2025 年　第 1 期

江苏省计算机学会

COMMUNICATIONS OF THE JSCS

——教学科研双驱的新时代 IT 人才培养模式探索与实践

——在“软件方法学”科研与育人领域奋楫笃行

江苏省计算机学会常务理事单位

南通大学人工智能与计算机学院

南通大学人工智能与计算机学院（阿里云大数据学院、人工智能研究院合署）是为积极响应国家重大战略，主动应对新一轮科技革命和产业变革，全面实施南通市“名城名校融合发展”战略、打造名城名校“彼此赋能、携手共进”典范样板而成立的新型工科学院。

学院聚焦人工智能领域核心技术和关键问题，建立长三角区域人工智能高端人才培养基地，创新人才培养体系和科技协同创新体系。学院现拥有一支能力突出、结构合理的高水平教学科研团队。学院致力于建设一流的人工智能、计算机科学和技术科学研究与产业孵化基地，形成“人工智能 + 新工科建设”人才培养新模式。

学院设有人工智能、人工智能（科大讯飞实验班）、计算机科学与技术、软件工程、数据科学与大数据技术、软件工程（中外学分互认联合培养项目） 6 个本科招生专业方向，目前设有信息与通信工程（人工智能与医学大数据方向）博士点、博士后流动站，系统科学（生命健康系统分析与调控学科方向）博士点、智能科学与技术学术型硕士点，计算机技术、人工智能两个领域电子信息专业学位硕士点。同时学院涵盖计算机科学与技术、软件工程等成人学历教育本科专业。学院计算机科学与技术专业为国家级一流本科专业建设点，软件工程专业为江苏省一流本科专业建设点，计算机科学与技术类专业为江苏省高校重点专业类。同时，计算机科学与技术专业、软件工程专业为江苏省卓越工程教育培养计划专业。《数据结构》课程为国家级一流本科课程。

2023 年“计算机科学”学科进入 ESI 全球学科排名前 1% 学科，学院学科建设进入崭新的发展阶段。

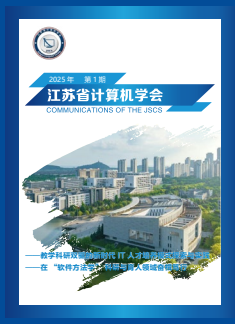
学院现有全日制在校本科生、研究生 1933 人，留学生 200 余人。经过多年的科学研究，学院形成了不确定人工智能与脑认知、新型计算理论与方法、智能软件工程与测试技术、网络与信息安全、数据库与信息系统、数字医学工程等多个研究方向。学院主持建设阿里云大数据学院政校企合作共建的产业学院，建有南通市虚拟现实技术公共技术服务平台、南通市高性能计算重点实验室等多个平台。

学院与美国、加拿大、英国、法国、澳大利亚、新加坡等国家和香港、澳门地区 30 多所高校开展学术交流与合作，与国内多所高校、科研院所建立紧密的合作关系。近年来累计派出教师到国外高校和科研机构进行学术交流与科研合作 90 余人次，师生参加国际学术交流 200 余人次。

学院秉承“祈通中西，力求精进”的校训精神、“学必期于用，用必适于地”的办学理念、“道德优美，学术纯粹”的价值追求，锐意进取，砥砺前行。

江苏省计算机学会通讯

COMMUNICATONS OF THE JSCS



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

新时代 IT 人才“铸强培优”–“三驱动五链条”的人才培养模式构建与实践

——河海大学刘凡教授

一、成果简介

近年来，教育部大力推进新工科建设，鼓励高校探索形成领跑全球工程 的中国教育模式，促进国家人才培养和强国建设。面对科技革命和 IT 产业变革带来的新变化新趋势，本成果以全面提高 IT 拔尖人才培养质量为重点，以解决关键核心技术“卡脖子”问题为导向，以强化课程思政与学科竞赛相互融通为主线开展 IT 拔尖人才的培养体系研究与设计，构建了 IT 人才“铸强培优”体系。并提出了围绕“技术发展、学生志趣、内外资源”的“三驱动”要素、落实“五链条”举措的创新模式，对 IT 人才“铸强培优”体系进行了实践。

针对课程教学模式对于学生自驱力不足，课程知识体系对于工科需求的前沿性、工程性欠缺，课程思政元素对于人才培养的挖掘不充分等问题，本成果积极开展教学改革，通过培优架构设计、教学内容重构、教学过程优化、智教水平提升、行业思政铸魂等举措，全面提升课程先进性、学生学习主动性，强化学生专业认知、科研创新能力，塑造学生正确三观，努力培养担当民族复兴大任的时代拔尖人才，并取得了如下教学成果：

（1）在人才培养方面，指导学生获国家级学科竞赛奖项 16 项、江苏省本科优秀毕业论文一等奖、二等奖；指导本科生第一作者在一区 TOP 期刊 IEEE TNNLS 上发表论文，并入选 ESI 热点论文、首届江苏省自然科学百篇优秀学术论文成果，已被引 3000 余次；指导本科生分别获国际会议最佳 Demo 论文、最佳 Dataset 论文、最佳 Presentation 论文。

（2）在教研教改方面，获全国高校人工智能教师教学创意竞赛二等奖、江苏省高校教师教学创新大赛二等奖等省部级及以上荣誉 12 项、教改课题 6 项、编写规划教材 4 部，自编教材与自研课程辅助小程序经多次迭代，在南京大学、东南大学等十余所高校成功推广应用，获得一致好评。

（3）在立德树人方面，连续担任班导师，所带班级获江苏省“先进班集体”2 个，指导学生获省优秀共青团员、省大学生年度人物提名奖、省优秀学生干部等 14 项省部级奖项，部分优秀毕业生被《光明日报》和新华社等媒体报道。

2. 成果主要解决的教学问题及解决教学问题的方法

2.1 成果主要解决的教学问题

（1）学生群体自驱力不足，志趣培养欠缺

与 IT 人才相关专业的理论课程知识点较为抽象且密集，需要思考如何采用更加合理、多样的教学方法和培养手段激发学生的探究意识和创新创造的激情。学生群体的自驱力不足，需重视学生志趣的培养，提升育人“温度”，帮助学生提高自我教育、学习的能力，给予学生充分施展拳脚的空间，同时鼓励学生将短期的兴趣发展成为终身志业的勇气和决心。

（2）课程体系工程性不足，前沿技术欠缺

传统的教学内容源于老师单方面构建，对专业技术发展前沿的跟进较为缓慢，忽视了专业课程教学内容与最新

的科研、竞赛等资源的协同。学生仅仅依赖课堂讲授是不够的，一方面学习过程过于被动，难以将理论知识应用到实际应用场景中，另一方面缺乏实践机会，知识的掌握仅停留在理论层面。在 IT 行业快速发展的背景下，学科领域的知识体系和方法论随着技术的发展而发生变化，因此专业课程体系的改革不仅要响应工程实践性的需求，还要应对专业技术发展的变化。

（3）思政元素挖掘不充分，资源协同欠缺

教学质量文化应该与专业特色充分融合，原有培养模式对学生的专业创新能力、综合素质和学术道德等方面质量文化建设力度还不够。在推行课程思政教学过程中，存在侧重专业知识而缺乏人文情感，忽视了创新精神和工匠精神，习惯于通用案例而轻视了学科竞赛与行业特色案例等校企内外资源的引入。课程引入思政素材空洞，缺乏实证调查数据，影响了思政元素的挖掘和思政教育的效果。

2.2 解决教学问题的方法

（1）培优架构设计：教学闭环与铸强培优的多元教育架构

本成果着眼于 IT 人才铸强培优的核心理念与创新教育思想，设计了学习闭环的多元教育架构。围绕“教研 - 教改 - 教评”落实了教学内容重构、教学过程优化、智教水平提升等三个方面的举措，重构了工程驱动与紧跟前沿的全局知识体系，优化了课赛激励与产教融合的内外资源协同，探索了数字化与智能化的教育教学技术。建立了集 CDIO 工程教育、产教融合、竞赛促教、科研促学、数字教评、行业思政资源及网络思政教育等要素于一体的多元教学过程，形成了微课线上自学 + 自建数字平台测验 + 线下翻转课堂 + 在线智能助教答疑相结合的教学闭环。在 IT 人才思想教育方面，践行了行业思政铸魂的方案与措施，打造了实践强能与思政培根的铸魂育人工程。

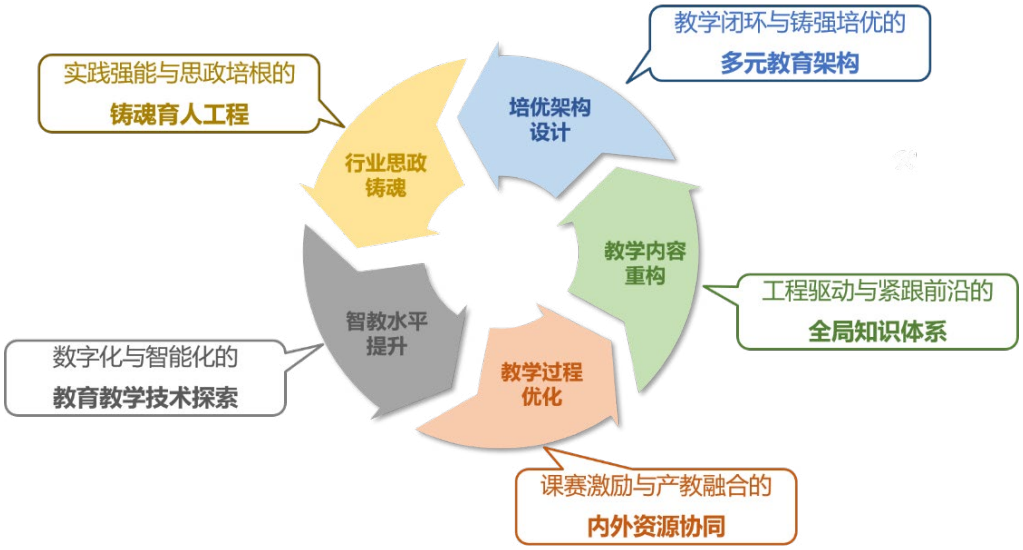


图 1 IT 人才培优架构设计示意图

（2）教学内容重构：工程驱动与紧跟前沿的全局知识体系

着眼“重局部轻整体”难题，以 CDIO 工程教育理念对每门专业课程的教学内容进行梳理，形成基于全局观的 IT 人才知识体系。教学内容的改革必须紧跟技术发展前沿，结合“新兴 IT 技术”，有针对性的选择实践案例，构建全局性多元化教学方案。同时，加强 CDIO 工程教育各个阶段的实践环节，培养学生工程实践能力、自我学习能力和创新能力。构建特色思政案例、行业工程案例，深挖课程背后思政要素，线上线下多场景、多角度渗透，考察锻炼学生在实践过程中的全局思维。

图 2 融合 CDIO 教育的工程案例

（3）教学过程优化：课赛激励与产教融合的内外资源协同

着重把学科竞赛和科研技能融入日常专业课程教学，提升学生实践能力，将原本抽象的知识变具体。以项目竞赛为导向，课程内容为载体，以赛促教。以科研为驱动，提升创新能力。充分发挥学校资源和创新优势，积极建设科研实验室与创新实践基地，同时开设科研课程，让学生了解科研的基本方法和技能，培养学生的科研意识和能力。

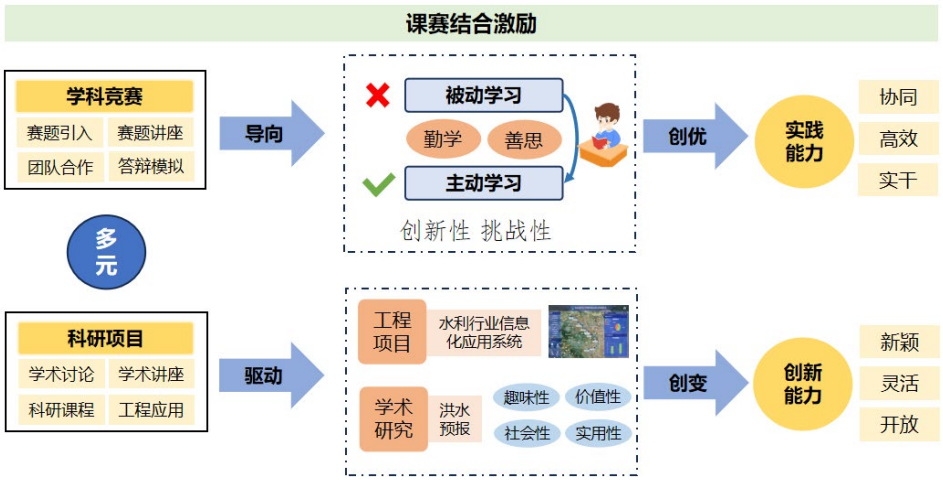


图 3 课赛结合激励示意图

根据新工科“问技术发展改内容，更新工程人才知识体系；问内外资源创条件，打造工程教育开放融合新生态”的要求，以解决高校实践平台与企业无法接轨、实践案例缺乏等问题为着眼点，融入产业案例，加强实践资源建设。与华为、阿里及百度等企业合作共建及一流课程等教育部产学合作课程及云计算开发实验平台，提高实践教学平台的实用性和创新性。

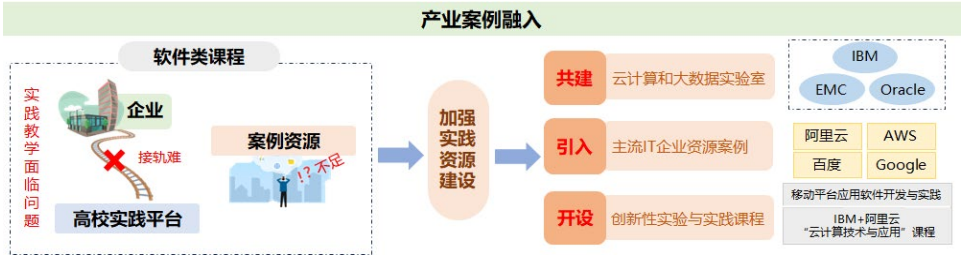


图 4 产教案例融入示意图

（4）智教水平提升：数字化与智能化的教育教学技术探索

自研辅助教学小程序 SmartClass，积极将数字技术资源应用于软件开发类的理论、实践课程的教学过程，实施数字技术与教育教学融合，持续开展数字化与智能化教育教学实践探索。自研辅助教学小程序除发布测验外，能够准确分析和追踪学生知识点的掌握情况，结合学生认知水平评估学生的综合表现。此外，通过引入课程在线助教与学生进行一对一沟通交流，以增强学生学习效果、提升智教水平与教学质量。

图 5 自研的微信小程序与引入的数字智教工具示例

2. 成果主要解决的教学问题及解决教学问题的方法

2.1 成果主要解决的教学问题

（1）学生群体自驱力不足，志趣培养欠缺

与 IT 人才相关专业的理论课程知识点较为抽象且密集，需要思考如何采用更加合理、多样的教学方法和培养手段激发学生的探究意识和创新创造的激情。学生群体的自驱力不足，需重视学生志趣的培养，提升育人“温度”，帮助学生提高自我教育、学习的能力，给予学生充分施展拳脚的空间，同时鼓励学生将短期的兴趣发展成为终身志业的勇气和决心。

（2）课程体系工程性不足，前沿技术欠缺

传统的教学内容源于老师单方面构建，对专业技术发展前沿的跟进较为缓慢，忽视了专业课程教学内容与最新的科研、竞赛等资源的协同。学生仅仅依赖课堂讲授是不够的，一方面学习过程过于被动，难以将理论知识应用到实际应用场景中，另一方面缺乏实践机会，知识的掌握仅停留在理论层面。在 IT 行业快速发展的背景下，学科领域的知识体系和方法论随着技术的发展而发生变化，因此专业课程体系的改革不仅要响应工程实践性的需求，还要应对专业技术发展的变化。

（3）思政元素挖掘不充分，资源协同欠缺

教学质量文化应该与专业特色充分融合，原有培养模式对学生的专业创新能力、综合素质和学术道德等方面质量文化建设力度还不够。在推行课程思政教学过程中，存在侧重专业知识而缺乏人文情感，忽视了创新精神和工匠精神，习惯于通用案例而轻视了学科竞赛与行业特色案例等校企内外资源的引入。课程引入思政素材空洞，缺乏实证调查数据，影响了思政元素的挖掘和思政教育的效果。

2.2 解决教学问题的方法

（1）培优架构设计：教学闭环与铸强培优的多元教育架构

本成果着眼于 IT 人才铸强培优的核心理念与创新教育思想，设计了学习闭环的多元教育架构。围绕“教研 - 教改 - 教评”落实了教学内容重构、教学过程优化、智教水平提升等三个方面的举措，重构了工程驱动与紧跟前沿的全局知识体系，优化了课赛激励与产教融合的内外资源协同，探索了数字化与智能化的教育教学技术。建立了集 CDIO 工程教育、产教融合、竞赛促教、科研促学、数字教评、行业思政资源及网络思政教育等要素于一体的多元教学过程，形成了微课线上自学 + 自建数字平台测验 + 线下翻转课堂 + 在线智能助教答疑相结合的教学闭环。在 IT 人才思想教育方面，践行了行业思政铸魂的方案与措施，打造了实践强能与思政培根的铸魂育人工程。

（2）教学内容重构：工程驱动与紧跟前沿的全局知识体系

着眼“重局部轻整体”难题，以 CDIO 工程教育理念对每门专业课程的教学内容进行梳理，形成基于全局观的 IT 人才知识体系。教学内容的改革必须紧跟技术发展前沿，结合“新兴 IT 技术”，有针对性的选择实践案例，构建全局性多元化教学方案。同时，加强 CDIO 工程教育各个阶段的实践环节，培养学生工程实践能力、自我学习能力和创新能力。构建特色思政案例、行业工程案例，深挖课程背后思政要素，线上线下多场景、多角度渗透，考察锻炼学生在实践过程中的全局思维。

（3）教学过程优化：课赛激励与产教融合的内外资源协同

着重把学科竞赛和科研技能融入日常专业课程教学，提升学生实践能力，将原本抽象的知识变具体。以项目竞赛为导向，课程内容为载体，以赛促教。以科研为驱动，提升创新能力。充分发挥学校资源和创新优势，积极建设科研实验室与创新实践基地，同时开设科研课程，让学生了解科研的基本方法和技能，培养学生的科研意识和能力。

根据新工科“问技术发展改内容，更新工程人才知识体系；问内外资源创条件，打造工程教育开放融合新生态”的要求，以解决高校实践平台与企业无法接轨、实践案例缺乏等问题为着眼点，融入产业案例，加强实践资源建设。与华为、阿里及百度等企业合作共建及一流课程等教育部产学合作课程及云计算开发实验平台 , 提高实践教学平台的实用性和创新性。

（4）智教水平提升：数字化与智能化的教育教学技术探索

自研辅助教学小程序 SmartClass，积极将数字技术资源应用于软件开发类的理论、实践课程的教学过程，实施数字技术与教育教学融合，持续开展数字化与智能化教育教学实践探索。自研辅助教学小程序除发布测验外，能够准确分析和追踪学生知识点的掌握情况，结合学生认知水平评估学生的综合表现。此外，通过引入课程在线助教与学生进行一对一沟通交流，以增强学生学习效果、提升智教水平与教学质量。

（5）行业思政铸魂：实践强能与思政培根的铸魂育人工程

实施集实践强能、思政培根及“行业 + 思政”铸魂育人工程，激发学生原动力，促进 IT 拔尖人才高质量特色化发展；立足重大工程项目案例，深入挖掘思政元素，通过行业特色精神熏陶、工程实践锤炼，提高 IT 拔尖人才培养质量。引入行业发展历程和成就、重大工程建设背后的故事等，将人文知识融入课程案例设计与讲解中，丰富课程内涵，增强课程感染力。科学探究理论知识在现实中的实践应用价值，培养学生解决实际工程问题的能力及社会责任感，提升其作为未来 IT 拔尖人才应具备的勇于担当、甘于奉献等职业素养，引导其积极投身于社会主义现代化强国建设的伟大事业中。

3. 成果的创新点

3.1 新时代 IT 人才“铸强培优”理念

本成果针对学生群体志趣培养的欠缺、专业课程体系前沿工程性不足以及思政元素挖掘不充分等问题，提出了“技术发展、学生志趣、内外资源”三驱动的 IT 拔尖人才培养理念。

（1）设计了 CDIO 工程理念下多元化的教学方法和培养手段，激发学生的探究意识和创新创造的激情。重视学生志趣的培养，提升育人“温度”，提高学生群体的自驱力，鼓励学生将短期的兴趣发展成为终身志业的勇气和决心。

（2）引入了学科竞赛和工程案例等内外资源，充分将理论知识应用到实际情境中。紧跟行业技术前沿，改革学科领域的知识体系和方法论以激发学生学习主动性，线上线下混合教学以突破教学时空限制。

（3）加强了对学生的专业创新能力、综合素质和学术道德等方面质量文化建设力度。充分引入具有特色的行业思政案例，提高了思政元素的挖掘和思政教育的效果。

3.2 “三驱动五链条”的 IT 人才培养模式

（1）“三驱动”与“五链条”的耦合。本成果将“学生志趣、技术发展、内外资源”的三元驱动与“培养架构设计、教学内容重构、教学过程优化、智教水平提升、行业思政铸魂”五链条进行了有机耦合。二者相互依赖，相互作用，使得新时代 IT 人才“铸强培优”的创新培养模式与举措得到了良好的应用效果。

（2）行业课程思政与教育教学的有机融合。强化教师课程思政的理念与能力，找准课程与思政元素的结合点，形成匠心独具的“行业思政元素”融入体系。同时引入水利行业发展历程和成就、重大工程建设背后的故事和事迹，将人文知识融入课程案例的设计与讲解中。

（3）满足新时代对人才培养的要求。从新时代经济社会发展对 IT 人才的需求出发，培养人才解决关键技术问题的创新实践能力，全面提高拔尖人才自主培养的质量，从而提高 IT 人才服务国家和区域经济社会发展能力和水平的“适配度”。

4. 成果的推广应用效果

刘凡自研的“智教”辅助教学小程序实现了教师课堂教学的教学分析与评价，在提供课程管理、定位签到功能的基础上，可精准地完成试题推荐和自动组卷功能，并准确分析和追踪学生知识点的掌握情况，量化处理学生的学习结果。该系统已在南大、东大等十多所高校推广应用。在 2020 年 2 月 24 日至 2020 年 6 月 26 日期间，共计为 4 所高校的 713 名师生提供服务，获得了一致好评。2021 年 1 月至今已为 2223 名师生提供服务，日均访问人数 9.83 次，日均打开 23.82 次，日均访问页面 222.52 次。

自编教材《JSP 基础入门》发布了全套微课视频，已被南京理工大学、中国矿业大学等多所高校采用，并在 MOOC 与河海大学网络教学平台连续开课 6 次，近 2000 余人在线选课，在 B 站、喜马拉雅播放量达 3 万余次。

加快建设教育强国的纲领性文件

——教育部负责人解读《教育强国建设规划纲要（2024—2035 年）

来源：中华人民共和国教育部

http://www.moe.gov.cn/jyb_xwfb/s271/202501/t20250119_1176197.html

近日，中共中央、国务院印发《教育强国建设规划纲要（2024—2035 年）》（以下简称《纲要》）。教育部负责人就《纲要》有关情况回答了记者提问。

1. 问：《纲要》出台有什么背景和意义？

答：教育是强国建设、民族复兴之基。党的十八大以来，以习近平同志为核心的党中央坚持把教育作为国之大计、党之大计，作出加快教育现代化、建设教育强国的重大决策，推动新时代教育事业取得历史性成就、发生格局性变化，我国教育现代化发展总体水平跨入世界中上国家行列，教育强国建设进入了蓄势突破、全面跃升的重要阶段。站在新的起点上，党的二十大明确提出到 2035 年建成教育强国的宏伟目标。

为加快推进教育强国建设，中央教育工作领导小组加强对《纲要》编制的统筹领导，教育部会同有关部门深入推进编制工作，认真学习习近平新时代中国特色社会主义思想，深入贯彻党的二十大和二十届二中、三中全会精神，全面学习领会习近平总书记关于教育的重要论述和重要指示批示精神，特别是在全国教育大会上的重要讲话精神，深入开展调研论证，广泛征求各地区各部门、各民主党派中央、有关学校和专家学者等意见建议。

此次印发的《纲要》，是在我国迈上全面建设社会主义现代化国家新征程、向第二个百年奋斗目标进军的关键时刻，党中央、国务院颁布实施的教育事业发展纲领性文件，是首个以教育强国为主题、以全面服务中国式现代化建设为重要任务的国家行动计划，是全面推进教育科技人才一体统筹发展、提升国家创新体系整体效能的顶层制度安排，对落实党的二十大重大部署，更好发挥教育强国建设在全面推进强国建设、民族复兴伟业中的先导任务、坚实基础、战略支撑作用，具有重大而深远的意义。

2. 问：《纲要》编制的主要思路是什么？

答：《纲要》以习近平新时代中国特色社会主义思想为指导，深入贯彻全国教育大会精神，紧扣中央关心、群众关切、社会关注，坚持目标导向、问题导向和效果导向，紧紧围绕教育的“三大属性”，以“六大特质”为主要特征、以“八大体系”为基本结构、以正确处理“五个重大关系”为关键要求，将深化改革贯穿全文，突出教育科技人才一体统筹部署，推出一系列创新举措，推动从教育大国向教育强国的系统跃升。

其中，“三大属性”，指的是教育的政治属性、人民属性、战略属性；“六大特质”，指的是教育强国应当具有强大的思政引领力、人才竞争力、科技支撑力、民生保障力、社会协同力、国际影响力。“八大体系”，指的是全面构建固本铸魂的思想政治教育体系、公平优质的基础教育体系、自强卓越的高等教育体系、产教融合的职业教育体系、泛在可及的终身教育体系、创新牵引的科技支撑体系、素质精良的教师队伍体系、开放互鉴的国际合作体系。

正确处理“五个重大关系”，指的是必须正确处理支撑国家战略和满足民生需求、知识学习和全面发展、培养人才和满足社会需要、规范有序和激发活力、扎根中国大地和借鉴国际经验的关系。

3. 问：《纲要》在总体目标设定上有哪些考虑?

答：《纲要》坚持远近结合，分 2027、2035 年“两步走”。

“第一步”面向开局起步阶段，重点是全方位打牢教育强国建设基础。《纲要》明确到 2027 年，教育强国建设取得重要阶段性成效。各级教育普及水平持续巩固提升，高质量教育体系初步形成，人民群众教育获得感明显提升，人才自主培养质量全面提高，拔尖创新人才不断涌现，关键领域改革取得实质性进展，教育布局结构与经济社会和人口高质量发展需求更加契合，具有全球影响力的重要教育中心建设迈上新台阶。

“第二步”面向中长期，深化重大战略布局，确保如期建成教育强国。《纲要》明确到 2035 年，党对教育事业全面领导的制度体系和工作机制系统完备，高质量教育体系全面建成，基础教育普及水平和质量稳居世界前列，学习型社会全面形成，人民群众教育满意度显著跃升，教育服务国家战略能力显著跃升，教育现代化总体实现。

4. 问：请介绍一下《纲要》的结构和主要内容。

答：《纲要》共 11 部分，分别对应总体要求、“八大体系”、综合改革和组织实施。

一是总体要求，明确了教育强国建设的指导思想、工作原则和主要目标。

二是塑造立德树人新格局，培养担当民族复兴大任的时代新人。提出加强和改进新时代学校思想政治教育，加强党的创新理论体系化学理化研究阐释和成果应用，拓展实践育人和网络育人空间和阵地，促进学生健康成长、全面发展，打造培根铸魂、启智增慧的高质量教材，推广普及国家通用语言文字。

三是办强办优基础教育，夯实全面提升国民素质战略基点。提出健全与人口变化相适应的基础教育资源统筹调配机制，推动义务教育优质均衡发展和城乡一体化，促进学前教育普及普惠和高中阶段学校多样化发展，统筹推进“双减”和教育教学质量提升。

四是增强高等教育综合实力，打造战略引领力量。提出分类推进高校改革发展，优化高等教育布局，加快建设中国特色、世界一流的大学和优势学科，完善拔尖创新人才发现和培养机制，构建中国哲学社会科学自主知识体系。

五是培育壮大国家战略科技力量，有力支撑高水平科技自立自强。提出实施基础学科和交叉学科突破计划，促进青年科技人才成长发展，提高高校科技成果转化效能，建设高等研究院开辟振兴区域发展新赛道。

六是加快建设现代职业教育体系，培养大国工匠、能工巧匠、高技能人才。提出塑造多元办学、产教融合新形态，以职普融通拓宽学生成长成才通道，提升职业学校关键办学能力，优化技能人才成长政策环境。

七是建设学习型社会，以教育数字化开辟发展新赛道、塑造发展新优势。提出提升终身学习公共服务水平，实施国家教育数字化战略，促进人工智能助力教育变革。

八是建设高素质专业化教师队伍，筑牢教育强国根基。提出实施教育家精神铸魂强师行动，提升教师专业素质能力，优化教师管理和资源配置，提高教师政治地位、社会地位、职业地位。

九是深化教育综合改革，激发教育发展活力。提出深化教育评价改革，完善人才培养与经济社会发展需要适配机制，提升依法治教和管理水平，健全教育战略性投入机制，构建教育科技人才一体统筹推进机制。

十是完善教育对外开放战略策略，建设具有全球影响力的重要教育中心。提出提升全球人才培养和集聚能力，扩大国际学术交流和教育科研合作，积极参与全球教育治理。

十一是加强组织实施。要求完善党委统一领导、党政齐抓共管、部门各负其责的教育领导体制，全面推进各级各类学校党的建设。充分发挥中央教育工作领导小组作用，各级党委和政府要切实扛起教育强国建设的政治责任，形成建设教育强国强大合力。

5. 问：如何抓好《纲要》贯彻落实？

答：贯彻落实好《纲要》，是当前和今后一个时期各级党委和政府的重要任务。教育系统要积极开展多形式、分层次、全覆盖的学习宣传培训，把全面实施《纲要》与学习贯彻习近平总书记关于教育的重要论述，特别是在全国教育大会上的重要讲话精神和习近平同志《论教育》结合起来，引导广大党员干部教师把思想和行动统一到中央决策部署上来，推动各项工作落地见效。

为了推动教育强国建设高起点高质量开局起步，教育部正抓紧研究启动加快建设教育强国三年行动计划，加强顶层设计，开展改革试点，强化监测评价，推动《纲要》重大部署落地落实。

教育关系千家万户，实施好《纲要》是全社会的共同责任。要健全学校家庭社会协同育人机制，动员全社会共同关心支持教育改革发展。广泛宣传报道各地各校学习贯彻《纲要》的进展成效，推广经验成果和先进典型，营造良好社会环境和舆论氛围。

学会动态

江苏省计算机学会大数据专委会 2025 年度第一次工作会议在南京召开

为了推进江苏省计算机学会大数据专委会工作的开展，3 月 28 日下午，江苏省计算机学会大数据专委会在江苏省联合征信有限公司召开了 2025 年度第一次工作会议。本次会议旨在谋划和讨论 2025 年度工作，明确 2025 年度工作规划以及重点任务。会议由大数据专委会主任吉根林教授主持，大数据专委会全体负责人出席会议。

本次工作会议聚焦 2025 年江苏省大数据学术会议举办事宜，与会的大数据专委会各位负责人围绕学术会议主题设定、议程规划、征稿审稿机制等核心议题展开深入讨论，多位专委会负责人积极发言，共同为 2025 年江苏省大数据学术会议的召开出谋划策。

为了充分发挥大数据专委会的优势，会议介绍了与中国铁道出版社的战略合作方案，拟与中国铁道出版社合作出版大数据系列教材，大数据专委会各位负责人对合作教材的内容以及目标定位提出了宝贵建议。

最后，大数据专委会主任吉根林教授对本次工作会议进行了总结，强调了 2025 年度大数据专委会的工作重点以及工作模式的创新，希望各位专委会负责人能够积极贡献，提升大数据专委会在省内影响力，带领大数据专委会在新的一年里更上一层楼！

基于机器学习的生物分子纳米孔数据识别方法研究

——2024 年江苏省计算机学会优秀博士论文奖

作者：关晓宇

单位：南京航空航天大学计算机科学与技术学院

指导老师：张道强

论文摘要

纳米孔单分子检测技术因其具有长检测时间和高分辨率等优点，是目前主流的单分子检测平台之一。纳米孔检测技术所生成的数据需经过数据分析手段分析和处理之后才可应用于后续的研究。然而，主流的纳米孔数据分析手段还停留在统计层面上，仍然需要大量的人工成本，并且不具备高精度识别的能力。最近，机器学习技术高速发展，已极大地促进了生物、化学和医学等学科的智能化发展。机器学习算法已应用于纳米孔数据的分析，可以减少生成数据所需的手动分析量、提高检测序列识别精度、指导遗传诊断和治疗各种疾病和病症。本论文基于机器学习的理论和应用技术，以纳米孔数据的特点开展四个方面的研究工作：（1）基于特征构建的机器学习方法减少生成序列所需的手动分析量；（2）基于深度学习的方法提高检测序列识别精度；（3）基于主动学习的方法降低样本标注成本；（4）基于无监督深度学习的方法识别突变位点来指导遗传诊断。主要工作和创新点如下：

（1）针对 RNA 的单分子识别问题，提出了一个基于特征构建的传统机器学习方法。根据检测序列数据的特殊性，结合纳米孔领域专家的意见，设计并构建了包括均值和标准差在内的总共 11 个特征。根据特征分布图可以明显观察到构建的特征之间有明显的判别界限，其中部分重要的特征会对分类结果起到直接的影响。基于构建的特征，利用几种传统的机器学习算法进行分类验证。实验结果表明，相对于其他几种机器学习算法，随机森林算法表现出了更好的性能。特征层面上，特征中值，标准差和均值对分类的性能起到了更有益的作用。模型对噪声信号的分类效果较差，主要是由于噪声序列相对复杂且没有规范的序列结构。提出的特征选择方案同样适用于其他类似的纳米孔数据分析任务中。所提出的手工特征构建方法配合传统机器学习分类的方案能够应用到更多的纳米孔数据分析中，从而减少生成的纳米孔检测序列所需的手动分析量，并且提高检测的准确度。

（2）针对 RNA 的单分子识别特征构建需要大量手工成本的问题，提出了一个基于序列转化策略的深度学习模型。由于检测的原始序列数据通过分段分割，导致每一个 RNA 分子的序列都是严格不等长的，而深度学习模型的输入却要求输入模型中的序列严格等长。因而，本文提出了序列转化方案，这两种序列转化的方式都实现了将不等长序列转化为可以通过深度学习模型进行学习的数据形式。基于等长的序列，提出了基于注意力机制的深度学习模型，避免了人工特征构建造成的成本增加。实验结果表明，该模型具有非常好的适用性，模型中使用的损失函数对分类

起到了积极的作用。基于等尺寸的图像，提出了基于 Transformer 的深度学习模型，同样实现了更高精度的分类。通过实验结果分析，二维图像数据相对于一维序列，有着更多的局部特征有利于模型的分类。将所提出的转化策略应用于牛津纳米孔编码数据集，验证了所提出策略的有效性。本文分析了传统机器学习的特征提取和深度学习特征提取的区别，探索了深度学习在纳米孔数据分析中的适用性。实现了原始纳米孔序列长度严格不等导致的无法使用深度学习模型进行分析的困境，提高了检测序列识别精度。

（3）针对纳米孔检测数据标注成本高的问题，提出了一个带约束的主动学习模型。本文将主动学习首次应用于纳米孔数据分析，根据纳米孔数据的特殊性，提出了偏置约束以改进主动学习中的样本选择策略。其可以有效的保证挑选的样本有利于模型的训练，使得模型更快的达到最优性能。实验结果表明，主动学习策略可以显著降低标注成本，并且可以达到全监督训练模型下的最优测试性能。设计的主动学习方法在其他纳米孔数据集上进行了验证，同样可以实现降低标注成本的目的。通过实验分析，主动学习策略可以帮助纳米孔专家了解哪些样本对分类任务至关重要，可以指导标注专家进行标注。提出的方法可以优化纳米孔领域的样本标注策略，使用主动学习技术可以有效的降低纳米孔庞大数据所带来的标注成本。本文分析了主动学习在样本标注成本上积极作用效果，探索了主动学习在纳米孔数据分析中的适用性。

（4）针对纳米孔测序的 DNA 羧甲基化数据识别问题，提出了一个无监督深度学习模型。该模型通过自编码器进行特征提取，在通过聚类方法将特征进行聚类，将数据点聚类到一个可以通过联合优化的特征空间中，最终实现精准识别突变位点的目的。该方法通过最小化编码器的聚类损失和解码器的重构损失进行迭代训练。实验结果发现，所提出的方法在提高性能的同时，对超参数设置也具有鲁棒性。这种无监督深度学习策略可以在减少人工计算成本的情况直接检测出甲基化突变位点，实验结果在一定条件下要好于全监督的机器学习方法。提出的方法可以方便生物专家对 DNA 突变位点的快速定位，对未来更多类似的从 DNA 测序中精准定位突变位点的任务提供一个解决方案，方便指导遗传诊断。所设计的方法可以看作是一种不需要任何人工干预的启发式工具，它可以用来处理原始纳米孔数据。

专家推荐语

非常荣幸向江苏省计算机学会推荐这篇题为《基于机器学习的生物分子纳米孔数据识别方法研究》的优秀博士论文。作者在该论文中深入探讨了纳米孔数据的特点，并从四个关键方面开展了创新性研究，展现了其卓越的科研能力和对前沿技术的深刻理解。

首先，论文通过基于特征构建的机器学习方法，显著减少了生成序列所需的手动分析量。这一方法的提出不仅提高了数据处理的效率，同时也为纳米孔数据的大规模应用奠定了坚实的基础。

其次，作者采用了深度学习的方法，大幅提升了检测序列的识别精度。在当前生物信息学研究中，数据的准确性至关重要。作者通过引入深度神经网络模型，开创性地解决了纳米孔数据中噪声干扰和复杂模式识别的难题，显著提高了序列分析的可靠性和准确性。

第三，论文引入了主动学习的方法，有效降低了样本标注成本。生物分子数据标注往往需要大量的人力和时间，而主动学习通过智能选取最有信息价值的样本进行标注，极大地减少了人工参与的工作量，从而提升了整个数据处理流程的效率。

最后，作者利用无监督深度学习的方法识别突变位点，进一步推动了遗传诊断的研究。无监督学习在不依赖标签数据的情况下，能够自动发现数据中的潜在模式和结构，这为遗传疾病的早期诊断提供了重要的技术支撑。

综上所述，本文的研究工作兼具理论深度与实际应用价值，不仅在机器学习和生物信息学领域具有重要的学术贡献，而且在生物分子数据分析和遗传诊断等实际应用中展现了巨大的潜力。作者通过严谨的研究方法和创新性的技术手段，为学术界和产业界提供了一套高效、精准的纳米孔数据分析解决方案。

因此，我强烈推荐这篇优秀博士论文参评江苏省计算机学会的优秀博士论文奖，相信作者的研究成果将为相关领域的发展做出重要贡献。

论文看点

本文结合机器学习和纳米孔技术的最新研究成果，通过机器学习方法充分分析生物学任务在纳米孔领域中的诸多研究问题。在已有的研究方法基础上，创新并发展更加高效的纳米孔分析模型和方法。本文主要解决如下三个核心问题：（1）可以减少生成序列所需的手动分析量；（2）提高检测序列识别精度；（3）指导遗传诊断和治疗各种疾病和病症。根据现有采集到的纳米孔检测数据，针对性的提出了若干关键性问题：（1）在分子识别中高效的引入机器学习有助于减少生成序列所需的手动分析量；（2）在分子识别中引入深度学习可以提高检测序列的识别精度；（3）引入主动学习技术能有效改进现有监督学习方法需要大量标注成本的困境；（4）甲基化的检测可以辅助医学进一步认识癌变致病机理，从而帮助医生更好的理解癌变。本文充分考虑如何有效的利用纳米孔检测数据的特点和数据格式，充分利用当前的机器学习先进技术，并致力于得到更广泛、可靠、易用的机器学习分析模型框架。探索更加高效的纳米孔数据的高精度识别模型和方法，探求纳米孔数据的深层次生物学原理。探寻机器学习分析生物分子识别的可靠性，验证深度学习在纳米孔检测数据中的适用性，验证主动学习在纳米孔数据中的可行性，利用无监督算法发现 DNA 羧甲基化在基因组序列的位置。

本文的主要看点如下：

一、基于特征构建的纳米孔检测 RNA 类型识别方法

RNA 在调节生命体基因表达和生理学方面发挥着关键的作用。RNA 结构在生命体的基因组中是普遍存在的，在生命体的各个方面均有显现。RNA 在生命体中具有多种功能，包括了蛋白质翻译、基因沉默、表观遗传调控、遗传信息存储和生物催化等。RNA 的功能多样性取决于其高度通用的三级结构变化。然而，现有的大多数结构和功能研究均需要复杂的预处理过程，在自然条件下直接检测 RNA 结构具有一定的难度。因此，研究 RNA 的功能和结构之间的关系是非常有必要的。原则上，高通量测序（High Throughput Sequencing，HTS）技术的进步开启了基因组学的新时代。DNA 和 RNA 序列可测量的容量的爆炸性增长使得全世界的研究人员受益匪浅。然而，构成这些序列的大多数 RNA 需要进一步研究其结构表征。一般来说，RNA 结构表征（即识别功能性 RNA 结构）需要结合热力学、系统发育学和实验分析。鉴于 RNA 种类的不断增加和研究详细 RNA 表征所需的昂贵设备等限制，快速且廉价的表征研究方法需要进行进一步的研究。

RNA 的不同折叠方式将导致生命体表现出不同的特性，使用 MspA 纳米孔来检测 RNA 分子可以根据检测信号的差异来判断不同的 RNA 分子类型。使用的 RNA 分子主要是 overhanged siRNA，blunt siRNA，tRNA 和 5s rRNA，其形态如图 1 所示。

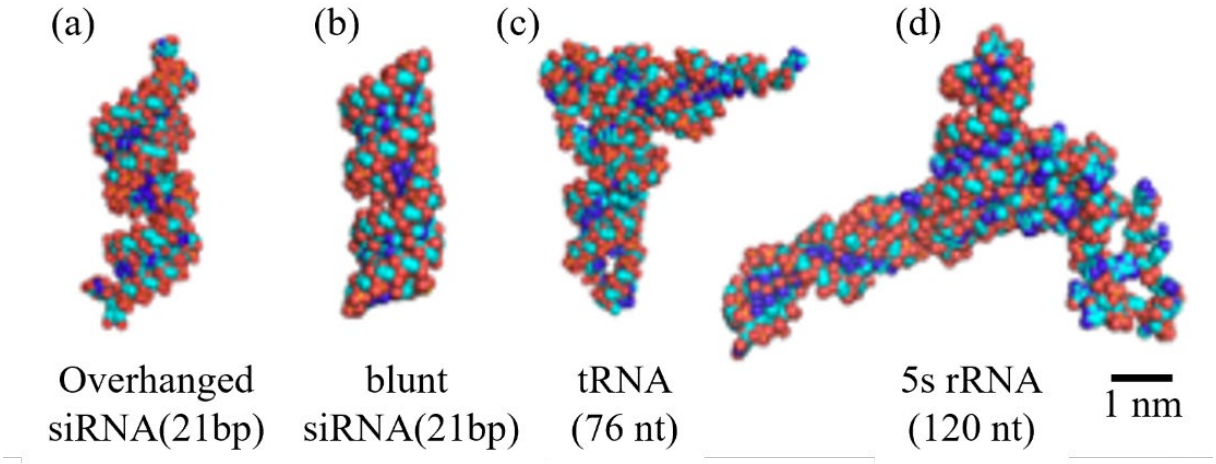


图 1 RNA 分子三维结构图：(a) overhanged siRNA；(b) blunt siRNA；(c) tRNA；(d) 5s rRNA

遵循传统的纳米孔传感策略，其中 1.5 M KCl 缓冲液（1.5 M KCl，10 mM HEPES，pH 7.0）顺式放置，1 M CaCl2 缓冲液（1 M CaCl，10 mM HEPES，pH7.0）反式放置。MspA 纳米孔具有大的前庭开口和整体呈锥形孔几何形状的特点，允许纳米孔捕获形态大的分析物（例如 RNA）。虽然 overhanged siRNA，blunt siRNA，tRNA 和 5srRNA 在结构上不同，但是可以通过相同的 MspA 纳米孔检测方法来进行区分。

tRNA 是 RNA 结构生物学中一个被广泛研究和熟知的一种 RNA 类型。MspA 纳米孔可以识别 tRNA 的两种结构。生成的 tRNA 信号的示例如图 2 所示。开孔电流（I_o）、堵塞幅度（I_b）、停留时间（t_{off}）和事件间间隔（t_{on}）的定义如图 2 所示。阻断百分比 $\%I_b$ 由（I_o- I_b）/I_o 确定。由于 blunt siRNA，tRNA 分子存在着不同的进孔方式，会产生两种不同的检测类型，这里将其定义为 type1 和 type2。横轴上的红色三角形对应于 tRNA type1，横轴上的紫色三角形对应于 tRNA type2。两个事件之间 $\%I_b$ 和 t_{off} 存在着显著差异。不同的 RNA 类型可以根据其不同的阻断特性来识别。MspA 呈圆锥形，可以有效的区分出不同的 RNA 类型或结构。

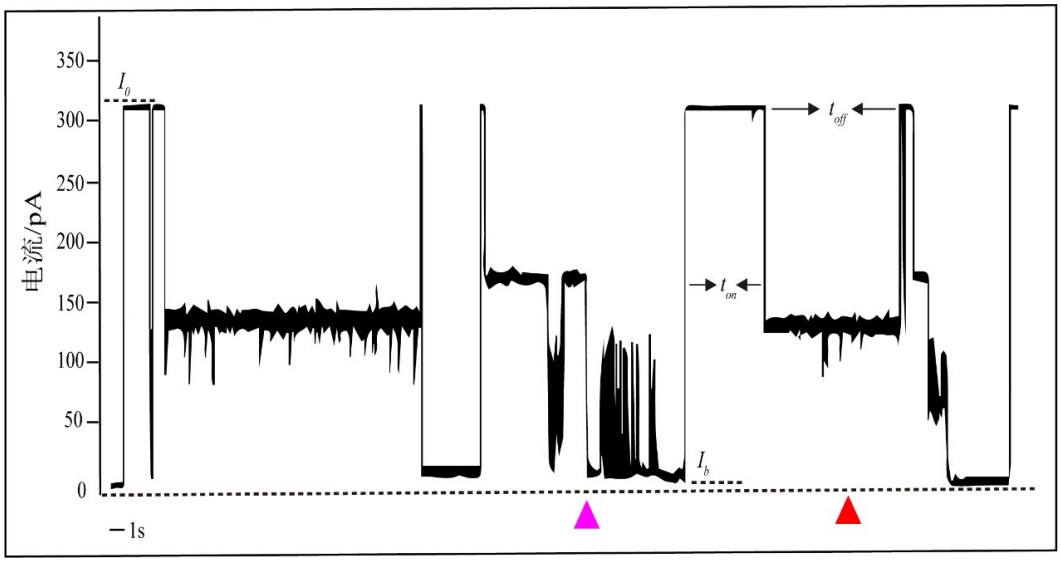


图 2 纳米孔检测 tRNA 分子产生的长序列

RNA 分子检测产生的信号如图 3 所示，其中图 3(a) 为单独 RNA 检测的长序列，长序列中有大量的噪声信号，存在部分小段的 RNA 分子检测序列，如图 3(b) 所示。从图 3(b) 可以看出不同的 RNA 检测信号存在着明显的差别。然而，完整的检测过程产生的信号却如图 3(a) 所示，其长序列容量非常的大且存在着大量的噪声信号，通过传统的分析方法处理往往需要巨大的劳动力成本。因此迫切需要一种可以从长序列中提取短的 RNA 分子序列的工具，并能够依顺序的对 RNA 分子的类型进行识别，机器学习技术可以完美解决这一难题。实现方法为：首先需要分割算法从长序列中分割出短的 RNA 分子序列，之后对 RNA 分子序列进行特征构建和机器学习，进而实现了对混合分析物的 RNA 分子类型的精准识别。

纳米孔检测技术已成为大分子识别的新兴技术，并已实现 DNA（或 RNA）测序、检测和数据存储等任务。现有的纳米孔数据分析策略取决于其具体的应用，因为纳米孔检测信号在传感和测序过程中往往是多样的。现有的分析纳米孔数据的方法主要通过以下两种策略实现的：（1）统计方法，通过统计分析技术识别数据间的统计差异；（2）基于机器学习的方法，这是一种新兴的强大工具。尽管这些方法为纳米孔数据分析提供了便利，但它们却受到了昂贵的计算资源的限制和缺乏专业特征选择的限制。

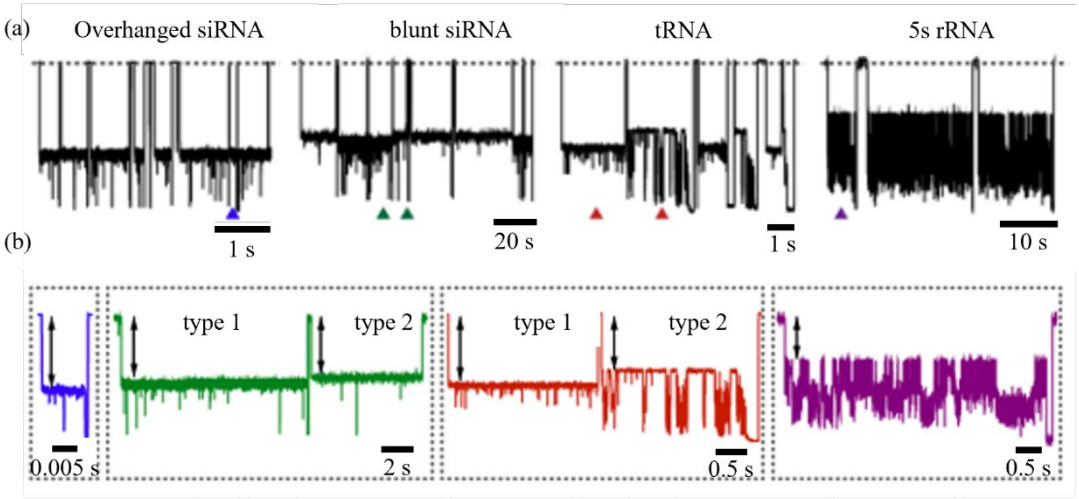


图 3 RNA 分子纳米孔检测结果：(a) 单独的 RNA 分子检测的长序列；(b) 提取的 RNA 分子检测信号

在使用机器学习进行训练之前需要对数据进行特征构建，使得构建出的特征有益于模型的训练。在本章中，分别构建了单个 RNA 序列的的 1 级位置（pos_level 1）、2 级位置（pos_level 2）、波动情况（noise）、停留时间（length）、最小值（min）、最大值（max）、中值（med）、平均值（mean）、标准差（std）、峰度（kurt）和偏度（skew）总共 11 个特征，特征的详细信息如表 1 所示。每个 RNA 序列构建的特征构成了特征向量，11 个特征的定义如图 4 所示。图中示例为 tRNA type2 分子的检测信号，图 4(a) 是原始序列，图 4(b) 为直方图。

表 1 特征构建的 11 个特征对照表

特征	特征简称	特征含义
1 级位置	pos_level 1	直方图对应的第一个高峰位置
2 级位置	pos_level 2	直方图对应的第二个高峰位置

特征	特征简称	特征含义
波动情况	noise	直方图对应的第一个高峰的宽度
停留时间	length	序列的总长度
最小值	min	序列的最小值
最大值	max	序列的最大值
中值	med	序列的中值
平均值	mean	序列的统计学均值
标准差	std	序列的统计学标准差
峰度	kurt	序列的统计学峰度
偏度	skew	序列的统计学偏度

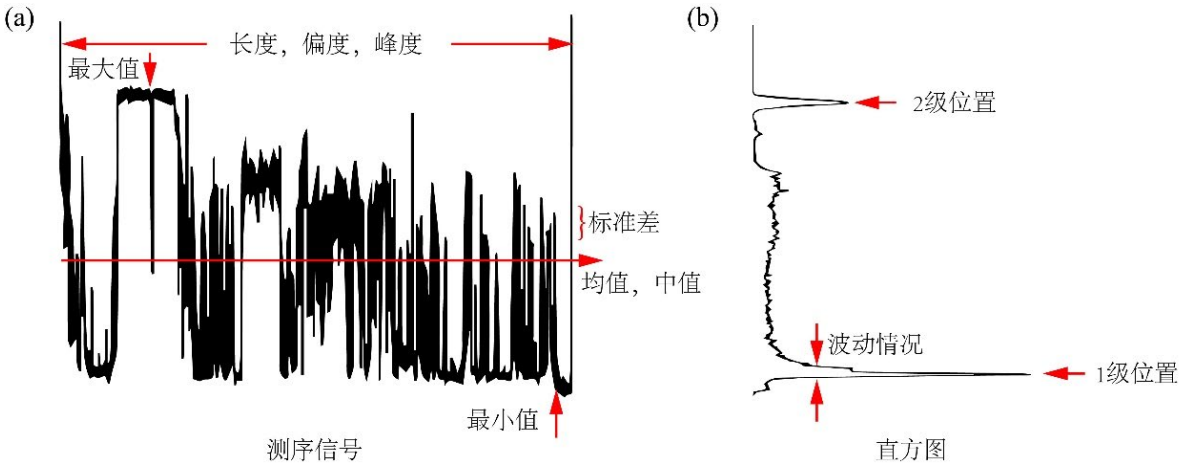


图 4 RNA 分子检测序列的 11 个特征：(a) tRNA type2 分子的检测信号原始序列；(b) 直方图

针对于 1 级位置，2 级位置和波动特征，特别采用了直方图映射的方式来获取，通过多峰值高斯拟合来构建每一个检测序列的特征。图 5 分别展示了 overhanged siRNA、blunt siRNA type1、blunt siRNA type2、tRNA type1、tRNA type2 和 5S rRNA 的 1 级位置和 2 级位置的获取示意图。如图 5(b)，5(c)，5(d) 所示，当直方图中只有一个可识别的高斯峰值时，2 级位置被设置为 0。如图 5(a)，5(e)，5(f) 所示，当识别到 2 个以上的高斯峰时，更接近开孔电流的峰被认为是峰 1，而另两个峰被认为峰 2 和峰 3。根据拟合结果分别确定各峰值的位置和波动。波动的计算通过计算峰 1 的标准差获得。

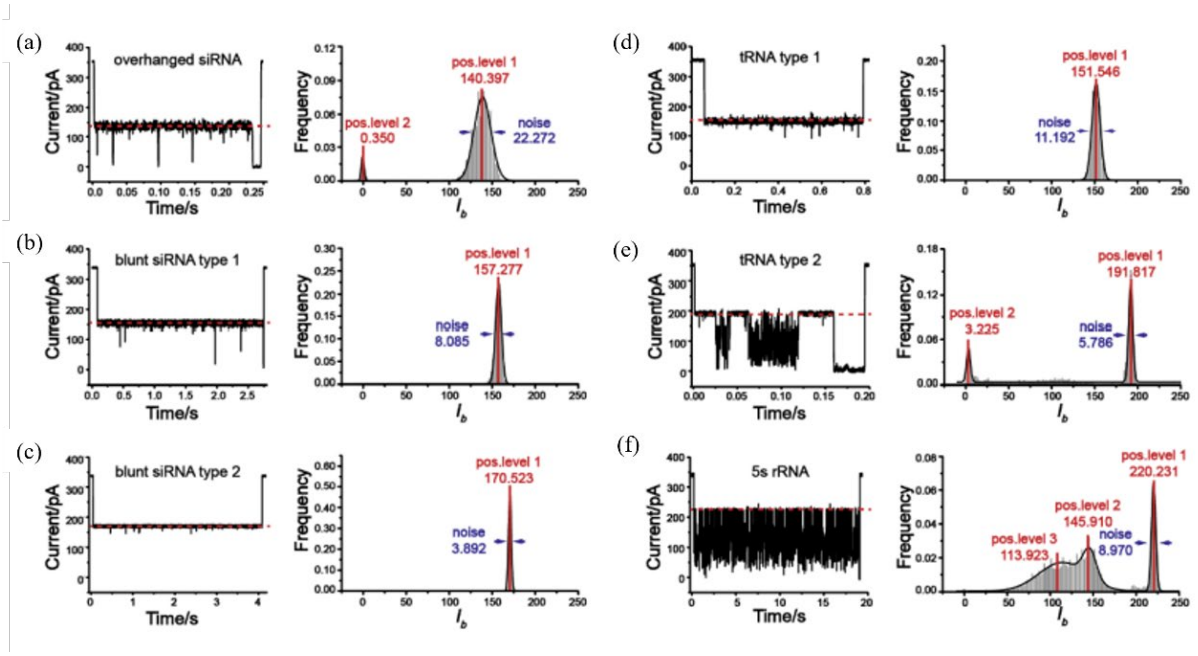


图 5 RNA 分子数据集的 1 级位置和 2 级位置获取示意图：(a) overhanged siRNA；(b) blunt siRNA type1；(c) blunt siRNA type2；(d) tRNA type1；(e) tRNA type2；(f) 5S rRNA

在实验的测试阶段分别输出了模型的分类准确度、特征重要性、混淆矩阵和学习曲线。其中分类准确度由正确分类的样本和总样本的商来计算。

分类准确度的实验结果如表 2 所示，其中随机森林要明显优于其他对比的机器学习算法。最佳算法随机森林在该数据集中获得了 0.939 的准确度，召回率和 F1 分数的度量趋势与准确度相同。

表 2 机器学习算法在 RNA 分子数据集上的性能比较

方法	准确度	召回率	F1 分数
KNN	0.878	0.852	0.850
CART	0.895	0.868	0.867
GradientBoost	0.916	0.910	0.912
Xgboost	0.928	0.917	0.915
Random Forest	0.939	0.937	0.939

本章针对 RNA 分类数据集，采用了特征构建配合传统机器学习的方案进行分析，得出如下结论：

（1）本章提出了针对 RNA 分类数据集的特征构建方案。根据 RNA 纳米孔数据集的特殊性，结合纳米孔领域专家的意见，设计并构建了包括均值和标准差在内的总共 11 个特征。根据特征分布图可以明显观察到构建的特征之间有明显的判别界限，其中部分重要的特征也将直接影响最终的分类结果。

（2）基于构建的特征，利用几种传统的机器学习算法进行分类验证。实验结果表明，相对于其他几种机器学习算法，随机森林算法表现出了更好的性能。因为随机森林是一种集成的分类树结构，其对于多维特征任务具有更好的分类效果。

（3）特征中值，标准差和均值对分类的性能起到了更重要的作用，而最大值却不会对分类性能产生特别大的影响。模型对 5S rRNA 和“其他 other”的分类效果要差，主要是由于这两种序列相对复杂。模型对 overhanged siRNA 和 blunt siRNA type1 的分类效果要好，主要是由于这两种序列相对识别简单。

（4）本章提出的特征选择方案同样适用于其他类似的纳米孔数据分析任务中。期望这种手工特征构建方法配合传统机器学习分类的方案能够应用到更多的纳米孔数据分析中，从而减少生成的纳米孔检测序列所需的手动分析量。

二、基于序列转化策略的纳米孔检测 RNA 类型识别方法

在前一章中，研究了利用 MspA 纳米孔的纳米腔在单分子水平上检测 RNA 类型。MspA 纳米孔可以直接区分许多低分子量的 RNA 类型，如 overhanged siRNA、blunt siRNA、tRNA 和 5s rRNA。在前一章中，应用了随机森林算法对不同 RNA 分子的检测信号进行自动分类。虽然随机森林算法可以获得显著的性能，但它仍需要领域知识来进行特征选择，这限制了算法的有效性。

经典机器学习方法通常需要复杂的特征提取方式，这往往需要巨大的时间成本。相比之下，深度学习是分析纳米孔数据的另一种有效策略。深度学习被认为是一种非常有前途的特征提取方法，因为该方案是端到端的，不需要人工参与特征提取的过程中。最近，深度学习在各种应用领域中取得了巨大成功，包括计算机视觉（Computer Vision，CV）、自然语言处理（Natural Language Processing，NLP）和大数据分析等。特别地，纳米孔检测信号存在着不等长的情况，因为不同的分子通过纳米孔保持时间存在差异。已有相关研究将深度学习用于纳米孔不等长序列分析中错误！未找到引用源。，其数据集由 ONT 生成，任务是对八类编码进行分类。由于其检测序列彼此长度差异较小，其通过插值技术来使其由不等长序列变成了等长序列。然而，对于第二章中的 RNA 分类数据集，其序列间的长度差异可能达到数万点，简单的通过插值技术会损失大量有用信息。

因此，为了克服以上这些限制，提出了一种策略，在保证不损失用于分类的关键信息的前提下，将不等长序列转化为等长序列或者图像。首先介绍下降不等长序列转化成等长序列的方法，这里将该策略定义为序列到序列（Sequence To Sequence，S2S）。基于此，设计了一个基于深度学习的网络模型（Sequence To Sequence Network，S2Snet），其可以实现 RNA 分子数据特征的自动提取和高精度分类。

生成的纳米孔数据需要进行特征提取和分析，这些往往受到了劳动成本的限制。并且生成的纳米孔数据具有非常高的复杂性和不规律的噪声，因此需要更专业的策略来提取其序列特征。迫切需要一种智能和用户友好的机器学习算法来应对这些挑战。

纳米孔数据分析领域有两个数据分析过程，如图 6 所示。在图 6 的顶部，传统的数据分析策略需要结合特征工程和特征选择，并使用经典的机器学习算法（如 RF 或 SVM）来构建模型。之前使用了这种纳米孔数据分析策略，其缺点是需要更多的领域专家来辅助特征提取。这增加了研究成本，如人力和财力。相反，在图 6 的底部，深度学习方法策略是端到端的，直接输入序列并输出分类结果。该策略具有良好的经济性，并能够取得更好的分类效果。

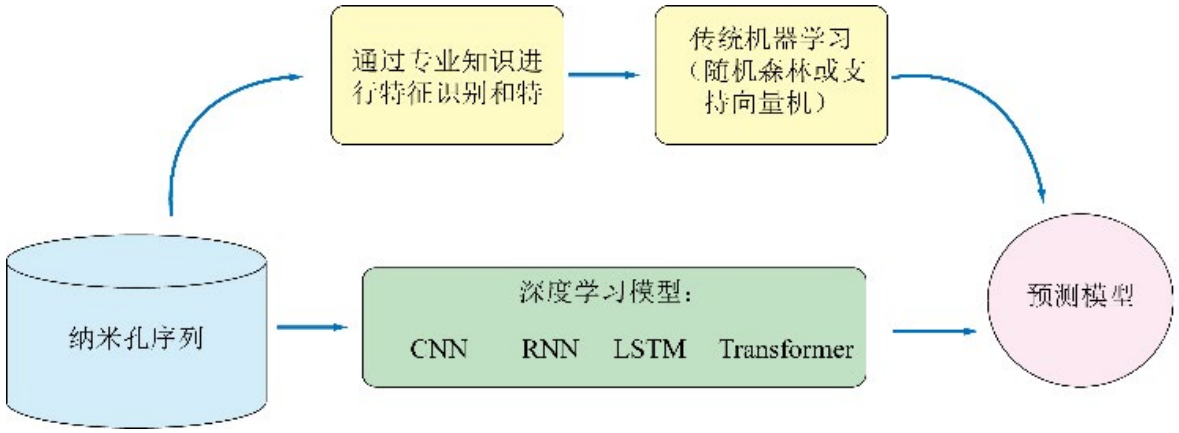


图 6 纳米孔数据分析示意图

只考虑单个 RNA 分子进入纳米孔的情况。不同 RNA 的子序列如图 7 所示。红线序列是 tRNA type1（T1），深蓝色线序列是 tRNA type2（T2），青色线序列是 overhanged siRNA（O），黄色线序列是 5s rRNA（R），紫色线序列是 blunt siRNA type1（B1），蓝色线序列为 blunt siRNA type2（B2）。图中展示了每个 RNA 分子的两个序列信号，因为来自相同 RNA 类型的不同 RNA 分子的序列信号可能不同（图 7 中的下标 1 和下标 2）。特别是，由于 RNA 的三级结构的不同，tRNA 和 blunt siRNA 具有不同的进孔方式，将产生两种不同的信号类型。因此，这两种 RNA 产生两种类型的信号，称为 type1 和 type2。当 RNA 分子进入纳米孔时，由于堵塞孔隙的分子碎片大小不同，堵塞幅度 I_b 也不同。从图中可以看出，tRNA type1 和 5s rRNA 序列形状与其他 RNA 类型有着显著的不同。对于其他 RNA 类型，可以使用阻断百分比 $\%I_b$ 和阻断幅度 I_b 来区分。blunt siRNA type1 的 $\%I_b$ 高于其他三种 RNA 类型，并且 blunt siRNA type1 在序列尾部具有阻断幅度 I_b 为 0 的特征。overhanged siRNA 和 tRNA type1 有许多毛刺，overhanged siRNA 的 $\%I_b$ 略高于 tRNA type1 的 $\%I_b$。

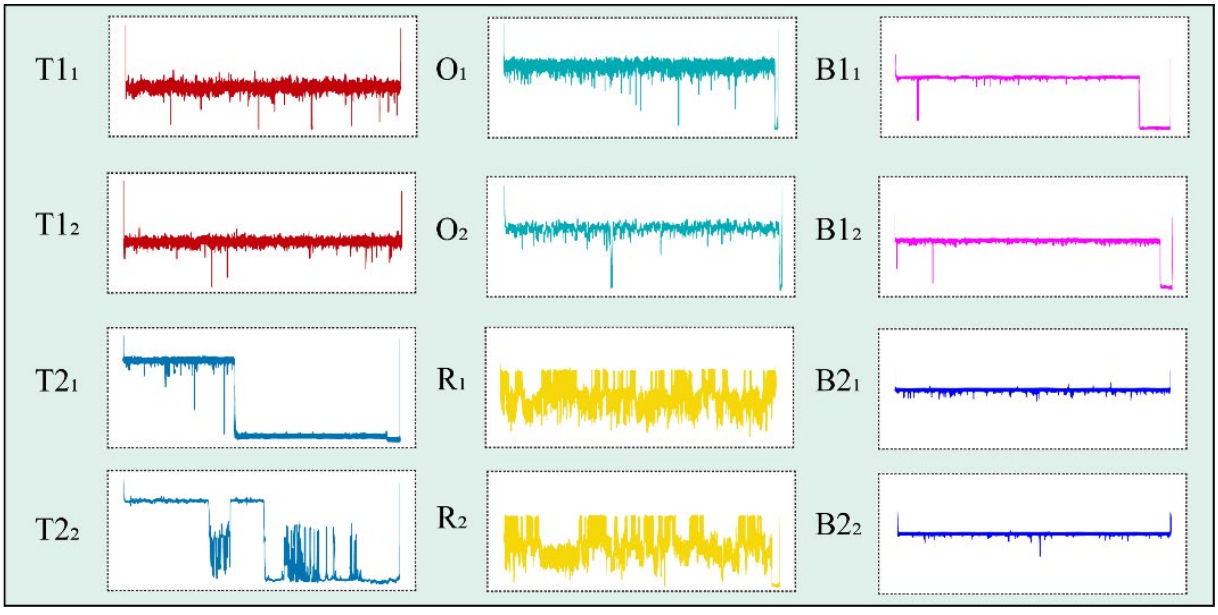


图 7 MspA 纳米孔检测产生的 RNA 原始序列检测结果例子

在数据预处理阶段，S2S 模块将不等长原始输入序列转换为等长序列。根据 S2S 模块，设计序列到序列神经网络 S2Snet 来预测 RNA 类型，其结构如图 8 所示。在深度学习阶段，选择卷积神经网络作为 S2Snet 基模块。如图 8(a) 所示 S2Snet 由预处理模块、卷积模块和输出模块组成。典型的卷积神经网络由两部分组成，如图 8(b) 和图 8(c) 所示。首先，将一系列卷积模块应用于 S2S 变换后的序列。在图 8(b) 中，“Conv”是提取具有局部结构（如峰值或步长）的特征的卷积。在卷积步骤之后，应用批量归一化（Batch Normalization，BN）将数据归一化为零均值和单位方差。最后，将激活函数应用于称为整流线性单元（Rectified Linear Unit，ReLU）的分段函数。该非线性函数对于学习特征之间的非线性关系是必要的。为了捕获变换后序列的关键部分，在第一和第二级卷积模块之间添加了注意力模块，其结构如图 8(d) 所示。矩阵 Q、K 和 V 是注意力模块的输入，MatMul 是矩阵乘法。Scale 可以减少相关矩阵的方差，有利于模型训练。Mask 用于屏蔽不必要的信息，SoftMax 用于输出归一化概率。输出模块用来对前端输入的特征进行类别的预测。

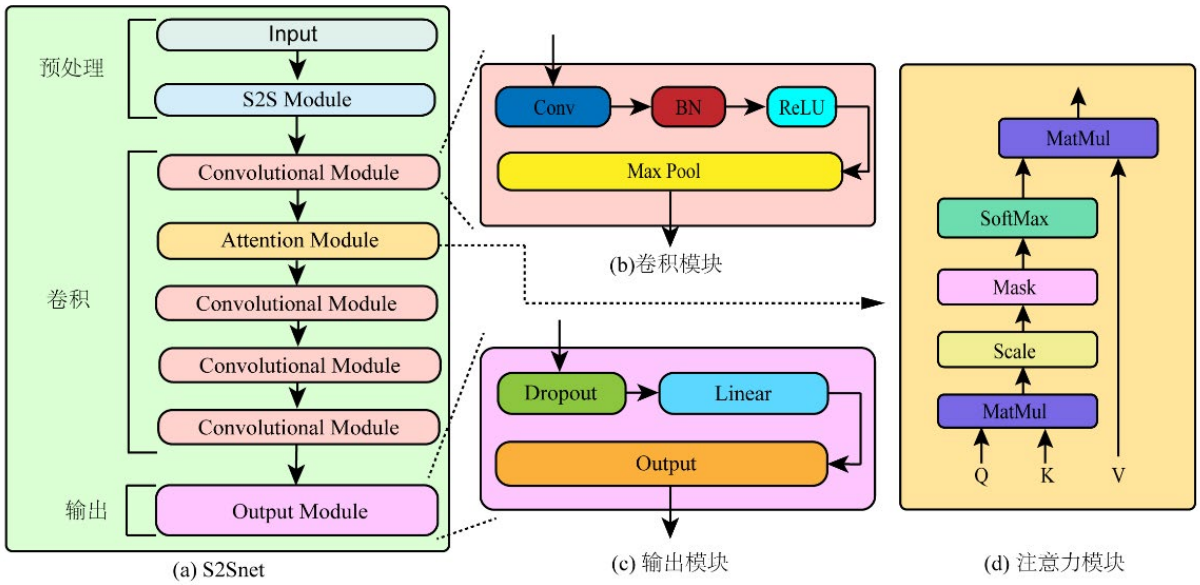


图 8 S2Snet 的网络结构图：(a) S2Snet 总体结构图；(b) 卷积模块的示意图；(c) 输出模块示意图；(d) 注意机制模块示意图

S2Snet 的 RNA 数据集上的分类准确度为 0.957。该性能高于第二章中的最佳算法 Random Forest 方法（在该数据集中获得了 0.939 的准确度）。召回率和 F1 分数的度量趋势与准确度相同。表 3 的最后一列是总训练和测试时间的平均值。S2Snet 的时间成本略长于 Random Forest，因为 S2Snet 是基于深度学习开发的，需要更多的训练时间来捕获输入序列中的局部模式。与手动提取特征的方式相比，S2Snet 减少了特征提取的人工和时间成本。针对于 RNA 数据集，在开始设计机器学习模型之前，首先要求专家挑选那些影响性能的特征，并基于此设计一种新的特征提取算法以获得其特征值，例如每个序列的阻塞幅度。该操作需要反复的迭代，需要 1-2 周才能实现期望的性能。相对而言，S2Snet 可以自动提取用于高精度分类的特征（该过程是实时的，整个数据集大约只需要几分钟）。因此，S2Snet 比需要手动提取特征的经典机器学习算法花费更少的时间。

表 3 S2Snet 与其他传统机器学习算法的性能比较

方法	准确度	召回率	F1 分数	时间（平均）
KNN	0.878	0.852	0.850	1m54s
CART	0.895	0.868	0.867	2m8s
GradientBoost	0.916	0.910	0.912	3m35s
Xgboost	0.928	0.917	0.915	9m43s
Random Forest**	0.939	0.937	0.939	11m53s
S2Snet**	0.957	0.953	0.953	36m38s

获得图像，需要使用深度学习模型，而 Transformer 比 CNN 更适合图像数据。在过去两年中，Transformer 框架已成为学习文本自我监督表示的强大架构。与 CNN 相比，Transformer 具有更好的优势，并在 NLP 任务中取得了良好的成功。Transformer 在计算机视觉领域的早期兴起主要用于挖掘序列信息，例如一些视频任务。近年来，ViT 方法因其具有更好的更大感受野特性而扩大了 Transformer 的使用范围。CNN 卷积的一个问题是感受野相对有限。为了扩展网络的感受野，需要卷积池堆叠多层结构。问题是，以某个中心为原点，感受野向外高斯衰减。因此，CNN 通常只关注图像中的一两个重要部分。上述的 S2Snet，则是使用注意力机制来抵消这种影响。Transformer 可以使用全局有效信息，多头注意力机制确保网络可以关注多个关键部分，每个关键部分都是独立的关注点。因此，这里使用 Transformer 处理纳米孔序列转换后的图像数据。

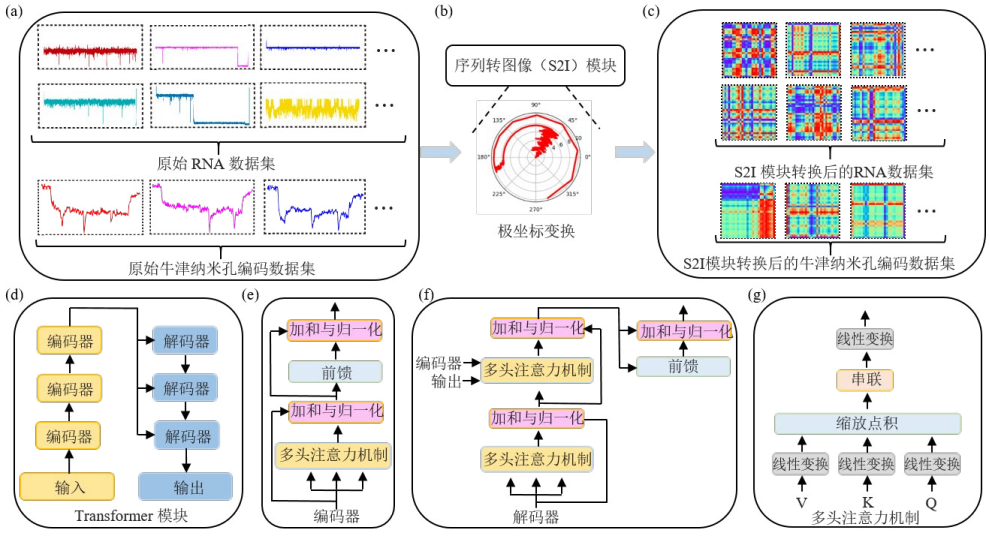


图 9 T-S2Inet 网络流程图：(a) 来自原始 RNA 纳米孔数据集和原始 ONT 编码纳米孔数据集中的示例样本；(b) S2I 模块的示意图；(c) S2I 模块转化的 RNA 数据集和转化的 ONT 编码数据集获得示例样本；(d) T-S2Inet 中的 Transformer 模块的示意图；(e) T-S2Inet 中的编码器模块的示意图；(f) T-S2Inet 中的解码器模块示意图；(g) T-S2Inet 中的多头注意力模块的示意图

因此，这里提出了一个从一维序列到二维图像的转换工具（Sequence to Image，S2I）模块，并为该转换模块设计了一个基于 Transformer 的网络结构 T-S2Inet。在 RNA 分子数据集和 ONT 编码数据集上进行了实验。

使用 S2I 模块，执行 T-S2Inet 来预测 RNA 类型和 ONT 编码类型。在数据预处理阶段，S2I 模块将原始输入序列转换为图像，如图 9(c) 所示。T-S2Inet 的特殊结构是图 9(d) 所示的 Transformer 模块，它由编码器和解码器组成。一个典型的编码器模块由三部分组成，如图 9(e) 所示。首先，将多头注意力模型应用于 S2I 变换后的图像。其次，“加和与归一化”和“前馈”应用于多头注意力模块的输出。解码器模块也由三个部分组成，如图 9(f) 所示。首先，两个多头注意力模块和“加和与归一化”模块串联连接到解码器的输入端。编码器的输出应用于第二个多头注意模块。第二，“加和与归一化”和“前馈”应用于第二个“加和与归一化”模块的输出。为了捕获转换图像的关键部分，使用 Transformer 模块。Transformer 的关键模块是多头注意力模块，如图 9(g) 所示。矩阵 Q、K 和 V 是注意力模块的输入，缩放点积注意力应用于三个输入。串联应用于缩放点积的输出，该操作将输入和输出保持在同一维度。

本章针对 RNA 分类数据集，提出了两个深度学习模型进行了数据分析，得出如下结论：

（1）提出了两个序列转化方案，一种是序列转化为序列，另一种是序列转化为图像。这两种序列转化的方式都实现了将不等长序列转化为等长的序列或等尺寸的图像，之后则可以通过深度学习模型进行学习。

（2）基于等长的序列，提出了基于注意力机制的深度学习模型，其可以避免人工特征提取造成的成本增加，也实现了更高精度的分类。实验结果分析表明，该模型具有非常好的适用性，模型中使用的 Focal loss 损失函数对分类起到了积极的作用。通过 t- 检测对性能提高进行了分析，结果表明比之前的随机森林方法提升明显。在 ONT 编码数据集上验证了该转化策略的有效性。

（3）基于等尺寸的图像，提出了基于 Transformer 的深度学习模型，其同样实现了更高精度的分类。实验结果分析表明，该方法比基于注意力机制的深度模型方法还要好。通过 t- 检测对性能提高进行了分析，结果表明基于 Transformer 的深度学习模型提升明显。二维图像数据相对于一维序列任务，有着更多的局部特征有利于模型的分类。在 ONT 编码数据集上验证了该模型的有效性。

（4）分析了传统机器学习的特征提取和深度学习特征提取的区别，探索了深度学习在纳米孔数据分析中的适用性。实现了原始纳米孔序列长度严格不等导致的无法使用深度学习模型进行分析的困境，提高了检测序列识别精度。

三、基于主动学习的纳米孔检测数据识别方法

最近，纳米孔检测技术作为单分子检测的主流技术之一，因为其对大分子具有长的读取时间和单碱基的高分辨率优势。由于纳米孔数据的量级巨大，传统的分析方法需要巨大的成本。因此，越来越多的机器学习算法开始逐步应用到纳米孔数据分析领域中，包括前面的几章提出的机器学习方法。这些机器学习算法在将纳米孔检测应用于各种生物任务方面实现了前所未有的突破。机器学习方法的成功通常是归功于对巨大标注数据的重复迭代训练。然而，数据标注通常需要较高的样本标注成本（即，需要具有广泛专业知识的专家进行手动标注）。为了实现标注效率的指数加速，率先将主动学习技术引入到纳米孔序列分析领域。主动学习是一种机器学习方法，它选择并标注复杂样本，以有限的成本获得高度准确的预测模型。主动学习已经应用于许多跨学科领域，如药物发现、材料设计和其他新兴学科。由于纳米孔数据分析是新兴的研究方向，因此在纳米孔领域基本上没有主动学习的相关研究。由于纳米孔原始序列的复杂性，样本标注需要更多的人力。纳米孔序列不仅包含有效的分子检测信号，还包含噪声信号。例

如，RNA 分子分类数据集包含六个 RNA 分子检测信号和一个噪声信号。结果表明，三种 RNA 分子的检测信号形状相似，噪声信号具有各种奇怪的形状。由于纳米孔原始序列的复杂性（信号混叠和噪声混淆），纳米孔数据分析领域中的样本标注需要更多的劳动力成本。例如，在标注 RNA 分子分类数据集的过程中，不仅需要从所有检测信号中过滤掉噪声信号，还需要对所有 RNA 检测信号进行分类。为了将主动学习技术应用于纳米孔领域的数据分析中，尝试将标注的样本更趋于真实标注。为了尽量减少这些样本被错误标注情况的发生，特别请了三位标注专家对整个 RNA 数据集进行标注。对于那些难以标注的样本进行头脑风暴，并结合三个专家的标注的建议，使标注结果尽可能准确。ONT 数据集也面临同样的困境，例如前面所述的编码数据集。由于其序列具有信噪比差、峰值振幅变化快、和峰值重叠等情况，人工标注也会存在误标注的情况。此外，整个数据集中有 58178 个样本，这大大的增加了标注专家的工作量。主动学习技术旨在从未标注的数据集中选择最有价值的样本，并将其交给代理（例如，人工标注）进行标注，在保证性能的前提下尽可能降低标注成本。因此，主动学习可能成为未来纳米孔领域各种生物任务的有效算法。

为了克服纳米孔数据集需要大量标注样本的高成本困境，应用了主动学习技术来验证其在纳米孔领域的有效性。将主动学习技术应用于 RNA 分子分类数据集和 ONT 编码数据集。由于纳米孔数据的复杂性，在主动学习技术中添加了偏置约束来优化其样本选择策略。

为了进一步说明主动学习策略如何在纳米领域进行应用，本小节将简要描述了主动学习策略应用于 RNA 数据集的整个流程，其流程如图 10 所示。在 ONT 编码数据集中应用主动学习策略的过程是类似的。

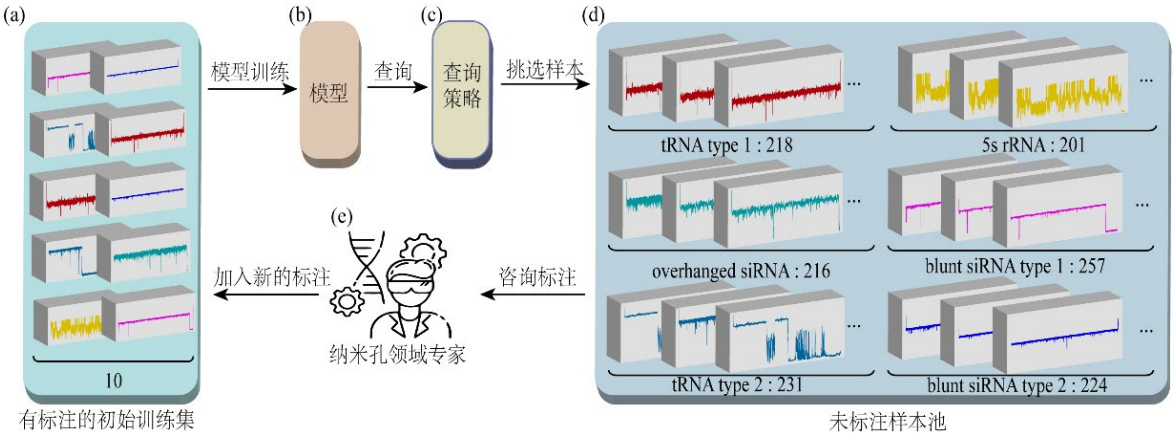


图 10 主动学习应用于 RNA 纳米孔数据集示意图：(a) 标注样本池 L；(b) 机器学习模型 C；(c) 查询函数 Q；(d) 未标注样本池 U；(e) 可以进行对查询后样本标注的纳米孔领域专家 S

具体而言，应用于 RNA 数据集的主动学习策略过程由五个部分组成，分别对应于图 10(a) 至图 10(e)。在本章的实验中，初始标注样本池 L 被设置为 10 个样本来训练模型，如图 10(a) 所示。在某些情况下，机器学习模型 C 被设置为 RNA 类型预测实验中的第二章的随机森林算法和第三章的 S2Snet 方法，如图 10(b) 所示。相应地，在 ONT 编码数据集中将 C 设置为 QuipuNet 方法。查询函数 Q 包含八种常见主动学习策略：委员会查询 QBC，查询信息和代表性示例（QUerying Informative and Repre-sentative Examples，QUIRE）是基于池的主动学习策略，密度（Density）是基于密度的采样策略，预期误差减少的查询（Expected Error Reduction，EER），学习主动学习（Learning Active Learning，LAL），自步主动学习（Self-Paced Active Learning，SPAL），不确定性采样

UNC，如图 10(c) 所示。显然，未标注的数据集池 U 是训练数据集的剩余部分（即没有最初的十个标注样本的剩余样本），如图 10(d) 所示。图中每个矩形块下方的数字表示每个类别中的样本总数。特别地，在纳米孔领域，监督员 S 是纳米孔领域的专家，具有高度的专业知识，能够标注未标注的样本，如图 10(e) 所示。整个学习过程是一个连续和迭代的过程，当达到最佳测试性能（准确度）时，学习过程将停止。

实验部分采用了两个数据集（三个模型）对不同主动学习策略进行验证。八种主动学习策略分别为：QBC、Random、QUIRE、Density、LAL、SPAL、EER 和 UNC。进行了一系列实验，以展示两个数据集（三个模型）在不同主动学习策略下的性能。

对于 RNA 分类数据集，采用随机选择的方法初始 10 个样本来训练随机森林模型和 S2Snet 模型，并使用不同的主动学习策略在每次迭代中选择 10 个样本。因此，随着迭代时间的增加，分类性能逐渐提高。两个数据集（三个模型）在不同主动学习策略下的实验结果如表 4 所示。结果表明，UNC 比其他方法表现的更好。其使用了两个度量 SR（与全样本相比标注成本降低的百分比）和时间。

表 4 在 RNA 分类数据集和 ONT 编码数据集中主动学习策略的性能比较

方法	RNA 分类数据集（RF）		RNA 分类数据集（S2Snet）		ONT 编码数据集（QuipuNet）	
	SR↑	Time↓	SR↑	Time↓	SR↑	Time↓
Random	0.012±0.003	79.04s±2.3s	0.024±0.001	32min42s±24s	0.058±0.007	2h39min36s±35s
QBC	0.834±0.021	762.58s±7.5s	0.836±0.143	4h21min55s±10s	0.448±0.063	17h39min18s±73s
QUIRE	0.023±0.002	1288.06s±10.3s	0.031±0.011	8h16min17s±36s	0.089±0.001	26h56min34s±65s
Density	0.033±0.005	101.58s±5.6s	0.037±0.002	59m11s±8s	0.134±0.021	3h41m23s±35s
LAL	0.045±0.017	154.83s±3.4s	0.054±0.017	1h15m21s±7s	0.198±0.011	4h56m43s±42s
SPAL	0.563±0.113	564.55s±6.3s	0.574±0.158	3h43m25s±15s	0.334±0.018	14h32m12s±51s
EER	0.322±0.134	379.54s±8.2s	0.342±0.129	2h11m43s±6s	0.212±0.042	8h14m32s±32s
UNC	0.844±0.045	88.88s±1.4s	0.855±0.062	40m10s±34s	0.553±0.033	2h49m10s±21s

分类器测试精度与迭代次数之间的关系如图 11 所示。从图中可以看出，UNC 在 RNA 分类数据集相比于其他主动学习算法要明显有优势，即需要很少的样本训练即可以达到最优性能。图 11(a) 和图 11(b) 中，当标注的数量约为 15% 时，UNC 就实现了随机森林和 S2Snet 在全数据集下训练所达到的性能。

图 11(c) 中展示了在 ONT 编码数据集上应用主动学习策略下的分类器测试精度与迭代次数之间的关系。结果表明，UNC 同样优于其他方法，但是并没有非常大的优势。当标注的数量约为 50% 时，UNC 就实现了 QuipuNet 在全数据集下训练所达到的性能。造成这种差异的两种主要情况：（1）数据集的不同容量；（2）数据集的任务不同。ONT 编码的任务相对困难，RNA 分类的任务相对容易。

显然，UNC 的性能优于其他主动学习方法。此外，实验结果表明，主动学习策略可以有效地降低样本的标注率，并保持全数据集训练下的性能。事实上，主动学习策略完全可以应用于纳米孔领域，其可以大量降低专家的标注成本。

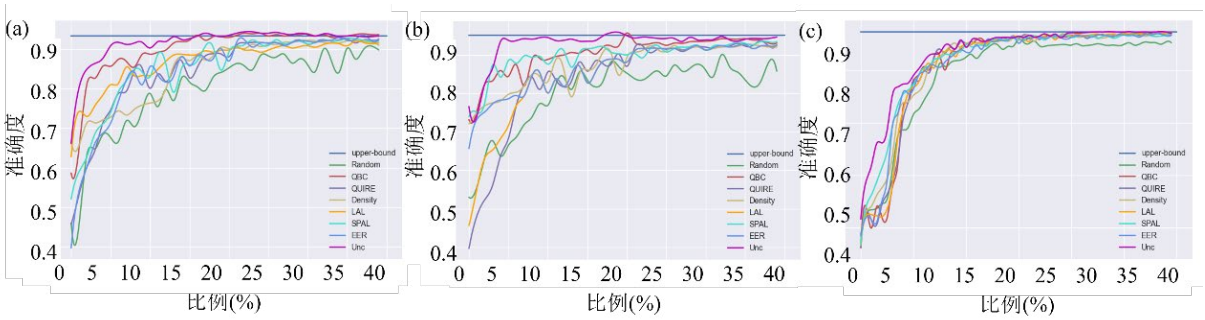


图 11 分类器测试精度与迭代次数之间的关系曲线：(a) RNA 分类数据集应用随机森林在主动学习上测试精度与迭代次数的变化曲线；(b) RNA 分类数据集应用 S2Snet 在主动学习上测试精度与迭代次数的变化曲线；(c) ONT 编码数据集应用 QuipuNet 在主动学习上测试精度与迭代次数的变化曲线

本章针对 RNA 分类数据集和 ONT 编码数据集，提出了一个主动学习模型进行了数据分析，得出如下结论：

（1）将主动学习首次应用于纳米孔领域，根据纳米孔数据的特殊性，提出了偏置约束以改进主动学习中的样本选择策略。其可以有效的保证挑选的样本有利于模型的训练，使得模型更快的达到最优性能。

（2）实验过程首先随机选择 10 个样本进行训练，然后每次迭代从未标注样本中选择 10 个样本进行标注，重新训练模型。实验结果表明，主动学习策略可以显著降低标注成本。并且可以达到全监督下训练模型下的最优测试性能。

（3）设计的主动学习方法在其他纳米孔数据集上进行了验证。实验结果表明，设计的主动学习方法同样可以实现降低标注成本的目的。在机器学习阶段，主动学习策略可以帮助纳米孔专家了解哪些样本对分类任务至关重要，可以指导标注专家进行标注。

（4）提出的方法可以优化纳米孔领域的样本标注策略，使用主动学习技术可以有效的降低纳米孔庞大数据所带来的标注成本。分析了主动学习在样本标注成本上积极作用效果，探索了主动学习在纳米孔数据分析中的适用性。

四、基于无监督深度学习的纳米孔测序 O6- 甲基鸟嘌呤识别

基因组测序有助于提高人类对疾病的认识，如何辨别与人类疾病相关的遗传危险因素至关重要。O6-CMG 在一定情况下会触发突变，引起 DNA 复制时的编码错误，这种错误往往与胃肠道肿瘤相关。近年来，纳米孔测序技术已经成为了一种新兴的大分子感知识别技术，可用于 DNA 测序。纳米孔装置由两个充满液体的储层组成，由纳米级的孔道相连接。储层上下两侧的带电粒子产生正负电位差，带电粒子由于电位差的驱动而通过中间的孔道，因而产生了可以反映分子特性的电流信号。

基于上述研究，提出了一种纳米无监督深度学习方法 （nano-unsupervised-deep-learning，nano-UDL）用于纳米孔数据分析，并在 O6-CMG 纳米孔数据集上验证了方法的有效性。提出的 nano-UDL 方法可以自动识别突变序列，且无需人工标注标签。nano-UDL 方法首先采用自动编码器从原始纳米孔数据中提取序列特征；然后，对提取的特征进行聚类；最后，实现了对 O6-CMG 的高精度识别。拟定的对比方法有 K 均值聚类（K-Means）、基于密度的噪声应用空间聚类（density-based spatial clustering of applications with noise，DBSCAN）、均值漂移（MeanShift）、凝聚层次聚类错误！未找到引用源。（Agglomerative）、谱聚类（SpectralClustering）等。期望通过与上述方法进行对比研究，能够证实方法的有效性，为今后纳米孔突变检测研究奠定理论研究基础，为更多

类似的生物信息相关研究问题提供研究思路。本章的主要创新思路如下：

（1）nano-UDL 方法是一种无监督学习方法，使用深度自动编码器提取特征，随后使用 MeanShift 聚类算法对提取的特征进行分类；

（2）对于 O6-CMG 纳米孔数据集，本章拟通过超参数敏感性验证和消融实验证实 nano-UDL 方法的稳定性，并通过对比实验的结果验证了 nano-UDL 方法可以达到提升精度的目的，并能够自动对 O6-CMG 位点进行识别。

nano-UDL 方法用于精确定位 O6-CMG 纳米孔数据流程示意图如图 12 所示。

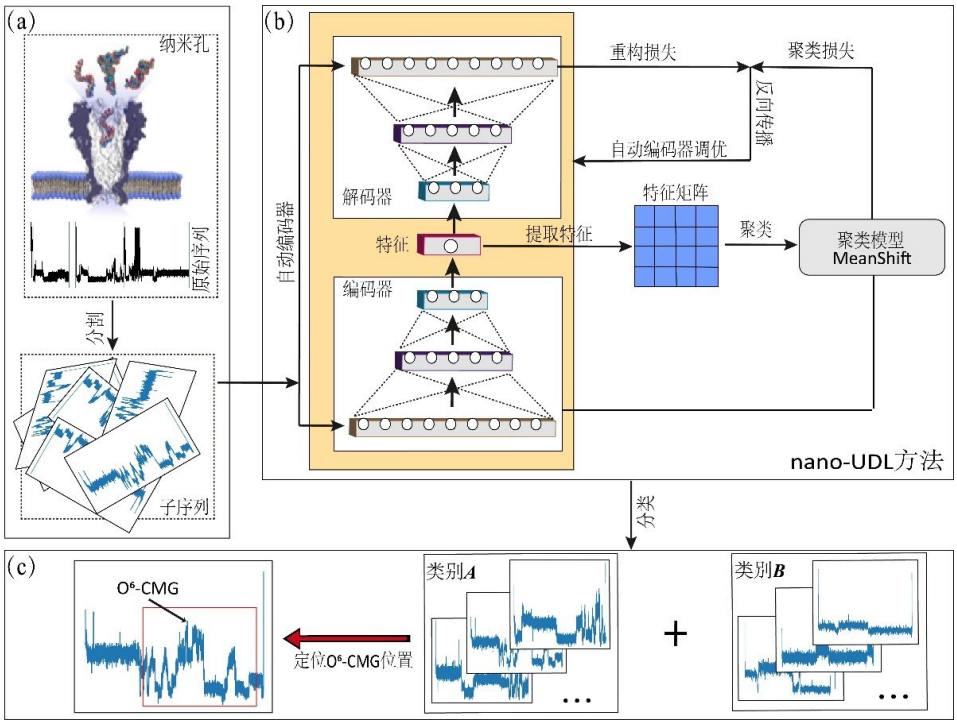


图 12 nano-UDL 用于精确定位 O6-CMG 纳米孔数据示意图：(a) 数据生成过程；(b) nano-UDL 方法的构建过程；(c) 对数据进行分类并精准定位 O6-CMG 位点

整体的流程分为以下三个阶段：

（1）首先，对纳米孔采集的原始序列进行分割，通过分割算法可以切割出若干子序列，如图 5.1(a) 所示；

（2）其次，nano-UDL 方法模型的构建阶段，包含了自动编码器提取特征用于聚类模型聚类，然后联合优化聚类损失与重构损失以实现自动编码器参数调优，如图 12(b) 所示；

（3）最后，将调优后的 nano-UDL 模型用于 O6-CMG 数据集聚类，分别形成了包含 O6-CMG 位点的序列簇（定义为类别 A）和不包含 O6-CMG 位点的序列簇（定义为类别 B）。只需对 A 类别中的序列进行计算即可实现对 O6-CMG 位点的精准识别，如图 12(c) 所示。

对不同的算法的性能进行了定量和定性的评估，如图 13 所示，其中前 5 种对比方法的输入特征是通过人工手动获取，提取的特征包括诸如均值、方差等统计信息特征。对比方法的输入特征是相同的，而所提出的 nano-UDL 方法输入的特征是通过如图 12 所示的自动编码器优化获取。通过观察发现，在 NMI、ARI、Kappa、ACC 等所有评价指标上，nano-UDL 方法均显著高于其他方法，并在 O6-CMG 数据集上表现出良好的聚类性能。这是因为 nano-UDL

方法使用多层自动编码器作为特征提取器，能够捕捉到非深度模型无法获得的局部特征。此外，图 13 中前 5 种方法都是非深度模型，nano-UDL 方法是基于深度学习的。由此可以观察到基于深度模型的方法明显比非深度的方法能发挥更好的性能。

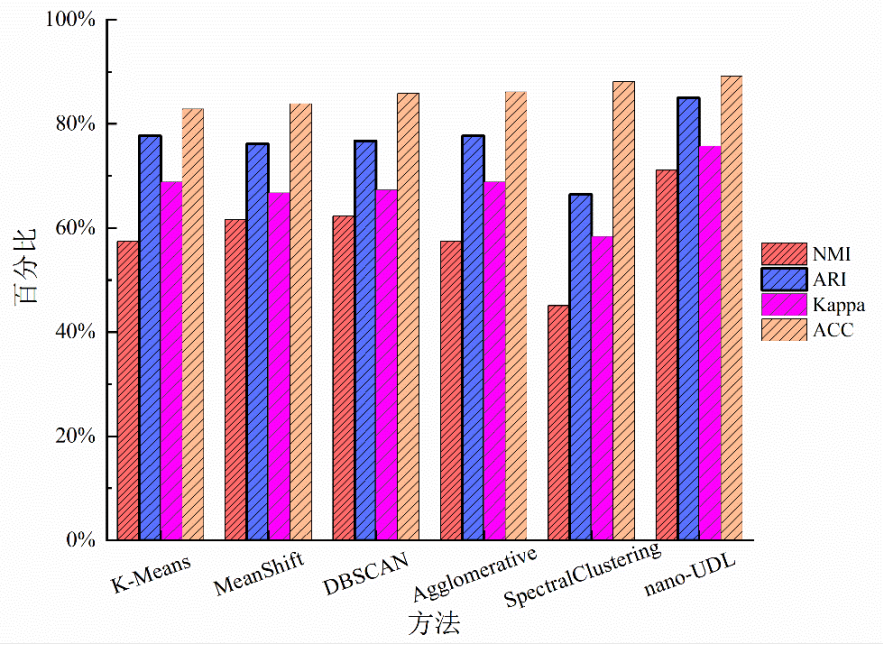


图 13 O_6-CMG 纳米孔数据集聚类性能比较

本章针对 O6-CMG 纳米孔数据，提出了一个无监督深度学习模型进行了数据分析，得出如下结论：

（1）提出了一个无监督深度学习模型对纳米孔测序的 DNA 羧甲基化数据进行识别。该模型通过自编码器进行特征提取，在通过聚类方法将特征进行聚类，将数据点聚类到一个可以通过联合优化的特征空间中，最终实现精准识别突变位点的目的。

（2）为了逼近目标分布，该方法通过最小化编码器的聚类损失和解码器的重构损失进行迭代训练。该方法可以看作是一种不需要任何人工干预的启发式工具，它可以用来处理原始纳米孔数据。这种无监督的策略可以用于更多的纳米孔数据中，可以实现无标注样本的机器学习。

（3）实验结果发现，所提出的方法在提高性能的同时，对超参数设置也具有鲁棒性。这种无监督深度学习策略可以在减少人工计算成本的情况直接检测出甲基化突变位点，实验结果在一定条件下要好于全监督的机器学习方法。

（4）提出的方法可以方便生物专家对 DNA 突变位点的快速定位，对未来更多类似的从 DNA 测序中精准定位突变位点的任务提供一个解决方案，方便指导遗传诊断。所设计的方法可以看作是一种不需要任何人工干预的启发式工具，它可以用来处理原始纳米孔数据。

作者简介

关晓宇，2023 年于南京航空航天大学获得计算机科学博士学位。研究领域包括机器学习、模式识别、生物信息学和分子医学分析。目前主要从事机器学习用于生物纳米孔数据分析。目前入职于南京航空航天大学计算机科学与技术、人工智能学院科研博士后。

指导老师

张道强，分别于 1999 年和 2004 年获得南京航空航天大学计算机科学学士和博士学位。他于 2004 年加入南京航空航天大学计算机科学与工程系担任讲师，现为教授。他的研究兴趣包括机器学习、模式识别、数据挖掘和医学图像分析。在这些领域，他发表了 200 多篇科学论文，被 Google Scholar 引用超过 12,000 次。他是国际模式识别协会 (IAPR) 的会员。

面向深度 Web 站点指纹攻击的防御技术研究

——2024 年江苏省计算机学会优秀硕士论文奖

作者：肖桂

单位：东南大学网络空间安全学院

指导老师：凌振

论文摘要

近年来，随着用户对隐私保护问题的日益关注，世界各地成千上万的用户采用 Tor 匿名通信系统来保护他们的通信隐私。然后，在机器学习技术快速发展的背景下，研究表明这种高度匿名性的网络也容易受到基于机器学习或深度学习模型的 Web 站点指纹（Website Fingerprinting，WF）攻击。为了进一步保护用户隐私，针对 Tor 的 Web 站点指纹防御技术成为了近年来的研究热点。

目前针对主流的 WF 攻击技术的防御方法主要分为基于流量特征抑制的 Web 站点指纹防御技术和基于流量特征整形的 Web 站点指纹防御技术。这些防御方法旨在通过主动改变流量特征抵御对 Web 站点指纹的识别，然而，这些防御方法只关注流量层面，没有针对性地利用深度学习模型的有效信息，导致所需的时间开销和带宽开销较大，难以满足实际部署需求。不仅如此，大部分防御方法只适用于离线分析，难以用于在线 Web 站点指纹防御的实际部署。为了解决这些问题，本文提出了两种高效的 Web 站点指纹防御方法，并实现了在线 Web 站点指纹防御系统，具体工作如下：

(1) 设计了基于遗传算法的流量特征整形技术高效地搜索伪元数据包插入的位置和方向，以生成能欺骗 WF 分类器的变异流量。利用遗传算法搜索技术，对原始需保护的流量随机执行一系列的变异操作以生成变异流量。为了更高效地搜索插入模式，利用了 DF 这个基于深度学习的 WF 攻击模型提取流量中的高维度特征向量，并根据不同 Web 站点间的特征向量间的距离设计了适应度函数和变异方向控制机制以快速搜索需保护 Web 站点流量的变异方向。实验表明，在更为真实的开放世界场景中只引入 8.1%的带宽开销下能将 WF 攻击准确率降至 0.4%以下。

(2) 设计了基于模糊测试的流量特征整形防御技术来抵御当前最新的基于深度学习模型的 WF 攻击，并实现为每个 Web 站点找到一种通用的伪元数据包插入模式。该防御方法设计了提高神经元覆盖率和最大化模型的误分类行为这个联合优化目标函数，并利用梯度上升法指导插入伪元数据包以最大化目标函数。种子在每次变异后，输入到 WF 攻击模型中，若模型没有发生误分类行为，则继续对当前变异的种子进行变异直到达到该种子的最大迭代次数。大量实验表明，在封闭世界场景中，本防御方法分别能仅以 14.18%，11.07% 的带宽开销将 DF、Var-CNN 的攻击准确性分别降低至 8.80%、4.43%，比当前已有的防御方法均要高效。

(3) 基于以上两种防御技术，设计并实现了在线 Web 站点指纹防御系统。实现该防御系统的核心思想是在客户端和 Bridge 处插入伪元数据包。首先是搭建私有 Bridge 节点，客户端与私有 Bridge 节点已在离线时知晓访问当前站点的防御方案，借鉴 WFPadTools 攻击实现 Bridge 在相应位置插入下行数据包，客户端在相应位置插入上行数据包，最终达到保护用户隐私的目的。

综上所述，本文研究并实现了 Web 站点指纹防御技术，通过深度学习模型提取的深度指纹特征设计遗传算法的适应度函数，实现了基于遗传算法的流量特征整形防御技术。从更细粒度角度利用模糊测试技术提高深度学习模型的神经元覆盖率以增加 WF 攻击模型的误分类行为，实现了基于模糊测试的流量特征整形防御技术。最终设计并实现了 Web 站点指纹防御系统，保护了用户隐私。

专家推荐语

首先，肖桂同学在学术研究方面表现出色。她在研究生三年期间，专注于暗网中的流量指纹分析方向，深入探索暗网流量的特征识别与分析方法。她以学生第一作者身份在国际顶级会议 IEEE INFOCOM 上发表了 2 篇学术论文，INFOCOM 作为中国计算机学会推荐的 CCF A 类会议，其学术水平和影响力广受认可。这些成果充分体现了她在科研中的创新能力和严谨的学术态度。

其次，她具备优秀的科研素养和团队合作精神。在研究过程中，她善于独立思考，能够有效解决科研中的难题。同时，她乐于分享自己的研究成果和经验，积极参与学术讨论和团队协作，促进了团队整体科研水平的提升。在学习上她认真学习专业知识，理论基础扎实，实践能力突出，多次获得校级“优秀三好学生”等荣誉称号。此外，她为人谦逊诚恳，待人友善，乐于助人，深受师生们的喜爱和尊重。他积极参与各类校园和社会公益活动，具有高度的社会责任感和奉献精神。

总之，肖桂同学在学术研究、学习成绩、科研能力和个人品德等方面均表现卓越，具备成为优秀科研工作者的潜质和实力。且她的硕士论文具有较高的学术价值和应用前景，为此，我诚挚地推荐肖桂同学申报“江苏省计算机学会优秀硕士论文”奖项，恳请评审委员会予以充分考虑。

论文看点

近年来，随着用户对隐私保护问题的日益关注，世界各地成千上万的用户普遍采用匿名通信系统来保护他们的通信隐私。匿名通信是一种通过采用数据转发、内容加密、流量混淆等措施来隐藏通信内容及关系的隐私保护技术。Tor、I2P、FreeNet 是其中典型的匿名通信系统，由于 Tor 较强的匿名保护能力和良好的用户体验设计，它成为了最受欢迎的匿名通信系统。Tor 通过多层加密和多跳转发技术，实现了用户身份的匿名。然而，这种高度匿名性的网络也容易受到流量分析攻击使其失去匿名性，其中最具代表性的一类攻击是 Web 站点指纹（Website Fingerprinting，WF）攻击技术。

WF 攻击技术是指攻击者能采集和分析用户与服务器端交互的流量，并抽取流量特征以识别用户正在访问的 Web 站点，其中 Web 站点的流量特征就是 Web 站点指纹。在 WF 攻击的假设中，攻击者不具有破解加密算法的能力，只能收集流量中数据包的大小、时间、方向等特征形成 Web 站点指纹，并与已知 Web 站点访问流量的指纹进行对比，以获知用户访问的隐私信息。当前的 WF 攻击技术主要分为两大类，一类是基于机器学习模型的 WF 攻击，攻击者

需要依赖专家知识来抽取有效的流量特征以识别不同的 Web 站点；一类是基于深度学习模型的 WF 攻击，攻击者利用深度学习模型自动抽取高维流量特征以识别流量属于哪个 Web 站点。

针对这些 WF 攻击技术，研究者们探索出了一系列的防御方法以隐藏流量的特征，进而达到保护用户通信隐私的目的。其中主要的两种操作是：添加伪元数据包和延迟真实数据包。添加伪元数据包的操作破坏了流量中数据包数量、方向等特征，但带来了额外的带宽开销。延迟真实数据包的操作破坏了流量中数据包的时间等特征，但给网络增加了额外的时间开销，可能会影响用户访问 Web 站点的体验感。因此如何有效的平衡防御开销和防御效果是所有设计 WF 防御技术的研究人员的主要目标。当前的防御方法主要分为两大类：一类是基于流量特征抑制的 Web 站点指纹防御技术，它是指采用流量混淆技术抹平流量特征使得属于所有类别的流量特征均相似，进而导致分类器无法正确分类；一类是基于流量特征整形的 Web 站点指纹防御技术，它是指采用插入伪元数据包方式使源流量特征被整形为另一个类的流量特征。以 WTF-PAD 和 Walkie-Talkie 为典型代表基于流量特征抑制的 Web 站点指纹防御技术分别是通过自适应填充和基于超序列的防御方法，但均至少带来了 50% 以上的带宽开销，这给实际部署防御策略带来了一定困难。因此现在大部分研究者开始转向基于流量特征整形的 Web 站点指纹防御技术研究，该类主要借鉴了计算机视觉领域中加入微小干扰生成对抗样本欺骗深度学习模型的方法，通过制定策略指导随机插入伪元数据包以对当前流量特征整形，进而使得模型误分类。相对于流量特征抑制的防御方法，研究者们已提出的流量整形防御技术，以 Abusnaina 等人提出 DFD 为代表其带宽开销显著降低到 14.4% 时能达到 86% 的误分类率抵御以 DF 为代表的基于深度学习的 WF 攻击技术。

然而当前的防御方法仍然不够高效，且当前的流量整形防御技术大多只针对流量特征本身，未对攻击者采用的深度学习模型进行深入研究，未利用深度学习模型中间结果信息和神经元信息来指导算法高效探寻伪元数据包插入模式，以更小的带宽开销实现更好的防御效果，即达到更高的误分类率。因此本文利用深度学习模型的中间结果提出了两种流量整形防御技术：（1）基于遗传算法的流量整形防御技术，（2）基于模糊测试的流量整形防御技术。前者是利用深度学习模型抽取的高维特征表示当前流量以指导遗传算法高效插入伪元数据包；后者是借鉴模糊测试（fuzzing）思想以提高神经元覆盖率为目标来寻找插入模式。前者技术利用了深度学习模型中的特征图（feature map）来指导遗传算法快速高效探寻伪元数据包插入模式，实现低防御开销高防御效果；后者技术是从更细粒度角度进一步深入分析深度学习模型，探究插入伪元数据包后对模型中神经元值的影响，进而反向指导如何插入伪元数据包，以达到用更低带宽开销成功抵御 WF 攻击。最后设计并实现在线 WF 防御系统，在 Tor 客户端和私有网桥端实现 Tor cell 级别插入伪元数据包以验证所提出技术方案的可行性和有效性，实现对用户隐私的保护。

1. 基于遗传算法的流量特征整形防御技术

针对基于深度学习模型的匿名通信网络 Tor 的 WF 攻击技术，提出了一种基于遗传算法的伪装流量搜索技术，在原始流量中通过有效的插入少量伪元数据包，在不影响用户正常通信的前提下使得攻击者无法识别出用户正在访问的 Web 站点，达到保护用户通信隐私的目的。为了更高效的插入伪元数据包就能使得深度学习模型分类失败，此防御方法利用了当前主流的 Web 站点指纹攻击技术叫深度指纹攻击 DF 中设计的卷积神经网络来提取流量中的高维度特征向量，根据目标保护 Web 站点与其他 Web 站点的流量特征向量间的距离，设计了变异方向控制机制和适应度函数以指导目标保护 Web 站点流量变异的方向，最终针对每个 Web 站点都找到一条成功逃逸深度学习模型的哑元数据包插入模式，并在封闭世界和开放世界验证此防御方法的有效性和高效性。

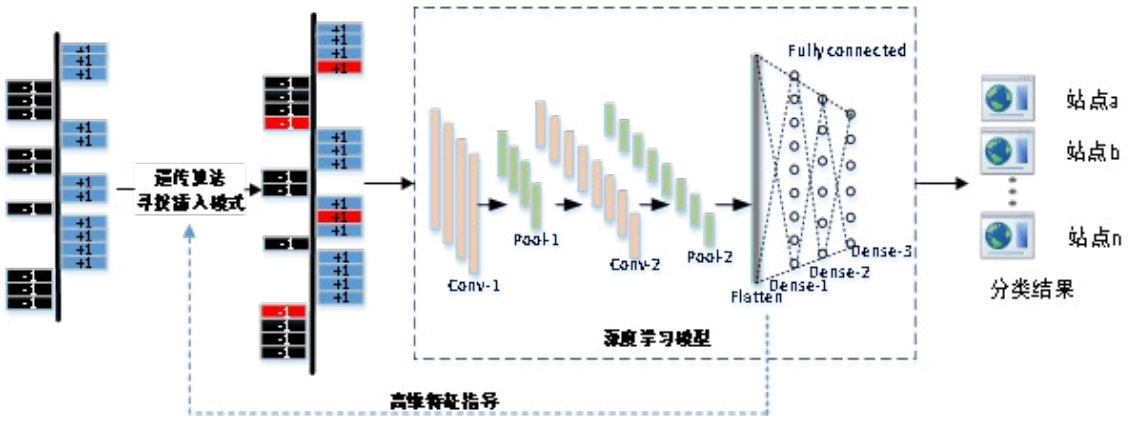


图 1-1 基于遗传算法的流量整形防御技术整体设计

针对以上的 WF 威胁模型，此防御方法的目标是在每个 Web 站点的流量中高效地搜索伪元数据包插入的位置和方向以生成能欺骗 WF 分类器的变异流量。位置是指在此条流量中何处插入伪元数据包，方向是指插入的伪元数据包是从客户端到出口节点的上行方向或从出口节点到客户端的下行方向。在 Web 站点指纹识别中，为了简化流量的表示，Wang 等人 [4] 将访问一个 Web 站点产生的流量被表示为 [+1,-1] 组合的序列。变异流量中所有伪元数据包的插入位置和方向构成了此条流量的伪元数据包插入模式。此防御方法的整体思路如图 1-1 所示，利用遗传算法搜索技术，对原始需保护的流量随机执行一系列的变异操作以生成变异流量。为了更高效地搜索插入模式，利用了深度学习模型 DF 提取流量中的高维度特征向量，并根据不同 Web 站点间的特征向量间的距离设计了适应度函数和变异方向控制机制以快速搜索需保护 Web 站点流量的变异方向。

本方法的工作流图如图 1-2 所示，首先挑选每个 Web 站点中能被 WF 分类器正确分类部分流量轨迹以初始化遗传算法种群。对初始化种群后得到的流量进行一系列的变异操作以选择在需保护流量的某个位置注入某方向（即 Tor 客户端到出口节点或相反方向）上伪元数据包。然后设计一个变异方向控制机制和适应度函数来指导整个搜索过程并确定是否找到成功逃逸的流量。对于成功逃逸的流量，记录此条流量的伪元数据包插入模式。当达到最大种群迭代次数时，停止所有搜索过程。根据上述防御方法的工作流程，本方案主要分为以下几个步骤：初始化遗传算法种群、设计变异操作、确定适应度函数、控制变异方向和选择变异流量。下面将分别讲述这几个步骤的详细过程。

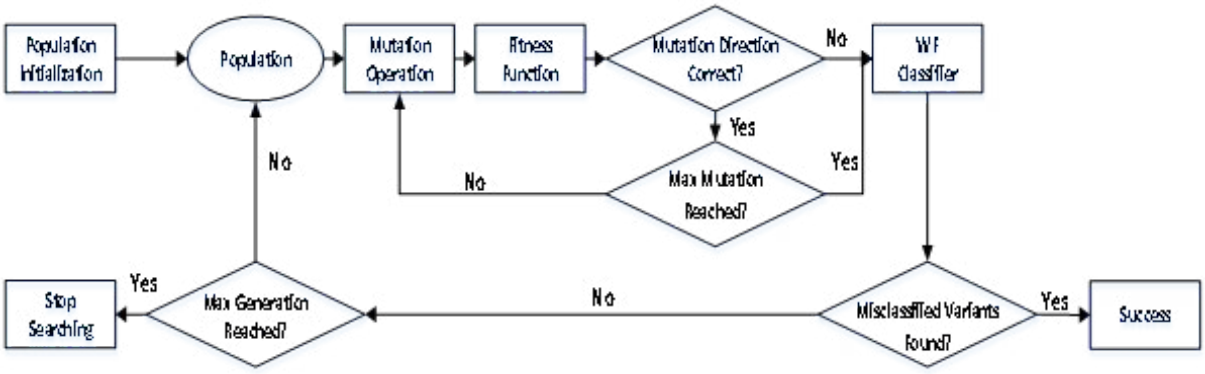


图 1-2 基于遗传算法的流量整形防御技术的工作流图

1.1 初始化遗传算法种群

针对数据集中的 N 个 Web 站点，为每个 Web 站点都精心挑选出 n 条被当前主流的深度学习模型正确识别的概率大于 90% 的流量，x_i^j 表示第 i ($i \in [1, N]$) 个 Web 站点的第 j ($j \in [1, n]$) 条流量。对这 n 条中的每一条流量均复制 m 次，因此初始化种群后一共 N*n*m 条流量，定义第 i 个 Web 站点的第 j 条流量的第 q 个复制样本为 $c_{i,j}^q (q = 1,2,\cdots,m)$。复制流量的目的是并发寻找流量样例的插入模式，高效快速探寻这 n 条流量的伪元数据包插入模型以得到成功逃逸的流量，当 m 条复制流量中的任意一条流量逃逸则认为当前流量已成功逃逸。

初始化种群后，为了对每个 Web 站点找到最佳变异方向，针对每个 Web 站点都选出与当前 Web 站点最不相似的前 k 个 Web 站点以构成该 Web 站点的种子池，让每个 Web 站点朝着自己种子池中的某个 Web 站点去变异。种子池的目的是为了让当前需逃逸流量朝着池中随机选中的目标 Web 站点流量特征变化。Web 站点间相似与否是通过计算不同 Web 站点间的欧几里得距离来衡量，Web 站点间越不相似，其距离越远。为了得到此距离，本防御方法利用了 DF 模型来提取流量中的高维度特征，通过 DF 的特征图（feature map）$\Phi(\cdot)$ 将输入的流量数据映射为相应的特征向量，因此得到每条流量 x_i^j 的特征向量 $\Phi(x_i^j)$。由于 DF 模型的攻击成功率能达到 98%，因此认为此模型能成功提取更具有代表性的高维度特征，所以选用 DF 模型作为特征提取器。本方法是采用平均特征向量 μ_i 代表第 i 个 Web 站点，因此需先计算每个 Web 站点的 n 每条流量的特征向量，再计算公式（2-1）出每个 Web 站点的平均特征向量。

$$\mu_i = \frac{1}{n}\sum_{j=1}^{j=n} \Phi\left(x_i^j\right) \#(2-1)$$

从种子池中随机选中某个 Web 站点 t，则根据平均特征向量计算当前第 i 个 Web 站点与第 t 个 Web 站点之间的距离，公式如下：

$$D(\mu_i, \mu_i(t)) = \ell_2(\mu_i, \mu_i(t)) \#(2-2)$$

$\mu_i(t)$ 表示第 i 个 Web 站点选中的 Web 站点的平均特征向量，$D(\mu_i, \mu_i(t))$ 表示 μ_i 与 $\mu_i(t)$ 之间的距离，ℓ_2 表示欧几里得距离。

1.2 设计变异操作

因为填充伪元数据包是当前针对各种 WF 攻击的有效防御技术，因此本文的变异操作也是在流量的不同位置填充 Tor 伪元数据包。对初始化种群后的流量设计了 3 种变异操作，第一种是随机添加上行数据包（+1），第二种是随机添加下行数据包（-1），第三种是随机添加上行 / 下行数据包（+1/-1）。针对需保护的目标流量，在选择了其中一种变异操作策略后，会在该流量中随机选择一个位置进行此变异操作进而产生新的变异流量。对于每条流量定义了最大的变异次数 $\mathcal{M}_I$，以控制防御方法的带宽开销。

1.3 确定适应度函数

适应度函数是用于设计变异方向控制策略以更快速的寻找到能欺骗基于深度学习模型的 WF 分类器的变异流量。本文利用 DF 攻击中神经网络来提取到流量的高维度特征向量，根据目标保护 Web 站点的某条流量的特征向量与其流量池中被选中的目标 Web 站点 t 的平均特征向量间的距离设计封闭世界和开放世界的适应度函数。定义 $x_i^j(o)$ 是第 i 个 Web 站点的第 j 条流量的第 $o(o \le \mathcal{M}_I)$ 次变异，则封闭世界的适应度函数为：

$$\mathcal{F}\left(x_i^j(o)\right) = D(\Phi\left(x_i^j(o)\right), \mu_i(t)) \#(2-3)$$

在开放世界，一共有 N+1 个类，N 个受监控 Web 站点类别，1 个不受监控 Web 站点类别，即所有的不受监控

Web 站点视为 1 类，第 N+1 类为不受监控 Web 站点类别。则在开放世界的适应度函数为：

$$\mathcal{F}'\left(x_i^j(o)\right) = D(\Phi\left(x_i^j(o)\right), \mu_i(t))\#(2-4)$$

1.4 控制变异方向

为了高效的插入伪元数据包找到成功欺骗攻击者分类器的变异流量，设计了变异方向控制机制。本文使用了滑动窗口$\mathcal{W}$和特征距离 D 来指导此次变异方向的正确与否，当发现此次搜索方向不正确时，可及时停止本次搜索。首先随机从当前需保护 Web 站点的种子池中随机选定一个 Web 站点 t，即确定了初始变异的距离 D。对需保护的 x_i^j 流量执行了$\mathcal{W}$ 次变异操作后，与第 t 个 Web 站点之间的特征距离应至少减少 $[W * D(\Phi(x_i^j), \mu_i(t))]/\mathcal{M}_I$ ，如图 1-3 所示。

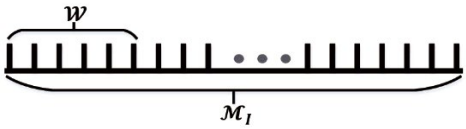


图 1-3 控制变异方向图

在变异次数逐渐增加且小于$\mathcal{M}_I$ 时，$x_i^j(o)$的与第 t 个 Web 站点间的特征距离应逐渐减小。基于此要求，在封闭世界下，变异方向控制机制设置为：

$$\mathcal{F}\left(x_i^j(h)\right) - \mathcal{F}\left(x_i^j(o)\right) \geq \frac{\mathcal{W}}{\mathcal{M}_I} D\left(\Phi(x_i^j), \mu_i(t)\right)$$
$$s.t.\ 0 < o - h < \mathcal{W}\#(2-5)$$

在开放世界，目标保护 Web 站点应要朝着不受监控 Web 站点类别变异，即第 N+1 类 Web 站点，因此其变异控制机制设置为：

$$\mathcal{F}'\left(x_i^j(h)\right) - \mathcal{F}'\left(x_i^j(o)\right) \geq \frac{\mathcal{W}}{\mathcal{M}_I} D(\Phi(x_i^j), \mu_{N+1})$$
$$s\ t\ 0 < o - h < \mathcal{W}\#(2-6)$$

只要满足公式（2-5）或公式（2-6）条件，说明此次搜索方向是正确的。否则停止搜索，并将此流量返回到下一次迭代的种群中等待再次变异。

1.5 选择变异流量

对要保护的 Web 站点流量执行变异操作后满足上述变异方向控制条件或者变异次数达到最大变异次数$\mathcal{M}_I$后，要挑选出成功变异的流量。此成功变异的流量不仅要能欺骗分类器，也同时也要满足以下条件：

$$r = \frac{D(\Phi\left(v_{i,j}^q\right), \mu_i(t))}{D(\Phi\left(c_{i,j}^q\right), \mu_i(t))}\#(2-7)$$

$v_{i,j}^q$ 表示复制样本流量 $c_{i,j}^q$ 变异后流量，r 表示变异后的流量与所选择的目标 Web 站点 t 之间的特征距离和变异前流量与第 t 个 Web 站点间特征距离的比值。当 r 小于某阈值 T 且能欺骗 WF 分类器时，则 $v_{i,j}^q$ 是成功逃逸的流量。当出现了成功逃逸的流量时时，会从原始的种群中删除该流量的其其余的复制流量，更新种群。

逃逸失败的流量需要放到种群中等待下一次的变异操作。为了加速遗传算法的探寻速度，建立了一个插入模式池。池中包括了成功变异的流量插入模式和潜在的能成功的插入模式。针对逃逸失败的流量，从池子随机挑选一个流量插入模式应用到当前流量中，插入完成之后新的变异流量重新放回种群中等待下一次迭代。当数据集中所有 Web 站点的 n 条流量均成功逃逸时停止整个探寻过程。

1.6 实验验证

在封闭世界和开放世界场景下进行了大量实验，验证本防御方法的有效性和高效性。封闭世界场景中，在获得最佳的滑动窗口大小$\mathcal{W}$和最大插入次数$\mathcal{M}_I$参数下，需计算针对 3 种 WF 攻击模型所需的带宽开销（Bandwidth Overhead, *BO*）。表 1-1 展示了这 3 种变异操作在 3 种攻击模型下的带宽开销的结果。表 2-2 中 IP 是指插入模式（Injection Patterns, *IP*），即 3 种不同的变异操作，*BO* 是指带宽开销。相比于另外两种变异操作，当插入 +1 时所需的带宽开销最小。然而在插入 -1 所需的带宽开销最大。因此同上述结论一样，插入 +1 的变异操作能获得最有效的插入模式。如表 2-2 结果所示，当插入 +1 时，DF、SDAE、和 CNN 模型的最佳检测率分别为 1.7%、1.4% 和 1.4%，而三个模型的带宽开销仅分别需要 12.0%、14.2% 和 16.8%。由于在表格 2-1 中使用未防御数据集时，DF 模型的检测率优于 CNN 和 SDAE 模型，因此表 2-2 中 DF 模型的检测率仍然高于 CNN 和 SDAE 模型。此外，当插入 +1/-1 时，本防御方法可以 5.4% 的检测率和 10.5% 的带宽开销欺骗 CNN 模型。尽管本方法使用 DF 模型的特征映射来计算特征向量，但仍可以高效的欺骗这 3 个模型，这意味着本方法具有强大的通用性。

表 1-1 封闭世界场景中不同插入模式下不同攻击模型的带宽开销和检测率

Models / IP	CNN				SDAE				DF			
	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR
Injecting 1	1000	500	16.7%	1.7%	1000	500	50.7%	2.7%	1000	500	10.1%	3.2%
	1200	500	16.4%	1.6%	1200	600	52.1%	1.5%	1200	600	12%	1.8%
	1400	500	16.8%	1.74%	1400	600	54.2%	1.4%	1400	600	12%	1.7%
Injecting -1	1000	500	19.1%	33.6%	1000	500	51.9%	37.6%	1000	500	15.8%	39.6%
	1200	600	21.8%	24.5%	1200	600	53.8%	34.6%	1200	600	17.9%	34.2%
	1400	600	21.7%	24.9%	1400	600	53.9%	32.1%	1400	600	18.1%	32.3%
Injecting 1/-1	1000	500	17.2%	9.6%	1000	500	54.1%	20.3%	1000	500	10.9%	26.7%
	1200	600	20.5%	5.4%	1200	600	52.7%	15.3%	1200	600	12.4%	19.7%
	1400	600	20.3%	3.2%	1400	600	52.4%	14.3%	1400	600	14%	19.3%

开放世界场景中，在获得最佳的滑动窗口大小和最大插入次数参数下，需计算在开放世界场景中针对 3 种 WF 攻击模型所需的带宽开销。表 1-2 描述了针对 3 种模型的 3 种变异操作的检测率和带宽开销情况。可观察到，插入 +1 时可以获得最佳的检测率和带宽开销，DF、SDAE 和 CNN 模型最佳检测率分别为 0.4%、0.1%和 0.1%，而带宽开销仅分别为 8.1%、10.2%和 10.2%。这表明，在 Tor 客户端到 Web 服务器的网页请求的流量模式特征上进行混淆可以有效地防御这 3 种基于深度学习模型的 WF 攻击。此外，当插入 +1/-1 时，CNN、SDAE 和 DF 模型的检测率可以显著降低至 0.5%、2.3%和 2.2%。因此，可得出结论：在开放世界场景中，这 3 个模型提取的高维特征使用了 Web 站点流量模式的上行数据包和下行数据包两个方向的特征。

表 1-2 开放世界场景中不同插入模式下不同攻击模型的带宽开销和检测率

Models / IP	CNN				SDAE				DF			
	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR	$\mathcal{M}_I$	$\mathcal{W}$	BO	DR
Injecting 1	1000	300	10.3%	01%	1000	400	8.9%	0.6%	1000	400	8.1%	0.7%
	1200	300	16.4%	01%	1200	400	8.4%	0.3%	1200	400	8.0%	0.6%
	1400	300	16.8%	0.2%	1400	500	10.2%	0.1%	1400	500	8.1%	0.4%
Injecting -1	1000	500	19.1%	15.1%	1000	500	13.1%	24.5%	1000	500	13.4%	29.8%
	1200	600	21.8%	9.5%	1200	600	13.4%	16.9%	1200	600	14.7%	28.1%
	1400	600	21.7%	8.7%	1400	600	13.2%	17.9%	1400	600	15.1%	26.2%
Injecting 1/-1	1000	500	16.7%	1.0%	1000	500	11.3%	4.1%	1000	500	10.2%	5.4%
	1200	600	16.9%	0.5%	1200	600	12.1%	3.5%	1200	600	12.4%	2.9%
	1400	600	16.1%	0.7%	1400	600	12.4%	2.3%	1400	600	12.2%	2.2%

大量实验证明本方法的可行性和有效性。在开放世界场景中，针对 DF 模型，检测率仅为 0.4%，带宽开销仅为 8.1%。此外，在封闭世界场景中，本防御方法对 DF 模型的检测率仅为 1.7%，带宽开销仅为 52.0%。本方法可用于对抗基于深度学习模型的流量分析攻击，以保护通信隐私。

2. 基于模糊测试的流量特征整形防御技术

为了从更细粒度角度分析深度学习模型，本文借鉴了模糊测试技术（fuzzing）的思想，提出了一种基于提高深度学习模型的神经元覆盖率指导插入伪元数据包的流量特征整形防御技术。从理论上分析，不同类别流量的特征不同，其神经元的激活情况不同。因此针对当前流量特征，若提高神经元覆盖率，则会引发深度学习模型误分类。此防御方法结合神经元覆盖率和模型分类出错两个目标设计了目标函数，为了让目标函数最大化，通过梯度上升法来指导插入伪元数据包到原始流量中。最终以底带宽开销为当前需保护的所有流量找到成功的流量对抗样本以欺骗深度学习模型，并在大量实验下验证此方法的有效性和高效性。

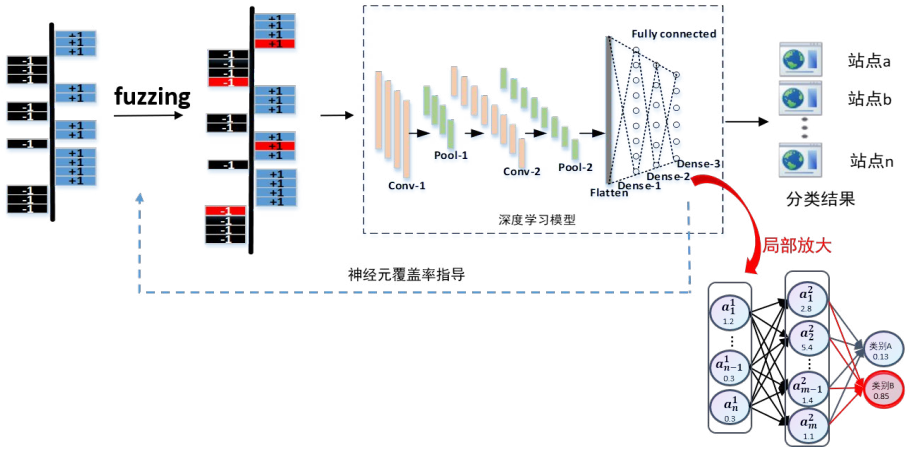


图 2-1 基于模糊测试的流量整形防御技术的基本思想

针对基于DL模型的攻击方法，Web站点指纹防御技术的核心是改变需保护的流量特征进而使得DL模型误分类，即通过改变需保护的流量特征生成新的测试样本以探寻DL模型的边界情况，而这正是模糊测试DL系统的核心目标。因此借鉴了模糊测试深度学习系统技术（fuzzing）的核心思想，从更细粒度角度分析深度学习模型，提出了一种基于模糊测试的流量整形防御技术（Fuzzing Defense，FuzzD）。此防御方法的核心思想是采用梯度上升法最大化DL模型的误分类行为和神经元覆盖率以指导插入伪元数据包，实现对每个站点找到一种高效的伪元数据包插入模式，如图2-1所示。由于DL模型的其决策逻辑通常表现在不同神经元的值和连接神经元之间的权重上，因此如果对于需保护的Web站点流量，通过改变流量特征以增大神经元覆盖率，有可能触及DL模型边界让模型产生错误决策，即模型分类出错。基于此原理，本防御方法将最大化神经元覆盖率和模型误分类行为这个两个目标进行联合优化，并设计了相应的优化目标函数。为了最大化目标函数，本文通过梯度上升法来指导插入伪元数据包到原始流量中，最终达到用低带宽开销为当前需保护的所有流量找到成功的流量对抗样本以欺骗深度学习模型的目的。

本防御方法的工作流图如图2-2所示。首先，需初始化种子队列，为每个站点挑选部分流量作为种子进行初始化，接着在从种子队列中迭代式选择种子进行后续变异。在种子变异阶段，先设计了联合优化目标函数，包括最大化DL模型的神经元覆盖率和模型的误分类行为这两部分。为了最大化目标函数值，本防御方法通过计算目标函数与输入流量之间的梯度指导伪元数据包在原输入流量中的插入位置和插入方向，得到变异后的流量样本。插入的位置需符合当前站点流量的基本特征。然后将变异后的流量输入到DL模型中计算神经元覆盖率和分类概率，判断是否能提高神经元覆盖率且使得模型误分类，若满足以上条件，则说明已找到原始流量的防御模式；否则，将重新对此流量进行变异直至达到该流量的最大迭代次数。根据上述防御方法的工作流程，本防御方法主要包括5个部分：初始化种子，选择种子，设计目标函数，制定变异策略，选择流量样本。下面将分别讲述这几个步骤的详细过程。FuzzD算法描述如算法1所示。

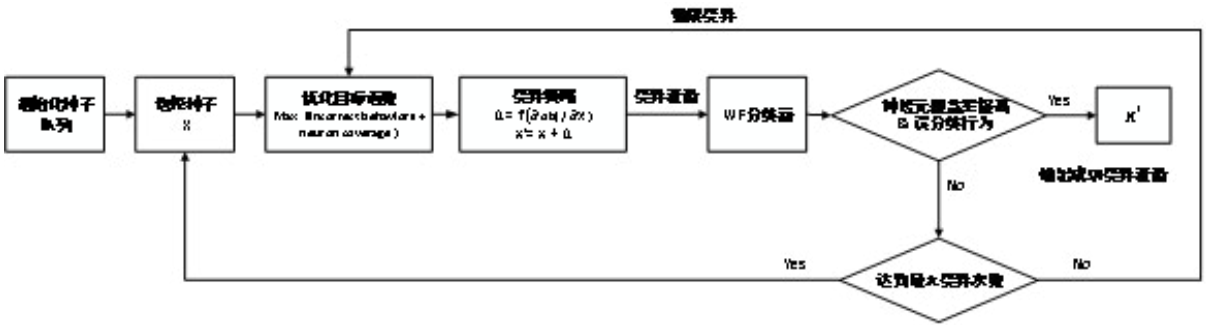


图2-2 基于模糊测试的流量整形防御技术的工作流图

2.1 初始化种子队列

本防御方法首先需要种子队列初始化，为每个Web站点选择相应的种子。针对测试数据集中的所有Web站点流量，首先为每个站点随机挑选n条均被攻击者的DL模型识别准确率超过95%的流量作为该站点的种子用于后续的种子变异。若按照本防御方法生成的插入模式（插入的位置和插入方向）插入到被DL模型识别准确率超过95%的种子产生的变异流量能让攻击者的DL模型失效，则更能说明本防御方法的有效性。

对于当前种子队列中的每个种子，需设置最大迭代次数以控制带宽开销。FuzzD防御方法设置最大迭代次数为

当前流量的 20% 的实际包长度，即控制当前带宽开销不超过 20%。带宽开销的定义见 3.3.2 节。

2.2 选择种子

模糊测试技术需要从种子队列中迭代式选择种子进行后续变异。从每次选择种子数量的角度看，当前种子选择策略分为单一选择和批量选择两种。前者是指每次从种子队列中选择一个种子作为候选种子，后者是指每次从种子队列中选择多个种子作为候选种子。本防御方法是采用单一选择策略，且每次选择种子队列中最新加入的种子作为下次变异对象，即在上次有效变异的基础上进行下一次变异。

2.3 设计目标函数

WF 防御方法的目标是能实现以最小的开销改变需保护站点的流量特征后，DL 模型无法准确识别出当前用户正在访问的 Web 站点。因此，本防御方法在结合模糊测试的基本思想后，设计了两个目标：提高 DL 模型的神经元覆盖率和最大化 DL 模型的误分类行为，即对输入流量能提高 DL 模型的神经元覆盖率的同时 DL 模型会误分类。因此设定了如下的目标函数：

$$obj = Max(\lambda_1 \, Neuron \, Coverage + \lambda_2 \, Incorrect \, Behaviors) \#(2-1)$$

其中 *Neuron Coverage* 指神经元覆盖率，*Incprrect Behaviors* 指模型的误分类行为。和是权重系数，衡量每个目标的重要性程度，在实验部分通过多次实验确定了这两个系数的最佳值。

本文参考了 Pei 等人 [46] 提出的神经元覆盖率 *NCov*，即测试用例中激活神经元数目占所有神经元的比例，第 *n* 个神经元被激活是指该神经元的值超过一个给定的阈值。通过改变输入流量输入到 DL 模型中以增大神经元覆盖率，可尽可能触及模型的边界，进而引发模型的误分类行为。

Incprrect Behaviors（*IB*）是指 DL 模型的误分类行为，定义为：

$$IB = c_1 - c_0 \#(2-2)$$

是指分类为原始正确类别（true label）的概率值，C_1 是指被模型分类为非原始类别标签中概率值最大的值，若 $C_1 > C_0$ 则认为当前攻击者的 DL 模型误分类。因此，增大该值可促进模型产生误分类行为。

因此基于模糊测试的流量特征整形技术的联合优化目标函数为：

$$obj = Max(\lambda_1 \, Neuron \, Coverage + \lambda_2 (c_1 - c_0)) \#(2-3)$$

2.4 制定变异策略

首先为了最大化神经元覆盖率 *NCov*，本防御方法提出了两种启发式策略以选择尽可能提高 *NCov* 的神经元，这两种策略分别是：

策略 0：选择过去被激活次数最多的神经元。受传统软件测试实践经验的启发，经常执行的代码片段更有可能引入软件缺陷。同理，经常覆盖的神经元可能会引发 DL 模型的边界行为，并激活更多的神经元。

策略 1：选择过去被激活次数最少的神经元，因为很少覆盖的神经元同样可能会引发 DL 模型的边界行为，并激活更多的神经元。

在采用不同策略选择好神经元的基础上，对于提高 DL 模型的神经元覆盖率和最大化 DL 模型的误分类行为这两个联合目标函数的优化，本防御方法采用梯度上升法实现最大化目标函数，并依据梯度信息确定当前变异过程中在输入流量中的待插入的位置和插入方向，以得到不同 Web 站点的插入模式。梯度上升法主要包括以下三个步骤：计算联合目标函数梯度、依据流量制约条件修改梯度和根据梯度添加伪元数据包。

（1）计算优化目标函数梯度：首先通过计算目标函数 *obj* 和输入样本 *x* 之间的梯度得到 *grads*，*grads* 是一个维度同输入样本的 *x* 向量，定义如下：

$$grads = \frac{\partial obj}{\partial x} \#(2-4)$$

（2）处理梯度 *f*(*grads*)：FuzzD 防御方法设计了两种变异操作，一是只插入上行伪元数据包（+1），二是可插入上行伪元数据包（+1）或者下行伪元数据包（-1）。因此对 *grads* 中的浮点数值有两种处理方式以确定伪元数据包的插入位置 insert_index 和插入值 insert_num。当只插入 +1 时，取当前计算的 *grads* 值最大的下标作为当前伪元数据包的插入位置；当可插入 +1/-1 时，则取当前 *grads* 绝对值最大的下标作为当前伪元数据包的插入位置，若插入位置处的 *grads* 为正数，则插入 +1，否则插入 -1。当只插入 +1 时，*f*(*grads*) 定义如下：

$$f_1(grads) = \{sign(grads_j), j | Max_{0 \le j \le n}(grads_j)\} \#(2-5)$$

当可插入 +1/-1 时，*f*(*grads*) 定义如下：

$$f_{+1/-1}(grads) = \{sign(grads_j), j | Max_{0 \le j \le n}(abs(grads_j))\} \#(2-6)$$

n 表示梯度维度，即输入数据的维度，*sign*(*grads*$_j$) 表示取梯度中的第 *j* 项的符号，*abs*() 表示求绝对值，*f*(*grads*) 的返回值为插入位置和插入值。

（3）添加伪元数据包：处理好梯度后，依据梯度向该位置插入伪元数据包后得到变异流量样本 *x*'。由于插入伪元数据包后仍必须符合 Web 站点流量基本特征，因此在通过 *grads* 获取的插入位置和插入值后，必须首先判断插入位置是否在当前流量的实际数据包个数 tra_len 之前，即不能将伪元数据包插入至流量中填充 0 的数据包部分。而且，在每执行一次插入伪元数据包操作后，需要对插入位置后的所有数据包位置整体右移，符合实际部署防御方案时插入伪元数据包后的情况。

$$x' = insert(x, f(grads)) \#(2-7)$$

2.5 挑选流量样本

经过插包操作后，将得到的变异流量样本 *x*' 输入到深度学习模型 DNN 中预测其分类，并可得到 DL 模型中激活神经元的值。若当前的变异流量样本 *x*' 能能使神经元覆盖率增加，且能成功欺骗 DNN 时，则保留并输出此条变异流量样本 *x*'。若不满足以上条件，且没达到此条流量的最大迭代次数时，则将改变异的流量加入种子队列中，后续在此条已变异的流量基础上继续 fuzzing，即选择最新变异的种子继续变异过程。

算法 1 FuzzD 算法

算法：FuzzD 算法
输入：ori_tra ← 测试集中的原始流量；
strategies← 神经元选择策略（策略 1 / 策略 2）；
dnn DL← 模型（DF / Var-CNN）；
BO ← 最大的带宽开销；
max_insert_times← BO * 每条流量的实际数据包个数 tra_len；
输出 : 伪元数据包插入位置 insert_index;
伪元数据包插入值 insert_num;
成功变异流量 gen_tra;
神经元覆盖率 neuron_coverage

算法：FuzzD 算法

```
gen_tra = []
tra_len[ ] = Count_trace(ori_tra)
max_insert_times[] = BO * tra_len[ ]
for x in ori_tra:
    seed_list = [x]
    while insert_time < max_insert_times[seed_list[0]]:
        seed_list.remove(seed_list[0])
        c1, c0 = dnn.predict(x)
        neurons = selection(dnn, strategies)
        obj = λ1(c1 − c0) + λ2 sum (neurons)
        grads = ∂obj / ∂x
        insert_num, insert_index = f (grads)
        x' = insert(x, insert_num, insert_index
        c1', c0' = dnn.predict(x')
        if (c1' > c0'& neuron_coverage improved
            gen_tra.append(x')
        else:
            seed_list.append(x')
```

2.6 实验验证

在封闭世界和开放世界场景下进行了大量实验，验证本防御方法的有效性和高效性。在封闭世界场景中，验证了本防御方法的泛化性。图 2-3 展示了 FuzzD 防御方法具有泛化性。防御方法的泛化性是指当针对一个已知晓模型的结构、神经元值等参数的情况下，利用针对该攻击模型产生的插入模式插入到原始流量中得到变异流量，将该变异流量输入到另一个未知参数的模型后仍能降低其检测率。为了验证 FuzzD 防御方法的泛化性，针对已知参数的 DF 模型，当所有参数均为最佳情况下，即神经元选择策略为策略 0，神经元数量为 45 时进行防御得到的插入模式在未知参数的 Var-CNN 模型中进行验证，图 3-8 结果表明，不论是插入 +1 操作还是插入 +1/-1 操作，FuzzD 均能有效降低 Var-CNN 模型的 DR，尤其是在插入 +1/-1 操作中，针对 DF 的插入模式在 Var-CNN 模型中更有效。

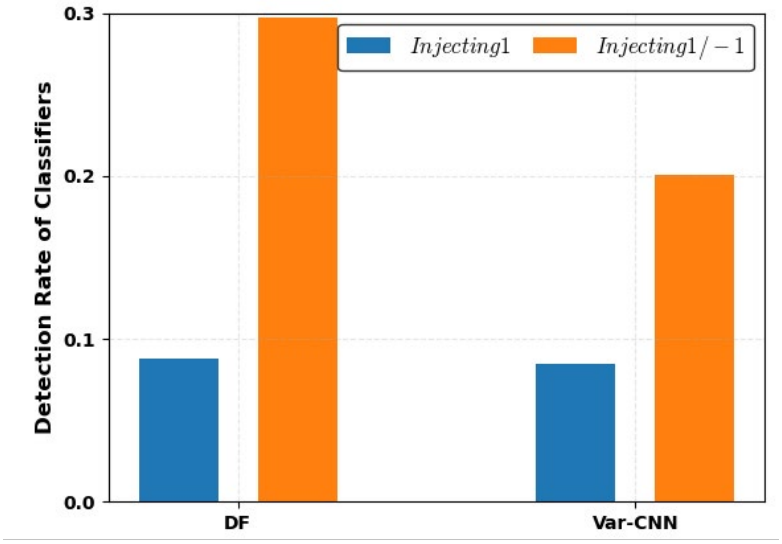


图 2-3 封闭世界中 FuzzD 方法的泛化性

表 2-1 展示了在封闭世界场景中，FuzzD 防御方法与其他防御方法的防御效果对比。表中结果表明，FuzzD-light 能以 14.18% 的带宽开销将 DF 和 Var-CNN 的 DR 降低至 8.8% 以下，比 BAND 的带宽开销低了 14% 左右，但防御效果与 BAND 相差无几；FuzzD-heavy 与 BAND 相比能以低 4% 以上的带宽开销实现更好的防御效果。与 Mockibird、WTF-PAD 相比，FuzzD-light 和 FuzzD-heavy 更是以低了 10%-40% 的带宽开销实现了低了 30%-82% 左右的检测率，均优于现有的防御方法。

表 2-1 封闭世界中 FuzzD 与其他防御方法的对比

Models \ Methods		FuzzD-light	FuzzD-heavy	BAND [53]	Mockingbird [18]	WTF-PAD [30]
DF	BO	14.18%	21.43%	25.02%	58.02%	63.23%
	DR	8.80%	5.62%	5.12%	38.11%	90.85%
Var-CNN	BO	11.04%	15.32%	25.07%	58.12%	63.12%
	DR	4.43%	2.15%	1.51%	35.21%	94.02%

同封闭世界一样 也许验证本防御方法的泛化性 图 2-4 展示了 FuzzD 防御方法在开放世界场景中具有泛化性 同封闭世界场景一样 为了验证 FuzzD 防御方法的泛化性 针对已知参数的 DF 模型 当所有参数均为最佳情况下 即神经元选择策略为策略 0 神经元数量为 45 时进行防御得到的插入模式在未知参数的 Var-CNN 模型中进行验证 不论是插入 +1 操作还是插入 +1/-1 操作 FuzzD 均能有效降低 Var-CNN 模型的 DR 尤其是在插入 +1/-1 操作中 DF 的插入模式在 Var-CNN 模型中的 TPR 降低至 10.73% 结果表明 在开放世界或封闭世界场景中 FuzzD 防御方法均具有泛化性 能让未知参数的模型产生误分类行为 。

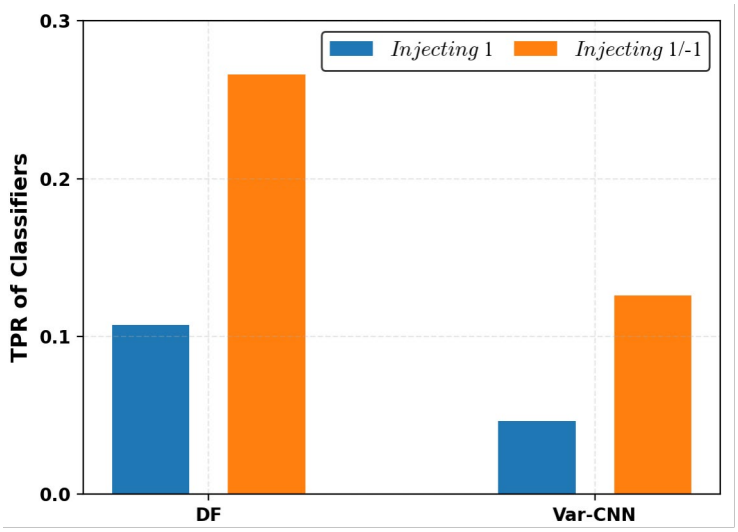


图 2-4 开放世界中 FuzzD 方法的泛化性

表 2-2 展示了在开放世界场景中，FuzzD 防御方法与现有防御方法的防御效果对比。表中结果表明，FuzzD-light 能以 14.18% 的带宽开销将 DF 和 Var-CNN 的 DR 降低至 10.73% 以下，比 Surakav-light 的带宽开销和 TPR 分别低了 40% 和 30%；FuzzD-heavy 与 Surakav-heavy 相比能以低 55% 左右的带宽开销实现更好的防御效果。与 WTF-PAD 相比，FuzzD 防御方法更是以低了 13% 的带宽开销实现了低了 81% 左右的 TPR，均优于现有的防御方法。

表 2-2 开放世界中 FuzzD 与其他防御方法的对比

Models \ Methods		FuzzD-light	FuzzD-heavy	Surakav-light [50]	Surakav-heavy [50]	WTF-PAD [30]
DF	BO	15.52%	25.85%	55.11%	81.02%	28.33%
	TPR	10.73%	6.72%	39.40%	8.14%	89.75%
Var-CNN	BO	11.22%	24.85%	55.21%	80.93%	27.02%
	TPR	6.00%	3.47%	39.70%	6.31%	88.80%

大量实验证明，在封闭世界场景中，FuzzD-light 分别能以 14.18%，11.07% 的带宽开销将 DF、Var-CNN 的攻击准确性分别降低至 8.80%、4.43%，比当前的 WTF-PAD、BAND、Mockingbird 方法均要高效。在开放世界场景中，FuzzD-light 能实现以 15.52% 的带宽开销将 DF 模型的 TPR 降低至 10.73%，而 Surakav-light 方法不仅需要 55% 带宽开销且仅降低 TPR 为 39.40%。

3.Web 站点指纹防御系统

本文利用 WFPadTools 协议设计并实现了一套在线 Web 站点指纹防御系统，实现了对流量序列的在线防御功能。该系统包括 3 个模块：离线伪元数据包插入模式生成、在线伪元数据包插入、可视化结果展示。在离线情况下生成每个 Web 站点的伪元数据包插入模式，防御者需提前在客户端和网桥端部署 Tor 插件，借助 WFPadTools 协议实现在客户端和网桥段进行在线伪元数据包填充操作，验证了本文提出的两种防御方法的可行性。

3.1 系统设计与工作流程

本 Web 站点指纹防御系统包括 3 个模块：离线伪元数据包插入模式生成、在线伪元数据包插入、可视化结果展示，系统工作流程如图 3-1 所示。

图 3-1 在线 Web 站点指纹防御系统工作流程

系统首先需在离线情况下生成每个 Web 站点的伪元数据包插入模式，插入模式包括了伪元数据包的插入位置和插入值。在进行在线伪元数据包插入前，防御者需提前在客户端和网桥端部署 Tor 插件，并将这些插入模式分别存储至客户端和网桥端。当用户开始访问 Web 站点时，防御者先选择相应 Web 站点的伪元数据包插入模式，并实时记录当前访问站点的流量情况，根据伪元数据包的插入模式，借助 WFPadTools 协议实现在客户端和网桥段进行在线伪元数据包填充操作。站点访问结束后，将加入防御后的流量输入至 WF 攻击模型中进行判别，最后将防御结果进行可视化。下面将对每个模块的设计与具体实现进行详细阐述。

3.2 模块设计与实现

离线伪元数据包插入模式生成

在离线情况下，可根据本文第二章第三章提出的基于遗传算法的流量特征整形防御技术和基于模糊测试的流量特征整形防御技术相应算法对未防御的数据集进行变异操作，以得到受监控站点数据集中每个 Web 站点的相应的伪元数据包插入模式。具体实现过程见第二章与第三章的描述。

在线伪元数据包插入

本文是在客户端和私有网桥节点处借助 WFPadTools 协议实现在线伪元数据包的插入。假设在线插入伪元数据包时，客户端和网桥端均已知晓当前需保护站点的伪元数据包插入模式。该插入模式首先需要依据原始流量处理成在原始流量的上行 Tor cell 的相对位置（C1，C2，C3，C4，C5，…）和下行 Tor cell 的相对位置（B1，B2，B3，B4，B5，…），如图 3-2 所示。根据本文第二章和第三章的两种防御方法，共设计了三种变异操作，分别是只插入上行伪元数据包、插入上行 / 下行伪元数据包、只插入下行伪元数据包。因此，客户端和网桥端均需记录实时流量，未到达伪元数据包填充点时，两端均正常发送或转发报文，不延迟或填充报文。

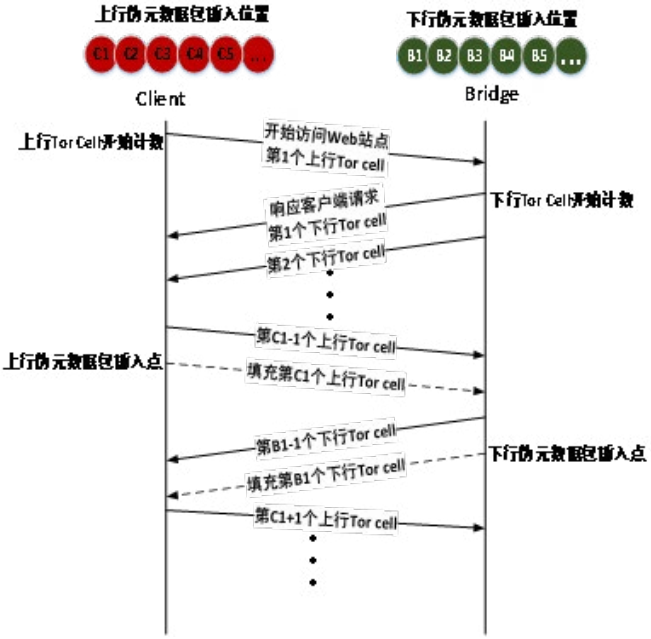


图 3-2 伪元数据包插入过程

当只插入上行伪元数据包时，只需要客户端对当前流量中上行 Tor cell 个数进行计数，到达上行伪元数据包的插入点时客户端填充上行伪元数据包即可。当插入上行 / 下行伪元数据包时，需要客户端对上行 Tor cell 计数，网桥端对下行 Tor cell 计数，到达各自的插入点时，两端分别填充伪元数据包，填充过程如图 3-2 所示。

基于以上插入过程，在线部署防御策略时需要对客户端和网桥端分别进行配置，修改两端的 torrc 文件使得 Tor 能使用 WFPadTools 协议进行防御。

3.3 防御结果展示

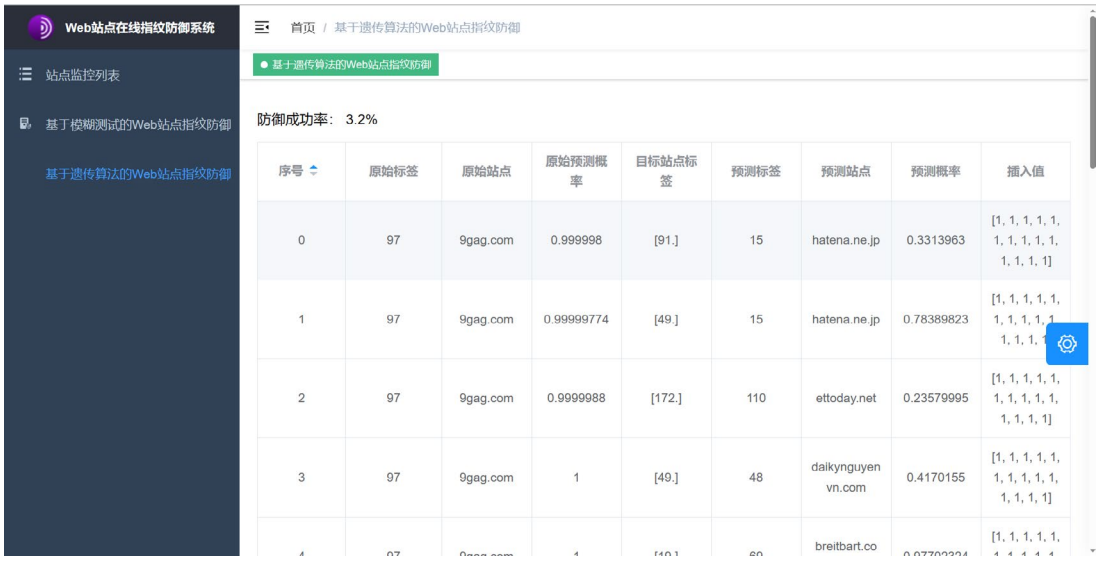


图 3-3 基于遗传算法的流量特征整形防御结果展示图

图 3-3 展示了基于遗传算法流量特征整形防御结果，包括了该流量的原标签、原始站点名、原始预测概率值、目标站点标签、防御后的预测标签、防御后的预测概率值、插入值的信息。

图 3-4 基于模糊测试的流量特征整形防御结果展示图

图 3-4 展示了基于模糊测试的流量特征整形防御技术的结果，包括了该流量的原标签、原始预测概率值、防御后的预测标签、防御后的预测概率值、带宽开销、伪元数据包的插入位置、插入值、当前模型的神经元个数等信息。

作者简介

肖桂，东南大学网络空间安全学院 2023 届硕士毕业生，所在科研团队为江苏省网络与信息安全团队（JSSEC）。团队共专职科研人员 11 人，其中教授 3 人、副教授 8 人。国家优青、国家优青（海外）各一人。团队指导在读博士研究生约 30 人，在读硕士研究生约 100 人。团队围绕“面向国家重大战略需求，开发具有自主知识产权和应用前景的技术成果和产品，开展原创性、系统性的科学技术研发工作，建设国内外知名的实验室与研究团队”的建设目标，团队从网络空间安全监测、无线移动网络安全、物联网三个研究方向展开研究。实验室长期从事匿名通信和流量分析、暗网探测和内容分析等领域的研究工作，承担完成了国家自然科学基金等多个科研项目，在 TPDS、TIFS、ComMag、CCS 等国际顶级期刊和会议上发表论文 50 余篇，并成功开发了一系列综合性工具，可以自动地从公共网络与暗网中采集暗网域名。

指导老师

凌振：国家优秀青年基金获得者，东南大学青年首席教授。于 2005 年和 2014 年分别获得南京工程学院计算机科学与技术学士学位和东南大学计算机应用技术博士学位。研究方向为网络安全、匿名网络、智能终端安全及物联网安全。近年来发表学术论文 80 余篇，包括重要国际期刊，如 IEEE/ACM ToN、IEEE TIFS、IEEE TPDS、IEEE TC、IEEE TDSC、IEEE JSAC、IEEE TIP 以及重要国际会议，如国际计算机安全会议 ACM CCS、USENIX Security、NDSS 和国际计算机网络会议 INFOCOM。先后主持国家自然科学基金青年、面上和优秀青年基金项目，国家重点研发计划课题，江苏省自然科学基金青年基金项目和优秀青年基金项目等。此外多次受邀在重要国内外学术会议作特邀报告，如国际著名黑客会议 BlackHat 上演讲 2 次等，相关研究成果还受到中央电视台科教频道 CCTV10《走近科学》栏目采访在第二届国家网络安全宣传周系列节目中进行报道。获得 2014 年 ACM 中国优秀博士论文奖、2015 年 CCF 优秀博士论文奖、2016 及 2020 年黑客极棒竞赛（GeekPwn）优胜奖、公安技术革新成果一等奖、国家教学成果二等奖、江苏省教学成果一等奖等奖项。

平台支撑情况

东南大学大数据计算中心，面向全校提供高性能计算与云计算服务，其硬件设施和计算能力包括 CPU 理论峰值浮点计算能力达到每秒 366.5 万亿次（共 9776 核），GPU 峰值计算能力为每秒 1107.4 万亿次（142 块 NVIDIA V100，8 块 NVIDIA K40），存储系统配置高达 5PB 容量的 SSD、SAS 和 SATA 磁盘混合结构模式，整个系统采用 10 台 36 口 40Gb 高性能 InfiniBand 交换机构建连接所有节点的全线速、无阻塞高速网络。除此之外，中心通过部署相应的云计算软件统一管理软硬件资源，以虚拟化和自动化的方式动态部署资源，用来统一提供计算和存储服务；所有这些条件为团队的研究提供了理想的实验与测试环境，可满足研究和开发的需要。

教学科研双驱的新时代 IT 人才培养模式探索与实践

——2024 年江苏省计算机学会教学成果奖刘凡教授

个人简介

刘凡，博士，河海大学教授，博士生导师，现任河海大学计算机与软件学院副院长，兼任江苏省计算机学会常务理事、江苏省人工智能学会常务理事、SCI 期刊 KSII TIIS 编委、Frontiers in Computer Sciences 客座编辑。入选江苏省“青蓝工程”优秀青年骨干教师、江苏省科协青年科技人才托举工程，先后主持国家自然科学基金面上与青年项目、装备预研教育部联合基金、江苏省自然科学基金面上项目等省部级及以上科研项目 20 余项；在 IEEE TNNLS、AAAI、ACM MM、IJCAI 等期刊或会议发表学术论文 100 余篇，ESI 高被引论文 5 篇，热点论文 1 篇。论文在 Google Scholar 中引用次数近 6000 次，H- 因子为 19，I10- 因子为 35，获首届江苏省自然科学百篇优秀学术论文；授权发明专利 30 项（第一发明人 22 项），获江苏省高等学校科学技术研究成果奖二等奖、江苏省自动化学会青年科技奖、江苏省信息技术应用学会科技奖一等奖、IEEE ICME 2021 最佳演示奖、IJCAI 2021 LTDL 最佳数据集论文奖。在教学方面，出版规划教材 2 部，主持教改项目 6 项，指导学生获江苏省优秀本科毕业论文一等奖，获国家级奖项 16 项，获得包括全国高校人工智能教师教学创意竞赛二等奖、江苏省高校微课教学比赛一等奖、江苏省高校教师教学创新大赛二等奖、中国计算机实践教育学术会议优秀论文一等奖等教学荣誉 30 余项。

图 1 刘凡教授

坚守独立探索精神，推动多模态人工智能在水利行业的落地应用

刘凡教授长期从事多模态大模型、跨模态视觉理解与分析等多个领域的研究，形成了如图 2 所示的“理论创新 - 算法研究 - 技术应用”的研究框架。理论创新层面：首创了多阶统计集成理论，证明了卷积张量特征的不同阶

的统计量之间具有互补性；提出泛化误差边界定理，揭示小样本学习模型误差由基类经验误差与假设空间 VC 维共同决定的规律；构建多模态大模型小样本迁移理论框架，指导域差异、模型容量与样本量的优化配置，有效提高模型的泛化性能并减少经验误差。算法研究层面，发表卷积神经网络权威综述（被引 4000 余次，入选 ESI 热点论文、首届江苏省自然科学百篇优秀学术论文），提出基于原型聚类的 ProtoCLIP 预训练方法，将多模态模型训练效率提升 4 倍；构建并公开了面向多模态图文特征对齐方法实际部署应用的 ITRA Codebase 开源框架；开源并发布了全球首个遥感视觉 - 语言基础模型 RemoteCLIP，数据集规模达同期数据的 10 倍以上，GitHub 代码仓库已获 352 个 Star，在给李国英部长《关于 AI for Science 在水利行业的探索实践与政策建议》中也介绍了 RemoteCLIP，西工大、航天宏图、中国空天院等相关单位知名学者都引用并肯定了此工作。技术应用层面，提出了基于心理学定律的特征表示方法 WLBP，并被成功应用于人脸识别、车辆主动安全等多个领域；研发了一套单兵视频动作捕捉系统以用于部队室内模拟训练，有效提升了部队的训练效率并获央视报道；开发了基于 RGB-D 相机及深度学习的疲劳驾驶检测技术，针对渣土车、泥浆车乱排乱放以及安全事故频发等问题，提出了基于物联网的运输车辆智能监控与主动安全一体化解决方案；设计新型无人机光电载荷支持多模态信息采集，载荷质量仅 368g，功耗仅 8.8w。成果已成功应用于北京、江苏、青海等省区以及水利部本级、黄河流域等水利管理部门或机构，在水资源管理、水旱灾害防御、水行政执法、水工程安全建设与管理等重要水利业务中发挥关键作用。

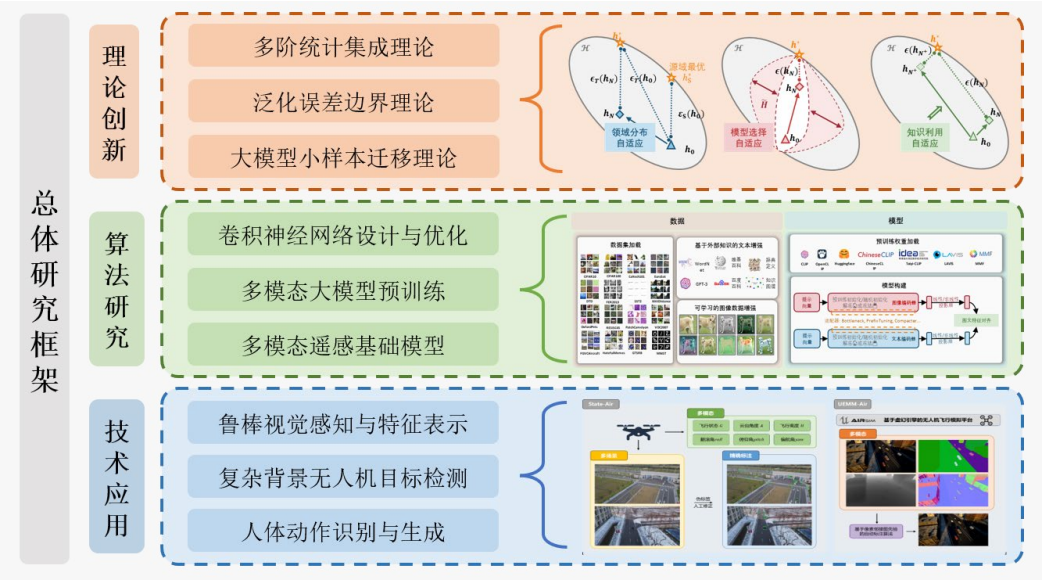


图 2 “理论创新 - 算法研究 - 技术应用”研究框架

落实立德树人任务，探索“三驱动五链条”的创新人才培养模式

面对科技革命和 IT 产业变革带来的新变化新趋势，刘凡教授以全面提高 IT 拔尖人才培养质量为重点，以解决关键核心技术“卡脖子”问题为导向，以强化课程思政与学科竞赛相互融通为主线开展 IT 拔尖人才的培养体系研究与设计，构建了 IT 人才“铸强培优”体系，并提出了围绕“技术发展、学生志趣、内外资源”的“三驱动”要素、落实“五链条”举措的创新模式，对 IT 人才“铸强培优”体系进行了实践。针对课程教学模式对于学生自驱力不足，课程知识体系对于工科需求的前沿性、工程性欠缺，课程思政元素对于人才培养的挖掘不充分等问题，刘凡教授积

极开展教学改革，设计了 CDIO 工程理念下多元化的教学方法和培养手段，引入了学科竞赛和工程案例等内外资源，加强了对学生的专业创新能力、综合素质和学术道德等方面质量文化建设力度。通过培优架构设计、教学内容重构、教学过程优化、智教水平提升、行业思政铸魂等举措，全面提升课程先进性、学生学习主动性，强化学生专业认知、科研创新能力，塑造学生正确三观，努力培养担当民族复兴大任的时代拔尖人才，具体如下：

（1）学科竞赛与科学研究双驱。指导学生获国家级学科竞赛奖项一等奖 5 项、二等奖 20 项、江苏省本科优秀毕业论文一等奖、二等奖；指导本科生第一作者在一区 TOP 期刊 IEEE TNNLS 上发表论文，并入选 ESI 热点论文、首届江苏省自然科学百篇优秀学术论文成果，已被引 4000 余次；指导本科生分别获国际会议最佳 Demo 论文、最佳 Dataset 论文、最佳 Presentation 论文。（2）教学研究与团队建设并举。获全国高校人工智能教师教学创意竞赛二等奖、江苏省高校教师教学创新大赛二等奖等省部级及以上荣誉 8 项、主持教改课题 8 项、产学研合作项目 5 项，编写规划教材 4 部，自编教材被南理工等 10 余所高校采用，获得一致好评。（3）专业教育与思政育人融合。指导的本科生曾获“江苏省优秀共青团员”和“2019 江苏省大学生年度人物”提名奖 1 人，江苏省“先进班集体”3 个、宝钢优秀学生奖 1 人、“江苏省三好学生”及“优秀毕业生”共 11 人，部分优秀毕业生的求职经历被《光明日报》和新华社等媒体报道。（4）智教工具与授课经验共进。自研的“智教”辅助教学小程序实现了教师课堂教学的教学分析与评价，可精准地完成试题推荐和自动组卷功能，并准确分析和追踪学生知识点的掌握情况，量化处理学生的学习结果。已获授权发明专利等知识产权证书 3 项，并在南大、东大等 10 多所高校推广应用。

图 3 “铸强培优”新时代 IT 人才培养体系

在“软件方法学”科研与育人领域奋楫笃行

——2024 年江苏省计算机学会优秀科技工作者宋巍教授

个人简介

宋巍，博士，南京理工大学计算机学院 / 人工智能学院 / 软件学院教授、博士生导师、南京理工大学第十届学术委员会委员，现为中国计算机学会杰出会员、传播工委委员、软件工程专委执行委员、服务计算专委执行委员，江苏省计算机会员，IEEE Senior Member，先后受邀担任 ISSTA、ICWS 等顶会的 PC (Senior) Member。2010 年毕业于南京大学软件新技术国家重点实验室，获计算机软件与理论博士学位，随后进入南京理工大学工作至今；曾先后访问香港科技大学、慕尼黑工业大学，曾历任 CCF YOCSEF 南京 AC 委员、秘书、副主席。十余年来一直从事软件工程与方法学、服务计算等方面的研究工作，先后主持或参与国家自然科学基金、国家重点研发计划、江苏省自然科学基金、江苏省 973 等科研项目。2016 年受邀参加国际计算机顶级论坛 Schloss Dagstuhl 并做报告。近年来，尤为关注过程软件分析及其应用研究，在 OOPSLA、ICSE、FSE、ASE、ISSTA、ICWS、ICSOC 等顶会以及 TSE、TSC、TKDE、TPDS、TDSC、TOSEM、中国科学 - 信息科学、软件学报等国内外顶刊上发表 CCF A 类论文 30 余篇。指导学生先后获得 2020 年第 22 届中国机器人与人工智能大赛“室外无人车智能挑战赛”冠军、2022 年江苏省优秀硕士论文、获顶会 ESEC/FSE 2023 杰出论文奖、2024 年“中国软件杯”大赛一等奖。部分科研成果已应用于字节跳动抖音产品线。

宋巍教授

聚焦软件方法学，致力于软件工程领域创新研究

宋巍教授长期从事“软件方法学”这一研究方向。在“软件定义一切”与“一切皆服务”的背景下，软件的运行环境逐步从封闭、静态、可控走向开放、动态、多变，开放环境为过程驱动软件的构建、运行、演化带来了新的挑战。

针对这些挑战，围绕“开放环境下的过程软件分析、挖掘及应用”这一主题，提出一种基于活动关系的过程软件分析方法，其特点在于以活动间的本质关系 (依赖、独立、互斥) 为单元来刻画过程行为和从事件日志中反映出的过程行为，从而区分过程的不变行为和变化行为。基于此分析方法，对基于过程软件全生命周期内所涉及的若干关键科学问题展开研究（图 1），取得了系统性的创新研究成果，并掌握了一套自主可控的关键技术，为开放环境下过程软件中间件的研制提供了坚实的技术储备。具体创新成果可归纳为三方面：1）基于活动关系的过程静动态分析：提出基于过程依赖图的过程重构方法，可使重构后的过程拥有最大的并发性和逼近最大程度的结构性；提出一种柔性过程动态演化方法，不但可保证过程实例迁移的合法性，而且显著提高了实例可迁移的比例。2）事件日志不完备情形下的过程挖掘：提出基于活动依赖关系的过程挖掘方法，能够从不满足局部完备性的事件日志中挖出正确过程模型；提出基于事件间传递先于关系的 DAG 科学工作流挖掘方法，当公开事件日志满足较低完备性要求时，所提方法即可保证挖出与正确模型路径等价的 DAG 模型。3）基于过程分析的软件缺陷检测：将过程分析理论应用于检测开放环境下软件中的缺陷，在真实移动应用和区块链智能合约中发现了众多软件缺陷和漏洞。

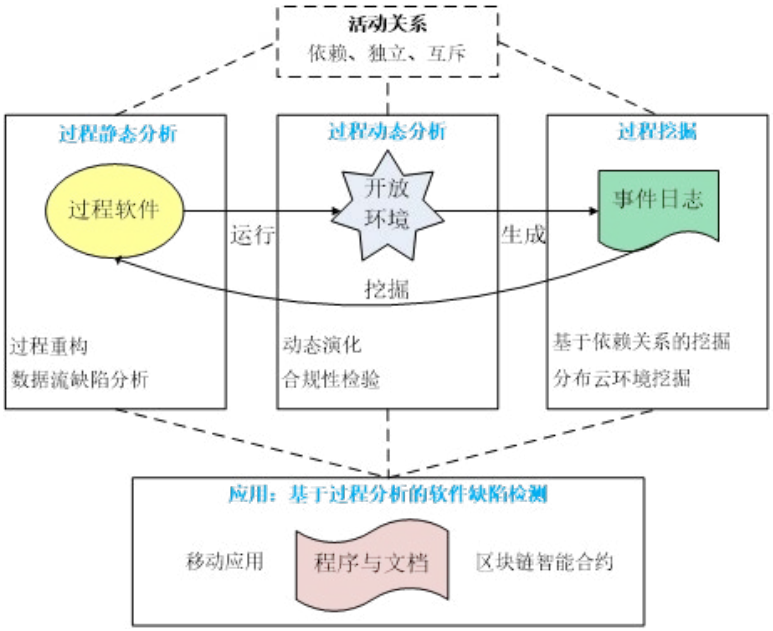


图 1“开放环境下的过程软件分析、挖掘及应用”研究框架

贯彻理论联系实际，服务于软件行业生态建设

宋巍教授注重研究成果的落地应用。在夯实理论研究的基础上，近些年来积极投身实践探索，从纷繁复杂的真实软件系统中精准锚定研究方向，聚焦于解决软件行业领域的实际问题。在移动应用与小程序领域，与“字节跳动”在 App 隐私合规方面开展合作，帮助自动化检验抖音等多个产品的隐私政策文件，帮助该企业有效规避了隐私政策不合规等重大风险。此外，开发了基于静态分析和测试的 App 分析工具集，可有效发现 App 中存在数据丢失、隐私泄露、JavaScript 误用、服务误用、图片低效加载、权限误用等缺陷或隐患，提高了 App 的质量与可信性（图 2）。在区块链智能合约领域，提出基于符号执行的交易竞争检测方法、基于变量行为分析的字节码变量及类型恢复方法、基于重构的合约调用 Gas 消耗优化技术等，为我国区块链软件生态建设提供技术储备。

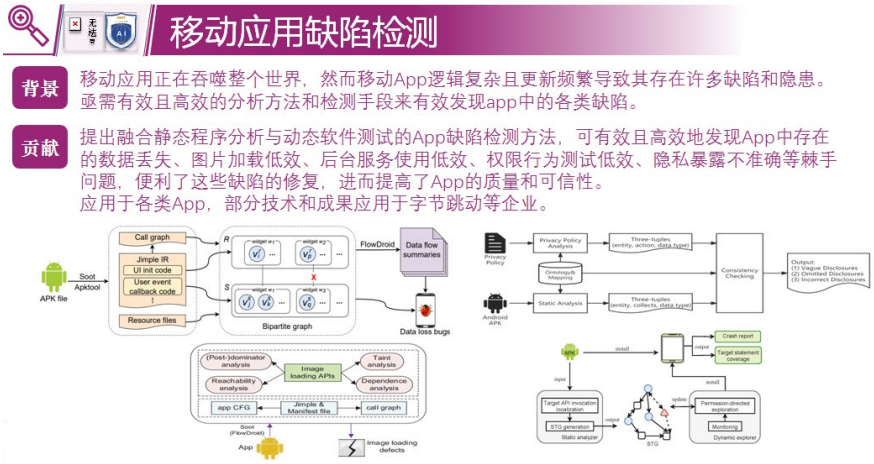


图 2 移动应用缺陷检测

坚持立德树人根本任务，专注于软件创新人才培养

宋巍教授以“传道受业、立德树人”为己任，注重知识传授与实践锻炼并行，深度挖掘学生创新思维，大力培养实操能力，为软件行业输送优质人才。在科研育人方面，依托国家自然科学基金与国家重点研发计划，近 5 年来共指导 20 余名博士 / 硕士研究生，指导的学生中有多人次在软件工程领域顶刊 TSE 与顶会 ICSE、FSE、ISSTA、OOPSLA 以第一作者身份发表论文，其中桂滨法同学的学位论文获 2022 年江苏省“优硕”论文、马辰阳同学获顶会 ESEC/FSE 2023 杰出论文奖（ACM SIGSOFT Distinguished Paper Award）。在实践育人方面，指导的学生团队获 2020 年第 22 届中国机器人与人工智能大赛“室外无人车智能挑战赛”冠军，获 2024 年“中国软件杯”大赛一等奖，获 2024 年中国软件大会第一届软件缺陷自动修复挑战赛一等奖。基于以上育人成果（图 3），宋巍获评教育部 - 华为“智能基座”栋梁之师称号。雄关漫道真如铁，而今迈步从头越。在今后的育人之路上，还需深耕细作，步履不停，持续为软件领域创新人才培养添砖加瓦。

图 3 指导学生多人次获奖

中国 AI 长卷（一）：大国重算

编者按：“中国 AI 到底发展得怎么样了？”在各种社交平台上，我们经常会看到这样的问题，也会看到各种各样的答案，但这些答案有着普遍的缺陷。它们往往只抽取一两个片段或案例，用非常取巧，甚至有点抖机灵的方式，极端唱好或者唱衰中国 AI。

事实上，所谓的中国 AI 产业覆盖面非常广泛。每个领域有各自的发展特点，产业优势以及产业局限性，很难用过分简单的方式，来概括事实上非常复杂的 AI 产业。

或许，复杂的问题就应该有详细的答案。就像一幅小画，画不尽中国广袤的山水。

想要探寻中国 AI 的底色，需要梳理来龙去脉，需要回看一步一景，需要去画一幅长卷。

今天我们都知道，驱动 AI 算法工作的“燃油”是 AI 算力。尤其当深度学习算法发展到了预训练大模型阶段，AI 算力已经成为整个 AI 领域的最大成本开销。根据相关数据，算力成本要占到大模型训练成本的 70% 左右，在大模型推理阶段则高达 95%。

如果说，AI 产业是一间工厂，那么 AI 算力就是工厂所需的煤和石油。更为致命的是，这些“煤和石油”的供应处在一种半垄断状态。在这次 AI 复兴当中，英伟达用 GPU 占据全球 AI 算力市场的主导地位。英伟达的高端 AI 算力不仅成本高昂，供不应求，但对于蓬勃发展的中国 AI 产业来说，能否确保其供应稳定都要打上大大的问号。

在算力贵且不稳的前提下，中国 AI 产业却涌现出了巨大的 AI 算力需求。根据相关数据预测，2030 年全球 AI 算力的需求将达到 2020 年的 500 倍。其中，中国 AI 算力的增长是主要驱动力。目前阶段，中美之间的 AI 算力差不多是 1 比 1.5。种种迹象显示，未来两国间的 AI 算力需求将拉平，甚至中国反超。

成本高昂、供应不稳，需求激增，这三点勾勒出了中国 AI 算力的整体发展背景。

AI 算力就是生产力。在种种令人不安的局面下，中国 AI 开始了聚沙成塔般的算力突围。

2017 年，是人工智能第三次兴起的第一年。在这一年里，AlphaGO 实现了对人类棋手的全面胜利，自动驾驶被广泛看好，深度学习算法四处开花。而这一切算法表现的背后，都离不开 AI 算力的支持。

这一年，英伟达拉开了股价飙升，AI 算力产品频繁迭代的大幕。谷歌开始在云上布局 TPU 等自研算力。全球半导体产业开始看到 AI 算力这个极具想象力的新方向。

而与此前历次半导体风口不同的是，这次中国的从业者们没有后知后觉，待产业成熟后再加油追赶，他们选择了抢跑。

在 2017 年 10 月，海思打造了麒麟 970，把端侧 AI 算力带到了华为手机。11 月，中国科学院和寒武纪共同发布了新一代产品，其中包括面向手机与云端的 AI 处理器。这在当时被称为全球首个深度学习专用处理器芯片。

如果说，这些芯片还更多集中在端侧场景，不能直接对标英伟达提供的高端 AI 算力，尤其是 AI 训练算力，那么到了 2018 年，情况就正式发生了改变。

2018 年 10 月，华为正式发布了全栈全场景 AI 解决方案。构成全站全场景 AI 主体的，是两款华为自研的 AI 芯片，也就是当年发布了用于推理的昇腾 310，以及预告中的昇腾 910，伴随着昇腾这个名字的出现，华为在 AI 基础设施领域的一系列布局开始浮现出来。

彼时，中美之间的贸易摩擦还没有开始。中国科技界不会料想到科技封锁的大棒即将迎面而来，更不会料想到 AI 算力这个还非常新颖、前沿的概念，居然会在几年后成为美国反复操纵，极力打击的中国科技“命门”所在。

如果没有华为对 AI 机遇的预判，昇腾在 AI 算力上的抢跑，或许后面的故事，就会是另一个走向。

2019 年到 2022 年，中国 AI 算力发展进入第二阶段。简要概述这个阶段的发展目标，就是把 AI 芯片变成了 AI 算力。

提及 AI 计算，很多朋友会有种疑惑，一方面国产 AI 芯片似乎非常多，时不时就能看到相关报道，但另一方面却又都说 AI 算力卡脖子。其中的问题，就在于芯片和算力是有区别的。

芯片需要能够量产，能够变成板卡、服务器、小站等计算产品，还需要具备全套的软件生态来帮助用户进行调用、开发，需要与各个应用场景进行适配，证明可用性。在这一系列问题都得到解决之后，还需要形成足够大的市场规模。

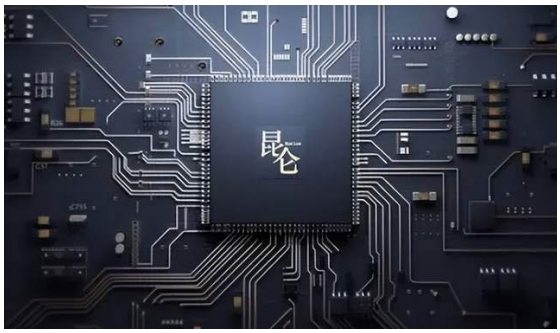
要顶着性能没有英伟达好，成本、生态、商业信任全都没有优势的逆境走向市场，国产 AI 算力这条路非常艰难。这也是为什么绝大多数国产 AI 芯片都只能停留在研制成功的新闻通稿里。

万幸的是，在中美贸易摩擦的背景下，这一阶段重要科技领域的自主可控成为各界共识，而 AI 算力在其中首当其冲。所以，国产 AI 算力没有像此前的算力国产化议题那样，反复被质疑是否有必要自研，全球化采购是否成本更优。因为理智的科技从业者都知道，AI 算力被当作美国的棋子只不过是早晚的问题。

在政策形势、市场需求，以及头部科技企业的带动下，国产 AI 芯片的算力转化虽然没有百花齐放，但也顺利完成了阶段性的目标升级。

2019 年 8 月，可用于 AI 训练，能直接对标英伟达高端产品的昇腾 910 芯片正式发布。其整数精度（INT8）算力可以达到 640TOPS，整体性能接近了英伟达的 A100。这标志着，中国 AI 算力的“拳头产品”来到了全球一线水平。

随后，昇腾生态的建设全面加速。深圳鹏城实验室基于昇腾 910 搭建了“鹏城云脑 II”，实现了中国首个自主可控的 E 级智能算力平台，可以提供不低于 1000Pops 的整机 AI 计算能力和 64PB 的高速并行可扩展存储。在武汉等 25 个城市，搭建了基于昇腾 AI 集群的人工智能计算中心，借助“东数西算”热潮，开启了云端 AI 算力这一新型基础设施的建设。



其他科技公司，同样也在这一阶段推动着 AI 芯片走向 AI 算力。百度在 2020 年量产了昆仑芯 1 代 AI 芯片，随后在百度搜索引擎、小度等业务中进行了部署。随后，基于百度自身业务与百度智能云庞大的 AI 算力需求，昆仑芯片达成了一定的量产规模。

先后布局 AI 芯片的，有华为这样的全产业链科技公司，也有阿里、百度等基于云计算业务拓展的 AI 芯片布局，同时还有寒武纪、海光信息、燧原科技、天数智芯、壁仞科技、摩尔线程、龙芯中科等半导体企业。中国 AI 算力的产业纵深，在一定程度上被拉开，IT 市场的国产化 AI 算力选择也开始多样了起来。

时间来到 2022 年，一个关键性指标开始浮出。根据 IDC 发布数据，2022 年中国 AI 加速卡出货量约为 109 万张，其中英伟达市场份额约为 85%，昇腾市场占有率 10%，百度昆仑为 2%，寒武纪和燧原科技均为 1%。

这意味着，中国 AI 计算市场上的国产化占比已经超过了 10%。虽然这个规模看上去依旧不够大，但它意味着国产 AI 算力已经获得了稳定的市场基数，成为除了英伟达之外，中国 AI 计算具有可行性的第二选择。

这是用极限速度跑出来的 10%，也成为中国 AI 产业的压舱石。

记得 2018 年，我与一些 AI 开发者、AI 公司的创始人聊过 GPU 供应问题。在问到他们是否认为英伟达 GPU 会

走向断供的问题时，大家普遍觉得不用担心，一方面是因为中国市场足够大，且增长足够快，英伟达不可能放弃，另一方面中美之间的 AI 技术差距还很明显，美国政府没有必要在这个领域出手干预。

然而事实证明，达摩克里斯之剑终会落下，我们永远不能乐观地认为科技铁幕上能打开一扇小窗。

2022 年国产 AI 算力能够走向规模化商用的另一重推动力，是因为英伟达高端 GPU 的禁售风波开始了。在此之前，英伟达雄踞了中国 AI 芯片市场超过 90% 的份额。但在 2022 年 10 月，美国商务部以担心军用转化为借口，对出口中国的 AI 芯片启动管制。其中，英伟达的 H100 和 A100 等高端 GPU 成为主要管制对象。

对于这个荒谬的新规，英伟达也并非没有寻找出路。作为禁令的对策，英伟达马上开发了两款专为中国市场设计的“平替”，也就是 A800 和 H800。这两款 GPU 性能都低于美国制裁措施规定的阈值，但在性能降低的同时，价格却进行了上涨。

然而即使这样的替代方案，也在一年后被宣告“此路不通”。美国商务部在 2023 年 10 月宣布禁止英伟达向中国供应 A800 和 H800，而且新的禁售令不仅影响英伟达，还将 AMD 和英特尔的芯片覆盖在内，并且影响了大量芯片设备厂商。这种做法，可谓是堵上了中国获取中高端 AI 算力供应的全部大门，甚至计划对使用亚马逊云、微软云等美国云计算平台来获取云端 AI 算力的中国企业进行限制。铁闸落下，空余无奈。

当然，英伟达也并没停下试试看的脚步。英伟达又一次设计了三款面向中国的“特供版”。其中，能够用于 AI 训练的 H20 在理论上只有 H100 的 20% 综合性能，缩水之严重令人惊叹。

至此我们或许可以说，依靠进口的中国 AI 算力之路已经被堵得水泄不通，接下来，只能路自己修，步自己走。

幸运的是，修出来的路还不止一条。在今天，国产 AI 算力已经可以通过多种方式供应市场。它们支撑着百模大战的繁荣，实现了英伟达禁令甚至没有激起太大的水花。当然，这些方式互有交叠，用户可以有多样化的搭配与选择。但整体而言，今天中国 AI 算力的来源有三条途径：

第一种，全国算力网络与云端 AI 算力。

在科技自立自强的大背景下，几年来中国极大程度上加强了 AI 算力设施的基础建设。作为“东数西算”的核心组成部分，中国兴建了大量智算中心，预计在 2025 年将提供超过 105EFLOPS 的 AI 算力，组成了一张庞大的 AI 算力网络。

而作为与国家 AI 算力基础设施结合相对紧密的运营商，也正在加紧提升对算力网络的利用与挖掘。在目前阶段，运营商纷纷加码云计算与 AI 大模型，逐渐形成了云端 AI 算力在技术上的成熟与长期成本上的优势。

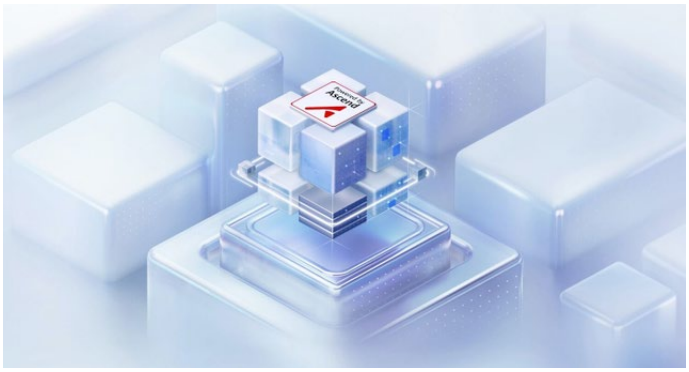
与此同时，各大云计算厂商也加强了 AI 算力的投入。一方面赶在禁令之前，大量囤积英伟达高端显卡，另一方

面也在探索自研 AI 芯片的使用，以及对其他国产 AI 算力的引入。

综合来看，全国一体化的 AI 算力网络，正在成为中国智能化的主要算力基石。

第二种，昇腾生态。

经过多年的发展，华为已经将昇腾建设为国内最为成熟，且完全没有英伟达 GPU 参与的 AI 计算生态。科大讯飞创始人刘庆峰曾经表示，华为的昇腾 AI 芯片可以达到与英伟达 A100 相当的性能。



相对来说，昇腾的优势在于软硬件体系较为齐备，可以广泛使用华为自研的技术进行支持；可以和同样由华为打造的鲲鹏生态结合，实现多元计算；整体产业生态较为繁荣，硬件、软件合作商丰富。弱势之处则在于，外界对昇腾芯片还是有性能不足与价格过高的质疑，并且昇腾生态相对封闭，与其他厂商的 AI 算力生态基本不打通。

今天，中国已经有一半的大模型由昇腾来支撑，并且华为云已经将昇腾算力带到了云端，推出了昇腾 AI 云服务。可以说，昇腾的出现和成长，让中国有了可以对标英伟达生态的 AI 算力选择。

第三种，异构智算。

昇腾之外，大多数 AI 芯片厂商还无法实现规模化出货，更多是以参与混合型算力的方式，加入数据中心、企业 AI 集群的建设当中。目前情况下，大多数企业与数据中心还是会选择英伟达来构建 AI 算力的主体，同时通过加入海光、寒武纪等国产芯片以及加速卡来构建 AI 算力，或者采取使用英伟达 GPU 进行训练，使用国产 AI 算力进行推理的模式。比如说，百度在文心一言训练中使用的是英伟达 GPU，推理侧则使用自主研发的昆仑芯 2 代。

类似策略，可以逐渐降低对英伟达的依赖度，并且发挥出多元化的 AI 芯片优势。由此，异构智算开始成为企业和数据中心新的需求。面向这种需求，IT 厂商也正在捕捉机会。比如联想推出了完全异构智算平台，来帮助实现异构化 AI 算力的管理与调配；新华三推出了面向异构智算的网络解决方案，解决异构智算带来的丢包与负载等问题。

这三根“足”，给中国 AI 算力带来了某种稳定性。经过极限情况下的多年经营与发展，今天中国 AI 算力谈不上充沛与廉价，至少有了可以遮风挡雨的稳固。

至少我们可以看到，中小企业应用 AI 算力的综合门槛正在降低，AI 算力的选择在增多，异构协同能力在加强，并且熟悉了昇腾与海光 DCU 这样能够直接替代英伟达 GPU 的存在。中国 AI 是否会因为算力而陷入生存僵局，已经不再是个问题。

总结一下，在 AI 算力层面，我们有办法，但办法不够好，其实也不够多。

然而换个角度想想，幸好我们有方法，否则麻烦就大了。

依靠精准的预判抢跑，在多重助力下超高速发展，在外部压力下极限成型。

智算，终成国之重器。

面向标记多义性数据的不确定性建模理论与方法研究

项目名称：面向标记多义性数据的不确定性建模理论与方法研究

完成单位：南京理工大学，南京航空航天大学

项目简介：

1、项目背景

人工智能技术近年来飞速发展，广泛应用于各个领域，如医疗、金融、交通等。然而，随着其复杂性和应用范围的扩大，人工智能中的不确定性问题也逐渐凸显。不确定性源自多个因素，如数据不确定性中的数据噪音、数据不完备数据、数据模糊等因素，模型不确定性中的模型结构等因素。因此，如何准确度量和应对这些不确定性成为人工智能领域的重要课题。在多个国家自然科学基金项目和江苏省自然科学基金项目的持续资助下，项目组在面向多义性数据不确定性建模方面取得了一系列进展。

理论层面上，构建了面向多义性数据的不确定性建模理论体系与建模框架。该框架从粒计算的视角出发，系统分析了不确定性产生的多种原因，进而定义了多义性数据建模中的不同类型的不确定性。通过深入研究每种不确定性的独特特点，基于模糊集理论、三支决策理论以及信息论等方法，提出了针对不同类型不确定性的多种度量方式。通过分析各种不确定性度量方式之间的相互关系，揭示了标记多义性空间中不确定性度量的关联机制。

方法层面上，设计了针对多义数据不确定性的各种不确定消解方法。通过对数据中的噪声成分进行分析，实现了对粗糙信息的细化，从而提出不准确数据的噪声消除及粗糙信息的细化方法。通过将数据分解为多个粒度层次，可以更加灵活地捕捉数据的结构信息，从而提出基于多粒度表示与直推模型的不完备数据补全方法。利用模糊隶属度函数将模糊数据转化为可解释的概率分布，设计出基于模糊隶属度的语义概率生成方法。

应用层面上，提出了高维空间图像粒化方法以及粒度自适应的不确定性人脸图像处理方法，构建了变粒度空间中的可解释知识表示与获取方法以及人脸轮廓和关键点的多分辨率表示和学习方法。提出了模糊熵的噪声消解方法，构建了基于多义性数据增强和标记相关性的优化模型。

主要科技创新

近年来，人工智能技术迅猛发展，并在医疗、金融、交通等众多领域得到了广泛应用。然而，随着技术复杂度的提升和应用范围的扩展，人工智能系统中的不确定性问题日益突出。不确定性主要来源于两大方面：数据不确定性和模型不确定性。数据不确定性包括数据噪声、数据缺失以及数据模糊等问题，这些都可能导致模型输入的准确性受到影响。另一方面，模型不确定性则来自于算法的设计和结构，例如模型假设和模型参数选择等因素，都可能使模型的预测结果

产生偏差或误差。因此，如何有效地度量和应对这些不确定性，已成为当前人工智能研究的重要挑战之一。例如基于粗糙度、精度、联合熵以及条件信息熵等对不确定性进行度量，进而从贝叶斯方法、概率推断和三支决策等方面切入，对不确定性进行消解。此外，改进数据质量、优化模型结构、提高模型的可解释性和鲁棒性，都是降低不确定性的重要手段。这些措施不仅能增强人工智能系统的性能，还能提升其在实际应用中的安全性和可信赖性。项目组在多个国家自然科学基金项目资助下，在面向多义性数据不确定性建模方面取得了一系列重要科学发现（如图 1 所示）。

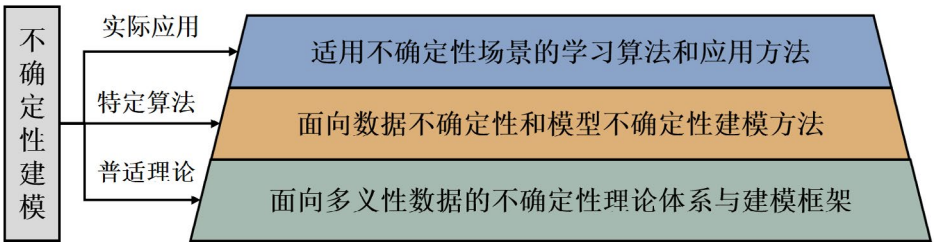


图 1 项目研究思路和科技创新发现

科技创新点一：构建了面向多义性数据的不确定性建模理论体系与建模框架。定义了多义性数据建模中的不确定性类型，并分析了其产生的原因。分析了数据不确定性与模型不确定性的特点。提出了多义性数据不确定性和模型不确定性度量，并分析了各种不确定度量之间的关系。

在现代生活中，数据无处不在，且常常充满多义性和不确定性，如医疗诊断、金融预测和智能交通等领域。传统建模方法难以全面处理这些复杂数据，并且在建模过程中存在模型的不确定性，可能会引发决策偏差。因此，亟需一套系统性的理论体系和建模框架来确保标记多义性数据的不确定性建模的可行性和有效性。

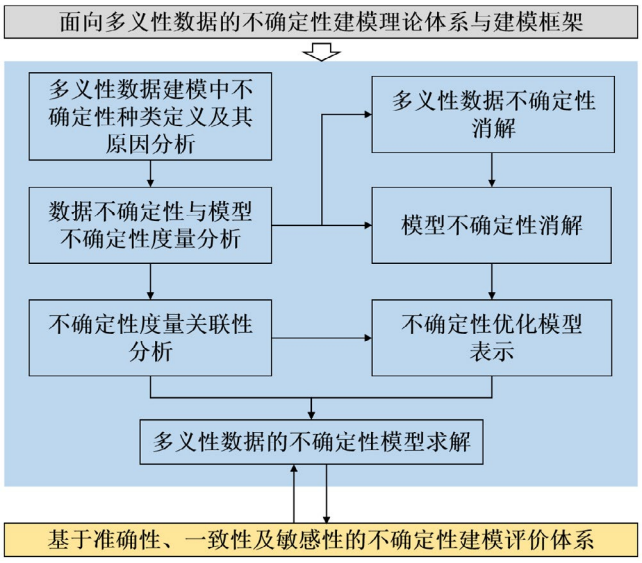


图 2 面向多义性数据的不确定性建模理论体系与建模框架

本发现的主要成果（如图 2 所示）包括：

1）构建了面向多义性数据的不确定性建模理论。基于大量实验观察，深入研究了多义性数据建模的不确定性问题的定义和产生机理。从数据来源、数据特点和数据用途等角度分析了数据不确定性，并定义了不确定性度量用以表示和

刻画数据的不确定性。从模型假设、模型推理和模型的泛化等角度分析了模型的不确定性，进一步研究了如何通过标记相关性和实例相关性等关联关系来降低模型不确定性。同时，提出了一系列评估指标，如置信区间、泛化误差和模型鲁棒性等，用于量化模型在不同情境下的不确定性表现。此外，归纳和总结了现有的各种不确定性度量，并分析了它们之间的内在联系，如等价关系、包容关系等。为后续模型的不确定性表示和求解提供理论性的保障。

2）提出了面向多义性数据的不确定性建模框架。基于前述的不确定性理论分析出发，基于粗糙集模型下的各种粗糙度量方式（包括粗糙度、精度度量、近似精度度量等）以及信息论下的度量方式（包括基于联合熵、条件信息熵、粗糙熵、模糊熵等）量化数据的不确定性，并基于前述的不确定度量分析，为多义性数据的不确定性消解提供理论性的指导。对于模型不确定性，通过在模型假设、模型推理和模型泛化等方面的分析，阐明模型不确定性消解机理以及模型不确定性消解的泛化误差，并基于标签关联关系挖掘和实例关联关系挖掘等方式为模型不确定性提供指导。基于上述不确定性模型的构建与表达，设计相应的算法求解模型，并从模型校准理论和度量理论方面对不确定性消解和模型输出进行概率准确性、排序一致性以及数据敏感性提出了相应的评价指标来衡量不确定性建模的效果。

科技创新点二：开发了针对多义性数据的数据不确定性和模型不确定性建模方法。提出了针对数据不确定性的噪声消除方法、不完备数据补全方法、模糊数据语义概率生成方法，以及基于标记相关性和标记增强的不确定性建模方法。

尽管不确定性的情况在现实世界中的存在形式呈现多样性和复杂化的趋势，但仍可以将所有的不确定性划分到数据不确定性和模型不确定性中，其中数据不确定性主要有噪声不确定性、数据不完整性和数据模糊性三种，通过不同的不确定性度量和消解方法处理该问题。对于模型中的不确定性，通过相应的关联关系挖掘来降低建模型中的不确定性。

本发现的主要成果（如图 3 所示）包括：

1）提出了针对多义性数据不确定性度量与处理方法。对于噪声不确定性，借助三支决策理论以及不确定性度量理论，研究了噪声不确定性的衡量方式，探讨了产生噪声数据的主客观原因。构建了不精确噪声数据的多粒度表示，并明确了粒度转换与精确度变化之间的定量关系。在前述不确定性建模框架内，分别从贝叶斯推理、极大似然估计以及集成学习的角度出发，构建了不精确数据的噪声成分识别与消除以及不精准成分的语义细化模型与算法。对于数据不完备性，研究了数据不完备性的度量方法并讨论了不完备性在不同场景中的潜在表现形式。基于不完备数据的统计特性，探索了不同形式的不完备数据对泛化风险和算法稳定性的影响。在前述不确定性建模框架内，分别借助插值理论、标记传播理论以及主动学习技术，构建了不完备数据的多粒度数据表示以及直推式数据补全方法。对于数据模糊性，提出了针对不同类型数据模糊性的多种不确定性度量方法，以便更精确地刻画数据模糊性带来的影响。基于模糊推理、三支决策理论以及信息论，系统分析了模糊数据的不确定性来源，并从理论上提出了适用于模糊数据建模的解决方案。在这些基础上，进一步通过模糊隶属度的语义概率生成模型，探讨了如何更准确地将模糊数据映射为概率分布，以便在建模过程中保持数据的语义一致性。

2）提出了针对模型不确定性建模方法。基于标记传播理论，构建了结合图模型的推理框架，分析和量化标记之间的相关性。通过学习标记之间的依赖结构，模型能够识别模糊或不确定标记的潜在含义，并通过传播机制推理未标记样本的标签。通过将实例分为多个粒度层次，区分不同层次上的实例间的差异，并根据不同粒度层次的数据，对实例的不确定性进行动态调整。基于贝叶斯推理技术，量化模型的不确定性，并通过估计实例的后验概率来调整其预测结果。此外，还引入了贝叶斯推理和集成学习方法，通过多个模型的预测结果，进一步消解模型中的不确定性。

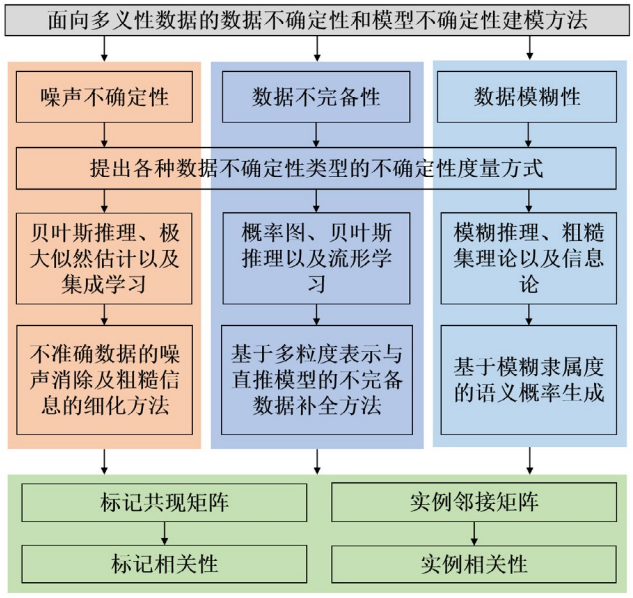


图 3 面向多义性数据的数据不确定性与模型不确定性建模方法

科技创新点三：提出了多义性数据和不确定性建模的学习算法和应用方法。针对人脸识别任务，探索了人脸噪声的多粒度表示对模型泛化能力的影响，提出了高维空间图像粒化方法以及粒度自适应的低质人脸图像增强方法。针对人脸面部情感分析任务，基于不确定性量化分析，提出基于模糊熵的噪声消解，探索了基础情感之间的关联关系，提出了基于低秩假设的人脸面部情感识别方法。

在不同的实际任务中，不确定性往往具有各自的表现形式。因此，如何将前述理论与方法的研究成果与实际任务的需求相结合具有深远的现实意义。本研究点主要讨论人脸识别和人脸面部情感表达分析任务场景，并深入阐发了前述方法与理论如何应用到这两个实际任务中。

本发现的主要成果（如图 4 所示）包括：

1）研究基于多粒度学习的人脸识别理论与方法。基于数据统计结果，分析了低质人脸识别数据的产生原因且归纳了构成低质数据的关键因素，并借助消融实验以及统计理论量化了不同因素对泛化识别误差的影响程度。进一步，从多个粒度层面上对噪声构成因素进行数据建模与表示，并分析了其对泛化能力的影响，从而明确了人脸粒度与泛化识别误差的数量关系。基于多粒度低质人脸数据表示，提出了高维空间图像粒化方法以及粒度自适应的低质人脸图像增强方法，构建了变粒度空间中的可解释知识表示与获取。进一步，人脸图像增强过程同时嵌入了人脸轮廓识别和关键点检测，构建了人脸轮廓和关键点的多分辨率表示和学习方法，从而根据人脸图像质量自适应确定最优计算层面并提取可解释的视觉表征。

2）研究基于不确定消解的人脸面部情感表达分析理论与方法。探索了在情感分析任务中数据不确定性和模型不确定性的不同类型。人脸面部情感数据不确定性可能源于光照变化、面部表情的模糊或情感表达的多义性等问题，而模型不确定性则与情感特征提取过程中模型架构的选择、标记偏差以及标签模糊性相关。紧接着，研究了不同不确定性度量之间的关系，分析它们如何共同作用，影响面部情感表达的识别精度与泛化能力。提出了基于模糊熵的噪声消解和标记相关性的人脸面部情感分析方法，通过计算样本的不确定性程度来消除数据中的噪声，从而增强情感识别的可靠性。此外，

在模型优化过程中，基于多义性数据的数据增强理论与标签矩阵的低秩假设，对模糊信息进行增强和对基本情感的表达进行相关性分析和度量。

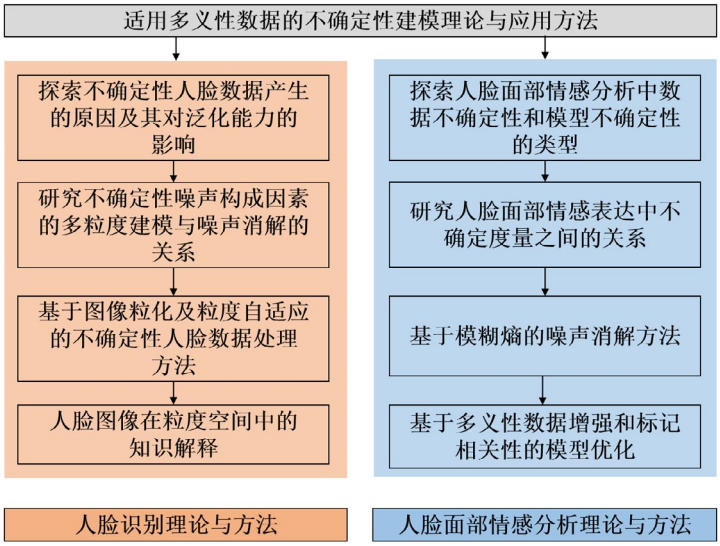


图 4 适用多义性人脸数据的不确定性建模应用

四、社会效益：

智能车载多音区音频管理系统的发布，已在业内形成了极大的影响力，间接引导整个车载音频行业向下一代智能化方向演进，车载智能音频系统在汽车行业的落地，与工信部把发展智能网联汽车作为重要战略方向的目标一致，协同促进国内汽车市场在智能化领域的蓬勃发展，将为中国企业提供在国际市场上竞争的机会，有助于提升国内企业的国际地位和影响力，随着智能音频系统的行业普及，将带动整个行业转型及发展，新兴产业规模预计达 300 亿 / 年。

学会动态

江苏省计算机学会九届二次常务理事会会议在宁召开

2024 年 1 月 2 日，江苏省计算机学会九届二次常务理事会会议在南京大学顺利召开，包括理事长、副理事长、秘书长在内的学会常务理事、监事会主席、工委代表、秘书处全体成员参加了本次会议。

首先学会理事长周志华致辞，他表示本次会议是新一届常务理事会召开的 2025 年的第一场工作会议，希望本届常务理事能为学会接下来的几年做好发展计划和工作安排。

争锋科技公司介绍

南京争锋信息科技有限公司（简称：争锋科技）成立于 2009 年，作为国内领先的金融科技解决方案服务商，争锋科技始终专注于证券、银行、基金、核心交易机构等金融行业，聚焦金融机构数字化免疫系统构建，依托自主创新的全栈式技术解决方案与全生命周期服务保障体系，助力客户及时、准确、低成本的发现自身平台系统可能存在的潜在风险，并迅速提高对不可知风险的抵抗能力、应对能力和系统韧性。争锋科技通过深度整合以混沌工程为核心的数字免疫技术体系，显著提升故障自愈率、应急方案自动化程度、平均修复时间、应急方案成熟度、可观测能力成熟度等信息系统韧性关键指标，为行业客户构建高可用、高可靠的金融业务生态提供重要保障。

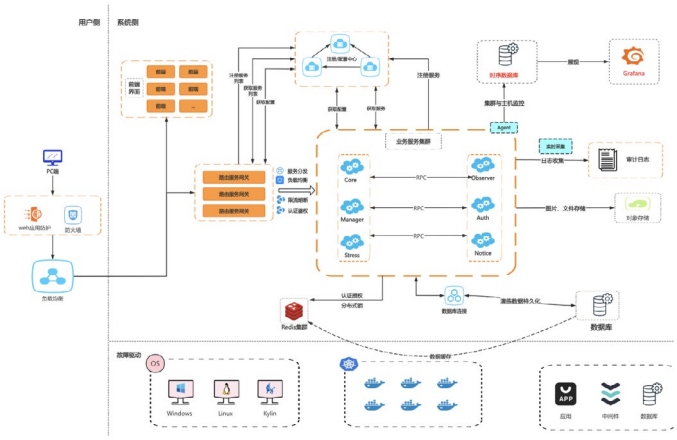
- ✓ 争锋科技数字免疫产品矩阵有助帮助企业在**开发过程尽早识别软件缺陷**，实现**“质量左移”、“安全左移”**，有效降低质量风险成本。
- ✓ 产品矩阵面向开发、测试、运维、SRE等各个部门，并妥善处理。
- ✓ 数字免疫产品矩阵产品联动组合形成更好的**质量保障效果**
- ✓ 为软件开发全生命周期的**各个阶段提供质量保证措施**
- ✓ 兼容DevQualOps理念，将质量管理融入开发流程

*争锋科技入选Gartner《2024年中国信息与通信技术成熟度曲线》最具创新领域之一“数字免疫（DIS）领导企业”。

争锋科技被 Gartner 评为 2024 年度中国数字免疫领导厂商，拥有软件开发成熟度 CMMI5 级、测试成熟度 TMMI5 级、质量管理体系、信息安全管理体系、信息技术服务管理体系、业务连续性管理体系等资质认证，获得国家高新技术企业、软件企业、江苏省专精特新企业、江苏省民营科技企业、南京市工程技术研究中心认定，是信通院稳定性保障实验室副理事长单位、证券基金行业信息技术应用创新联盟成员单位、中国软件行业协会成员单位、信创工委会成员单位。

核心技术：

1、精准故障注入技术



原子故障覆盖范围：公司自主研发的故障注入引擎，能够精准模拟各类复杂的故障场景，包含但不局限于虚拟机、物理机、常用数据库、Java 应用、中间件、容器、k8s 集群故障注入的原子故障。同时，平台支持动态扩展，既可以集成三方工具，亦可以支持用户自定义故障，可精确控制故障的类型、强度、持续时间和发生频率。同时，平台兼容统信、麒麟等国产化信创平台，支持容器、虚拟机、物理机环境，兼容 JDK1.6 版本。争锋科技一直在持续更新迭代故障原子库，目前数量超过 600 个，开箱即用，操作便捷。

爆炸半径控制：采用轻量级探针技术，在不影响被测环境的前提下实施故障注入，故障出发后，系统自动化记录异常响应，在爆炸半径控制上具备多种控制故障影响范围的手段，如设置故障作用的时间阈值、影响的资源数量阈值、业务指标变化阈值等，实现超时终止、阈值终止。还具备智能终止功能，基于对系统整体状态的实时评估，自动判断并终止可能导致系统不可恢复的故障注入，确保故障影响在可控范围内，极大的保障了被测系统业务的快速恢复。

2、智能化场景编排与实验

拓扑编排模式：在复杂的故障场景中，通常涉及多个应用的交互，这些应用或服务可能因各种原因（如网络问题、硬件故障、软件缺陷等）而出现故障，那么故障可能以并发或串行的形式出现，导致整个系统或部分服务的功能受损。拓扑编排模式能够使故障设计者可以清晰地看到应用服务之间的依赖关系，从而更有针对性地进行故障注入和演练，验证系统在不同故障场景下的表现。同时，用户直可以接在应用拓扑图上选择应用节点，并直接在该节点上配置故障模拟选项。这种交互方式极大地简化了故障演练的创建流程，提高了操作的便捷性和直观性，减少了传统方式中需要频繁切换界面和手动输入配置信息的繁琐步骤。

风险场景批量自动生成：根据编排完成的故障场景，轻松选择多个目标系统批量创建故障演练活动。这一过程中，通过批量指定多种故障影响范围（即爆炸半径类型），包括单点故障、单机房故障乃至全机器故障，确保通过丰富的演练活动对系统进行深入且全面的稳定性风险排查，有效增强系统的可靠性和韧性。风险场景封装“演练内容、配置默认参数（参数默认值、参数是否自动获取基础数据、端口是否合并）”，即选择哪些原子故障来模拟，平台

自动设置合理的故障参数规则以达到预期的演练效果。

演练结果自动反馈：平台对自动化采集演练过程数据，涵盖业务连续性、应急响应流程等多个维度的数据，并引入自动分析功能，以减轻人工反馈负担，提高反馈效率与精确度。

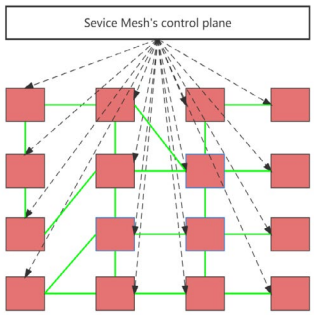
3、多维度监控与韧性评估

多维度观测：为了全面评估系统在混沌实验中的表现，平台具备强大的数据采集能力，它能够从多个维度收集系统数据，包括性能指标（如 CPU 使用率、内存占用、响应时间等）、日志信息、业务指标（如交易成功率、订单处理量等）。

稳定性度量分析：平台内置系统韧性评估体系，根据演练结果对故障定位、业务连续性影响情况、应急过程有效性等方面进行数据指标采集与评估，通过韧性评估体系对系统、团队进行持续评估。

应用场景介绍

1、应用架构韧性验证



通过混沌工程平台开展混沌工程实验，从系统部署架构的角度验证系统应用抵御故障的能力，评估其稳定性是否满足企业需求。

金融行业：在金融行业，系统的稳定性和数据的准确性至关重要。通过混沌工程平台，金融机构可以在上线新的交易系统、支付系统或核心业务模块之前，进行全面的混沌实验。模拟网络故障、服务器宕机、高并发交易等场景，提前发现并解决系统在极端情况下可能出现的数据不一致、交易失败、服务中断等问题，确保金融业务的安全稳定运行，保护客户资金安全和企业声誉。

电商行业：电商平台在促销活动期间往往面临巨大的流量压力，系统的稳定性直接影响用户购物体验和企业的销售额。利用混沌工程平台，电商企业可以在日常运营中持续进行混沌实验，模拟高并发访问、部分服务器故障、网络拥塞等场景，优化系统的弹性和扩展性。确保在 “双 11” “618” 等购物狂欢节期间，系统能够稳定承载海量的用户请求，避免出现页面加载缓慢、下单失败等问题，提升用户满意度和企业竞争力。

云计算行业：云计算服务提供商需要保障其基础设施和云服务的高可用性和可靠性。通过混沌工程平台，对云平台的计算、存储、网络等资源进行混沌实验，模拟节点故障、资源耗尽、网络分区等故障场景，优化云平台的容错机制和自愈能力。同时，帮助云服务用户在将应用迁移到云平台之前，对应用的云适应性进行全面测试，确保应用在云环境中能够稳定运行。

2、监控告警准确性验证

对 IaaS、PaaS、SaaS 构建丰富的混沌实验场景并在预生产 / 生产环境中进行实施，验证整个运维监控体系的完备性、即时性。

验证监控全面性：通过对业务系统注入故障，验证监控的发现能力，以及监控范围覆盖是否全面。

验证监控时效性：通过对业务系统注入故障，验证系统故障和性能瓶颈是否实时监控获取，告警是否及时有效的发送，接收人是否及时有效的接受。

验证监控准确性和有效性：通过对业务系统注入故障，验证监控指标是否准确无误，告警阈值的设置是否合理，告警事件的描述和发送是否准确。

3、应急预案有效性验证

以应急预案为蓝本，以故障注入为驱动，以动态可视化的事故情景为主线，构建多层级多角色协同应急演练，为企业应急处置人员提供一个故障高仿真、情景自定义、应急效果可度量的互动式数字化演练平台。

平台提供应急预案编写录入功能，帮助用户按照规范和标准编制应急预案，并支持与演练故障场景联动，评估预案有效性。用户可以根据具体情况填写设计应急预案的演练内容，包括应急响应流程、责任分工、资源调配、评估度量等。配置完成后，可以基于平台各类故障注入能力，构建应急预案的演练场景，如基础设施、网络、容器、应用、中间件等相关故障风险。

4、红蓝对抗

红蓝对抗演练模式是一种模拟真实攻击与防御的实战化训练模式，旨在通过未知故障的挑战，提升系统的稳定性和团队的应急响应能力。

在红蓝对抗演练中，演练的时间范围、系统范围和参与人员是预设且明确的，但故障类型、具体发生时间和影响范围却保持未知，从而大大增加演练的挑战性和实战性。这种模式下，团队被划分为蓝方团队（攻击组）、红方团队（防守组）和紫方团队（监督评审组和组织保障组），各自承担不同的职责和任务。

蓝方团队（攻击组）利用混沌工程演练平台，融合其压测流量管理、观测指标管理以及监控告警管理等功能，精心编排多种高可用故障演练场景。通过载入预制的参数，快速实例化出各种对抗性的故障演练任务，以挑战业务系统的稳定性。

红方团队（防守组）依靠运维监测系统，密切关注业务系统的运行状态。一旦发现故障，需迅速定位故障原因，

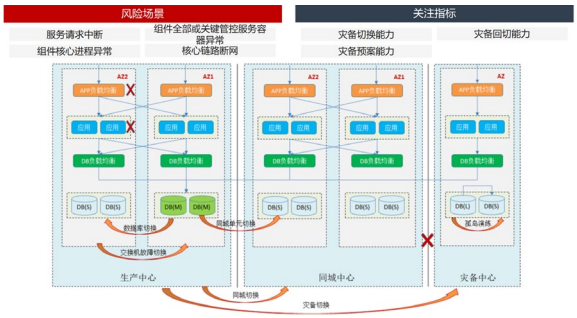
并启动应急处置流程，以确保业务系统的快速恢复。

紫方团队（监督评审组和组织保障组）细分为监督评审与组织保障两组。监督评审组主要负责每轮攻防结果的公正评定；组织保障组主要负责演练协调通知及整体演练流程的组织保障。

5、灾备切换方案验证

通过混沌工程平台模拟基础设施故障导致的整个可用区故障，观察灾备切换是否正常执行，自动及手工容灾切换是否正常，灾备预案是否有缺失、是否可用、是否适用，并通过演练总结报告进行优化整改，以充分保障灾备方案的适用性和可用性。

6、信创组件稳定性验证



通过在信创组件正式投身生产环境前，利用混沌工程模拟多样化故障场景，能够敏锐洞察组件在设计、开发及部署阶段潜藏的各类问题。

混沌工程对信创组件稳定性的验证主要从故障注入、实验执行和结果优化三方面着手。故障注入涵盖硬件层面的CPU、内存、硬盘等故障模拟，网络层面的延迟、丢包、分区等设置，以及软件层面的内核崩溃、数据不一致等情况模拟；实验执行包括依据业务重要性和架构规划实验，利用混沌工程平台自动化注入故障并采集性能、日志、业务等多类数据；结果优化通过分析数据定位问题，制定如优化代码、调整配置等解决方案，并再次实验验证，以此循环改进信创组件稳定性。

混合架构验证：针对信创国产化替代中常见的 ARM 与 x86 混合部署架构，设计故障场景以验证跨芯片架构的兼容性。例如模拟 CPU 资源争用、内存分配异常等，确保不同硬件平台下的组件协同稳定。

信创生态链测试：验证组件与上下游生态（如麒麟操作系统 + 达梦数据库 + 东方通中间件）的集成稳定性，模拟数据同步延迟、接口调用失败等场景。

单位信息

联系人：马宁

联系电话：18066071919

Email 地址：maning@njzfit.cn

公司官网：https://www.njzfit.cn

简　介

江苏省计算机学会理事单位

博智安全科技股份有限公司

博智安全科技股份有限公司（以下简称：博智安全）成立于 2009 年，于 2015 年 1 月 1 日正式运营。博智安全专注以网络靶场细分领域为核心方向，是国家规划布局内重点软件企业、国家高新技术企业、工信部认定的国家级专精特新“小巨人”企业、国家发展改革委认定的“国家企业技术中心”。博智安全秉承“为国家安全锻矛铸盾”为使命，为国防安全网络对抗领域提供产品装备和能力，为加强国家网络边疆防护能力、训练相关人员提升网络对抗能力水平和捍卫民口关基行业网络安全底线而贡献力量。

博智孪生仿真靶场是基于国防安全进行高精拟合，以孪生仿真技术为基础、威胁模拟生成为手段、攻防推演验证为目标，基于电力、智能制造、轨交、卫星、物联网、低空经济等 20 余种行业场景，借助博智靶域智教 AI 大模型前沿算法驱动的智能赋能体系，构建集教学实训、比武竞赛、攻防演练、安全评估、试验鉴定、防御训练、红蓝对抗、钓鱼邮件、实网演练于一体的综合演训平台。

博智安全在国家安全领域耕耘多年，目前已获得 CNVD 原创漏洞发现贡献单位、工信部网络安全威胁和漏洞信息共享平台漏洞报送最具贡献单位、工信部网络安全技术应用试点示范单位、中国网安产业竞争力 50 强、CNCERT 网络安全应急服务支撑单位、工业信息安全监测预警网络建设支撑机构、国家工业信息安全漏洞库优秀成员单位、国家工业信息安全应急服务支撑单位、首批工控安全防护能力贯标技术服务机构、江苏省工控安全工程研究中心、江苏省网络靶场工程技术研究中心、CMMI 五级、ITSS 二级标准化认证等多项荣誉认定，并按照认证要求对公司的研发、生产、交付和内部管理各环节进行严格控制，帮助客户从国家安全体系建设投资中获取收益。

博智安全十分重视人才队伍建设，公司领军人物傅涛博士是国家中组部重点人才工程、江苏省“333 高层次人才培养工程”、南京市人民政府“南京市有突出贡献中青年专家”、南京市科技人才局“中青年拔尖人才”。公司运营和技术骨干获得“南京市五一劳动奖章”、“南京市技术能手”等多项殊荣。

博智安全凭借深厚的技术积累和完备的技术服务，赢得了中国电子科技集团、中国电子信息产业集团、中国航空工业集团、中国兵器工业集团、国家工信部、国家公安部、国家能源集团、中国石化、中国移动等逾千家客户的信赖。

博智安全知识产权体系完备，拥有 70 余项发明专利和 320 项软件著作权。公司多项产品和技术获得省市科技进步奖和金慧奖，并有多项产品获得省高新技术产品认定和省高价值发明专利认定。

博智安全已经引入达晨资本、鼎兴量子、中科科创、清科集团和多家国资委基金等知名投资机构股东。博智安全持续加大产品研发和市场运营投入，加速构建专业化产品及市场运营服务体系，以更加优质的研发成果和安全守护实力回报社会，坚实履行行业先驱企业的社会责任，为我国国家安全的发展添砖加瓦。