



2024年 第4期

# 江苏省计算机学会

COMMUNICATIONS OF THE JSCS



——迈向自主协同的网络化无人系统  
——在“稳健可信人工智能”的科研道路上砥砺前行

# 江苏省计算机学会常务理事单位

## 金陵科技学院计算机工程学院

金陵科技学院于 1997 年开展计算机科学与技术本科专业教学，致力于培养服务地方经济发展的高素质应用型工程技术专门人才。计算机工程学院现有教职工 79 人，其中教授 7 人、博士 23 人，拥有省级中青年突出贡献专家、省级优秀科技创新团队带头人、江苏省“333 工程”、“青蓝工程”培养对象、省级学术技术带头人、江苏省“科技副总”等优秀人才，以及省级、市级、校级科技创新团队、教学团队等。学院设置计算机科学与技术、数字媒体技术、数据科学与大数据技术等三个本科专业，其中计算机科学与技术专业为国家级一流专业，数字媒体技术专业、数据科学与大数据技术专业为教育部产学合作协同育人项目支持建设专业。

学院建有计算机科学与技术南京市重点学科，拥有数据科学与智慧软件江苏省重点实验室、江苏省智慧企业管理软件工程中心、南京市企业管理软件行业化工程技术研究中心、数据智能产教融合示范基地、大数据可视化协同研究中心等科研平台，以及专业实验室和计算机基础实验教学中心，总面积 3760 平米，总资产 3079 万元，能充分满足专业的实验教学需要。目前已与省内外 IT 品牌企业开展产学研合作，联合培养大数据、人工智能、嵌入式系统、游戏开发和 AR/VR 研发等领域专业人才。近年来，学院完成国家级自然科学基金、省重点研发计划、省自然科学基金等纵向项目 40 余项，承担产学研合作横向课题 70 余项，获得专利、软件著作权 200 余项，教师发表高水平 SCI/EI 论文 230 余篇。

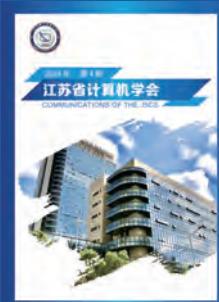
目前，学院全日制在校生 1400 余人。学院以学生为中心、注重学生工程系统实践能力培养和人文素质的综合提高，重视高等学历教育与职业素能教育相结合，支持学生在取得本科毕业文凭的同时还参加计算机相关职业岗位技能认证、行业认证或计算机应用技术软件资格与水平考试证书，为学生成长成才创造了良好环境。毕业生工程实践能力强，岗位适应性好，历年学生就业率均在 97% 以上，毕业生深受社会及用人单位的广泛好评。近年来，学生在省级及以上各类竞赛中有 430 余人次获奖，多次在全国大学生云计算创新大竞赛、嵌入式系统竞赛、计算机设计大赛、程序设计大赛、蓝桥杯全国软件和信息技术专业人才大赛、AR/VR 比赛、大数据人工智能竞赛、中国工程机器人大赛暨国际公开赛等学科竞赛中获得优异成绩。





# 江苏省计算机学会通讯

COMMUNICATIONS OF THE JSCS



## 顾问委员会

主任：周志华

副主任：武港山 耿新 刘昊

陈兵 李千目 李凡长

周勇 肖甫 李斌

吴小俊 毛启容

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

李畅

## 编委会

主编：路通

副主编：金莹 申富饶 聂长海

张洁

编委：徐大华 石克 吴春雪

严诚

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

## 学生记者团专栏

01 | 无人系统时敏智能感知 | 傅忱忱副教授专访

03 | 大模型时代下云原生计算系统研究展望 | 顾荣研究员专访

## 教学成果交流

05 | 面向工程能力培养的“4213”程序设计教学方法 | 皮德常

## 政策解读

11 | 关于开展“网络去 NAT”专项工作 进一步深化 IPv6 部署应用的通知解读

## 学术交流

13 | RFID 异常检测技术研究 | 陈星宇

21 | 基于变异测试的测试用例集有效性评价技术研究 | 张鹏

## 会员风采

30 | 迈向自主协同的网络化无人系统 | 朱琨

33 | 在“稳健可信人工智能”的科研道路上砥砺前行 | 郭兰哲

## 科学普及

36 | OpenAI 最强竞对官宣 大模型可以使用电脑了

39 | 为什么说光量子计算是面向未来的计算范式

43 | 研究实锤 别让大模型“想”太多 OpenAI 准确率竟下降 36.3% ?

## 科创成果

49 | 多维无源感知关键技术及应用

58 | 基于 AI 的社会安全态势感知平台



## 无人系统时敏智能感知 ——傅忱忱副教授专访

傅忱忱，东南大学计算机学院副教授、博导，获江苏省优秀青年基金、ACM 分会新星奖、国际会议 EMSOFT 最佳论文提名等。于 2017 年获得香港城市大学电脑科学系博士学位。2019 年起加入东南大学罗军舟教授团队，长期从事无人系统智能感知等领域的研究，累计在 TMC、JSAC、TC、TKDE、INFOCOM、RTSS 等国际著名学术期刊和会议上发表学术论文 40 余篇，其中 CCF A 类论文 26 篇。主持包括江苏省优秀青年基金、国家自然科学基金面上、军工及企业产学研等在内的项目十余项，研究成果在中国兵器北方信息控制研究院集团、西安地铁、江苏南钢等企业应用。

近年来，随着无人系统应用需求和规模的快速增长，如何实现高性能、低延迟、强自主的智能无人系统成为科研界和工业界广泛关注的问题。傅忱忱基于物联网“感 - 传 - 算”体系结构，提出了“无人系统时敏智能感知架构”，重点围绕时敏数据获取、时敏感知计算、时敏智能决策三大核心内容展开研究。她的研究旨在解决传统无人系统在实时性、自主性和效率上的不足，为无人系统在复杂环境中的应用提供全面的技术支持。

在具体实践中，傅忱忱联合中国兵器工业集团和中电科集团等国内顶尖科研机构，围绕智能无人系统的关键技术进行了多项应用验证工作。这些研究聚焦于精准定位、导航、三维建模等实际应用场景，成功构建了一套体系化的时敏智能感知无人系统原型。通过系统的优化和验证，该原型在数据采集、计算速度、智能决策等方面均表现出卓越的性能，为无人系统的实时性和自主性问题提供了创新性解决方案。傅忱忱的研究成果不仅在理论上拓展了无人系统的技术边界，同时也为工程化实现奠定了坚实的基础。未来，这一架构有望在无人驾驶、工业自动化、应急救援等多个领域发挥重要作用，展现出广阔的应用前景和社会价值。

**Q1：您的研究基于物联网的感 - 传 - 算体系结构，构建了无人系统时敏智能感知架构。您认为这一架构的核心创新点是什么？相较于传统的无人系统感知技术，它有何优势，是哪些因素带来的这些优势？**

傅忱忱：这一架构的核心创新点在于感 - 传 - 算一体化设计，突破了传统无人系统中感知、传输和计算各自为政的局限性，通过网络资源与计算资源的联合调度，提高了整体效率。相比传统技术，该架构的优势在于其系统化设计，从整体上优化了资源调度，以及任务依赖性优化，能在感知、计算、决策等链式任务中实现更高效的协作。这些优势得益于多资源协同调度和系统级的整体设计，既提升了系统性能，也增强了扩展性。

**Q2：智能无人系统强调高性能、低延迟、强自主，这对感 - 传 - 算体系的各个环节提出了很高要求。您认为当前实现这些目标面临的主要技术瓶颈是什么？未来的研究会在哪些方面取得突破？**

傅忱忱：当前面临的主要技术瓶颈包括：性能与延迟的权衡，不同场景对性能的需求存在差异（如无人机强调通讯延迟，自动驾驶更注重计算准确性和环境安全性），以及现有技术在安全性上的不足，距离政策要求的高标准安全性仍有差距。未来的研究突破点在于：提升通讯技术的稳定性与带宽以适应动态场景、改进算法的实时预测能

力以支持复杂环境中的导航与规划，以及通过增强系统的鲁棒性与快速响应能力来满足安全性需求。

**Q3：在无人系统的应用中，安全性和可靠性始终是重要议题。您提出的时敏智能感知架构是通过何种方式来提升系统的鲁棒性和安全性的？**

傅忱忱：该架构通过实时性优化（Real-Time）与延迟容忍性增强（Delay-Aware）来提升鲁棒性和安全性。一方面，通过快速响应和适应延迟变化，确保系统在动态环境中的稳定性；另一方面，结合传统算法、强化学习和神经网络等多种技术，优化系统在不同场景下的表现。此外，根据不同的应用需求，定制化部署不同的优化方案，从多方面提高系统的延迟容忍能力和实时反馈性能。

**Q4：您参与了多项产学研合作项目，这些项目涵盖军工、地铁、钢铁等领域。作为科研工作者，您如何看待这种合作模式对技术转化和应用的价值？您如何看待科研成果与产业化需求之间的平衡？**

傅忱忱：产学研合作模式在技术转化和应用中具有很大价值。高校更多专注于预研项目，突破技术瓶颈，而企业则注重落地应用，两者结合可以实现技术的迭代优化和实际部署。合作有助于科研聚焦产业的真实需求，避免方向偏离，并通过应用场景反馈，发现新的研究方向，推动技术发展与创新。在科研与产业化平衡方面，基础研究需保持前瞻性，但也要关注实际需求，通过不断的反馈和迭代，让科研成果更好地服务于产业和社会发展。

本文采编：学生记者团（郎宇（南京农业大学）、杨鹏（南京农业大学）

## 学会动态

### 第四届应用型高校计算机教育年会在江苏南通成功举办

2024年12月6—8日，第四届应用型高校计算机教育年会在江苏南通成功举办，本次大会由江苏省计算机学会应用型高校计算机学科建设专家委员会、信息技术新工科产学研联盟计算机工程与应用工作委员会主办，南通理工学院承办，南京理工大学紫金学院、用友集团新道科技股份有限公司、南京京秉蔚信息科技有限公司、江苏源本科技（集团）股份有限公司协办。来自国内50余家单位100余位代表参加了本次大会，本次大会主题为“数智赋能应用型高校高质量发展”，旨在深入学习习近平总书记关于教育的重要论述，全面贯彻全国教育工作会议精神，扎实推动新时期应用型高校计算机学科的改革与发展。

### 第四届应用型高校计算机教育年会（2024年12月）





# 大模型时代下云原生计算系统研究展望

## ——顾荣研究员专访

顾荣，南京大学计算机学院特聘研究员 / 博导，教育部青年长江学者、达摩院青橙奖获奖者（2023），江苏省计算机学会会员，中国计算机学会（CCF）分布式计算与系统专委会常务委员、大数据专家委员会执行委员、系统软件专委会执行委员及开源发展委员会执行委员。当前主要研究方向为云计算与AI系统，已发表 / 录用论文60余篇，包括USENIX ATC、EuroSys、SIGMOD、VLDB、ICDE、KDD、WWW、VLDBJ、TPDS、TKDE、TON等领域顶级会议和期刊。主持多项国家级和企业合作项目，包括国家自然科学基金面上项目、国家重点研发计划子课题、中国博士后基金特别资助项目，以及华为、阿里巴巴、腾讯、支付宝、字节跳动、中兴通讯、中国移动、中国石化和中国中车等企业的科研合作项目。获江苏省科学技术一等奖、IEEE TCSC Early Career Excellence Award、HPCC会议最佳论文奖、CCF分布式计算与系统专委会青年创新先锋、南京大学青年五四奖章，并获得阿里、华为、腾讯和中兴等多家企业的学术合作奖项。多项研究成果已应用于或贡献于知名开源项目。担任云原生计算基金会（CNCF）Fluid开源社区主席和Alluxio开源项目PMC Member。获评中国信通院OSCAR“尖峰开源人物”，入选“2023开源创新榜”优秀开源人物（由中科院和中国计算机学会等组织评选），并荣获中国开源大赛一等奖。

在10月26日南京大学费彝民楼举办的“2024年AI+开源论坛暨技术研讨会”中，顾荣老师作了题为《基于开源模式的云原生AI系统研究初探》的报告，在报告结束后，我们有幸采访到了顾荣老师。

**Q1：您深耕云计算与大数据系统多年，取得了众多优秀成果，了解到您最近关注分布式AI模型训练与推理系统，您能否介绍一下您在该领域的最新研究成果。**

顾荣：大模型通常具有高算力资源消耗和广泛的推理服务需求，这使得分布式AI模型的训练与推理系统成为我们当前研究的重点方向，也是国内外学术界和研发机构关注的前沿热点。我们专注于解决数据编排、任务调度以及推理机制优化等核心问题。首先，在训练任务的分配方面，我们重点研究调度算法的优化，使其能够根据任务的优先级和资源需求，智能地分配计算资源和节点，从而提升系统的整体资源利用率。其次，在推理优化方面，我们致力于提升推理执行效率，以增强大模型在实际大规模应用场景中的快速响应能力。

**Q2：目前AI和云原生技术都是比较火热的领域和话题，您如何看待云原生技术在推动AI发展中的作用？**

顾荣：近年来，随着云计算技术的迅猛发展，越来越多的企业和研究机构意识到，服务应用的规模化发展离不开云平台的支持。通过云上部署，不仅可以提升访问的广泛性和跨域访问的速度，还能显著简化运维流程。借助容器及其编排技术，系统能够实现灵活的弹性扩缩容，从而达到降本增效的目的。

随着AI技术的迅速普及，AI上云已经成为未来发展的重要趋势，“AI+云原生”因此吸引了众多企业和学术界的高度重视。云原生技术的引入，旨在推动AI系统的更广泛应用，特别是在推理阶段的优化与应用方面。云原生架构在促进AI应用的规模化和商业化落地方面展现出巨大潜力。深入研究这一领域，将为AI技术的普及和应用提供更加坚实的基础支持。

**Q3：您认为当前云计算和大数据领域面临的最大挑战是什么，未来发展趋势如何？**

顾荣：在过去十年间，大数据和云计算技术实现了飞速发展，逐渐成为推动各行业创新和数字化转型的核心动力。以 Hadoop 和 Spark 为代表的大数据解决方案已趋于成熟，并被广泛应用于各类大数据场景。

进入新的发展阶段，技术重心逐渐转向支持大规模人工智能（AI）训练和推理的能力。如何进一步优化大数据架构，提升计算效率，并实现大数据与 AI 模型的深度融合，成为当前技术发展的重要方向。然而，这一过程中仍面临诸多挑战和技术瓶颈。例如，如何有效迭代优化现有系统架构，实现大数据与 AI 框架的无缝衔接，是该领域亟需解决的关键问题之一。

**Q4：将技术转换为生产力一直是科研发展的目标，在学术界和工业界之间，我们的研究和应用应该如何平衡？**

顾荣：学术研究与实际应用场景的问题解决并非相互独立，而是相辅相成、相互促进的关系。因此，建议大家在选择科研课题时，综合考虑二者的平衡与优化。在科研过程中，一方面需要从实际应用场景中抽象出具有学术价值的科学问题，建立系统性和深度的科研方法论，并产出高质量的学术论文；另一方面，也应注重原型系统的开发，通过实践验证理论成果，从而推动理论与实践的有机结合，实现科研工作的双重价值。

**Q5：您作为云计算与大数据领域有一定积累的研究者，对刚进入该领域发展的年轻研究者，你有何建议？他们在科研和职业生活中应该如何关注这些重点问题？**

顾荣：大数据和云计算系统作为基础设施级系统软件，为众多行业应用提供了强有力的支撑。建议大家在进入这一领域时，从具体行业应用入手，深入挖掘和分析大数据与云计算系统中的不足，尤其是在计算、存储和调度等核心环节。通过对这些问题的深入剖析，往往能够提炼出一些具有共性的技术挑战，这不仅有助于推动系统优化，也为科研工作带来了宝贵的创新机会。

本文采编：学生记者团（王益镔（东南大学）、郎宇（南京农业大学）、杨鹏（南京农业大学）

## 学会动态

### 2024 年江苏省人工智能学术会议在宿迁召开

由江苏省计算机学会、江苏省信息技术应用学会、江苏省人工智能学会联合主办，宿迁学院、宿迁市工业和信息化局、宿迁市科学技术协会联合承办，江苏省计算机学会人工智能专委会、江苏省信息技术应用学会人工智能专委会协办的 2024 年江苏省人工智能学术会议于 2024 年 11 月 15 日至 17 日在宿迁召开。





# 面向工程能力培养的“4213”程序设计教学方法

## ——南京航空航天大学皮德常教授

### 简介

“面向对象 C++ 程序设计”是面向南京航空航天大学工科专业大一学生所开设的一门计算机编程基础课。

针对大一新生建模能力较弱、算法思维不足和项目实践缺乏的问题，皮德常结合自己 30 年的教育教学经验和丰富的工程项目科研经历，提出了以学生学习为中心，面向工程能力培养的科教融汇课程育人的教学理念。在教学改革过程中，编写了 14 本教材，建设了国家一流课程，研发了在线自动程序评测 OJ 软件平台，提出了一套“从理论到实践”四环节推进基础教学、“从程序到工程”二阶段深化教学、“从课内到课外”延伸拓展教学、“主观客观相结合”三维考核的教学创新方法。

创新模式自实施以来，学生编程能力得到了显著提升，课程团队指导学生获“中国软件杯”一等奖、程序设计竞赛金奖等。创新成果获国家精品在线开放课程、国家级一流本科课程、国家级规划教材、省精品课程、省精品教材、省高校重点教材等，2023 年获国家教学成果二等奖，2024 年获全国高校教师教学创新大赛一等奖等。

### 主要解决的教学问题：

课程以 C++ 编程思想为主线，以培养初学者的程序设计能力为目标，理论与实践结合，循序渐进地学习 C++ 程序设计方法，注重分析问题和编程解决工程问题能力的培养。

#### 2.1 成果主要解决的教学问题

##### (1) 建模能力较弱

调查发现，约 95% 的学生在进入大学之前没有接触过 C++ 编程，绝大多数学生在初学编程时难以理解代码之间的逻辑，感觉很抽象，不会建模，逻辑思维较弱。

##### (2) 算法思维不足

算法是采用计算机解决问题的方法。大一新生对算法很迷茫。有的学生采用背英语课文的方式背 C++ 算法的代码。对于存在错误的代码，不知如何修改，这是学生缺乏算法思维的表现。对于大一新生，这是一个比较普遍性的问题。

##### (3) 项目实践缺乏

C++ 程序设计是一门理论和实践紧密结合的课程，但学生在学习时往往偏重理论，忽视知识的灵活应用，尤其对综合性的工程项目无从下手。学生的编程实践能力与国家对拔尖创新人才的需求，还存在着较大差距。

#### 2.2 解决教学问题的方法

皮德常在教学内容、教学内涵与外延、教学评价等方面提出了四大创新方法，将其贯穿程序设计课程教学，达到全流程培养与全过程评价的创新。

### (1) “从理论到实践”四环节推进基础教学

传统程序设计教学中的编程案例过于抽象、距离工程实践较远，无法满足编程思维与工程思维等高阶思维培养的需求，且难以引起学生的兴趣。

为克服传统上述缺陷，本课程针对课堂教学环节提出“从理论到实践”四环节推进基础教学。以皮德常完成的工程科研项目为主线，贯穿课堂单元教学内容，实现四个思维的培养，引导学生利用所学知识尝试解决实际工程问题。

从相关工程科研项目中提取出典型的工程项目场景，并从相关的软件系统中提炼出适用于不同教学阶段的工程案例。每个教学单元以1~2个工程案例为指引，采用如图1所示的四环节教学模式组织教学，提升对逻辑、算法、编程和工程四个思维的培养。



图1 “从理论到实践”四环节推进基础教学



图2 工程案例与教学单元对应关系



构建的工程案例包括航天器遥测数据预测系统、卫星故障诊断系统、机载数据管理系统、卫星电源多参数预测系统、可靠性试验精确分析系统和太阳翼展开可靠性评估系统等。

每个教学单元从工程项目中存在的真实问题情境出发，引导学生结合单元知识，在一段时间内持续开展学习和探究，尝试创造性地解决问题，实现基于项目案例的“做中学”，工程案例与课程知识单元对应关系如图 2 所示。

该学习方式不仅提供了案例支持，而且还结合了科研成果的技术前沿，将工程问题“真实的力量”引入课堂，激发了学生的学习兴趣。使学生不知不觉地开阔了思路、提高了学习兴趣，明白了所学知识与国家需求密切相关。

在完成每个教学单元的学习之后，引导学生回归工程实践，将所学知识应用于软件开发，并在工程情景下领悟工程内的设计取舍与质量评价，完成基于工程思维的反思与改进。同时，为了在有限的教学时间内完成上述四个环节的教学，本课程同时引入了线上线下混合教学模式。

学生在课前通过线上自主学习相关教学单元知识点，从而缩减课堂教学用时，实现课堂翻转，把课堂时间用于培养编程思维与工程思维，深度促进线上线下融合学习。



图 3 线上线下混合教学资源

## (2) “从程序到工程”二阶段深化教学

传统的程序设计教学过于关注单独的知识点讲解，学生经过课程学习后能够编写一个个独立的小程序，但是难以掌握如何使多组程序模块组合起来，构成一个具有一定复杂度的项目，即编程缺乏高阶性。针对此问题，我们将教学内容组织为基础教学与项目设计两个阶段，如图 4 所示。

基础教学属于课程的第一个阶段，该阶段采用四环节教学模式，完成程序设计基础知识点的讲解与训练，本阶段侧重培养逻辑思维、算法思维和编程思维。

在课程第二个阶段，将与教学同步开展项目设计与实现，每位同学需要完成一个较大型的编程任务。项目设计旨在通过程序设计方法，培养学生针对较大型复杂工程问题的设计与编程实现能力，本阶段侧重培养工程思维。项目设计同时注重思政导向，紧密结合学科和社会热点，引导学生树立正确的人生观、社会观和价值观。

项目设计具有研究性、创新性和综合性，加大了学生学习的投入，让学生体验到“跳一跳才能够得着”的挑战。每年都有部分学生的项目设计经完善后，会直接演变成后续的学科竞赛项目，为学生创新能力的培养进一步助力。



图 4 两阶段教学组织模式

### (3) “从课内到课外”延伸拓展教学

课程组引导 / 指导学生参与有关学科竞赛，实施从课内到课外延伸拓展，进一步培养和提高学生的工程思维和创新能力。鼓励学生在学期间，主持或参与至少一个学科竞赛项目或工程项目。

我们以中国大学生程序设计竞赛、计算机设计大赛、中国软件杯、ACM/ICPC 编程竞赛为指引，引导学生参与竞赛。学生可以自主选择参加由导师主持的工程项目，或者是由学长领队的学科比赛，通过“传帮带”培养学生的创新能力，近几年指导学生获奖 12 项。

课程团队在现有课内教学内容的基础上，通过设计延伸教学内容，与程序设计学科竞赛联动，将程序设计有关的学科竞赛，作为培养学生创新能力和综合素质的重要抓手之一。

通过高强度、高挑战度的学科竞赛或工程项目，不但培养了学生的逻辑思维、算法思维、实践创新及团结协作能力，同时也夯实学生的编程思维和工程思维，提升了人才培养质量。以赛促学、赛教融合、科教融汇的方法，不但促进了学生的全面发展，而且也提高了教师的教学水平。

### (4) “主观客观相结合”三维考核教学

传统书面作业与卷面课程考试机制，存在实践性弱、主观性强两大弊端，对学生算法思维与编程思维的考察存在不足。为了考核学生真正的编程能力，我们在全校率先取消了纸质的卷面考试，提出了上机作业评分、期中期末机考与项目设计三维考核机制，如图 5 所示。

为了客观高效地评价代码，课程团队带领优秀学生研发了在线自动程序评测 OJ 平台作为过程性考核平台。从 OJ 平台导出的学习大数据，通过聚类算法实现对课程知识点的聚类分析，发现学生学习的薄弱环节，进行个性化的学习推荐，进而实现因材施教。



图 5 三维考核机制

## 创新点

本成果的创新点体现为两个方面：

### 教学理念创新

针对教学中的问题，皮德常总结 30 年的教育教学经验和科研背景，把工程科研项目融入课程，提出了面向工程能力培养的科教融汇课程育人的教学理念。

皮德常主持完成的 40 多个科研项目为科教融汇奠定了坚实基础。项目中的半数为前沿探索类项目，如国家自然科学基金、国家 863、国家重点研发计划等，其余的 20 多个项目是与研究所合作完成的工程软件项目。研究成果用于航空、航天等不同的工程装备，先后获得 5 项省部级科技进步奖。

总结工程软件项目经验，结合 C++ 编程特点，皮德常提炼出程序设计教学中最核心的培养要素，即四个思维能力的培养。按照培养难度依次为：逻辑思维、算法思维、编程思维和工程思维，皮德常又进一步提出了面向工程能力培养的科教融汇课程育人的教学理念。

在教学过程中，以四个思维的培养贯穿课程教学，以实际工程项目求解为问题背景，结合课堂单元教学内容，设计工程案例，激发学生学习兴趣，按照课程由浅入深的顺序，递进培养四个思维，使学生应对未来工作中的科学挑战和工程挑战。

### 教学方法创新

为了落实提出的教学理念，解决教学中存在的问题，皮德常在教学内容、教学内涵与外延、教学评价等方面提出了四大创新方法，即“从理论到实践”四环节推进基础教学、“从程序到工程”二阶段深化教学、“从课内到课外”延伸拓展教学、“主观客观相结合”三维考核的教学创新方法，并将它们贯穿 C++ 程序设计课程教学的始终，达到全流程培养与全过程评价的创新。

## 成果的推广应用效果

### 应用成效

(1) 学生学习成绩与创新能力显著提升。教学创新方法实施以后，学生已具备解决复杂工程问题的初步能力，夯实了专业基础。以皮德常课程班 23 届的学生为例，学生大创项目的主持参与全覆盖，获得省部级以上学科 / 创新创业竞赛人均 4 次。

(2) 教学建设成效显著。皮德常与教学团队同成长，先后获得国家级课程、教材和荣誉称号等奖励 5 项，省级奖励 10 项。

(3) 教学方法赢得专家和学生肯定。皮德常近年的授课评教结果全为优秀，一直稳居全校全院评教结果的前列，教学深受学生好评。

### 推广成效

(1) 皮德常主讲课程在国内影响大。皮德常负责的国家精品在线开放课程“面向对象 C++ 程序设计”提供了丰富的教学资源，全部内容已在国家高等教育平台、中国大学 MOOC、超星名师讲坛等平台开放，有 70 多万人次学习，选课学校有 150 多所，课程热播在超星平台位居工学领域第一，具有较大的社会影响力。

(2) 出版教材院校选用广。皮德常及时总结教学经验，编著教材，先后有 10 多个版本，曾获省高校重点教材、省精品教材、国家级规划教材等称号，国内 100 多所院校选作授课教材。

(3) 传播教学经验积极性高。皮德常曾在全国高等院校教育与教学研讨会、计算机教指委主办的计算机核心课程教师培训班、高等学校计算机程序设计课程研讨会等做过 20 多次的教学经验交流报告。

## 学会动态

### 第六届江苏省计算机网络与云计算学术会议在张家港举办

第六届江苏省计算机网络与云计算学术会议于 2024 年 11 月 8-10 日在张家港中油泰富国际酒店成功召开。本次会议由江苏省计算机学会、江苏省计算机学会网络与分布计算专委会、江苏省计算机学会云计算专委会共同主办。会议邀请了贾小华教授、李宣东教授、张燕咏、朱燕民教授等著名专家学者作了大会特邀报告。为省内外相关领域专家提供了一个碰撞思想的平台。





# 《关于开展“网络去 NAT”专项工作 进一步深化 IPv6 部署应用的通知》解读

来源：信息通信发展司

近日，工业和信息化部办公厅、中央网信办秘书局联合印发《关于开展“网络去 NAT”专项工作 进一步深化 IPv6 部署应用的通知》（以下简称《“网络去 NAT”通知》）。为推动《“网络去 NAT”通知》落地实施，回应社会关切，现将有关内容解读如下：

## 一、出台《“网络去 NAT”通知》目的是什么？

党中央、国务院高度重视 IPv6 规模部署和应用发展，自 2017 年中共中央办公厅、国务院办公厅印发《推进互联网协议第六版（IPv6）规模部署行动计划》以来，中央网信办、工业和信息化部认真贯彻落实党中央、国务院决策部署，组织产业各方大力推动 IPv6 部署和应用，取得积极成效，我国网络基础设施和应用基础设施 IPv6 服务能力已全面具备。截至 2024 年 5 月，我国 IPv6 活跃用户达到 7.94 亿，各类网络中已分配 IPv6 地址的终端数达到 17.65 亿，移动网络 IPv6 流量占比达到 64.56%。

同时也要看到，当前我国 IPv6 规模部署和应用仍存在堵点和瓶颈，如 NAT44 设备的大规模部署，降低了企业开发 IPv6 设备和应用的主观意愿，增加了用户使用 IPv4 网络的粘性，客观上阻碍了 IPv6 规模部署和应用水平的进一步提升。为贯彻落实《推进互联网协议第六版（IPv6）规模部署行动计划》，2023 年 4 月我部会同中央网信办等八部门联合印发《关于推进 IPv6 技术演进和应用创新发展的实施意见》，明确提出实施“网络去 NAT”专项行动。近期，我们组织专家深入调研，围绕“网络去 NAT”的目标、路径、方案等问题，编制形成《“网络去 NAT”通知》，旨在向全社会释放我国加快向 IPv6 演进升级的明确信号，凝聚网络、应用、终端等产业各方合力，加速提升 IPv6 规模部署和应用水平。

## 二、《“网络去 NAT”通知》的工作思路是什么？

在《“网络去 NAT”通知》推进过程中要统筹做好“三个协调”：一是协调全局与局部，推动基础电信企业制定整体工作方案和时间表，同时选取部分区域开展试点，探索经验，提前达标。二是协调网络与应用，一方面基础电信企业持续拓宽 IPv6 网络通路，另一方面互联网企业深化应用改造，加大放量引流力度，“建”“用”双方协调互促，发挥乘数效应。三是协调改造与安全，合理有序推进 NAT44 设备使用规模逐步降低，同时要求做好网络安全监测和应急处置，保障网络安全稳定运行。

## 三、《“网络去 NAT”通知》中提出了哪些工作任务？

《“网络去 NAT”通知》部署了五方面工作任务：

一是细化工作方案，有序实现网络升级。基础电信企业要认真摸排 NAT44 设备部署应用情况并建立信息台账，制定“网络去 NAT”工作方案和时间表，有序推进全网 NAT44 设备使用规模降低。工业和信息化部、中央网信办组织开展试点，在试点区域提前完成任务。

二是紧抓关键环节，持续拓宽 IPv6 通路。基础电信企业新增互联网专线默认开通 IPv6 功能，加快实施家庭网关 IPv6 地址前缀二次分发功能升级。到 2024 年底，基础电信企业自有环境固定宽带用户 IPv6 连通率不低于 80%。推动新增和存量家庭无线路由器提升 IPv6 支持率。

三是深化应用改造，主动引导流量迁移。互联网企业要深化应用服务升级改造，优化放量引流策略，提升固网环境下 IPv6 流量占比。内容分发网络运营企业要提升用户加速、业务调度和内容回源等 IPv6 流量占比。云服务企业向用户提供服务时默认启用 IPv6 功能。

四是强化运行维护，确保网络安全稳定。基础电信企业、互联网企业要强化网络日常运行维护管理和重要指标监测。要强化网络安全防护管理，做好预警监测和应急处置。要制定预案，保障网络安全稳定运行。

五是加强督促评测，促进工作实效落地。基础电信企业对开展试点工作的子（分）公司加强倾斜与保障。构建监测体系，开展重点指标统计，工业和信息化部、中央网信办适时公布监测结果。各省（区、市）通信管理局、工信主管部门加强工作指导，各省级网信办督促互联网企业加强协同。

## 学会动态

### 第一届江苏省计算机教育大会职业教育改革高峰论坛成功举办

2024 年 12 月 1 日，第一届江苏省计算机教育大会职业教育改革高峰论坛——暨 AI 赋能职业教育应用论坛在南京顺利举行。本次论坛由江苏省计算机学会网络空间安全职业教育专业委员会和江苏省计算机学会职业院校产教融合专业委员会联合主办，旨在推动职业教育改革与创新，促进人工智能等新一代信息技术与职业教育的深度融合发展。





# RFID 异常检测技术研究

## ——2024 年江苏省计算机学会优秀博士论文奖

作者：陈星宇

单位：南京大学

指导老师：陈力军

### 论文摘要

继计算机、互联网之后，物联网引发世界信息产业的第三次浪潮，射频识别技术 (Radio Frequency Identification, RFID) 作为物联网应用的关键支撑技术之一，因其远距离识别、非视距和无源通讯的特点，正逐渐替代传统识别手段，渗透到供应链、零售、医疗等多个领域，大幅提升了生产效率以及改善了用户体验。随着 RFID 研究不断深入，RFID 系统规模不断扩大，复杂密集型 RFID 场景日渐普及，这为物联网应用的安全性带来严峻挑战。因此，研究真实场景下的 RFID 异常检测技术，对于提升系统可靠性与服务质量具有重要意义。然而，现有异常检测工作在鲁棒性、时效性、以及抗噪性方面存在一些尚待解决的关键挑战。首先，传统异常检测通常利用射频信号作为感知媒介，然而射频信号极易遭受环境影响，导致系统缺乏鲁棒性；其次，大规模 RFID 系统中，少量的异常数据往往淹没于海量的正常数据之中，如何最小化正常数据的干扰从而快速收集检测异常数据亟待解决；最后，真实系统的复杂性导致 RFID 射频信号存在较大噪声，如何提升微弱信号的信噪比 (SNR) 从而降低异常感知的误差也是一项关键挑战。

为解决上述问题，本文致力于真实场景下的 RFID 异常检测技术研究，重点围绕身份异常、数据异常以及行为异常三个角度进行展开。为更好地适用于真实场景，一方面，本文兼容广泛普及的 RFID 国际标准，在不进行任何硬件与协议修改的前提下实现 RFID 异常检测；另一方面，为应对复杂大规模场景，研究鲁棒、高效、精准的异常检测技术。本文的主要研究内容与贡献总结如下：

1) 针对身份异常检测的鲁棒性问题，本文提出了一种全新基于能量的身份认证异常检测技术。区别于传统射频信号指纹，该技术将标签电路中存储的电能所反映出的 RC 电路差异作为指纹进行身份认证，大幅提升身份异常检测的鲁棒性。同时，本文设计了一种与商用 RFID 系统及其现行国际标准兼容的能量指纹测量方法，保障身份异常检测在真实场景下的适用性。

2) 针对数据异常检测的时效性问题，本文提出了一种兼容 RFID 国际标准的高效范围检测协议，以解决密集型 RFID 场景中数据异常检测的通信开销和实时性问题。本文设计了选择查询 (SQ) 和与区间查询 (Q) 两种协议，利用 RFID 国际标准规定的选择 (select) 和查询 (query) 以及数据编码可以快速静默非目标标签，从而快速筛选出包含异

常数据的标签，将异常数据检测的时效性提升一个数量级。

3) 针对行为异常检测的抗噪性问题，本文提出了一种自适应随机共振方案(ASR)，可在各种普适计算场景中自适应提升信噪比，增强弱信号。通过结合频谱分析与顺序二次规划技术，ASR能够快速收敛到近乎最佳的系统参数，从而输出低噪声、高质量检测数据，为提升诸如机械故障诊断、呼吸异常监测等异常行为检测的性能提供有力保障。

## 专家推荐语

陈星宇博士具有扎实的专业基础、开阔的研究视野以及突出的学术能力。近年来围绕物联网射频异常检测技术开展了深入研究，提出身份异常检测、数据异常检测、行为异常检测三位一体的理论体系，取得了多项创新性研究成果。在 USENIX NSDI, ACM MOBICOM, ACM MOBISYS, IEEE/ACM TNET 等国际重要会议与期刊上发表学术论文 19 篇，其中 CCF A 类论文近 11 篇；研究成果受到 10 多位 IEEE Fellow, ACM Fellow 以及外籍科学院 / 工程院院士的关注与积极评价。申请获得并主持江苏省自然科学基金青年项目。申请 / 授权中国发明专利 5 项，美国专利 1 项。曾荣获 2024 年国家资助博士后研究人员计划，2024 年江苏省卓越博后计划，2021 年南京大学优秀博士研究生创新能力提升计划 A 等荣誉。

陈星宇博士申请人注重成果转化与落地应用，参与研发 RFID 移动定位系统，解决了大型仓储环境下物资定位困难的问题。荣获第 46 届日内瓦国际发明展特别金奖、第 22 届中国国际工业博览会高校展区—特等奖、世界机器人大赛 - 共融机器人挑战赛二等奖等荣誉。系统已在 20 省、直辖市广泛应用，涉及行业包括军需物资、电力物资、图书档案、仓储物流，具有较高的科研价值与显著的社会效益。

综上，陈星宇博士理论基础扎实、系统经验丰富，是计算机领域具有突出研究成果的青年科技创新人才，郑重推荐其申请 2024 年江苏省计算机学会优秀博士论文奖。

## 论文看点

### 1 研究内容

物联网异常检测是网络安全的重要一环，关乎国家政治、经济、社会安全。本文深入探讨并解决实际物联网系统中普遍存在的身份异常、数据异常和行为异常等核心问题，为提升物联网系统的安全性、可靠性和时效性提供保障，具体而言，针对身份异常，着重研究鲁棒性强的身份异常检测技术，以应对标签伪造或克隆等安全风险，并特别关注方案在真实系统中的结果一致性，确保检测结果的准确无误；针对数据异常，开发高效的数据异常检测技术，以在海量数据中迅速而精确地识别异常数据，减少高延迟带来的经济损失和检测失效；针对行为异常，探索具备抗噪能力的行为异常检测技术，以期在复杂的物联网环境中自适应参数调整，改善目标对象的数据质量，提高物联网系统整体的可靠性。本文的研究内容能够显著提升物联网系统异常检测的整体性能，为智能物联网在实际应用中的落地保驾护航。

#### (1) 鲁棒的身份异常检测

身份异常检测指的是判断系统中是否存在被伪造的异常标签。传统身份异常检测研究采用标签的信号物理层指纹特征来识别真实性，导致相关应用在面对环境变化时可靠性不足。本文深入探索了环境无关的指纹信息用于身份异常检测。特别的，指纹需要达到较高的检测精度和环境鲁棒性，满足真实场景下 RFID 身份异常检测的能力。在



此基础上，找到一种标准兼容的指纹测量方法，避免增加额外的硬件成本，促进身份异常检测的落地与实现。

### (2) 高效的数据异常检测

数据异常检测指的是判断系统中是否存在存有离群值的标签。传统数据异常检测研究采用提升并发度的方式来改善数据异常检测的时间效率。这种方式虽然可以极大提升 RFID 系统的检测效率，但是却与现行 RFID 通讯标准不兼容，难以在开放系统中部署实现。本文提出一种非均衡采样机制，通过减少收集正常标签的时间来缩减收集全部数据集合所需的总时间。特别的，非均衡采样协议需要考虑通讯标准问题，从而在真实系统中部署实现。

### (3) 可靠的行为异常检测

行为异常检测指的是判断系统中 RFID 标签微小的微动数据。目前传统研究假设噪声和微弱信号具有不同的频率，通过滤波器的方式将一定频率的微动信号提取出来。但是噪声去除的方式一定程度上也会伤害已经非常微弱的微动信号，进一步降低 RFID 系统的异常检测能力。为克服这一不足，本文基于随机共振模型，提出一种噪声抑制机理，避免在提取微动信号时削弱目标信号。在此基础上，本文研究如何打破传统方法已知信号频率的假设，拓展检测的使用场景。

本文以国家网络安全战略部署为背景，以物联网全面发展驱动我国数字化经济转型为目标，针对当下物联网系统存在的数据易泄露、标签规模大、环境噪声强等问题，提出“可靠检测，高效检测，增强检测”的技术路线，形成身份异常检测、数据异常检测，行为异常检测三位一体的 RFID 异常检测理论体系与系统原型，为保障物联网服务应用在真实系统中的安全性与稳定性提供理论依据与技术支持。

## 2 文章解决的科学问题

### (1) 如何解决环境依赖问题，实现鲁棒的身份异常识别。

针对真实场景下环境复杂变化，寻找一种环境无关的指纹用以区分真实标签和克隆标签是研究需要解决的科学问题之一

### (2) 如何解决低效传输问题，实现高效的异常数据收集

真实场景下系统规模大，如何设计一种标准兼容的方式来提高海量标签场景下异常数据传输效率是本项目的一科学问题。

### (3) 如何解决信号微弱问题，实现可靠的行为异常检测

如何克服真实场景的动态变化，设计一套自适应方式去过滤动态噪声，以达到最优或接近最优的行为异常检测效果是本项目的科学问题之一。

## 3 研究方案

### (1) 基于鲁棒性能量指纹的身份异常检测方式

#### ■ 基于物理层信号特征的标签数据聚类算法

验证标签身份是否异常需要标签的特征唯一且不会改变，因此环境因素通常对标签异常检测存在负面影响。为克服该局限性，本文利用标签所处的环境差异性和标签自身差异性导致的信号不同来统计对应的标签个数，从而发现系统中是否存在克隆标签。在 RFID 系统中，标签的物理层信号特征包括信号强度（Received Signal Strength Indication, RSSI），相位（Phase），多普勒频移等。其中信号强度和相位被绝大部分射频设备支持，因而是用于标签异常检测的优选特征。RFID 标签的物理层信号特征综合反映了标签位置和自身属性的差异，因而可以被用于

发现系统中存在的异常标签。

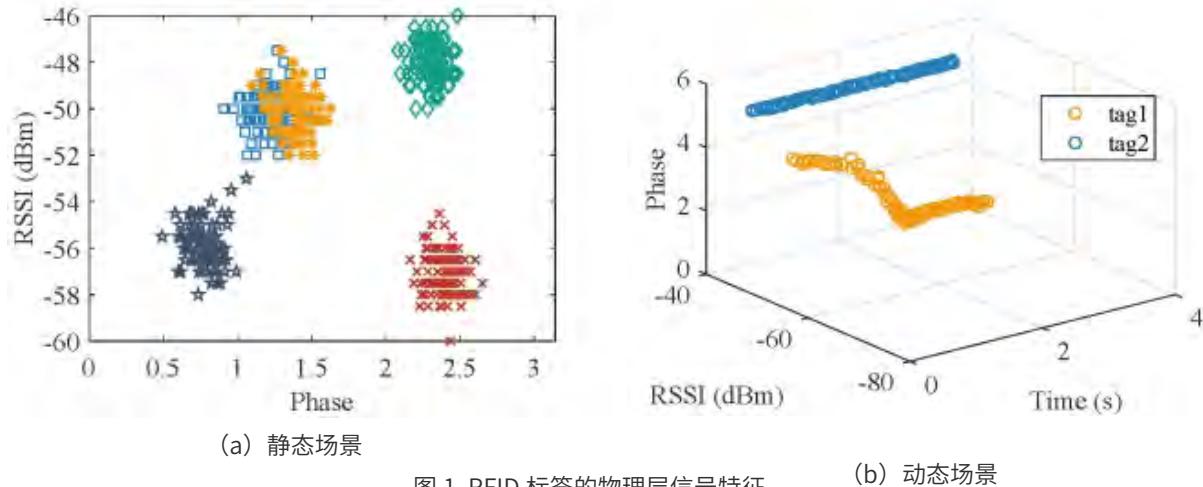


图 1. RFID 标签的物理层信号特征

研究提出一种基于聚类的机制，通过对标签物理层信号的 RSSI-Phase 特征二元组进行聚类，从而判断系统中是否存在异常标签。图 2 (a) 展示了采集于静止标签的 RSSI-Phase 二维图像。在 RSSI-Phase 坐标空间内来自一个标签的数据在平面上围绕一个中心随机扰动。通过基于密度的聚类算法可以得到类中心的个数，从而判断系统中是否存在异常标签。在开放场景下，系统内的标签具有很强的移动性，采用密度聚类可能导致一个运动标签的数据被判断为多个不同的聚类中心。针对这一问题，研究在动态场景中进一步考虑采集时间，构建 RSSI-Phase-Time 三维空间对标签数据进行聚类。如图 2 (b) 所示，一个标签的物理层数据在 RSSI-Phase-Time 三维空间中构成一个数据链。通过基于传递的聚类算法，可以将一个标签的数据归为一类。此外，在开放场景下一个运动标签的数据可能因为遮挡和噪声断为多个收尾相连的子链。研究将时域上将首尾相连的数据进行合并。合并后的每一条子链代表一个运动标签，从而实现所有情况下的标签异常检测。

#### ■ 基于标签电量指纹的异常标签认证方式

基于标签的物理层信号特征可以帮助判断标签的具体个数，从而发现异常标签。但是该方法无法判断出多个数据相同的标签中哪个是真实标签。受标签位置和信道环境的影响，一个真实标签的物理层信号数据也会改变，从而极大概率被判断为伪造的异常标签。针对这一问题，本文首次提出使用标签的电量指纹特征来判断标签的真实性。研究发现被动标签没有任何内置电源，完全通过阅读器发出的射频信号获取运转所需的能量。服从这一设计，标签电路可以等效于一个电阻电容 (RC) 串联电路。因为标签制造的差异性，异常标签和真实标签的电阻和电容存在差异性，可以被用于区分不同标签。相较于标签信号，标签的电容电阻不受标签环境和运动状态的影响，具有更高的鲁棒性和可靠性。

然而检测标签的电容电阻通常需要使用专用检测设备（如欧姆表，电压表等）。如果标签的电容电阻特征发生改变，则检测到一个异常标签。然而，基于专用检测设备的测量方式需要将标签芯片从电路中取出，这会造成标签永久性的损坏。同时，检测设备成本高昂，难以在开放大规模场景下广泛使用。针对这一问题，本文提出一种间接的测量法，基于当前的 RFID 通讯标准来测量标签的电路差异。本文使用标签的持续时间作为特征来构建唯一的



标签指纹特征。标签的持续时间指的是标签失去电能供应以后维持功能的最大时间。持续时间受电容电阻影响，因而存在标签差异性。本文测量持续时间的基本思路如图 3 所示。第一步中，阅读器释放电磁波激活标签的电路，并对标签电容充电。在第二步中，关闭阅读器（停止供电），并等待时间  $T$ 。第三步，重新打开阅读器，查看标签的功能（轮询控制，存储数据等）是否失效。通过不断增加等待时间  $T$ ，最终可以测出标签维持功能的最大时间  $T_{max}$ 。通过  $T_{max}$  可以间接测量一个标签的电容电阻，从而判断一个标签是否真实。如果一个标签的  $T_{max}$  与数据库里的记录不符，就可以将其判断为一个异常标签。不同于传统标签指纹，该标签指纹被当前 RFID 协议支持，无需专用检测设备，具有更为优异的实用性。

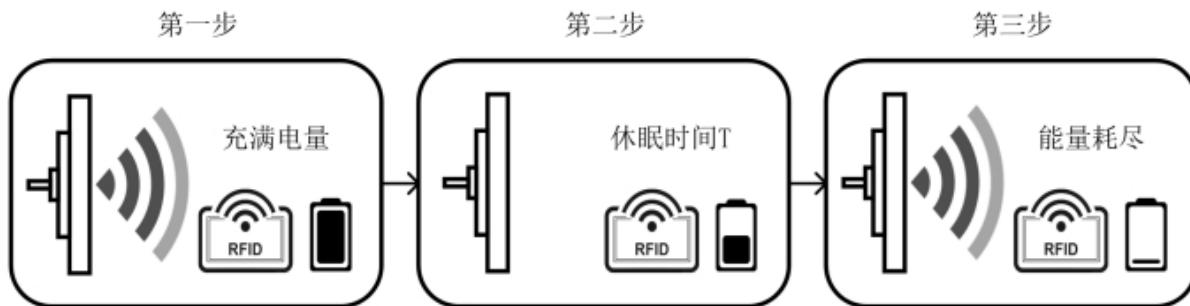


图 2. RFID 电量指纹获取过程

### (2) 基于非均衡采样的数据异常检测方法

#### ■ 基于基数估计的数据异常检测方法

大规模系统中异常数据的个数是评价系统安全性的重要指标。如果异常数据的总数大于阈值，就需要及时排查相关物品，从而提升系统的安全性。目前统计异常数据的个数需要对每个标签进行重复性数据收集（即便其数据正常），造成极大的时间浪费。针对这一问题，本文采用概率模型来粗粒度地估计异常数据的总数，从而较快检查出异常数据的大致数目范围。一种常见的概率模式是标记重捕法。通过标记几个存储异常数据的标签，然后再次采集。通过查看标记的标签在二次采样中所占的比例，就可以大致可以估计出包含异常数据的标签数目。但是在真实系统中通常难以提前知道异常标签的总数，难以确定标记重捕法中标记标签的最优个数。如果标签的基数很大而标记的标签过少，阅读器将无法短时间找到已经标记的标签。

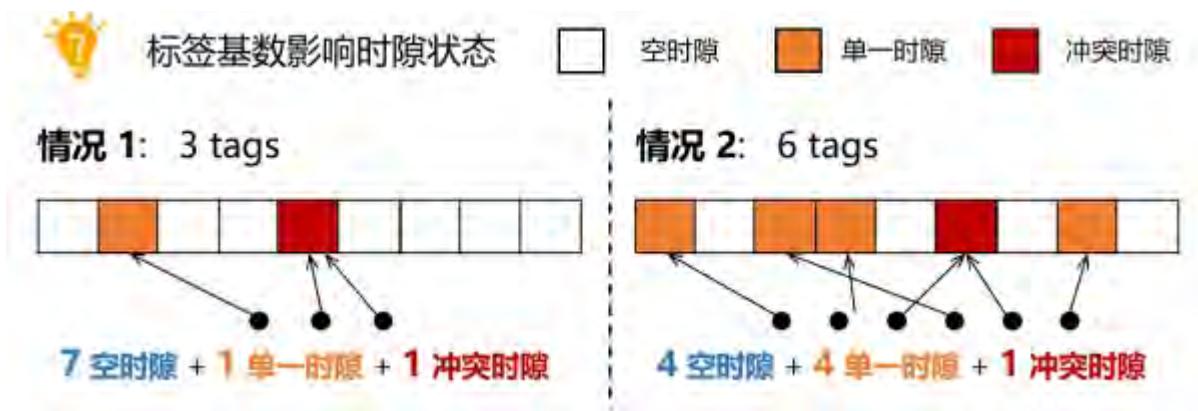


图 3. 标准兼容的 RFID 基数估计算法

针对这一问题，本文提出一种哈希模型来判断标签的大致基数。如图 4 所示，通过将时间分成等长的时隙然后安排标签在不同的时隙回复，RFID 阅读器将接收到三种不同的情况，分别为单一时隙，冲突时隙和空时隙。在单一时隙只有一个标签回复，阅读器可以顺利接收到该标签的信号。在冲突时隙有大于一个标签回复，标签之间的信号会相互干扰，导致阅读器无法接收任何数据。在空时隙没有任何标签回复数据，阅读器也不会接收到数据。显然图 4 中三种时隙的占比与标签的数目又很强的关联性。在左边的例子中，标签数目少，空时隙占比会较高。在右边的例子中，标签数目多，空时隙和冲突时隙的总数会相应增加。因而通过分析对应时隙的个数可以被用于估计标签总数。但是当前的 RFID 通讯协议并不支持获取标签的 MAC 层传输状况，无法统计各个种类时隙的具体个数。针对该问题，本文使用标签的平均回复间隔来估计标签的总数。平均回复间隔受两次单一时隙之间的冲突时隙和空时隙的个数影响，间接反应了异常数据标签的大致规模，因而可以被用于数据异常检测。基于这一思想，本文构建标签个数与回复间隔之间的函数，通过输入测量的回复间隔来推断对应的标签个数，从而在当前 RFID 协议的框架下实现基于基数规模的数据异常检测协议。

#### ■ 基于 C1G2 选择查询指令的区间查询协议。

异常数据收集的一大瓶颈是对大部分正常标签的无效数据采集。如果阅读器只读取异常数据的标签而不收集存有正常数据的标签，数据异常检测的速度将大幅度降低。基于这一思想，本文提出一种基于广播字符串的标签选取方式。例如生产日期的检测中，10月1日到10月7日之间的商品是系统中存在的异常数据。阅读器通过逐个广播数据每一天对应的数据，让数据相同的标签回复数据，就可以通过七条指令找到系统中的异常标签。该方式非常适合在异常值较少的系统中使用，但是异常值较多时的需要广播较多的字符串。

针对该问题，本文通过设计合理的编码方式来辅助阅读器选取系统中的目标标签。如图 5 所示，假设我们要选取数据为 3-4 的目标标签。通过在标签内部写入如图所示的 0-1 比特串，可以通过 5-8 位的共同字符串“0000”选出 2 个目标标签  $t_3$  和  $t_4$ ，从而将查询标签的指令个数从 2 条降低到 1 条。但是通过连续的 0-1 比特的编码方式消耗较多的标签存储空间。当前的 RFID 标签通常只有不到 256 比特的存储区域，支撑 128 个标签的编码。对于该问题，本文进一步将标签的存储空间分为两部分，例如采用 10 个比特来进行类别标记，其余 246 个比特来进行编码，从而将编码的覆盖范围扩大到  $210 \times 123$ ，从而满足绝大部分系统的数据总数。在此基础上，本文研究了协议在当前的 RFID 通讯设备上的部署问题，在实际系统中提升了 RFID 系统的传输速率。

#### (3) 基于自适应随机共振的行为异常检测方式

#### ■ 基于自适应势垒调整的随机共振参数控制

目前的随机共振算法通常采用暴力搜索或者遗传算法来调整随机共振系统的噪声或参数。这种方式虽然可以找到全局最优解，但是因为搜索没有方向，导致算法的收敛速度缓慢，无法应用于需要动态适应环境变化的开放场景。针对这一问题，本文将基于参数调整的随机共振模型，构建参数调整与随机共振之间的关联，从而给搜索指定方向，快速确定随机共振系统的最优参数。

	1	2	3	4	5	6	7	8	9	10	11	12
$t_1$	1	0	0	0	0	0	1	0	0	0	0	0
$t_2$	0	1	0	0	0	0	0	1	0	0	0	0
$t_3$	0	0	1	0	0	0	0	0	1	0	0	0
$t_4$	0	0	0	1	0	0	0	0	0	1	0	0
$t_5$	0	0	0	0	1	0	0	0	0	0	1	0
$t_6$	0	0	0	0	0	1	0	0	0	0	0	1

图 4. 基于二进制编码的区间采样



为达到这一目标，本文从随机共振的生成机理的角度设计调整方式。以如图 6 (a) 所示的双稳态函数系统来说，系统将微弱信号作为小球的一阶倒数去驱动小球在两个数据低点（势井）之间来回跳动。如果双稳态系统的形状与微动信号的噪声匹配。小球在两个势井之间的变换频率将和微动信号的频率相符，并具备

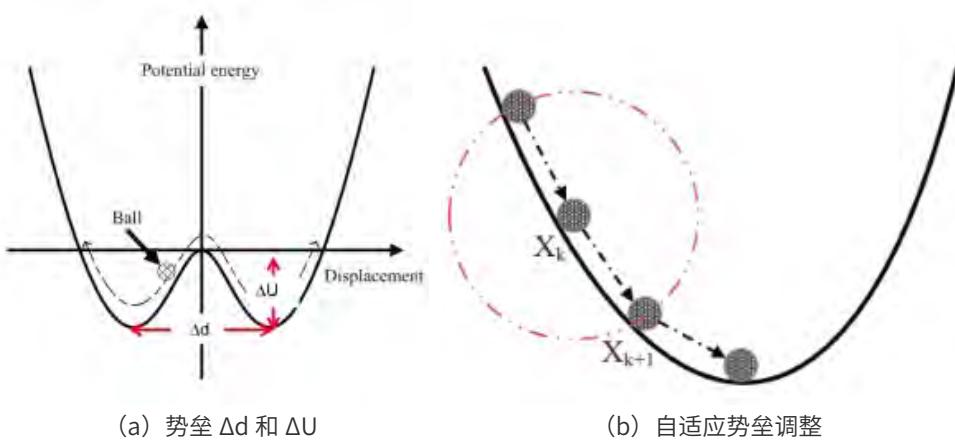


图 5. 基于二进制编码的区间采样

更优的信噪比。当两个势井的间隔和深度不足时，小球会在两个势井之间随机跳动，输出的信号以噪声为主，称为随机震荡。当两个势井的间隔和深度太大时，小球处于其中一个势井保持不变，称为单势井震荡。根据这一原理，本文将势井的深度和距离作为调整参数来生成随机共振。如果双稳态系统输出的震动不足，则减小势垒（增大距离或者减小深度）。如果双稳态系统的震动过强，则增加双稳态系统的势垒。在此基础上，本文通过序列二次优化算法去动态调整势垒的调整步长。首先初始化一套系统参数，如果势垒不足，则增加双稳态函数的势垒。如果增加以后输出依然时随机震荡，则将调整幅度变为两倍。如果调整完毕输出处于单势井震荡，则将调整幅度变为原来的一半。通过以上过程不断调整步长，最终参数将收敛于最优值附近，从而快速产生一次随机共振。

#### ■ 基于频谱分析算法的随机共振评估方式

为了获得最佳信噪比，检测 RFID 系统的微弱信号需要确定滤波器或者随机共振算法的效果。本文设计一种噪声评估函数来判断系统的噪声。以随机共振为例，为了验证参数或者噪声设置的有效性，目前的随机共振算法通常需要预先知道周期信号的频率。传统方法使用短时傅里叶变换获取标签在信号频率处的信号强度与其它所有频率幅值之和的比值（信噪比）来评估信号的噪声，从而判断随机共振是否发生。如果信噪比改善则认为随机共振已经发生。这种方式目前只适用固定频率的信号检测，针对实际场景中常见的动态信号，如呼吸心跳等感知数据，方法的效果会显著下降。

针对这一问题，本文从频谱分析的角度来评估信号的还原质量。图 7 展示了一段随机共振算法的时频分析图（通过连续小波变换获得），当随机共振发生时，标签的信号的能量将集中在某一个区域。当双稳态系统的势垒不足的时候，标签的信号处于随机震荡，频谱的能量分布在所有的频段，没有显著的增强。当随机共振的势垒过高时，输出信号只加强低频部分。根据这一观测，本针对标签的频谱图设计一种人工智能识别算法，根据频谱图像像素点的

分布特点来判断随机共振是否发生。不同于传统评估方法，基于分布的效果评估更加适合频率未知的开放场景，为噪声处理方法在行为异常场景下的落地提供有力支撑。

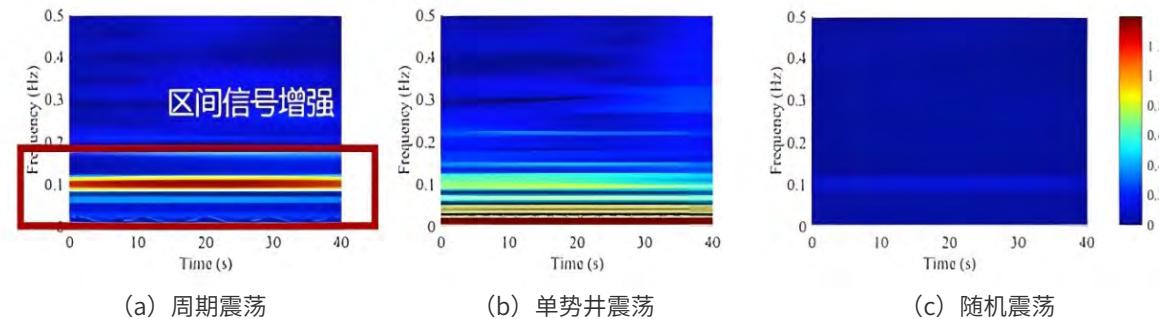


图 6. 基于频谱分析的信噪比分析方法

## 作者简介



陈星宇博士现担任南京大学博士后研究员。2016年至2023年期间，陈星宇博士于南京大学计算机科学与技术系就读博士学位。在博士期间，陈星宇围绕工业互联网、新基建等国家战略需求，在物联网射频感知领域开展较为深入研究，在国际会议与期刊上发表高水平学术论文 20 篇，包括中国计算机学会 A 类（CCF A）论文 11 篇，部分创新成果以第一作者发表在计算机顶会 USENIX NSDI（大陆独立单位第二篇、江苏省首篇）、ACM MOBICOM、ACM MOBISYS，研究成果受到多位院士、Fellow 的正面评价。此外，陈星宇博士注重科研成果转化，授权发明专利 5 项、美国专利 1 项，并参与研发物联网移动定位机器人，获得 2024 年日内瓦国际发明展金奖，2018 年全国高校物联网应用创新大赛二等奖等荣誉。

## 指导老师



陈力军教授现担任南京大学计算机学院、计算机软件新技术国家重点实验室（南京大学）教授、博士研究生导师，南京大学智能机器人研究院院长。陈力军教授于 1993 年和 1998 年在中国矿业大学分别获得自动控制专业硕士学位和博士学位。1998 年至 2000 年，他担任南京大学博士后研究员；2001 年至 2002 年，在美国密歇根州立大学担任博士后研究员；2007 年，他成为香港理工大学访问学者。近年来，陈力军教授在人工智能和机器人多模态融合感知技术开展深入。课题组设计研发图客物联网移动定位机器人，在全国 23 个省份与国外落地应用，获 2022 年世界智能制造十大科技进展，2018 年日内瓦国际发明展特别金奖等奖项。



# 基于变异测试的测试用例集有效性评价技术研究

## ——2024年江苏省计算机学会优秀博士论文奖

作者：张鹏

单位：南京大学

指导老师：周毓明

### 论文摘要

评估测试用例集的有效性是软件测试过程中的重要任务。通过评价当前测试用例集的有效性，测试人员可以了解当前软件被测试的充分性，以便于进一步地安排后续测试工作。当前，围绕测试用例集有效性评估，主要有两大类测试技术：覆盖测试和变异测试。覆盖测试统计多少比例的覆盖目标被执行到，得到一个覆盖率作为测试用例集有效性的评估指标。变异测试模拟了真实场景下，待测试软件可能出现的缺陷（即变异体），然后计算出仿真缺陷被检测出的比例，即变异得分，作为评估测试用例集有效性的评估指标。一般认为覆盖测试技术开销小但评估准确性低，变异测试技术开销大但准确性高。基于这样的事实，探寻一种开销小且准确性高的技术是研究的热点。

但现有技术也存在一些问题。一方面，通过有监督机器学习方法来预测一个模拟缺陷能否被检测出，可以将变异测试的开销大大减少，但相比覆盖测试还有额外开销。能否在保证准确性的前提下，将之进一步优化到与覆盖测试的开销相当？另一方面，面对覆盖测试和变异测试的种种优化技术，研究人员缺乏统一的标准化评估框架。例如在比较多个变异约简策略时，使用现有不合理的评价指标，会使得胜出的约简策略未必在评价测试用例集时是准确的。为解决上述问题，首先，本文提出了一种轻量的无监督预测性变异测试技术，用于进一步减少变异测试的开销。其次，针对变异测试中变异体约简策略现有评价指标不准确、不合理的问题，本文提出了新的基于保序的评估指标，以帮助进行更加合理有效约简变异体。最后，针对任意测试用例集有效性评估技术的评估框架缺乏问题，本文使用提出的新的基于保序的指标，延展出一套标准化的面向测试用例集有效性的度量评估框架。本文的主要贡献如下：

1) 针对现有预测性变异测试技术需要额外开销的问题，本文提出了一种无监督预测性变异测试技术。该技术利用伯努利模型对变异测试建模，基于覆盖信息预测变异体是否被杀死。该技术在进一步减少了预测性变异测试技术开销的基础上，取得了和现有技术相近的预测效果。此外，本文指出了由于有监督方法依赖于训练集，当训练集与测试集存在较大差异时，使用有监督预测方法有风险得到与事实相反的结论。

2) 针对现有变异约简技术评估不合理的问题，本文首先深入分析了现有的两种主流评估指标，指出现有评估指标忽视了测试用例集评估技术的初衷：反映客观世界中测试用例集间有效性的序关系。为了解决这一问题，本文提出两个指标：保序性以及工作量感知的相对保序性。这两个指标直接评估了变异约简策略保持客观世界中测试用

例集间有效性的序关系的能力，为变异约简评估流程提供了客观的评估指标。

3) 针对测试用例集有效性评估不规范不统一的问题，本文还拓展了保序性的应用范畴，提出了一套标准化的测试用例集评估技术的评估框架。无论是覆盖测试还是变异测试，乃至于未来新的测试技术，只要该技术可以用于反映测试用例集的有效性，本文的评估框架就可以提供一种标准化的评估指导方案。

## 专家推荐语

该博士论文以变异测试为核心，围绕测试用例集的有效性评估，提出了多项创新性技术与标准化框架。论文主要探讨了在软件测试领域，如何通过改进变异测试技术，以实现高效且准确的测试评估，解决现有技术中开销大、评估指标不准确等难题。

首先，作者提出了一种轻量的无监督预测性变异测试方法，利用伯努利模型对变异体存活概率进行预测。该方法突破了传统变异测试的高昂开销问题，达到了与有监督方法相似的预测效果，但避免了训练数据依赖性强、部署困难等局限性。此外，该方法在预测过程中只需少量特征，大大提高了技术的通用性。

其次，论文提出了两个基于保序的创新评估指标，解决了现有变异约简策略评估方法不准确的弊端。这些指标不仅能有效衡量约简前后测试用例集有效性的序关系变化，还能显著提升变异约简策略的评估精度。通过这些新指标，研究者可以更客观地比较不同变异约简策略的优劣。

最后，作者设计了一套标准化的测试用例集有效性评估框架。该框架基于保序性，能够为不同类型的测试技术提供统一的评估标准，不仅适用于覆盖测试和变异测试，还可扩展至未来的新型测试技术。这一贡献为测试技术的比较研究提供了极具实践意义的参考。

总体而言，该论文在理论创新与技术实践上均取得了重要进展，具有显著的学术价值和应用前景，值得在相关领域获得高度关注与评价。

## 论文看点

### 论文看点：基于变异测试的测试用例集有效性评价技术研究

张鹏博士的论文《基于变异测试的测试用例集有效性评价技术研究》围绕软件测试中的核心任务——测试用例集有效性评价，提出了多项创新性技术，解决了当前覆盖测试与变异测试技术中的开销与准确性问题，并建立了标准化的评估框架，提升了变异测试的实用性和科学性。以下将详细介绍论文中的几个重要研究亮点和技术创新。

#### 一、测试用例集有效性评价的核心问题

软件测试是确保软件质量的关键环节，而测试用例集的有效性直接决定了测试过程的效果。传统的测试技术主要包括覆盖测试和变异测试。覆盖测试通过统计被执行到的代码比例来评估测试用例集的有效性，覆盖率是该技术的核心指标。然而，覆盖测试的局限性在于它并不能有效检测出潜在的缺陷，只能提供被执行代码的比例信息，评估准确性较低。

相比之下，变异测试是一种模拟实际软件缺陷的技术，它通过引入故意的“变异”（即人为生成的小缺陷），来验证测试用例集能否检测出这些变异体。变异测试的评估指标是“变异得分”，即变异体被测试用例集成功“杀死”的比例。该技术被认为能更真实地反映测试用例集的检测能力。然而，变异测试的开销较大，尤其是在大型软件项



目中，执行大量变异体会导致计算资源的大量消耗。

因此，张鹏博士的研究首先明确了一个核心问题：如何在保证评估准确性的前提下，减少变异测试的开销，使其应用更加广泛？这一问题成为论文研究的基础，驱动了一系列创新技术的提出。

## 二、无监督预测性变异测试的提出

变异测试虽然能够有效评估测试用例集的缺陷检测能力，但由于生成和执行大量变异体，计算开销十分高昂。为了减少这种开销，近年来的研究集中在通过有监督模型预测变异体存活情况。然而，这些模型的训练依赖于大量标注数据，并且假设训练数据与目标数据分布相同。然而，现实软件系统往往存在概念漂移，这使得有监督模型在不同分布下的预测结果不准确。为了应对这一问题，张鹏博士提出了一种基于覆盖的无监督预测方法（Coverage-Based Unsupervised Approach, CBUA），该方法不依赖任何训练数据或变异体执行，大幅降低了变异测试的计算成本。

### 1. 传统方法的局限性

传统的有监督学习方法需要通过大量历史数据训练模型，这些模型利用代码特征和测试特征预测变异体是否会被杀死。然而，训练这些模型需要执行所有变异体以生成标注数据，这无形中增加了计算开销。此外，这些方法假设训练数据与目标系统的数据分布相同，但现实情况往往并非如此。随着时间的推移，软件系统可能发生变化，导致训练模型在新的系统上无法准确预测变异体的存活情况。基于此，张鹏博士指出，现有的有监督方法在概念漂移的情况下表现不佳，并且成本较高，因此需要一种新的方法来解决这些问题。

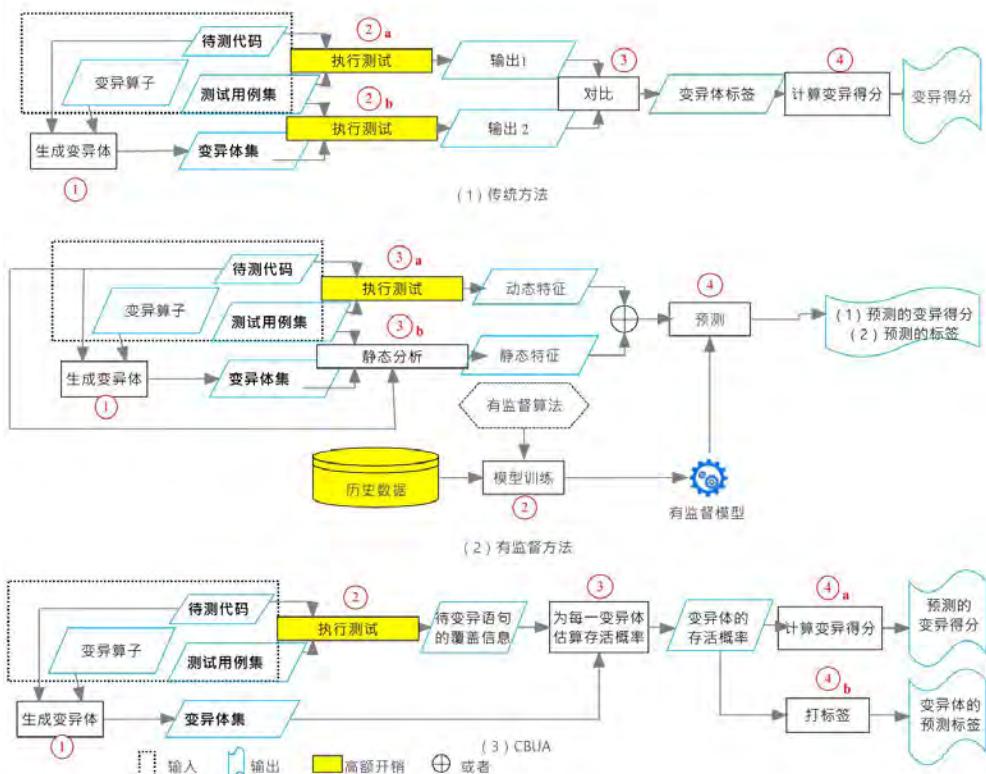


图 1. CBUA 与其他方法的对比

## 2. CBUA 方法的核心思想

为了解决上述问题，CBUA 方法应运而生。CBUA 通过使用代码覆盖信息来预测变异体的存活概率，而不依赖于任何标注数据或变异体执行。具体来说，CBUA 首先收集测试用例集在目标代码上的覆盖信息，特别是变异语句的覆盖情况。然后，该方法根据覆盖信息估算每个变异体的存活概率，并据此计算变异得分，进而评估测试用例集的有效性。相比于传统方法，CBUA 只需一次对原始代码的测试执行，不需要对变异体进行任何额外的操作，大大降低了计算成本。

CBUA 的一个重要特性是得分单调性，即向测试用例集中添加更多的测试用例不会降低其变异得分。这一特性在许多有监督方法中并不一定得到保证，而 CBUA 能够确保添加更多测试用例时，测试用例集的变异得分只会增加，不会减少。这为 CBUA 在测试用例集评估中的应用提供了更强的理论保障。

## 3. 预测流程与实验验证

图 1 对比了传统的变异测试方法，有监督预测性变异测试方法和无监督预测性变异测试方法。对于传统方法，首先要在原始程序以及全部变异体上分别执行测试用例集。然后通过对比输出，获取每一个变异体的标签。最终计算出一个变异得分。对于有监督方法，首先要利用有标签的历史数据训练出一个有监督模型。第二步在测试集上搜集特征，包括了执行覆盖测试时的动态特征，以及直接从代码，变异体和测试代码中提取的静态特征。第三步将这些特征数据作为输入放入预测模型获取预测结果，最终获得一个变异得分。从图中的对比可知，一方面，CBUA 方法相较传统方法，消除了变异体的执行过程；另一方面，相较有监督方法，CBUA 消除了训练过程。只需要如下的三步即可轻松构建：

覆盖信息的收集：首先，执行测试用例集并收集变异语句的覆盖信息。变异语句的覆盖情况反映了测试用例集中各个测试用例是否执行了代码中的变异部分。

存活概率的估算：根据收集的覆盖信息，CBUA 使用概率公式来估算每个变异体的存活概率。具体来说，覆盖的测试用例数量越多，对应的变异体越有可能被杀死。通过概率估算，每个变异体将被分配一个“存活”或“被杀”的标签。

变异得分的计算与标签预测：根据每个变异体的存活概率，CBUA 计算出测试用例集的预期变异得分。此外，通过对比存活概率与 0.5 这一阈值，可以判断变异体是否会被杀死。实验结果表明，CBUA 在预测变异得分的准确性上，与现有的有监督方法相当，甚至在某些场景下表现更为出色，特别是在预测那些被覆盖但未被杀死的变异体时具有更强的能力。这一能力能够帮助开发者识别测试用例集的弱点，从而指导新测试用例的生成。

## 4. CBUA 的优势

CBUA 方法的主要优势在于其低成本、高效性和不依赖训练数据。首先，该方法只需要一次对原始代码的测试执行，而不需要执行所有变异体，从而大大降低了变异测试的计算开销。其次，CBUA 不依赖任何标注数据，避免了有监督模型在概念漂移条件下表现不佳的问题。此外，CBUA 在保证预测准确性的前提下，确保了测试用例集的得分单调性，这使得其在不同场景下具有广泛的应用潜力。

## 5. 实验与应用前景

张鹏博士通过大量实验验证了 CBUA 的有效性。实验中，研究人员使用了多个真实软件项目的数据，评估了 CBUA 在不同场景下的表现。结果显示，CBUA 不仅能够在预测变异得分上与现有有监督方法相媲美，而且在处理



概念漂移时具有更强的适应性。特别是在识别那些被覆盖但未被杀死的变异体时，CBUA 展现出了更高的准确性和实用性。

总体而言，CBUA 为测试用例集的有效性评估提供了一种低成本且高效的替代方案，特别适用于需要减少计算开销的场景。未来的研究可以进一步扩展 CBUA 的应用范围，例如结合更多的代码特征或动态执行信息，以进一步提升其预测能力。

### 三、变异体约简评估指标的创新

变异测试虽然在测试用例集有效性评估中表现出色，但执行全部变异体的开销非常高昂。因此，研究人员提出了多种变异体约简策略，旨在减少需要执行的变异体数量。变异体约简的核心目标是通过选择具有代表性的变异体，减少测试开销的同时，仍然能保持对测试用例集有效性的准确评估。然而，现有的评估指标往往存在局限，难以全面评估变异体约简策略的效果。因此，张鹏博士在论文中提出了新的评估指标，专注于保序性和工作量感知的相对保序性，从而改进对变异体约简策略的评估。

#### 1. 现有评估方法的局限性

在变异体约简评估中，主流的评估方法包括全局变异得分（Global Mutation Score, GMS）和变异得分的变化（Variation in Mutation Score, VMS）。全局变异得分通过最小化测试用例集覆盖变异体的子集来最大化原始变异集的变异得分，而 VMS 则通过比较原始变异集和约简后变异集的变异得分，评估约简策略的有效性。然而，这些指标主要侧重于测试用例集的“绝对”得分变化，忽略了测试用例集之间的“相对”顺序保留问题。

张鹏博士指出，变异体约简策略的真正目标应是保持不同测试用例集在原始变异集下的相对有效性顺序。在评估不同测试用例集的有效性时，重要的不仅是绝对得分的变化，更是这些测试用例集在相对顺序上的保持。因此，传统的评估指标可能会导致误导性的结论，尤其是在不同测试用例集的效果对比中。

#### 2. 保序性与工作量感知的相对保序性

为了克服现有指标的不足，张鹏博士提出了保序性（Order-Preservation, OP）和工作量感知的相对保序性（Effort-aware Relative Order Preservation, EROP）两项评估指标。这两项指标的核心思想是，通过衡量变异体约简后，测试用例集在原始变异集和约简变异集下的相对有效性顺序是否保持不变，从而评估约简策略的有效性。图 2 形象地解释了各种测试用例集间的得分序关系。

保序性（OP）：该指标评估在原始变异集与约简后变异集下，测试用例集的相对顺序是否得到保持。具体而言，如果测试用例集在原始变异集下的得分较高，那么在约简后变异集下也应表现得更好。通过计算这种顺序关系的保持程度，OP 能够更加准确地评估变异体约简策略的效果。

工作量感知的相对保序性（EROP）：该指标进一步考虑了测试成本的因素。即在保持顺序关系的基础上，结合测试用例集执行所需的工作量，评估约简策略的成本效益比。通过引入工作量感知，EROP 能够更好地平衡测试成本与有效性之间的关系。

#### 3. 实验验证与效果分析

为了验证保序性（OP）和工作量感知的相对保序性（EROP）在评估变异体约简策略中的有效性，张鹏博士进行了大量实验。在这些实验中，使用了多个开源项目的数据，对比了几种主流的变异体约简策略，包括随机变异体选择（RMS）、某些特定操作符选择（COS）、子变异体选择（SMS）和聚类变异体选择（CMS）。实验的核心目

标是通过计算不同策略在 OP 和 EROP 指标下的表现，评估其在保持测试用例集有效性顺序上的能力。

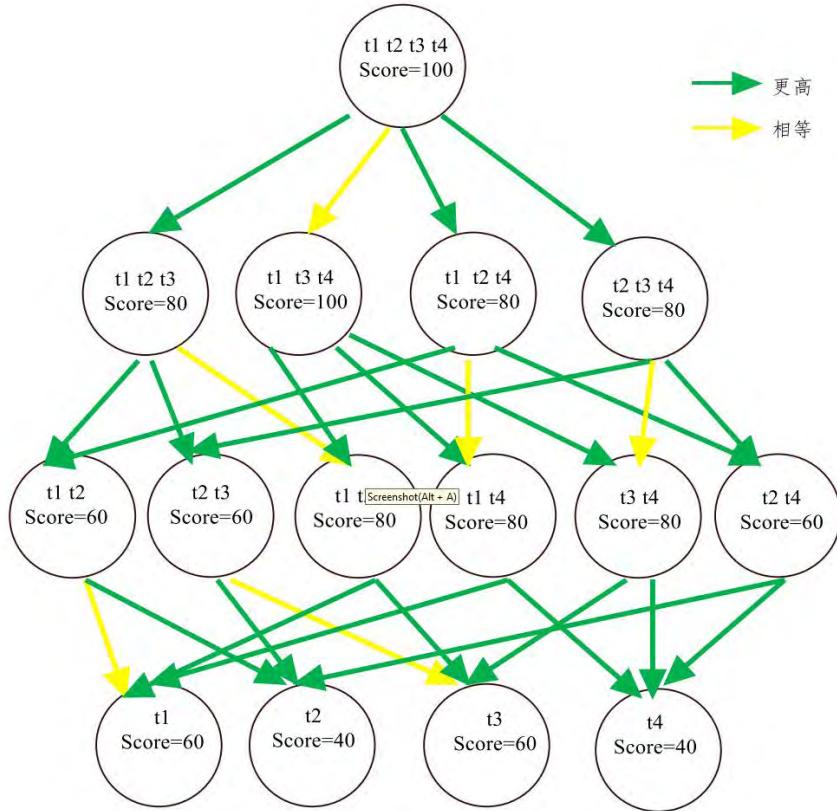


图 2 保序性中的序的形象解释

实验结果显示，保序性（OP）和工作量感知的相对保序性（EROP）能够显著区分不同的变异体约简策略。具体而言，子变异体选择（SMS）和聚类变异体选择（CMS）在这两个指标下的表现最佳，它们在保持测试用例集的相对有效性顺序方面表现优异，能够有效减少执行开销的同时，保证测试效果的准确性。相比之下，随机变异体选择（RMS）和某些特定操作符选择（COS）虽然在减少变异体数量上表现较好，但在顺序保持能力上较为薄弱，容易导致测试用例集的有效性评估出现偏差。

#### 4. 保序性评估的应用前景

保序性（OP）和工作量感知的相对保序性（EROP）作为全新的评估指标，不仅在学术研究中展示了出色的表现，也为工业界的实际应用提供了重要参考。随着软件系统的复杂度增加，如何在保证测试准确性的前提下减少变异测试的执行成本，成为了越来越多软件工程师关注的问题。通过使用 OP 和 EROP，开发者可以在不牺牲评估准确性的情况下，显著减少测试的工作量和资源消耗。

此外，保序性评估方法还可以用于优化现有的变异体生成和选择策略。例如，开发者可以结合 OP 和 EROP 来设计新的混合策略，动态选择合适的变异体子集，从而在不同场景下灵活应用。这不仅提高了测试效率，还确保了软件缺陷检测的有效性。

#### 5. 总结与未来工作



通过实验证，保序性 (OP) 和工作量感知的相对保序性 (EROP) 证明了其在评估变异体约简策略中的独特优势。与传统的全局变异得分 (GMS) 和变异得分变化 (VMS) 相比，OP 和 EROP 更加关注测试用例集的相对有效性顺序，从而为变异体约简策略提供了更全面的评估标准。未来的研究可以进一步扩展这两项指标的应用场景，例如在更大规模的软件项目中进行实验证，或者将其应用于新的自动化测试技术中，以进一步优化测试流程并提升软件质量。

#### 四、标准化评估框架的构建

在软件测试领域中，测试用例集的有效性评估至关重要，而现有的评估方法往往各自为政，缺乏统一的标准。这种局面导致了研究人员在评估不同测试技术时，结论往往不一致，甚至相互矛盾。因此，张鹏博士提出了一套名为 ASSENT 的标准化评估框架，旨在解决这一问题。

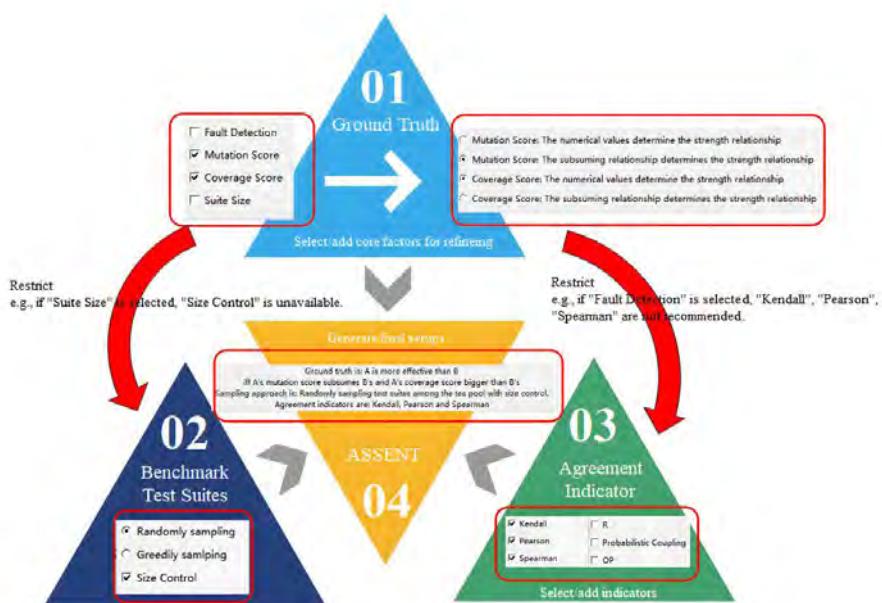


图 3. ASSENT 组件的英文解释和相互关系

#### 1. 评估框架的动机

当前的测试用例集评估方法主要依赖于覆盖率和变异得分等指标。然而，覆盖率仅能反映代码的执行情况，而不能有效衡量测试用例集的缺陷检测能力。变异测试虽然能模拟真实缺陷，但其开销较大。此外，许多研究过于依赖单一指标，没有综合考虑测试成本、缺陷检测能力等多维度因素，导致结果难以通用或在实际应用中效果不佳。因此，构建一个包含多维度、标准化的评估框架成为了亟待解决的问题。

#### 2. ASSENT 框架的组成

为了应对上述挑战，张鹏博士提出了 ASSENT 框架（图 3）。该框架由三个核心组件构成，分别是真实值 (Ground Truth)、基准测试用例集 (Benchmark Test Suites) 和一致性指标 (Agreement Indicator)。

**真实值：**真实值是评估测试用例集是否能够有效检测出缺陷的标准。在该框架中，真实值通常基于测试用例集是否能捕捉到真实缺陷，进而确保测试效果的评估基于统一的标准。

**基准测试用例集：**这是用于生成测试用例集之间有效性顺序的基准。研究人员可以通过不同的抽样策略生成这些测试用例集，例如通过覆盖率或变异体进行随机抽样，来确定它们在不同评估指标下的表现。

一致性指标：该指标用于衡量不同评估指标与真实值之间的一致性。例如，某一评估指标与真实值的顺序关系是否一致，能否准确反映测试用例集的实际效果。常用的一致性指标包括 Kendall's Tau、Spearman 相关系数等。

### 3. ASSENT 框架的应用与验证

ASSENT 框架不仅在理论上提出了标准化的评估方法，还通过实际实验验证了其有效性。在实验中，研究人员使用 ASSENT 框架对多种测试指标进行了系统对比，包括变异得分和覆盖率。实验结果显示，变异得分和子变异得分在反映测试用例集有效性方面表现最佳，与真实缺陷检测能力高度相关。

具体来说，实验针对真实软件项目中的缺陷，生成多个测试用例集，并计算它们的覆盖率、变异得分等指标。随后，使用 ASSENT 框架评估这些指标与实际缺陷检测结果的相关性。结果表明，变异得分能够准确反映测试用例集的有效性，而覆盖率在某些情况下存在高估的风险。这些结果验证了 ASSENT 框架在实际应用中的准确性和有效性。

### 4. 标准化评估框架的优势

ASSENT 框架的提出，填补了测试用例集有效性评估领域中的空白。与传统方法相比，该框架具有以下显著优势：

多维度评估：ASSENT 框架不仅考虑了缺陷检测能力，还结合了测试成本、测试顺序等多维度因素，使评估结果更加全面和准确。

一致性比较：通过一致性指标，研究人员可以更加客观地比较不同测试指标的有效性，避免了由于评估标准不同导致的结果不一致问题。

应用广泛：无论是现有的覆盖测试、变异测试，还是未来的新型测试技术，ASSENT 框架都能为其提供统一的评估方法。因此，该框架不仅具备学术意义，还具有广泛的实际应用前景。

## 五、实验验证与实际应用

为了验证所提出技术和指标的有效性，张鹏博士进行了大量实验。实验结果显示，无监督预测性变异测试技术在大多数场景中，预测精度与现有有监督方法相当，但其执行开销显著降低，证明了该技术在实际应用中的高效性。此外，基于保序性的评估框架在不同变异数体约简策略的对比中，展现出较高的区分能力，能够更加准确地反映测试用例集的有效性。

值得一提的是，张鹏博士的研究成果已经得到了广泛的应用和认可。一些软件测试工具已经采用了论文中提出的无监督变异测试技术，用于提升测试效率。论文的研究不仅具有显著的学术价值，还在实际软件开发和测试过程中发挥了积极作用。

## 六、总结与展望

张鹏博士的论文通过一系列创新技术和评估框架的提出，显著推动了软件测试用例集有效性评估领域的发展。无监督预测性变异测试技术的提出，大大减少了变异测试的开销，提升了其应用可行性；新的评估指标和标准化评估框架的构建，为变异测试技术的评估与推广提供了科学的依据。论文的研究成果既具理论创新性，又有广泛的实际应用前景，必将在软件测试领域产生深远影响。

总的来说，张鹏博士的研究为软件测试技术的发展提供了重要的理论基础和实践指导，是该领域中的一次重大突破。



## 作者简介



张鹏在 2022 年在南京大学计算机科学与技术系获得博士学位，导师周毓明教授。作者目前在卢森堡大学担任博后研究员。研究方向主要是软件测试，学术上围绕变异测试作出了一些突破性的推进，工业上针对深度学习计算芯片相关软件栈精度测试的方法论和工具提供了革命性的解决方案。

## 指导老师



周毓明，男，1974 年生，博士，教授，博士生导师，教育部新世纪优秀人才，江苏省杰出青年基金获得者。主要研究方向为实证软件工程，承担或者作为主要人员参与 973 课题、国家自然科学基金重大研究计划重点项目、863 高技术计划项目和国家自然科学基金项目多项，先后担任多个国际会议的程序委员。与合作者一起在《ACM Transactions on Software Engineering and Methodology》、《IEEE Transactions on Software Engineering》等期刊和 ICSE、FSE、ASE 和 IJCAI 等会议上发表 / 录用论文多篇，研究工作被 60 多个国家的学者他引 1000 多次。获 2007 年江苏省科学技术奖一等奖（第二完成人）、2012 年江苏省科学技术奖一等奖（第七完成人）。2008 年获得教育部新世纪优秀人才计划支持，2010 年获得“CCF 青年科学家”奖，2013 年获得江苏省杰出青年基金支持。

# 迈向自主协同的网络化无人系统

## ——2023年江苏省计算机学会青年科技奖朱琨

### 个人简介

朱琨，南京航空航天大学计算机科学与技术学院教授、博士生导师。入选江苏特聘教授、江苏省333高层次人才培养工程、江苏省科技副总、南航长空学者人才计划、南航杰出人才培养计划，获江苏省计算机学会青年科技奖、ACM南京分会新星奖、国际会议IEEE WCNC最佳论文等奖项。担任江苏省网络与分布式计算专委会主任、南京智慧交通智能网联研究中心主任。主要研究方向为智能网络与计算、网络化无人系统。以第一或通信作者发表CCFA、IEEE/ACM Trans及中科院一区期刊论文六十余篇（包括IEEE JSAC、IEEE TMC、TKDE、TCOM等），在领域旗舰会议发表论文五十余篇。Google Scholar引用4500余次。撰写教材2本，在三十多所高校使用。担任多个期刊编辑及客座编辑，担任MLICOM、WASA等国际会议程序委员会及本地主席。指导学生获江苏省网络空间安全学会优博，互联网+大赛全国金奖，中国机器人大赛-无人机挑战赛全国一等奖、中国高校计算机大赛-网络技术挑战赛二等奖。目前主持JKW创新特区、国家自然科学基金面上项目、重点研发计划课题任务、江苏省优秀青年基金等项目二十余项。



### 坚守初心，科研学术取得突出成果

朱琨在智能网络与计算领域做出了突出的学术贡献，入选江苏特聘教授。近年来以智能网联为中心，围绕感知、算控一体化协同，开展自主协同的网络化无人系统研究，实现了多维感知、智能通信、协同计算、自主决策的初步融合，有效降低了网络化无人机集群系统的信息传输和处理时延，提升了无人机集群系统的智能协作能力。基于上



述成果，与国防单位合作研发体系化反无人机系统，实现无人机反制从单装应用向网络化协同的跨越发展。先后参与建党 100 周年、北京冬奥会等重大活动低空监视任务，并经过海外严酷环境下的实战检验。

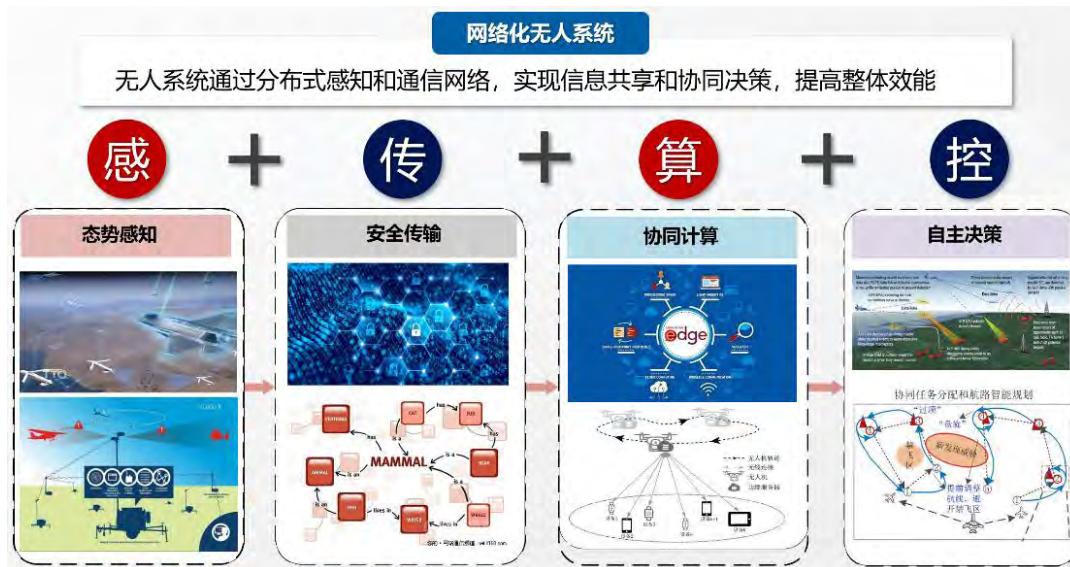


图 1 网络化无人系统

### 产学研用结合，实现成果落地

朱琨坚持产学研用结合，积极服务中航科工、中电集团、中船集团等国防企业以及中兴通讯、南京地铁等重点企业。研发了具有自主知识产权的轨道交通 LTE-M 车地通信网络优化与健康监控平台，一体化集成了基站优化部署、智能频谱分析、干扰识别、故障检测及预警等功能，应用于南京地铁 S1 号线和徐州地铁 1 号线，实现了原有西门子系统的替代。研发的“基于边缘智能的用户侧数据驱动 5G 接入网故障诊断技术”应用于中兴通讯，大幅降低了系统响应时延。同时由于大幅降低了传统方法对输入数据在质和量方面的要求，显著提升了技术适用性。



图 2 地铁电磁环境可视化监测平台

朱琨谨记“科学研究必须服务于人才培养”的教育宗旨，围绕“系统能力、国际视野、创新精神”这一培养目标，以国家重大需求为引领建设“顶天立地、服务国防”的团队文化。工作以来已毕业4名博士，其中三人获校优博，一人获江苏省网络空间安全学会优秀博士论文。指导研究生为第一作者发表CCF A类及IEEE Trans论文三十余篇，多名研究生获国家奖学金等荣誉。基于研究成果，积极带领学生参加高水平竞赛。指导学生获得“互联网+”创新创业大赛全国金奖（2021年）与江苏省一等奖（2020年）、中国机器人大赛-无人机挑战赛全国一等奖（亚军）、江苏省网络技术挑战赛全国二等奖（2023年）。在此基础上，朱琨还担任了江苏省计算机学会网络与分布计算专委会主任，以及多个国际会议的程序委员会主席等职务，为智能化网络与计算等相关领域的科研发展、人才培养以及学术平台的建设持续做出贡献。



## 学会动态

### 第一届江苏省计算机教育大会在南京顺利召开

11月30日至12月1日，第一届江苏省计算机教育大会在南京举行，会议由江苏省计算机学会主办，江苏省计算机学会计算机教育专委会、南京航空航天大学计算机科学与技术学院等单位承办。本次大会以“面向新质生产力的计算机教育”为主题，会议吸引了来自省内外50余所院校及单位的代表参加。江苏省科学技术协会副主席夏军、南京航空航天大学副校长施大宁、江苏省计算机学会秘书长金莹、大会主席陈兵等嘉宾出席开幕式。





# 在“稳健可信人工智能”的科研道路上砥砺前行

## ——2024年江苏省计算机学会青年科技奖郭兰哲

### 个人简介

郭兰哲，南京大学智能科学与技术学院助理教授（特聘研究员），博士生导师。2022年博士毕业于南京大学机器学习与数据挖掘研究所，导师为黎铭教授、李宇峰教授。主要研究方向为开放环境机器学习、神经符号学习，在 ICML/NeurIPS/ICLR/TPAMI 等机器学习领域顶级会议期刊发表论文 30 余篇，研究成果被包括欧洲科学院院士、ACM/AAAI/IEEE Fellow、KDD Innovation Award（数据挖掘最高奖）获得者在内的多位国际知名同行专家好评引用，认为具有重要性和实用性。获吴文俊人工智能科学技术奖优博（当年全国共 9 位）、第九届百度奖学金（全球华人博士生共 10 位）、2024 年江苏省自然科学百篇优秀论文等荣誉，受邀担任国际人工智能联合会 IJCAI 2025 Workflow Chair、南京大学 - 帝国理工大学机器学习联合中心秘书长，亚太数据挖掘会议 PAKDD 2024 稳健机器学习研讨会程序主席，《Frontiers of Computer Science》预备青年编委等。



### 坚守初心，科研学术取得突出成果

郭兰哲面向开放环境下的半监督学习，从开放环境下半监督学习面临的数据分布失配、数据动态流式、先验知识不足、类别比例失衡等问题，提出了多项创新理论与算法，初步建立了一套适用于开放环境的高稳健半监督学习解决方案，工作发表于 ICML、NeurIPS、ICLR、KDD、TPAMI 等国际顶级会议，被包括欧洲科学院院士、IEEE/ACM/AAAI Fellow、NeurIPS 程序主席、数据挖掘最高奖 KDD Innovation Award 获得者在内的多位国际同行专家好评引用，认为具有重要性和实用性。主持研发的半监督学习工具包 LAMDA-SSL 下载量 19K+，研究成果在滴滴出行智能评价、华为金融风控等实际业务中转化应用，入选滴滴雅典娜精英人才培养计划、获华为 - 南京大学人工智能

能联合实验室突出贡献奖。

围绕现有人工智能模型面临的可信性不足、可靠性不足的问题，郭兰哲探索神经符号学习，尝试以数据驱动的机器学习与知识驱动的符号推理融合的方式，提升现有以数据驱动为主的AI模型的可信可靠性。面向智慧司法任务，通过融合法律法条知识，提升现有大语言模型在诸如案情特征识别、争议焦点识别等司法裁判任务中的推理能力。开源司法大模型 Law GPT 获 Github Star 5.9K+。



[pypi v1.0.2](#) [Anaconda.org 1.0](#) [stars 191](#) [visitors](#) [last commit 3 months ago](#)  
[license MIT](#) [coverage 92%](#) [downloads 18k](#) [Platforms osx-64,linux-64,win-64](#)

[Documentation](#) | [Paper](#) | [Examples](#) | [Slide](#)

半监督学习工具包 LAMDA-SSL: <https://github.com/YGZWQZD/LAMDA-SSL>

## LaWGPT：基于中文法律知识的大语言模型



[docs](#) [Wiki](#) [Hugging Face](#) [enutry303](#) [version beta.1](#) [os Linux](#) [last commit June](#) [Stars 5.9k](#)

LaWGPT 是一系列基于中文法律知识的开源大语言模型。

该系列模型在通用中文基座模型（如 Chinese-LLaMA、ChatGLM 等）的基础上扩充法律领域专有词表、大规模中文法律语料预训练，增强了大模型在法律领域的基础语义理解能力。在此基础上，构造法律领域对话问答数据集、中国司法考试数据集进行指令精调，提升了模型对法律内容的理解和执行能力。

详细内容请参考[技术报告](#)。

LawGPT：基于中文法律知识的大语言模型。



## 立德树人，培养人才初步取得成绩

郭兰哲承担南京大学智能科学与技术学院本科生《人工智能导论》与研究生《高级机器学习》课程，结合当前研究趋势，对课程内容进行了创新与改革，采用“问题驱动”和“践履笃行”的方式努力提升学生的学习热情和求知欲。通过与前沿科学问题结合，让学生更多地通过主动思考来分析问题和解决问题，根据课程内容与技术发展趋势，设计多项个人或团队合作的实验题目，让学生更多地投入到实践的过程中来，真正将所知所学付诸实践，提高动手能力，实现知行统一。

郭兰哲担任南京大学本科生新生导师、本科生班主任，每周设立办公室开放时间至少两小时，积极与学生进行交流，答疑解惑，帮助学生确定未来的职业目标和大学学习的努力方向，对个别困难的学生加以特别关注，帮助学生正确面对学习、生活中的挫折，使他们树立信心去迎接挑战，获南京大学优秀本科班主任。

在指导研究生之外，郭兰哲面向本科生开放实验室，指导多项大学生创新创业实践项目，以顶会顶刊论文分享的方式让学生了解人工智能国内外最新的研究方向和前沿进展，指导有兴趣的高年级本科生参与科研活动，锻炼知识运用与实践能力，并通过挑战性问题激发本科生的学术激情与创造力，指导本科生参与全球人工智能算法精英选拔赛，获优秀指导教师。

## 学会动态

### JSCS 计算机伦理与职业修养专委会 2024 年学术年会暨换届选举大会在南京举办

2024年11月30日，江苏省计算机学会计算机伦理与职业修养专业委员会(以下简称“伦理专委会”)在南京大学仙林校区计算机学院111报告厅隆重举行了其2024年学术年会暨换届选举大会。此次会议旨在深入贯彻落实《深化新时代教育评价改革总体方案》，探讨计算机类院系在立德树人根本任务下，如何有效推动课程思政与教育教学改革，并交流课程思政改革中的成功经验，以推动江苏省计算机教育事业的创新发展。全省各高校计算机学科所在院系的党委(党总支)书记、副书记、分团委书记、学工组长、辅导员等80余名教育工作者齐聚一堂，共同参与了此次盛会。会议内容丰富，议程紧凑，既有深入的理论探讨，也有丰富的实践经验分享。



# OpenAI 最强竞对官宣：大模型可以使用电脑了！

——科普中国

最近，OpenAI 最强竞对 Anthropic 宣布：现在，Claude 可以使用电脑了。

据介绍，最新版本的 Claude 3.5 Sonnet 在通过适当的软件设置运行后，可以按照用户的指令在电脑屏幕上移动光标，点击相关位置，并通过虚拟键盘输入信息，模拟人们与电脑进行交互的方式。

Anthropic 认为，这项技能 -- 目前处于公开测试阶段 -- 代表了人工智能（AI）领域的重大突破。

在最新博客文章中，他们分享了在开发计算机使用（computer use）模型过程中的一些研究心得，以及如何让这些模型更加安全。

## 为什么要开发 computer use？

为什么这项新功能很重要？大量的现代工作都是通过计算机完成的。让人工智能能够像人类一样直接与计算机软件进行交互，将开启大量应用，而这些应用对于目前的人工智能助手来说根本无法实现。

在过去几年里，强大的人工智能发展已经取得了许多重要的里程碑式成果 -- 例如，能够进行复杂的逻辑推理，能够看到和理解图像。下一个前沿领域是计算机应用：人工智能模型无需通过定制工具进行交互，而是可以根据指令使用任何软件。

## 研究过程

Anthropic 表示，他们以前在工具使用和多模态方面的工作为这些新的计算机使用技能奠定了基础。操作计算机需要具备查看和解释图像的能力，这里指的是计算机屏幕上的图像。它还要求推理如何以及何时根据屏幕上的内容执行特定操作。结合这些能力，他们训练 Claude 解读屏幕上的内容，然后使用可用的软件工具执行任务。

当开发人员让 Claude 使用一款计算机软件并赋予其必要的访问权限时，Claude 会查看用户可见内容的屏幕截图，然后计算光标需要纵向或横向移动多少像素才能点击正确的位置。训练 Claude 准确计算像素至关重要。如果没有这项技能，模型就很难下达鼠标指令 -- 这就好比模型在回答“‘香蕉’这个词中有多少个 A？”这样看似简单的问题时经常会感到吃力一样。

令人感到惊讶的是，Claude 在接受了计算器和文本编辑器等几款简单软件的 computer-use 训练后（出于安全考虑，模型在训练期间无法访问互联网），竟然能够迅速地掌握这些技能。结合 Claude 的其他技能，这种训练使它具备了非凡的能力，能够将用户的书面提示转化为一连串的逻辑步骤，然后在计算机上进行操作。他们观察到，该模型甚至会在遇到障碍时进行自我纠正并重试任务。



虽然他们在取得初步突破后很快就取得了后续进展，但这需要大量的尝试和错误才能实现。Anthropic的一些研究人员指出，开发 computer use 模型的过程与他们初入人工智能领域时想象的“理想化”人工智能研究过程非常接近：不断迭代，反复回到绘图板（drawing board），直到取得进展。



目前，Claude 是像人一样使用计算机的 SOTA 模型，即通过观察屏幕并采取相应行动。在 OSWorld 为测试开发者让模型使用计算机的尝试而创建的一项评估中，Claude 目前的得分率为 14.9%。尽管远未达到人类水平（一般为 70%-75%），但却远远高于同类产品中排名第二的人工智能模型的 7.7%。

## 安全使用计算机

人工智能的每一次进步都会带来新的安全挑战。computer use 主要是降低人工智能系统应用其现有认知技能的门槛，而不是从根本上提高这些技能，因此 Anthropic 对 computer use 的主要关注点是当前的危害而非未来的危害。他们发现，更新后的 Claude 3.5 Sonnet（包括其新的 computer use 技能）仍处于人工智能安全等级 2 级，也就是说，它并不需要比 Anthropic 现有的安全和安保措施更高的标准。

当未来的模型因存在灾难性风险而需要人工智能安全等级 3 级或 4 级保障措施时，computer use 可能会加剧这些风险。Anthropic 判断，在模型还只需要人工智能安全等级 2 的保障措施时，现在就引入 computer use 可能会更好。这意味着，他们可以在风险过高之前开始处理任何安全问题，而不是在风险更为严重的模型中首次添加 computer use 功能。

因此，Anthropic 的信任与安全团队对新的 computer use 模型进行了广泛的分析，以找出潜在的漏洞。他们发现的一个问题是“提示注入”--这是一种网络攻击，即向人工智能模型输入恶意指令，使其推翻先前的指令或执行偏离用户初衷的意外操作。由于 Claude 可以解读来自互联网计算机的屏幕截图，因此有可能接触到包括提示注入攻击在内的内容。

使用公开测试版 Claude computer-use 版本的用户应采取相关预防措施，将此类风险降至最低。作为开发人员的资源，Anthropic 在参考实现中提供了进一步的指导。

与任何人工智能功能一样，用户也有可能故意滥用 Claude 的计算机技能。Anthropic 开发了分类器和其他方

法来标记和减少这类滥用。

Anthropic 表示，根据他们对数据隐私采取的标准方法，默认情况下，他们不会对用户提交的数据（包括 Claude 收到的任何截图）训练他们的人工智能生成模型。

## computer-use 的未来

Computer use 是一种完全不同的人工智能开发方法。到目前为止，LLM 开发人员一直在让工具与模型相匹配，创造定制环境，让人工智能使用专门设计的工具来完成各种任务。现在，Anthropic 可以让模型适应工具 -- Claude 可以适应人类日常使用的计算机环境。他们的目标是让克劳德能够像人一样使用已有的计算机软件。

当然，还有很多事情要做。Claude 使用计算机的速度仍然很慢，而且经常出错。人们在计算机上经常做的许多操作（拖动、缩放等），Claude 还无法尝试。Claude 的屏幕视图具有“翻书”的性质 -- 截图并将它们拼凑在一起，而不是观察更细粒度的视频流 -- 这意味着它可能会错过短暂的操作或通知。

Anthropic 表示，即使在为今天的发布录制 Computer use 演示时，他们也遇到了一些有趣的错误。例如，Claude 不小心点击了停止长时间屏幕录制，导致所有镜头丢失，以及 Claude 突然中断了编码演示，开始浏览黄石国家公园的照片。

Anthropic 预计，Computer use 将迅速改善，变得更快、更可靠，对用户想要完成的任务更有用。对于软件开发经验较少的人来说，它也将变得更容易实现。

(文章来源科普中国;[https://www.kepuchina.cn/article/articleinfo?business\\_type=100&classify=0&ar\\_id=537510](https://www.kepuchina.cn/article/articleinfo?business_type=100&classify=0&ar_id=537510))

## 学会动态

### 江苏省计算机学会大数据专委会工作会议在南京召开

2024 年 6 月江苏省计算机学会大数据专委会进行了领导换届工作。为了进一步推动新一届大数据专委会更好的工作，9 月 14 日下午，江苏省计算机学会大数据专委会在南京审计大学召开了 2024 年度工作会议。本次会议旨在研讨新一届大数据专委会的工作思路，明确新一届专委会的工作任务和工作计划。会议由大数据专委会主任吉根林教授主持，来自计算机学会、大数据专委会及南京审计大学计算机学院的多位领导与专家参加了此次会议。





# 为什么说光量子计算是面向未来的计算范式？

——科普中国

想象一下，未来某一天，我们不再依赖那些辛勤工作的电子，而是召唤出光子——光的魔法精灵，来构建光量子芯片，那将是多么神奇的景象！

光量子芯片与传统计算机中的电子芯片不同，它使用光子作为信息的传递者。光子不仅是光的最小单位，在量子世界里，它们还有一个酷炫的名字——光量子。

这些光量子芯片拥有超快的传输速度、几乎不消耗能量的低能耗，以及能够处理海量数据的大带宽。这些特点预示着光量子芯片有潜力打破电子芯片的极限，满足未来对计算能力爆炸性增长的需求。

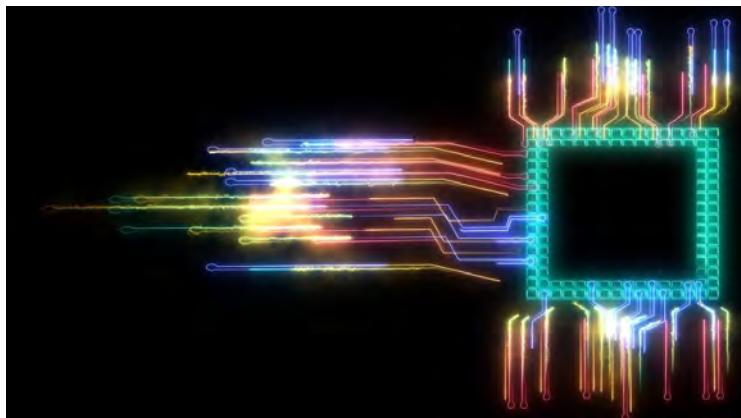


图 1 光量子芯片的艺术示意图

(图片来源：VEER 图库)

## 一、传统的计算机体系架构——冯·诺依曼计算范式

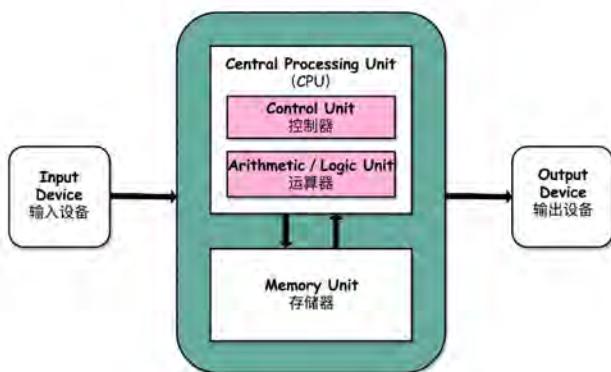


图 2 “冯·诺依曼计算范式”的示意图

(图片来源：wikipedia)

的电子计算机都遵循这种体系架构，它有两个显著的特点：一个是“存算分开”，另一个是“顺序执行”。

提到电脑和手机，大家肯定再熟悉不过了，它们都是我们日常生活中的电子计算机。

但你可能不知道的是，目前市场上几乎所有的电子计算机都采用了一种叫做“冯·诺依曼计算范式”的体系架构。

[https://en.wikipedia.org/wiki/Von\\_Neumann\\_architecture#/media/File:Von\\_Neumann\\_Architecture.svg](https://en.wikipedia.org/wiki/Von_Neumann_architecture#/media/File:Von_Neumann_Architecture.svg)

这个听起来有点拗口的术语，其实是由 20 世纪 40 年代的匈牙利数学家和物理学家约翰·冯·诺依曼 (John von Neumann) 提出的。现在，几乎所有

让我们用一些简单的关键词来揭开传统电子计算机的神秘面纱：

首先，是大家相对熟悉的中央处理单元（CPU）和存储器。

CPU 是计算机的大脑，负责执行计算和逻辑操作；而存储器则像是计算机的书架，用来存放大量的数据和指令，方便随时取用。这种 CPU 和存储器分离的设计，就是我们所说的“存算分开”。

虽然这种设计让编程变得简单，但它也有缺点。数据和指令需要在 CPU 和存储器间来回穿梭，这不仅拖慢了速度，还增加了能耗和延迟。就像在繁忙的街道上，如果车辆频繁往返于两个地点，交通就会变得拥堵。

接下来，是对大家来说，可能有些陌生的“串行运算”和“数据潮汐”。

“串行运算”就像排队到收银台来结账，每个人都得等到前一个人完成后才能轮到自己。在电子计算机中，每条指令也得等前一条执行完毕才能开始。而“数据潮汐”则像是计算机版的潮涨潮落。有时候，数据和指令像潮水一样汹涌而来，CPU 忙得不可开交；有时候，CPU 却闲得发慌，等着新的数据和指令到来。不管怎样，CPU 一次只能处理一条指令，面对海量数据和指令的冲击，它也会感到力不从心。

这种“顺序执行”的方式，对于并行处理大量任务来说，就像是单行道上的交通堵塞，难以应付。这就是为什么科学家们正在探索新的计算模式，比如前面提到的光量子芯片，它们可能会带来革命性的改变。

## 二、光量子计算——面向未来的计算范式

幸运的是，光量子计算方案为我们带来了新的希望。

它采用光子作为信息的载体，能够充分利用光的高速、低能耗和大带宽等独特优势，有望绕开冯·诺依曼计算范式中“存算分开”和“顺序执行”的难题，从而打破传统电子计算机架构的性能瓶颈。

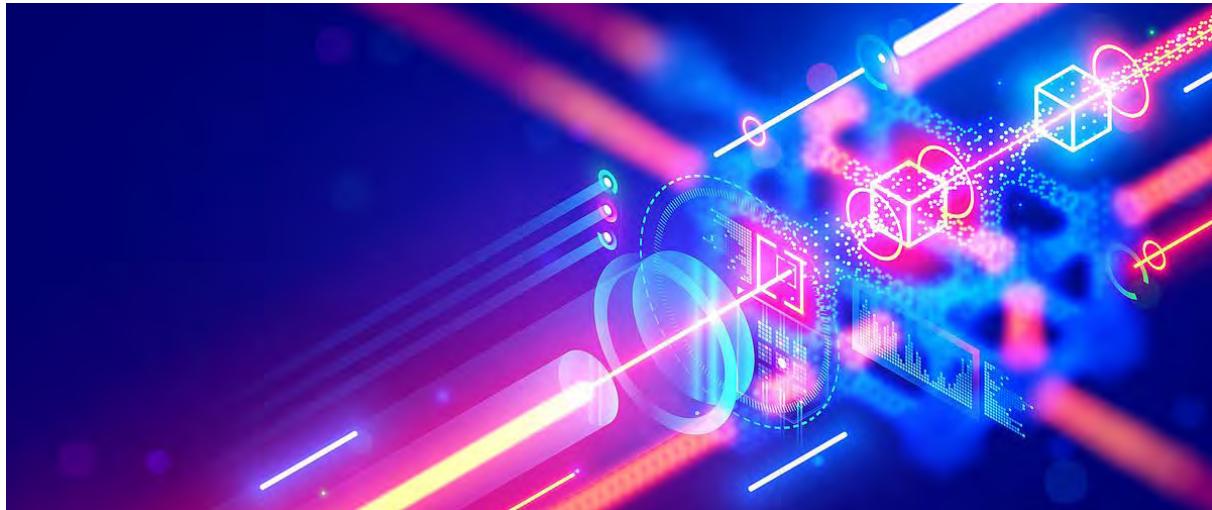


图 3 光量子计算方案的艺术示意图

(图片来源：VEER 图库)

让我们用两个简洁的短语来概括光量子计算方案的核心特征：那就是“传输即计算”和“结构即功能”。

“传输即计算”意味着，在光量子计算方案中，光子在传输的同时就能完成计算任务，无需像传统计算机那样在不同的硬件单元之间来回传输数据。这是因为光子在传输过程中始终以光速进行信息传递，在光量子芯片的厘米



尺度上，光信号的延迟仅为纳秒级别。计算可以在光子通过光学网络或在光子芯片上传输时进行，极大地提高了运算效率。

“结构即功能”则表明光量子芯片的结构设计直接决定了它所实现的功能。光量子芯片使用光波导替代电子芯片中的铜导线，实现光子在芯片中的信息传输。当光子在不同的光波导中传输时，它们之间会产生光信号的干涉，科学家们可以利用这一物理现象来模拟线性计算等计算过程，即通过光子在传播和相互作用中的信息变化来进行运算。

因此，通过设计不同结构的光量子芯片，科学家们可以实现各种功能的量子算法任务，充分发挥光子的独特性能优势，有效地解决电子芯片所面临的挑战。

### 三、光量子芯片：人工智能的超速引擎

在人工智能的竞技场中，速度是关键。

想象一下，如果中央处理器（CPU）只能像超市收银台前的队伍一样，一个接一个地处理任务，那效率得多慢！但幸运的是，图像处理器（GPU）就像是拥有超能力的收银员，能够同时处理多个顾客，大大加快了任务处理的速度，为人工智能的发展注入了活力。

说到并行运算，光量子计算就像是拥有分身术的超级英雄，能够在瞬间处理无数任务。光子，这些以光速奔跑的粒子，它们的自由度就像是多功能的瑞士军刀，可以同时完成多种任务。无论是不同的偏振状态，还是不同的路径，甚至是轨道角动量，这些特性都能被用来提升运算的并行性。

在光量子芯片的神奇世界里，科学家们就像魔术师一样，能够对光子进行各种魔法般的操控。他们使用“波分复用”技术，就像指挥交通一样，让光子在光波导中井然有序地并行传输，利用路径信息来实现光量子态的编码。

更令人兴奋的是，光量子芯片还能兼容现今主流的半导体制造工艺，这意味着制造成本更低，未来的光量子计算机可能会像家用计算机一样普及。这不仅能够满足人工智能硬件的性能需求，还可能开启一个全新的计算时代。

### 四、光量子计算已不只停留于方案

当我们畅想未来的计算方案时，光量子计算方案无疑是一位拥有超能力的竞争者。它不仅拥有闪电般的速度、几乎不消耗能量的神奇能力，还有着处理海量信息的超宽带。



图 4 人工智能技术的艺术示意图

（图片来源：VEER 图库）

想象一下，如果冯·诺依曼架构是一辆老旧的蒸汽火车，那么光量子计算方案就像是现代的高速磁悬浮列车，轻松突破了速度和能耗的极限。

而且，光量子计算方案的并行运算能力，就像是拥有分身术的超级英雄，能够同时解决多个问题，这对于人工智能来说，简直是如虎添翼。

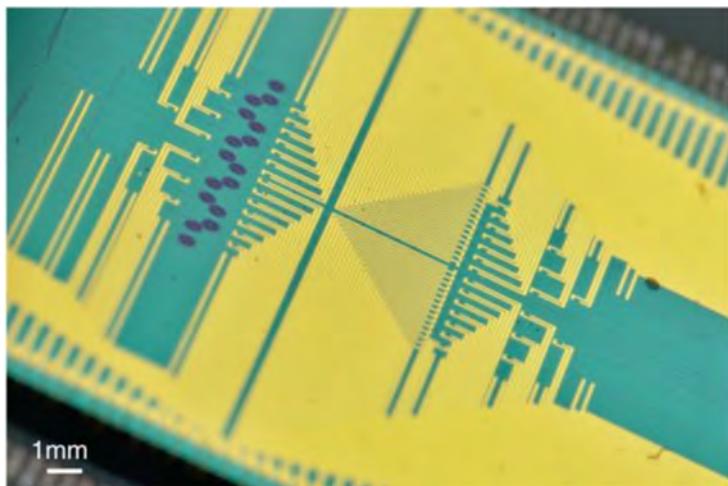


图 5 基于光波导的光量子芯片实物图

(图片来源：Science) 参考文献 [4]

Multidimensional quantum entanglement with large-scale integrated optics | Science

现在，让我们来点小剧透：光量子芯片的研究已经取得了令人兴奋的进展！

科学家们已经不仅仅是在实验室里摆弄小实验，他们已经在用这些神奇的光子芯片创造出一些令人惊叹的成果。

那么，这些光量子芯片到底能做什么呢？科学家们是如何利用它们来探索未知的量子世界的呢？保持你的好奇心，让我们一起在接下来的文章中揭开光量子芯片的神秘面纱，探索它们如何改变我们的世界吧！

## 参考文献

- [1] Von Neumann J. Von neumann architecture[J]. Online [http://en.wikipedia.org/wiki/Von\\_Neumann\\_architecture](http://en.wikipedia.org/wiki/Von_Neumann_architecture), 1945, 8.
- [2] 王剑威 , 丁运鸿 , 龚旗煌 . 大规模集成光量子芯片实现高维度量子纠缠 [J]. 物理 , 2018, 47(5): 317-319.
- [3] 包觉明 , 陈晓炯 , 丁运鸿 , 等 . 硅基光量子芯片上量子调控技术和量子信息应用 [J]. 科学 : 物理学 , 力学 , 天文学 , 2020, 50(8): 47-51.
- [4] Wang J, Paesani S, Ding Y, et al. Multidimensional quantum entanglement with large-scale integrated optics[J]. Science, 2018, 360(6386): 285-291.

(文章来源科普中国 [https://www.kepuchina.cn/article/articleinfo?business\\_type=100&classify=0&ar\\_id=532211](https://www.kepuchina.cn/article/articleinfo?business_type=100&classify=0&ar_id=532211))



# 研究实锤：别让大模型“想”太多！ OpenAI 准确率竟下降 36.3%？

——科普中国

思维链（CoT）已被证明可以在许多任务（如多步骤推理）上显著提升大模型的性能。然而，在哪些情况下，CoT 会系统性地降低大模型的性能，这仍然是一个有待进一步讨论的问题。

如今，来自普林斯顿大学和纽约大学的研究团队，参照思考对“人类性能”的影响，提出了新的见解。

他们认为，虽然模型的认知过程与人类的认知过程并不完全相同，但可以参照思考对人类“性能”产生负面影响的情况，假定思考会对模型产生负面影响的环境。

他们从心理学中选择了 6 项已被充分研究的任务类型来探讨 CoT 对 LLM 性能的影响，并验证了 CoT 在一些任务中甚至可能导致模型准确率下降。

这一发现不仅为未来优化 LLM 的提示策略提供了新思路，还为理解人类与模型在推理过程中的相似性与差异性带来了新见解。

## MIND YOUR STEP (BY STEP): CHAIN-OF-THOUGHT CAN REDUCE PERFORMANCE ON TASKS WHERE THINKING MAKES HUMANS WORSE

Ryan Liu<sup>\*1</sup>, Jiayi Geng<sup>\*1</sup>, Addison J. Wu<sup>1</sup>, Ilia Sucholutsky<sup>2</sup>, Tania Lombrozo<sup>3</sup>, Thomas L. Griffiths<sup>1,3</sup>

<sup>1</sup>Department of Computer Science, Princeton University

<sup>2</sup>Center for Data Science, New York University

<sup>3</sup>Department of Psychology, Princeton University

{ryanliu, jiayig}@princeton.edu

论文链接：<https://arxiv.org/abs/2410.21333>

研究表明，CoT 并非在所有任务中都能提高模型性能，在隐性统计学习、面部识别、含例外模式的数据分类三种情况下，各种 SOTA 模型的性能都会明显下降。此外，研究本身进一步揭示了通过人类心理学研究大模型的可行性。

### 研究方法

为分析 CoT 对大语言模型（LLM）与多模态大模型（LMM）性能的影响，该研究的方法框架基于以下两个关键条件：

- (1) 言语思考或深思熟虑会损害人类“性能”的情况。

(2) 将制约人类“性能”的因素推广到语言模型的情况。

之后，为验证“CoT 在一些任务中会导致模型表现下降”的假设，研究团队在上述两个条件的指导下基于人类心理学设计了以下 6 种任务场景：

隐性统计学习 (Implicit Statistical Learning)：考察模型在隐含语法结构的分类任务中使用 CoT 是否会降低表现。基于心理学中的实验结果，该研究假设人类在进行语言推理时往往表现较差，因此 CoT 在该场景下应有类似的效果。

面部识别 (Facial Recognition)：在该任务中，模型需要识别图像中的人脸。基于人类在口头描述面部特征后识别率下降的现象，研究假设 CoT 会影响模型的面部识别准确性。

含例外模式的数据分类 (Classifying Data with Patterns that Contain Exceptions)：该任务模拟模型在含有异常标签的数据中学习的表现。研究假设 CoT 会导致模型在遇到例外情况时增加学习轮次，因为人类通常会倾向于建立简单规则，从而忽视个别特例。

解释逻辑不一致 (Explaining a logical inconsistency)：在逻辑一致性判断任务中，模型需要识别出两句话之间的逻辑冲突，该任务通常会引发人类的语言推理困难。

空间直觉 (Spatial Intuitions)：模型需要推断液体在倾斜容器中的位置。该任务依赖空间和运动直觉，心理学研究表明人类在使用语言推理时效果不佳，该研究假设模型也会遇到类似问题。

特征聚合决策 (Aggregating Features for a Decision)：模型在多维度决策情境中聚合信息并做出决策。由于信息过载通常会导致人类在 CoT 模式下表现不佳，因此研究假设在该任务中，CoT 将不会提高模型性能。

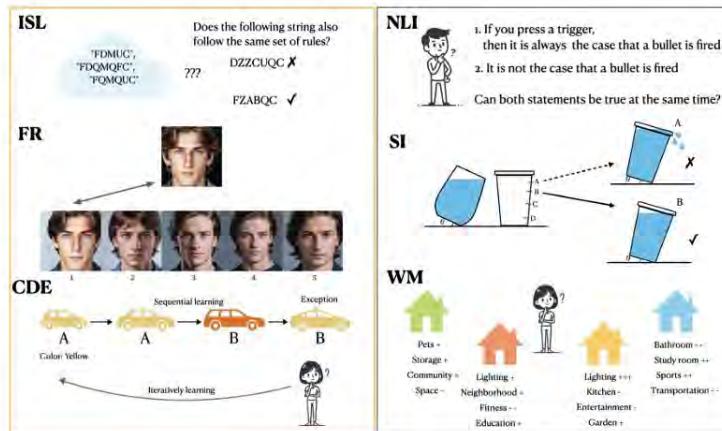


Figure 1: Tasks evaluated for reductions in performance from CoT prompting. Implicit Statistical Learning (ISL): Classification of strings generated by an artificial grammar. Face Recognition (FR): Recognition of a face from a set that shares similar descriptions. Classification of Data with Exceptions (CDE): Learning labels in the presence of exceptions. Natural Language Inference (NLI): Recognizing a logical inconsistency. Spatial intuitions (SI): Tilting water glasses. Working Memory (WM): Aggregating features for a decision. Humans show reductions in performance when engaging in verbal thinking in all tasks, we show that the first three have similar effects on LLMs and VLMs, while the last three differ between humans and models in meaningful ways.

图 | 对 6 项任务进行评估，以确定 CoT 提示是否会降低任务的绩效。（来源：该论文）

针对每个任务场景，研究团队分别构建了零样本 (zero-shot) 和 CoT 提示条件，并在多个主流 LLM 和 LMM 上进行测试，包括 GPT-4o、Claude 3.5、Llama 等，通过对不同条件下模型的准确率，量化 CoT 提示的效果，从而验证他们的假设。



## 实验结果

研究团队首先对满足上述两个关键条件的 3 类任务场景进行实验验证。

### 隐性统计学习

针对该情境，该研究考察了模型在分类基于特定语法结构的序列时的表现。任务包含 4400 个分类问题，基于 100 种有限状态语法 (FSG) 结构，每个测试提供 15 个样例，再要求模型对新序列进行分类。

实验结果显示，使用 CoT 提示的模型表现显著下降，尤其是 OpenAI o1-preview 模型的准确率下降了 36.3%。这表明当模型过度依赖逐步推理时，CoT 可能会抑制其对隐性统计模式的学习能力。

Table 1: Results contrasting zero-shot and CoT for artificial grammar learning.

	Zero-shot	CoT	Performance decrease	p-value
GPT-4o (subset)	94.00%	-	36.30%	< 0.0001
OpenAI o1-preview (subset)	-	57.70%	36.30%	< 0.0001
GPT-4o	87.50%	64.40%	23.10%	< 0.0001
Claude 3 Opus	70.70%	62.70%	8.00%	< 0.0001
Claude 3.5 Sonnet	65.90%	67.70%	-1.80%	0.969
Gemini 1.5 Pro	68.00%	61.95%	6.05%	< 0.0001
Llama 3 8B Instruct	59.70%	57.90%	1.80%	< 0.05
Llama 3 70B Instruct	60.50%	58.30%	2.20%	< 0.05
Llama 3.1 8B Instruct	53.52%	51.54%	1.98%	< 0.0001
Llama 3.1 70B Instruct	65.90%	57.10%	8.80%	< 0.0001

图 | 人工语法学习中 zero-shot 和 CoT 对比结果。 (来源：该论文)

### 面部识别

在该任务情境中，该研究测试了 CoT 是否会影响模型的面部识别能力，这是基于心理学中“语词遮蔽”现象进行的任务情境设计。模型需要在 500 项任务中从 5 个候选中匹配初始人脸。

结果表明，当被要求执行 CoT 时，每个被测试的 LMM 都显示出性能下降，与假设一致。

Table 2: Comparison of zero-shot and CoT prompts for facial recognition.

	Zero-shot	CoT	Performance decrease (absolute)	Performance decrease (relative)	p-value
GPT-4o	64.00%	51.20%	12.80%	20.00%	< 0.01
Claude 3 Opus	44.00%	29.60%	14.40%	32.73%	< 0.0001
Claude 3.5 Sonnet	97.80%	94.80%	3.00%	3.07%	< 0.05
Gemini 1.5 Pro	66.00%	54.60%	11.40%	17.27%	< 0.05
InternVL2 26B	9.20%	6.00%	3.20%	34.78%	< 0.05
InternVL2 Llama3 76B	15.77%	13.77%	2.00%	12.68%	0.44

图 | 面部识别中 zero-shot 和 CoT 提示的对比。 (来源：该论文)

### 含例外模式的数据分类

该任务通过包含多个主次特征的分类任务来测试模型在处理含例外情况时的表现，任务要求模型在多次分类中逐步学习，目标是尽可能减少迭代次数。

实验在 GPT-4o、Claude 3.5 Sonnet 和 Claude 3 Opus 上进行，结果表明，CoT 显著增加了学习轮次。平均来看，GPT-4o 在 CoT 条件下完成正确分类所需的轮次为直接提示的四倍，而 Claude 3.5 Sonnet 和 Claude 3 Opus 的轮次需求也分别增加至直接提示的两倍多。

Table 3: Average number of rounds for models to learn labels using either direct or CoT prompting.

	Direct	CoT	# Rounds increase (absolute)	# Rounds increase (relative)	p-value
GPT-4o	2.9	12.5	9.6	331%	< 0.0001
Claude 3.5 Sonnet	2.3	6.4	4.1	178%	< 0.0001
Claude 3 Opus	2.4	5.5	3.1	129%	< 0.05

图 | 使用直接或 CoT 提示，模型学习标签的平均轮数。（来源：该论文）

在 GPT-4o 的进一步分析中发现，直接提示使模型在第二或第三轮就能达到完美分类，而使用 CoT 时模型在第四到第五轮仅能正确分类 8/10 的对象。这表明 CoT 提示会引导模型偏向基于规则的推理方式，而忽视了已知的正确答案，导致分类效率大幅下降。

之后，研究团队又对满足条件（1）但不满足条件（2）的三类任务情境开展实验。

## 解释逻辑不一致

在该任务中，模型需要识别句子对中的逻辑矛盾性。该任务基于 SNLI 和 MNLI 数据集以及合成数据集。

研究发现，CoT 增加了模型忽视矛盾的可能性，模型在逐步推理时更倾向于关注复杂的逻辑结构，从而忽视了直接矛盾判定。这表明在需要精确逻辑验证的任务中，CoT 提示存在局限性。

Table 4: Results comparing zero-shot and CoT across the logical inconsistency task using stimuli from MNLI, SNLI, and synthetic LLM generation.

	MNLI		SNLI		Synthetic	
	Zero-shot	CoT	Zero-shot	CoT	Zero-shot	CoT
OpenAI o1-preview (subset)	-	-	-	-	-	86.5%
GPT-4o	53.2%	93.9%	51.4%	94.3%	51.0%	74.0%
Claude 3.5 Sonnet	65.2%	67.5%	67.4%	69.8%	56.7%	57.8%
Claude 3 Opus	62.7%	58.8%	66.2%	58.7%	54.5%	51.8%
Gemini 1.5 Pro	73.2%	68.2%	68.8%	63.9%	60.5%	61.5%
Llama 3.1 70B Instruct	55.6%	81.6%	50.4%	82.3%	50.0%	65.8%

图 | 逻辑不一致任务中比较 zero-shot 和 CoT 的结果。（来源：该论文）

## 空间直觉

在该情境中，模型需要通过“倾斜杯子”的问题来推断水面的位置。这类任务依赖于人类的空间或运动直觉，而人类通常在非言语思维下表现更好。

模型接收了视觉提示和多项选择答案，实验结果显示，使用 CoT 提示对模型表现无明显影响。这说明在依赖空间或运动直觉的任务中，模型的推理方式与人类的直觉差异较大，因而 CoT 提示的负面影响较小。



Table 5: Results comparing zero-shot and CoT on the spatial intuition task.

	Zero-shot	CoT	Performance change (absolute)	Performance change (relative)	p-value
GPT-4o	38%	40%	+2%	+5.00%	0.61
Claude 3.5 Sonnet	42%	38%	-4%	-10.53%	0.28
Claude 3 Opus	42%	38%	-4%	-10.53%	0.28
Gemini 1.5 Pro	35%	36%	+1%	+2.78%	0.99
InternVL2 Llama3 76B	39%	31%	-8%	-25.81%	0.67

图 | 空间直觉任务中 zero-shot 和 CoT 的比较结果。 (来源: 该论文)

## 特征聚合决策

此任务模拟了基于多项特征的决策过程 (如选房) , 用于测试信息超载对决策的影响。人类在类似任务中由于记忆限制, 往往在 CoT 模式下表现较差。相对地, 模型保留了所有上下文信息, 能够无损地聚合和评估每项特征。

结果显示, CoT 提示在高上下文记忆任务中提高了模型表现, 说明在信息保留至关重要的场景下, CoT 提示能够发挥正向作用。

Table 6: Results for apartment selection task across four models and three ranges of  $\Delta$ .

$\Delta$	[0.1, 0.3]		[0.3, 0.5]		[0.5, 1]	
	Zero-shot	CoT	Zero-shot	CoT	Zero-shot	CoT
GPT-4o	47%	45%	57%	56%	80%	87%
Claude 3.5 Sonnet	50%	62%	62%	72%	81%	95%
Claude 3 Opus	35%	50%	57%	58%	72%	84%
Llama 3.1 70B Instruct	42%	6%	44%	5%	43%	20%

图 | 四种模型和三种范围内的公寓选择任务结果。 (来源: 该论文)

## 不足与展望

当然, 该研究也存在一些局限性, 如下:

inference-time 推理的类型。自从 CoT 提示被提出以来, 研究人员开发了多种特定于应用领域的提示方法, 以及更复杂的多次前向传递的通用提示方法, 如思维树 (tree-of thought) 和自一致性 (self-consistency) 。他们在 GPT-4o 模型上测试了思维树方法在隐式统计学习任务中的有效性, 发现其确实提高了分类准确率 (64.55% vs. 62.52%) , 但仍远低于零样本推理的 94.00% 准确率。未来的研究仍需探索此方法是否可以适用于其他任务领域和模型中激发语言思维的方法。

应用范围。尽管这一研究基于心理学的启发式方法提供了一种识别 CoT 失败案例的策略, 但这无法涵盖所有可能导致 CoT 表现下降的情况。现有的心理学研究基于多种理论和实际考量来研究人类, 并不能提供涵盖所有任务的详尽或代表性样本, 且会遗漏一些仅在模型中具有研究价值的特殊案例。

关于 CoT 未能复制人类结果的替代解释。对于 CoT 在后面三个任务中没有观察到表现下降, 存在一种替代解释——在 LLM 中实现这些任务的方式消除了表现下降的效果。虽然研究对后三个任务情境进行了多种变体的探索, 但由于提示的变化几乎是无穷无尽的, 这些探索并不详尽。

研究团队表示，虽然该研究聚焦于 CoT 推理，但所提出的框架为利用人类心理学研究评估和改进模型表现提供了一种通用策略。

他们认为，未来还需要更多的跨学科合作，通过将自然语言处理方法、心理学见解与人类和模型表现比较的相关研究相结合，可以形成更全面的 AI 评估和改进策略。

(文章来源科普中国 [https://www.kepuchina.cn/article/articleinfo?business\\_type=100&classify=0&ar\\_id=555995](https://www.kepuchina.cn/article/articleinfo?business_type=100&classify=0&ar_id=555995))

## 学会动态

### 第八届江苏省青少年信息机器人科技素养实践活动城市赛区选拔工作落下帷幕”

2024 年 8 月由全国高等院校计算机基础教育研究会、全国工商联教育商会、江苏省科学技术协会指导，江苏省计算机学会主办，第八届江苏省青少年信息机器人科技素养实践活动暨白名单赛事预选观摩交流活动，在全省淮安、连云港、苏州、常州、无锡、宿迁、扬州、泰州、盐城、徐州、南通、南京等城市圆满落下帷幕。

本次大赛以“数字教育 科创未来”为主题，吸引了省内近上万青少年积极参与，同时吸引了近 300 多万人次的观众通过直播观看比赛盛况。百余名经验丰富的教练老师参与了此次科普活动，为青少年选手们提供了宝贵的经验和辅导。活动还得到了相关省市级领导、专家的深切关怀和专业指导。





# 多维无源感知关键技术及应用

## ——2024 年江苏省计算机学会科学技术奖

项目名称：多维无源感知关键技术及应用

完成单位：南京大学，苏州大学

项目简介：

当前，基于泛在信号的无源感知技术已成为智能化社会发展的重要支撑。该技术利用环境中广泛存在的无线信号（如 WiFi、5G、LoRa 等）实现目标的无源检测与感知，具备低成本、护隐私、范围广、非侵扰等优势，特别是集成感知与通信（ISAC），在物联网、智能家居和自动驾驶等领域获得了广泛关注。具体而言，基于泛在信号的无源感知技术通过分析信号传播的变化，感知环境中的物体及其动态行为；近年来随着通信技术的快速发展，相关技术研究取得了长足进展。项目组在该领域持续开展近十年深入研究，重点探索如何通过信号的多节点感知、多域的统一处理、及多模态融合，进一步提升无源感知的精度和鲁棒性。项目组的研究成果不仅在理论上取得了突破，也在实际应用中得到了验证。

在多节点精确感知方面，项目组研究了利用超声波声学信号、WiFi 信号、毫米波射频信号等无线信号进行手势跟踪、呼吸检测、头部运动跟踪及唇语识别等多节点高精度感知任务。其中，项目组研究了利用商业音频播放 / 录音设备实现高精度温度场测量，实现了环境温度分布监测。项目组使用分布式设备覆盖较大的室内区域，并为系统设计了一种分布式测距算法，使用多个接收节点以微秒级的分辨率测量多条声波路径的飞行时间。然后，提出了一种 dRadon 变换算法，从沿不同路径测得的声速中重建温度分布。

在多域统一感知方面，项目组研究了结合多个信号域实现的统一感知技术，提升了无源感知的感知范围及感知精度。结合相位域和时频域的无设备手势跟踪技术，实现了毫米级的手指跟踪精度。同时针对声学传感感知范围有限且测试用户单一的问题，开发了同时使用空间域波束形成和频域波束形成的 Dualforming 算法，大大扩展了声学传感范围，并且能够同时感知多达 6 个目标的心跳数据。

在多模态融合感知方面，项目组积极探索不同模态数据之间的互补性，获取不同方面的环境或目标信息，并设计了新颖的感知系统来验证所提出的理论。提出了一种细粒度的手指运动追踪方案，使用基于轻质超声波的传感，以及一个商用单目摄像头，以实现对用户手指的 3D 跟踪。单目摄像头在二维空间中跟踪用户的手指，同时利用基于超声波的传感技术获取用户手指在三维空间中的深度信息，最终得以准确跟踪三维空间中手指运动。此外，研究了利用无线耳机内置磁铁的磁场信息辅助来进一步提高目标的跟踪和声学传感性能。这两种模态的融合有效抑制了累积时钟偏移误差和多路径噪声干扰，使得跟踪误差降低到毫米级。

项目组在该领域深耕多年，相关成果发表在 MobiCom、UbiComp、INFOCOM、MobiSys、SECON、TMC、

TNET 等国际一流会议期刊上，包括 CCF A 类刊物 37 篇，获得了 ACM Fellow、IEEE Fellow、AAAS Fellow、美国国家科学院院士、美国国家工程院院士、哥德尔奖得主、欧洲科学院院士等国内外知名专家学者的正面引用和评价，其中主要代表作 SCI 总他引 778 次，谷歌学术总他引 1204 次。相关成果曾获得 CCF A 类会议 UbiComp 杰出论文奖（Distinguished Paper Award），CCF B 类会议 SECON 最佳论文奖亚军（Best Paper Award Runner-up），CCF C 类会议 WASA 最佳论文奖（Best Paper Award）。同时，部分成果入选 ESI 高被引论文。项目组在相关领域已申请发明专利 10 项，授权 4 项。此外，项目组还参与制定了 IEEE 802.11bf WiFi 无线感知通信标准。

依托上述研究成果，项目组成功获得多项国家及省部级科研项目资助，包括国家自然科学基金项目四项、江苏省基础研究计划（自然科学基金）项目资助一项、江苏省高等学校基础科学（自然科学）项目一项以及企业合作项目四项。同时，理论联系实际，项目组与华为、海思、国家电投集团、南京脑科医院等行业领先单位开展深度合作，例如，国电投江苏新能源公司通过引进项目组研发技术，在马鞍山、滨海等新能源场站部署了项目组研发的基于无源感知的智能监测和预警系统，近年累计增加利润 3922.01 万元。

## 主要科技创新

基于泛在信号的无源感知是一种新兴的、具有重要应用价值的技术，这种技术能够在不依赖视觉信息的情况下，通过环境中的自然信号或现有的信号源对环境变化进行精确感知和检测。该技术通过巧妙利用已有信号的反射、折射、散射等特性，实现对目标物体或环境状态的感知和监测。近年来，无源感知技术日益受到广大学者的关注，并得到了广泛的研究和应用。特别是在信息技术较为发达的美国、日本、加拿大和欧洲等国家和地区，众多学者在无源感知的应用研究等方面开展了系统而深入的研究。

无源感知技术的关键问题包括多节点协同感知、多域信号处理、以及多模态数据融合等方面。其核心问题是如何在不依赖视觉信息的情况下，精确地获取和处理来自多节点、多域、多模态的感知数据，从而实现对复杂环境中目标的高效、可靠感知。这需要解决信号弱化、干扰抑制、数据融合与分析等一系列技术挑战，并最终形成对环境或目标的科学判断。

无源感知技术的研究不仅在应用领域具有重要价值，同时在理论上也具有深远的科学意义。无源感知技术的发展历程可以追溯到早期的雷达反射和无线电波传播研究，而随着信息技术的发展，尤其是无线通信和信号处理技术的进步，无源感知逐渐演变为一个独立且多学科交叉的研究领域。研究者们在多节点感知、跨域数据融合、信号处理等方面不断取得突破，形成了一系列创新性成果，推动了无源感知技术的迅速发展。

无源感知技术的应用不仅限于特定的技术领域，而是对社会各个层面都产生了积极影响。首先，无源感知技术在公共安全领域的应用，如智能监控、灾害预警等，有效提升了社会的安全保障水平。其次，在智能交通、智慧城市等领域的应用，提高了资源利用效率，推动了社会的可持续发展。再次，在环境监测、健康医疗等领域，无源感知技术提供了低成本、高效益的解决方案，有助于社会福祉的提升。近年来，项目组在多节点精确感知、多域统一感知、多模态融合感知等方面进行了系统、深入和细致的研究，取得了一系列成果，具体为：

### 主要科技创新 1: 多节点精确感知

多节点精确感知是一种结合多个分布式节点感知源的信号处理方法，通过对更多维度传感器信号的获取，实现



在动态和复杂环境中维持系统的高感知精度。在人际交互领域，项目组提出结合 Zadoff-Chu 序列和频分复用技术，实现了高精度的 3D 头部方向追踪；在医疗健康感知领域，项目组提出利用 ZC 序列和信道脉冲响应（CIR）技术，实现多用户呼吸信号识别；在环境感知领域，项目组提出了一种创新的分布式测距算法和 dRadon 变换算法，在分米级空间分辨率下实现了温度分布重建；在结构感知领域，项目组设计了一种捕捉不同介质传播信号的分析方法，为移动设备提供了全新的交互体验。多节点感知通过更多维度的信号处理技术、高精度的时间同步和节点定位技术，确保多个分布式感知节点在同一时间基准下工作，从而实现数据的精确同步与节点定位，提高各节点之间的协同效率，避免数据冗余和时序误差。这种方法不仅扩展了系统的空间覆盖范围，还提升了感知的灵敏度和响应速度，使得系统在面对更多、更复杂的环境变化时具有更强的适应性和可靠性（代表性论文 4）。

在人机交互领域的研究中，考虑到头部方向追踪在在线课程和体感游戏等多个领域具有广泛的潜在应用。通过获取用户的头部方向信息，这些应用将有更多机会提升性能并提供更好的用户体验。然而，现有的头部追踪研究，基于计算机视觉（CV）的方法存在跟踪角度范围有限和隐私问题，而基于惯性测量单元（IMU）的方法则存在累积误差。这些方法都无法提供准确和稳定的用户头部方向信息。项目组设计并实现的系统 [1] 通过在普通耳机上布置多个超声波收发器，实现了 3D 头部方向追踪。如何解决多个发射节点之间发射信号的同步是一个挑战。项目组使用 Zadoff-Chu 序列作为基带信号，将其调制到超声波频段作为发送信号，并在此基础上设计了一种频分复用方法，实现了两个发射器的同时定位。项目组使用神经网络设计了一种特殊的头部方向追踪算法，基于头部运动识别，仅使用头部的两个坐标实现大约 6 度的头部方向追踪。

在医疗健康感知领域的研究中，呼吸是重要的生命体征之一，但是临床仪器（如呼气二氧化碳监测仪或容积描记仪）需要专业操作人员，无法在家庭环境中进行长期监测。而长期监测对早期诊断慢性疾病（如阻塞性睡眠呼吸暂停综合症和慢性阻塞性肺病）至关重要。因此，近年来家庭连续呼吸监测系统的发展迅速。然而，用户研究表明，由于隐私问题或高成本和长期身体接触的要求，人们不愿意使用基于摄像头或使用特殊设备（如集成电容传感器的腰带或带有气压传感器的智能坐垫）的家庭呼吸监测系统。项目组设计并实现了 RESPTRACKER[2]，这是第一个在家庭环境中使用基于声学的商业现成设备进行的连续多人的呼吸追踪系统。为了更好的分离各个用户的信号，项目组使用 ZC 序列以高于 10 厘米的分辨率区分不同的声反射路径，并通过信道脉冲响应（CIR）测量各个接收节点信号传播路径的特征，减少了每条路径上其他对象干扰，提出了多维信号组合方案来选择和重组用户的呼吸信号，提高了感知用户呼吸信号的精度。评估结果表明，在单用户场景中，系统能够在不同环境下以低于 0.6 次每分钟（BPM）的误差稳健地估计呼吸频率。在三米的距离内，RESPTRACKER 的误差小于 1 BPM，并且在用户移动时保持低于 0.8 BPM 的误差。在多用户场景中，RESPTRACKER 能够分离同一房间内超过四个用户的呼吸信号，并为每个用户实现低于 1 BPM 的误差。

在环境感知领域中，环境温度分布监测在室内温度控制和建筑能量管理等多种实际应用中非常重要。项目组开发和设计了 VECTOR[3]，一个温度场监测系统，利用商业音频播放 / 录音设备实现高温度感知精度和快速响应时间。传统温度传感器测量的是传感器探头的温度，由于热传导过程，它可能与空气温度不同。而声学信号的速度由空气温度物理决定，几乎没有延迟。在许多场景中，由于空调、阳光或数据中心服务器的工作负载，空气温度可能会迅速变化，而这种变化可以通过声学感知捕捉到。此外，声学感知还支持在较长距离内测量平均温度，并可以使用分布式设备覆盖较大的室内区域。相比之下传统传感器只能捕捉特定点的温度，通常放置在墙壁或靠近屋顶的地方，

远离目标区域。项目组为系统设计了一种分布式测距算法，使用多个接收节点以微秒级的分辨率测量多条声波路径的飞行时间。然后，项目组提出了一种 dRadon 变换算法，从沿不同路径测得的声速中重建温度分布。实验结果表明，项目组可以通过单条声波路径测量温度，误差为  $0.25^{\circ}\text{C}$ ，并在分米级空间分辨率下重建温度分布。相关成果荣获 CCF A 类会议 UbiComp 杰出论文奖（Distinguished Paper Award）。

在结构感知领域中，使移动设备的所有表面都能够感应触摸手势，而不仅限于触摸屏区域，可以带给用户全新的交互体验。新的非显示屏触摸手势可以解决触摸屏遮挡问题。例如，背面手势（BoD）使用在智能手机背面敲击或滑动作为补充输入界面。当使用背面滚动手势滚动内容时，屏幕不再被遮挡，还能允许游戏玩家使用背面作为触摸板。项目组设计并实现了 VSkin[4]，一个基于声学信号的系统，支持在移动设备背面进行精细的手势感应。VSkin 捕捉结构传播的声音（即通过设备结构传播的声音）和空气传播的声音（即通过空气传播的声音）。由于触摸表面会显著改变设备的结构振动模式，因此结构传播声音的特征是触摸检测的可靠特征，即手指是否接触表面。因为单一麦克风能够接收到的反射路径有限，为了能够得到细粒度、多方向的手指移动信息，项目组利用了商业移动设备上的多个麦克风通道，实现多传输路径的感知。通过测量每一条路径接收信号的相位和幅度，VSkin 得以捕捉快速的滑动和敲击事件。实验结果表明，VSkin 实现了 99.65% 的触摸检测准确率和 3.59 毫米的手指移动距离准确率。

进一步，项目组探索了利用结构传播的声音（即通过移动设备结构传播的声音）来识别移动电话背面的滑动手势。设计和实现了 StruGesture[5]（代表性论文 4），全面检测各个麦克风通道的接收信号，并提出了一种基于峰值选择算法来提取混合通道脉冲响应（CIR）中的结构传播成分。项目组还开发了一种深度对抗学习架构，以学习手势特定的表示，从而对移动设备背面的滑动手势进行稳健和有效的识别。项目组进行了大量实验来评估在九种部署场景下的鲁棒性。结果表明，StruGesture 对 10 种手势平均识别准确率达到 99.5%，优于现有最先进分类器。

[1]Song J, Dai H, Shi S, et al. HeadTracker: Fine-Grained Head Orientation Tracking System Based on Headphones[C]//International Conference on Wireless Algorithms, Systems, and Applications (WASA). Cham: Springer Nature Switzerland, 2022: 317-329. (CCF C)

[2]Wan H, Shi S, Cao W, et al. RespTracker: Multi-user room-scale respiration tracking with commercial acoustic devices[C]//IEEE INFOCOM 2021-IEEE Conference on Computer Communications (INFOCOM). IEEE, 2021: 1-10. (CCF A)

[3]Wan H, Wang L, Zhao T, et al. VECTOR: Velocity based temperature-field monitoring with distributed acoustic devices[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2022, 6(3): 1-28. (CCF A)

[4]Sun K, Zhao T, Wang W, et al. VsSkin: Sensing touch gestures on surfaces of mobile devices using acoustic signals[C]//Proceedings of the 24th Annual International Conference on Mobile Computing and Networking (MobiCom). 2018: 591-605. (CCF A)

[5]Wang L, Zhang X, Jiang Y, et al. Watching your phone's back: Gesture recognition by sensing acoustical structure-borne propagation[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2021, 5(2): 1-26. (CCF A)



## 主要科技创新 2: 多域统一感知

多域统一感知是一种将单一类型的传感器感知数据，通过多域变换技术在多个域中展开，分别提取每个域中的独特特征，并进行统一分析与融合的信号处理方法。在人际交互领域，项目组提出使用信号的相干性以及信号的信道脉冲响应 CIR，实现毫米级精度的无设备手势跟踪和动作的高精度识别；在医疗健康感知领域，项目组通过创新的双重形成（Dualforming）技术，扩大了声学感知的应用范围；在工业设备感知领域，项目组提出使用信号谱的特征金字塔来进行特征匹配，实现了旋转物体瞬时角速度（IAS）的非接触式测量；在安全认证感知领域，项目组结合了口腔和声带运动特征，实现了对语音认证攻击的高精度识别。通过分析时频域、空间域、相位域、信道响应域的信息，提高系统对复杂信号的处理能力，扩展感知系统在多域、多维度上的应用潜力，使得系统能够在多样化场景中保持高效、精准的感知能力（代表性论文 1、2、5）。

在人机交互领域的研究中，人类活动的识别对于人机交互、潜在危险识别、运动健康检测等领域有着重要意义。传统基于 WiFi 信号的人类活动识别系统，它们的主要局限性在于缺乏一种能够定量关联 CSI 动态变化 [6,7] 和人类活动的模型。提出并实现了 CARM[8]（代表性论文 1），一种基于 CSI 的人类活动识别和监测系统。对于 CARM，项目组提出了两个理论基础：一个 CSI- 速度模型，量化了 CSI 值动态变化与人类运动速度之间的相关性；一个 CSI- 活动模型，量化了不同人体部位运动速度与特定人类活动之间的相关性。通过这两个模型，项目组定量建立了 CSI 值动态变化与特定人类活动之间的相关性。CARM 使用 CSI 域中的这种相关性作为剖析机制，通过将给定活动与最佳拟合剖析进行匹配来识别该活动。结果表明，CARM 在几种不同环境中的活动识别平均准确率超过 96%。

无设备手势跟踪是一种增强人机交互（HCI）机制，因为手指太大，无法控制如此小屏幕上的图形用户界面元，所以适用于小型可穿戴设备。项目组设计并实现了 LLAP[9][6]，利用大多数移动设备上已有的扬声器和麦克风来实现对手的无设备跟踪。传统基于声音的测距系统要么使用到达时间 / 时间差（TOA/TDOA）测量，要么使用多普勒频移测量，它们的距离测量精度通常在厘米级。为了实现毫米级的手 / 指跟踪精度，项目组利用人手反射的声音与移动设备发出的声音之间的相干性，通过相干检测器将接收到的声音信号转换为复值基带信号。项目组提出一种结合复值基带信号幅度域和相位域的静态分量去除算法，并将信号的相位变化转换为人手运动的距离。结果表明，对于一维手部运动和二维空中绘图，LLAP 的跟踪精度分别为 3.5 毫米和 4.6 毫米。使用 LLAP 跟踪的手势轨迹，项目组可以以 92.3% 和 91.2% 的准确率识别空中绘制的字符和短词。

随着增强现实（AR）/ 虚拟现实（VR）技术的兴起以及移动设备的小型化，手势识别在人机交互领域变得越来越重要。一些基于超声波的手势识别系统已经被提出。然而，它们大多依赖于低分辨率的多普勒效应，主要关注整个手的运动，无法处理细微的手指动作。项目组设计并实现了 UltraGesture[10]（代表性论文 2），一种基于通道脉冲响应（CIR）[11] 的超声波手指运动感知和识别系统。CIR 测量可以提供 7 毫米的分辨率，足以识别细微的手指动作。UltraGesture 将 CIR 域的测量结果封装成图像，并构建卷积神经网络模型将这些图像分类为不同类别，以对应不同的手势。UltraGesture 可以在大多数移动设备上运行现有的商业扬声器和麦克风，无需任何硬件修改。结果表明，UltraGesture 可对包括手指点击和旋转在内的 12 种手势实现超过 99% 的平均准确率，相关成果荣获 CCF B 类会议 SECON 最佳论文奖亚军（Best Paper Award Runner-up），并获得全国高校物联网应用创新大赛一等奖及最佳指导教师奖。

在医疗健康感知领域的研究中，心跳监测的声学传感已成为无线传感领域的一个热门研究主题 [12,13]。现有的声学传感系统存在两个局限性——感应范围有限，且只能对单一用户进行心跳监测，这阻碍了应用的大规模部署。项目组提出并实现了 DF-Sense[14]，一种基于双重形成（Dualforming）的多用户声学传感系统，旨在家庭环境中进行心跳监测。为了能够感知多用户所在位置，项目组提出了算法，利用来自连续帧的动态接收信号，增加 MUSIC 算法 [15] 中相关矩阵的秩，以检测多个受试者。在已经了解多个受试者位置的条件下，提出了 Dualforming 算法，利用空间域波束形成和频域波束形成，推导出与接收阵列中每个麦克风的子载波相对应的相位，然后通过补偿多个子载波和麦克风之间的相位差异来创建构造叠加。Dualforming 可以将传感信号与噪声比（SSNR）提高  $MN$  倍（其中  $M$  和  $N$  分别是麦克风和子载波的数量），大大扩展了声学传感范围。实验结果表明，DF-Sense 在 10 米范围内实现了瞬时心率的高精度测量，满足大多数日常需求，且能同时监测多达 6 个对象的心跳。

帕金森病（PD）是一种脑部疾病，可能导致运动症状，包括行走困难、运动迟缓、身体失衡和震颤。行走困难是帕金森病的典型症状，帕金森步态的识别是统一的帕金森病评定量表（UPDRS）中重要的一个子项。为了能够以非接触和保护隐私的方式测量帕金森病患者的步态参数，项目组提出了 PD-Gait[16]。项目组是第一个使用不可听的声学信号测量帕金森病步态的工作。为了实现稳健的测量，项目组结合多普勒频移和距离谱的信息，提出了一种新颖的声学测距方法，以避免接收数据中的“断音”和“峰值分布不均”；项目组还提出了自动聚焦微多普勒特征，以提取稳健的步幅周期时间长度。验证结果表明，PD-Gait 在周期时间长度和测距方面分别可以达到 0.052 秒和 0.1 米的精度。项目组在一家脑科医院的老年神经科部署了 PD-Gait，最终测量结果与医生的判断高度相关，能够有效检测患者在服药前后步态的细微变化。

进一步的，项目组设计并实现了 mP-Gait[17] 系统，一种将步态分析与毫米波技术相结合的帕金森病情评估方法，主要针对现有临床诊断中医生通过如 HY 分级和 UPDRS 评分衡量帕金森步态情况的方法，存在精准度低、误差大、容易受到诊断医生主观影响等问题，展开了深入研究。旨在为医生提供患者的量化步态特征数据，并提供预测的 UPDRS 步态分数来辅助诊断，该系统使用毫米波雷达阵列计算得到的步长误差平均为 7.16 厘米，对于 UPDRS 步态分数的预测误差为 0.36 分。该系统于 2021 年 6 月起，应用于南京脑科医院的多个门诊、会议室和患者病房等共 4 类不同实际场景，收集了 243 位帕金森病人的步态数据，并在中国大学生计算机设计大赛中获得江苏省特等奖和全国二等奖。此外，和南京脑科医院成功申报南京市卫生科技发展专项资金重点项目“基于无源感知技术的新型帕金森病诊疗模型研究”。

在工业设备感知领域的研究中，测量旋转物体的瞬时角速度（IAS）在工业和日常生活中无处不在。工程师通过 IAS 诊断发动机的运行状态。风速计利用旋转杯的 IAS 获取瞬时风速。传统的 IAS 测量系统在安装、精度和成本等方面存在局限性。项目组设计并实现了 PCIAS[18]（代表性论文 5），利用智能手机的声学信号以非接触方式测量旋转物体的 IAS。因为不同场景下对 IAS 的测量要求不同，从每分钟 10 转（RPM）到 10000 RPM，项目组结合了声音信号多个域中的特征来保证 PCIAS 可以适用于各种 IAS 测量范围。项目组使用距离特征域上的周期来测量低 IAS 的物体，使用多普勒特征域的周期来测量中 IAS 的物体，使用噪声特征域的周期来测量高 IAS 的物体。为了能够得到各种不同几何形状的旋转物体 IAS，项目组提出了一种基于信号谱图的特征金字塔匹配算法，使得 PCIAS 适用于不同形状的旋转物体的同时，其性能优于几乎所有现有的商业 IAS 测量设备。实验结果表明，PCIAS 在低 IAS 范围内的相对准确度超过 92%，在中 IAS 范围内超过 94%，在高 IAS 范围内超过 96%。此外，该工作在知乎等公



众媒体获得超过 16 万的阅读量，吸引了众多相关从业人员关注和交流。

在安全认证感知领域的研究中，由于不同人的声道形状和语调不同，语音已被用作人类认证的生物特征。然而，传统基于语音的人类认证容易受到四种攻击：模仿、语音转换、合成和语音重放。项目组设计并实现了 SpeakPrint[19]，一种基于超声波的智能手机人类语音认证方案，可抵御这些攻击。与传统语音认证系统不同，SpeakPrint 通过同时记录超声波信号中的口腔和声带运动，捕捉用户说话的方式。项目组认为，对于正常用户来说，从语音信号中提取的特征应该与从超声波信号记录的口腔和声带运动保持一致，而模仿者或音频播放器无法在超声波领域产生相同的信号。项目组设计并训练了一个 SVM 分类器，综合考虑从正常语音域中提取的 MFCC 特征和从超声波域中提取的 MMSI 特征，来实现对语音认证攻击的识别。实现了 SpeakPrint，并在 40 名用户身上进行了实验。实验结果表明，SpeakPrint 可以 100% 准确地检测重放攻击，对于超过 5 个单词的密码短语，可以 99.12% 准确地检测带有嘴唇同步的重放攻击。这项技术可用于多因素认证系统，在该系统中使用多种认证机制来实现深度防御。

[6]Wang W, Liu A X, Shahzad M. Gait Recognition Using WiFi Signals[C]//Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing (UbiComp). 2016: 363-373. (CCF A)

[7]Yu N, Wang W, Liu A X, et al. QGesture: Quantifying Gesture Distance and Direction with WiFi Signals[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2018, 2(1): 1-23. (CCF A)

[8]Wang W, Liu A X, Shahzad M, et al. Understanding and Modeling of WiFi Signal based Human Activity Recognition[C]//Proceedings of the 21st annual international conference on mobile computing and networking (MobiCom). 2015: 65-76. (CCF A)

[9]Wang W, Liu A X, Sun K. Device-free Gesture Tracking Using Acoustic Signals[C]//Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom). 2016: 82-94. (CCF A)

[10]Ling K, Dai H, Liu Y, et al. Ultragesture: Fine-grained Gesture Sensing and Recognition[J]. IEEE Transactions on Mobile Computing (TMC). 2020, 21(7): 2620-2636. (CCF A)

[11]Wang L, Sun K, Dai H, et al. WiTrace: Centimeter-level Passive Gesture Tracking Using OFDM Signals[J]. IEEE Transactions on Mobile Computing (TMC). 2019, 20(4): 1730-1745. (CCF A)

[12]Wang L, Wang X, Zhang D, et al. Knowing Your Heart Condition Anytime: User-Independent ECG Measurement Using Commercial Mobile Phones[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2023, 7(3): 1-28. (CCF A)

[13]Wang L, Huang K, Sun K, et al. Unlock with Your Heart: Heartbeat-based Authentication on Commercial Mobile Phones[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2018, 2(3): 1-22. (CCF A)

[14]Wang L, Gu T, Li W, et al. Df-sense: Multi-user Acoustic Sensing for Heartbeat Monitoring with Dualforming[C]//Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services (MobiSys). 2023: 1-13. (CCF B)

[15]Wang L, Li W, Sun K, et al. LoEar: Push the Range Limit of Acoustic Sensing for Vital Sign Monitoring[J].

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2022, 6(3): 1-24. (CCF A)

[16] Li Z, Pan Y, Dai H, et al. PD-Gait: Contactless and Privacy-Preserving Gait Measurement of Parkinson's Disease Patients Using Acoustic Signals[J]. Software: Practice and Experience (SPE). 2024, 54(9): 1733-1753. (CCF B)

[17] Zhang W, Dai H, Xia D, et al. mP-Gait: Fine-grained Parkinson's Disease Gait Impairment Assessment with Robust Feature Analysis[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2024, 8(3): 1-31. (CCF A)

[18] Li Z, Dai H, Wang W, et al. Pcias: Precise and contactless measurement of instantaneous angular speed using a smartphone[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2018, 2(4): 1-24. (CCF A)

[19] Dai H, Wang W, Liu A X, et al. Speech based human authentication on smartphones[C]//2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). IEEE, 2019: 1-9. (CCF B)

### 主要科技创新 3: 多模态融合感知

多模态融合感知是一种先进的信息处理方法，它结合来自不同类型感知源或信号模式的数据，以实现对目标或环境的全面、准确理解。项目组提出在 3D 目标追踪中，使用超声波信号捕捉摄像头无法提供的深度信息；提出利用磁场信号弥补声学信号的不足，克服了累积时钟偏移误差和多路径噪声干扰。通过整合多种传感器（如视觉、音频、超声波、磁场等）的数据，多模态融合感知克服了单一感知模式局限性，提高了系统的可靠性和鲁棒性（代表性论文 3）。

首先，项目组提出了一种细粒度的手指运动追踪方案，使用基于轻质超声波的传感，以及一个商用单目摄像头，以实现对用户手指的 3D 跟踪 [20]（代表性论文 3）。该系统利用单目摄像头在二维空间中跟踪用户的指，同时利用基于超声波的传感技术获取用户手指在三维空间中的深度信息，并从超声信号的相位变化，准确地测量手指在深度方向上的运动。通过开发快速、轻量级的信号处理算法，该系统可以准确跟踪手指运动，并测量手指在两个视频帧之间的弯曲角度。实验结果表明，所提方案达到了 98.4% 的手指敲击检测精度，检测延迟为 17.69 ms，比纯视频方案少 57.7 ms。

除此之外，项目组观察到无线耳机是一种普遍存在的声学传感平台，可用于许多应用程序，如运动跟踪和手写输入。然而，无线耳机在所连接的智能设备之间遭受时钟偏移，这将随着时间的推移迅速积累错误。此外，与智能手机和语音助手相比，无线耳机传输的声信号由于频率响应差而要弱得多。因此，为了解决这一问题，提出了 MagSound [21]，它使用无线耳机内置磁铁辅助来提高商用耳机的跟踪和声学传感性能。利用磁场强度，MagSound 可以预测无线耳机的位置，并且不受时钟偏移。这可以用来重新校准声学跟踪。此外，这两种模态的融合有效抑制了累积时钟偏移误差和多路径噪声干扰。同时，为了进一步提高对噪声的鲁棒性，MagSound 采用了精心设计的正交频分复用（OFDM）信号进行测距跟踪。实验结果表明，MagSound 在二维跟踪中保持了毫米级误差，并将手写识别准确率提高了 49.81%。



[20]Sun K, Wang W, Liu A X, et al. Depth Aware Finger Tapping on Virtual Displays[C]//Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys). 2018: 283-295. (CCF B)

[21]Wang L, Wang W, Dai H, et al. MagSound: Magnetic Field Assisted Wireless Earphone Tracking[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp). 2023, 7(1): 1-32. (CCF A)

## 社会效益（限 200 字）

国电投江苏新能源公司通过引进南京大学研发的先进技术在马鞍山、滨海等新能源场站采用了多节点分布式的温度场检测系统，实时监测设备运行状态，提高了安全隐患的预警能力。同时，公司还应用了人类活动识别和检测技术，以及人体动作感知技术，大幅提升了对工人作业行为的监测和分析能力，有效防范了人员伤害事故的发生。这些技术创新不仅提高了安全管理水品，也为企业节省了大量的人力、物力投入。近二年累计增加利润 3922.01 万元，充分体现了技术创新在提高经济效益和社会效益方面的重要作用。

## 学会动态

### 江苏省计算机学会携手微软在宁举办 AOAI 高校人工智能师资培训

8月25日至26日，由江苏省计算机学会主办，微软（中国）有限公司和赛尔教育科技发展有限公司承办的“江苏省计算机学会 - 微软 AOAI 高校人工智能师资培训”在南京大学仙林校区圆满落幕。此次培训旨在积极响应国家关于复合型人才培养的号召，推动高校人工智能教育深入发展。



# 基于 AI 的社会安全态势感知平台

项目名称：基于 AI 的社会安全态势感知平台

完成单位：南京擎天科技有限公司

项目简介：

一、项目主要技术内容

## (一) 项目目标

围绕城市运行“一网统管”和城市数字化治理模式，基于监测预警数据、网格化治理机制等，聚焦城市安全管理的主要问题和突出风险，充分利用科技手段提高社会安全态势感知和应急处置能力，依托多部门多灾种的智能化灾害监测网络和城市燃气、桥梁、承灾体、风险隐患等基础信息，基于云计算、大数据、移动互联网、人工智能等新技术、新方法，建设基于人工智能的社会安全态势感知平台，实现预警大数据的智能化采集、整理、分析、计算、服务，形成较为完善的社会安全态势感知与运营模式。

## (二) 技术应用场景

1、经营性自建房监测预警

将安全风险从高到低划分为四个等级，超过安全阈值的，进行安全风险分级分类预警。

2、气象灾害靶向预警

运用全国自然灾害承灾体调查数据，对短临暴雨发生风险的评估预警。

3、商业综合体监测预警

对排油烟道设计、燃气泄漏、电气设备安全故障等场景，通过物联感知、自查自纠，结合研判分析，科学辨识商业综合体餐饮场所安全风险，从源头预防和减少突发事件的发生。

4、高层住宅（老旧小区）监测预警

利用电气安全监测，向事前预知预判为重点的精准预防管理升级。

## 二、授权专利情况

基于 AI 的社会安全态势感知平台在建设过程中产生与平台相关的多项专利成果。

## 三、技术经济指标

项目建设完成后，已逐步向多个涉及应急领域的客户推广，市场潜力巨大。

## 四、应用推广及效益情况

### (一) 应用推广



项目在多个行业、部门孵化形成应用推广。

## (二) 项目效益

### 1、经济效益

#### (1) 提高行政管理效能，降低政府行政成本。

该平台有效提高城市的社会安全态势感知能力，积累大量基于 AI 的预测预警模型，促进行业部门之间的高效协同，提高处置效率，降低行政管理成本。

#### (2) 提高风险防范能力，降低经济损失。

该平台提升城市社会安全综合监测预警能力，努力用最小代价实现最大的风险防范效果，提高风险隐患处置工作的有效性，在降低突发事件所造成社会经济损失方面将发挥重要作用。

### 2、社会效益

#### (1) 提升治理效能和安全管理能力。

针对现阶段社会安全综合监测预警存在的不足，通过管理创新与信息化技术相结合，相较于原有依靠大量人力完成的任务，以 AI 优化能力为核心，大幅提升治理效能和安全管理能力。

#### (2) 有助于保障群众生命财产安全，增强群众安全感。

可大幅提升城市应对较大、重大、特别重大突发事件的预防预警、资源调度、响应处置等方面的工作能力。

## 主要科技创新

### 一、项目立项背景

近年来，社会安全新兴风险、传统产业风险等积聚滋生、易发多发，一些城市相继发生重特大影响恶劣的社会安全事故（灾害），然而传统的风险预警以单一变量预警方式为主，在科学性、准确性、高效性上有待提升，亟需借助人工智能等新一代信息技术，对城市的社会安全重点领域进行综合量化评估。

聚焦社会安全管理的主要问题和突出风险，充分利用科技手段提高社会安全综合监测预警和应急处置能力，依托多部门多灾种多风险的智能化监测网络和城市燃气、桥梁、承灾体、风险隐患等基础信息，建设基于 AI 的社会安全态势感知平台，从城市生命线工程安全、公共安全、生产安全、自然灾害四个板块分专题开发跨部门的社会安全风险耦合预警场景，形成较为完善的社会安全态势感知与运营模式，基本实现“能监测、会预警、快处置”三个功能，努力实现社会安全运行“一网统管、智能化联动处置”，构建系统立体的“前置防线”。

### 二、关键技术能力

#### (1) 底层技术框架

该平台采用擎天 SkyEmpower 数字建模一体化平台，该平台构建了一套具有业务属性的智能模型全周期管理和复用体系，重新定义数据价值的应用方式，通过数据模型的作用和成果进行引导，明确模型需求从而改变传统的建模流程和方式；通过设计模型算法的柔性指标，满足政企单位个性化需求的同时实现模型的加速进化，提升数据服务的敏捷性和效率；通过突出业务能力的思路和经验的延伸，将成果生态化实现横向拓展，从而打造业务模型的共享复用机制。

#### (2) 跨境追踪技术

在对人员的识别中，相比原先只对人脸特征进行识别，项目应用了跨境追踪技术，增加了对人员体态、衣着特征的识别，包括是否背包、衣服颜色、款式等信息。进一步聚合形成对人员特征的完整描绘和动态追踪，多模态提升人员识别的准确性和全面性。

在车辆识别中，除了车牌号码识别，利用车辆识别算法实现了对车辆的二次结构化，识别出车身颜色、车牌颜色、车辆类型、车辆品牌等信息，提升车辆抓拍不清晰情况下的车辆识别能力。

同时，基于对人和车的特征识别，通过对视图库结构的升级优化及感知数据的融合分析，实现人员、人脸、人体与车辆的全面关联，基于关联后的数据建设了实战性强、数据准确的感知应用。

### (3) 论文发表情况

该平台设计核心围绕人工智能、智慧城市、风险隐患、数字模型为核心进行设计，参考并借鉴了公司多项论文。

### 三、国内同类技术对比

相对于国内目前传统数字建模平台，我们有以下对比，如下：

1、建模流程方面：传统产品以业务需求的理解情况驱动模型的构建；我们采用的 SkyEmpower 数字建模一体化平台是通过模型的作用和成果进行引导，对于个性化的需求有更强的实现能力。

2、数据效率方面：传统产品的模型个性化程度低，难以适应千人千面的政务数字化需求；我们采用的 SkyEmpower 数字建模一体化平台设计柔性指标体系，满足个性化需求的同时实现模型加速进化，具备快速建模能力，在非后台介入的情况下实现一体化建模、数据流转和结果输出校验。

3、模型共享机制方面：传统产品主要服务于当前场景，模型的思路及经验容易形成信息茧房；我们采用的 SkyEmpower 数字建模一体化平台突出思路和经验的延伸，将成果生态化实现，有助于围绕应急、公安、气象等多行业相似模型之间互补互促、优化改进。

4、跨镜追踪技术：增加了对人员体态、衣着特征的识别，包括是否背包、衣服颜色、款式等信息。进一步聚合形成对人员特征的完整描绘和动态追踪，多模态提升人员识别的准确性和全面性。基于底层坚实的数据底座封装了一系列通用且高性能的算法服务，使得用户能够快速响应业务需求，搭建个性化的专题应用，融合模型的命中率相较于传统单个人像识别模型中识别效果最好的模型高出 8%。

### 四、技术创新点





## (一) 平台技术创新

### 1、平台对于多样化模型的接入能力和鲁棒性的增强

传统人工智能建模平台对于模型和算法的接入能力主要取决于具体应用目标，而应对越发复杂的技术需求，平台的模型层和算法层可能面临更大的挑战。该平台可以更加快速的完成多样化模型的接纳能力，提高多样算法融合过程中的平台鲁棒性。

### 2、平台可以实现对于当前模型和算法的可视化管理过程

传统人工智能建模平台对于模型和算法的管理的主要方式是后台管理的方式，对于非技术人员或一线工作人员来说并不友好。该平台实现对于当前的平台模型和平台算法进行可视化的动态管理过程，这种管理的范围包括但不限于：模型数据输入 / 输出管理、模型适用范围管理、相似模型快速构建能力、多模型核心数据管理和调用等、模型关键步骤的调整等，在实际应用中提高 30%-50% 的建模效率。

### 3、平台可以实现对于模型和算法的嵌套

传统人工智能建模平台对于模型和算法之间的嵌套和重用能力较弱，对于嵌套和重用的算法都需要通过后台进行重新编译和校正才可以使用，客观上增加了相似模型重用的代价和条件。该平台以可视化的方式，对整体的模型和算法实现嵌套和重用的管理，一方面增加平台效率，另一方面也可以实现相似模型的快速复用，在同类领域中复用度最高约为 80%。

## (二) 平台功能创新

### 1、预测预警系统

围绕城市生命线工程安全、公共安全、生产安全、自然灾害四大板块，实时汇聚经营性自建房、商业综合体、气象灾害、电动自行车等社会安全重点领域应用场景的风险监测、预警信息以及多渠道群众上报、舆情监测数据，通过预测预警子系统提供的工具化算法库，提供场景化的模型整合能力，实现应急业务数据价值的快速变现，高效挖掘数据内在价值，推动业务数据应用到智能化数据价值挖掘。

### 2、应急预案系统

通过 AI 助手提取、配置、推荐等能力实现预案要素标准化、预案演练可视化、预案评估数据化，政府、企业端在预案编制、预案演练、应急救援处置上，实现风险评估、资源调度等能力的对接与优化，实现政府各层级之间、政府部门之间以及政府与企业之间在风险、资源等信息上的相互衔接。

### 3、靶向治理系统

拓展网格化社会治理模式，结合城市应急预案，规范预警信息的处置流程，实现与权属责任单位、行业主管部门、应急管理部门、各级地方政府，应急联动响应、自动派单处置，实现与横向业务部门之间，纵向市、区、街（镇）、网格、企业等多级联动，实现预警任务从接收到响应处置、信息反馈，归档管理等工作的全过程线上闭环管理。

### 4、态势感知系统

围绕城市生命线工程安全、公共安全、生产安全、自然灾害四个板块，叠加场景三维可视化模型应用，打造立体的、多级分类的大数据分析监测预警一张图，对能够影响社会安全态势发生变化的因素，可快速的感知过去、现在、未来可能造成的风险，发现风险管理的脆弱点，对盲点进行持续跟踪，对风险异常进行实时识别，对突发事件进行实时监测，实现对社会安全风险综合监管的全域概览与精准把控。

### (三) 应用思路创新

#### 1、拓展数据共享和数据融合

通过多手段汇聚及治理社会安全风险相关数据，为社会安全态势感知业务应用建设提供数据支撑。

#### 2、探索社会安全风险综合预警

围绕城市生命线工程安全、公共安全、生产安全、自然灾害四大板块，探索商业综合体、经营性自建房、气象灾害、电动自行车等风险全面感知、综合分析、智能预警，依据算法融合和模型训练提升预警的精准度。

#### 3、提升预警应急处置能力

基于网格化治理机制，实现应急联动处置的横向指派、纵向调度，提升综合预警应急处置保障能力。

## 社会效益（限 200 字）

#### (1) 提升治理效能和安全管理能力。

针对现阶段社会安全综合监测预警存在的不足，通过管理创新与信息化技术相结合，相较于原有依靠大量人力完成的任务，以 AI 优化能力为核心，建立高效的问题发现和响应机制，联动多方力量，改善治理模式，拉动相关责任主体共同参与社会安全治理，不断织密安全防护网，构筑立体“前置防线”，大幅提升治理效能和安全管理能力。

#### (2) 有助于保障群众生命财产安全，增强群众安全感。

可大幅提升城市应对较大、重大、特别重大突发事件的预防预警、资源调度、响应处置等方面的工作能力，助力防范、化解、降低重大突发事件对人民群众生命财产安全的影响，增强群众安全感。

## 学会动态

### 会校联动，JSCS 走进金陵科技学院

为加强与高校的合作，推动计算机科学与技术的创新发展，江苏省计算机学会（JSCS）秘书长金莹教授于 2024 年 9 月 26 日走进金陵科技学院，面向青年教师和研究生做了一场题为“会校联动，



共绘未来——江苏省计算机学会助力研究生科研成长与职业发展”的报告。本次报告旨在展示学会的全面资源与支持，助力青年学者在学术和职业道路上取得更大成就。



## 简介

### 江苏省计算机学会理事单位

#### 南京中新赛克科技有限责任公司

南京中新赛克科技有限责任公司（以下简称：“中新赛克”），成立于 2007 年，母公司于 2017 年在深圳交易所上市，股票代码 002912。中新赛克是国内专注于数据提取、数据融合计算及其在信息安全等领域的领军企业，目前中新赛克在南京软件谷、吉山软件园、长沙绿地中心、重庆互联网智能产业园共拥有四个研发基地。同时，在北京、南京、广州、长沙、重庆、济南等十余个区域重点城市设有分支机构。

中新赛克主营产品分为五大体系，包括宽带网产品、移动网产品、网络内容安全产品、大数据运营产品、数据与网络安全产品，覆盖了网络空间数据提取、数据存储和计算、数据分析和挖掘、数据应用及展示等领域。主营业务为网络可视化基础架构、网络内容安全、大数据运营及数据与网络安全等产品的研发、生产和销售，以及相关产品的安装、调试和培训等技术服务，为政府、运营商、企事业单位及个人等提供产品和服务。

#### 荣誉资质

- 国家知识产权优势示范企业
- 国家专精特新小巨人企业
- 国家规划布局内重点软件企业
- 高新技术企业
- 部级信息技术应用创新优秀解决方案
- 江苏省规划布局内重点软件企业
- 江苏省正版正货承诺企业
- 江苏省质量信用 AAA 级企业
- 南京市市长质量奖