



2023年 第4期

江苏省计算机学会通讯

COMMUNICATIONS OF THE JSCS



—漫谈 ChatGPT
—“硬件与人工智能安全”的坚守者

江苏省计算机学会常务理事单位

南京邮电大学计算机学院、软件学院、网络空间安全学院

学院始于 1985 年设立的计算机通信系。目前设有计算机科学与技术系、软件工程系、网络空间安全系、数据科学与工程系、计算机基础教学中心、计算机实验教学中心（江苏省计算机基础实验教学示范中心，江苏省计算机软件与服务外包校企合作工程实践教育中心），拥有江苏省无线传感网高技术研究重点实验室、江苏省大数据安全与智能处理重点实验室、江苏省高性能计算与智能处理工程研究中心、计算机技术研究所、高性能计算与大数据处理研究所等科研平台，挂靠有南京邮电大学大数据研究院、南京邮电大学 - 盐城大数据研究院等机构。

学院师资力量雄厚，现有专任教师 173 人，其中教授 40 人，副教授 68 人，博士生导师 22 人，硕士生导师 110 人。拥有中国科学院院士、国家级和省级教学名师、国务院特殊津贴获得者、国家“杰出青年基金”获得者、教育部“长江学者”特聘教授和青年学者、国家高层次领军人才、国家海外优青、江苏省特聘教授、江苏省杰青等高层次人才，入选省“333 工程”、“青蓝工程”和“六大人才高峰”40 余人次；建有江苏省领军型人才团队、江苏省高校优秀科技创新团队、江苏省“青蓝工程”科技创新团队、江苏省“青蓝工程”优秀教学团队、江苏省“六大人才高峰”创新团队和江苏省优秀基层教学组织等。

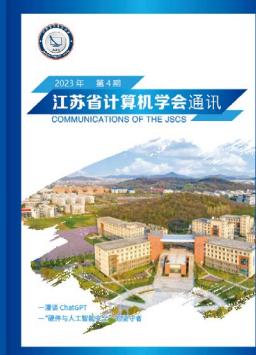
学院拥有“网络空间安全”一级学科博士后科研流动站和一级学科博士点，并共建“电子信息”专业学位博士点，一级学科硕士点 3 个，本科专业 5 个，其中国家一流本科专业建设点 3 个、国家特色专业建设点 2 个。

近年来，学院大力提升科技创新水平和社会服务能力，在网络信息安全、物联网与传感网、模式分析与机器智能、云计算与大数据处理、数据挖掘理论与应用、软件分析与服务组合等领域开展了深入研究。近五年，学院获批国家级项目 93 项，省部级项目 43 项，市厅级项目 49 项；获省部级科技奖 19 项。

学院一贯注重学生的创新意识和创新能力的培养，在国内享有良好声誉。本科生在国际和国内各项竞赛中成绩突出，全国大学生电子设计大赛最高奖“索尼杯”、美国大学生数学建模竞赛最高奖 SIAM 奖、中国国际“互联网+”大学生创新创业大赛国赛金奖、全国大学生数模竞赛一等奖、“挑战杯”全国大学生课外学术科技作品竞赛一等奖、2018 年“创青春”全国大学生创业大赛金奖、ACM-ICPC 亚洲区域赛金奖和亚洲东大陆赛区总决赛金奖、“问鼎杯”全国大学生网络信息安全与保密技能大赛特等奖、中国大学生程序设计竞赛（CCPC）金奖、全国大学生信息安全竞赛一等奖，“蓝桥杯”全国软件和信息技术专业人才大赛特等奖、第二届全国“红帽杯”网络安全攻防大赛冠军、厦门国际大数据大赛唯一的一等奖、全国大学生英语竞赛特等奖、江苏省大学生程序设计大赛冠军等。近三年，共获得国家级及以上重要奖项 500 多人次，省级重要奖项 600 余人次。

学院现有在校生 4509 人，其中博士生 62 人，硕士生 1208 人，本科生 3239 人。学院本科毕业生理论基础扎实，综合素质高，就业率始终保持在 99% 以上。本科生可以通过校际合作和交换计划赴国内外著名大学进行联合培养和合作交流；优秀毕业生可免试攻读硕士、博士学位研究生，近年来学生获得全国高校百名研究生党员标兵、中国青年志愿者优秀个人、中国电信奖学金、江苏省疫情防控先进个人等国家级、省级荣誉称号。





顾问委员会

主任: 周志华

副主任: 武港山 耿新 胡江溢
李斌 夏士雄 李凡长
陈兵 詹永照 程光
委员: 罗军舟 肖亮 申富饶
陶先平 吉根林 胡孔法
张道强 黄强 邓建明

编委会

主编: 路通

副主编: 聂长海 张洁

编委: 钱柱中 游辉敏 石克

地址: 中国江苏省南京市栖霞区仙林大道 163 号

邮编: 210023

电话: 025-89680909

邮箱: jscs@njn.edu.cn

江苏省计算机学会通讯

COMMUNICATIONS OF THE JSCS

学生记者团专栏

01 | 智能物联网普及之路在何方——郭忠文教授专访

03 | 漫谈 ChatGPT——王崇骏教授专访

07 | 大小模型端云协同智能计算——刘生钟老师采访

教学成果交流

10 | 《知识表示与处理》课程体系建设教学成果交流 | 赵一铮

政策解读

14 | 推动科技向善 把好伦理“方向盘”

学术交流

16 | 基于应用需求的量子密钥分发网络路由技术研究 | 赵梦楠

22 | 贝叶斯网络结构学习优化算法研究

会员风采

33 | “硬件与人工智能安全”的坚守者 | 薛明富

36 | 在“边缘智能计算系统”的科研道路上求实创新 | 周俊龙

科学普及

38 | 觉悟 ChatGPT, 科研第五范式

科创成果

46 | 组合测试理论与方法研究

学会动态

02 | 第 9 届 EAI 物联网即服务国际会议在金陵科技学院隆重召开

06 | 聚势启新·创引未来 | 2023 国产高性能计算技术与应用论坛在南京召开

09 | “数字赋能 智慧职教” -- 职业教育数字化转型发展研讨会成功举办

13 | JSCS 老科委到南京邮电大学开展科技服务

32 | “江苏工匠” 岗位练兵职业技能竞赛 - 物联网安装调试职业技能竞赛、工业互联网技术应用职业技能竞赛在南京、南通举办

35 | JSCS 走进泰州学院

59 | 第五届江苏省计算机网络与云计算新技术研讨会在盐城成功召开

60 | 江苏省大学生计算机设计大赛产学研赛创新高峰论坛暨十周年纪念大会成功举办



智能物联网普及之路在何方

——郭忠文教授专访

郭忠文，二级教授，博士生导师，中国海洋大学信息学部副部长，国家基金委重大仪器项目首席科学家，ACM SIGAPP主席，山东省计算机科教联合体副理事长，山东省软件工程领域学科评议组成员，CCF物联网专委会以及网络与数据通信专委会委员。主持863、908、国家自然科学基金等多项国家级项目，以及国家重点研发、973、国家自然科学重点基金、国家海洋公益专项、国家海洋可再生能源专项等多项国家级课题。在包括ACM Computing Survey、TPDS、TVT、TIM等知名国际期刊和INFOCOM、ICDCS、RTSS、SENSYS等国际会议发表多篇具有影响力的学术论文，累计发表SCI、EI收录学术论文150余篇，主持制定4项物联网领域IEEE国际标准。荣获青岛市科技进步一等奖和山东省教育成果一等奖各一项，讲述的《物联网系统设计与开发》课程荣获国家级一流课程。取得的研究成果已在多个大型企业事业单位中得到推广应用，产生较大经济效益和社会效益。

在11月11日的第五届江苏省计算机网络与云计算新技术研讨会中，郭教授作了题为《智能物联网普及之路在何方》的报告，针对智能物联网技术在实际应用过程中面临的诸多难题，例如传感设备多样、应用场景复杂、原有系统升级改造成本高、缺少标准规范等，给出了独特的见解和解决思路，期望加速我国智能物联网的普及进程，让广大民众和企业更便捷地享受物联网智能化带来的红利。

在报告结束后，我们有幸采访到了郭忠文教授。

Q1：边缘计算也是一个目前非常热门的研究方向。在边缘计算与智能物联网之间，它们有何交集和差异呢？

郭忠文教授：边缘计算和智能物联网之间存在显著的交集，但也有一些关键差异。首先，边缘计算的核心在于将数据处理、存储和分析功能从中心化的数据中心转移到网络的边缘，即靠近数据源的位置。这样做的主要好处是减少延迟，提高处理速度，同时减轻中央服务器的负担。

在智能物联网系统中，边缘计算扮演着至关重要的角色。物联网设备通常会产生大量数据，如果这些数据都发送到云端进行处理，将会造成巨大的网络负担和处理延迟。而边缘计算能够在数据产生的地点即时处理这些信息，从而实现更快的响应时间和更高

效的数据管理。

然而，边缘计算和智能物联网之间也存在一些差异。例如，物联网强调的是设备的互联和通信，以及不同物联网节点的数据互操作，而边缘计算更注重于数据处理和响应的速度。

总体而言，边缘计算为智能物联网带来了重要的技术优势，使得物联网设备能够更加智能和高效地工作。随着这两个领域的不断发展和融合，我们可以期待出现更多创新的应用和解决方案。

Q2：您在报告中提出了一套完整的体系，您认为时下的一些新兴技术，如人工智能和区块链，能否融入您提出的智能互联网体系，并为之提供参考和帮助？

郭忠文教授：当然，人工智能和区块链技术都是革命性的创新，它们在智能互联网体系中扮演着关键角色。AI技术能够对大量数据进行分析和学习，从而提供更加智能的决策支持和自动化服务。在智能物联网的各种应用场景中，AI能够使设备更加智能，提供个性化和更高效的服务。

区块链技术在提高安全性、增加透明度和促进数据共享方面发挥重要作用。在物联网的背景下，区块链可以确保数据的安全和不可篡改。此外，区块链的分布式账本特性可以帮助建立去中心化的智能物联网结构，增强系统的抗干扰能力和可靠性。

因此，这些新兴技术不仅能够融入智能物联网体系，而且将极大地推动其发展，使其更加高效、安全和智能。

在智能物联网领域迅猛发展的当下，郭忠文教授的深刻见解为我们展现了一个极具价值的视角，面对“智能物联网普及之路在何方”这一问题，郭教授给出了“体系结构 + 标准规范 + 操作系统 + 开发平台 + 商业模式”的回答，为加速智能物联网发展提供了切实有效的指导。如今，智能物联网的兴起为我们带来前所未有的机遇，同时也提出了新的挑战，让我们以开放的心态和创新的精神，共同探索智能物联网在未来社会中的广阔应用前景。

学会动态 •

第9届EAI物联网即服务国际会议在金陵科技学院隆重召开

2023年10月27至29日，由欧洲创新联盟主办，金陵科技学院、江苏省通信学会和江苏省计算机学会共同承办，中山大学、西北工业大学和北京邮电大学技术协办的第9届EAI物联网即服务国际会议在金陵科技学院隆重召开。来自全球近50位专家、学者出席活动。





漫谈 ChatGPT

——王崇骏教授专访

1、您认为 ChatGPT 在从 AI 到 AGI 的发展过程中起到了什么作用或者说有什么意义？

在这个转变的过程中，存在哪些挑战和机遇？

ChatGPT 是将 GPT 技术应用于 Chat 场景的人机对话系统，它为传统 NLP 中的许多问题提供了“一统化”的解决方案。这意味着使用同一模型可以实现多个任务，似乎表明当模型规模达到一定程度时，可能会出现“涌现”现象（尽管目前仍有很多研究者对 ChatGPT 的“涌现”持保留意见）。

ChatGPT 代表了语言处理和对话交互领域的技术进步。虽然，就目前而言，我个人认为 ChatGPT 本身还不足以被称为 AGI，但是 ChatGPT 插件生态的开放和构建为实现 AGI 提供了一种潜在途径。例如，ChatGPT 可以借助插件生态，通过赋能人机交互来实现多个面向特定领域的专用 AI 系统的融合。这些融合的 AI 系统实际上是面向多个领域的通用人工智能，因此，从这个意义上说，ChatGPT 为 AGI 提供了一条可行的路径示范。

此外，我认为 ChatGPT 在推动对话系统的发展、探索语言理解和生成技术方面具有引领和促进的作用，同时也面临着更多的挑战，例如改进对话系统的技术，加强人机交互的理解和合作等方面。

对于转变过程中存在的挑战与机遇，我认为有以下几个方面：

(1) GPT 本身是一个数据驱动的 AI 模型。数据驱动的本质是归纳逻辑，但归纳逻辑存在天然缺陷，即可能从真实的样本中归纳出错误的结果。针对这个问题，目前主流的方法是在数据驱动的研究和建模中引入逻辑驱动的知识。ChatGPT 也允许用户提供提示 (Prompt) 和思维链 (COT) 等来优化生成结果。在思维链的输入中，一些研究将一些规则样本 (训练) 思维链然后植入 ChatGPT，以获得更好的效果。这代表了数据驱动和常识知识相结合的路径。

(2) 由于 ChatGPT 是数据驱动的，不可避免地会面临隐私泄露、偏见和歧视、道德伦理等问题；且 ChatGPT 训练过程中使用的数据来源复杂且规模庞大，训练样本的复杂性使得模型输入输出更难以解释。

(3) 对于以 ChatGPT 为代表的生成式人工智能，很难界定其生成内容的版权归属和内容的“内涵”（是否源自于某些受版权保护的训练样本）。此外，作为一个生成式系统，无法保证其输出内容的真实性以及是否符合人类世界的道德、法律、伦理和主流价值观。

(4) 归根结底，AI 是赋予人类工具理性的工具。就目前的研究来看，所有的 AI 算法

都无法具备人类的价值理性，AI的应用场景也仅限于仅需要展现工具理性的场景。然而，是否应该赋予AI价值理性以及如何赋予AI价值理性呢？这是个难题，前者涉及伦理道德问题，后者则是技术路径上的难题。

总体来说，ChatGPT是AI向AGI发展过程中一个重要步骤，它不仅代表了技术的进步，也暴露出了一些重要的挑战，这为未来的研究和探讨提供了有力的启示和思考。

2、随着数字经济的发展，数据安全和隐私问题越来越受到重视。您认为ChatGPT等人工智能模型在数据处理和使用过程中应如何更好地保护用户隐私和数据安全？

隐私保护和数据安全必须基于法律保障和技术支持：

(1) 目前，隐私保护、个人信息保护和数据安全受到多部法律的保护。在中国，涉及到的相关法律至少包括《民法》、《个人信息保护法》、《数据安全法》、《网络安全法》等；在我们使用类ChatGPT等技术时，首先要具备法律意识，做到“涉密不上网，上网不涉密”。

(2) 在法律保护的前提下，当数据被授权使用时，存在一系列隐私计算技术可以确保数据在使用过程中不被窃取、篡改等，实现所谓的“可用不可见”。目前的隐私计算技术包括密码学（如数据脱敏、差分隐私、同态加密等）、联邦学习、安全多方计算、可信计算平台等。同时，可以采取一系列安全监管类软件和措施，例如法律合规性检查、数据安全监测和审计、匿名化和脱敏、数据加密、访问控制和权限管理等。此外，也可以采用安全的开发和测试实践，包括漏洞扫描、代码审查和安全测试等，以确保模型本身的安全性，并减少潜在的漏洞和风险。

3、您认为以ChatGPT为代表的预训练模型在推动数字经济发展，尤其是在助力传统行业数字化转型方面，可能会起到什么作用？又会面临怎样的挑战和机遇？

(1) 对于像ChatGPT这样的大型模型，我们无法采用传统的“预训练+微调”模式。相反，我们必须使用提示学习技术来实现生成的内容与需求目标的匹配。因此，在类ChatGPT的大型模型中，提示学习的意义将变得越来越重要

(2) 在ChatGPT盛行的当下，出现了一种以“斯坦福羊驼”为代表的小型模型实现路径。该路径利用更小的预训练模型（“斯坦福羊驼”使用的是Meta的大语言模型LLaMA 7B）通过垂直数据（数据其实来自于GPT，类似于知识蒸馏）的微调来满足更垂直领域的需求。“斯坦福羊驼”的成功一方面展示了大型模型和知识蒸馏的能力和应用潜力，另一方面也为那些通常无法开发大型模型的中小型企业，在大型模型时代提供了一种可能的技术实现路径

(3) 对于企业而言，数字化转型通常采取以下一般做法：开发创新的数字化产品和服务、替代或拓展原有的数字化产品或服务、数字化客户界面与工具、优化管理实践（组织结构）、用户培训（新技术、新思路）。从实践角度来看，像ChatGPT这样的大型模型能够胜任几乎所有的岗位，包括人机交互、工具性和记忆性教育教学、多边交流协作、业务系统的RPA调用等。



4、您在数字经济与实体经济融合创新发展论坛中提出“元宇宙将成为下一代互联网的新战场”，您认为 ChatGPT 在元宇宙中可能会扮演怎样的角色？对于元宇宙和人工智能的结合，我们需要关注哪些重要的研究问题和关键挑战？

(1) ChatGPT 的出现进一步推动了人工智能生成内容 (AIGC) 的发展，吸引了更多人的关注。从互联网内容的发展历程来看，最早期的专业生产内容 (PGC) 和后来的用户生成内容 (UGC)，AIGC 应该是下一步的发展必然。而从用户需求的角度来看，在未来的元宇宙世界中，PGC 和 UGC 已经远远无法满足用户对内容个性化和多样性的需求，因此 AIGC 也必然会成为 PGC 和 UGC 的重要补充。

(2) 按照“元宇宙未行，虚拟数字人先行”的说法，未来的虚拟数字人，无论是面向商业 (to B) 还是面向消费者 (to C)，都需要与各方进行交互。例如，作为虚拟助手，它可以通过与用户对话，提供信息、回答问题和提供导航服务，帮助用户更好地理解元宇宙的环境、功能和交互方式。另一方面，作为社交互动伙伴，它可以扮演虚拟朋友、合作伙伴或虚拟角色，与用户进行交谈、分享经验并提供情感支持。因此，由于 ChatGPT 在语言处理、语义理解和对话管理方面具有优势，在未来的元宇宙中，ChatGPT 将发挥重要作用。

而对于元宇宙和人工智能的结合，我认为以下几点值得关注：

(1) 从技术角度来看，元宇宙在底层技术上依赖于区块链，用于构建可信的互联网环境。而在元宇宙的体验层面，主要依赖于 VR、AR、MR 等技术，通过这些技术实现用户的体验感、沉浸感和自然交互感等。而在两者之间，有许多 IT 技术共同支撑，包括大数据、交互技术、游戏技术、人工智能技术、网络技术、物联网技术、云计算等等；

(2) 用于元宇宙的 AI 技术涉及到至少 CV、ML、NLP 等方面，具体的研究内容包括但不限于：

- 1) 虚拟数字人：包括人物合成、人物表达、合成显示、识别感知、分析决策等方面的研究；
- 2) 虚拟身份和用户认证：如何建立可信、安全的虚拟身份验证和认证系统，以保护用户隐私并防止虚拟身份的滥用；
- 3) 内容生成和管理：如何有效地生成、管理和筛选元宇宙中的内容，以确保内容的质量、多样性和适应性；
- 4) 人机交互和用户体验：如何设计直观、自然的交互界面，以及如何提供个性化、情感化的用户体验；
- 5) 法律、道德和伦理问题：包括隐私保护、权益保护、道德标准、版权保护等方面的问题。

5、总体来讲，您对于 AI 的未来发展有何看法，抱有怎么的期待或是存在何种顾虑？

总体而言，关于人工智能的未来发展，我抱有理性的乐观态度。我认为在很长一段时间内，人工智能将成为人类的优秀助手，大大弥补人类在工具性理性方面的不足，使人类胜任以前难以应对的工作，并实现与 AI 的有效合作。有人认为人工智能或将导致许多工作岗位的消失，这种说法或许是正确的，但需要注意的是：随着技术和经济的发展，工作岗

位在人类发展过程中一直在消失，并同时涌现出新的岗位。因此，取代人类并使其失去工作岗位的不是人工智能等新技术，而是那些会运用人工智能和高科技的人；

我期待看到 AI 在医疗、能源、交通、环境等领域的进一步突破，改善人类的生活；企业可以通过智能化和数字化转型，实现成本优化、效率提升、产品创新和销售促进，释放人们的时间和精力，进而专注于创造性和高价值的工作；而作为智能辅助工具，AI 能够帮助人们拓宽思维，增强决策和问题解决能力。

然而，对于 AI 的发展，我们也有一些顾虑。例如，过度依赖 AI 可能会削弱人的决策主体性，引发各种风险。此外，AI 无法判断对错、善恶，如果被不良人士利用，可能会带来负面效应。当前，AI 在训练过程中大多无法受到人类法律、道德和伦理的约束，这使得 AI 伦理成为一个 important 且迫切需要关注的问题。而现有关于人工智能的法律和监管机制还不够完善，使得 AI 可能处于无监管的“野蛮生长”状态，这是危险的。因此，我们需要深度科普 AI，当 AI 犯错误时，人们能有正确的认识，防止因恐惧和误解产生的塔西陀陷阱，即人们对 AI 失去信任。

学会动态 •

聚势启新 · 创引未来 | 2023 国产高性能计算技术与应用论坛在南京召开

10月20日，一场科技盛会如期而至。由江苏省计算机学会高性能计算专业委员会主办，江苏奥工信息技术有限公司、南京大学高性能计算中心、东南大学大数据计算中心联合承办的“2023 国产高性能计算技术与应用论坛暨江苏省计算机学会高性能计算专委会 2023 年秋季会议”圆满召开。





大小模型端云协同智能计算

——刘生钟老师采访

在网络与先进计算前沿技术论坛中，刘生钟老师代表吴帆老师团队做了题目为《大小模型端云协同智能计算》的报告。人工智能的快速发展与广泛应用已成为经济社会发展的强大引擎。大小模型端云协同进化作为人工智能关键前沿技术，可有效克服云侧集中式学习范式在可扩展性、实时性、个性化、负载成本、隐私安全等方面的不足，已成为产学研的焦点，并被 **Gartner** 和阿里巴巴达摩院等国内外知名机构预测为革新智能计算范式的科技趋势。在报告中，刘生钟老师追寻端智能技术的发展脉络，分享了团队在端侧智能推理、大规模联合学习以及端云协同分布式智能支撑系统等方面的研究进展，对该领域的未来发展做出了展望。

在报告结束后，我们有幸采访到了刘生钟老师。

Q1: 您认为，如何在端云协同智能计算中更好地平衡效率、性能、成本和隐私？这方面有哪些新的策略或技术值得探索？

刘生钟：在端云协同的情境中，综合考虑效率、性能、成本和用户隐私是很重要的。首先，成本是非常重要的因素，特别是如果想要商业化。在这种情况下，端设备的成本必须控制在一个很低的范围内，而端设备的计算能力是有限的。因此，我们尽可能在端设备上执行少量的轻量级计算，而将更多的计算留给边缘服务器或云服务器。然而，这会带来另一个问题，即如何保护用户的隐私。我们可以首先对数据进行压缩，然后卸载一部分计算。在 2020 年，我们进行了一项工作，使用单层卷积层类似于压缩感知的方式对数据进行压缩。这种方式计算非常轻量，我们可以将压缩后的数据传输到边缘服务器或云服务器上，然后使用深度网络对其进行解码，并进行后续的推理。这种机制是非常灵活的，可以在输入层进行，也可以在特征层进行，但关键是它是非对称的，使用云服务器上的计算资源来节省通信资源。关于深入探索的新策略，一种可能的方向是考虑视频流的情况，如何结合时空冗余性和相关性来进一步减少压缩数据的尺寸。另一方面，随着端设备上的计算能力的提升，我们可以在端设备上执行更多的计算，但必须考虑能耗情况。我们需要在端和云之间做出权衡，同时考虑传输的稳定性，如果网络信号不稳定，我们可能会选择在端设备上执行更慢的计算，但仍然能够获得推理结果。换言之，端设备需要具备一定的计算性能，即使平常不会用到。

Q2：随着物联网、自动驾驶和智慧城市等领域的发展，端云协同智能计算将如何推动这些领域的创新和变革？您在报告中提到了一项多相机利用空间冗余实现减轻负载的工作，这是出于什么样的考虑呢？

刘生钟：端云协同智能计算的目的不是完全取代人的工作，而是使人的工作变得更简单。例如，在物流调度中，使用智能模型和算法可以大大提高效率和准确性，同时减少人的工作量。但人仍然需要在监管角色中出现。端云协同应该是一个整体，整体的效率和精度需要提升，同时也要提高人的体验。同样以物流调度为例，在外卖配送环节，现在通过使用智能算法，调度可以变得更快更准确，这不仅仅是因为配送员数量的增加，更是因为调度算法的效率有了显著的提升。在端云协同的环境下，我们可以优化配送路径，减少往返时间，以提高整体效率。在更广泛的层面上，端云协同智能计算涉及到组合优化、实时调度，以及模型设计和资源分配等问题。这些问题非常复杂，不仅仅是单个团队能够解决的。但是，总的来说，目标是提高效率和精度，同时改善人的体验。关于多设备协同的问题，通过多设备协同，可以实时监测所有角度和位置，从而获得更广的覆盖范围。相比于使用单个摄像头，这种方法可以避免在某个方向上监测时忽略其他方向的问题，同时也利用空间冗余性，减少了单个设备计算所需的数据和计算的复杂度，进而单个设备的成本也能得到控制，可以更容易地大规模部署设备。

Q3：从教育和人才培养的角度来看，未来端云协同智能计算领域需要哪些新的技能和知识？研究型高校应如何适应这些变化？

刘生钟：现在很多学生都倾向于涌向深度学习和人工智能这些热门领域。但我认为，教育不仅仅是追求当前的热点，而是要培养学生具有深厚的基础和广泛的视野。首先，高校应该强调基础知识的教育。无论是深度学习、边缘计算还是其他领域，扎实的数学、统计和计算机科学基础是至关重要的。通过掌握这些基础知识，学生将能够更加灵活地适应不同领域的变化。其次，我们应该教育学生批判性思维和问题解决能力。这包括教育他们如何从实际应用出发看问题，而不仅仅是在特定数据集上优化模型的性能。这涉及到对深度学习模型的通用性、可扩展性和在实际环境中的应用效果进行深入思考。另外，我认为我们应该鼓励学生探索AI之外的领域。AI领域可能会变得饱和，而其他领域，如边缘计算，也有很多值得探索的问题。这就要求我们在教育过程中培养学生的多样性和多领域的兴趣。在边缘计算领域，学生应该关注如何在有限的资源下实现高效的模型推理，这包括考虑计算时间、内存和能耗等因素。这一领域的目标不同于传统深度学习要让模型变得更大，而是要让它们变得更小、更高效，适应性能受约束的设备和应用，从而提高智能计算的普适性。

在这个日益数字化和智能化的世界中，端云协同智能计算将成为推动各行各业进步的关键动力。刘生钟老师的观点为我们揭示了端云协同智能计算领域的挑战和机遇，展示了在效率、性能、成本和隐私之间寻求平衡的重要性。他强调了通过创新技术以及教



育和人才培养来推动这个领域的发展。同时，刘生钟教授强调基础教育、批判性思维和多领域兴趣的培养，为我们指明了适应不断变化的科技环境和应对未来挑战的道路。在这个充满无限可能的时代，让我们积极拥抱创新、追求卓越，在端云协同智能计算的大潮中勇往直前，携手共创智慧、高效和人性化的未来世界。

学会动态 •

“数字赋能 智慧职教”--职业教育数字化转型 发展研讨会成功举办

10月21日，职业教育数字化转型发展研讨会暨职业院校产教融合专委会年会在江苏海事职业技术学院邮轮培训中心成功举办。此次会议由江苏省计算机学会、中国计算机学会南京分部及教育部职业教育信息化教学指导委员会联合主办，江苏海事职业技术学院及江苏省计算机学会职业院校产教融合专家委员会承办，教育部教育技术与资源发展中心职业教育发展处处长赵林、教育部职业教育工业与信息化指导委员会副主任眭碧霞、国家数字化学习资源中心副主任吕英娜、江苏省计算机学会秘书长金莹、江苏省职业技术教育学会秘书长李振陆、江苏海事职业技术学院校长温华兵等领导出席大会，56所职业院校、行业协会、企业代表共130余人参加此次会议。本次大会吸引了众多高校代表积极参与，为各高校数字化转型搭建了重要的交流平台。与会代表们结合当下职业技术教育发展的新变化、新方式、新技能等多方态势，就职业教育数字化背景下资源建设与应用、管理与服务、教学能力与实践、产教融合创新及数据基层建设等层面进行深入解读和探讨。大会为推动高校教育信息化高质量发展贡献了实践经验和创新活力，有利于各高校共同推进信息化建设和数字化创新发展，开启高校数字化转型精彩篇章。



《知识表示与处理》课程体系建设教学成果交流

——南京大学赵一铮副教授

简介

知识表示与处理是人工智能早期发展的核心领域，它的目标是让机器拥有人类的经验和知识，并且能够按照某些规则推理得到新的知识，最终运用这些知识来辅助机器进行计算和决策。把人类的理性行为变成计算模型是该领域的核心研究问题。大模型时代，知识表示与处理将在增强 AI 系统的可解释性、可靠性、鲁棒性等方面起到至关重要的作用，被认为是未来 AI 系统不可缺少的组成元素。

赵一铮博士在南京大学开设的《知识表示与处理》课程是国内首个系统性地涉及该领域理论与技术、设有完整教学大纲的专业课程。此课程从知识表示与推理的历史脉络切入，深入讲解知识表示与处理领域的基础概念、经典理论、技术及应用，并重点探讨领域的前沿问题。它包括了基于集合论、模型论、证明论、递归论以及计算理论的各种分析方法。学生将全面掌握这些理论和方法，并得以在实际中展现其初步工程实践能力。课程旨在引导学生深入挖掘事物的内在逻辑和哲学，而不仅仅是将知识作为记忆的目标。

该课程已连续开设四年，期间获得了学院与学校的鼎力支持。在教学过程中收到了学生们的积极反馈。经过多次优化和调整，最终确定了目前的教学大纲。其中的教学内容也已整理为教材，这本由机械工业出版社即将推出的教材将成为国内首部知识表示与处理领域的专业书籍。南京大学于 2019 年 5 月发布全国首个《人工智能本科专业教育培养体系》，该课程与相关教材成为该培养体系的核心组成部分。

主要解决的教学问题及解决教学问题的方法：

知识表示与处理不仅是人工智能领域的起源和核心子领域，更是知识驱动范式的基石。然而，截至 2019 年为止，国内尚未有高校开设相关课程。由此导致的问题是：我们在课程建设方面缺乏经验，并没有成熟完善的课程体系可供参考，需要从零开始搭建完整的课程体系。在这样的背景下，我们不可避免地遇到了许多教学问题。然而随着实践教学经验的积累，我们已经总结出了一套行之有效的解决方法。以下我们将对这些方法进行详细阐述：

教学问题一：

尽管许多国外一流高校，如牛津大学、普林斯顿大学、麻省理工学院和斯坦福大学等，都在不断推进该课程的建设，但直接借鉴它们的建设经验并不可行。主要原因是，



这些国外高校的课程体系中，存在一定的前序和后序课程进行铺垫和衔接，而这些课程在国内的体系里或是缺少，或是存在各方面的差异。例如，虽然“数理逻辑”这门课程为学生奠定了基础逻辑学知识，但其主要是从数学角度进行教授，而非从人工智能或知识表示与处理的视角出发。这种教学方式导致许多知识点无法有效连接，使得“数理逻辑”成为了形式化的前序课程，不能有效铺平“知识表示与处理”的学习道路，从而增加了学生的学习难度。

对应解决方法：

为适应人工智能的多个不同子领域，我们精心设计了一系列升阶课程路径，并将每一路径组织为一个独立的课程模块。每个模块将涵盖3至4门相关专业课程，且某一课程可在多个模块中重复出现。通过此方法，学生可以清晰地识别不同课程之间的联系，以及它们与人工智能各专业方向的相互关系，从而更科学地规划自己的学习计划，有计划、有预见地完成四年学业。

以《数理逻辑》、《知识表示与处理》和《智能规划与推理》为例，这三门课程被整合为一个模块下的三个课程单元。在其中，《数理逻辑》作为其他两门课程的先修课，将在大一下学期作为专业必修课开设。该课程不仅涵盖了命题逻辑、一阶谓词逻辑、模型论、集合论和证明论等数理逻辑的经典理论和知识，还将介绍这些理论在计算科学和人工智能领域中的关键应用。这一点区别于数学系或其他学院开设的同名课程内容。简而言之，该课程将基础理论与专业知识相结合，展示数理逻辑基础知识在计算机科学和人工智能实际应用中的体现，旨在激发学生的学习兴趣和研究热情。而《知识表示与处理》作为《数理逻辑》的后续课程，《智能规划与推理》作为前两门的后续课程，分别安排在大二下学期和大四上学期开设。

教学问题二：

专业课程的内容往往深受其主讲教师个人研究兴趣的影响。这不仅涉及领域内的经典理论和技术，还包括许多独特且属于较为细分领域的知识点。然而，这样的教学方向可能并不有利于学生全面、系统地掌握该学科领域的基础知识和核心概念。因为这种方法可能导致学生的知识体系偏向某一方面，从而缺乏一个更为全面和均衡的认识和理解。再者，由于国外《知识表示与处理》课程体系建设历史悠久，其中不可避免地蕴含了某些时代的烙印。举例来说，一些使用的符号与当前的符号体系有显著差异，而某些知识点也已经逐渐陈旧，不再是该领域的核心研究内容。因此，如何审慎选择和权衡保留哪些传统知识点，以及如何融入最新的基础知识，并在有限的教学时间内将这些知识点以逻辑清晰的方式串联起来，无疑构成了一项巨大的教学挑战。

对应解决方法：

我们对国外高校课程体系和内容进行了深入调研，评估了每个知识点的讲授必要性、前沿性以及难度。在此基础上，我们将教学内容划分为两大部分：一部分是不随时代变化的核心基础知识；另一部分是随研究进展而发展的技术变革。对于核心基础部分，我

们制定了严格的课程标准和核心指导方针，并明确列出了学生必须掌握的基础知识和核心概念。为确保这些标准和方针得到有效实施，我们还设定了相应的评估机制。针对技术变革部分，我们定期更新课程内容，以准确反映最新的研究发展和技术变革。此外，我们还提供丰富的参考书目和在线资源，支持学生进行自主学习，从而拓宽其知识面。这样的设计旨在为学生提供一个全面而深刻的学习体验，帮助他们牢固掌握本学科的基础和前沿知识。

创新点

创新点一：国内首个《知识表示与处理》课程建设和教材建设项目

“知识表示与处理”作为人工智能的起源和重要子领域，经过近 70 年的发展，已成功构建一套完整的理论体系，成为整个人工智能学科体系中不可或缺的一环。国际上，诸如牛津、斯坦福和 MIT 等顶尖院校，均有卓越的科研团队和课程支持服务于此领域。然而，国内对该领域的关注程度相对较低，涉及相关课题研究的团队和提供相关课程的高校相当稀少，大部分工作主要集中在数据驱动的人工智能技术上，而相关的中文教材更是几乎空白。

南京大学率先开设《知识表示与处理》课程，成为国内首家提供此类课程的高校，课程内容覆盖知识表示与处理多个前沿研究领域。此课程的推出，打破了现有以数据驱动为主的人工智能课程的单一格局，构建了以知识驱动为核心的人工智能课程。这不仅帮助学生更好地了解知识表示与处理的发展历史，还引导学生从更广阔的视角审视所学的人工智能相关知识，同时也使学生和从事人工智能其他领域研究的学者能够更深入地理解知识表示与处理领域的前沿科学问题，为其未来的研究工作提供启发。目前，申报人目前已与机械工业出版社签约，计划在 2023 年之前出版《知识表示与处理》中文教材一册。该教材的内容将根据学科前沿动态与社会发展需求而有所调整，旨在契合课程目标，力求成为国内高校该课程的标杆教材。

创新点二：模块化课程建设方案

模块化课程建设的主要目标是解决当前人工智能学院专业课程间联系薄弱、各自独立、零散无序且缺乏统一规划的问题。通过实施模块化课程建设，我们能够避免课程内容的重复，增强具有先后修习顺序的课程之间的联系和互适配性，使学生能够清晰地理解所学知识点之间的内在关联，从而为学生的选课过程提供清晰和可靠的指导方案。人工智能专业课程的模块化建设在国内尚未有成熟的实践成果可供参考。作为国内首批建立人工智能学院的高校之一，以及一所 C9 联盟成员高校，南京大学无疑肩负着在人工智能课程体系建设方面进行探索、领导和示范的责任和义务。此项目旨在为未来人工智能学院的课程体系建设提供更加科学和实用的实践论证及指导方案，从而推动该领域教育的持续改进和发展。

推广应用效果

南京大学推出的《知识表示与处理》课程，荣登国内首个综合涵盖该领域理论知识



和技术、并配套完整系统教学大纲的人工智能专业课程。尽管有部分高校开设了与知识表示、知识图谱等相关的课程，它们仍主要以数据驱动为核心，以向量为底层计算模型基础，呈“知识为表象，数据为基础”的特征。此创新课程深入探索了知识表示与处理的基础理论和核心知识，同时也包括了以逻辑语义（包括集合论和模型论）为基础的计算模型等数理基础知识。在过去三年中，作为接待团队成员，赵一铮接待了来自复旦大学、浙江大学、厦门大学、哈尔滨工业大学、北京师范大学、中央音乐学院、中国传媒大学等多所国内一流学府的考察团队，并多次接到各参访院校领导及负责人关于该课程建设的详细询问。

值得一提的是，《知识表示与处理》课程成功入选 2022 年度教育部产学合作协同育人项目名单。通过与腾讯公司的紧密合作，赵一铮充分利用腾讯公司的专业技术、丰富资源和宽广平台，为国内缺乏相关师资的高校成功打造了一门人工智能专业核心课程，为其他国内高校提供了宝贵的经验和参考，成为了建设此类课程和培训相关师资的极佳途径。

此外，赵一铮还通过教材建设进一步提升知识表示与处理领域在国内的影响力，从而吸引更多国内院校的科研团队和学者投身于该领域的研究与探索，激励更多学生投入到该领域的学习和科研工作。值得强调的是，数据驱动和知识驱动的人工智能范式并非互斥，它们的底层哲学和逻辑理论有许多共通之处，可以相互协作和补充。有效结合这两种范式可以解决许多单一范式无法克服的计算问题，比如结合数据驱动模型的学习能力（归纳能力）和知识驱动模型的推理能力（演绎能力）实现它们之间的有益互补和交互作用。

2023 年 7 月赵一铮应邀参加全国高校人工智能专业建设暨系列课程导教班，作为主讲人分享了该课程的建设经验，引起了与会者的热烈反响和讨论。

学会动态

JSCS 老科委到南京邮电大学开展科技服务

2023 年 10 月 30 日，江苏省计算机学会老科学家委员会一行十多人，到南京邮电大学仙林校区考察调研并开展科技服务活动。参加本次活动的老科学家有：王汝传、张云鹤、张福炎、潘金贵、傅德胜、秦军、王和珍、陈汉武、孙志挥、刘凤玉、许满武、陈世福等来自南京大学、东南大学、南京邮电大学、南京理工大学、南京信息工程大学等高校的计算机各学科领域的老专家。江苏省计算机学会秘书长金莹教授，省学会石克和严诚秘书也参加了本次活动并提供全程服务。

这次活动是为了贯彻落实习近平总书记关于科技创新和科技人才的重要论述，大力弘扬科学家精神，发挥老科技工作者余热的实际行动。同时也是进一步学习习近平新时代中国特色社会主义思想和党的二十大精神，深入贯彻落实党和国家关于经济高质量发展，建设科技强国、教育强国、数字强国等一系列重要方针的具体措施。

推动科技向善 把好伦理“方向盘”

新一轮科技革命和产业变革加速演进，新兴技术突破和应用给经济社会发展带来的影响日益深刻，伴随产生的伦理问题成为全世界面临的共同挑战。促进科技向善，迫切需要加强科技伦理治理，完善科技伦理监管规则，强化科技伦理审查作用，切实有效防控科技伦理风险，实现科技创新高质量发展和高水平安全的良性互动。

近日，科技部会同教育部、工业和信息化部、国家卫生健康委等十部门联合印发《科技伦理审查办法（试行）》（以下简称《审查办法》）。科技部有关负责人就《审查办法》的出台背景、基本考虑和重要举措，以及如何保障各项措施扎实落地等回答了记者提问。

记者：《审查办法》出台的主要背景是什么，有何重要意义？

有关负责人：习近平总书记2021年5月在两院院士大会、中国科协第十次全国代表大会上发表重要讲话时指出，科技是发展的利器，也可能成为风险的源头。要前瞻研判科技发展带来的规则冲突、社会风险、伦理挑战，完善相关法律法规、伦理审查规则及监管框架。2022年3月，中办、国办印发《关于加强科技伦理治理的意见》，提出了完善科技伦理审查规则流程，健全科技伦理（审查）委员会设立标准、登记制度，建立科技伦理审查结果专家复核机制等具体要求。

科技部坚决贯彻落实党中央、国务院的决策部署要求，会同有关部门深入调研、总结经验、广泛征求意见，研究制定了《审查办法》。《审查办法》是覆盖各领域科技伦理审查的综合性、通用性规定，对科技伦理审查的基本程序、标准、条件等提出统一要求，为各地方和相关行业主管部门、创新主体等组织开展科技伦理审查提供了制度依据。

记者：《审查办法》有哪些基本考虑和具体举措？

有关负责人：起草过程中，我们注重把握3个方面。一是要促进科技向善，防范科技活动可能带来的潜在风险，保证科技创新活动的正确方向；二是要遵循科技规律，立足我国科技发展阶段和社会文化特点，推动科技创新与科技伦理的协同发展、良性互动；三是要坚持问题导向，重点解决科技伦理审查职责不明确、程序不规范、机制不健全等问题。

为此，《审查办法》从健全体系、规范程序、严格标准、加强监管等方面提出了一系列措施、作出了相关规定。

一是划定了科技伦理审查的主要范围，提出要坚持促进创新与防范风险相统一，客



观评估、审慎对待不确定性和技术应用风险。科技伦理审查要重点针对可能影响人的合法权益和动物福利以及对生命健康、生态环境、公共秩序、可持续发展等带来伦理风险的科技活动。

二是明确了科技伦理审查的责任主体、科技伦理（审查）委员会的设立标准和组织运行机制，并对委员会的制度建设、监督管理等提出具体要求。

三是明确了科技伦理审查的基本程序，确定了伦理审查内容和审查标准，明确了需要开展伦理审查复核的科技活动清单内容及调整更新机制。

四是明确了各相关部门、地方和各类创新主体的监督管理职责，建立了科技伦理（审查）委员会和科技伦理高风险科技活动登记制度，对科技伦理违规行为及调查处理分工等作出规定。

记者：制定“需要开展伦理审查复核的科技活动清单”的主要目的是什么？纳入清单管理的科技活动主要考虑哪些因素？

有关负责人：制定清单的主要目的是为了有效应对生命科学、人工智能等新技术加速突破和应用所带来的伦理风险与挑战。纳入清单管理的科技活动主要考虑3个方面的因素。一是科学技术自身的伦理风险，包括科学知识和安全信息的充分程度，技术的成熟度、操作难易程度、安全性、有效性和可控性；二是科技活动伦理风险发生的可能性、风险种类、严重程度、影响范围等；三是科技活动的必要合理性、目标人群或目标应用场景等。

记者：贯彻落实《审查办法》，下一步有哪些具体要求和部署？

有关负责人：一是相关行业主管部门和地方按照职责权限和隶属关系建立本系统、本地方科技伦理审查的监督管理机制，制定和修订本系统本地方的科技伦理审查办法、细则等制度规范，建立健全对纳入清单管理的科技活动的专家复核机制。科技部加强国家科技伦理管理信息登记平台建设，为相关行业主管部门、地方加强科技伦理监管提供信息化支撑。

二是各类创新主体要切实履行科技伦理管理主体责任，健全本单位科技伦理审查监管机制，加强科技伦理（审查）委员会制度建设和能力建设，加强对本单位科技伦理委员会委员和科技人员的教育培训，开展负责任的研究与创新。

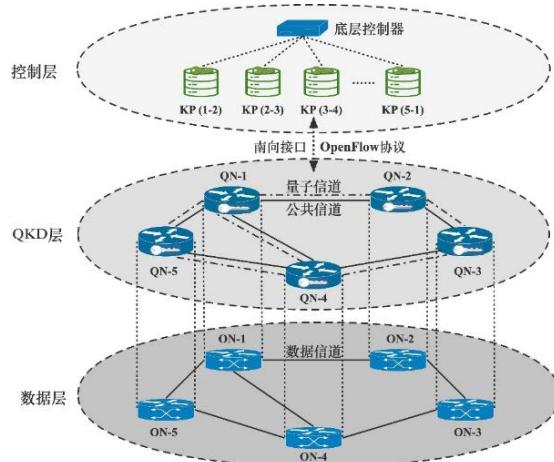
三是科技人员要自觉遵守科技伦理规范，学习科技伦理知识，提高科技伦理意识，按要求申请伦理审查，关注科技活动中伦理风险变化，遇到问题及时报告。

四是科技类社团可制定本领域的科技伦理审查具体规范和指南，为创新主体和科技人员提供细化指导。

基于应用需求的量子密钥分发网络路由技术研究

作者：赵梦楠

单位：东南大学网络空间安全学院



当今时代信息技术发展迅猛，数据充斥在人们的日常生活中，保障用户在网络中的隐私安全十分重要。依托于传统密码学进行密钥交换的通信网络已经无法抵抗量子计算带来的严峻挑战，量子密钥分发技术凭借量子自身特性能够实现理论上无条件安全的保密通信，将成为未来信息安全领域的关键技术。QKD(Quantum Key Distribution, QKD)网络作为量子保密通信的底层基础，与传统通信网络相比有许多不同之处，因此在搭建过程中无法照搬传统网络的运行机制。目前随着QKD网络规模的不断扩大，如何选择密钥中继的最优路径成为影响网络效率的关键。针对上述问题，本文将应用需求作为重要的考虑因素，充分衡量各应用对密钥的需求情况，提出了两种QKD网络路由方案：应用需求自适应的量子密钥分发网络 (Application Deamnd Addaptive Quantum Key Distribution Network, ADA-QKDN) 路由方案和应用优先级排序的量子密钥分发网络 (Application Priority Sorted Quantum Key Distribution Network, APS-QKDN) 路由方案。

专家推荐语

随着计算机技术的发展，经典的密码学算法在未来将无法抵抗量子计算机的攻击。量子密钥分发技术依赖于量子力学基本原理，为经典密钥分发算法面临的上述安全威胁



提供了解决思路。目前，世界各国都在探索 QKD 网络这一领域，研究的 QKD 网络规模不断扩大，从局域、城域范围逐步扩展到广域，甚至是天地一体化网络。在 QKD 网络的实际应用中，如何实现高效、高质量服务是扩大应用规模面临的关键性问题。该文针对 QKD 网络中的路由策略展开研究，提出了应用需求自适应的路由方案以及在特定场景下的应用优先级排序路由方案，两种方案均实现了一定的性能提升，对提升 QKD 网络的服务质量具有重要意义。

论文看点

- 1、针对现有 QKD 网络路由方案中存在的量子节点数据处理压力大、链路状态分析不充分以及链路成本计算不合理等问题，提出了 **ADA-QKDN** 方案。该方案引入软件定义的思想，将 QKD 网络拓扑信息的收集以及构建统一交予路由管理中心，并将应用自身的密钥需求作为制定路由策略的重要衡量指标。
- 2、针对应用密钥请求短时间内大量抵达以及应用之间密钥需求差异大的高并发特殊场景，提出了 **APS-QKDN** 方案。该方案对大量并发请求进行优先级排序，并在制定排队策略时综合考虑当前时刻的网络资源状况。
- 3、以北京投入使用的量子城域网架构作为底层网络拓扑，设计并实现了 QKD 网络路由方案验证系统。实验结果表明，**ADA-QKDN** 方案在请求参数动态变化的过程中性能表现稳定，应用密钥请求成功率至少提升 10%；**APS-QKDN** 方案在高并发的特殊场景下的应用密钥请求成功率至少提升 15%。

研究背景

随着网络通信技术的日益发达，数据安全受到的挑战也日益严重。密码学作为信息安全技术的核心，其重要性也愈来愈凸显，其中安全高效地获取密钥是密码学的研究重点。传统保密通信交换密钥的安全性通常依赖于大数分解等数学难题，随着量子计算逐渐成熟，在不久的将来这些数学难题会被轻易攻破，彼时传统的密钥交换算法完全没有安全性可言。量子密钥分发 (**Quantum Key Distribution, QKD**) 技术，作为一项具有颠覆性的新兴技术，能够解决上述提及的密钥安全分发的难题。它利用量子态来编码需要传递的密钥信息，双方在一定的机制下进行密钥协商，最终能够获得理论上无条件安全的通信密钥。

随着 QKD 组网技术的不断成熟，量子密钥已经逐渐走向实用化。然而，目前的 QKD 网络还处于实验环节，要真正融入人们的日常生活还需要考虑到网络节点以及用户数量剧增的问题，这会导致网络拓扑结构的复杂化，这时高效的路由方案显得尤为重要。同时，QKD 网络在承载用户密钥时消耗的链路资源与传统网络的带宽资源有较大区别，因此不能照搬传统网络的路由方案。

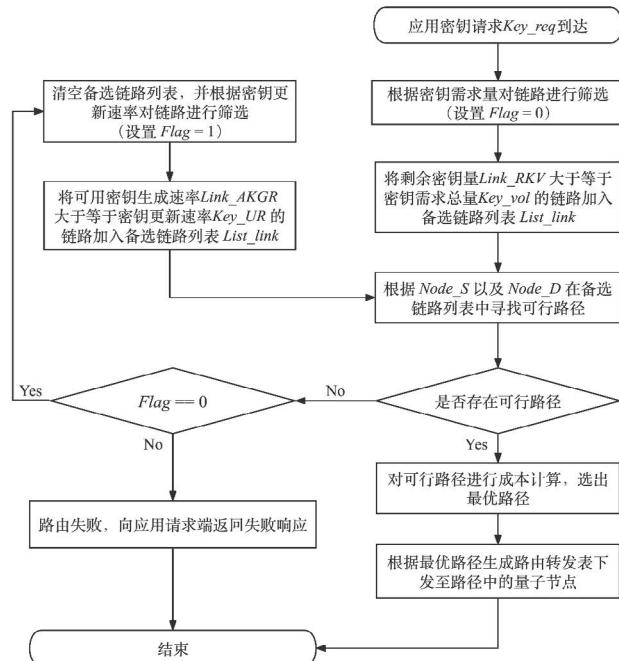
目前已有一些针对 QKD 网络中路由问题的研究，但现有研究大多未考虑量子密钥的

实用化场景，而这一因素是量子密钥投入实际使用的前提。目前，QKD 网络路由问题的研究成果中，大多通过借鉴经典网络中的成熟路由算法，并结合 QKD 网络自身特点进行路由方案的改进。

基于上述问题，本文展开 QKD 网络路由技术研究，致力于提升 QKD 网络的服务性能，主要体现为提升用户的密钥请求成功率以及减少链路资源的浪费。具体研究内容为：针对现有的 QKD 网络路由方案中存在的链路数据包泛洪、忽略应用自身的密钥需求、链路成本计算不合理等问题，提出了应用需求自适应的 QKD 网络路由方案 (APS-QKDN)；在该机制基础上，针对高并发以及应用密钥请求差异大的特定场景，设计了应用优先级排序的 QKD 网络路由方案 (ADA-QKDN)。

应用需求自适应的 QKD 网络路由方案

APS-QKDN 路由方案的中心思想是根据应用自身的优先级特性以及密钥需求来调整应用密钥请求的处理顺序，具体的算法流程如图所示。下面从三个角度对应用自身特性以及需求进行分析，明确各个因素对应用处理优先级产生的影响。



图：APS-QKDN 路由算法流程图

1) 应用自身优先级：根据重要程度的不同各个应用会具备不同的自身优先级。同一个应用在不同时刻的数据传输紧急程度可能不同，因此自身优先级也是可以动态变化的。优先级较高的应用密钥请求意味着该应用的安全要求较高或者是当前该应用急需获得密钥来进行保密通信。因此路由管理中心在对请求进行优先级排序时需要优先考虑该应用的密钥请求，从而提升重点应用的服务质量。



2) 密钥需求总量: APS-QKDN 路由方案是针对应用密钥需求相差较大的特定场景,因此密钥需求总量是影响请求优先级的关键。当应用的密钥需求总量较高时,意味着为当前应用分配资源的难度较大,但是也需要综合考虑当前的网络资源状况。如果当前网络中各个量子节点处的剩余密钥量都远高于各应用的密钥需求总量,则应用之间的需求差异对请求优先级的影响可以忽略;但是如果当前 QKD 网络资源紧缺,则将需求量较高的请求在合理的延时范围内进行推迟可以利用有限的资源满足更多的请求,从而提升密钥请求成功率。

3) 密钥更新速率: 密钥更新速率与上述提到的密钥需求总量对优先级排序造成的影响类似。密钥需求总量主要影响网络中各个量子节点的剩余密钥量的消耗,而密钥更新速率更多的是影响量子链路时间片的消耗。当应用要求的密钥更新速率较高时,意味着它需要占据更多的时间片资源。然而,链路的密钥生成速率是有限的,时间片资源一旦被某个应用占据,很难在短时间内被释放。因此需要慎重决定密钥请求的处理顺序,从而在短时间内满足更多的请求。针对密钥更新速率较高的请求会采取适当的延后操作,使得更容易被满足的轻量级请求得到优先处理。

应用优先级排序的 QKD 网络路由方案

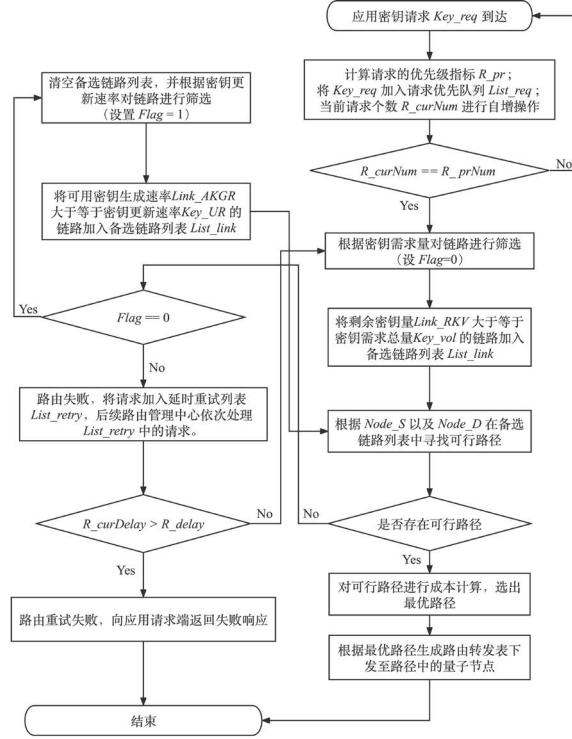
ADA-QKDN 路由方案从以下几个方面进行了改进:

1) 考虑到实际使用过程中,不同应用的重要程度以及通信紧急程度不同,在方案中将为应用的密钥请求设定不同的优先级标识。应用可以在运行过程中根据当前的通信状况动态改变密钥请求的优先级。当短时间内达到大量请求时,优先级更高的请求更有可能被优先处理,这也比较符合应用运行的实际情况。

2) 针对高并发的特定场景,对短时间内达到的一定数量的应用密钥请求进行优先级排序来决定这些密钥请求的处理顺序,而不是简单的依据密钥请求的达到顺序。其中,请求优先级的判定标准需要综合考虑三个因素,分别是应用自身优先级、应用的密钥需求总量以及密钥更新速率。同时,也要结合网络的实时资源状况来判断应用密钥需求总量的相对大小,实现请求优先级的评判与当前时刻的网络状态自适应。密钥需求量相对较小的请求更有可能被提前处理,从而为后续请求预留更多的资源空间;重要程度不高且密钥需求量较大的请求可能会在当前应用可接受的延时范围内暂时积压,从而提升整体资源利用率以及请求成功率。

3) 根据应用自身可接受的延时范围,给予路由失败的应用密钥请求延时重试的机会,而不是简单的直接拒绝或极端的无限等待。此改进可以避免在路由失败后应用在短时间内频繁发起密钥请求的情况,并将请求失败的后续操作统一交予路由管理中心处理,从而减轻应用的数据处理压力。同时,为了避免应用长时间等待密钥响应,在延时即将到期时向应用反馈,使得应用及时了解密钥请求的处理进度,并针对重试失败的请求采取相应的备选策略。

ADA-QKDN 路由方案的具体算法流程如图所示。



图：ADA-QKDN 路由算法流程图

总结与展望

本文的主要工作总结如下：

- (1) 对于 ADA-QKDN 方案，详细阐述了 ADA-QKDN 方案采用的底层网络架构、系统模型，应用特性以及网络拓扑构建方式；详细分析了 ADA-QKDN 方案的链路成本以及路径成本的计算方式，即针对不同的应用密钥需求制定不同的路由策略，路径成本的计算方式也会根据路由策略进行动态调整；接着详细描述了路由策略的运行流程。
- (2) 对于 APS-QKDN 方案，详细阐述了 APS-QKDN 方案的系统模型以及特定场景下的应用特性；详细介绍了 APS-QKDN 路由方案采取的应用请求优先级评判标准以及延时重试机制。其中，应用密钥请求的优先级会根据应用自身特性以及当前网络环境进行自适应调整，使得系统能够为短时间内到达的大量请求设置合理的处理顺序；延时重试机制从实际使用出发，改进了现有路由方案不合理的请求拒绝机制。
- (3) 设计并实现了 QKD 网络路由方案验证系统，对本文提出的 ADA-QKDN、APS-QKDN 路由方案以及其他方案进行了方案验证与性能分析。

综上所示，本文提出的 ADA-QKDN 方案和 APS-QKDN 方案，可以有效地提升 QKD 网络的服务质量。在网络资源有限的情况下，满足尽可能多的应用密钥请求。同时，考虑到高并发的特殊场景也制定了相应的路由策略，更贴近目前大数据时代应用数量剧



增的现实场景，具有很好的实践意义。

此外，本文提出的 QKD 网络路由方案为应用密钥请求选出最优路径后，在后续的应用运行期间该路径相对固定，可以进一步进行优化。未来研究可以对运行时间较长的保密会话增加传输路径的随机性。例如不定时开展路径选择，为应用选出符合当前网络状况以及应用自身需求的最优路径，从而进一步提升密钥协商的安全性以及网络资源的利用率。

作者简介



第一作者：赵梦楠，东南大学硕士研究生，主要研究方向为物联网安全、量子密钥分发技术等。

E-mail: 220194603@seu.edu.com

通信作者：



陈立全，东南大学教授，主要研究方向为密码与安全协议、物联网安全、区块链技术等。

E-mail: Lqchen@seu.edu.cn

贝叶斯网络结构学习优化算法研究

作者：颜克斐

单位：江南大学人工智能与计算机学院

论文摘要

贝叶斯网络（Bayesian Network, BN）作为一种概率图模型，得益于图论和概率论的结合，可以清晰地表现变量间的因果关系，在对不确定关系进行表示和推理中发挥了重要的作用，因此在人工智能、医疗、故障诊断、数据挖掘等领域有着广泛的应用。对BN的建模分为两个部分，结构学习和参数学习，其中结构学习是参数学习乃至对BN进行应用的基础和前提，从数据中对结构进行学习具有重要意义，也是本文的主要研究方向。由于搜索空间随着问题规模的增加而呈现超指数级增长，贝叶斯网络结构学习（BN Structure Learning, BNSL）的任务已被证明是NP-难的，引入智能优化算法是一种解决此类问题的思路，其中遗传算法（Genetic Algorithm, GA）近年来取得了诸多成果，但仍然面临着搜索效率低、准确度不高、与问题的结合不够紧密等问题。本文通过对基于GA的BNSL算法进行研究，通过利用搜索空间中的结构信息对GA的搜索行为进行改进；利用互信息和支持度挖掘种群中的优势信息引导GA进行快速收敛；此外还研究了BNSL算法在食品安全事故成因分析上的应用方法。本文的主要工作如下：

(1) 针对GA冗余搜索过多导致的搜索效率低下的问题，提出了一种基于结构信息改进的GA用于BNSL。首先利用精英个体的马尔科夫链来指导交叉算子，通过局部的社会学习来提高算法的收敛速度；在此基础上利用评分等价原理进行有偏好的变异操作，随着算法迭代的停滞次数自适应调整偏好概率，加强了局部搜索能力进而提升了算法的探索性。通过两个算子的结合，算法在探索性和收敛性得到了更好的平衡且均在一定程度的提高。

(2) 受互信息的原理和在此领域中运用的启发，针对演化算法特有的种群优势，提出了基于互信息引导的GA。首先结合条件独立测试和互信息分别用于限制搜索空间和种群初始化空间，利用随机的个体参数保证种群多样性，提高了初始种群的评分；随后利用互信息和支持度挖掘种群中潜在的优势基因，并用于指导交叉算子中差异边的处理；此外在对无效解的处理上，利用互信息作为环路去除的指导，避免了优势基因在个体修复过程中的误删除。在11个数据集上的实验结果验证了互信息在该领域强大的指导作用，



通过结合以上三个改进点的优势，算法最终在收敛速度和所得结构的准确度上均有较大幅度的提升。

(3) 基于上一项工作提出的混合 BNSL 算法，设计了一套基于食品样品检测数据的食品安全事故原因分析方法，包括：将检测数据转化为训练数据的预处理及特征工程方法；基于 BN 模型的事故原因分析及风险预测方法；最后提出了一个有关食用油安全风险的 BN 数据集。通过以上流程将 BNSL 算法应用于实际，实现了由结构化检测数据到利用 BN 进行风险预测的全流程分析方法。

综上所述，本文针对基于 GA 的混合结构学习算法进行了两部分工作，分别从结构信息和互信息角度入手优化了算法的性能和效率；为食品安全事故分析和风险溯源提供了一套基于 BN 的解决方案；最后提出了一个 BNSL 数据集，可用于后续的算法研究。

专家推荐语

论文以贝叶斯网络结构学习为研究对象，选题具有较好的理论意义和应用价值。论文首先对贝叶斯网络的结构学习研究和发展现状进行了介绍，分析了基于演化计算的 BN 结构优化的研究进展，随后针对该研究领域所存在的问题提出了基于结构信息改进的混合 BNSL 算法和基于互信息引导的混合 BNSL 算法，并设计了基于食品样品检测数据的食品安全事故原因分析方法。基于结构信息改进的混合 BNSL 算法在遗传算法的研究基础上，通过局部的社会学习来提高算法的收敛速度，并利用评分等价原理来加强算法的局部搜索能力；基于互信息引导的混合 BNSL 算法结合条件独立测试和互信息来提升初始种群的质量，挖掘种群中潜在的优势基因，并利于互信息指导环路去除；对提出的算法在实际问题场景中进行了应用，证实了 BNSL 算法在面对结构化数据的有效性。论文概念叙述基本正确，结构完整，格式规范。以第一作者身份在 CCF A 类期刊发表学术论文 1 篇，以共同作者身份发表中科院 1 区论文 1 篇，EI 中文期刊 1 篇，有非常好的研究成果，反映作者有扎实的基础理论知识和科学研究能力，推荐申报优秀硕士生学位论文。

论文看点

本文针对基于遗传算法在解决贝叶斯网络结构学习 (Bayesian Network Structure Learning, BNSL) 问题中存在的收敛慢、准确度低等问题展开研究，分别从结构信息和互信息的角度出发对遗传算法的流程和算子进行了改进优化；在此研究基础上，从实践角度出发设计了一套基于食品样品检测数据的食品安全事故原因分析方法。

基于结构信息改进的遗传算法 (SIGA)

所提出的基于结构信息改进的遗传算法 (Structural Information-based Genetic Algorithm, SIGA) 流程如图 1 所示。其中基于马尔科夫毯的精英交叉算子以及基于 V-结构的自适应偏好变异算子为 SIGA 的改进点。算法以使用 CHC 选择、均匀交叉以及单点变异的标准 GA 作为基于评分搜索阶段的基准；对于基于约束的阶段，采用独立性检验方法缩减搜索空间。

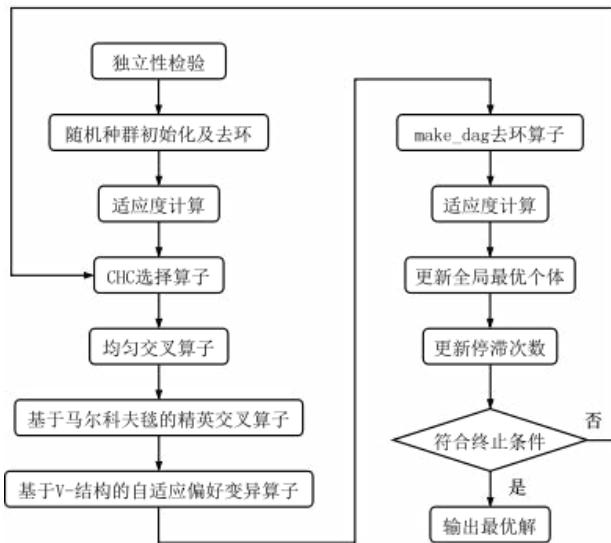


图 1 SIGA 的算法流程图

基于马尔科夫毯的精英交叉算子：在算法的交叉阶段利用精英个体中的马尔科夫毯 (Markov Blanket, MB) 进行局部的社会学习，使算法能够从高评分个体中得到借鉴因此算法收敛更快。具体来说，前一代的种群经过选择后得到当前种群，在种群中先执行均匀交叉，为算法提供收敛性；随后依据个体的适应度评分建立精英集，为个体演化提供更高效的引导。为了避免陷入局部最优，我们从种群中选取一半的个体进行基于 MB 的精英交叉。图 2 展示了一个拥有 7 个节点的 BN 局部结构中应用 MB 精英交叉算子的例子，原始个体结构向精英个体学习节点 B 的 MB 局部结构信息，经过交叉的后代个体如图 2(c) 所示。

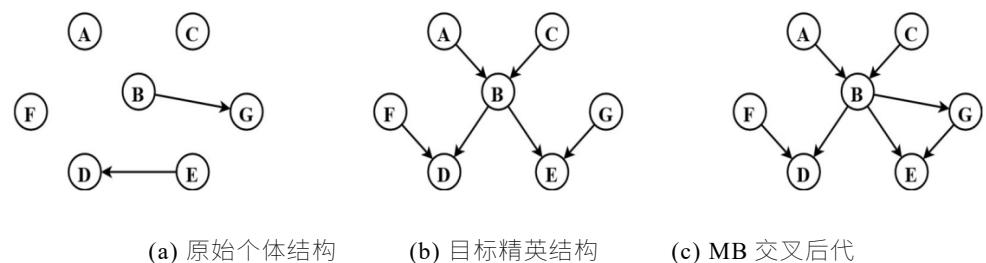


图 2 基于马尔科夫毯的精英交叉算子

基于 V- 结构的自适应偏好变异算子：SIGA 提出了一种新的变异算子，在变异阶段利用对局部结构的改动进行偏好性选择，加大了评分的局部扰动，进而提升了算法的探索性。图 3 展示了一个例子说明变异算子可能进行的操作，对于 DAG 搜索空间中的变异操作，会在 3 种情况间进行改动，即图 3(a)-(c) 中 A、B 两节点间的 3 种连接方式，分别为 $A \rightarrow B$ 、 $A \leftarrow B$ 和 $A \leftrightarrow B$ 。对于每种可能的变异结果，我们统计对应的结构改动引起父节点集发生变化的节点数，并选取其中变化更大的变异操作。

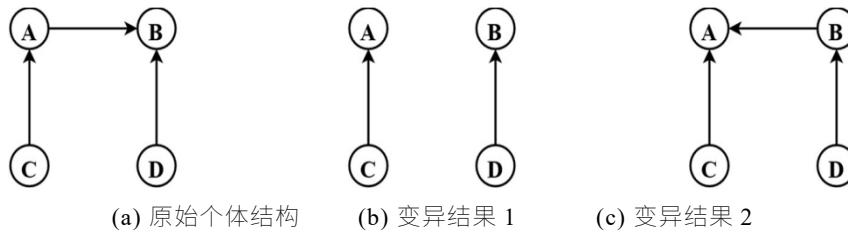


图 3 V- 结构变异类型

此外,为了避免算法陷入带有偏好的局部最优,我们对于变异操作进行偏好选择的概率,设置了一个与迭代次数相关的偏好概率 P_{prefer} :

$$P_{prefer} = (e^{-Nostag/4} + 1) / 2 \quad (1)$$

P_{prefer} 从 1 开始缓慢下降,其目的是更好地平衡算法的探索性和收敛性。

SIGA 通过上述两种改进方法的配合,使得遗传算法在 BNSL 问题中的探索性和收敛性均得到了提高且取得了更好的平衡。

基于互信息引导的混合 BNSL 算法 (MIGA)

受到互信息度量相关性的启发,本文提出了基于互信息指导的遗传算法 (Mutual Information guided GA, MIGA) 用于贝叶斯网络结构学习,属于包括基于约束阶段和基于评分搜索阶段的混合结构学习算法。基于互信息引导的混合 BNSL 算法流程如图 4 所示。

其中包含了基于互信息改进的三个算子,首先在基于约束的阶段和种群初始化时,分别利用条件独立测试和互信息对搜索空间和初始化空间进行限制,在保证搜索空间较为准确的情况下提升了初始种群的评分;随后在迭代搜索过程中,使用互信息和支持度对种群的潜在优势进行挖掘,并用于设计新的交叉算子,提高了后代个体中优势基因的保留概率;在此基础上将互信息用于改进环路去除算子,在修复非法解时避免了优势结构的误删除。

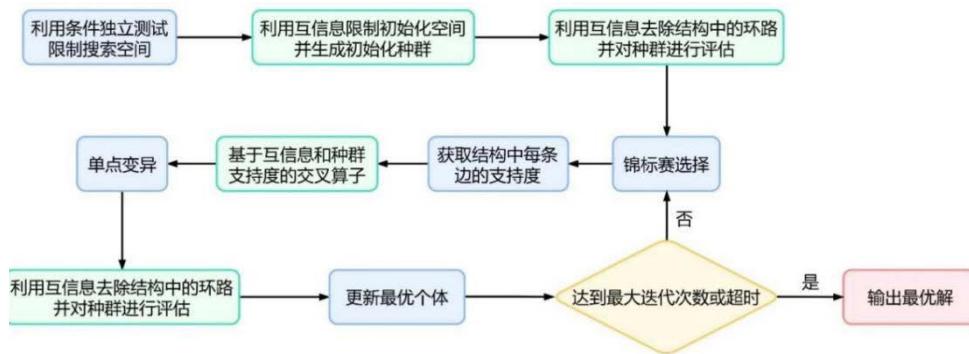


图 4 MIGA 算法整体流程图

基于互信息的种群初始化方法:当利用互信息和固定的参数 α 对搜索空间进行限制时, α 取值过大或过小都会造成问题,进而影响算法最终的准确度或搜索效率。而基于

条件独立测试的方法能获得更为准确的搜索空间。为了使得算法迭代的起点更高，即获得更高评分的初始种群，MIGA 结合了独立性检验和互信息，在基于约束的阶段利用条件独立测试对搜索空间进行限制的同时，在种群初始化阶段利用互信息对于种群的初始化空间进行了额外的限制。具体的操作如图 5 所示：

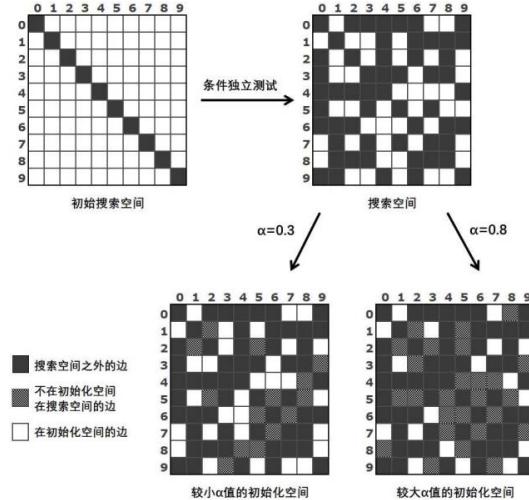


图 5 种群初始化方法

该方法利用条件独立性测试排除了独立节点间的连接确定了搜索空间，又参考互信息为每个个体设置了不同的初始化空间。经过自上而下的排除和自下而上的选择，两次筛选使得算法在保证搜索空间有效性的同时最大程度上限制了种群的初始化空间。同时不同个体拥有不同的参数 α ，因此具有不同的初始化空间，进而保证了种群的多样性。

互信息和支持度指导的交叉方法：本文将每条有向边的在种群中出现的频率定义为边在种群中的支持度，支持度越高意味着这条边在前一代选择中得到了更多的保留，也即对应的基因在种群中的优势更大。为了降低随机遗传的盲目性，我们参考支持度，结合互信息在描述节点间相关性上的指导作用，提出了一种更好的继承策略使得更多的优势基因被遗传到后代种群中。

该算法的伪代码如下：

算法 互信息和支持度指导的交叉算子

输入：种群规模 N ；种群 Pop ；搜索空间 SS ；边的种群支持度 spt ；归一化的互信息矩阵 MI_{norm}

输出：经过交叉操作的种群 Pop

```

1:   for 每个种群  $Pop$  中的个体  $p1$  do
2:       从  $Pop$  中随机选取另一个个体  $p2$ ；
3:       初始化一个新的个体  $p$  作为  $p1$  和  $p2$  的后代;
4:       for  $SS$  中的每条边 do
5:           if 父代个体  $p1$  和  $p2$  中都存在这条边 then
6:               将这条边保留至后代  $p$ ;
7:           else if  $p1$  和  $p2$  中存在这条边 then
8:               获取这条边的保留概率  $P_{return}$ 
9:               按照概率  $P_{return}$  将这条边保留至后代  $p$  ;
10:          end if
11:      end for
12:      将后代个体  $p$  加入种群  $Pop$  中
13:  end for

```



在所提出的算子中，我们将种群支持度 spt 和归一化的互信息矩阵 MI_{norm} 作为输入。为种群中的每个个体随机分配另一个个体组成父代，对于两个父代个体中的相同基因直接遗传给后代个体。而对于父代个体基因表现不同的位置，即其中一个个体结构中存在某条边而另一个个体中不存在这条边的情况下，我们按照设置的保留概率 P_{retain} 随机决定这条边是否继承给后代个体，保留概率表达式如下：

$$P_{retain} = 0.25 + \frac{spt + MI_{norm}}{4} \quad (2)$$

经过改进的交叉算子可以在迭代过程中为后代种群保留更多的优势基因，从搜索效率来说，也可以为算法带来更好的收敛性。

基于互信息的结构环路去除方法：在算法的搜索过程中，每次对个体结构进行修改都有可能在结构中引入环路，这违反了结构作为解的定义。因此，为了在修改后及时对无效解进行修复，我们设计了一种基于互信息的去环算子。提出的算子包含两个主要步骤，图 6 展示了算法去除一个具有 5 个节点的有向图中环路的例子。

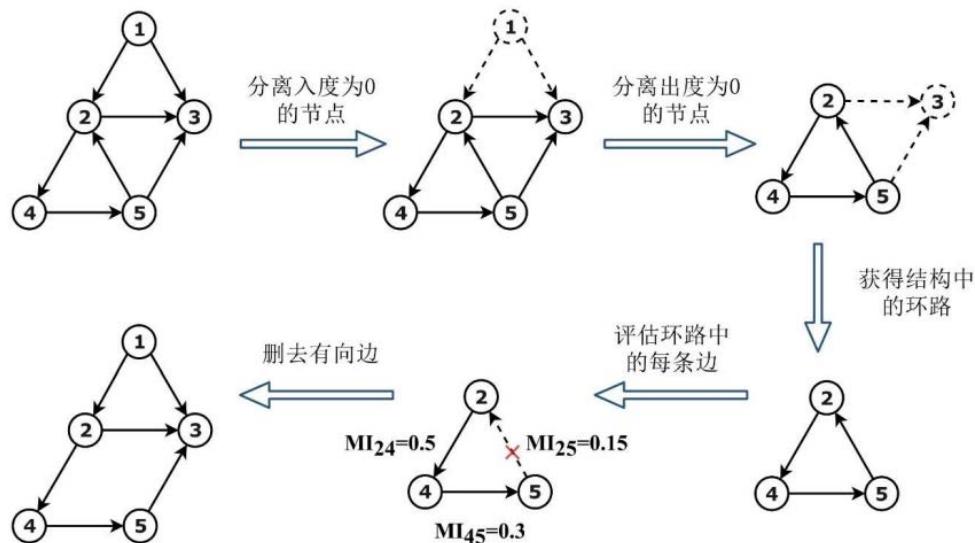


图 6 基于互信息的结构环路去除方法

其中，第一步是将环路从个体结构中剥离出来，方法是循环删除结构中入度或出度为 0 的节点和与之相连的边，直到结构中没有节点可以被删除；第二步是在获得的环路中删去一条边，删除时将互信息作为对各条边的评价指标，对有向边按照 MI 进行排序，选择其中互信息值最小的一条边进行删除。循环执行以上的两个步骤，每次循环删除结构中的一条边，直到网络中不存在有向环，个体成为一个有效解。

MIGA 通过结合以上三个策略，遗传算法在解决贝叶斯网络结构学习问题上的收敛速度和输出模型准确度有了显著提升。

基于贝叶斯网络的食品安全事故原因分析方法

本文中基于贝叶斯网络的食品安全事故原因分析方法的流程图如下：

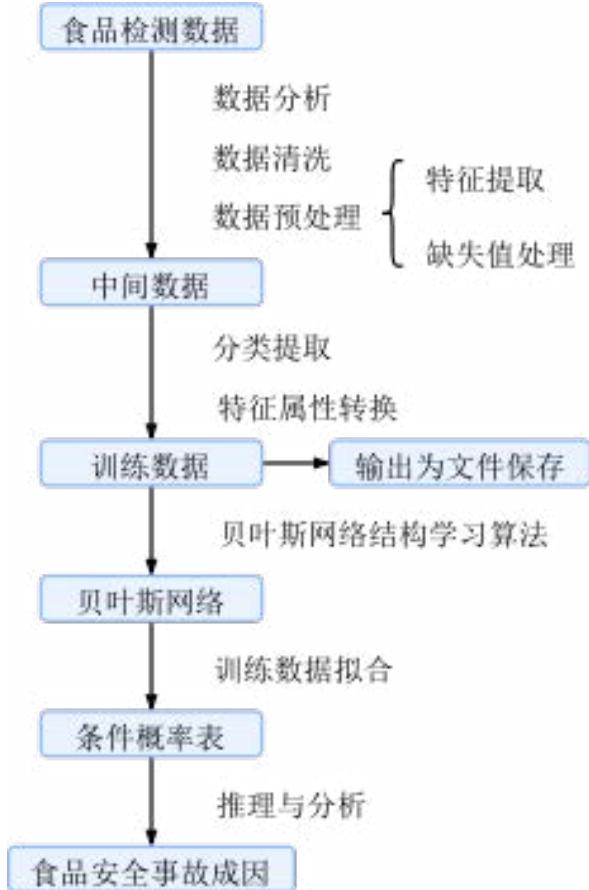


图 7 基于互信息的结构环路去除方法

流程整体可分为三个部分：前期的数据预处理、中间的贝叶斯网络建模方法、和后续基于 **BN** 的食品安全事故原因分析方法，其中输出的训练数据可作为 **BNSL** 算法研究中的样本数据集使用。

详细步骤如下：

食品检测数据经过特征工程等数据处理，处理步骤包括数据预处理、特征提取、特征属性转换和缺失值处理，得到可供贝叶斯网络结构学习算法使用的训练数据。随后使用前面提出的 **MIGA** 从训练数据中学习 **BN** 网络结构，经过数据拟合得到完整的 **BN** 模型。最后利用 **BN** 进行推理分析，包括正向的推理预测食品安全风险发生的概率，以及采用反向推理的方法分析事故发生原因。利用这样一套分析流程与方法，可以帮助相关部门对食品安全风险进行掌控，并且在食品安全事故发生后对事故原因进行溯源。

图 8 展示了对于得到的 **BN** 网络结构使用可视化展示后的结果：

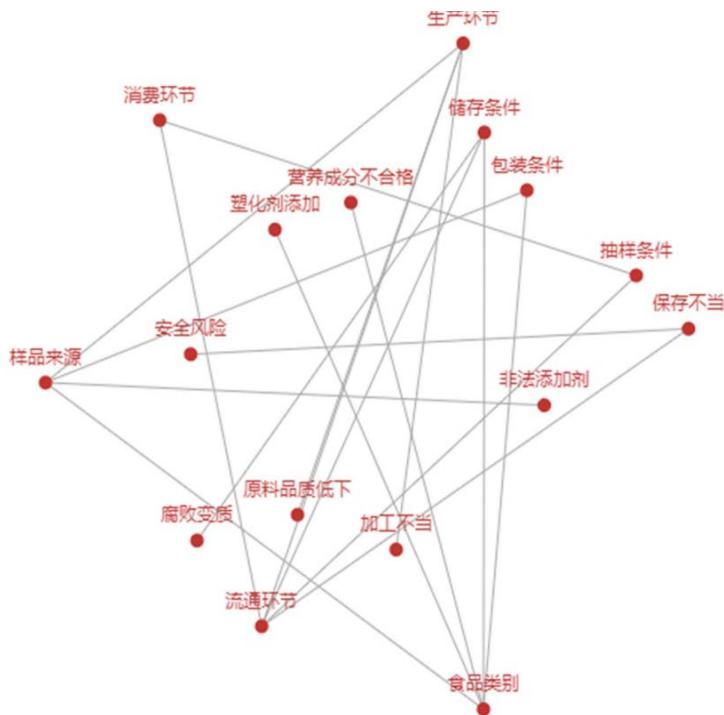


图 8 由食用油检测数据得到的贝叶斯网络结构模型

此外，还依据在课题中获得的食用油等检测数据提出了一个贝叶斯网络数据集 Oil Net，包括了完整的网络结构和对应的条件概率表，为今后的算法研究提供了参考和有效的工具。Oil Net 的 DAG 网络结构如下：

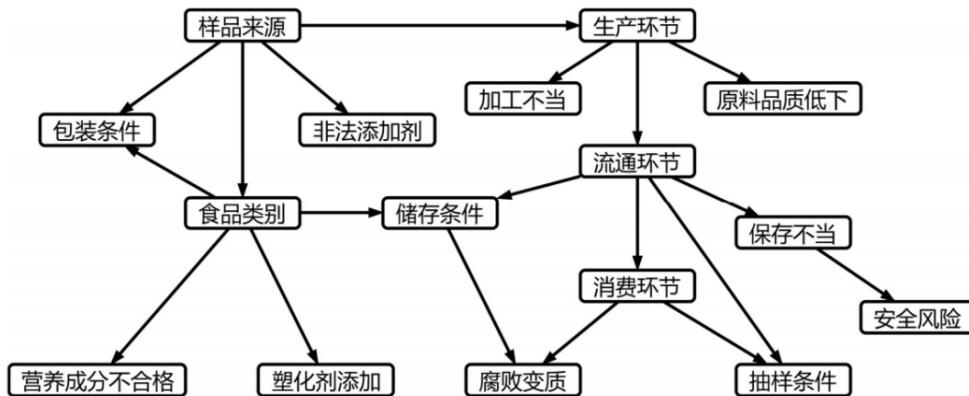


图 9 Oil Net 的 DAG 网络结构

综上所述，本文针对基于 GA 的混合结构学习算法进行了两部分工作，分别从结构信息和互信息角度入手优化了算法的性能和效率；为食品安全事故分析和风险溯源提供了一套基于 BN 的解决方案；最后提出了一个 BNSL 数据集，可用于后续的算法研究。

总结与展望

贝叶斯网络作为对不确定关系进行表示和推理的重要工具，可以用于解决很多现实世界中的问题，从数据中自动学习 **BN** 的 **DAG** 结构虽然是一个 **NP**- 难的问题，但是作为 **BN** 应用和研究的前提，一直是研究者重点关注的问题。贝叶斯网络结构学习的三大类算法中，混合方法能在一定程度上结合其余两种方法的优势，而演化算法作为一种通用的随机搜索算法，可以在搜索空间中进行相对全面的探索学习，但仍存在着搜索效率不高、收敛速度慢、学习结果不够准确等现实问题。

本文通过深入研究 **BNSL** 问题的实质，将结构学习问题与遗传算法结合从而对上述问题进行改进，此外还将 **BN** 用于解决实际应用场景下的具体问题。主要研究内容总结如下：

(1) 基于结构信息改进的遗传算法 (**SIGA**)：为了实现更快的算法收敛速度，使用精英个体的马尔科夫毯指导个体的局部交叉；为了减小由评分等价原理带来的冗余搜索，我们在变异算子中偏好选择那些局部改动更大的操作。

(2) 基于互信息引导的遗传算法 (**MIGA**)：为了获得更高评分的初始种群，我们分离了算法的搜索空间在和种群的初始化空间，并限制后者的大小；为了挖掘和利用种群经过选择后的潜在优势，我们使用互信息和支持度对边的重要程度进行评估并用于指导交叉算子；为了减小环路破除算子的破坏性，我们提取环路并删除其中不重要的边。

(3) 基于贝叶斯网络的食品安全问题分析方法：通过特征工程和数据预处理，将食品样品检测数据转化为算法的训练数据；利用提出的 **MIGA** 学习其中的贝叶斯网络，进而分析事故原因或潜在的风险；提出了一个数据集可供后续算法研究使用。

展望：本文在理论研究层面，在经过了多次实验尝试后，针对基于演化计算的贝叶斯网络结构学习算法进行了一定程度的优化和改进，其中主要涉及贝叶斯网络的局部结构特征和信息论的相关原理，在研究和实验的过程中，我们有了一些发现和对未来工作的想法如下：

(1) 从评分函数的理论限制和不同评分间的差异中，可以看出在评估网络结构与数据的拟合程度时，评分函数仍有改进的空间，例如 **AIC** 和 **BIC** 评分的差异仅在惩罚项的系数上，因此如何找到更合适的系数或许是一个值得研究的方向，或是通过多目标算法来避免对参数具体取值的依赖，使得评分函数能在更普遍的场景下对模型进行更合适的评估。

(2) 遗传算法的变异操作主要目的是探索更多未知结构，在这个过程中如何确定对结构的探索朝着增加种群多样性的方向进行，我们在第三章中进行了一些基于评分等价性的尝试，在此基础上也可以考虑用其他形式表示算法未探索的结构，并有针对性地指导变异操作，或许可以进一步提升算法的探索性。



作者简介



颜克斐，本科毕业于中南大学土木工程专业，2019-2022 年于江南大学师从方伟教授并获得计算机科学与技术硕士研究生学位，现于南京大学计算机科学与技术系攻读博士学位，主要研究方向：数据挖掘、因果推断、智能优化方法与应用、图结构学习。

E-mail：6191910036@stu.jiangnan.edu.cn

导师：



方伟，江南大学人工智能与计算机学院教授、博导，副院长（主持工作），入选江苏省高校青蓝工程中青年学术带头人，江苏省“六大人才高峰”高层次人才。获教育部科学技术奖一等奖 1 项、中国商业联合会科学技术奖一等奖 1 项，出版专著 1 部，主持国家重点研发计划子课题，国家自然科学基金面上项目、青年项目，JWKJW 开放基金，江苏省重点研发计划、面上项目，中国博士后基金（一等）项目，授权美国发明专利 2 项、中国发明专利 10 项。

主要研究方向是智能优化算法，在算法的设计、改进及应用等方面展开具体研究，已在 IEEE TKDE、TCYB、SMC、TEVC、TII、Information Sciences、Information Fusion、KBS、ESWA、ICDE、GECCO 等国内外重要刊物及国际会议上发表学术论文

90 余篇，其中 SCI 检索论文 40 篇。学术成果被国内外同行广泛引用并正面评价，论文目前总被引 3800 多次，H-index=32，单篇最高被引 380 多次。

E-mail: fangwei@jiangnan.edu.cn

学会动态 •

“江苏工匠”岗位练兵职业技能竞赛 - 物联网安装调试职业技能竞赛、工业互联网技术应用职业技能竞赛在南京、南通举办

为促进物联网、工业互联网技术人才的选拔和培养，推动相关技术的应用和产业发展，发掘物联网、工业互联网技术在赋能经济社会数字化转型的潜力价值，江苏省计算机学会协同江苏新大陆时代科技有限公司举办了 2023 年“江苏工匠”岗位练兵职业技能竞赛 - 江苏省物联网安装调试职业技能竞赛决赛（以下简称“物联网竞赛”）、江苏省工业互联网技术应用职业技能竞赛决赛（以下简称“工业互联网竞赛”），两项赛事均由江苏省人力资源和社会保障厅指导，江苏新大陆时代科技有限公司提供技术支持。10 月 20-21 日，物联网竞赛在南通举办，本赛项南通职业大学承办，共有来自江苏省 22 所院校的 60 支队伍同台竞技。10 月 28 日，工业互联网竞赛在南京举办。本赛项由江苏南京技师学院承办，共有来自江苏省 20 所参赛单位的 60 位参赛选手前来竞技。本次竞赛新大陆时代科技将企业先进技术融入到大赛中，促进人才培养和产业发展相结合，推动专业课程内容与职业标准有效衔接，助力江苏各院校物联网与工业互联网专业建设及人才培养。未来，新大陆时代科技将继续携手院校及行业企业，共建新一代信息技术产教融合人才培养生态，为培养更多高素质产业技术技能人才、能工巧匠、大国工匠提供支撑。





“硬件与人工智能安全”的坚守者

——2023年江苏省计算机学会优秀科技工作者薛明富副教授

个人简介

薛明富，南京航空航天大学计算机科学与技术学院/人工智能学院副教授。IEEE Senior Member、CCF 高级会员。获 2023 年度 CCF 科技进步奖二等奖。获 2023 年江苏省计算机学会优秀科技工作者。任中国人工智能学会人工智能与安全专业委员会委员；ACM 南京分会执行委员会委员；中国图象图形学学会数字媒体取证与安全专委会委员；中国人工智能学会青年工作委员会委员；CAAI 智能信息网络专委会委员；江苏省计算机学会信息安全专委会委员；江苏省人工智能学会智能与安全专委会委员；江苏省计算机学会网络与分布计算专业委员会委员；江苏省网络空间安全学会数据安全专委会委员。研究方向为人工智能安全、深度学习模型版权保护、硬件木马检测。已主持科研项目 17 项，含 4 个国家级项目、4 个省部级项目、入选 3 个 CCF 基金等。另参与 5 个国家级项目。参与两项人工智能安全国家行业标准制定。发表论文 60 余篇，获 3 项论文奖。专著 2 部。申请发明专利 13 项，已授权 7 项。任 25 个国际会议的程序委员会委员。任 70 个 SCI 国际期刊（包括 20 余个 Transactions/ 一区）、2 个 CCF 中文 A 类学报、35 个国际会议的审稿人。



砥砺前行，科研学术硕果累累

薛明富在人工智能安全和硬件安全领域作出重要学术贡献。“AI+ 安全”智能安全融合关键技术研究和规模化应用成果获 2023 年度 CCF 科技进步奖二等奖。已主持科研项目 17 项，含 4 个国家级项目（两项国家自然科学基金、两项国防项目）、4 个省部级项目、入选了 3 个 CCF 基金等。另参与承担 5 个国家级项目（3 个国防项目及 2 个国家自然科学基金）。与绿盟科技联合申报入选为工信部人工智能安全揭榜挂帅项目。发表论文 60 余篇，其中 30 余篇 SCI，含 14 篇 CCF A 类 / 中科院一区 /Transactions。获 3 项论文奖：IET 高影响力论文奖、IET CDT 年度编辑推荐论文奖、会议最佳论文奖。申请发明专利 13 项，已授权 7 项。



图 1. 获 2023 年度中国计算机学会 (CCF) 科技进步奖二等奖

研以致用，校企合作成效显著

薛明富的研究工作得到了各研究院所、各行业单位的认可，与 20 余家科研院所、安全企业有过良好的合作，较好地服务了国家与行业需求：（一）在硬件木马检测方面，与多个研究所开展了长期密切合作，主持国家自然科学基金、多项国防项目、江苏省自然科学基金、2016 CCF- 启明星辰鸿雁科研计划、2017 CCF- 绿盟科技鲲鹏科研基金等十余个项目，自主研发有多款硬件木马检测原型系统、硬件木马特征提取与分析平台、硬件木马样本库等。（二）在人工智能安全方面，研究涵盖了 AI 系统全周期的各类攻防，研发了多款原型系统。在绿盟科技股份有限公司、深信服科技股份有限公司等企业的人脸识别、自动驾驶、信创操作系统杀毒引擎技术等业务中得到了实际应用。主持及承担国家自然科学基金、国防项目、航空科学基金航空人工智能专项、2021 CCF- 绿盟科技鲲鹏科研基金等多个项目。



图 2. 入选 2016 年 CCF- 启明星辰鸿雁科研计划、2017 年 CCF- 绿盟科技鲲鹏科研基金、

2021 年 CCF- 绿盟科技鲲鹏科研基金

立德树人，教书育人成果斐然

薛明富坚持“立德树人”的人才培养理念，做学生学习和科研路上的引路人。从教至今，获校教学创新二等奖、校微课竞赛二等奖、“奇安信杯”校园网络文化创意设计大赛微课赛道二等奖。教学评估连续获优秀。主持 3 项校教育教学改革研究项目。发表教改论文两篇。指导本科生科创项目多项，3 组获评优秀，2 组获院科创二等奖。指导本科生申请发明专利多项、发表高质量论文多篇。所指导的硕士生多人次获校优秀硕士学位论文。所指导毕设多次获成绩优秀。担任信息安全专业 1618201 班主任，该班先后获 2018 年度优秀班级、2019 年度优秀班级、2019 年度标兵班级、2020-2021 学年优秀班级、2020-2021 学年度江苏省省级先进班集体、2022 年本科毕业生先进班级等称号。



学优秀奖之教学创新二等奖、校微课竞赛二等奖

学会动态

JSCS 走进泰州学院

2023 年 11 月 1 日，江苏省计算机学会组织专家团来到泰州学院，围绕学科竞赛与泰州学院教务处、信息工程学院、美术学院、创新创业学院相关领导和学科竞赛指导教师一起，共同举办了学科竞赛赋能创新型人才培养专题研讨会暨 2024 年度学科竞赛备赛启动会。



在“边缘智能计算系统”的科研道路上求实创新

——2022年江苏省计算机学会青年科技奖获得者周俊龙副教授

个人简介

周俊龙，南京理工大学计算机科学与工程学院副教授，博士生导师，院长助理，美国圣母大学联合培养博士，主要研究方向为计算机体系结构与高性能计算、边缘计算、嵌入式系统等，在其研究领域国内外权威期刊和会议（例如 IEEE TC、TCAD、TPDS、TDSC、TSC、TASE、DATE、ACM CSUR、软件学报、计算机研究与发展等）发表论文 100 篇，包括 IEEE /ACM Transactions 论文 40 篇，ESI 高被引论文 10 篇，国际会议最佳论文 2 篇，授权发明专利 7 项。主持国家和省部级科研项目 10 余项，例如国家自然科学基金面上项目 / 青年项目、江苏省自然科学基金优青项目 / 青年项目、科技委基础加强项目课题、中国博士后科学基金特别资助项目 / 面上项目一等资助等，参与国家自然科学基金重点项目、国家重点研发计划课题、国家外专局项目、美国国家科学基金会项目等；曾获上海市计算机学会优博（2019）；连续四年（2020-2023）入选斯坦福大学发布的“全球前 2% 顶尖科学家”榜单；江苏省一流本科课程（2021）、江苏省优秀青年基金资助（2022）；江苏省计算机学会青年科技奖（2022）、国家级一流本科课程（2023）。担任中国计算机学会（CCF）嵌入式系统专委会副秘书长 / 常委，江苏省计算机学会嵌入式与物联网专委会副主任 / 体系结构专委会秘书长，CCF 体系结构 / 分布式计算与系统 / 容错计算专委会执行委员。担任含 CCF B 类期刊 JSA 在内的 4 个国际 SCI 期刊领域编委；10 余次国际重要会议的程序委员会主席，常年担任 DAC、RTSS 等 CCF A 类会议的程序委员会委员。

砥砺奋进，学术科研取得突出成果

周俊龙在计算机体系结构与高性能计算、嵌入式系统领域做出重要学术贡献，入选江苏省优青、斯坦福大学发布的“全球前 2% 顶尖科学家”榜单、多次获得国际会议最佳论文奖、上海市计算机学会优秀博士学位论文等奖项。近年来紧密围绕“创新驱动发展”以及研发“自主可控、安全可靠的软硬件产品”的国家战略需求，面向安全攸关的嵌入式系统与软件领域，以异构多核处理器为研究重点，致力于研究高可靠的多核处理器软硬件协同设计方法和应用技术，取得了一系列创新型研究成果。在一流国际期刊会议发表论文 100 篇（例如 IEEE TC、TCAD、TPDS、TDSC、TSC、TASE、TCC、TCAS、



TSUSC、TR、TII、TCBB、TAES、TNSE、DATE、ACM CSUR、ACM TCPS、软件学报、计算机研究与发展等），成果收到 30 余位院士 /Fellow 正面引用评价。

知行合一，研究成果付诸实际应用

周俊龙积极促进研究成果向实际应用转化，响应国家在研发“自主可控、安全可靠软硬件产品”的战略需求，以嵌入式系统与软件为应用切入点，初步形成了一套业务化、工程化的多核处理器软硬件协同设计方法和应用系统开发技术。相关成果作为软硬件协同设计技术与应用教育部工程研究中心、国家可信嵌入式软件工程技术研究中心、先进通信与数据工程研发中心、上海国茂数字技术有限公司等单位研制的多款嵌入式仪器配套软件及工业制造、智慧交通等行业应用服务软件，进行了成功推广和批量销售，产生了显著的社会经济效益。

立德树人，教学工作获得多项荣誉

周俊龙坚持以培养德才兼备的优秀人才为教学工作目标，在向学生传授前沿理论和创新技术的同时，特别注意引导他们不断提升思想品德和个人修养。此外，作为本科生核心课程的主讲教师和课程负责人，《计算机组成原理》获第二批国家级一流本科课程和首届江苏省一流本科课程，参与省部级虚拟教研室建设试点项目，参编《计算机组成原理》教材第3版（“十一五”国家级规划教材、“十三五”江苏省高等学校重点教材、面向新工科专业建设计算机系列教材、国家级一流本科课程配套教材），指导研究生发表CCF A类论文多篇、获IEEE iThings2020最佳论文奖、IEEE CPSCom2022最佳论文奖、南京理工大学优秀博士生培养基金等奖励。



觉悟 ChatGPT，科研第五范式

摘要：科研范式是常规科学所赖以运作的理论基础和实践规范，随着科学的发展以及外部环境的推动不断发生变化，人类的科学研究已经经历了经验科学、理论科学、计算科学以及数据科学四类范式。**ChatGPT** 取得的瞩目成果展现了人机融合驱动的研发逻辑的巨大潜力，可以看出，如今科研范式的变革正处在“哥白尼革命”的前夜，其基本理论和方法将会发生巨大变化。本文从 **ChatGPT** 的研发逻辑出发，分析了 **ChatGPT** 对科研范式变革的启示，对结合非线性抽象思维（人）和逻辑推理（机）的人机融合新科研范式进行了展望。

1. 科研范式总是随着科技的重大进步而变革

科学研究范式，指的是常规科学研究所赖以运作的理论基础和实践规范，是从事研究的科学家群体所共同遵从的世界观和行为方式。科学研究范式建立的目的是帮助科学们以最有效的方式完成科学研究，提高科学的研究的准确性和可靠性。

科学研究范式不是一成不变的，它会随着科学的发展以及外部环境的推动不断发生变化。由于科学家对科学研究范式的信奉，受到时代认知的局限性，某种科学研究范式总会在科学发展到一定程度后显示出不足而无法解决一些问题，出现困难、矛盾和困惑，这种矛盾推动了科学家们的反思和进一步探索，进而逐渐形成新的科学研究范式。美国科学哲学家 **Kuhn** 认为，科学家从一个旧范式跳到一个新范式时，往往是从一个预测性较弱的范式跳到一个预测性较强的范式，只不过他们在这样做的时候，根本无法理解新范式优越的预测力背后的根本原因。根据图灵奖得主 **Jim Gray** 的观点，人类的科学研究已经经历了四类范式，分别是经验范式、理论范式、模拟仿真范式和数据挖掘范式。科学研究范式的演进对于科学家理解和解决更复杂的科学研究问题发挥着重要指导性作用。

第一范式：经验范式：以经验主义和人的深度思考为主导的科学研究范式，实验是开展研究的主要手段。实验法最早可追踪到古希腊和中国，数千年文明史中，人类绝大多数技术发展源于对自然现象观察和实验总结。典型案例如：人类记录各种自然现象（钻木取火、摩擦起电等），柏拉图、亚里士多德对哲学的思考等。经验对于科学研究是必不可少的，但经验范式本质上是在已经猜测的理论之间进行选择，而不是发现理论。

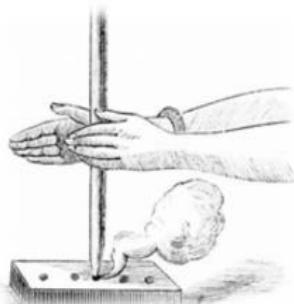


图2 古代钻木取火

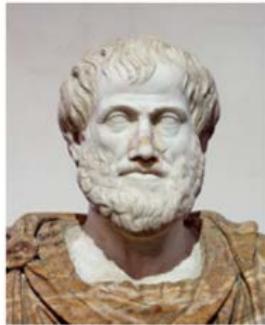


图3 亚里士多德“第一哲学”形而上学

第二范式：理论科学：当实验条件不具备的时候，第一范式难以为继，此时为了研究更为精确的自然现象，催生出了新的科学的研究范式。第二范式是以建模和归纳的理论学科和分析为主导的科学的研究范式。相比于依赖观察和实验的第一范式可以做到“知其然”，第二范式的科学理论需要做到“知其所以然”，对自然界某些规律做出背后原理性的解释，不再局限于描述经验事实。如：数学中的集合论、图论、数论和概率论；物理学中的经典力学，相对论、弦理论、圈量子引力理论；地理学中的大陆漂移学说、板块构造学说；气象学中的全球暖化理论；经济学中的微观经济学、宏观经济学以及博弈论；计算机科学中的信息论等。



图4 牛顿三大定律



图5 科学方法之父—伽利略

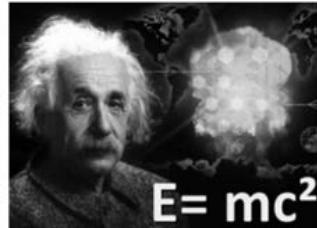


图6 爱因斯坦相对论

第三范式：计算科学：随着理论研究的深入，以量子力学和相对论为代表的理论对超凡的头脑和复杂的计算提出了超高要求，同时验证理论的难度和经济投入也越来越大，第二范式面临重大瓶颈和挑战，迫切需要提出新的科学的研究范式。第三范式是以计算和模拟为主导的科学的研究范式，由 1982 年诺贝尔物理学奖获得者肯尼斯·威尔逊 (Kenneth Wilson) 提出并确立。20 世纪后半叶，伴随高性能计算机和基于大规模并行计算的计算机体系结构的发展，科学家尝试在理论模型指导下，利用计算机设计数值求解算法、编写仿真程序来推演复杂理论、模拟复杂物理现象，如弹药爆炸、病毒传播、天气预报、人口增长、温室效应等。借助计算机的巨大算力，科学家可以精确地、大规模地求解方程组，进而去探索那些无法通过实验法和理论推导法解决的复杂问题。经过合理的模型

假设和简化，第三范式主导的计算科学在经济学、心理学、认知科学等缺乏简单、直观分析解决方案的软科学问题中同样取得了成功应用。

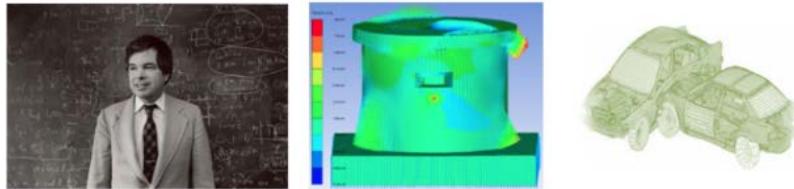


图 7 肯尼思·威尔逊，美国物理学家，因建立相变临界现象理论获 1982 年诺贝尔物理学奖

图 8 Autodyn 模拟弹药爆炸冲击波对地堡的破坏效果

图 9 基于有限元数值模拟的汽车碰撞测试

第四范式：数据科学：第三范式是先提出可能的理论，再搜集数据进行仿真计算和验证，然而随着科学的发展和环境的变化，人们可能已经拥有了大量的数据，但难以直接提出可能的理论，此时第三范式的指导意义有限，需要开发或总结新的科学的研究范式。第四范式是以数据驱动为主导，通过数据和算力探索前沿的科学的研究范式，由 1998 年图灵奖获得者吉姆·格雷（Jim Gray）提出。第四范式以数据考察为基础，联合理论、实验和模拟等方法于一体，也称数据密集型科学。其与第三范式的区别在于，随着数据量的高速增长，计算机不仅仅局限于按照科学家设定的程序规则开展模拟仿真，还能从海量数据中发现规律，形成基于关联关系的科学理论，其本质是通过海量数据的收集代替人类传统的经验观察过程，借助机器的高算力代替人类的归纳推理，从而实现远超经验范式的理论归纳能力。这种通过程序 + 数据发现规则的过程，可以在一定程度上代替过去由科学家才能完成的工作。当今处于数据爆炸的时代，数据科学成为技术发展的前沿领域，第四范式强调借助并行计算、数据挖掘、机器学习等技术去发现隐藏在数据中的关系与联系，将数据作为解决问题的工具而非问题本身。

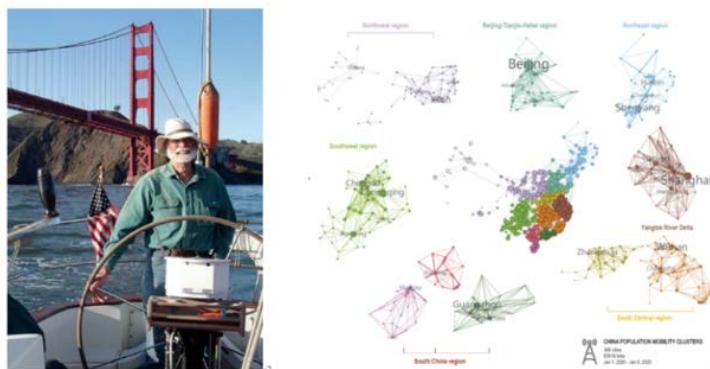


图 10 吉姆·格雷（James Gray），因推动数据库技术发展获 1998 年图灵奖，喜爱户外运动，2007 年 1 月 28 日独自驾船在海上失联至今

图 11 移动大数据揭示新冠肺炎疫情全球大流行背景下人口出行模式规律——吕欣 National Science Review 成果



2. ChatGPT 正在接近人类研究问题的逻辑

2.1 ChatGPT/GPT4 研发逻辑

ChatGPT 是由 OpenAI 最新发布的一种对话式自然语言处理模型，其前身是 InstructGPT 模型，与 BERT 等大语言模型（Large Language Model, LLM）相同，GPT 采用了以 Transformer 为核心的编码器 - 解码器结构。通俗来讲，该类模型的原理类似于人脑的语言处理过程，基于自注意力机制编码器能够获取文本内部的关联关系，进而将其编码为可存储、更抽象的语义向量。解码器则可根据该语义向量生成文本序列。鉴于自然语言任务的内在相似性，通过在海量文本数据上进行自监督训练，可以得到具有丰富知识表达的预训练模型，达到“理解”语言的目的，通过在机器翻译、知识问答等不同下游任务上进行微调训练，即可完成多种自然语言处理任务。

GPT 第一代到第三代采用的都是监督学习的思路，但数据规模和模型规模爆炸式增长，GPT-1 的参数量是 1.17 亿，预训练数据量约 5GB，GPT-2 分别使用了 15 亿参数和 40GB 的训练数据，而 GPT-3 的模型参数量多达 1750 亿，预训练数据量多达 45TB，模型效果也明显提高，但是仍然存在内容胡编乱造、不遵循用户的明确指示、内容偏见有害等显著问题。

而 ChatGPT 的火爆，在于其文本生成结果的准确性、完整度、连贯性、交互性相较于 GPT-3 有了显著提升，在问答对话、代码生成、文本摘要等不同领域都接近甚至超过了人类水平。相对于 GPT-3，ChatGPT 核心创新点主要有三个方面。

2.1.1 指示学习（Instruct Learning）

大语言模型的主流训练模式经历了数次变迁。首先是预训练 + 微调（Pre-training + Fine-tuning），接着是提示学习（Prompt Learning）。但真正改变规则的是指示学习的前身——上下文学习（In Context Learning）。

OpenAI 自 GPT-2 的研发中就开始探索大语言模型的泛化迁移（即零样本、小样本）能力，希望能够实现训练一个模型作为基座支撑，用户只需调用接口，输入想让其完成的任务要求即可，并不需要对大语言模型进行领域微调。但由于 GPT-2 仅抛弃微调阶段，其余基本延续 GPT-1 的两模式，其效果并不好。当到了 GPT-3 时，上下文学习被开创性的提出，部分解决了上述困境。与之前的训练模式相比，上下文学习在预训练阶段将输入样本组织成“任务描述 + 样例 + 提示”的模式，通过海量样本和 1750 亿训练参数的支撑，GPT-3 基本实现了零样本能力。

OpenAI 并未止步于此。在他们对人工智能的设想中，应该是人工智能优化改善人们的生活，而不是用户倒过来迁就它。上下文学习的确有效，但不符合人类的使用模式。日常中，非人工智能从业者们很难给出一个规范的“任务描述 + 样例 + 提示”，以要求人工智能给出答案。因此，OpenAI 在 ChatGPT 的训练中继续往前走了一步，雇用了 40 名标注人员进行提示（Prompt）设计等工作，以人的认知经验引导模型的训练。在

指示学习中，所有输入样例都符合真实应用场景中，人们的语言习惯和行文风格。比如执行翻译任务，训练样本大致上是“帮我把这个英文句子翻译成中文：Today is a good day”。通过这一设计，真正让人工智能服务于人类日常应用，并在最终性能表现上证明了其优异性。

2.1.2 人类反馈的强化学习 (Reinforcement Learning with Human Feedback, RLHF)

自 Deepmind 2017 年研制出 AlphaGo 系列横扫围棋界后，强化学习破圈走进大众视野。但其在自然语言处理领域始终未能得到良好应用，根本原因之一在于目标难定义。不像棋类游戏那样有明确的规则约束，一段人工智能生成的文字或是给出的回答，很难直接判定“好坏”，这也就导致了强化学习中的奖励机制难以界定。

OpenAI 在 ChatGPT 的实践中给出了一个精妙的设计思路，即人类反馈的强化学习。整个过程包含三个阶段。一是有监督学习阶段。先是通过专业标记者设计出小规模、高质量、全领域的提示语 (Prompt)，然后利用这些数据，微调 GPT-3。二是奖励模型训练阶段。利用上一步获得的 GPT-3，对收集到的真实问题生成系列答案，不同于以往给予生成文本固定奖励值的模式，OpenAI 采用了答案排序机制：以是否符合人类认知作为标准对文本进行排序。让专业标记者对这些生成的答案进行排序，并用这些答案数据训练一个奖励模型。三是策略网络学习阶段。此处基于 PPO 的强化学习模式，针对输入的问题，让新模型 (另一个 GPT-3) 生成答案，接下来通过前一步获得的奖励模型对答案打分，反馈更新策略网络。完成了强化学习的自训练迭代模式。

这一过程中最重要的创新点在于引入了人类的认知，答案不再是机器自己判断，或者少数专家给出“黄金解” (Ground Truth)。而是整合了用户的阅读习惯、言语风格等多方位的背景知识后，给出了相对合理的排序让机器去理解。

2.1.3 思维链 (Chain of Thought, CoT)

思维链的技术通俗易懂，将以往期望模型能够一步学会的过程进行拆解，逐步输入到机器里。譬如计算一个复杂的数学逻辑题目，将计算过程分拆为一个个具体的步骤，指导模型一步步学习。

ChatGPT 于 2022 年 11 月投入用户使用后，OpenAI 不断对其在实践中暴露出的偏见、伦理、安全等方面的问题进行不断优化。但 ChatGPT 基座的 GPT3.5 模型本质上是仅能处理自然语言的单模态模型，未能将音视频、图像等多模态信息纳入其处理能力范畴。然而，随着 2023 年 3 月 14 日 GPT4 的发布，这些短板也被攻克。OpenAI 总裁兼联合创始人格雷格·布罗克曼 (Greg Brockman) 表示：“GPT4 不仅仅是一个语言模型，它还是视觉模型。”在官方演示视频中，格雷格用笔和纸画了个网站草图，拍照上传给 GPT-4，让其自主编写一个网页，后者仅用 1 到 2 秒的时间，就生成了网页代码、制作出了几乎与图中一模一样的网站。



在具体性能方面，**GPT4** 可接受的文字输入长度提升到了 2.5 万个单词，允许长内容创建、扩展对话以及文档搜索和分析等。可以更准确地解决高级推理难题，具有更广泛的常识和解决问题的能力，在专业和学术方面更是表现出接近于人类的水平。在模拟律师考试中，**GPT4** 的得分能排进前 10% 左右，而 **GPT3.5** 的得分只能排在倒数 10% 左右。此外，**GPT4** 不仅能对文本或图片进行单独的识别，还可以接受图文混排的内容。遗憾的是，OpenAI 此次仅给出了 98 页关于 **GPT4** 的技术性能报告，未透露任何技术实现方面的细节。

2023 年 3 月 24 日，OpenAI 更是进一步宣布在 **ChatGPT** 中实现了对插件的初始支持，能够帮助 **ChatGPT** 访问最新信息、运行计算或使用第三方服务。此外，不仅亲自下场提供了两款插件：网络浏览器和代码解释器，还开源了知识库检索插件的代码，由任何开发人员自行托管。换句话说，OpenAI 破除了 **ChatGPT** 能力边界的限制，通过第三方插件的形式，逐步打造新的 AI 生态圈，能够为人类衣食住以及工作、学习等方方面面提供更加优质的服务。

2.2 科研范式变革之前夜

复杂模型、海量数据、超大算力是当前科技的鲜明特征。据谷歌统计，自数字化时代以来全球产生的数据量基本服从“大数据摩尔定律”，即每两年就会翻一倍。尤其是近年来，随着一些复杂系统的出现，以及物联网技术的发展，人类社会积累的数据量呈加速增长趋势。例如，欧洲核子研究中心大型强子对撞机每秒钟就可以产生 40 兆兆字节的数据，一架波音喷气机引擎每分钟即产生 20 兆兆字节运行信息。国际数据公司（IDC）曾预测，全球数据总量将从 2018 年的 33ZB（1ZB=10²¹ 字节）增长到 2025 年的 175ZB。

通过超大算力对海量数据进行知识挖掘，通过人工智能技术实现复杂系统的推理，使得科学的研究的效率得到了空前提高。在应用科学领域，李飞飞团队大型图数据库 **ImageNet** 涵盖约 2 万类、1400 万张标注了类别的图像，通过使用深度学习技术成功地提高了图像识别的准确率，将错误率降低到了历史最低水平；谷歌旗下 **DeepMind** 公司的 **AlphaZero** 使用了深度学习和强化学习方法，在 40 天内进行了约 4900 万次围棋自我对弈，并达到了超越人类和其他计算机程序的水平。在基础科学领域，**DeepMind** 使用人工智能来帮助证明或提出新的数学定理，辅助数学家形成对复杂数学的直觉；**AlphaFold** 基于生命科学积累的大量数据，能在几分钟内预测出一个典型蛋白质的结构，无需对蛋白质进行结晶或使用昂贵的冷冻电镜进行研究，颠覆了传统结构生物学的研究模式。在分子生物学、天文学、地理科学、大气科学等数据积累丰富、结构化程度高、问题定义清晰的领域，数据和人工智能的结合已经带来了科研范式的巨大变革。

当前，科学的研究正在进入人机融合的时代，**ChatGPT** 采用人机融合的设计模式。借助人类反馈的强化学习技术实现了对话这一自然语言处理任务的突破，2023 年 3 月 14 日发布的 **GPT-4** 模型，拥有图像识别、高级推理等功能，其单词处理能力是 **GhatGPT** 的

8 倍，在模拟律师考试、SAT（美国高考）数学考试中成绩分别超过 90% 和 89% 人类考生，可见，在具有主观、非线性、不良结构特征的研究对象上，仅靠机器已经难以实现有效的科学发现，基于人机融合的科学的研究方式开始展现其强大的潜力。

近年来，大数据、人工智能等相关技术的更新换代呈现出越来越快的趋势。以自然语言处理（NLP）领域为例，其自诞生之日起历经五次研究范式转变，其中小规模专家知识方法发展了 40 年（1950-1990），浅层机器学习模型发展了 20 年（1990-2010），深度学习算法发展历时 10 年（2010-2017），而大语言模型从 2018 年概念提出到 ChatGPT 产品落地才用了 5 年。新一代技术何时出现不得而知，但也许，我们已经到了科研范式发生变革的前夜。

3. 人机融合的科研第五范式渐露端倪

随着数据获取以及计算能力的不断提高，科学的研究从经验范式发展到了数据科学范式，可以从海量数据中挖掘出仅凭人脑难以发现或总结的科学规律，但经过多年的科学实践可以发现，不论是计算科学还是数据科学范式，在面对社会、经济、人脑智能等复杂巨系统科研对象时，都存在数理模型难以构建、数据学习效率低下、内在机理不明等局限性，数据科学范式的主要作用是发现关联规律，而忽略了因果关系，而很多情况下，简单的因果推断就能胜过大量数据分析所得到的经验规律，如果仅靠数据的统计分析，不赋予机器常识和深度理解能力，则很难在复杂对象的科学理论发现上达到人类的智能水平。GPT1 到 GPT3 等 ChatGPT 的前代版本通过典型的数据科学范式进行设计，虽然在很多自然语言处理任务上取得了较好的效果，但难以实现类人响应，ChatGPT 进行了颠覆式迭代，通过将人类经验融合到模型设计当中，使其具备了从人类的反馈中强化学习并重新思考的能力，在对话这一自然语言处理任务上实现了突破性的效果。

可以看出，通过将人的直觉性经验或专家性经验融合到数据模型或者计算模型当中，以人类专家经验引导改进“机器”的低效探索，发挥“机器”的计算能力优势和人类的直觉性优势，以人机融合、人在回路的形式进行科学实践，可以弥补“机器”无法感知或推理某些难以量化的科学规律上的局限性。随着研究对象从客观对象向主观对象不断发展、从简单系统向复杂巨系统不断拓展，科学的研究的范式也正在从数据科学范式向人机融合的第五范式转变。

在已有的四个科研范式中，科技进步为人们提供了新的理论、方法和工具，但人类始终没有被作为生产力列入科研范式之中。事实上，人在科研范式中的主导地位一直存在，例如第一范式中人类观察总结，第二范式中人类归纳推导，第三范式中人类建模分析，第四范式中人类设计框架等。人在科研范式中的主导地位是因为人类具有非线性抽象思维，人的大脑是开放变化的复杂系统，能够轻松实现整体大于部分之和；善于处理关联演化性问题，哪怕对于无迹可寻的事物也可能顿悟致知。科研范式发展到今天，借鉴 ChatGPT 的研发逻辑，人的定位应该从幕后走向台前，为擅长逻辑推理的机器赋予人



类特有的非线性抽象思维，机器积累量变，人脑触发质变，螺旋升级共同促进科学技术的进步和发展，实现真正意义上的人机融合，称之为科研第五范式。

人机融合的科研第五范式可简要概述如下：首先，人类深度参与、促进并主导机器进化，创造更加开放、更加智能、更加强大的工具；其次，新工具将人类从一部分“体力”劳动中解放出来，让人集中精力于非线性抽象思维，同时新工具还可能提高人类的整体认知水平，为人脑进化创造新环境；然后，认知或思维进化后的人类再次 / 同时促进擅长逻辑推理的机器进化，如此迭代，相互纠缠，螺旋上升。

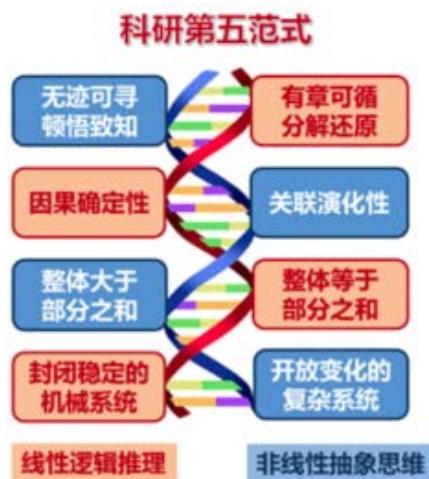


图 13 科学研究的第五范式

4. 总结

如美国著名哲学科学家 Thomas Samuel Kuhn 所述，科研范式是常规科学所赖以运作的理论基础和实践规范。ChatGPT 的横空出世，让我们看到，如今科研范式的变革正处在“哥白尼革命”的前夜，其基本理论和方法将会发生巨大变化。如同牛顿引力理论取代亚里士多德物体运动理论，相对论和量子物理取代引力理论，达尔文进化论取代神创论。在经历经验科学范式（直观描述现象）、理论科学范式（基于模型或归纳表征现象）、计算科学范式（仿真模拟复杂现象）、大数据科学范式（数据分析挖掘）之后，我们相信以逻辑推理（机）结合非线性抽象思维（人）的人机融合新科研范式无疑将会大放异彩，它更加强调创新导向、价值导向、贡献导向；更有利子推动科研模式从无价值负载的信马由缰式研究转向负责任、有组织的研究转变，更能够促进如曼哈顿工程、阿波罗登月计划、引力波探测、中国的神舟飞船等大科学、大工程的建设，在未来科学技术发展中占据主流位置，成为推动科技发展的重要基础。

组合测试理论与方法研究

基本情况：

项目名称：组合测试理论与方法研究

完成人：聂长海、吴化尧、钮鑫涛

完成单位：南京大学

项目简介：

软件已经成为社会发展的基础设施，软件质量正在成为制约发展速度和质量的重要因素，软件测试是软件质量保证的重要手段，其关键问题是寻找尽可能小的测试用例集、以最少的成本发现待测试软件中潜在的问题。人们从不同角度、基于不同的理论、针对软件开发的不同阶段和不同类型的软件提出了很多软件测试方法，这些方法的综合使用为提高软件质量做出了重要贡献。组合测试方法就是其中一种，它旨在系统测试软件中存在的各种可能的“组合”，检查各种“组合”可能触发的软件故障。组合测试由于其广泛的应用价值越来越受到人们的广泛重视，自 1985 年第一篇研究论文发表以来，到目前为止已经累计有近 800 篓相关研究论文。早在 2000 年初期，组合测试已经被列为美国国防部测试标准，2015 年被正式列入国际软件测试标准（ISO/IEC/IEEE 29119），随后我国也将其列入了国家标准。从 2012 年开始，组合测试开始有个专门的国际研讨会（IWCT），一直到现在，我们项目组成员一直参与组织工作，并承担过联合主席。

项目负责人自 2001 年跟华为公司合作测试数据生成项目，开始关注到这项技术，应用项目驱动，系统地研究了这项技术，并为华为公司开发一个组合测试工具，具体技术内容后来发表在计算机学报 2004 年第 3 期上，当年可以查到的参考文献只有 20 多篇。通过对基于组合覆盖的软件测试技术、最小测试用例集生成技术，以及对基于组合测试的软件故障诊断方法等研究积累，于 2004 年获得国家自然科学基金青年科学基金项目“测试用例集优化约简的若干问题研究（60403016）”，并于 2005 年获得江苏省自然科学基金青年科学基金项目“最小测试用例集生成技术研究（BK2005060）”，这两个项目的重要研究内容就是组合测试的测试用例生成、约简和故障诊断，项目取得了丰硕的研究成果，期间发表学术论文 16 篓。

在以上两个项目的支持下，经过 3 年的扎实工作，项目负责人凭借取得的成果，在 2007 年获得国家自然科学基金面上项目“组合测试技术中的关键问题研究（60773104）”，



2008 年获得国家 863 科技专项“组合测试技术及其自动化支持工具（2008AA01Z143）”，2010 年获得江苏省自然科学基金面上项目“组合测试的理论、方法和应用（BK2010372）”支持，项目执行期间各项成果获得重大突破，2013 年又得到教育部博士点基金项目“组合测试的理论、方法与实证（20130091110032）”。

项目培养了一批硕士和博士研究生，最为突出的吴化尧博士和钮鑫涛博士，他们在攻读博士期间对组合测试进行了深入系统的研究，产出了一批创新性成果和一系列高水平论文，毕业留校工作后分别获得了国家自然科学基金项目“智能化组合测试方法及应用研究（61902174）”、“组合测试中极小特征模式的理论及应用（62102176）”和江苏省自然科学基金项目“自适应组合测试理论和方法研究（BK20190291）”、“组合测试技术中故障定位的理论及方法（BK20221439）”的支持，为组合测试的研究和发展注入了新的重要的活力。

主要科技创新

组合测试是一种充分考虑各种因素及其相互作用的科学实用的软件测试方法，这种方法克服了正交实验设计，均匀设计等传统方法的不足，设计一组数量较少的测试用例，直接检测各种影响因素及其组合对软件产生的影响，更适合于软件测试。近年来，组合测试技术日益受到人们重视，并得到了人们广泛的研究和应用。特别是在信息技术比较发达的美国、日本、加拿大和欧洲等国家，一些学者在组合测试的测试用例生成和组合测试的应用等方面进行了系统深入的研究。

组合测试的关键问题分布在软件系统组合模型的抽取、测试数据生成、测试结果的分析评估、组合测试故障诊断以及组合测试过程改进等方面，其核心问题是如何以最快捷的速度精心设计一组科学的、最小规模的组合测试数据实现对待测试软件科学有效的测试，并形成对该软件质量的科学判断，具有重要的现实意义。

组合测试技术的研究不仅具有重要的应用价值，同时在理论上也具有重要的科学价值，组合测试技术领域的研究充分反映了计算机科学领域抽象、理论和设计的三个学科形态的特征。例如，组合测试数据生成问题一经提出，人们关于这个问题的研究很快就超出了应用的范围，而把这个问题当成了一个具有应用背景的数学问题来研究，引起了很多数学家的研究兴趣，已有不少研究成果刊登在国内外的一些重要数学杂志上，这些研究成果无疑对于组合测试技术的应用研究起到积极的推动作用。

组合测试技术的研究结果不仅可以对软件测试领域产生积极作用，由于计算机软件是一种特殊的产品，不仅其他领域关于产品检查测试的方法可以被软件测试借鉴和使用，同时，在软件测试技术领域有效的研究结果也可以为其他领域的质量检查提供新的思路，特别是组合测试方法，这种方法一开始实际上是来源于工农业生产中实验方法的改进，关于这种方法的研究结果，可以广泛地应用到各行各业，从而对经济建设产生积极的影响。近 20 年来，我们在组合测试数据生成、组合测试故障诊断、组合测试应用等方面进行了

系统、深入和细致的研究，取得了一系列成果，具体包括以下方面：

1 组合测试数据生成

软件测试的一个基本问题就是如何设计有效的测试用例以揭示系统中各种因素及其相互作用所触发的软件故障，这正是组合测试的目标。组合测试所要解决的关键问题就是从系统参数庞大的组合空间中选择少量组合进行测试以有效地发现软件故障，而引起软件故障的参数组合则称为“故障模式”。针对这个基本问题，我们提出了一系列的 2 维覆盖表组合测试用例生成方法，包括 数学方法 [1,2]，启发式贪心算法 [3-6]，演化搜索方法 [7-13] 等，还研究 n 维覆盖表 [14]，相邻因素覆盖表 [15]，可变力度覆盖表 [16,17] 等不同类型组合测试用例集的生成进行了系统的研究 [18]，覆盖表生成是组合测试研究的关键问题之一，其中，贪心算法因为速度快、生成的覆盖表规模小而得到人们的青睐。人们提出了很多基于不同策略的贪心算法，其中，多数算法可以归结到一个统一的算法框架，即形成一个可配置贪心算法，从该框架又可以衍生出很多新的算法。如何科学地配置优化受多个因素影响的算法框架、有效生成覆盖表是一个新的挑战。针对具有 6 个决策点的贪心算法框架，设计了 3 条不同的实验路线，系统地探索各个决策点以及它们之间相互作用对生成覆盖表规模的不同影响，寻找最佳配置，从而可以有效地生成规模更小的覆盖表，为覆盖表生成的贪心算法的设计和优化提供理论和实践基础。

我们总结了目前组合测试领域中搜索技术的研究现状，提出了基于搜索的组合测试这一概念，并讨论了其未来的研究方向 [19]。提出了一种新的用于覆盖表生成的离散粒子群算法，该方法克服了传统粒子群算法在离散空间中搜索的局限性，能生成比已有搜索技术规模更小的覆盖表。这一研究不仅能帮助我们更高效地自动生成测试数据，同时也进一步验证了搜索技术在软件测试领域的实用性 [20]。由于蚁群算法也是一种能够有效求解组合优化问题的演化搜索算法，已经被应用到求解覆盖表生成问题中。已有的研究工作表明蚁群算法适于求解一般覆盖表、变力度覆盖表生成以及覆盖表排序等问题，但算法结果与其他覆盖表生成方法相比并不具有优势。为了进一步探索与挖掘蚁群算法生成覆盖表的潜力，我们进行了四个层次的改进工作：1) 算法变种集成；2) 算法参数配置优化；3) 演化对象结构调整及演化策略改进；4) 利用并行计算优化算法时间开销。实验结果表明，通过以上四个层次的改进，蚁群算法生成覆盖表的性能有了显著提升 [21]。

[1] 聂长海, 徐宝文. 基于接口参数的黑箱测试用例自动生成算法. 计算机学报, 2004, 27 (03): 382-388

[2] 聂长海, 徐宝文, 史亮. 一种新的二水平多因素系统两两组合覆盖测试数据生成算法, 计算机学报, 2006,29(6):841-848

[3] 史亮, 聂长海, 徐宝文. 基于解空间树的组合测试数据生成, 计算机学报, 2006,29(6):849-857

[4] 聂长海, 蒋静, 覆盖表生成的可配置贪心算法优化, 软件学报,



2013,24(7):1469~1483 .

[5] Changhai Nie, Baowen Xu, Liang Shi, Ziyuan Wang. A new heuristic for test suite generation for pair-wise testing, seke' 2006, San Fransisco: 517-522

[6] Nie, Changhai1, Jiang, Jing1, Wu, Huayao1, Leung, Hareton2, Colbourn, Charles J.3, 4, Empirically identifying the best greedy algorithm for covering array generation, ICSTW 2013, 卢森堡, 2013.3.18-2013.3.20

[7] 查日军, 张德平, 聂长海, 徐宝文. 组合测试数据生成的交叉熵与粒子群算法及比较. 计算机学报, 10 期, pp 1-14, 2010/10/15

[8] 梁亚澜, 聂长海. 覆盖表生成的遗传算法配置参数优化. 计算机学报, 2012, (07):1522-1539

[9] Changhai Nie, Hareton Leung, and Baowen Xu. Using computational search to generate 2-way covering array. In Proceedings of the 1st International Symposium on Search Based Software Engineering (SSBSE' 09), Fast Abstract. Cumberland Lodge, Windsor, UK May 13-May 15, 2009.

[10] Changhai Nie, Baowen Xu, Ziyuan Wang and Liang Shi. Generating Optional Test Set for Neighbor Factors Combinatorial Testing, QSIC' 2006. Beijing, 259-265.

[11] Changhai Nie, Huayao Wu, Liang Yalan, Hareton Leung, Fei-Ching Kuo and Zheng Li Search Based Combinatorial Testing, The 19th Asia-Pacific Software Engineering Conference (APSEC 2012). 2012, Dec. Hong Kong.

[12] 孙文雯, 聂长海. 一种组合测试用例生成的可配置混合算法. 计算机科学, 2011, 38 (8) : 130-135.

[13] 吴化尧 聂长海, 覆盖表生成的粒子群算法: 参数优化和自适应算法, 小型微型计算机系统, 2012 (10) :2259-2267

[14] Changhai Nie, Baowen Xu, Liang Shi, Guowei Dong, Automatic Test Generation for N-way Combinatorial Testing. Second International Workshop on Software Quality (SOQUA 2005), Lecture Notes in Computer Science 3712, 2005, 203-211

[15] 王子元, 聂长海, 徐宝文, 史亮. 相邻因素组合测试用例集的最优生成方法, 计算机学报, 2007,30(2):200-211

[16] Ziyuan Wang, Baowen Xu, Changhai Nie: Greedy Heuristic Algorithms to Generate Variable Strength Combinatorial Test Suite. QSIC 2008: 155-160

[17] Ziyuan Wang, Changhai Nie, Baowen Xu: Generating combinatorial test suite for interaction relationship. SOQUA 2007: 55-61

[18] 王子元, 徐宝文, 聂长海. 组合测试用例生成技术. 计算机科学与探索, 2008, (06): 571-589

[19] Huayao Wu, 聂长海, An overview of search based combinatorial testing, SBST, 印度, 2014.6.10- 2014.6.15

[20] Huayao Wu, Changhai Nie, FeiChing Kuo, Hareton Leung and Charles J. Colbourn, A Discrete Particle Swarm Optimization for Covering Array Generation, IEEE Transactions on Evolutionary Computation, 19(4): 575-591, 2015

[21] 曾梦凡; 陈思洋; 张文茜; 聂长海, 利用蚁群算法生成覆盖表: 探索与挖掘, 软件学报, 2016.1.01, (04): 855~878

2 组合测试故障诊断

我们一直研究组合测试的故障定位问题, [22,23] 将发现错误的测试用例中包含的故障模式与未发现故障的测试用例所包含的故障模式进行比对排除, 这样可以缩小可能的故障模式范围, 但这种方法效果很差, 特别是在使用覆盖表进行测试时, 几乎无法缩小这个可疑故障模式范围。然后用一些附加测试用例进行补充测试, 进一步缩小这个范围。

[24] 将该方法应用于网络软件的故障定位。[25] 将一条发现故障的测试用例所包含的可能的故障模式表示成一个偏序集, 这样通过发现故障的测试用例和未发现故障的测试用例形成的偏序集的做差, 形成的新的偏序集, 我们认为这个偏序集的极小元是最大可能的故障模式, 这个结果尽管比之前的结果改进很多, 但我们找出来的还只是可能的故障模式。最重要的成果是论文 [26] 提出了组合测试中的“最小故障模式”概念。最小故障模式是软件系统中众多可能故障模式的内核, 是使软件系统发生组合相关的故障的根本原因, 它精确地描述了哪些因素间的相互作用会触发故障。最小故障模式中涉及的每个因素都是触发软件故障必不可少的条件之一, 组合测试以及其后的故障定位过程本质上就是发现和寻找最小故障模式的过程。我们还利用差异调试法 [27,28] 和元组关系树方法对故障原因进行定位 [29,30]。[31] 系统研究了组合测试下多故障问题对组合测试的故障定位造成的影响, 并给出了一种轻量级测试用例替代算法来削弱这种消极的影响。

[22] 聂长海, 徐宝文, 史亮. 一种基于组合测试的软件故障诊断方法. 东南大学学报(自然科学版) 2003/06

[23] Liang Shi, Changhai Nie, Baowen Xu: A Software Debugging Method Based on Pairwise Testing. International Conference on Computational Science (3) 2005: 1088-1091

[24] Lei Xu, Baowen Xu, Changhai Nie: Testing and Fault Diagnosis for Web Application Compatibility Based on Combinatorial Method. ISPA Workshops 2005: 619-626

[25] 徐宝文, 聂长海, 史亮, 陈火旺. 一种基于组合测试的软件故障调试方法. 计算机学报, 2006, 29(1):124-131

[26] Changhai Nie, Hareton Leung. The minimal failure-causing schema of combinatorial testing, ACM Transaction on Software Engineering and Methodology



(TOSEM), 20(4), 2011.

[27] Jie Li, Changhai Nie, Yu Lei: Improved Delta Debugging Based on Combinatorial Testing. QSIC 2012: 102-105

[28] 李杰, 聂长海, 梁亚澜 软件兼容性测试的故障定位, 第七届中国测试学术会议, 2012, 6月 23-24, 中国, 杭州, 294-301

[29] 钮鑫涛, 聂长海, CHAN Alvin, 组合测试故障定位的关系树模型, 计算机学报, 2014.01.01, (12): 2505~2518

[30] Xintao Niu, Changhai Nie, Lei Yu and Alvin Chan, Identifying Failure-Inducing Combinations Using Tuple Relationship, accepted by International workshop on combinatorial testing (IWCT) in International conference on software testing, verification and validation, ICST 2013

[31] Xintao Niu, Changhai Nie, Yu Lei, Hareton Leung and Xiaoyin Wang, Identifying Failure-Causing Schemas in the Presence of Multiple Faults, TSE, 46(2): 141-162, 2020 (CCF-A)

3 组合测试理论与模型

针对传统的组合测试方法在测试初期无法给出准确的模型, 无法及时、充分利用在测试过程产生的信息以及故障定位和约束处理相对滞后等不足, 提出了自适应组合测试方法, 给出了可以动态修改测试模型和测试策略的自适应组合测试框架 [32], 并进行了实证研究 [33]。[34] 形式化定义了除故障模式和健康模式之外的第三种模式: 待定模式 (pending schema), 该模式能够确认组合测试故障定位算法的不完整性, 给出了待定模式和其它模式和测试用例之间的关系, 以及一个高效的算法能够之间从已知的测试用例集中抽取出待定模式。建立了组合测试的待定模式 (pending schema) 理论, 利用待定模式的基本偏序集的性质, 将需要 $2n$ (其中 n 为系统参数个数) 复杂度 (存储和计算) 的定位算法降低到只需要多项式的级别。

[32] Nie, Changhai, Leung, Hareton, Cai, Kai-Yuan, Adaptive combinatorial testing, 13th International Conference on Quality Software, QSIC 2013, 中 国, 2013.7.29-2013.7.30

[33] Nie, Changhai, Chen, Siyang, Leung, Hareton, Cai, Kai-Yuan, A case study of a daptive combinatorial testing, 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops, COMPSACW 2013, 日本, 2013.7.22-2013.7.26

[34] Xintao Niu, Huayao Wu, Changhai Nie, Yu Lei, and Xiaoyin Wang. A Theory of Pending Schemas in Combinatorial Testing. TSE, 48(10): 4119-4151, 2022[CCF-A]

4 组合测试优化

在组合测试的测试用例集优化方面, 我们首先针对相邻测试用例间参数 重配置时所

产生的切换成本这一问题，提出了一种基于切换成本的测试用例集排序策略，该方法可以有效地降低测试执行时的参数重配置次数，从而节约测试开销 [35]。特别在测试用例优先级排序方面，提出了一种基于权重的方法来衡量测试用例间的切换成本，并基于此对测试用例集进行优化来提高故障检测率。我们发现贪心算法在这一问题上易陷入局部最优，而能求得最优解的动态规划算法又受限于测试用例集规模，因此需要考虑演化算的使用。同时，我们还发现现在切换成本和组合覆盖率间存在权衡，仅优化其中的一个因素并不一定能得到最优的故障检测率。因此我们进一步使用了多目标演化算法，并通过与单目标方法的比较验证了多目标优先级排序的有效性 [36]。其次，针对测试用例集中存在一定的组合冗余这一问题，我们提出了一种随机后优化覆盖表约简算法，该方法能在满足组合覆盖的前提下进一步约简所需的测试用例数目 [37]。针对传统组合测试生命周期存在冗余的问题，提出了一种全新的交互式框架，将传统的测试到故障定位线性流程变成了一个测试和定位交互的流程，避免了重复定位，减少了测试用例冗余，使得测试用例生成和故障定位可以互相利用已有信息，提高测试效率 [38]。针对传统组合测试缺乏 Oracle 的支持，首次给了结合蜕变关系增强组合测试的测试用例 oracle 自动判定框架，搭建了第一个适用于 CT 和 MT 的评估数据集，评估了算法有效性 [39]。

[35] Wu, Huayao, Nie, Changhai, Kuo, Fei-Ching, Test suite prioritization by switching cost, ICSTW 2014, 2014.3.31-2014.4.4

[36] Wu, Huayao, Nie, Changhai, Kuo, Fei-Ching, The optimal testing order in the presence of switching cost, Information and Software Technology, 2016.12.01, 80: 57~72

[37] Li, Xiaohua, Dong, Zhao, Wu, Huayao, Nie, Changhai, Cai, Kai-Yuan, Refining a randomized post-optimization method for covering arrays, ICSTW 2014, 2014.3.31-2014.4.4

[38] Xintao Niu, Changhai Nie, Hareton Leung, Yu Lei, Xiaoyin Wang, Jiaxi Xu and Yan Wang, An Interleaving Approach to Combinatorial Testing and Failure-inducing Interaction Identification, IEEE Transactions on Software Engineering (TSE), 2020, 46(6):584-615 (CCF-A)

[39] Xintao Niu, Yanjie Sun, Huayao Wu, Gang Li, Changhai Nie, Yu Lei, and Xiaoyin Wang. Enhance Combinatorial Testing with Metamorphic Relations. IEEE Transactions on Software Engineering (TSE), 48(12): 5007-5029, 2022

5 组合测试比较研究

人们通常认为组合测试要优于较为简单的随机测试和自适应随机测试，但有一些研究认为组合测试和随机测试之间没有显著区别，同时也有一些研究显示，自适应随机测试要显著优于随机测试，因而也优于组合测试。为了对这些问题进行探究，我们针对组



合测试、随机测试与自适应随机测试在最小故障模式的检测能力在理论和实证方面分别进行了系统的比较研究 [40]。已有比较研究没有考虑到测试模型的完整性、以及待测故障的故障率对 CT 相对有效性的影响，我们在 9 个真实软件的 1683 个测试场景上开展实验。研究发现强调了约束识别对 CT 故障检测能力的影响，并给出了 CT 表现最优的故障率区间研究组合测试 (CT) 、随机测试 (RT) 和自适应随机测试 (ART) 在不同测试场景下的相对有效性，为辨识组合测试的适用范围提供参考依据 [41]。

[40] Changhai Nie, Huayao Wu, Xintao Niu, FeiChing Kuo, Hareton Leung and Charles J. Colbourn, Combinatorial Testing, Random Testing, and Adaptive Random Testing for Detecting Interaction Triggered Failures, *Information and Software Technology*, 62: 198-213, 2015

[41] Huayao Wu, Changhai Nie, Justyna Petke, Yue Jia and Mark Harman. An Empirical Comparison of Combinatorial Testing, Random Testing and Adaptive Random Testing. *IEEE Transactions on Software Engineering (TSE)*, 46(3): 302-320, 2020 (CCF-A)

6 组合测试的应用研究

我们已经将组合测试应用到软件配置测试 [42]，手机的功能测试 [43]，浏览器兼容性测试 [44] 等方面，特别是服务测试和服务容错性测试，以及智能化软件测试。

现代企业通常以 Web 服务的方式向用户提供软件服务，当前流行的 Web 服务中通常存在 Web API 数量众多、API 输入参数类型复杂、以及不同 API 之间依赖关系复杂的问题。针对上述问题，我们围绕 RESTful APIs 的组合测试方法开展了研究，提出一种能自动确定 API 测试序列及其对应参数取值的组合测试数据生成方法 RESTCT，可以通过文档自动建立测试模型并生成测试用例，让 RESTful API 测试更加自动化，能对服务深层状态进行有效测试 [45]。

服务系统常通过部署冗余组件提升系统的可靠性，以确保系统即使在面对各种无法预测的故障时，仍能正常为用户提供服务。可靠性已成为微服务系统设计和实现中的一个关键问题。故障注入测试是验证和评估软件系统容错机制的常用方法。然而由于故障注入空间的大小随系统冗余组件的增加呈指数级增长导致的效率问题，以及开发和测试团队的分隔性及系统版本的频繁变更导致的即时、精确、完整的系统业务结构信息难以获取的问题，现有的故障注入测试方法均不适用于微服务系统这一测试场景。为此，我们提出了基于容错瓶颈的故障注入算法 FBFI，能够在无法提前获知待测系统的业务结构（如每个服务中的业务节点数量、相邻服务的节点之间的调用关系）的情况下，设计故障注入测试用例以验证待测系统中所有冗余组件的正确性 [46]。

针对多标签图像分类软件给出一个自动化测试框架，具体而言解决三个关键研究问题：测试充分性，测试输入生成，测试结果判定；探究对于深度学习软件而言蜕变关系组合的 cost-effectiveness；将标签空间覆盖问题转化为全二值属性的覆盖表生成问题，

并使用图片拼接来生成测试输入 [47]。

[42] 聂长海, 徐宝文. 软件配置测试的测试数据生成 (英文) *Journal of Southeast University(English Edition)* 2004/01

[43] 黄海波, 吴化尧, 聂长海 手机通话功能的组合测试, 第七届中国测试学术会议, 2012, 6月 23-24, 中国, 杭州, 49-54

[44] Lei Xu, Baowen Xu, Changhai Nie, Huowang Chen, Hongji Yang: A Browser Compatibility Testing Method Based on Combinatorial Testing. ICWE 2003: 310-313

[45] Huayao Wu, Lixin Xu, Xintao Niu, and Changhai Nie. Combinatorial Testing of RESTful APIs. International Conference on Software Engineering (ICSE), 2022: 426-437

[46] Huayao Wu, Senyao Yu, Xintao Niu, Changhai Nie, Yu Pei, Qiang He, and Yun Yang Enhancing Fault Injection Testing of Service Systems via Fault-Tolerance Bottleneck IEEE Transactions on Software Engineering (TSE), 2023: 49(8): 4097-4114 [CCF-A]

[47] Shengyou Hu, Huayao Wu*, Peng Wang, Jing Chang, Yongjun Tu, Xiu Jiang, Xintao Niu, and Changhai Nie. ATOM: Automated Black-Box Testing of Multi-Label Image Classification Systems International Conference on Automated Software Engineering (ASE), accepted, 2023 [CCF-A]

7 组合测试的约束处理

现实存在的组合测试模型一般都是有约束的, 约束覆盖表的生成 (Constrained Covering Array Generation, CCAG) 是组合测试研究领域的一个关键挑战, 众多研究者开发了各种约束覆盖表生成算法。然而, 现有算法通常使用约束求解器或基于禁止元组的技术来处理约束, 这可能会限制它们搜索较小覆盖表的潜力。我们系统总结组合测试覆盖表生成的约束识别、处理和维护技术; 研究适用于不同覆盖表生成算法的最优约束处理技术, 构造规模更小的组合测试用例集 [48]。我们致力于探索更有效的 CCAG 约束处理技术, 以便进一步缩小约束覆盖表的大小, 为此, 提出了一种新的基于自适应惩罚的并行禁忌搜索 (Adaptive Penalty based Parallel Tabu Search, APPTS) 算法来解决 CCAG 问题。APPTS 在适应度函数中引入惩罚项来处理约束搜索空间, 并采用自适应惩罚机制来动态调整不同搜索阶段的惩罚权重。此外, APPTS 采用 Java 并行流来计算候选解决方案的适应度值, 以加快生成过程 [49,50]。

[48] Huayao Wu, Changhai Nie, Justyna Petke, Yue Jia, and Mark Harman. Comparative analysis of constraint handling techniques for constrained combinatorial testing. IEEE Transactions on Software Engineering (TSE), 2021, 47(11): 2549-2562. [CCF-A]

[49] Yan Wang, Huayao Wu*, Xintao Niu, Changhai Nie, and Jiaxi Xu. An Adaptive Penalty based Parallel Tabu Search for Constrained Covering Array Generation.



Information and Software Technology (IST), 143: 106768, 2022

[50] Yan Wang, Huayao Wu, Xintao Niu, Changhai Nie, and Jiaxi Xu. A Constrained Covering Array Generator using Adaptive Penalty based Parallel Tabu Search. International Workshop on Combinatorial Testing (IWCT), 2022: 82-86

8 组合测试研究的总结和基础设施建设

经过近 40 年的发展，组合测试已经发展成为一种独立的和颇具影响的软件测试方法。该方法利用组合覆盖表作为测试用例集，可以有效检测软件系统中各种因素及其相互作用而促发的各种故障。我们在充分分析已有文献的基础上，结合我们在该领域的研究实践，撰写了《A survey of combinatorial testing》，首次系统介绍组合测试的演化发展，总结出其中的关键问题、方法、重要应用和未来发展方向，为组合测试的进一步研究和发展提供支持和参考。该论文 2011 年发表在国际最具影响力的一流期刊《ACM Computing Survey》上 [51]，到目前为止共有 857 次引用，具体有 IEEE & ACM Fellow Mary Jean Harrold (JSS 2013)、IEEE & ACM Fellow Lionel Briand (TSE 2011, JSS 2016)、IEEE Fellow Gregg Rothermel (FOSE 2014, TC 2016)、ASE Fellow Sven Apel (FSE 2013, CSUR 2014, ICSE 2015, ICSE 2016, TOSEM 2018, ICSE 2021)、IEEE Fellow Yves Le Traon (TSE 2014, ICSE 2016) 和 IEEE & ACM Fellow Tao Xie (ICST 2012, TSE 2018) 等众多软件工程领域的国际知名学者在各自论文中介绍组合测试方法时都将本工作列举为本领域的代表性工作。2015 年在科学出版社出版专著《组合测试》，该著作是国内唯一一本系统介绍组合测试方法的专业著作 [52]。2017 年，项目负责人又对组合测试研究进展进行的总结 [53]。自 2013 年以来，我们建设维护了组合测试研究文库 (repository) [54]，该文库收集了所有的组合测试相关研究论文，目前有近 800 篇，同时我们也建设了一个线上组合测试集成工具，我们建设的组合测试微课获得江苏省微课大赛一等奖，国家三等奖，任何时候可以线上播放 [55]。

[51] Changhai Nie, Hareton Leung. A survey of combinatorial testing, ACM Computing Survey, vol. 43, issue # 2, Jan. 2011.

[52] 聂长海, 组合测试, 科学出版社, 北京, 2015.4

[53] 聂长海, 组合测试研究进展, 中国科技论文在线, 2017 年 : 1-15.

[54] http://gist.nju.edu.cn/ct_repository/

[55] <https://weike.enetedu.com/play.asp?vodid=179973&e=3>

附 2：主要论文列表（按照年代顺序排）

[1] Shengyou Hu, Huayao Wu*, Peng Wang, Jing Chang, Yongjun Tu, Xiu Jiang, Xintao Niu, and Changhai Nie. ATOM: Automated Black-Box Testing of Multi-Label Image Classification Systems International Conference on Automated Software Engineering (ASE), accepted, 2023 [CCF-A]

- [2] Huayao Wu, Senyao Yu, Xintao Niu, Changhai Nie, Yu Pei, Qiang He, and Yun Yang Enhancing Fault Injection Testing of Service Systems via Fault-Tolerance Bottleneck, TSE, 2023: 49(8): 4097-4114 [CCF-A]
- [3] Xintao Niu, Yanjie Sun, Huayao Wu, Gang Li, Changhai Nie, Yu Lei, and Xiaoyin Wang. Enhance Combinatorial Testing with Metamorphic Relations. TSE, 48(12): 5007-5029, 2022[CCF-A]
- [4] Xintao Niu, Huayao Wu, Changhai Nie, Yu Lei, and Xiaoyin Wang. A Theory of Pending Schemas in Combinatorial Testing. TSE, 48(10): 4119-4151, 2022[CCF-A]
- [5] Huayao Wu, Lixin Xu, Xintao Niu, and Changhai Nie. Combinatorial Testing of RESTful APIs. International Conference on Software Engineering (ICSE), 2022: 426-437[CCF-A]
- [6] Yan Wang, Huayao Wu*, Xintao Niu, Changhai Nie, and Jiaxi Xu. An Adaptive Penalty based Parallel Tabu Search for Constrained Covering Array Generation. Information and Software Technology (IST), 143: 106768, 2022
- [7] Yan Wang, Huayao Wu, Xintao Niu, Changhai Nie, and Jiaxi Xu. A Constrained Covering Array Generator using Adaptive Penalty based Parallel Tabu Search. International Workshop on Combinatorial Testing (IWCT), 2022: 82-86
- [8] Huayao Wu, Changhai Nie, Justyna Petke, Yue Jia, and Mark Harman. Comparative analysis of constraint handling techniques for constrained combinatorial testing. IEEE Transactions on Software Engineering (TSE), 2021, 47(11): 2549-2562. [CCF-A]
- [9] Xintao Niu, Changhai Nie, Yu Lei, Hareton Leung and Xiaoyin Wang, Identifying Failure-Causing Schemas in the Presence of Multiple Faults, TSE, 46(2): 141-162, 2020 (CCF-A)
- [10] Huayao Wu, Changhai Nie, Justyna Petke, Yue Jia and Mark Harman. An Empirical Comparison of Combinatorial Testing, Random Testing and Adaptive Random Testing. IEEE Transactions on Software Engineering (TSE), 46(3): 302-320, 2020 (CCF-A)
- [11] Xintao Niu, Changhai Nie, Hareton Leung, Yu Lei, Xiaoyin Wang, Jiaxi Xu and Yan Wang, An Interleaving Approach to Combinatorial Testing and Failure-inducing Interaction Identification, TSE, 2020, 46(6):584-615 (CCF-A)
- [12] 王燕, 聂长海, 钮鑫涛, 吴化尧, 徐家喜. 覆盖表生成的禁忌搜索算法, 软件学报, 2018, 29(12): 3665-3691.
- [13] 聂长海, 组合测试研究进展, 中国科技论文在线, 2017 年 : 1-15.
- [14] 曾梦凡, 陈思洋, 张文茜, 聂长海. 利用蚁群算法生成覆盖表: 探索与挖掘. 软件学报, 2016, 27(4):855-878



[15] Huayao Wu, Changhai Nie, Fei-Ching Kuo, The Optimal Testing Order in the Presence of Switching Cost, *Information and Software Technology*, 80: 57–72, 2016

[16] Changhai Nie, Huayao Wu, Xintao Niu, Fei-Ching Kuo, Hareton Leung and Charles J. Colbourn. Combinatorial Testing, Random Testing, and Adaptive Random Testing for Detecting Interaction Triggered Failures. *Information and Software Technology* 62 (2015) 198–213.

[17] Huayao Wu, Changhai Nie, Fei-Ching Kuo, Hareton Leung and Charles J. Colbourn, A Discrete Particle Swarm Optimization for Covering Array Generation, *IEEE Transactions on Evolutionary Computation*, 19(4): 575-591, 2015

[18] Huayao Wu, Changhai Nie, Fei-Ching Kuo. Test Suite Prioritization by Switching Cost. In: Lisa O' Conner ed. *Proceedings of the 7th International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. Cleveland, USA: IEEE, 2014: 133-142

[19] Xiaohua Li, Zhao Dong, Huayao Wu, Changhai Nie, Kai-Yuan Cai. Refining a Randomized Post-optimization Method for Covering Arrays. In: Lisa O' Conner ed. *Proceedings of the 7th International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. Cleveland, USA: IEEE, 2014: 143-152

[20] Huayao Wu, Changhai Nie. An overview of search based combinatorial testing. In: Phil McMinn eds. *The 7th International Workshop on Search Based Software Testing (SBST)*. Hyderabad, India: ACM, 2014, 27-30

[21] 钮鑫涛, 聂长海, Alvin Chan, 组合测试故障定位的关系树模型, *计算机学报*, 2014, 37 (12) : 2505-2518

[22] Changhai Nie, Siyang Chen, Hareton Leung, Kai-Yuan Cai: A Case Study of Adaptive Combinatorial Testing. *COMPSAC Workshops 2013*: 47-52

[23] Changhai Nie, Jing Jiang, Huayao Wu, Hareton Leung, Charles J. Colbourn: Empirically Identifying the Best Greedy Algorithm for Covering Array Generation. *ICST Workshops 2013*: 239-248

[24] Xintao Niu, Changhai Nie, Yu Lei, Alvin T. S. Chan: Identifying Failure-Inducing Combinations Using Tuple Relationship. *ICST Workshops 2013*: 271-280

[25] Changhai Nie, Hareton Leung, Kai-Yuan Cai: Adaptive Combinatorial Testing. *QSIC 2013*: 284-287

[26] 聂长海, 蒋静, 覆盖表生成的可配置贪心算法优化, *软件学报*, 2013/7, 1469-1483

[27] 吴化尧 聂长海, 覆盖表生成的粒子群算法:参数优化和自适应算法, *小型微型计算机系统*, 2012 (10) :2259-2267

- [28] 李杰, 聂长海, 梁亚澜 软件兼容性测试的故障定位, 第七届中国测试学术会议, 2012, 6月 23-24, 中国, 杭州, 294-301
- [29] 黄海波, 吴化尧, 聂长海 手机通话功能的组合测试, 第七届中国测试学术会议, 2012, 6月 23-24, 中国, 杭州, 49-54
- [30] Changhai Nie, Huayao Wu, Liang Yalan, Hareton Leung, Fei-Ching Kuo and Zheng Li Search Based Combinatorial Testing, The 19th Asia-Pacific Software Engineering Conference (APSEC 2012). 2012, Dec. Hong Kong.
- [31] Jie Li, Changhai Nie, Yu Lei: Improved Delta Debugging Based on Combinatorial Testing. QSIC 2012: 102-105
- [32] 梁亚澜, 聂长海. 覆盖表生成的遗传算法配置参数优化. 计算机学报, 2012, (07):1522-1539
- [33] Changhai Nie, Hareton Leung. A survey of combinatorial testing, ACM Computing Survey, vol. 43, issue # 2, Jan. 2011.
- [34] Changhai Nie, Hareton Leung. The minimal failure-causing schema of combinatorial testing, ACM Transaction on Software Engineering and Methodology (TOSEM), 20(4), 2011. [CCF-A]
- [35] 孙文雯, 聂长海. 一种组合测试用例生成的可配置混合算法. 计算机科学, 2011, 38 (8) : 130-135.
- [36] 聂长海, 徐宝文. 基于接口参数的黑箱测试用例自动生成算法. 计算机学报, 2004, 27 (03): 382-388
- [37] 聂长海, 徐宝文, 史亮. 一种新的二水平多因素系统两两组合覆盖测试数据生成算法, 计算机学报, 2006,29(6):841-848
- [38] 徐宝文, 聂长海, 史亮, 陈火旺. 一种基于组合测试的软件故障调试方法. 计算机学报, 2006, 29(1):124-131
- [39] 史亮, 聂长海, 徐宝文. 基于解空间树的组合测试数据生成, 计算机学报, 2006,29(6):849-857
- [40] 王子元, 聂长海, 徐宝文, 史亮. 相邻因素组合测试用例集的最优生成方法, 计算机学报, 2007,30(2):200-211
- [41] 查日军, 张德平, 聂长海, 徐宝文. 组合测试数据生成的交叉熵与粒子群算法及比较. 计算机学报, 10 期, pp 1-14, 2010/10/15
- [42] Changhai Nie, Baowen Xu, Ziyuan Wang and Liang Shi. Generating Optional Test Set for Neighbor Factors Combinatorial Testing, QSIC' 2006. Beijing, 259-265.
- [43] Changhai Nie, Baowen Xu, Liang Shi, Guowei Dong, Automatic Test Generation for N-way Combinatorial Testing. Second International Workshop on Software Quality (SOQUA 2005), Lecture Notes in Computer Science 3712, 2005, 203-211



[44]Changhai Nie, Baowen Xu, Liang Shi, Ziyuan Wang. A new heuristic for test suite generation for pair-wise testing, seke' 2006, San Fransisco: 517-522

[45]Ziyuan Wang, Baowen Xu, Changhai Nie: Greedy Heuristic Algorithms to Generate Variable Strength Combinatorial Test Suite. QSIC 2008: 155-160

[46]Ziyuan Wang, Changhai Nie, Baowen Xu: Generating combinatorial test suite for interaction relationship. SOQUA 2007: 55-61

[47]Liang Shi, Changhai Nie, Baowen Xu: A Software Debugging Method Based on Pairwise Testing. International Conference on Computational Science (3) 2005: 1088-1091

[48]Lei Xu, Baowen Xu, Changhai Nie, Huowang Chen, Hongji Yang: A Browser Compatibility Testing Method Based on Combinatorial Testing. ICWE 2003: 310-313

[49] 聂长海,徐宝文,史亮.一种基于组合测试的软件故障诊断方法.东南大学学报(自然科学版) 2003/06

[50] 聂长海,徐宝文.软件配置测试的测试数据生成(英文) Journal of Southeast University(English Edition) 2004/01

[51] 王子元,徐宝文,聂长海.组合测试用例生成技术.计算机科学与探索,2008,(06): 571-589

[52] 周晓宇,聂长海,徐宝文,陈火旺.两两组合覆盖测试数据自动生成集成工具的设计与实现.计算机科学,2005,32 (01): 118-120 (8) : 130-135.

学会动态 •

第五届江苏省计算机网络与云计算新技术研讨会 在盐城成功召开

第五届江苏省计算机网络与云计算新技术研讨会于 2023 年 11 月 10 日 -12 日在盐城成功召开, 100 多位来自全省产学研各界专家学者参加了本次会议。本次会议由江苏省计算机学会, 以及学会云计算专委会和网络与分布计算专委会共同主办, 盐城工学院、盐南高新区管委会共同承办。会议就计算机网络与云计算领域相关理论与技术的最新研究进展和发展趋势开展深入、广泛的学术交流, 并邀请蒋昌俊院士等著名专家学者作大会报告, 为省内外相关领域的专家提供了一个碰撞思想的平台。



学会动态

江苏省大学生计算机设计大赛产学研赛创新高峰论坛暨十周年纪念大会成功举办

在江苏省教育厅的指导下，在江苏省高教学会的支持下，由江苏省大学生计算机设计大赛组委会和北京新大陆时代科技有限公司主办，江苏省计算机学会和南京大学共同承办的江苏省大学生计算机设计大赛产学研赛创新高峰论坛暨十周年纪念大会于 2023 年 11 月 4 日在南京隆重举行。国家总督学顾问、中国教育国际交流协会副会长、江苏省政协原副主席、南京大学原党委书记胡金波教授，南京大学本科生院院长王骏教授，中国大学生计算机设计大赛组委会秘书长李吉梅教授，江苏省计算机学会副理事长、南京大学计算机科学与技术系、人工智能学院、软件学院党委书记武港山教授，江苏计算机学会原秘书长南京大学杨献春教授，以及江苏省大学生计算机设计大赛（以下简称省赛）组委会委员南京大学陶先平教授和南京大学金莹教授、执委会主任南京师范大学吉根林教授、专家委主任扬州大学殷新春教授，江苏新大陆时代科技有限公司总经理张兵先生，淮阴工学院副校长严云洋教授、南京中医药大学副校长胡孔法教授，安徽省级赛组委会安徽大学杨勇教授，以及来自省内外 100 多位嘉宾和专家来到现场参加了这次大会。





简介

江苏省计算机学会理事单位

盐城工学院信息工程学院

盐城工学院是江苏省属全日制普通本科高校，2018年，获批“硕士学位授予单位”，2022年，获江苏省博士学位授予立项建设单位。盐城工学院信息工程学院成立于2008年，2021年学院与华为技术有限公司共建华为信息与网络技术学院，同年与盐南高新区共建西伏河人工智能学院、西伏河人工智能学院产业研究院。

学院现有信息与通信工程“十四五”江苏省重点学科，设有电子信息专业学位硕士授权点，涵盖人工智能、计算机技术、通信技术、新一代电子信息技术、自动控制技术等研究方向。本科专业设有人工智能、计算机科学与技术、软件工程、网络工程、电子信息工程、光电信息科学与工程。其中，电子信息工程专业入选国家级一流本科专业建设点，同时获批江苏省产教融合型品牌专业；计算机科学与技术专业获批江苏省一流本科专业建设点。

学院现有江苏省多源信息智能感知与处理关键技术省级科研平台、盐城市光纤传感及应用工程技术研究中心、光电信息技术研究所、软件应用技术研究所、多源信息智能处理技术研究所等多个研究机构，聚焦模式识别与智能信息处理、信息安全、信息系统架构与体系、传感器技术与应用等四个研究方向。近五年，主持国家自然科学基金、江苏省自然科学基金等纵向项目20余项，横向到账经费5000余万元，发表高水平论文100余篇，出版学术专著10余部，授权专利50项，成果转化产生经济效益500多万元，获省部级奖5项。

学院现有在编教职工84人，其中正高职称10人，副高职称30人，具有博士学位教师32人，硕士研究生导师18人，江苏省“333工程”培养对象2人；省高校“青蓝工程”中青年学术带头人、优秀青年骨干教师5人，优秀教学团队1个；省“双创计划”5人；省“六大人才高峰”高层次人才项目资助4人。

学院现有在读本科生2200余人，硕士研究生160余人。近五年，先后获得江苏省微课比赛一等奖、二等奖等多项，主持江苏省高等教育、教育教学指导委员会、教育部产学合作协同育人等省部级以上教改项目20余项，指导省级以上大学生实践创新训练计划项目20余项，荣获江苏省普通高校本科优秀毕业论文（设计）一等奖1项、二等奖2项等。学院学生在省级以上大学生学科竞赛中屡获佳绩，先后获得全国大学生数学建模竞赛国家一等奖、全国大学生智能汽车竞赛国家一等奖、中国大学生计算机设计大赛国家一等奖，累计获国家级奖项20余项、省级以上奖励50余项。

学院电子信息及计算机两大学科始终围绕长三角一体化发展国家战略、江苏省及盐城市“十四五”科技创新规划，瞄准电子信息、网络与通信、人工智能等新一代信息技术产业，以强化学科方向特色和提升人才培养质量为中心，全面提升学院学科建设水平和质量，推动跨领域的协同创新和多学科协同发展新格局，为区域电子信息、人工智能等产业发展培养高水平应用型工程技术人才。学院与盐城国家级高新技术产业开发区、华为技术有限公司、立铠精密科技（盐城）有限公司等政、企深度合作，构建产教融合长效机制，建立起政、校、企紧密协同的人才培养生态系统。紧密对接长三角及江苏地区信息技术产业，将人工智能、电子信息及计算机技术等相关知识融入课程体系，形成“能力导向、跨界融合”人才培养特色。充分发挥新四军铁军精神在专业建设中的作用，将铁军精神融入课程教学，培养“特别能吃苦，特别有韧劲，特别讲奉献”的高素质信息技术人才。