用 CPU+GPU/FPGA/ XPU 的异构计算架构取代 CPU 处理核心，从而加速每个节点的计算速度。但是在数据混洗的过程中，主机端与 FPGA 设备之间数据路径存在冗余，主机之间采用 TCP/IP 通信的方式存在着不必要的数据拷贝。分布式应用开发人员利用高层次综合工具（High Level Synthesis，HLS）进行计算与网络集成时存在着资源与算法并行度不协调的问题。

为解决上述问题，设计实现了一个基于 Spark 的分布式异构计算架构。在该架构中，节点间数据混洗（Shuffle）过程被卸载到 FPGA 硬件设备上，单机节点计算得到的中间数据结果将直接发送给远端节点做数据聚合，解决了单机环境下数据路径循环的问题，缩短了分布式环境下数据传输的关键路径。同时将混洗过程中的 TCP/IP协议栈替换为 RDMA 协议栈，进一步减少通信过程中的数据拷贝过程。此外，设计了 Spark 算子的 FPGA 硬件实现优化方法，包括指令级深度流水线、函数级流水线以及独立读写的数据分区等。基于优化方法重构了经典 Spark 算子，提出了算子分布式实现通用解决方案，通过硬件数据分区与软件逻辑分解一一映射的方式提高算子并行度，减少 FPGA 硬件资源使用。

实验结果中，基于设计架构下的各算子（K-means、PageRank 和 Bitonic Sort）对比原生 Spark 系统分别获得了 3 倍、3.5 倍和 3.5 倍的加速效果，同时，随着数据集的增大，执行时间增长缓慢，体现了系统的可扩展性。实验结果表明，设计的分布式 CPU+ FPGA 架构对于未来计算中心的设计具有参考意义。

基于 GPU 的图神经网络训练加速方法研究

姓　　名：吴敏康
研究方向：高性能计算、体系结构
导　　师：邵志远教授
指导老师：邵志远教授
E-mail：kyle_email@qq.com
QQ：865495620
联系电话：15327129856
毕业去向：摩尔线程智能科技（北京）有限责任公司

近年来，图神经网络获得了广泛的关注和应用，如何加快图神经网络的训练速度，是当前计算机系统领域研究的热点问题。目前已经有大量工作致力于通过加速聚合阶段来加速图神经网络训练，但现有的工作对图邻接矩阵内的所有数据都采用同一种方法进行计算，没有注意到矩阵内数据分布可能对性能造成的影响。

为加速图神经网络训练，聚焦于在图形处理器平台上，加速聚合阶段的稀疏矩阵乘法。具体来说，利用矩阵内部的数据分布差异，对同一矩阵的不同部分采用不同的计算方式完成稀疏矩阵乘法，从而达到提升整体计算性能的目标。首先在小规模矩阵上研究稀疏矩阵乘法加速，提出了一个二分计算模型，并验证其在计算加速上的有效性。进一步地，对规模更大的真实图上的数据分布进行分析，结合真实图数据的特点对二分计算模型进行扩展和改进，提出适用于真实图邻接矩阵上的混合计算模型，以加速图神经网络训练中的聚合阶段。同时，提出了与混合计算模型相配套的图分割方法。

实验结果表明，所提出的方法在小规模稀疏矩阵上进行稀疏矩阵乘法时，能够达到现有方法1.42~3.56倍的加速比；在图神经网络训练中，聚合阶段相较于GPU上现有的图神经网络训练系统GNNAdvisor，能够达到1.51~3.47倍的加速比，图神经网络整体训练性能相较GNNAdvisor能达到1.13~1.37倍的加速比。

动态图处理硬件加速机制研究

姓　　名：杨　赟
研究方向：图计算、体系结构
导　　师：姚德中
指导老师：张　宇
E-mail：1348046572@qq.com
QQ：1348046572
联系电话：15029991528
毕业去向：华为技术有限公司

为分析现实世界中随着时间不断变化的图数据，动态图处理被广泛应用。现有动态图处理方案，如 KickStarter、GraphBolt、DZiG，通常采用增量计算的方式，通过修正最近一次图快照的计算结果来获得最新图快照的计算结果。然而，由于受到图更新影响的图顶点状态值通常会沿着图拓扑结构不规则地传递给其它图顶点，现有方法仍面临着严重的冗余计算开销和不规则内存访问行为，致使现有处理器架构的硬件资源难以有效利用。

为了解决现有方法存在的问题，可编程硬件加速器 TDGraph 被提出来，它采用拓扑驱动的增量执行机制来加速动态图的处理，从而实现更加规则化的状态传递和更好的数据访问局部性。首先，为了解决冗余计算开销问题，加速器实现了拓扑驱动的遍历规则化单元，其以受图更新影响的图顶点为根，沿着图拓扑结构预取其它图顶点，并在运行时同步它们的增量计算。通过这种方法，多个受图更新影响的图顶点状态传递可以沿着图拓扑结构一起进行，从而有效减少了冗余计算和数据访问成本。其次，为了解决图顶点状态值的不规则内存访问问题，加速器实现了图顶点状态合并单元，通过将频繁访问的图顶点状态值数据访问行为进行动态合并，以此提高数据访问局部

性，进而提高缓存和内存带宽的利用率。

基于量化更新价值的高效多版本图计算系统

姓　　名：叶翔宇

研究方向：图计算

导　　师：廖小飞

指导老师：郑　龙

E-mail：xyye@hust.edu.cn

QQ：982582989

联系电话：15856883884

毕业去向：支付宝（杭州）信息技术有限公司

大数据时代下，图计算在许多领域得到大规模应用。现实中图的动态演化会产生一系列不同的历史版本，在这些版本上执行图计算的过程被称为多版本图计算。由于图演化在实际问题中的普遍性，多版本图计算的性能优化问题引起研究者广泛关注。多版本图计算的过程通常包括三个步骤：选择时间点进行图版本采样；切换到指定时刻的图版本；在该图版本上执行图计算。通过分析现有多版本图计算系统的工作模式，归纳出当前研究存在的一个普遍假设：对某个时刻的图版本请求，在该时刻都存在历史上被采样的版本快照。然而上述假设缺乏对现实需求的考量，即目标版本在历史上没有被采样。这种情况下，现有系统只能切换到该时刻附近被采样的版本，并在该版本上执行图计算。上述现象导致图计算结果出现不可控的误差，严重降低多版本图计算系统的实用性。针对上述问题，通过对真实动态图数据的实验分析观察到：在图演化过程中，只有少数的图更新会对图计算结果产生较大的影响。基于上述结论，提出了一个新颖的多版本图计算系统，该系统引入一个新的概念“更新价值”，用以描述图更新对图计算结果的影响程度。在此基础上，提出了有效的更新价值量化方法，实现了基于更新价值的图更新数据差分管理，有效提升了图计算结果的准确性。进一步，在更新价值量化划分和更新数据差分管理的基础上，提出了基于价值划分的更新单元查询映射和局部性感知的定制化多级动态缓存结构，有效提升了系统的计算效率，并显著降低了额外存储开销。在现实世界图数据集上的实验结果表明，该系统在图计算结果准确性方面比现有先进系统提升了74.35%~95.72%，在内存开销方面减少57.03%~90.44%，为此引入的额外计算时间不到2.96%。

面向分离式内存系统的智能网卡

姓　　名：于泽霖

研究方向：分布式内存系统、智能网卡

导　　师：刘海坤

指导老师：刘海坤

E-mail：1034346166@qq.com

QQ：1034346166

联系电话：15827561060

毕业去向：渤海银行股份有限公司

分离式内存系统可以提供高容量、高带宽、低能耗的共享内存服务，已在数据中心中广泛应用。当前，智能网卡作为广泛应用于分离式内存/存储系统任务卸载的新兴硬件，存在着硬件资源受限、存储访问开销大、任务卸载机制低效等多方面的问题，极大的限制了分离式存储系统的整体性能。

为了解决智能网卡存在的问题，面向分离式存储系统的智能网卡（TIGA）从传输和任务卸载两个方面对现有分离式架构中的智能网卡进行针对性的设计。TIGA通过定制的硬件设计和高效的任务分发策略来解决分离式存储系统中智能网卡卸载存在的问题。首先，TIGA通过修改传统智能网卡的硬件布局并设计了配套的软件栈，以消除冗余数据通路和软件栈开销，有效的降低了卸载过程中的数据传输开销。同

时，TIGA根据当前的运行监测数据动态调整计算资源在各卸载任务间的分配，以使智能网卡的总卸载吞吐最大化。最后，TIGA在智能网卡负载饱和时将额外负载交给内存节点的弱CPU，以进一步提高整体性能。实验结果表明，相较于最新的智能网卡任务卸载框架FairNIC 和iPipe，TIGA有效地提升了分离式存储系统的性能并降低了构建成本和能耗，为建设高效节能的数据中心提供了参考实现。

大 数 据 组

增量和减量最优间隔分布学习机的研究

姓　　名：陈丽君
研究方向：机器学习、增量学习
导　　师：张　腾
指导老师：张　腾
E-mail：643336157@qq.com
QQ：643336157
联系电话：18890027383
毕业去向：华为技术有限公司

增量和减量学习是在数据变化时高效更新模型的有效手段，现有研究主要针对具有线性损失函数的大间隔模型。而最新研究表明以最优间隔分布学习机（Optimal margin Distribution Machine, ODM）为代表的新型模型通过优化间隔分布取得了更好的泛化性能，目前仍缺乏对这类模型进行增量和减量学习的研究。

为填补此空白，针对单个训练样本变化的场景，提出单样本增量和减量最优间隔分布学习机（Single Incremental and Decremental ODM，SID-ODM）以在数据变化时高效地更新模型，避免从头开始训练。增量和减量更新关键在于通过KKT条件构建原始样本与变化样本之间的等式关系，从而能够实时更新ODM的最优解与分类超平面。其中，ODM的二次损失函数导致其拉格朗日乘子无边界，难以追踪。为此，提出了一种确定无界拉格朗日乘子更新方向的方法，进而不断迭代直至更新到最优解。在多个训练样本变化的复杂场景下，SID-ODM失效，考虑所有变化样本对最优解的影响，设计多样本更新策略，并提出更通用的增量和减量最优间隔分布学习机（ID-ODM）。为进一步提高算法效率，基于Sherman-Morrison-Woodburry公式实现了矩阵逆的快速更新。此外，相关的理论分析证明了ID-ODM 的收敛性。

实验结果显示在所有数据集上，ID-ODM较重新训练的方法的更新速度至少提升4倍，平均加速比达9.1倍，且几乎没有泛化损失。

基于日志结构合并树的键值数据库性能优化研究

姓　　名：贺　康
研究方向：键值存储、日志结构合并树
导　　师：华强胜
指导老师：石宣化
E-mail：363352333@qq.com
QQ：363352333
联系电话：15827075119
毕业去向：阿里巴巴达摩院（杭州）科技有限公司

日志结构合并树（Log-Structured Merge-Tree，LSM-tree）作为键值存储数据结构，广泛应用于分布式存储系统。LSM-tree合并过程会抢占 CPU 资源，导致写暂停。

针对合并过程导致的写暂停问题，提出一种定变长融合键值长度LSM-tree 森林数据库系统（FeasDB）。FeasDB 包含一个动态键值长度选择器、多个统一键值长度LSM-tree，以及一个变长键值长度LSM-tree。统一键值长度存储可在后台合并过程中实现键值对的批量处理，从而缩短时间消耗。针对长度分布范围集中的键值对，FeasDB 通过采用多个统一键值长度的LSM-tree 组合存储方案，在提高系统整体性能的同时，尽量降低空间开销。对于极短或极长

的键值对，FeasDB 使用变长键值长度的 LSM-tree 进行处理。动态键值长度选择器的主要功能是根据键值长度将其分配到相应的 LSM-tree 进行处理。

FeasDB降低了 LSM-tree 后台合并过程的 CPU 开销，合并效率提升了1.4至1.9倍。在处理大规模数据的高并发场景时，FeasDB与RocksDB相比，其写性能最高提升可达3.7倍。此外，在模拟多尺寸键值长度负载的实验中，FeasDB以牺牲10%的额外平均空间开销为代价，使系统整体写性能提升了1.2至2倍。

高效可扩展的图流概要结构研究

姓　　名：蒋志国
研究方向：流数据处理系统、概要结构设计
导　　师：陈汉华教授
指导老师：陈汉华教授
E-mail：15200947573@qq.com
QQ：3365596503
联系电话：15200947573
毕业去向：华为技术（深圳）有限公司

图流是一种描述实体间实时关联的流数据模型，具体表现为一个连续到来的边 流序列，在大数据应用中被广泛使用。图流的巨大规模和极快的更新速度对数据存储 结构提出了严格的要求，包括亚线性的空间成本、计算高效的操作支持和高可扩展 性。因此，概要存储结构成为一种实际可行的技术手段。现有图流概要结构主要使用 基于哈希且大小固定的压缩矩阵，实现了对规模预知的图流数据的高效存储。然而大 小固定的压缩矩阵无法支持规模动态扩展的图流数据的高效管理。 针对现有图流概要结构不可扩展的问题，提出了时空高效扩展的图流概要结构 ——Auxo。Auxo 基于二分搜索及公共二进制前缀嵌入方法设计了一种高效可扩展的指纹前缀嵌入树结构。通过将指纹的前缀信息嵌入在二分树型扩展结构中，一方面省略了前缀信息的存储开销，将内存成本降低了O(log|E|)级别。另一方面在查询时，嵌入树型结构中的前缀信息还可以作为索引降低查询操作的时间开销，将查询开销降低到了O(log|E|)级别。进一步，为了解决树型结构指数扩展时内存效率波动的问 题，提出了比例渐近扩展策略，该策略使得 Auxo 以一种渐近的方式逐层生成新的一 层树型结构。 实验使用了真实系统的大规模图流数据集来测试 Auxo 的性能。实验结果表明， 与目前最先进的工作相比，Auxo 将各项操作的时间开销降低了一到两个数量级，同时节省了平均 15%的存储开销。比例渐近扩展策略将 Auxo 的内存利用率进一步提 高了 20%以上，实现了真实系统 80%以上的平均内存利用率。

基于霍克斯过程的时序异质信息网络嵌入研究

姓　　名：史瑞泽
研究方向：社交网络分析、数据挖掘
导　　师：黄　宏
指导老师：黄　宏
E-mail：924458793@qq.com
QQ：924458793
联系电话：15629109030
毕业去向：华中科技大学攻读博士

近年来，由于异质信息网络嵌入的优异性能，它已成功应用于许多现实任务。然而，以往的研究大都关注于静态异质信息网络，或将时序异质信息网络按照时间戳进行离散化处理，无法完全捕捉异质信息网络的动态演化过程，这降低了异质信息网络嵌入的实际应用价值。

因此，基于霍克斯过程的时序异质信息网络嵌入算法（THINE算法）被提出，用于建模时序异质信息网络的演化过程、保留其中的全

部动态信息。THINE算法首先利用基于元路径的随机游走理解异质信息网络中的静态结构和静态语义，进而结合霍克斯过程对网络的时序演化过程进行模拟，并建模了历史事件的衰减效应，以维护网络中不断变化的动态信息。此外，THINE算法还应用了语义级别和结构级别的注意力机制，其中一个用于捕捉不同元路径的语义，另一个则区分了不同邻居节点的距离。因此，THINE算法能够考虑异质信息网络演化过程中的所有时序信息，并获得多类型节点之间的影响，有助于理解和分析现实世界中的时序异质信息网络。

在三个真实世界的时序异质信息网络上的广泛评估表明，THINE算法在两个静态任务（节点分类、链接预测）和一个动态任务（时序链接推荐）中的表现都优于以往的基线方法，且性能提升高达25%。

基于语义增强的时序知识图谱补全模型研究

姓　　名：魏昊辉
研究方向：知识图谱、自然语言处理
导　　师：肖江
指导老师：黄宏
E-mail：1192899669@qq.com
QQ：1192899669
联系电话：13338156177
毕业去向：江苏省某机关

时序知识图谱的补全工作极大缓解现实中时序知识图谱所面临的稀疏、不完整问题。现有工作受限于单一知识的内部交互导致的局部性特征感知，无法全面考虑图谱的结构信息和时序依赖之间的关系，另一方面知识图谱的隐性语义信息被极大的忽略。

针对以上问题，基于时序感知的语义增强补全方法被提出，模型最大限度地获取时序知识图谱结构信息和时序信息，通过信息融合理解时序知识图谱中的显性语义信息，同时搭建自注意力机制框架封装时序知识图谱内的丰富邻域信息，以实现实体的时序感知；另一方面构建适配的时序感知解码器，在时序知识预测过程中激活实体 表征的时序知识完成实体关系间的语义交互，提升解码能力。为了捕捉时序知识图谱中更多的隐性语义关系增强实体表征语义表达能力，自适应构建动态超图以完成时序知识图谱中深层语义的生成，获取非既定关系决定的实体间信息。在两种不同角度对时序知识图谱进行表示学习后，通过表征融合可以获得基于时序依赖以及隐性语义生成的实体关系一致性表征，利用解码器完成时序知识图谱上的知识补全。

GPU加速的神经网络训练运行时系统关键技术研究

姓　　名：吴文超
研究方向：深度学习系统、异构计算
导　　师：金　海
指导老师：石宣化
E-mail：1366199330@qq.com
QQ：1366199330
联系电话：13297096614
毕业去向：阿里巴巴

深度神经网络作为人工智能的重要支撑，在现实领域中取得了重要进展和广泛应用。随着 GPU 新的硬件特性不断推出，现有的神经网络训练运行时系统无法充分将网络负载的计算特征与 GPU 底层硬件特性进行有效适配，存在并行效率低、负载不均衡严重、多任务协调难等问题，造成训练效率低、GPU 资源利用率差，严重制约了模型开发和工业化落地。鉴于此，论文以负载特性为指导，充分挖掘深度神经网络模型的并行性，规避负载的不均衡性、探索多任务的协调性，系统耦合应用特征与

GPU硬件特性，填补上层负载和底层硬件之间的语义鸿沟，具体贡献为：

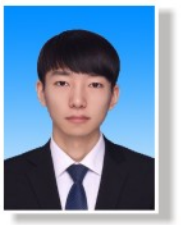
（1）提出了基于关键路径的细粒度并行策略，将负载特征、软件多流、硬件并行进行系统性结合，有效地减少了训练迭代时间和提升了GPU资源利用率。

（2）提出了图结构指导的负载均衡方法，进行选择性的数据放置、度数指导的任务分配，和多阶段流水线调度，从而最大化访存效率、并行效率、实现多粒度负载均衡。

（3）设计了一个面向多任务训练的高效GNN系统，提出了一种细粒度、通用、高效的跨任务算子融合方法，消除了任务之间冗余的图数据访问开销，大幅度提升了kernel性能，实现了无序、高竞争资源的合理利用。

面向隐私保护的学习型索引及空间范围查询方法研究

姓　　名：张　亮
研究方向：隐私保护、查询处理
导　　师：蒋文斌
指导老师：丁晓锋
E-mail：wezanglang@163.com
QQ：1076403280
联系电话：18291917335
毕业去向：中国农业银行（陕西省分行）



当前，随着物联网的快速发展和各类移动互联设备的日益普及，空间范围查询技术被广泛地应用到各类复杂的查询需求中。然而在查询处理的过程中，用户的数据隐私面临严重威胁。面向隐私保护的空间范围查询技术常采用数据加密技术，但是加密并不能够完全解决隐私问题。攻击者通过对加密数据的访问模式、搜索模式的分析可以获取数据隐私。同时，数据加密也带来了存储与计算上的性能损耗，使得查询处理与数据共享变得困难。

针对上述问题，提出了一种安全的学习型索引SDI以及基于SDI的空间范围查询算法SDBRQ。首先，SDI基于特殊的空间曲线编码对空间数据进行映射，并在映射空间上采用轻量级的线性分段模型对数据分布进行学习；SDI针对经过映射的空间范围查询的特点，对学习模型的作用域以及预测类型进行扩充；同时，SDI通过二叉树的形式组织预测器来加速预测过程，并采用提出的安全索引访问算法CAP，通过半随机的方式遍历SDI。其次，SDBRQ采用CAP算法预测查询对应的数据区域，并在该区域内通过盲取的方式获取查询结果，进一步保护了数据的访问模式。最后，提出了基于密文压缩的优化方案用于提升查询效率，降低通信代价，并形式化地分析了索引构造与所提算法的复杂度和安全性。

最后，基于C++实现了索引SDI以及查询算法SDBRQ。在两个真实世界数据集CAR、GOWALLA以及一个合成数据集SYN上，评估了SDI在不同参数设置下的构造时间与存储开销，验证了索引构造的可行性；接着，评估了不同参数对于SDBRQ查询效率的影响。在默认参数设置下，SDBRQ在CAR上的最短查询时延达到了2.49s，在GOWALLA上则达到了2.72s，经过优化后，其查询时延降分别降低至2.75s和0.79s，对比同类工作，该结果表明了所提算法在效率方面的有效性。

代码预训练语言模型的可解释性结构化方法研究

姓　　名：赵　清
研究方向：自然语言处理、代码智能
导　　师：郑　龙
指导老师：万　瑶
E-mail：1395943316@qq.com
QQ：1395943316
联系电话：15674816619

毕业去向：成都联想信息技术有限公司

代码预训练模型在可解释研究方面进展甚微，尚不清楚这些模型为何起作用以及它们可以捕获哪些代码相关特征。这会降低用户对预训练模型的信任度，阻碍模型在实践中的进一步发展。针对代码预训练语言模型缺乏可解释性的问题提出了三种基于代码语法信息的可解释性分析方法，旨在从三个不同的角度详细分析代码预训练模型是否捕获了代码语法信息：（1）提出了一个分析模型自注意力分布的方法。该方法通过计算与抽象语法树结构对齐的注意力所占比例，探讨了代码预训练语言模型中注意力如何与源代码中的语法信息对齐。（2）提出了一个分析模型词向量的结构探针。该方法通过计算模型词向量映射到语法空间中词与词之间的距离，探讨了代码预训练语言模型是否隐式地嵌入了代码语法信息。（3）提出了一种基于代码词素之间的语法距离来归约语法树的方法。该方法根据语法距离指定拆分点的选择顺序，以自上而下的方式递归地划分输入以归约抽象语法树结构。通过对比归约树和标准语法树之间的相似度，从而判断模型能多大程度捕获语法信息。

基于时序支持度的高阶时序社区快速搜索方法研究

姓　　名：朱春雪
研究方向：图数据库、时序社区挖掘
导　　师：袁平鹏
指导老师：袁平鹏
E-mail：3133166408 @qq.com
QQ：3133166408
联系电话：18729667384
毕业去向：西安欧珀通信科技有限公司

图通常随时间演变，发现高阶时序社区结构可以为好友推荐、广告投放等一系列实际应用提供理论及方法支撑，具有重要意义。为探索高阶时序社区，首先将静态truss模型扩展到时序图中，时序三角形的概念被提出，并以此为基础来定义了极大–δ–truss（Maximal-δ-truss, MDT）高阶时序社区。为搜索特定顶点周边的极大–δ–truss社区（Searching MDT for q, q-MDT），提出了目标点驱动的脉动式局部搜索方法（Local search, LS）。LS先以查询顶点q为中心向外扩展得到时序子图，并在此子图上搜索目标社区。根据搜索结果决定是调整扩展参数继续搜索还是直接返回当前结果作为目标社区。为了进一步提升在大型复杂时序网络搜索q-MDT的性能，采用自底向上的方法增量式构建了时序–truss（Temporal-based truss, TBT）索引。基于此TBT，提出了更加高效的索引辅助的高阶时序社区搜索方法（Index search, IS）。IS以TBT索引为指引可快速搜索出目标顶点所在的MDT。

为评估算法的效率和有效性，在9个真实世界数据集上进行实验。在与其他时序社区搜索模型的对比实验中，验证了LS和IS方法的效率。参数敏感性测试和可扩展性测试证明了IS算法的稳定性。与其它模型的有效性对比实验中，q-MDT模型在两个有效性指标上分别排名第一和第二。基于真实背景的案例分析也证明了q-MDT模型的实际价值和意义。

基于中间数据复用的深度学习预处理加速方法研究

姓　　名：朱占洋
研究方向：深度学习、大数据
导　　师：金　海
指导老师：石宣化
E-mail：752858854@qq.com
QQ：752858854
联系电话：18801356519
毕业去向：阿里云计算有限公司

深度学习在多个轮次的训练中需要进行反复的数据预处理，包括数据读取，解压缩，数据增强等等步骤，以提高模型训练效果。然而，这些复杂的数据预处理计算正在逐步成为深度学习模型训练过程中主要的瓶颈之一。

因此，多模型数据加载器（Multi-Model Data Loader，MMDL）被提出，以针对现有的模型训练加速工作所普遍存在的仅专注于单一模型加速的不足，面向多模型并行训练任务，实现数据预处理步骤的加速。MMDL通过在服务器级别向同时运行的多个不同的模型训练任务提供统一的数据预处理服务，使用分支流水线实现多个模型间读取以及预处理之后的数据的共享，减少重复读取和计算。同时，MMDL对经过部分预处理的数据进行缓存，用于后续轮次中读取和使用，从而减少了需要的预处理操作数量。此外，MMDL使用自动化的流水线构建与跨模型的数据批生成等方法实现了更高效率的模型间数据共享。因此，MMDL实现了多模型并行训练场景下系统整体的开销降低和预处理加速。

MMDL实现的加速效果依赖于预处理工作流的设置，随着模型间相同的预处理操作增加而提升。在使用ImageNet数据集同时训练三个模型的情况下，对比各自独立进行的数据预处理，MMDL实现了最差1.6倍和最好3.15倍的加速效果，证明了MMDL能够显著降低预处理开销，加速预处理过程。

安　全　组

健壮的低存储开销可搜索对称加密研究

姓　　名：但镇武
研究方向：可搜索加密
导　　师：徐　鹏
指导老师：徐　鹏
E-mail：913347780@qq.com
QQ：913347780
联系电话：18971253331
毕业去向：湖北省某保密单位

动态可搜索对称加密能够使隐私数据安全地存储在外包服务器上，通过对称加密支持关键字安全搜索并允许用户动态地添加删除文件。目前有客户端存储和健壮性两个衡量方案性能的重要指标。低客户端存储让用户在资源受限的设备上能进行大规模更新与搜索操作，健壮性针对重复添加或删除不存在的关键字–文件对等误操作情形仍能保证安全性和正确搜索结果。然而低客户端存储和健壮性目前仍无法兼容。为了适应更加丰富的应用场景，构造一个安全、高效、低客户端存储和健壮的方案迫在眉睫。

为了解决上述问题，针对低客户端存储与健壮性的融合展开研究。针对低存储开销方案效率低，基于比传统不经意数据结构更加高效的映射方式将客户端状态外包存储到服务器上，实现低客户端存储，提高更新和搜索效率。针对目前融合低客户端存储与健壮性的矛盾，构造出能快速检测出误操作的警报检查机制。通过该机制，客户端安全检索存储在服务器上的本地状态，高效精准地识别出误操作的所有情形，对其操作进行混淆，保证方案的健壮性。结合满二叉树构造出基于树状结构的关键字搜索协议，实现拟最优关键字搜索复杂度，同时限制误操作产生的冗余数据不影响实际搜索性能。融合上述具体方法，构造出低客户端存储与健壮性相融合的DSSE方案MASE，并证明其满足对应的安全性。

在方案MASE性能测试阶段，与Orion、Mitra*、QOS和Horus*等利用不经意密码原语实现低客户端存储的方案进行比较，进行了健壮性测试与更新–搜索对照实验。在随机误操作模拟实验中，只有方案MASE满足健壮性，返回

了正确的搜索结果。相比其他方案，方案MASE在搜索阶段最高快14280倍，在更新阶段最高快12倍。此外，在大规模删除情形下，方案MASE的搜索优势更加明显。

高安全与健壮的可搜索对称加密研究

姓　　名：窦昊辰
研究方向：动态可搜索对称加密、前向安全、后向安全、健壮性
导　　师：徐　鹏
指导老师：徐　鹏
E-mail：530317348@qq.com
QQ：530317348
联系电话：18189299202
毕业去向：CISPA亥姆霍兹信息安全研究中心攻读博士

随着云计算与大数据技术的不断发展，海量数据的安全存储与检索已经成为当前研究人员亟待解决的重要课题之一。动态可搜索对称加密技术(Dynamic Searchable Symmetric Encryption，DSSE)的提出，使得用户可以根据自身需求添加或删除已经加密上传的数据。如果客户端发起搜索请求，服务器可以直接以密态形式检索数据。这样，在保证用户数据安全性的同时，也极大地提升了搜索效率。

目前在所有的DSSE方案中，FB-DSSE方案实现了目前最强的安全等级：前向和Type-I-后向安全。然而，在FB-DSSE和其他的大部分方案中，一旦客户端无意中发出不合理的更新请求，那么这一方案便无法保持其原本所声称的搜索结果正确性和安全性。这一性质被叫做DSSE的健壮性。目前，完全健壮的DSSE方案只实现了最低的安全等级：前向和Type-III后向安全。因此，如何在保证DSSE方案的高安全性下实现健壮性，就是本文的核心内容。

本文所提出了两个新的DSSE方案（分别命名为SR-DSSEa和SR-DSSEb）。在文件搜索聚合方面，两个方案在完成搜索后，采用了一种基于逻辑电路形式的聚合算法，使得其在面对客户端发出不合理的请求时，均仍能保证搜索结果的正确性。不仅如此，两个方案分别采用了不同的安全计算策略，由此达到了DSSE的高安全性。在实验测试方面，SR-DSSEa比SR-DSSEb具有更有好的的客户端开销，更低的通信带宽和交互轮次，适用于本地计算能力较弱的场景。而在整体搜索性能上，SR-DSSEb具有比SR-DSSEa更好的表现，且已经接近现有的FB-DSSE方案，适用于对搜索效率有较高要求的场景。

基于行为图的智能漏洞检测增强技术研究

姓　　名：方一霖
研究方向：软件工程、软件安全
导　　师：袁　斌
指导老师：袁　斌
E-mail：qzfyl98@qq.com
QQ：935474206
联系电话：18957020938
毕业去向：俄亥俄州立大学攻读博士

随着深度学习技术的快速发展，越来越多的研究者开始将深度学习技术应用到源代码漏洞检测中，检测效果较传统的基于静态分析的方法有了显著提升。但是这些检测系统检测的视角都局限在了函数个体上，没有关注函数之间的关联信息，实际上存在漏洞的函数往往存在一定的关联性，从而导致检测系统在人工构建的数据集中效果好，但却在真实数据集上效果差。针对这些问题，主要研究内容和贡献如下：

1）提出了软件漏洞检测中的行为图概念。已有的漏洞检测方法把检测的视角局限在了函数个体上，而忽略了函数之间可能存在的联系关系，从而导致在人工数据集上效果较好，但

是在真实数据集上检测效果不理想。为了发现和利用这种关系，提出了行为和行为图的概念，论述了借助行为图可以高效的发掘函数之间的联系、函数行为和函数安全性的可行性。

2）开发了一个基于行为图的增强智能漏洞漏洞检测框架VulBG。框架由行为图构建、行为图特征提取和模型融合三个阶段构成。行为图构建阶段主要通过对程序源码通过函数行为特征提取、特征聚类分析和行为相识度等来构成行为图；行为特征提取阶段通过使用结点嵌入技术将函数结点转化成特征向量，然后通过一个多层感知器来得到有效的函数行为特征；在模型融合阶段，将行为特征和传统的漏洞检测模型提取的特征进行拼接，然后通过一个分类器进行漏洞判定。通过在两个真实数据集中进行的实验表明，采用VulBG框架的漏洞检测准确率有显著提升。

3）对行为图与VulBG的有效性进行全面的评估。通过对不同种类的数据集来验证评估行为图和VulBG在漏洞检测中的性能，使用VulBG对其他基准模型带来地性能改善也进行了评估，对VulBG中的关键模块的选择和重要参数的设置对性能的影响也进行了讨论。结果表明VulBG不仅自身就可以优秀地胜任漏洞检测任务，还可以显著地增强其他基于深度学习的漏洞检测系统的检测能力。

基于知识图谱的软件供应链安全分析研究

姓　　名：贾亚茹
研究方向：知识图谱、软件供应链
导　　师：马晓静
指导老师：马晓静
E-mail：409321342@qq.com
QQ：409321342
联系电话：13930185698
毕业去向：中建三局信息技术研究院

近年来，随着开源软件的复用不断增多，针对软件供应链攻击的事件也愈加频繁，软件供应链的安全分析及维护受到越来越多学者的关注。然而，组件依赖的开源数据庞大复杂，导致供应链构建过程复杂，难以进行全面的分析。知识图谱灵活、丰富、可扩展的特性，解决了处理大规模开源数据的难点。目前基于知识图谱的软件供应链安全研究中，直接使用推理功能获得的信息冗杂，缺乏进一步数据过滤及关键信息提取。由此导致安全分析结果不够准确、修复流程过于复杂的问题，不利于维护供应链安全。

基于此，针对Java生态提出了供应链知识图谱推理结果处理方法，通过过滤冗杂信息及提取安全风险组件信息获得更加准确的数据。首先，采集并处理多源漏洞和开源组件数据，构建包含漏洞、组件两大实体的供应链知识图谱。然后，查询得到漏洞、组件依赖以及漏洞传播路径等供应链知识图谱推理结果。最后，结合噪声依赖关系过滤方法、冗余依赖关系过滤方法过滤供应链知识图谱推理结果中的冗杂信息，并对过滤后漏洞传播路径中的组件进行滞后库预测，得到存在安全风险组件的相关信息。

所构建的供应链知识图谱包含了千万规模的漏洞和组件信息，具有较好的可视化展示效果。通过实验测试，证明了供应链知识图谱推理结果处理方法可以有效过滤推理结果中60.4%的冗杂影响组件以及86.2%的冗杂传播路径，预测存在安全风险组件的准确率达到80.6%。最后，以一条具体漏洞传播路径为例给出了软件供应链安全分析及维护流程。供应链知识图谱推理结果处理方法为后续基于知识图谱的软件供应链安全分析奠定了数据基础，有利于保障软件供应链安全。

基于过程间分析的源代码漏洞检测技术研究

姓　　名：李　妍

研究方向：漏洞检测、机器学习
导　　师：羌卫中
指导老师：羌卫中
E-mail：861280680@qq.com
QQ：861280680
联系电话：15827315891
毕业去向：中国农业银行

近年来，由软件漏洞引发的安全事件层出不穷，及早发现并修补漏洞能够有效降低损失。传统的基于规则的漏洞检测方法依赖于专家定义规则，存在较高的漏报率，基于深度学习的方法能够自动学习漏洞程序的潜在特征，然而随着软件复杂程度的提升，该类方法在面对真实软件时存在精度下降的问题。一方面，现有方法在执行漏洞检测时大多在函数级工作，无法处理跨函数的漏洞样例，且无法识别具体的漏洞触发位置；另一方面，常用的深度学习模型不擅长捕捉程序语句中的长期依赖关系，图神经网络传播节点信息的深度受限于模型层数，循环神经网络模型在输入序列过长时性能会下降。

为了提升基于深度学习的漏洞检测器的实用性，实现了一个基于过程间分析识别复杂漏洞的细粒度漏洞检测系统。首先，在源代码的预处理方面，优化了现有的程序切片算法，结合过程内和过程间切片对跨函数的漏洞进行全面的上下文分析以捕获漏洞触发的因果关系。其次，在漏洞检测的粒度方面，分析总结了13种漏洞类型的sink语句特征，在程序切片中识别潜在的漏洞触发位置，实现了精确的漏洞定位。最后，在深度学习模型的设计方面，应用了包含多头注意力机制的Transformer神经网络架构，共同关注来自不同表示子空间的信息来提取节点的深层特征，解决了传统神经网络模型信息衰减的问题，能够更有效地学习程序语句之间的长期依赖关系，识别潜在的漏洞模式。

在来自linux kernel、openssl和wireshark等17个开源软件的漏洞数据集上开展实验，与现有的静态漏洞检测器进行对比，实验结果表明该方法在真实软件数据集上的F1分数达到了70.7%，相较于对比方法提升了9.7%~39.0%，并成功检测出多个开源软件漏洞，证明了其有效性与实用性，验证了系统方案设计的合理性。

面向多方协同的链上数据修复机制

姓　　名：刘烨凯
研究方向：区块链、变色龙哈希
导　　师：代炜琦
指导老师：代炜琦
E-mail：762384204@qq.com
QQ：762384204
联系电话：15827351787
毕业去向：中兴通讯股份有限公司（深圳）

区块链因其支持公开透明的数据交换和不可篡改的数据记录而在物联网、数字版权等众多领域广泛应用。然而，公开透明和不可篡改的特性可能被滥用，导致谣言或其他误导性信息在链上传播，而这些违规信息却无法被删除。因此，需要一种链上数据修复方案来修复违规信息。目前已有的链上数据修复方法包括硬分叉、合约覆盖和基于变色龙哈希算法的可编辑区块链。但这些方法在实际应用中存在一些限制，例如时间成本过高、无法完全删除数据、需要管理算法陷门、修复操作难以审计等。

针对上述问题，提出了一个面向多方协同的链上数据修复机制（On-chain Data Repair Mechanism for Multi-party Collaboration, ODRMMC）。该机制基于变色龙哈希算法，具有公开算法陷门、易于审计修复操作的特点，并且能够实现数据完全删除。首先，提出一种新型结构的区块链作为基础架构用于实现数据修复机制。其次，提出一个基于合约的锁定投票方

案，用于支持链上协同审计投票。再次，提出一种区块增加的数据修复机制，该机制结合投票的结果，能轻易审计和记录每次修复操作的内容。最后提出一个交易完整性验证算法，以验证未修复交易和已修复交易的正确性。

此外，在以太坊上实现ODRMMC作为原型进行评估。实验表明，通过公开陷门优化变色龙哈希函数后性能有显著提升。ODRMMC具备适应不同变色龙哈希算法的能力。性能测试结果表明，添加模块的总时间消耗在毫秒级，对原始系统的性能影响可以忽略不计。同时，评估结果还表明，启动特殊交易的成本与正常以太坊交易的消耗相当，并且在可控范围内。

针对恶意人脸感知模型的隐私保护技术研究

姓　　名：刘晓耕
研究方向：可信人工智能
导　　师：胡胜山
指导老师：胡胜山
E-mail：liuxiaogeng@hust.edu.cn
QQ：2673723046
联系电话：18674003851
毕业去向：威斯康星大学麦迪逊分校攻读博士

尽管深度人脸感知系统在身份识别和验证方面表现出惊人的性能，但近年来，有报告指出人脸感知模型存在被滥用的情况，这些智能系统会被利用并对诸如社交媒体等网络平台上的用户进行过度监控。该问题引发了学术界与产业界对于这项技术在人脸图像隐私方面的担忧。最近，一些研究指出，采用对抗样本技术可以对人脸图像施加保护，从而让其免受未经授权的人脸感知系统的识别。然而，现有的生成人脸对抗样本图像的方法存在许多局限性，如视觉效果差、白盒访问权限的依赖性高、对抗可迁移性弱等缺点，使该项技术难以真正地应用于社交平台上的人脸隐私保护这一场景。

针对上述问题，提出了一种新颖的人脸图像隐私保护方法，即基于生成网络妆容迁移的人脸对抗样本算法AMT-GAN，该方法旨在构建具有更强迁移性和更好视觉质量的人脸对抗样本图像，同时摆脱了对于目标模型的白盒访问依赖。AMT-GAN采用了生成式对抗网络（Generative Adversarial Network, GAN）作为核心架构，能够将指定参考图像的妆容迁移到用户人脸图像的同时，将该人脸图像转化为对抗样本图像。具体而言，在妆容迁移对抗网络的基础上，引入了一个新的正则化模块，并采用联合训练策略来平衡对抗性噪声和妆容迁移中的循环一致性损失之间的冲突，实现攻击强度和视觉变化之间的理想平衡。

大量实验验证，与现有技术相比，AMT-GAN不仅可以保持令人满意的视觉质量，而且可以在包括旷视科技Face++、阿里云和微软Azure等商业人脸感知编程接口（Application Programming Interface, API）中实现更高的黑盒攻击成功率，这意味着在实际使用中，AMT-GAN能够具有更强的人脸隐私保护效率，大大增加了基于对抗样本的人脸隐私保护技术在现实场景下的实用性。

基于梯度矫正的联邦学习拜占庭鲁棒性研究

姓　　名：陆建荣
研究方向：联邦学习
导　　师：胡胜山
指导老师：胡胜山
QQ：2195953269
联系电话：15085693861
毕业去向：攻读博士

联邦学习（Federated Learning，FL）能够在不泄露边缘客户端隐私数据的情况下协作客户端学习得到一个知识丰富的全局模型。然

而，联邦学习极容易受到拜占庭攻击的影响，即攻击者可以通过制造有毒的本地模型来破坏全局模型的性能。许多工作致力于设计鲁棒的聚合算法来抵抗拜占庭攻击，但他们都无法在数据高度异构的场景下为联邦学习提供有效保护。此外，最先进的防御依赖于一个与客户端具有类似分布的私有数据集，这违反了联邦学习的隐私保护原则。因此，设计一个适用于高度数据异构场景下的联邦学习防御方法仍然具有挑战性。

鉴于此，本研究旨在弥补联邦学习防御算法和其实际应用之间的差距。首先，通过深度分析现有防御方案的局限性和失效原因，提出了一个理解毒性攻击如何影响联邦学习的通用框架。基于理论的启发，进一步提出了一种新的适用于数据高度异构场景下的联邦学习防御方案：HeteroFL。HeteroFL的核心思路是压缩恶意梯度的生成空间并最后矫正他们的方向。具体来说，HeteroFL结合了四个简单但互补的防御模块，依次压缩恶意梯度的合法方向空间，随后利用矫正性聚合纠正聚合的模型，使其在遭到攻击的情况下向全局最优更新方向移动。为了进一步评估其鲁棒性和有效性，利用了当前最强的攻击框架来设计一个专门针对HeteroFL的自适应攻击方案，同时评估了三种情形下自适应攻击有效性和可检测性对HeteroFL的影响。

在四个高度异构的数据集上进行了大量的实验，结果表明，HeteroFL在击败各种拜占庭攻击方面优于所有先进的防御措施。具体而言，HeteroFL在不同数据集和不同模型上实现了与基线相似的全局模型准确率，而现有防御措施的全局模型准确率反而下降了34%到79%。

基于符号执行的车联网终端固件漏洞检测系统

姓　　名：陆弈帆
研究方向：系统安全、软件安全
导　　师：邹德清
指导老师：袁　斌
E-mail：842825214@qq.com
QQ：842825214
联系电话：13511636369
毕业去向：华为技术有限公司

随着车联网技术的飞速发展，车辆终端的网联化与智能化水平快速提升，车辆系统更高的开放性与复杂性在提高驾乘体验的同时也带来了网络安全风险。网联汽车的网络安全问题将导致严重的生命安全与社会安全问题，亟需对车辆终端进行安全测试以提高其安全性。然而，现有的固件漏洞检测技术在车辆终端的电子控制单元（Electronic control unit，ECU）上存在适配不佳、效率低下与检测不全面的问题，无法有效地进行大规模、自动化的固件安全测试。

针对以上问题，提出的车联网终端固件漏洞检测系统SEAR提供自动化的ECU固件预处理功能，并针对符号执行的路径爆炸问题优化测试效率，支持对多种类型漏洞进行检测。为解决符号执行对ECU固件的适配问题，SEAR利用自动化的固件解析与反编译器的静态分析对不同硬件与架构的ECU提供广泛支持，并针对固件优化了符号执行的状态建模。SEAR基于Binaryninja的中间语言进行模拟执行，针对漏洞检测任务改进了符号化地址的具体化策略，并设计了循环摘要技术以提高符号执行的效率。此外，SEAR将符号执行与污点分析相结合，并设计了一系列的漏洞模型以实现全面的漏洞检测。

实验证明SEAR具有更高的符号执行效率与准确的漏洞检测能力。在真实场景下对两款CANBOX固件的测试中SEAR发现了7处越界访问与1处整数溢出漏洞，在对OpenHarmony操作系统的LiteOS-A内核的测试中发现了3处内存信息泄露漏洞，在向厂商报告后得到3个CVE编号。

基于图卷积神经网络的抗混淆安卓恶意软件检测

姓　　名：齐　蒙
研究方向：软件安全、深度学习
导　　师：邹德清
指导老师：邹德清
E-mail：113231780@qq.com
QQ：113231780
联系电话：13125089896
毕业去向：华为技术有限公司杭州研究所

安卓操作系统作为移动设备上最广泛使用的开放源代码操作系统，具有着丰富的应用市场与客户人群，但其普及率与快速的增长也为恶意软件的发展提供了温床。在当前已被研究者提出的恶意软件检测技术中，忽视程序语义而直接提取简单程序特征的方法检测速度快但精确度不理想，将程序语义转换为图模型并采用图分析的方法精确度高但开销大且扩展性低。因此，设计出一种精度高、检测速度快且最好能抗混淆的安卓恶意软件检测方法是一个非常重要的研究课题。

针对以上问题，SriDroid采用了图卷积神经网络，实现了对安卓恶意软件的抗混淆检测。首先，通过静态分析提取出安卓应用程序的函数调用图从而保留应用的程序语义，然后采用抽象API技术将函数转化成对应的类、包或者家族，以此构建成更为简单的抽象图从而减少运行开销并增强鲁棒性。基于得到的抽象图，构建基于图卷积神经网络的分类器，通过聚合邻居特征提升检测的准确性，并通过引入Triplet Loss对比损失使得SriDroid在抵抗代码混淆方面具有更强的鲁棒性。

经过对20246个安卓应用的实验分析，实验结果表明SriDroid可以达到99.17%的恶意软件检测精确度，并在对抗代码混淆的实验测试中达到97.4%的召回率。相较于现有的安卓恶意软件检测方法，SriDroid具有最高的检测精度、最快的检测速度和最好的抗混淆能力。

面向多维传播模式的开源软件漏洞管理研究

姓　　名：石　建
研究方向：软件安全、漏洞检测
导　　师：邹德清
指导老师：邹德清
E-mail：2441255356@qq.com
QQ：2441255356
联系电话：13006376656
毕业去向：国家互联网应急中心（待定）

开源软件中的漏洞伴随开源软件的使用而广泛传播，对软件安全影响深远。随着开源软件的普及和复杂度的不断提高，软件漏洞管理变得越来越重要。围绕开源软件中的漏洞传播问题，在不同维度对开源软件中的漏洞进行管理。

在面向软件演化的漏洞传播方面，提出一种管理软件演化中漏洞的方法。定义漏洞的属性和安全性指标，以评估软件的安全性；针对漏洞属性缺失的问题，设计确定漏洞引入时间的算法。应用漏洞管理方法对OpenBSD和Firefox进行实验分析，证明了该方法的有效性。实验结果表明，软件的安全性在软件的演化过程中不断提高。

在面向代码克隆的漏洞传播方面，提出自动化的克隆漏洞管理方法，使用空间克隆关系图和时间克隆关系图的概念描述软件系统中的克隆状况。其中，空间克隆关系图用于描述程序之间的克隆关系，时间克隆关系图用于描述程序中克隆随时间的演变，并将克隆管理和漏洞管理紧密的联系为一个整体，实现了克隆漏洞的追踪。使用克隆漏洞管理方法对8个Ubuntu版本进行实验分析，以追踪代码克隆中的漏洞。实验结果表明，克隆普遍存在，约六分之一的代码是克隆的；有漏洞的克隆和无漏

洞的克隆寿命相似。

在面向开源组件依赖的漏洞传播方面，提出基于知识图谱的开源软件漏洞管理方法。设计基于Simhash的文件依赖识别算法，识别修改的开源组件；提出文件特征索引的开源组件仓库构建方式，识别文件的来源；构建了C/C++ 开源软件的漏洞知识图谱管理软件依赖关系以及漏洞与软件的关联关系，追踪漏洞在组件依赖中的传播。在Github C/C++开源软件数据集上，开源组件识别的准确率达到94%。实验结果表明，开源组件中的漏洞数量不断增长，开源组件中的漏洞随着开源组件依赖链传播。

设备驱动程序内核模糊测试语法模板生成研究

姓　　名：侍　言
研究方向：内核模糊测试
导　　师：羌卫中
指导老师：羌卫中
E-mail：395224259@qq.com
QQ：395224259
联系电话：15371365016
毕业去向：阿里云计算有限公司

覆盖率引导的内核模糊测试已经成为设备驱动程序测试的重要方法。模糊测试器使用语法模板生成有效的测试用例，以更好地覆盖设备驱动程序的代码路径。然而，现有的驱动程序语法模板数量有限，且只涵盖少量设备种类，无法完全支持设备驱动程序测试的需要。驱动程序语法模板是一组系统调用序列，其中的入口点声明是特定于驱动程序的，需要从驱动程序中获取入口点函数名称和参数类型。语法模板中的设备节点路径指定了设备驱动程序的操作目标，但设备节点路径由系统动态生成，因此无法直接获取。此外，设备驱动程序的测试需要有设备的响应，否则无法对驱动程序的功能和交互进行测试。

针对上述问题，设计了一种设备驱动程序语法模板生成方法，并实现了模拟设备来为驱动程序的测试提供支持。首先基于设备类型和规范使用静态分析提取语法模板所需的各种设备驱动程序信息，例如入口点、设备标识和设备注册函数等。其次，在QEMU仿真环境下实现了模拟设备，它们可以像真实硬件一样注册在操作系统中，并响应来自设备驱动程序的数据和控制请求。然后根据真实设备节点路径总结模式匹配的规则，针对不同类型的设备实现了静态生成设备节点路径的方法。最后根据语法模板的声明式语言的语法将设备的静态信息与设备节点路径转换为规范的语法模板，实现了驱动程序语法模板的生成。

为USB、PCI、属性设备等设备驱动程序共生成了628个语法模板，基于Syzkaller实现了添加新的语法模板和模拟设备支持的CoSyzkaller。实验结果表明，生成的语法模板符合规范，并且能够被模糊测试器正确地应用。在三个版本的Linux内核上进行测试，CoSyzkaller相比于Syzkaller平均提高了13%的代码覆盖率，并且覆盖了Syzkaller无法覆盖的驱动程序代码路径。因此，为驱动程序生成语法模板是一种有效提高设备驱动程序模糊测试效率和有效性的方法。

基于历史缺陷用例的求解器测试技术研究

姓　　名：孙茂林
研究方向：软件测试
导　　师：文　明
指导老师：文　明
E-mail：694899730@qq.com
QQ：694899730
联系电话：17681096164
毕业去向：南京大学攻读博士

可满足性模理论（Satisfiability Module Theory，SMT）求解器可以检查一阶逻辑公式的可满足性，被广泛应用于软件工程和安全实践中，如软件验证、漏洞检测和服务访问控制。隐藏在SMT求解器中的缺陷会对其下游应用产生误导，从而造成严重后果。因此，确保SMT求解器的可靠性和健壮性是至关重要的。尽管研究者已经提出了许多测试SMT求解器的技术和方法，但如何有效地发现SMT求解器中的缺陷仍然是一个挑战。此外，现有研究对于求解器缺陷管理系统中的历史缺陷触发用例关注较少。

因此，为提高求解器测试的效果，首先对开源SMT求解器的缺陷跟踪系统中曾触发过缺陷的公式进行了实证研究，观察到在历史缺陷用例中包含着大量有价值的信息，例如这些公式的骨架，即公式的核心结构，以及它们所包含的原子公式。通过研究发现这些信息是触发求解器缺陷的关键因素。基于该发现提出一种新技术，利用从历史缺陷用例中提取的骨架，并在原子公式中挖掘关联规则，以指导使用原子公式枚举骨架。将该方法实现为一个实用的模糊测试工具HistFuzz，并使用它在最先进的SMT求解器Z3和cvc5上进行广泛的测试。截至目前，HistFuzz已经为Z3和cvc5发现111个已被确认的新缺陷，其中的108个已经被开发者修复，这些数据在一定程度上说明该技术不仅具有有效性，也从实际上改善了先进的SMT求解器的可靠性和正确性。此外还通过实验表明，与其他SMT求解器模糊测试工具相比，HistFuzz在缺陷检测能力，代码覆盖率以及效率等方面均有更为出色的表现。

物联网云平台中敏感API滥用风险分析与检测

姓　　名：万　俊
研究方向：物联网安全、智能家居安全
导　　师：袁　斌
指导老师：袁　斌
E-mail：737071824@qq.com
QQ：737071824
联系电话：13946610810
毕业去向：比亚迪股份有限公司

近年来，随着物联网技术的发展，智能家居成为物联网应用场景中极为重要的一部分。为了帮助用户更好地实现对智能设备的自动化操作，新兴的物联网云平台为用户和开发人员提供了大量开放的应用程序接口（Application Programming Interface，API），使其能够开发符合自己需求的自动化应用程序。然而，开放的物联网云平台API也带来了新的安全隐患。为此，开展了针对物联网云的API安全研究，发现当前广泛使用的SmartThings云平台中存在API的滥用问题，使得攻击者可以对物联网设备进行非法访问，或利用云平台作为其他攻击的中介。调研结果表明，上述安全风险影响了73.7%的SmartThings云平台应用程序。

为了评估并抵御上述攻击，设计了4种概念验证方案进行实验验证，实验结果表明这4类攻击是可以真实存在的，并可能造成较大危害。在实验过程中，归纳了攻击的可能流程和可能利用的API，确定了发现攻击需要检查的API范围以及相应的检查原则，进而设计并实现了SmartNotify工具对此类新型攻击进行检测。SmartNotify以待检测的SmartApps作为输入来自动分析源代码，提取需要检查的API信息，并进行安全的参数检查和循环检查，最后输出关于SmartApps的威胁报告。用户能够根据威胁报告决定是否安装使用特定的SmartApp，平台则能够根据威胁报告决定是否上架该SmartApp。

为了测试提出工具的有效性以及高效性，对SmartNotify进行了相应的实验评估。在有效

性方面，验证了基于参数检查和循环检查的不同检测结果的正确性。在效率方面，使用SmartNotify对SmartThings云平台可用的371个SmartApp进行了检查，平均检测时间在500 ms以内，开销较低，即所设计的方案能够高效的实现对SmartApps的安全检测。

可搜索对称加密的低选择性查询恢复 攻击方法研究

姓　　名：王　淼
研究方向：可搜索加密、数据安全
导　　师：徐　鹏
指导老师：徐　鹏
E-mail：1162521578@qq.com
QQ：1162521578
联系电话：15213122797
毕业去向：重庆长安汽车股份有限公司

随着云计算的迅速发展，用户对云服务器上加密数据的有效搜索越来越关注。可搜索对称加密技术（Searchable Symmetric Encryption，SSE）是一种密码学原语，它允许用户在不解密数据的情况下，直接在云服务器上根据关键字检索加密后的数据文件。然而，现有的SSE技术在检索过程中仍然暴露了一些信息，使得攻击者可以利用用户搜索历史中的信息来发起查询恢复攻击，推断查询对应的关键字。

在过去研究者所提出的查询恢复攻击中，通常都有两个很强的限制：它们只能针对那些匹配了大量文档的查询起作用，这样就忽略了实际中更常用的只包含少量精确结果的查询；同时，攻击都需要获得大量的用户背景知识，例如用户文档的明文信息。针对以上两个问题，提出了两种基于文档容量泄露的新型查询恢复攻击。第一种攻击作为基础攻击，利用文档大小将关键字和查询进行匹配，并进一步利用本文所提出的筛选规则来找到查询的底层关键字，这使得攻击可以恢复匹配文档数量较少的查询，即低选择性查询，同时所需的用户文档先验知识比现有攻击更少；第二种攻击在基础攻击上做了一些改进，它需要一些额外但容易获取的攻击条件，可以在掌握相同数量级的用户文档先验知识的情况下恢复更多查询。

通过在真实数据集上进行的大量实验，验证了两种新型攻击的效果。实验结果表明，在相同条件下，新的攻击比之前最好被动攻击的查询恢复率要高得多，并且可以有效地恢复匹配文档数量较少的查询查询。在一个大规模的真实数据集实验中，新型攻击可以实现比子图攻击（目前最有效的同类型攻击）高出25倍左右的查询恢复率。

基于学习的JavaScript模糊测试优化技术研究

姓　　名：王勇聪
研究方向：软件测试、模糊测试
导　　师：文　明
指导老师：文　明
E-mail：1789120321@qq.com
QQ：1789120321
联系电话：13088723169
毕业去向：南京中兴新软件有限责任公司

JavaScript（JS）引擎被广泛使用在各类浏览器中。引擎漏洞可能会导致严重的后果，因此，确保JS引擎的安全性和可靠性是非常重要的。虽然模糊测试是一种简单且有效的检测JS引擎漏洞的方法，然而现有的JS模糊测试工具更关注于设计有效的变异机制来生成不同的和有效的种子，而经常忽略了种子语料库的重要性，也很少能够很好的利用到JS种子特征指导模糊测试。

本研究主要从以下几个角度系统评估了种子选择对JavaScript引擎模糊测试的影响，分别是种子的收集来源、种子数量、种子大小和种

子包含的代码结构。实验结果表明，种子的来源和种子包含的代码结构对现有模糊测试工具的效果可以产生显著的影响。CVE或者包含更多相关代码结构的JS种子可以使代码覆盖率更高，检测到更多崩溃。

受实证研究结果启发，本研究首先设计了一个简单的启发式学习方法来优先选择更优质JS文件作为种子语料库。实验结果表明，当使用该种子库时，现有的最先进模糊测试工具能够实现明显更高的代码覆盖率并且识别出更多的崩溃。本研究还进一步设计了一种结合语言模型，通过学习种子特征指导种子生成的模糊测试变异策略DLFuzz，实验结果证明了DLFuzz在漏洞检测能力的先进性，反映了基于学习的JS模糊测试技术优化的可行性。

基于多路径探索的恶意软件逃逸检测

姓　　名：徐钫洲
研究方向：恶意软件分析、逃逸检测
导　　师：羌卫中
指导老师：羌卫中
E-mail：abc_000@qq.com
QQ：593506000
联系电话：18190204665
毕业去向：成都市高新区事业单位

分析恶意软件时，静态分析较为高效但易受加密、加壳等混淆技术的影响，而动态分析捕获具体执行的行为所以更具鲁棒性。但恶意软件开始使用逃逸技术获取环境信息，并在检测为动态分析环境时执行不同的路径以阻碍动态分析。虽然已知的逃逸技术可以明确进行应对，但是缺乏完整先验知识的逃逸技术则需要采用更通用的应对方法，如采用不确定条件的逻辑炸弹，以及其他没有部署应对策略的和未知实现的未支持逃逸技术。

为此，提出了一种自动探索逃逸恶意软件的方法，利用污点分析和概率计算指导的多路径探索来高效应对逃逸技术。首先基于动态覆盖信息计算路径的执行概率，并使用污点分析识别使用不确定条件的逃逸技术的相关分支。随后优先选择具有较低概率的分支以及被污点信息影响的分支进行探索，而探索由强制执行直接设置所选路径的分支结果实现。此外，利用主动反逃逸对策应对已知的逃逸技术来减少多路径探索的开销。还提供了敏感行为的相关信息以辅助进一步的人工分析。

根据该方法实现了一个原型系统。通过在一组使用了多种逃逸技术的样本上的实验，表明系统能够通用而高效地应对逃逸技术。其中概率计算在不需要先验知识的情况下指导多路径探索，进而应对未支持的技术，并且在复杂的控制流上能以显著少于线性探索的轮数应对逃逸技术。此外，污点分析能够准确地识别逻辑炸弹相关的分支并优先进行探索。

健壮的非交互式可搜索对称加密研究

姓　　名：徐润泽
研究方向：可搜索加密
导　　师：徐　鹏
指导老师：徐　鹏
E-mail：987836098@qq.com
QQ：987836098
联系电话：13236507798
毕业去向：上海蔚来汽车有限公司

随着计算机网络的飞速发展，信息量与数据量快速增长，计算机的计算能力和存储能力无法满足人们的需求。在这种情况下，越来越多的数据被集中存储在云服务器上，这就带来了数据隐私的问题。动态可搜索对称加密是一种实用的加密原语，可以允许客户端更新存储在云服务器上的加密数据，例如添加或删除文档。然而，现有的许多DSSE方案在更新时可能

会造成客户端隐私泄漏，并且也没有考虑到客户端发出不合理请求对方案造成的影响。

为了解决上述问题，基于对称可撤销加密和不经意映射协议构造了动态可搜索对称加密方案，命名为方案ROSA。方案采用逻辑删除和物理删除相结合的方式，由服务器来进行非交互式物理删除避免传输冗余数据，并降低客户端重加密的计算开销。此外，方案ROSA实现了满足健壮性下的Type-II后向安全，可以正确处理客户端发出的不合理更新操作，例如添加重复记录或删除不存在的记录。

复杂度对比分析表明方案ROSA实现了亚线性级的搜索复杂度，搜索计算成本仅与包含被搜索关键字的密文数量相关。通过真实数据集对方案ROSA进行了性能测试，并与现有的实现非交互式安全的前向和后向安全DSSE方案进行了比较（即方案Janus++和Aura）。结果表明方案ROSA达到了实际性能，与方案Janus++相比，本方案生成可搜索密文至少节省了90%的时间开销；就搜索性能而言，与方案Janus++相比，本方案节省了至少98.81%的时间，并且与方案Aura相比，查找一个匹配的密文只需要额外的44.73微秒。

基于细粒度优化选项差分的编译器缺陷定位技术研究

姓　　名：杨　静
研究方向：软件测试
导　　师：文　明
指导老师：文　明
E-mail：1549152742@qq.com
QQ：1549152742
联系电话：15711040272
毕业去向：快手科技有限公司

为了控制生成的目标代码的性能，编译器提供大量不同优化选项。几乎所有软件 产品都是在编译优化后发布的，因此编译器优化类的缺陷对软件系统的正确性具有 重大影响。然而，作为规模庞大的复杂软件系统，编译器支持大量不同粒度级别的优 化选项，准确定位编译器优化类的缺陷非常困难。为此，现有研究提出基于见证测试 程序的编译器缺陷定位方法。现有方法因需要突变生成大量见证程序，导致缺陷定位效率较低且效果受限。为了解决现有方法的局限性，提出了一种基于细粒度优化选项 配置的自动化编译器缺陷定位方法 ODFL。

编译器的优化组件支持不同粒度的优化选项，粗粒度的优化选项默认启用上百 种细粒度的优化选项，而真正触发编译器缺陷的优化仅仅是细粒度优化中的某一种 或某几种。根据此观察，针对编译优化类的缺陷定位，ODFL 基于粗粒度优化选项与 细粒度优化选项的层级关系来配置细粒度优化序列，以更轻量级更准确定位编译优 化类的缺陷。具体而言，首先，筛选出由触发缺陷的粗粒度优化选项启用的细粒度的 优化选项，识别出细粒度的缺陷相关和缺陷无关的优化选项。其次，基于细粒度优化 构建有效的失败优化配置和通过优化配置。最后，利用基于频谱的缺陷定位技术评估 编译器的代码行是缺陷行的可疑值，并将代码行的可疑值聚合到编译器文件，根据编 译器文件的可疑值排名进行编译器缺陷定位。

为评估 ODFL 的缺陷定位效果，在 60 个 GCC 编译器缺陷上开展实验。实验结果显示，ODFL 在所有评估指标方面都显著优于最先进的编译器缺陷定位方法 RecBi， 展示了ODFL 的有效性。此外，ODFL 比 RecBi 更加高效，平均可以节省超过 88% 的缺陷定位时间。

面向物联网设备共享中授权数据泄露的风险分析与抵御研究

姓　　名：杨茂根
研究方向：物联网安全

导　　师：袁　斌
指导老师：袁　斌
E-mail：yangmg1101@qq.com
QQ：1282860473
联系电话：17775466516
毕业去向：中国电信天翼云科技有限公司

为满足用户多样的设备管控需求，当前的物联网云平台不仅支持设备远程访问等基本功能，还支持设备共享等复杂功能。然而，由于缺少统一的安全规范和行业标准，各个平台提出的设备共享方案存在显著的异构性，尤其是在授权数据的管理方面。例如，设备ID、产品ID和访问令牌等不同类型的数据被不同的物联网云平台作为授权数据。另一方面，已有研究表明，当前物联网云平台在数据安全管理方面存在安全风险，可导致授权数据泄露，进而造成设备非法控制、用户隐私泄露等安全问题。

为此，首次系统地研究了当前物联网云平台的授权数据管理机制，揭示了其普遍存在的安全风险。具体而言，对ThingsBoard、ThingsKit、JetLinks等6个物联网云平台进行了系统性的安全分析，发现了6个可导致授权数据泄漏的安全缺陷，并在实验环境中验证了上述缺陷的可利用性及其安全后果。

进一步地，为了解决物联网云平台中的上述授权数据泄露问题，提出了一种基于影子数据替换实际数据的授权数据保护方法。该方法将真实授权数据保留在云平台内部，而将影子授权数据分发给被授权用户，以此来保证实际数据授权不被泄露。基于上述方法，实现并开源了缺陷代码自动修复工具SecHARE。该工具可以为缺陷物联网平台生成安全强化补丁，并在启动时自动加载，实现对物联网授权数据的有效保护，进而为物联网设备和用户提供有力的保障。

为了验证所提防御方法和SecHARE工具的兼容性、有效性和高效性，使用SecHARE对ThingsBoard、JetLinks、Kaa等3个流行开源物联网云平台的授权泄露缺陷进行了修复。实验结果显示，提出的防御方法可以有效地防止授权数据的泄露，且仅引入14.25 ms的设备控制时延和0.14%的CPU计算负载等额外开销。

基于排队论的软件定义网络中拒绝服务攻击抵御机制

姓　　名：张　帆
研究方向：软件定义网络安全
导　　师：袁　斌
指导老师：袁　斌
E-mail：zhangf743@qq.com
QQ：540873712
联系电话：13660363990
毕业去向：中国电信股份有限公司广东研究院

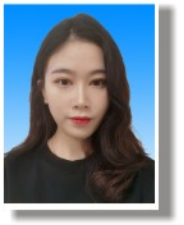
软件定义网络是一种数据面和控制面分离的新型网络架构，具有突出的可编程性、可扩展性和灵活性，被广泛用于云数据中心的部署。然而，软件定义网络也面临着许多安全威胁，其中最典型的便是分布式拒绝服务攻击。攻击者利用软件定义网络数据平面转发机制的缺陷，大量消耗软件定义网络设备有限的处理资源，影响正常网络转发服务，进而严重影响数据中心网络的可用性、可靠性和安全性。现有的防御方法多倾向于使用攻击检测，而这种被动式防御方案耗时较长且准确性不高，还可能出现在采取应对措施之前，攻击就已导致网络服务不可用，而无法按原计划施行防御策略的情况。

为了避免上述问题，提出了一种基于排队论的软件定义网络中拒绝服务攻击抵御机制。首先，在网络流量分布方面，网络中不同位置的转发节点所承担的网络流量压力不同，基于网络拓扑和路由选择分析，对网络流量分布情

况进行建模，估算各转发节点需要承担的流量大小。其次，对软件定义网络的数据包处理机制进行分析，利用M/M/c排队论模型对数据包在交换机和控制器中的处理过程进行建模，估算数据包处理过程在交换机和控制器中的平均等待时间。最后，将软件定义网络中数据包处理过程的排队论模型与网络流量分布估算模型相结合，分析得出抵御攻击成功的约束条件，并利用约束条件对抵御攻击所需的设备资源量进行估算，从而指导网络管理者合理配置和投入网络资源，实现服务质量可保障的拒绝服务攻击抵御。

基于理论模型，使用模拟仿真环境搭设软件定义网络实验平台，对比真实网络中的数据与模型计算结果，证明了模型的可行性。对攻击抵御机制进行了量化分析，证明了机制在评估拒绝服务攻击与防御效果上的有效性，以及对网络管理者合理配置资源的指导意义。

基于多源代码表示的智能漏洞检测技术研究



姓　　名：张雨婷

研究方向：软件安全、漏洞检测

导　　师：文　明

指导老师：文　明

E-mail：534569157@@qq.com

QQ：534569157

联系电话：15971501968

毕业去向：北京快手科技有限公司

随着软件的开源趋势成为主流，软件漏洞随着开源软件或包的使用和分发而广 泛传播，这对于软件系统的安全造成了严重的威胁。考虑到源代码中包含了丰富的 语法和语义信息，学术界和工业界提出了大量基于源代码分析和代码表示的漏洞检 测方法。

目前，代码表示方法主要包括三种：代码度量信息、代码文本序列和代码图结 构。本文通过总结不同代码表示的特征定义与提取方法，分别构建了三种漏洞检测 系统。基于代码度量信息的漏洞检测方法主要利用一系列代码复杂度指标来衡量代 码包含漏洞的可能性，基于代码文本序列的漏洞检测方法根据词序列来保留代码的 上下文语义信息，基于代码图结构的漏洞检测方法中主要剖析了多种代码图结构在 图表示学习中的贡献。本文通过对三种代码表示方法的比较，开展基于多源代码表 示的智能漏洞检测技术评估，从而得出了最优的代码表示结构。

本文在收集了大规模真实软件漏洞数据集和人工构造的数据集后，对三种不同 的代码表示技术和十三种学习模型进行了系统的实验评估。实验结果表明了基于代 码度量信息的漏洞检测方法和基于代码文本序列的漏洞检测方法在人工数据集上的 有效性和在真实软件漏洞数据集上的局限性。在基于代码图结构的漏洞检测方面， 本文剖析了多种代码图结构在图表示学习中的贡献，得到了最优的代码图表示。最 终，本文提出了基于图切片的漏洞检测方法，将该方法与最先进的漏洞检测系统相 比，在真实软件漏洞数据集上，漏洞检测的平均准确率等指标都得到了显著的提升， 表明了该方法的有效性、实用性和可拓展性。

深度包检测系统漏洞挖掘框架研究

姓　　名：张哲畅

研究方向：网络安全、模糊测试

导　　师：邹德清

指导老师：袁　斌

E-mail：1095936449@qq.com

QQ：1095936449

联系电话：15071231830

毕业去向：宾夕法尼亚州立大学攻读博士

深度包检测系统（Deep Packet Inspection，

DPI）能记录且阻断异常的网络连接，是现代网络系统中至关重要的安全设备。然而，DPI系统往往由各厂商自主设计，缺乏统一的安全实现规范，因此，DPI系统中容易出现设计或实现不当导致的安全漏洞，如攻击者可以构造攻击数据包来绕过DPI系统对特定网络连接的检查，实现非法数据传输与服务访问。现有工作通常使用人工检查、符号执行与模糊测试等方法来检测DPI系统中绕过检测规则的漏洞，存在需要过多人工干预、路径爆炸、反馈信息较少等不足，难以实现高效的漏洞挖掘。

针对以上问题，设计状态差异引导反馈机制，提出一种基于差分测试的自动化DPI系统绕过漏洞检测方法。具体而言，通过对已知的DPI系统绕过漏洞的分析，发现在非法数据包通过DPI系统之前，DPI系统通常会处于异常状态。因此，可以通过诱使DPI系统进入异常状态来扰乱其正常功能，触发其存在的绕过漏洞。由此，设计并实现了基于差分测试的DPI 系统漏洞检测框架StateDiver，该框架以DPI系统的状态转移情况为引导，为测试用例的变异和筛选提供指导，实现对绕过检测规则漏洞的高效挖掘。

共发现了Snort、Snort++、Suricata开源DPI系统中的16个规则绕过漏洞，其中8个为已知漏洞，8个为新发现的漏洞。向厂商报告了所有发现之后，获得1个CVE与1个CNVD漏洞编号。对新发现漏洞进行分析并揭示了新的绕过漏洞路径。此外，将StateDiver与同类先进工作Geneva进行了详细对比，结果表明StateDiver在发现漏洞数量、速度与诱发状态迁移的丰富性等方面均优于Geneva。

面向异质数据的联邦学习算法研究

姓　　名：赵瑞欣
研究方向：联邦学习
导　　师：马晓静
指导老师：马晓静
E-mail：913510499@qq.com
QQ：913510499
联系电话：18697348960
毕业去向：浙江菜鸟供应链管理有限公司

数据安全是国家数据治理的重要议题。联邦学习作为具有隐私约束的分布式机器学习框架，是一种自然的保护数据的机器学习技术。但是在现实中，联邦学习面临着数据异质性的挑战，导致联邦学习的性能出现灾难性下降。本文围绕数据异质性问题展开研究，旨在不增加通信开销的前提下，实现模型的高准确度和高泛化性。

针对数据异质性导致的客户端模型漂移问题，提出基于动态知识蒸馏的联邦学习算法FedDKD，以纠正本地模型的优化目标。允许客户端以灵活的方式调整知识蒸馏比例，有效学习全局模型表示，减轻局部模型的差异，使得聚合后服务器能够生成性能更优的全局模型。在MNIST、CIFAR10和CIFAR100三个数据集上评估了FedDKD的有效性，证明了其对数据异质性问题的普遍适用性，模型准确度普遍高于现有的最好方法。

针对数据异质性导致单个全局模型无法适用于所有客户端的问题，提出基于自适应聚类的联邦学习算法FedAC，为每个客户端提供最适合的模型。FedAC可以在不了解客户端数据分布的情况下，学习数据特征表示，自适应识别聚类结构，将相似的客户端归到一起，促进相似用户互相学习。另外，充分考虑类内类间关系，进一步提高了模型准确度和泛化性。经实验验证，FedAC可以以较少的通信开销实现准确的集群划分，为用户提供高质量的中心模型。

综上所述，本文从约束客户端模型漂移和自适应聚类两个角度展开研究，有效改善了数据异质性带来的模型精度下降问题。

论人工智能的低碳转型

（储朝阳 https://blog.csdn.net/qq_44009891/article/details/131124210）

近些年，人工智能技术的巨大突破重塑了人类社会的方方面面，然而这种成功的背后却是巨量的能源消耗和二氧化碳排放。本文从实际数据出发探讨了人工智能模型产生巨量能耗和碳排放的现象和原因，并说明了人工智能技术需要低碳转型，要走可持续的绿色人工智能之路。

一、问题的由来

近十年来，人工智能（AI）领域飞速发展，已经给世界带来了翻天覆地的变化，各种高性能AI大模型纷至沓来，深度学习、人脸识别、对话系统、AlphaGo等技术和产品的成功令世人瞩目。然而，在这辉煌的背后，一种令人担忧的负面趋势也开始凸显：硬件设施的性能、参数规模的大小、训练数据的多少成为影响AI模型表现的重要因素之一，但似乎也逐渐成为业界和学界研究人员心目中的“决定性因素”。更可怕的是，主流舆论对此尚未给予足够的关注。

人脑是一种效率极高的智能来源，但目前的AI还达不到这样的水平。现代AI模型需要消耗大量电力，而且对电力的需求正以惊人的速度增长。OpenAI于2018年发布的一份关于AI计算量增长趋势的分析报告显示，自2012年以来，构建一流AI模型所需要的计算资源消耗平均每3.5个月翻一番，至2018年这个指标已经增长了30万倍以上[1]。

更为重要的是，庞大的电力需求意味着巨量的二氧化碳排放。根据美国环保署(EPA)公布的数据，在美国，一千瓦时电力平均对应0.954磅二氧化碳排放量。2019年7月，来自马萨诸塞大学阿姆赫斯特分校的研究人员进行了一项广泛研究，研究表明训练一套优秀的深度学习模型可能产生高达284吨的二氧化碳排放，这相当于美国汽车从制造到报废所产生的总二氧化碳排放量的近五倍[2]。然而，这并不是AI模型所能产生的碳排放的极限。一年后的6月，OpenAI宣布搭建了有史以来（至2020年）规模最大的AI模型GPT-3[3]，这套惊人的模型被认为是自然语言处理领域目前最成功的语言模型之一，已经成为令人印象深刻的伟大技术成就。GPT-3能够根据简单的创意写出完整的小说，把晦涩艰深的法律条文用平易近人的语言阐述，也能回答从柴米油盐到天文地理的提问，表现远超此前的任何自然语言处理模型。然而，与卓越性能相对的是高昂的代价：GPT-3训练的硬件和电力成本高达1200万美元（约7500万人民币）。根据谷歌和美国加州大学伯克利分校于2021年联合发表的一项研究，GPT-3在训练过程中消耗了1287兆瓦时的电力，产生了552吨的碳排放，相当于一辆汽车120年的排放量[4]。

美国麻省理工学院的研究员Neil Thompson指出，深度学习的发展令人工智能革命成为可能，但其不断增长的成本值得警惕。当规模的膨胀成为了一股无法阻挡的潮流，同步骤增的金钱成本、能源消耗以及碳排放，都成为了学界和业界无法忽视的问题。

AI技术带来的碳排放已经不容忽视，如果

行业趋势继续下去，那么情况将很快失去控制。除非我们愿意重新评估并改革当今的AI研究议程，否则人工智能领域很可能在不久的未来成为引发气候变化的罪魁祸首。

二、为什么AI模型会产生如此多的碳排放

在当今以深度学习为中心的研究范式当中，人工智能的主要进步主要依赖于模型的规模化扩展：训练数据集更大、模型参数规模更大、计算资源更大。

近些年，训练人工智能模型的数据集在快速增肥。在使用包含30亿个单词的数据集进行训练之后，BERT模型[5]在2018年实现了同类最佳的自然语言处理(NLP)性能。而在利用包含320亿个单词的训练集完成训练之后，XLNet[6]又超越了BERT。不久之后，GPT-2[7]开始在包含400亿个单词的数据集上接受训练。最终是我们前面提到的GPT-3，它使用的是一套包含约5000亿个单词的加权数据集。尽管数据规模的扩大带来了模型性能的提升，但是模型在训练过程中需要为每一条数据执行一整套冗长的数学运算，并以复杂的方式更新模型参数，这就使得数据集越大，与之对应的算力与能源需求也在飞速增长。

此外，AI理论不断进步，随之膨胀的，是模型参数的规模，这种趋势带来了更准确的结果和更优秀的性能。在研究和实践中，专家们发现，大大增加模型的参数量，使其超过数据数量的过程，也就是“过参数化”，能够提高训练的效率，加强模型的泛化能力。前文提到的GPT-3 模型就包含多达1750 亿个参数，而它的前身GPT-2模型(在2019 年发布时，同样创下了体量层面的纪录)只有15亿个参数。GPT-2在拥有千万亿次算力的设备上训练了几十天，相比之下，GPT-3 的训练时长增长至数千天。中国的研发团队在大规模模型训练的道路上也没有落后，智源研究院2020 年发布的悟道2.0模型已经拥有1.75万亿个参数，相当于GPT-3的10倍，阿里达摩院2021年发布的M6模型参数量更是超过了10万亿。这种靠“每况愈大”模型推动AI技术进步的问题在于，模型中包含的参数量越大，计算需求就越大，所带来的能源消耗就越夸张，并由此产生巨量碳排放，这最终会转化为对环境的影响。

三、展望未来：可持续的绿色AI

虽然人工智能模型规模的膨胀至少在未来数年是不可阻挡的潮流，但我们仍有必要控制它们的能耗和碳排放。

目前的学界和业界，尤其是科技巨头公司，研发模型的模式有过于粗放之嫌，在降低消耗方面还有很大提升空间。IBM研究员、麻省理工学院–IBM沃森人工智能实验室成员约翰·科恩表示：“如果要让人工智能继续快速发展，我们需要减少它对环境的影响，需要开发方法使人工智能模型更小、更高效。”

专门针对深度学习场景设计的芯片是降低AI研发和部署成本的一个关键。由于深度学习模型训练需要大量并行的浮点运算和读写，CPU 难以胜任，目前主要使用 GPU或TPU进行。英特尔，英伟达，谷歌等公司都在设计和生产这些芯片。

通过政策的激励和管控，更合理地安排人工智能研发和部署，也是管控成本，提高整体效率的重要手段。例如，如果人工智能模型使用主要来自可再生能源的电力进行训练，其碳排放就会相应降低；在可再生能源密集的区域，可再生电力的时间分布不均匀，在电力冗余的时段进行训练，也是提高人工智能产业能源利用效率的有效方法。中国2022年2月正式启动的“东数西算”工程就被学界认为是促进产业绿色发展的强力手段。由于东部资源日趋紧张，大规模发展计

算中心难以为继，而西部可再生能源充沛，可以很好地承接东部算力需求。华为、百度、腾讯等众多科技公司都响应了这一工程，在西部建立了计算或数据中心。“东数西算”所提升的不只是人工智能训练，而是整个算力密集产业的能效，不仅能立竿见影地降低研发部署成本，更有希望让整个行业的发展变得更加高效更加绿色，推动东西部协调发展。

算法研究者、硬件开发者和政策制定者都正在实现“绿色AI”的道路上努力着。但这些硬件和软件支持，终究只是支持。过于关注准确性等性能指标，而以牺牲速度和模型大小等效率指标为代价的观念和评价体系，普遍存在于整个人工智能领域。这样的评价观念亟待转变，AI模型也需追求碳中和，性能不应成为唯一指标。毕竟模型做出来，不只是为了在学术期刊和会议上攀比的，更多时候是为了转化为应用，服务业界的，而任何高能耗高排放的应用产品，都是不符合时代要求的。只有认识到这一点，才能真正让“绿色AI”像今天的“绿色化学”一样，从文献里走出来，走进每一个实验室，走进每一家公司，走进每个人的生活。

参考文献

[1] Amodei, D. and Hernandez, D. AI and compute, 2018. Blog post.

[2] Strubell, E., Ganesh, A. and McCallum, A. Energy and policy considerations for deep learning in NLP. In Proceedings of ACL, 2019.

[3] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. Advances in neural information processing systems, 2020, 33: 1877-1901.

[4] Patterson D, Gonzalez J, Le Q, et al. Carbon emissions and large neural network training. arXiv preprint arXiv:2104.10350, 2021.

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

[6] Yang Z, Dai Z, Yang Y, et al. Xlnet: Generalized autoregressive pretraining for language understanding. Advances in neural information processing systems, 2019, 32.

[7] Radford A, Wu J, Child R, et al. Language models are unsupervised multitask learners. OpenAI blog, 2019, 1(8): 9.

精彩言论（转载）

人生终有不期而遇的温暖和生生不息的希望。只有甩掉身上沉重的包袱，才会静生慢长，不负努力和时光。

（林立成 https://baijiahao.baidu.com/s?id=1627178952976891231&wfr=spider&for=pc）

宣父犹能畏后生，丈夫未可轻年少。

（于跃 https://t.bilibili.com/795768800675889256?spm_id_from=444.41.0.0）

于高山之巅，方见大河奔涌；于群峰之上，更觉长风浩荡。

（甘 芮 http://lianghui.people.com.cn/2023/n1/2023/0311/c452482-32641676.html）

追求的后面没有句号，人生也永远没有太晚的开始，只要你听从内心的召唤，勇于迈出第一步，人生的风景就永远是新奇的、美妙的。

（甘 芮 http://opinion.people.com.cn/n1/2016/0503/c1003-28320992.html）

人生的态度是，抱最大的希望，尽最大的努力，做最坏的打算。——柏拉图《理想国》

（江宗泽 https://www.juzikong.com/posts/5a09d716-a878-4469-a364-f93faec384d6）

解释型语言生态中的供应链攻击与防御

（郭啸辰 https://zhuanlan.zhihu.com/p/636488017）

一、背景介绍

许多现代网络应用依靠解释型编程语言，因为它们有丰富的库和包。像PyPI、Npm和RubyGems这样包管理器提供了一个集中的仓库，开发者可以搜索和安装代码包，以帮助开发。包管理器不仅使开发过程更加有效，而且还创建了一个大型社区，进行合作和分享开源代码。不幸的是，攻击者已经找到方法渗透到这些社区，用恶意代码感染良性的流行软件包，窃取凭证，安装后门，甚至滥用计算资源进行加密货币挖掘。

这个问题的影响不仅局限于小型个人网络应用程序，而且还涉及到大型的网站、企业，甚至是依赖于开源解释型编程语言进行内部和外部应用程序开发的政府机构。攻击者可以通过向这些包管理器投毒恶意包，制造软件供应链攻击来渗透到一些防御良好的组织中。例如，在NPM中，eslint-scope是一个每周下载量达数百万的软件包，被攻击者篡改以窃取开发者的凭证。同样地，在RubyGems中，rest-client已经被下载了超过一亿次，被攻击者篡改以在Web服务器上留下了一个远程代码执行（RCE）后门。这些攻击行为悄无声息，用户极难发现。

考虑到包管理器投毒攻击的巨大影响，学术界和工业界也提出了多种恶意包检测的方法。但是现有方法的误报很高，数量巨大的待审核的可疑包无法处理，导致方法无法真正投入使用。例如PyPI官方在2020年推出的检测工具PyPI malware checks就在上线两年后由于未能达到预期效果被撤销。

二、攻击方法

由于第三方库开发和管理涉及到众多的环节，而且规则较为灵活，容易被攻击的角度有很多，主要分为上游和下游两类，分别对应着攻击中央仓库和攻击用户。有的攻击者目标是安装第三方库的使用者，他们上传大量有害的第三方库。用户在安装的过程中执行命令盗取用户信息，或者是从远程下载恶意软件植入后门，开启挖矿进程占用资源。部分攻击者更是将目标直接瞄准了中央管理仓库，PyPI就曾被爆出存在远程代码执行的漏洞，如果管理仓库被控制，攻击者可以随意修改包中的代码，后果不堪设想。也有攻击者对中央仓库进行DDos攻击，造成服务瘫痪。常见的攻击方式有以下几种：

错别字劫持（typosquatting）：在PyPI/NPM等代码管理仓库中typosquatting攻击是一种常见的攻击方式。攻击者会注册一个与目标软件包名称相似但存在拼写错误的软件包名称，然后上传一个恶意软件包到这个软件包名称下，使得用户在安装软件包时误输入错误的软件包名称，从而下载并安装恶意软件包。比如，假设一个开发者想要下载名为"requests" 的Python软件包，但是他因为疏忽打错了一个字母，将软件包名称误输为"reqeusts”，而攻击者正好注册了这个名称的软件包并上传了一个恶意的软件包到其中。当这个开发者使用pip安装这个软件包时，他实际上下载并安装了攻击者上传的恶意软件包。typosquatting攻击可以利用用户的拼写错误和疏忽大意，骗取用户下载恶意软件包，从而对用户的电脑和数据

造成危害。

依赖混淆（dependency confusion）：依赖性混淆攻击发生在从公共包管理器而不是预定的私有/内部包管理器下载依赖性库，因为恶意攻击者可以欺骗软件包管理从他控制的公共仓库下载恶意软件包。攻击者利用了开发者在使用多个软件库时的依赖管理机制，例如在使用Python的pip包管理器时，pip会首先查找本地环境中已经安装的软件包，若没有找到则会去公共的Python软件包仓库PyPI中寻找。攻击者通过上传一个与公共Python软件包仓库中同名但版本号更高的恶意软件包到一个私有的软件仓库，就能够欺骗pip去下载该恶意软件包，而不是从公共的Python软件包仓库中下载。

账户劫持（account hijacking）：由于开源项目的合作性质，攻击者很容易猜测到管理者账户的薄弱身份凭证或者是通过社会工程学获得账户的管理权，并在管理的项目中插入恶意代码进行攻击。在最近发生的管理器投毒事件中，这类情况频频发生。目前包管理器的维护者们已经注意到了这个问题，主动实施了一些安全改进措施，例如软件包签名和双因素认真（2FA）等。

三、防御方法

考虑到整个第三方库供应链可以攻击的环节很多，防御检测的角度也有很多。因为我们无法参与中央仓库的管理，这里我们只考虑下游的防护，即作为第三方库的使用者如何检测出危险的库，防止受到伤害。

在检测思路上，第三方库和其他软件的不同之处在于会提供有关的基本信息，如开发者姓名，邮箱，源码地址等等。可以将这些基本信息纳入检测的范围之内，因为正常的开发者会留下尽可能全面的信息方便后续的沟通，恶意攻击者上传的库往往缺乏这些基本信息，或者是随意填写虚假的姓名，邮箱。一些基本信息中带有anonymous，hacker等词的几乎可以直接判定为恶意包。

攻击者为了以假乱真，会使用正常开发者的信息迷惑检测。附上正常项目的代码仓库和star数，其实用户实际下载的代码和仓库中的完全不同。可以通过检测仓库名和下载的包名是否一致，开发者的名称是否一致，包中的代码和仓库中的是否一致进行判断。对于在安装过程中提供可执行程序的包，可以从仓库中的源码构建可执行程序，将两者进行比较。攻击者可能提供和开源代码完全不同的恶意可执行程序供下载。

除了通过基本信息进行大致的筛选和辅助判断，还可以使用静态分析的方式。构建控制流图对敏感的api进行提取。因为在危险的第三方库中，攻击者为了隐藏痕迹或者是获得窃取的敏感数据，会使用网络相关的api下载木马程序或上传信息，如python中的requests。还有加密相关的api，用来加密攻击代码不易被关键词检测发现。还有对于文件的读写操作也很值得注意，很多攻击是将危险代码下载写入本地文件中，或是从本地的机密文件中读取用户信息密码等敏感信息。可以使用污点分析的方式，判断是否有进行过滤，是否存在敏感信息泄露，危险代码下载的情况。

静态分析的检测方式速度会比较快，但是检测效果并不能很理想。因为静态检测会产生大量的假阳性的情况，而且很多精心设计的有害包并不能发现。例如，有的攻击者除了发布有害包，还会发布新的包依赖有害包，在新发布的这些包中并没有使用任何的危险代码。但是在安装的过程中还是完成了攻击者预想的攻击操作。为了避免这种情况，动态检测是必要的。可以使用沙箱，防止有害的包造成实际的

破坏。有害包为了攻击能够真正被执行，攻击者往往会将有害代码写入安装的环节，而不是在实际调用的过程中执行。所以需要在沙箱环境中执行安装的环节，通过工具检测安装过程中执行的系统调用进行判断是否有有害行为的执行。安装完成后，比较安装后的环境和初始环境的差别是否有除了包文件以外的文件写入。重点关注这些写入的文件，有一些是代码运行所必须的代码配置或是证书，有一些则是危险代码和事先编写好的后门程序。对这些文件使用常用的方式进行扫描可以有效检测出有害包。

除了这些传统的检测方式，还可以采用机器学习的方法。提取包的关键特征进行分类学习。采集的特征包括发布者的信息是否完善，网络、加密、文件读写相关api的数量，动态执行代码的数量等等。

四、总结与展望

包的中央管理仓库管理十分松散，影响的范围又极其广泛，安全保护十分重要。需要定期对包进行安全扫描，去除有害包。先通过基本信息进行初步筛查，然后使用静态分析的方法辅助扫描，动态分析最终确定是否有害。除了这种方法，现代的机器学习也可以提供帮助进行分类，但是缺乏足够的数据集进行训练是现在面对的一大难题。

我们作为开发者，同时也是第三方库的使用者。尽管我们无法控制上游的软件包，但可以遵循最佳策略来补救安全问题。开发者可以用已知安全的版本，以避免来自上游的供应链攻击。还应该定期检查安全公告并及时更新以避免已知的漏洞。对于不受信任的包，开发人员可以手动检查，部署一个审查管道来检查代码，并在运行时将其隔离，以避免潜在的危险。

参考文献

[1] Duan R, Alrawi O, Kasturi R P, et al. Towards measuring supply chain attacks on package managers for interpreted languages[J]. arXiv preprint arXiv: 2002.01139, 2020.

[2] Sejfia A, Schäfer M. Practical automated detection of malicious npm packages[C]//Proceedings of the 44th International Conference on Software Engineering. 2022: 1681-1692.

[3] Vu D L, Pashchenko I, Massacci F, et al. Typosquatting and combosquatting attacks on the python ecosystem[C]//2020 ieee european symposium on security and privacy workshops (euros&pw). IEEE, 2020: 509-514.

[4] Zimmermann M, Staicu C A, Tenny C, et al. Small World with High Risks: A Study of Security Threats in the npm Ecosystem[C]//USENIX security symposium. 2019, 17.

[5] Ladisa P, Plate H, Martinez M, et al. Taxonomy of attacks on open-source software supply chains[J]. arXiv preprint arXiv:2204.04008, 2022.

精彩言论（转载）

非淡泊无以明志，非宁静无以致远。太过重视结果与收获，就会失去生命的厚重与人生的层次，有时连想要的结果也得不到了。

（于 跃 https://t.bilibili.com/ 736047061291499522?spm_id_from=333.999.0.0）

与己为善，与人为善，与世界为善

（石月鑫 https://user.qzone.qq.com/450825297/main）

如果你独处时感到寂寞， 这说明你没有和你自己成为好朋友。

（孙一飞https://m.gmw.cn/baijia/2020-09/02/1301519508.html）

你永远不知道，你的未来有多强大!

（宋诚 https://new.qq.com/rain/a/20201011A00CFO00）

为庆祝计算机科学与技术专业建设50周年实验室举办多期高端学术讲座

燕　燕

为庆祝计算机科学与技术专业建设50周年，由实验室主办的两期高端学术讲座分别于4月21日和5月26日在武汉成功举行，两次活动分别由科学技术发展院院长廖小飞和实验室副主任吴松主持。

计算机学院书记易辉代表学院向与会专家表示诚挚欢迎和衷心感谢。易辉书记在开场词中强调，习近平书记指出教育、科技和人才是支撑国家发展的重要力量。华中科技大学计算机学院自1973年专业建设以来，一直紧密围绕国家需求，积极探索并行分布式计算、信息存储和网络传输等领域的创新突破。在全面贯彻党的二十大精神的开局之年，易辉书记鼓励全体师生认真学习贯彻习近平新时代中国特色社会主义思想，抢抓机遇，共建辉煌，为推动教育强国、科技强国、人才强国做出更大的贡献。

北京航空航天大学李昂生教授报告的主题是“信息世界的数学原理”，浙江理工大学丁佐华教授报告的主题是“智能制造中的软件理论与方法”，江西师范大学薛锦云教授报告的主题是“新概念循环不变式及其在软件正确性验证中的应用研究”，重庆大学廖晓峰教授报告的主题是“分布式机器学习的优化与安全”，中国科学院软件研究所张健研究员报告的主题是“软件测试与分析”，中国科学院软件研究所詹乃军研究员报告的主题是“模型驱动的复杂安全攸关嵌入式系统形式设计理论及开发环境”，新加坡国立大学何丙胜教授报告的主题是“联邦学习系统：数据孤岛上的有效和高效机器学习系统”，澳大利亚斯威本科技大学杨耘教授报告的主题是“Cost-effective big data management in cloud”，中国科学技术大学陈恩红教授报告的主题是“面向智能教育情境的学习者能力建模：研究与展望”，清华大学陶建华教授报告的主题是“个性化人物音频生成与伪造鉴别技术”，天津大学胡清华教授报告的主题是“低质多模态数据动态融合”，东南大学耿新教授报告的主题是“机器学习的“基因”：让机器像人一样学习”，华中科技大学张腾副教授报告的主题是“基于间隔分布优化的学习方法”。

两次活动分别围绕分布式计算、智能制造、软件正确性验证、模型驱动设计以及联邦学习和人工智能等研究领域，汇集各方智慧，开展了富有基础性、前瞻性的思想交流。通过学术讲座的举行，不仅增进了与各领域专家之间的交流，而且让师生们更多地了解了计算机科学与技术方面的知识，开拓了视野，期待未来更多的交流与合作。

燕　燕
负责事务：实验室宣传、科研项目管理等
Email：yany@hust.edu.cn

Combining Optimization Objectives: New Modeling Attacks on Strong PUFs

万才学 推荐

Combining Optimization Objectives: New Modeling Attacks on Strong PUFs是密码工程与硬件安全领域的国际会议CHES于2021年收录的一篇论文。该论文中提出了一种新的对PUF的建模攻击算法，该算法同时利用了响应信息与可靠性信息，结合了两种攻击算法，使得建模攻击更有效率，甚至能攻破之前普遍被认为无法被攻破的iPUF。

物理不可克隆函数（PUF）是一类硬件设计原语，其是一个输入响应系统，利用由制造元件时的噪声导致的元器件之间的细微误差，而产生独特的，不可预测的输出或响应。由于其轻量，独特，且物理不可克隆的特性，其常被作为一些经典加密情景，如设备身份认证等场景下的轻量级方案。

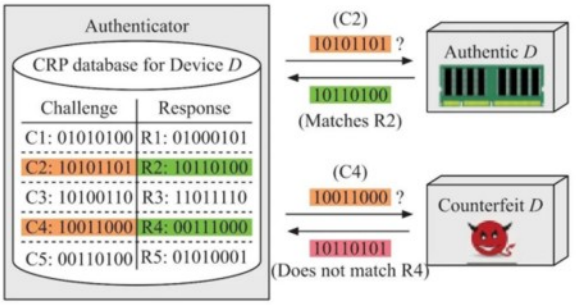


图1 基于PUF的设备认证方案示例

然而，由于简单的APUF能够被建模为一个线性的数学模型，因此可以被逻辑回归等经典的机器学习攻击算法成功攻击。于是研究者们常基于对多个APUF的组合以及构建非线性且不可微的行为，通过增大其解空间复杂度与防止梯度下降攻击的方式，保证其设计的安全性。例如iPUF即是在组合多个APUF的基础上，将上层PUF的输出插入到下层PUF的输入中，这种非线性且不可微的的行为极大地增大了对iPUF进行建模的难度，使得其几乎无法被常用的逻辑回归算法攻击。同时，上下层PUF对可靠性的贡献不一致的特点使得基于可靠性的攻击算法也无法成功攻击iPUF，于是iPUF的提出者在论文中宣称其是安全的。

但是在CHES的这篇文章中，作者从可靠性攻击和逻辑回归攻击两个角度，反驳了iPUF作者的宣称，提出了一种改进的，结合两种攻击算法的攻击方案，成功攻破iPUF，并且对其他PUF的攻击也十分有效。

由于环境噪音的存在，PUF在同一输出下的多次输出可能会产生不一致，即可靠性信息。这种现象一定程度上反映了其隐含的数学模型的结果的特性，于是可以利用PUF的可靠性信息对其进行攻击。常用的基于可靠性的攻击算法是协方差矩阵自适应进化策略（CMA-ES），其本质上是一种遗传算法，依照个体与目标可靠性的相关性评估适宜度，最终迭代进化得到最优个体。该算法每一轮都会收敛到一个APUF，因此对于一般的复杂PUF来说，其堆叠APUF造成的困难度增加对CMA-ES算法来说是线性的而不是指数级的。但是由于算法原理的特性，CMA-ES每次攻击都会更大概率地落在对可靠性影响最大的那一个APUF上，而

iPUF作者宣称iPUF中上下层PUF结构对整体的可靠性影响是有较大的差距的，其攻击结果必然会落在下层PUF上，但是未知上层输出便意味着未知下层的输入，因此无法在未攻破上层的情况下攻击下层。于是其宣称iPUF是无法被CMA-ES攻破的。

但这篇文章中作者推测并且通过实验证明了：在实际情况中，上下层对可靠性的影响程度并不会严格按照理想情况分布，因此一定程度上是有可能攻破iPUF的，只不过概率较低。于是为了防止CMA-ES因各APUF对可靠性贡献不均匀而难以攻击，作者引入了一个新项来评估遗传个体的适应程度：不与已成功攻击的个体相关。于是，在两项适应度的评估下，CMA-ES的攻击每一次都能落在不同的APUF上，因此能攻克影响程度差距大的情况，能够更有效地攻破iPUF。

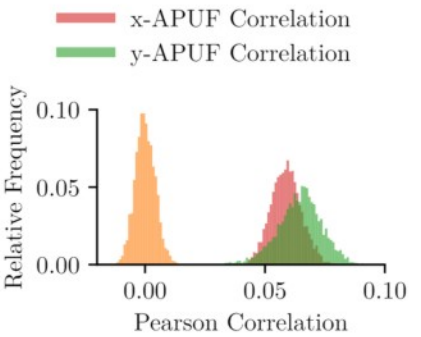


图2 iPUF的上下层与整体的可靠性相关度

此外，作者将基于可靠性的攻击与逻辑回归算法结合，不再使用稳定度较差，收敛速率较低的遗传算法，通过将可靠性相关程度的信息与PUF响应值的信息写入损失函数然后进行反向传播，利用梯度下降的方式使得模型快速收敛。同时，上文提到的多个APUF之间不相关的限制同样能够写进损失函数，于是能够将上述改进的方案由遗传算法完全转换成了梯度下降的如下形式。

$$
\begin{aligned}
\mathrm{loss}_{\mathrm{combined}} = & \sum_{j} \mathrm{loss}_{\mathrm{bin}}(\mathrm{model}_{\boldsymbol{W}}^{\mathrm{total}}(\boldsymbol{\Phi}[j,:]), \boldsymbol{r}[j]) \\
& - \epsilon_{\mathrm{a}} \sum_{j_1} \sum_{j_2} \mathrm{loss}_{\mathrm{PC}}(\mathrm{model}_{\boldsymbol{W}[j_1,:]}^{\mathrm{Arbiter}}(\boldsymbol{\Phi}[j_2,:]), \boldsymbol{h}[j_2]) \\
& + \epsilon_{\mathrm{b}} \sum_{j_1=1}^{k-1} \sum_{j_2=j_1+1}^{k} |\mathrm{loss}_{\mathrm{PC}}(\boldsymbol{W}[j_1,:], \boldsymbol{W}[j_2,:])|
\end{aligned}
$$

损失函数中的第一项即为模型输出与label之间的损失，后两项中，前者为当前模型与整体的相关程度，期望相关程度尽量高，因此作为loss的负项；后者为当前各个APUF之间的相关程度，希望相关程度尽量低，因此作为loss的正项。值得注意的是，对PUF而言参数W的正相关和反相关本质上均为相似，因此对后者的相关程度应取绝对值。

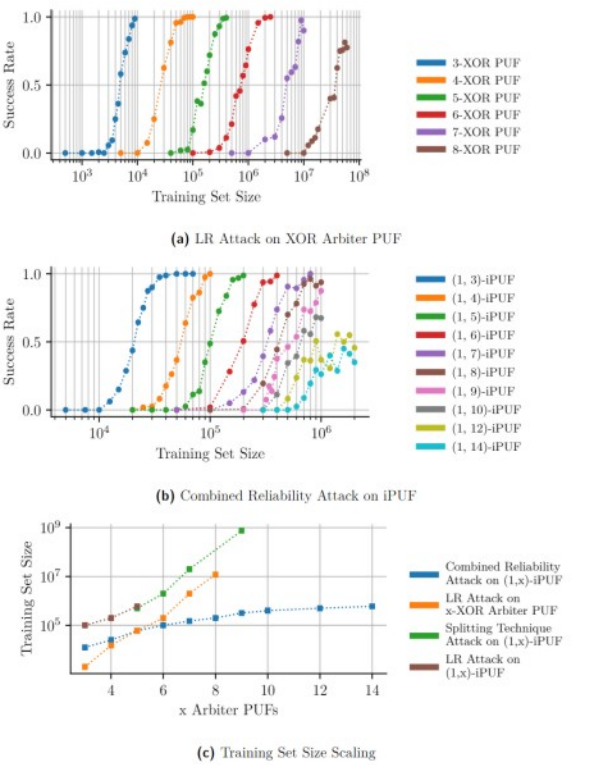


图3 不同攻击方法对各PUF的攻击对比

但是对iPUF而言，由于其包含将上层输出

GraphSPD：Graph-Based Security Patch Detection with Enriched Code Semantics

谢子凡　推荐

“GraphSPD：Graph-Based Security Patch Detection with Enriched Code Semantics”是网络与信息安全领域的顶级会议IEEE Symposium on Security and Privacy（S&P）在2023年收录的一篇论文。

随着开源软件的日益普及，许多上游软件的漏洞已经传播到下游软件。由于不同的维护政策，软件供应商可能会默默地发布安全补丁，而不发布安全公告（如CVE）。这使得用户不知道安全补丁的存在，并为攻击者提供了利用未修补的漏洞的机会。因此，检测这些未正式公开的安全补丁对安全软件的维护是非常必要的。作者在论文中提出了一个基于图神经网络的安全补丁检测系统，名为GraphSPD，它将补丁表示为具有更丰富语义的图，并利用为补丁特别设计的图卷积网络进行预测。作者首先开发了一种名为PatchCPG的新型图结构，通过合并补丁前和补丁后源代码的两个代码属性图（Code Property Graphs，CPG），以及保留补丁的上下文、删除和添加的代码来表示软件补丁。通过应用切片技术，作者保留了与补丁最相关的上下文并减少了PatchCPG的大小。然后，作者开发了一个端到端的深度学习模型，称为PatchGNN，以确定一个补丁是否是安全补丁。PatchGNN包括一个嵌入（embedding）过程，将PatchCPG转换成数字向量，以及一个新的多属性图卷积机制，以适应PatchCPG中的各种关系。实验结果表明，GraphSPD在安全补丁检测方面的性能明显优于最先进的方法。

开源软件的补丁记录了两个不同版本的源代码之间的变化。对一个维护良好的 Git 仓库而言，开发者在每次提交（一个commit）都应

接上页

插入至下层输入这一行为，导致iPUF的数学模型是不可微的，使得其无法在梯度下降算法中作为损失函数中的一项。于是，作者对iPUF的数学模型进行了分析和变型，在不改变输出的前提下去除了原本模型中的不可导部分，并且将下层输入进行了分片，将插入行为抽象为了可导的数学模型。在此基础上便可以将其应用于上述结合后的攻击算法。

作者进行了多种条件下的实验测试，实验结果表明其提出的方案无论是从攻击的成功概率，还是最难能攻克的目标，以及攻击所需的数据量，都是显著优于已有算法的。

总结来说，该论文结合两种攻击方式，能够同时利用响应信息与可靠性信息的同时使得新的攻击方案具有超出前二者的攻击效果。这种对多种攻击手段的适当的结合方案能够取各者之所长，避各者之所短，是一个非常有效的思路。

万才学
2022级博士研究生
研究方向：机器学习与芯片安全
Email：caixuewan@hust.edu.cn

是单一目的，例如安全修复、解决功能错误（非安全修复）、增加新功能。在这些补丁中，安全补丁通常被认为比非安全补丁更紧急，并被赋予更高的优先级。如果一个补丁属于任何通用弱点列举规范（CWE）类型的漏洞，就应当认为它是一个“安全补丁”。图1和图2分别举例说明了一个安全补丁和一个非安全补丁。补丁表示为一组连续删除和/或添加的语句（即以单个–或+开头的行）。在表图1中，一个if语句（第8行）被添加到原始赋值语句之前，以检查指针tcon是否有效。这个安全补丁缓解了NULL指针解除引用的漏洞，在这个漏洞中，值为NULL的无效指针会导致软件崩溃或退出。图2显示了一个非安全补丁的例子，一个过时的标识符hack被删除，因为它在后续代码中不再使用，但这个变量定义（第7行）不会引起任何安全问题。

```
1  commit 18f39e7be0121317550d03e267e3ebd4dbfbb3ce
2  diff --git a/fs/cifs/smb2pdu.c b/fs/cifs/smb2pdu.c
3  @@ -907,7 +907,8 @@ tcon_exit:
4   tcon_error_exit:
5      if (rsp->hdr.Status == STATUS_BAD_NETWORK_NAME) {
6          cifs_dbg(VFS, "BAD_NETWORK_NAME: %s\n", tree);
7  -       tcon->bad_network_name = true;
8  +       if (tcon)
9  +           tcon->bad_network_name = true;
10     }
11     goto tcon_exit;
12  }
```

图1：安全补丁的例子

```
1  commit 6cf575b1ad989fbb8a239dd6acc26d72286eb4cb
2  diff --git a/src/path.c b/src/path.c
3  @@ -1145,7 +1145,6 @@ int git_path_diriter_init(
4      unsigned int flags)
5   {
6      git_win32_path path_filter;
7  -   git_buf hack = {0};
8
9      static int is_win7_or_later = -1;
10     if (is_win7_or_later < 0)
```

图2：非安全补丁的例子

默认情况下，补丁包含6个相邻的代码行（例如图2中的第4-6行和第8-10行）作为每个代码修订的上下文。然而，这些上下文语句可能没有提供足够的语义来理解补丁。为了对一个开放源码软件补丁进行代码分析，作者检索应用这个补丁之前和之后的源代码，并分别称之为pre-patch和post-patch文件/函数。当pre-patch的代码中存在漏洞，而相应的修复语句在post-patch的代码中时，一个补丁被认为是安全补丁。通常情况下，安全补丁会涉及到对关键值的安全检查，例如，对于修复“use-after-free”和“double-free”漏洞的补丁，安全检查是增加一个条件语句，以检查指针或内存的可用性。

图3说明了GraphSPD的整体工作流程。GraphSPD的输入是软件的补丁，它首先被映射到一个包含控制/数据依赖和程序语法的中间图表示PatchCPG。接下来，PatchCPG的节点和边被转换为嵌入矩阵，并被送入基于图的深度神经网络PatchGNN，以预测输入补丁是安全补丁还是非安全补丁。

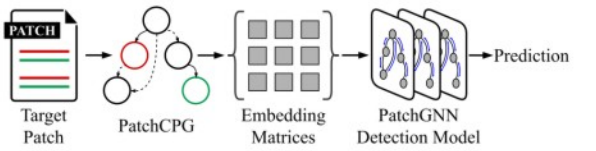


图3：GraphSPD的整体架构

作者首先假设pre-patch和post-patch版本的源代码都可以通过版本控制系统（如Git）进行访问。第一步是构建pre-patch和post-patch版本的代码属性图。给定pre-patch和post-patch版本中与补丁有关的函数，作者应用Joern解析器来生成两个CPG（即G_{pre}和G_{post}）。对于一个CPG，作者用两个集合来描述该图：G=(V, E)。V是一个2元组（id, code），表示的节点集合，其中id是识别节点的数字，code是这个节点所描述的源代码组件，即抽象语法树（Abstract Syntax Tree, AST）中的代码标记或控制依赖图（Control Dependency Graph，CDG）、数据依赖图（Data Dependency Graph，CDG）中的语句。E是有向边的集合，用3元组（id1, id2, type）表示，其中id1和id2代表开始和结束节点的ID。type ∈ {AST, CDG, DDG}是边的类型，表

示该边是属于AST还是属于控制/数据依赖边。

生成两个CPG图之后，作者将两个图合并为一个统一的图结构，称为PatchCPG，作者用两个集合（V'，E'）描述PatchCPG，其中$V'=V_{pre}\cup V_{post}$，$E'=E_{pre}\cup E_{post}$。然后给每个节点/边的元组附加一个额外的元素，表示版本。因此，节点集V'的元素用三元组（id, code, version）表示，边集E'的元素用四元组（id1, id2, type, version）表示，其中version∈{deleted, added, context}是版本信息，表示节点/边属于PatchCPG的哪种类型的组件。

在PatchCPG中，没有必要在一个函数中包括所有的语句，因为其中有些语句与漏洞无关。为了找到与漏洞相关的上下文语句，作者以删除/增加的语句为标准，对源代码进行程序切片。作者将与补丁相关的切片语句定义为补丁上下文，因为只有这些语句与被修改的语句有依赖关系。切片可以在两个方向上进行：反向和前向分片。反向切片是为了定位漏洞的来源，作者从补丁增加行开始进行反向切片，即分析与增加行有数据依赖/控制依赖关系的语句；前向分片是为了找到受脆弱性影响的语句，作者从补丁的删除行开始进行前向切片。最终在经过切片后，可以获得PatchCPG，图4就是一个PatchCPG的例子。

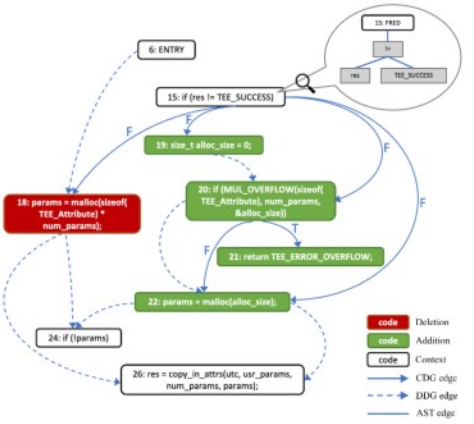


图4：一个合并后CPG的例子

切片完成后，作者对两个CPG中的边和节点进行嵌入（embedding）操作。边嵌入被用来反映两个节点之间的关系，节点嵌入是每个PatchCPG节点中代码的数字表示，它可以是CDG/DDG中的语句或AST中的代码标记。

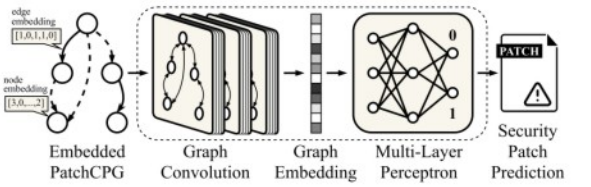
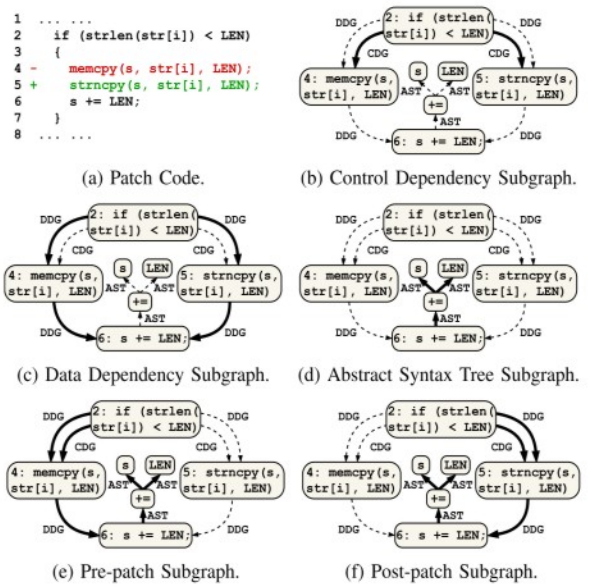


图5：PatchGNN的整体架构

对边和节点进行嵌入之后，作者设计了PatchGNN模型，用于预测给定的补丁是否是安全补丁。PatchGNN是基于图神经网络的，由于边属性的多样性（5种属性），作者设计了多属性图卷积机制构建PatchGNN模型。图5显示了PatchGNN模型的架构。在将节点/边嵌入转换为数字向量后，PatchCPG实例首先被送入三个多属性图卷积层。图卷积层将用不同子图中的邻域信息更新PatchCPG的节点嵌入。作者在PatchGNN中只使用3个卷积层，因为更多的卷积会导致图形过度平滑。在3层图卷积之后，图嵌入可以通过图池化和向量连接得到。图嵌入是一种图的表示方式，所有的节点、边和它们的特征都被转化为一个统一的向量。最后，作者利用多层感知器将图嵌入转换为二分类结果。

在多属性图卷积网络中，对于PatchCPG中的每个节点，PatchGNN中的卷积层倾向于通过不同类型的边从其邻居处收集信息。由于边属性的不同，所以不能使用一组统一的权重来进行模型训练。边嵌入的每个维度都可以表示PatchCPG中的不同关系，即前向切片/后向切片连接、控制/数据依赖和AST图。如图6所示，每条边可以有多个属性，而每一类属性都可以用来构建一个子图。由于每个维度的信息都对应着一个子图，作者把每个维度看作是一个单独的卷积通道。在PatchGNN的卷积层中，可以单独处理每个通道

的结构信息，并汇总所有子图的处理信息。



(a) Patch Code.
(b) Control Dependency Subgraph.
(c) Data Dependency Subgraph.
(d) Abstract Syntax Tree Subgraph.
(e) Pre-patch Subgraph.
(f) Post-patch Subgraph.

图6：PatchCPG的子图，每个子图是 PatchGNN中的一个单独的卷积通道

由于卷积层的输出是图，所以需要进一步处理以获得最终的预测结果。首先，作者利用池化层来减少数据维度并获得图嵌入。作者使用平均池化和最大池化来获得两个图表示，每个图表示都是一个高维向量。然后，这两个高维向量被串联起来，构建一个最终的图嵌入，它包含了PatchCPG中的所有节点、边和它们的特征信息。为了确定一个补丁是否与安全有关，作者建立了一个3层感知器，将图嵌入转化为2个单元的softmax输出（p0, p1），其中p0+p1=1。输出中的每个单元表示一个补丁实例属于非安全/安全补丁类别的概率。

在实验评估中，作者选择了两个基准数据集：PatchDB和SPI-DB。PatchDB是一个安全补丁数据集，包含12K安全补丁和24K非安全补丁。它包括一个从美国国家通用漏洞数据库（NVD）爬取的数据集和一个从GitHub提交中收集的数据集。PatchDB中的样本来自311个开源库（如Linux内核、MySQL、OpenSSL、Vim、Wireshark、httpd、QEMU等），可用于测试安全补丁分类能力。SPI-DB是另一个安全补丁数据集，最初从Linux、FFmpeg、Wireshark和QEMU收集，包括38291个补丁。

实验结果如图7所示，GraphSPD 在PatchDB上可以达到80.39%的准确性，F1分数为0.557。在SPI-DB上，GraphSPD的准确性为63.04%，F1分数为0.503。

Method	Dataset	General Metrics		Special Metrics	
		Accuracy	F1-score	Precision	FP Rate
TwinRNN [8]	PatchDB	69.60%	0.461	48.45%	19.67%
	SPI-DB	56.37%	0.512	49.07%	41.57%
GraphSPD	PatchDB	80.39%	0.557	77.27%	5.05%
	SPI-DB	63.04%	0.503	63.96%	19.16%

图7：基于GraphSPD和TwinRNN的安全补丁检测的性能比较

为了评估GraphSPD的适应性，作者其他流行的开放源码软件进行了案例研究，作者从NGINX的GitHub仓库中收集相邻主要版本之间的提交。如图8所示，NGINX 1.19.0和1.21.0（两个相邻的主线版本）之间有180条提交。总的来说，GraphSPD从787个输入提交中检测出27个潜在的安全补丁，27个中的21个被确认为真正的安全补丁。检测精度为78%（只有22%的误报），这与PatchDB等基准数据集上的表现一致，显示了GraphSPD相当大的泛化能力。

Changes w/	CVE	Total Commits	Valid Commits	Detected S.P.	Confirmed S.P.	Precision
1.19.x	3	180	127	7	6	86%
1.17.x	3	134	82	4	3	75%
1.15.x	1	203	120	7	4	57%
1.13.x	1	270	157	9	8	89%
Sum.	8	787	486	27	21	78%

S.P. = security patches

图8：在真实代码仓库中的安全补丁预测能力

谢子凡
2022级博士研究生
研究方向：软件安全
Email：xzff@hust.edu.cn

On Sharding Across Heterogeneous Blockchains

刘存扬　推荐

这篇文章是发表在2023年ICDE上的一个工作，其主要聚焦于解决异构区块链的跨链交易瓶颈问题。

随着目前，异构区块链在现实生活中的应用变得越来越广泛，不同区块链用户通过跨链交易相互交易，例如有一个在比特币和以太坊之间的交易，以太坊上的用户A在花费以太币的同时位于比特币上的用户B增加的资产是比特币。这其中就涉及到了越来越多的跨链交易，而目前，跨链交易的确认通常是通过中继链来完成的，所以中继链的性能就是跨链交易性能的瓶颈所在。本文提出了将分片应用于中继链的方式来提高系统的性能。

现有的工作比如Rapidchain[CCS'18]、Monoxide[NSDI'19]、Byshard[VLDB'21]这些工作是基于比特币UTXO模型的分片协议，这种模型并不适用于中继链上。其他分片协议通过智能合约来支持自定义的事务逻辑，例如Pyramid [INFOCOM'21]、Chainspace [NDSS'18]。这些相关工作存在三个问题：一是由于跨分片交易导致的吞吐量的下降，二是两阶段提交协议导致的长时间的确认延迟，三是一些顶部分片需要存储更多的交易数据，需要更多的存储容量。

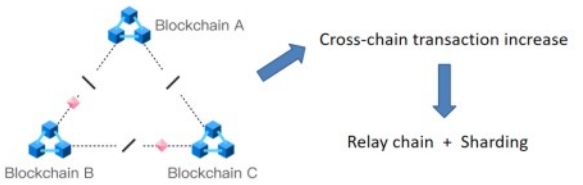


图1 跨链交易现状

现有的问题是，将分片技术应用于中继链上会导致大量的跨分片交易，这同样影响系统的交易性能。为了提高性能，本文通过将与跨链交易相关联的三种交易分配到一个分片中，如图2所示，完成这三个跨链交易的时间就是三轮共识，性能提高了三倍，理论上来说，这样可以完全消除跨分片交易。但是上述情况都是在一种前提下进行的，也就是一个跨链交易对应的三种交易都只需要验证一次就行。现在假设有六个跨链交易，前四个分别需要30次验证，后两个分别需要3次验证。前四个跨链交易分配给分片1，后两个分别分配给分片2和3，那么完成这六个交易需要120轮共识的时间。而如果采用最原始的办法，需要126轮共识的时间，仅仅提高了6轮共识的效率，这样的性能提升是不满足要求的。为了使性能最大化，本文的工作需要让分片大小最小。跨链交易的分配问题是一个NP困难的问题，本文将分配问题表示为一个具有最小分片大小的字典最小化目标，现有的整数优化求解器无法求解字典化最小化目标。所以，如何解决这个问题也是本文的重点之一。

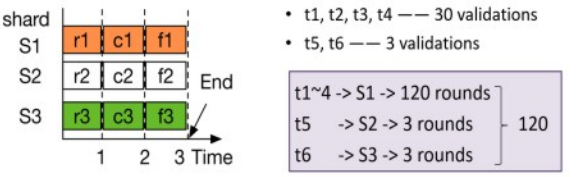


图2 分片跨链交易

本文的分片协议：Sliver。首先，通过一个可验证的随机函数VRF选出的Leader节点，Leader节点会生成并广播一个随机数。每个矿工利用这个随机数执行一种随机算法，确保矿

工在分片之间是均匀分布的。分片中的矿工m会获得分配到的跨链交易的子集，记为Am。当矿工接收到与交易ID关联的交易时，它会搜索子集，检查交易是否属于这个子集。如果是，矿工把这个交易添加到任务队列中；否则，删除这个交易。在实际实现中，这样的分配可以频繁地执行，因此确保能够有效地获得分配解决方案是很重要的。

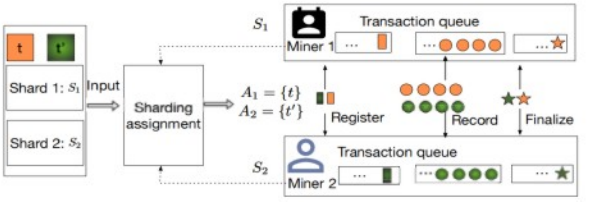


图3 系统流程图

分片中的分配方案可以显著影响分片大小，从而影响中继链的吞吐量。下面是相关的一个具体的例子。假设有两个分片S1和S2，以及4个跨链交易{T1、T2、T3、T4}，每个交易分别需要6、4、3、5次验证，共需要18次验证，不同的分配方案带来不同的共识时延。当分配方案中，最大分片大小最小时，共识时延最低，中继链的吞吐量最高。

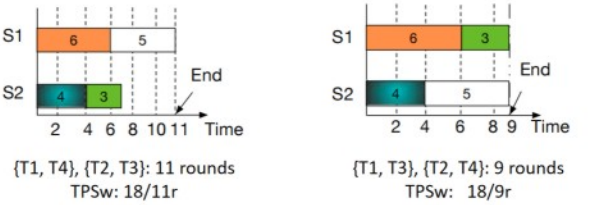
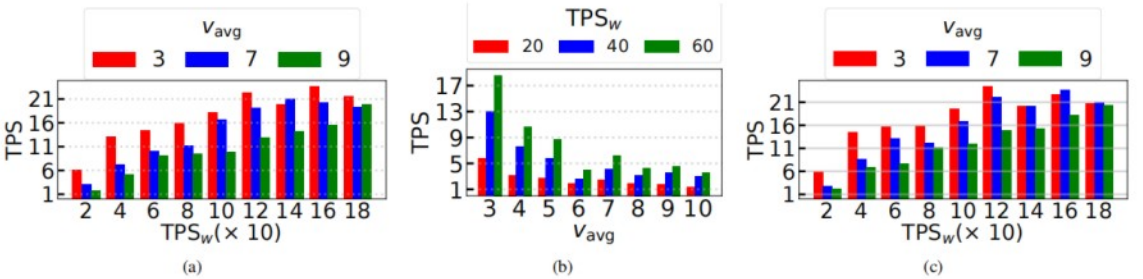


图4 交易分配样例

上述提到的分配问题被证明是NP难度的，经过作者的推导这个问题可以被转化成一个具有字典最小化目标的整数优化问题。字典最小化目标导致这个问题无法有效地解决，具有以下三个性质的整数优化问题可以被转化成一个线性规划问题：完全单模线性约束矩阵、标准最小目标、可分离凸目标函数。本文的分配问题满足这三个性质，则通过标准的线性规划求解器即可进行求解。



实验一 系统性能瓶颈影响因素

Vavg表示中继链上用于确认跨链交易的平均交易数量，TPSw表示中继链的吞吐量，TPS表示跨链交易的吞吐量。

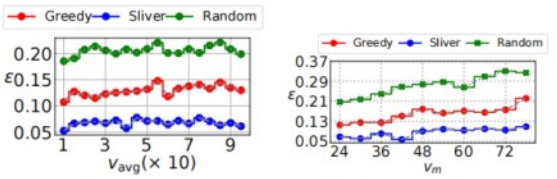
THE NUMBER OF CROSS-SHARD TRANSACTIONS IN PYRAMID.

v_{avg}, $S = 19$	3	5	7
Number of cross-shard transactions	3663	5973	8541
v_{avg}, $S = 19$	9	11	13
Number of cross-shard transactions	11283	14439	17493
v_{avg}, $S = 19$	15	17	
Number of cross-shard transactions	19275	22873	
S, $v_{avg} = 19$	3	5	7
Number of cross-shard transactions	0	6030	8847
S, $v_{avg} = 19$	9	11	13
Number of cross-shard transactions	10392	11019	11988
S, $v_{avg} = 19$	15	17	
Number of cross-shard transactions	12324	12387	

实验二 不同系统中继链上的跨分片交易数量测试

图(a)说明如果中继链的TPS增加，系统的TPS也会增加。图(b)进一步说明，无论平均交易数量有多大，当中继链的TPS从20增加到60时，TPS都会上升。图c是模拟实验，结果跟实际实验是一致的。可以确定，中继链是跨链交易确认的瓶颈。

该模拟实验比较了在Sliver和Pyramid之间中继链上的跨分片交易的数量。无论平均交易数量有多大或一个系统中存在多少个分片，Sliver中的跨分片交易数始终为0。如果Vavg或分片数量增加，Pyramid中的跨分片交易数量就会增加。

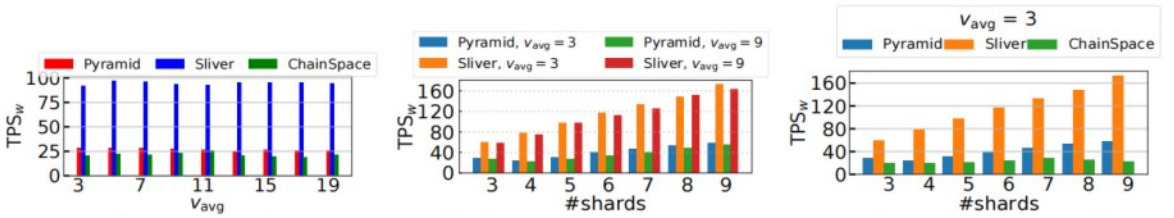


(a) ϵ with respect to v_{avg}. (b) ϵ with respect to v_m.

实验三 本文与贪心算法、随机算法的性能比较

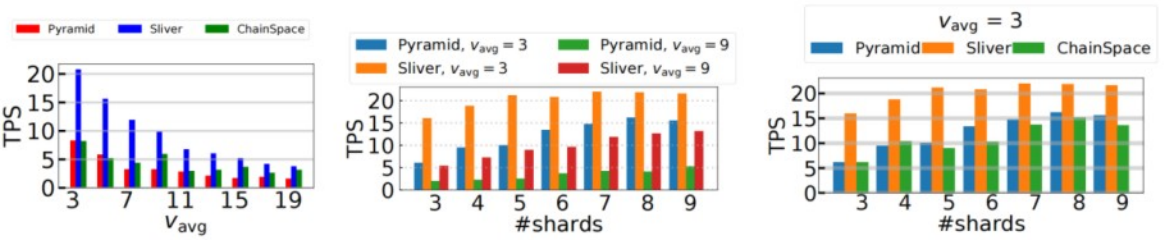
贪心算法迭代所有的跨链交易。在每次迭代中，它为每个跨链交易选择工作负载最轻的分片，并相应地更新所有分片的工作负载。随机算法为每个跨链交易随机选择一个分片。

纵坐标是一个与性能损失相关的因子，数值越小代表性能损失越少。相比其他两种算法，Sliver的性能损失是最少的。图(b)中可以看到性能损失随着vm的增大而增大，说明了系统需要更小的vm来提高性能。



(a) TPS_w with respect to v_{avg}. (b) TPS_w with respect to # of shards. (c) TPS_w with respect to # of shards.

实验四 不同系统中继链吞吐量比较



(a) (b) (c)

实验五 不同系统跨链交易吞吐量比较

图(a)可以看出Sliver的中继链吞吐量相比Pyramid和Chainspace有显著提高。在Sliver中，中继链的吞吐量与Vavg关系很小。在图b和图c中，随着分片数量的增加，Sliver的中继链的吞吐量呈线性增加，而Pyramid和Chainspace中增加要慢得多，这是因为跨分片交易会对吞吐量产生负面影响。

图(a)说明这三个系统的跨链交易的TPS均随着Vavg的增加而降低。Vavg越高，说明确认跨链交易需要的时间越长，从而间接减少了每秒确认的跨链交易的数量。图(b)中选取了两个不同Vavg的数值进行实验，分片数量分组增加。当分片数量不变时，跨链交易的TPS随着Vavg的增加而减少，验证了图(a)中Vavg对TPS产生负面影响的结果。结合图(b)(c)可以知道，无论Vavg有多大，Sliver都能显著改善TPS，原因是Sliver支持并行交易验证，消除了跨分片交易。

本文通过把相互依赖的交易分配给同一个分片从而完全消除跨分片交易；通过把NP-hard的跨分片交易分配问题转化成一个带有字典最小化目标的整数优化问题，再通过λ技术转化成线性规划问题，通过标准线性规划求解器进行求解，大幅提高了跨链系统存在的吞吐量问题，提升了系统的性能。

刘存扬
2022级硕士研究生
研究方向：区块链
Email：cunyangliu@hust.edu.cn

科研之路在于发现问题解决问题

陈　铭

在实验室做了几年科研，经历过低谷，也有过论文被顶会录用的喜悦。总的来说，我还是深感幸运。在毕业之际，和师弟师妹们分享一下我对科研的心得体会，希望能助大家找到适合自己的科研之路。

找到问题

科研的过程无非是找到问题再解决问题，剩下的便是水到渠成的事。万事开头难，找到问题应该是科研中最重要的一环，因为确实很困难，尤其是对于刚步入科研的同学来说。从我自身经历来说，找到问题的方法有两种：一是从论文中找问题；二是从实践中找问题。从论文中找问题就是看论文，看自己感兴趣的论文，而且最好是方向较新的论文，因为可做的东西会多点，也更容易找到问题。从自身角度来说，首先要搞懂问题背景以及支持的应用，从应用背景的角度去思考有哪些是现有工作不能解决的；其次就是把现有方法搞懂，理解为什么它们能够解决这类问题，再去思考现有方法的不足。当然，自己发现问题还是会比较难。所以最好是和老师做汇报，去讲讲你感兴趣有意思的论文。老师毕竟科研敏感度较强，会比你更容易看到问题。在汇报的过程，这样的交互及讨论也会促进发现问题。第二种实践中找问题，就是把现有工作的代码在不同的数据集、不同的设置下中跑一跑，分析分析。我的第二个工作就是在实验中发现了图流中高度点的问题。

解决问题

对于解决问题，首先最重要的自己要多思考。直接想到一个新颖的解决办法是很难的，也是要运气的。所以一开始可以先提出一个基本思路，能解决问题但是还不够好；或者先找找现有的能解决这类问题的方法。再在这基础上，针对自己问题的特性去提出一些创新。慢慢地，解决方法能够越来越完善，更有创新性。自己一直思考问题，有时候会陷进去跳出不来。这个时候需要改变一下，第一种方式是把这个问题先放一放，一天都不要去管。等自己放松下来（比如晚上洗澡或者跑步的时候），再去想想，有时候会突然蹦出灵感。第二种方式就是去给同学讲讲现在的一个思路，讲的过程你要让同学听懂，这样的梳理过程也会让自己发现一些不足；同时你的同学因为不懂你这个东西，会问你一些很奇怪但是你从来没有考虑到的问题，甚至给你的思路提出一些很荒诞但是确值得尝试的建议。这会对于优化自己的思路很有帮助。

总之，科研需要自己多努力，也需要同学老师的帮助，缺一不可。希望大家都能出好的成果，顺利毕业！

陈　铭
2019级硕士研究生
研究方向：图流概要结构
Email：mingc@hust.edu.cn