

深圳市大数据研究院

Shenzhen Research Institute of Big Data

Autumn 2018



目录

研究院新成员

02

新闻动态

05

讲座报告

10



研究院新成员



程磊博士

程磊 博士分别于 2013 年和 2017 年在浙江大学和香港大学获得学士和博士学位。期间，他曾担任美国北卡罗来纳州立大学(NCSU)的研究助理(2012 Summer)和阿里云数据科学研究院的实习算法工程师(2016 Summer); 他因学业优秀被授予浙江大学优秀毕业生，研究优秀被授予香港大学 P.K.Yu 纪念奖学金，教学优秀被授予香港大学最佳助教奖。2017 年–2018 年，他任职于阿里巴巴菜鸟网络，担任高级算法工程师，负责设计了智能能仓内机器人拣选系统的资源分配算法与路径规划算。2018 年 11 月，他加入深圳市大数据研究院，担任研究科学家。

程博士的研究兴趣主要包括张量数据挖掘，大规模贝叶斯推断与优化，以及混合整数规划。主要应用包括社交网络分析，图像/视频处理，阵列信号处理，无线通信，资源分配与网络规划等。他已经在国际学术刊物及会议上发表论文 10 多篇，包括信号处理学会顶级期刊 IEEE Trans. on Signal Processing。他长期担任主流学术期刊如 IEEE TSP, TNNLS, TCOM, TVT 等的审稿人。

杨柳莎 博士于 2012 年获得北京邮电大学学士学位，于 2018 年获得香港科技大学博士学位。她于 2015 年在法国巴黎中央理工–高等电力学院进行学习和访问。她于 2018 年 12 月加入深圳市大数据研究院。她的研究方向是高维随机矩阵理论在协方差估计和信号处理中的应用。具体应用领域包括投资组合优化的算法设计，雷达阵列信号处理和都市供水管道故障检测等。



杨柳莎博士

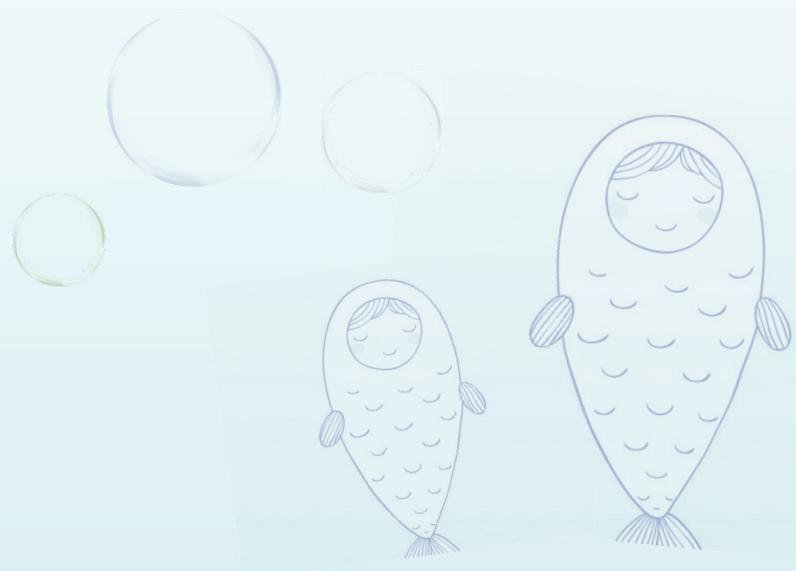


访问教授



Oleg Burdakov 教授

Oleg Burdakov 自 1999 年来于瑞士林雪平大学任职教授，同时兼任美国盖恩斯维尔佛罗里达大学应用优化中心的教职员。1998 年，他担任巴西坎皮纳斯大学的客座教授。他还曾担任俄罗斯科学院计算中心（1980–1994）和法国图卢兹 CERFACS（1995–1997）的研究科学家。他分别于 1977 年和 1980 年获得了莫斯科物理科学与技术学院的硕士学位与博士学位。他是期刊《优化方法和软件》的主编。2014 年他参加由美国离散数学和理论计算机学中心（DIMACS）组织的解决最小生成树问题国际挑战赛并获奖。2017 年，他获得了中国科学院院长国际奖学金计划下的访问科学家奖。他的研究方向包括：解决优化问题和非线性方程系统的数值方法，尤其是牛顿型、稳定割线和插值方法、全球化策略、基数约束优化、反问题、多线性最小二乘、非光滑优化和方程、线性和非线性鞍问题和单调方程、单调回归、数据拟合和插值、跳限制最短路径和最小生成树问题。



访问学生



胡雯静
香港中文大学（深圳）
计算机科学
研究方向：学习分析



叶昊林
香港中文大学（深圳）
计算机科学与工程
研究方向：服务器架构与运维，软件工程

王钦
中山大学
计算机科学与技术
研究方向：计算机视觉



沈雅龄
香港中文大学（深圳）
计算机科学与技术
研究方向：大数据医疗



肖婕
香港中文大学（深圳）
数据统计
研究方向：商业分析



张浩凝
香港中文大学（深圳）
计算机科学与工程
研究方向：深度学习、图像处理



新闻动态

深圳市机构编制委员会及深圳市事业单位登记管理局访问我院

2018年9月7日，深圳市机构编制委员会办公室主任徐波，副主任王耀文，综合处处长李智及深圳市事业单位登记管理局长查竞春一行访问香港中文大学（深圳）暨深圳市大数据研究院，并与研究院的成员举行了交流座谈。出席此次座谈的研究院成员包括副院长崔曙光教授，研究员韩晓光博士、陆彦辉博士、李钰鹏博士及周健君博士。

崔曙光教授首先介绍了深圳市大数据研究院的基本情况和发展状况。随后，双方针对未来合作的可能性进行了讨论，尤其说明了深圳市大数据研究院的前世今生，以及深圳市大数据研究院依托于香港中文大学（深圳）的特别优势。座谈结束后，深圳市编办一行参观了香港中文大学（深圳）校园，并与校领导进行了另一轮座谈。



(深圳)理工学院本科学生成获天池 IEEE ICDM 2018 全球气象 AI 挑战赛冠军



近日，在刚刚落幕的天池 IEEE ICDM 2018 全球气象 AI 挑战赛决赛轮中，来自香港中文大学（深圳）理工学院的四位本科学生成组成的参赛队伍一举夺魁！

天池 IEEE ICDM 2018 全球气象 AI 挑战赛由阿里云·天池平台联合深圳气象局和香港天文台联合举办，

旨在利用人工智能技术帮助粤港澳大湾区解决暴雨防范的问题。





比赛分为初赛和决赛两轮，从五月到十月一共持续五个月，共有来自全球的 1700 多支队伍参加比赛。除了奖金奖励之外，前五名队伍还被受邀参加 IEEE ICDM 2018 新加坡大会进行展示与颁奖。参赛队伍通过天池平台下载 SRAD2018 训练数据集和初赛测试集。每一个训练样本有 61 张雷达图像，每张雷达图像间隔 6 分钟。参赛队伍需要自己设计算法，根据前面的 31 张图片 (3h, 6 mins/image) 去预测后面 6 张图片 (3h, 30 mins/image)。模型确定好后，跑测试集并提交预测结果。

比赛意义

天气预报对人们的出行有着决定性的作用。气象人员能够准确的预测出天气的变化，对社会的意义是不容小觑的。小到普通大众的旅行游玩，中到台风的航线预测，大到飞机的起飞降落，每一步都需要准备的天气来决定。一旦失误，有可能是旅途不快，更严重的将是生命之星的陨落。气象局能够通过举办这个比赛来号召 AI 团队用最先进的技术去更准确的预测天气，这对社会的意义是十分重大的。



比赛过程

初赛刚开始时，在参赛同学们拿到数据后，指导老师把他们分成了两组，让他们尝试从不同方向进行攻坚与实践，到后来确定较好的模型后，同学们开始对数据进行筛选和分类，并不断尝试新的思路和模型优化，最后成功进入了决赛。顺利地进入决赛让同学们更加自信。在各位指导老师的帮助下，他们成功从回归模型预测转到分类预测，这一次他们完成了质的飞跃，从 20 多分的成绩一下跃居第二。此后他们更加小心，珍惜每一次提交数据的机会，最终拿到冠军！

选手感言

宁述亮：计算机科学与技术专业 大四 逸夫书院 毕业于章丘四中

作为这次比赛队伍的队长，很高兴能够和队友一起拿到比赛的冠军。一开始大家都不是很懂气象这一块图像应该是什么样的情形，也对 video prediction（图像预测）一知半解，真的



很感谢教授们的帮助，把我们带进科研的大门。经过这次比赛，我学到了很多关于网络搭建和模型优化的知识，也第一次真正了解了一个项目应该怎样零的突破到厚积薄发。模型效果不好是没事的，只要我们还有改进的 idea（点子），多花时间去尝试，一定会得到一个较好的结果。很荣幸能够受邀参加 ICDM 的会议，在会上听了很多大牛的演讲，学了很多知识，也开阔了自己的视野。但是我们还存在很多的问题和不足，比如不熟悉气象图像，以致于我们的数据处理看起来貌似都不是很 professional（专业）。这次会议的完美体验更加坚定了我去探索，去钻研的决心，期待下一次的我，能够带着论文参加会议。



洪一舸：数学与应用数学专业 大三 学勤书院 毕业于杭州第二中学

抱着重在参与的初心加入了团队，不想成了团队的技术核心，更没想到竟拿了冠军。从5月到10月，近半年的参赛过程历经坎坷，一次又一次山穷水尽，一次次峰回路转。全新的领域、几乎全新的知识，30多篇论文的学习借鉴，一次又一次算法尝试和参数调整，几千行代码的反复修改，似乎快要突破，却又擦肩而过。我比较骄傲的是从七月开始的五周，每周我都能有进展，基本上都是一口气把程序改通。特别感谢指导老师韩晓光教授，李镇教授和黄锐教授的倾力指导，在我感到无助和迷茫时，给予大力支持；也感谢队友们的携手共进，做出很多有价值的探索，并相互分担压力。这次比赛让我体会到要解决问题，要充分理解业务特点，这样才能发现问题的关键，少走弯路，思路也会更开阔。今后继续努力。

杨天行：计算机科学与技术专业 大三 学勤书院 毕业于淄博第四中学

这次比赛给我的最大收获是学会变通，同时要不停地多尝试新方法。一开始我们尝试了三种方法，到最后保留下来一种，再加以各种形式的拼改，最终形成现在的模型；会议中我们和一队气象从业人员交流了很多，了解到不少实用经验，都可以改善我们的测试结果。我们一直在尝试，改变，也错过了很多可能性。错过了或许会有些遗憾，可如果不尝试，不及时改变，理想的结果就更不可能。

陈浩宇：计算机科学与技术专业 大二 逸夫书院 毕业于温岭中学



有很多需要改进的地方，同时也扩充了不少知识面，掌握到了一些方法，这让我在以后能够更好地去实现一些目标。当然也认识到了三位非常好的队友，非常感谢三位学长，每一个人真的都非常的好，都带给了我很大的帮助。

指导教授点评



韩晓光教授：

气象预报问题具有非常重要的社会意义，其本身也是数据科学领域多年来密切关注的研究课题。6个月前，我们的同学还没有这方面的专业背景知识，对气象图像处理以及机器学习算法也知之甚少。到现在已经能够非常熟练的进行数据处理以及掌握最先进的深度学习技术，甚至将现有方法改进以适用于气象预测任务并一举夺得冠军。我见证了他们的成长，也为他们感到骄傲，并希望各位同学能再接再厉，继续接受未来一个又一个的挑战！

韩晓光博士，现任理工学院研究助理教授。他在2017年9月获得香港大学计算机科学专业博士学位。在此之前，他于2009年本科毕业于南京航空航天大学数学系，于2011年在浙江大学应用数学专业获得硕士学位，并于2011年至2013年间在香港城市大学担任研究助理。韩博士于2017年9月加入香港中文大学（深圳）和深圳市大数据研究院，从事计算机科学领域的教学科研工作。他的主要研究方向包括计算机视觉，计算机图形学，人机交互以及医学图像处理。



李镇教授：

我们几位同学在成功进入决赛后，准确地分析总结初赛中遇到的问题，果敢地在决赛中把比赛策略从图像回归问题切换到粗粒度分类问题。不骄不躁，在遇到新问题时不怕困难，不怕吃苦，与三位导师及时沟通，队员之间紧密合作，做事踏实刻苦。天道酬勤，冠军是对我们同学最好的嘉奖和鼓励。

李镇博士，现任理工学院研究助理教授。分别在2011年和2014年于中山大学获得学士和硕士学位，在2018年于香港大学



获得博士学位。李镇博士同时在 2016 年和 2018 年于芝加哥大学，丰田芝加哥研究院进行访问学者研究工作。李镇博于 2018 年九月加入香港中文大学（深圳）和深圳市大数据研究院。

他研究方向主要是利用数据挖掘和深度学习算法进行蛋白质结构预测，从序列层面到折叠层面。他是蛋白质结构预测竞赛 CASP12 冠军的主要成员，并获得 PLOS CB 2018 最新突破和创新奖项。他同时从事机器学习算法和三维计算机视觉问题的研究，例如 RGB-D 语意分割，形状补全等。

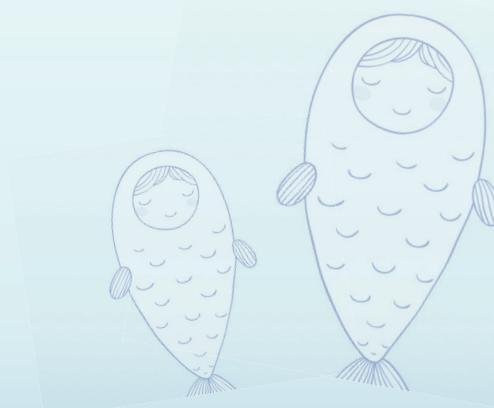


黄锐教授：

我希望同学们再接再厉，把在这次比赛中学到的知识和技能充分消化，灵活运用到学校学习和其它实践中去，再创佳绩！

黄锐博士，现任理工学院副教授。先后毕业于北京大学（理学学士，1999），中科院自动化所（工学硕士，2002），和美国罗格斯大学（博士，2008）。博士毕业后他在罗格斯大学继续从事了两年博士后研究并于 2010 年回到中国加入华中科技大学担任教职。2012 年至 2016 年，他曾在 NEC 中国研究院担任研究员，现任香港中文大学（深圳）理工学院副教授。

黄教授在数据降维和子空间分析、可变形模型（deformable models）、概率图模型等方面以及它们在计算机视觉、模式识别、（医学）图像处理中的应用等方面做过大量研究工作，近年来主要从事智能视频监控相关的研究，包括行人的检测、跟踪与识别等。目前研究兴趣集中在计算机视觉在机器人领域的应用。黄教授已发表相关领域学术论文 50 余篇，主持过包括国家自然科学基金在内的多项科研课题。





讲座报告

一、高维数据的可视化分析

时间：2018/8/17 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：成生辉博士，深圳市大数据研究院研究科学家

讲座摘要：

高维度数据与生活与物质的许多方面息息相关，例如医疗记录、环境监测、商业市场及社交网络等。对于人类而言，理解数据之间错综复杂的关系是一项挑战。可视化分析可以提供强大的机制、通过从原始数据中挖掘关系并将其雕刻为与交互相关的影像来获取洞察力，从而帮助人们探索和利用这些复杂的数据。所有可视化工具都使用了某种投影策略以达到在两个屏幕尺寸的范围内传达高维空间的目的。由于这一预测在除了最微不足道的案例之外都是一个固有的不适当问题，所有方法都会有一定的优劣。了解现存范例的优缺点可以为设计最适合的可视化策略提供信息。它可以帮助从业者在众多可用工具中选出最佳工具，并且可以帮助研究人员设计新工具来推进最新技术发展。

嘉宾介绍：

成生辉博士是深圳市大数据研究院的研究科学家。他的研究重点是大数据、数据科学、数据可视化、可视化分析、数据挖掘等。他于 2016 年至 2018 年期间在美国布鲁克海文国家实验室担任客座研究员，2015 年还担任德国莱比锡大学医学信息学、统计学和流行病学研究所的科学研究员。他分别于 2016 年和 2018 年获得美国石溪大学计算机科学硕士和博士学位。

二、基于中文电子病历的医疗知识挖掘与应用

时间：2018/8/20 星期一 上午 10:30–11:30

地点：诚道楼 207

嘉宾：关毅教授

讲座摘要：



中文电子病历是具有较高可信度的医疗大数据之一，蕴含了丰富的医疗知识，是医疗人工智能研究重要来源。本报告系统地介绍了演讲人及其团队在中文电子病历的医疗知识表示、医疗知识挖掘以及医疗知识的应用等方面的研发工作。

嘉宾介绍：

关毅，男，1970 年生。长期从事自然语言处理相关研究。曾主持、参加过多项国家自然科学基金重点及面上、八六三项目重点及面上项目，以及国际国内合作项目的研究。在 IEEE TRANSACTION on SMC 等国际期刊，《中国科学》等国内一级期刊和核心期刊，ACL 等国际国内会议上发表论文一百余篇，提出了系统相似及其测度理论等多项创新成果。中国专利七项，美国专利两项。拥有面向移动平台的自主产权输入法产品 WI 输入法，用户量数十万。作为主要研制人参加微软拼音输入法、BOPOMOFO 智能汉字输入系统、Wenwen 智能搜索引擎等项目研发。出版专著一部。主要研究方向包括：医疗健康信息学，自然语言处理，知识工程。承担哈工大研究生专业必修课《自然语言处理》等课程的教学工作。

三、分布式（深度学习）作业的在线网络调度

时间：2018/8/24 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：李钰鹏博士，深圳市大数据研究院研究科学家

讲座摘要：

在这个大数据时代，机器学习工作——尤其是深度学习工作——通常以分布式方式执行。这些工作通常由 MapReduce, Dryad 和 Spark 等计算框架支持，这些框架在云机器学习应用程序中非常流行。在这些框架中，一个作业的数据流可以共享一个共同的性能目标。但是，当云提供商旨在优化网络级指标时，这种应用级需求通常会被忽视。为了提高应用程序级性能，我们设计了有效的方案来智能地调度数据流来满足作业的要求。

嘉宾介绍：

李钰鹏博士于 2017 年获得香港大学计算机科学系博士学位。随后，他在香港大学担任博士后研究员。现在，他担任深圳市大数据研究院的研究科学家，专注于研究网络、AI 多智能体系统和经济学之间的跨学科主题。他专注于具有性能保证和算法博弈论的在线算法技术，并且倾向于理论和应用问题。



四、来自部分矩阵采样的低秩近似：无假设局部最优分析和在内存高效内核 PCA 中的应用

时间：2018/8/29 星期三 上午 9:30–10:30

地点：诚道楼 207

嘉宾：李晓东，加州大学戴维斯分校统计系

讲座摘要：

在本次演讲中，我们从无假设逼近的角度研究非凸矩阵的填充，即对半正定矩阵的秩、特征值或特征向量没有任何假设情况下，我们建立了基于目标函数任意局部最小值的低秩近似误差。作为有趣的发现，当对秩、特征值、特征向量和采样率施加某些假设时，我们的方法能改进现有文献针对非凸矩阵填充的最新结果，没有假的局部最小值。我们还讨论了如何将所提出的低秩近似框架应用于存储效率高的内核 PCA。数值实验表明，与众所周知的 Nystrom 算法相比，我们的方法在近似精度方面具有竞争力。

嘉宾介绍：

李晓东博士是加州大学戴维斯分校统计系助理教授。在此之前，他曾在宾夕法尼亚大学沃顿商学院统计系工作了两年。他分别于 2008 年和 2013 年获得北京大学学士学位和斯坦福大学数学博士学位。他对机器学习、统计学、优化和信号处理有着广泛的研究兴趣。另外，他还对优化/光谱方法与潜在的低秩/光谱结构之间的联系拥有浓厚兴趣。他的论文发表在各种统计学、数学和工程学期刊上，例如 AoS、ACHA、FOCM、JACM、IEEE TIT 等。

五、物体检测和场景文本检测简介

时间：2018/8/31 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：齐宪标，深圳市大数据研究院研究科学家

讲座摘要：

对象检测是计算机视觉中的一项具有挑战性的任务。最近，深度卷积神经网络显著提高了物体检测的性能。在本次演讲中，我将介绍最先进的物体检测方法，包括 SSD (Single Shot Multi-boxes Detection)、YOLO (You Only Look Once)、更快的 R-CNN 等。另外，我还讨论一个重要且极具挑战性的物体检测场景 – 自然场景中的文本检测，伴随着一些关于工业界和学术界最新作品的讨论。



嘉宾介绍：

齐宪标博士分别于 2008 年和 2014 年获得了北京邮电大学 (BUPT) 信息工程学士学位、以及信号和信息处理博士学位。从 2011 年 1 月到 2012 年 5 月，他担任了微软亚洲研究院 (MSRA) 网络搜索和挖掘小组的实习生。2012 年 7 月至 2013 年 5 月，他访问了深圳先进技术研究院 (SIAT)。从 2014 年 1 月至 2016 年 4 月，他先在芬兰奥卢大学担任过访问研究员和博士后。2016 年 5 月至 2018 年 5 月，他在香港理工大学担任博士后研究员。从 2018 年 6 月起，他在深圳市大数据研究院的研究科学家。他的研究兴趣主要在于计算机视觉和模式识别，其中包括面部分析、物体检测、医学图像分析等。

六、信号处理中鲁棒统计的研究进展

时间：2018/9/18 星期二 上午 11:00–12:00

地点：诚道楼 207

嘉宾：Abdelhak M. Zoubir 教授，德国达姆施塔特工业大学

讲座摘要：

随着脉冲测量环境和实际工程系统中的异常值的增加，鲁棒统计学在持续变得越来越重要。经典估计或检测理论不适用于这种情况，因此需要鲁棒统计方法。该讲座旨在讨论鲁棒统计的最新进展，并展示其解决信号处理问题的能力。首先，我们强调在现实生活中使用鲁棒统计的动机，以及鲁棒统计如何在这些实际系统中解决问题。然后，我们介绍了一些鲁棒性的定义，并讨论了一些鲁棒估计。讲座的第二部分涉及鲁棒检测。信号及天线阵列处理多个领域的应用支持这些理论处理。

嘉宾介绍：

Abdelhak M. Zoubir 是 IEEE 院士，也是 IEEE 杰出讲师 (Class 2010–2011)。1992 年，他在德国波鸿鲁尔大学获得了博士学位。1992 年至 1998 年，他在澳大利亚昆士兰科技大学担任副教授。1999 年，他加入澳大利亚科廷科技大学担任教授。2003 年，他加入德国达姆施塔特工业大学，担任信号处理教授和信号处理小组负责人。他的研究方向主要是信号处理的统计方法，重点关注 bootstrap 技术、鲁棒检测和估计、以及应用于电信、雷达、声纳、汽车监控和安全以及生物医学的数组处理。他在上述领域发表了 400 多篇期刊和会议论文。Zoubir 博士担任众多国际 IEEE 会议和研讨会的主席和技术主席。最近，他担任意大利佛罗伦萨 ICASSP-14 的技术



联合主席。此外，他还曾担任各种期刊的出版委员会成员，尤其是 IEEE 信号处理杂志（2012–2014）的主编。Zoubir 博士是 IEEE 信号处理协会（SPS）技术委员会信号处理理论和方法（SPTM）的主席（2010–2011）。他曾担任 IEEE SPS 理事会成员，并且是欧洲信号处理协会（EURASIP）的主席。

七、面向虚拟机功率计量 – 一种决策树方法

时间：2018/9/21 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：顾崇林博士，香港中文大学（深圳）博士后

讲座摘要：

近年来，世界各地都建立了大量的云数据中心。它为数据中心的电源管理带来了新的挑战，例如电源监控和节能调度。如果我们知道每个虚拟机的功耗，那么所有这些挑战都可以更容易地解决。由于 VM（Virtual Machine 虚拟机）在软件级别运行，我们需要采用建模方法来测量其功率。本次讲座将探讨如何测量或估算每个虚拟机的功耗。此外，我们还将讨论虚拟机电量计量在实现云计算中的按需付费模式中的重要性。

嘉宾介绍：

顾崇林博士于 2018 年 4 月在哈尔滨工业大学（深圳）计算机科学与技术系获得博士学位。之后，他在香港中文大学（深圳）担任访问学生。现在他是这里的博士后。他目前的研究重点是绿色云数据中心的在线或实时调度。

八、医学图像分析中的深度学习

时间：2018/9/28 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：王昌淼博士，香港中文大学（深圳）博士后

讲座报告：

深度学习正在迅速成为最先进的技术，为多种医疗应用的性能提供改良。具体而言，对 X 射线、计算机断层扫描（CT）和磁共振成像（MRI）等高级模式的分析从深度学习的数据驱动性质



中获益最多。在本次报告中，我们涵盖了医学图像分类、定位、检测、分割、注册等关键研究领域和应用。最后，我们讨论了未来研究的开放性挑战和可能的方向。

嘉宾介绍：

王昌淼博士于 2018 年 6 月获得中国科学院大学博士学位。现在他与崔曙光教授一起研究，担任香港中文大学（深圳）的博士后。他的研究兴趣集中在开发机器学习和深度学习模型以解决医学图像分析挑战。

九、智能电网的流式分析

时间：2018/9/24 星期一 上午 10:30–11:30

地点：诚道楼 207

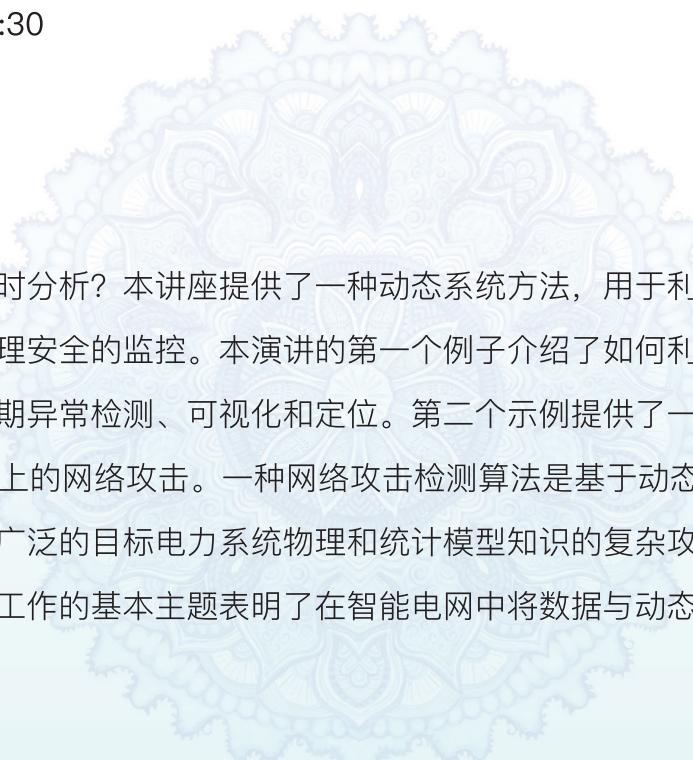
嘉宾：谢乐教授，德克萨斯州农工大学

讲座摘要：

如何对电网中的流量测量数据进行实时分析？本讲座提供了一种动态系统方法，用于利用不同时间尺度的数据来改进对网格网络和物理安全的监控。本演讲的第一个例子介绍了如何利用同步相量数据降维和鲁棒主成分分析进行早期异常检测、可视化和定位。第二个示例提供了一个在线框架，用于检测自动生成控制（AGC）上的网络攻击。一种网络攻击检测算法是基于动态水印的方法而设计出来的。该检测算法为具有广泛的目标电力系统物理和统计模型知识的复杂攻击者发起的网络攻击检测提供了理论保证。该工作的基本主题表明了在智能电网中将数据与动态情境感知模型相结合的重要性。

嘉宾介绍：

谢乐教授现为美国德州农工大学电子与计算机工程系副教授。谢教授于 2004 年在清华大学电子系获学士学位；于 2005 年在哈佛大学获工程科学硕士学位；于 2009 年在卡耐基梅隆大学电子与计算机系获博士学位。他的行业经验包括 2006 年在美国新英格兰独立调度公司（ISO-New England）实习，2007 年在美国爱迪生能源公司营销和贸易实习。谢教授的研究领域主要为电力系统大数据建模，电力系统运行与控制，低碳电力技术、电力市场。他曾经获得美国国家自然科学基金杰出成就奖，美国能源部 Ralph E. Powe Junior 杰出成就奖，2016 IEEE SmartGridComm 最佳论文奖，以及 2012 North American Power Symposium 最佳学生论文





奖。谢教授担任多个国际会议的程序委员会主席、组织者或成员，并且现在是 IEEE Transactions on Smart Grid 的编委。

十、目标跟踪的非线性集成员估计融合

时间：2018/10/12 星期五 下午 3:00–5:00

地点：诚道楼 207

嘉宾：王治国博士，香港中文大学（深圳）博士后

讲座摘要：

随着数据科学和信息科学的迅速发展，多传感器信息融合技术已成为航空航天、机器人、自驾车等多个领域的重点研究课题。精确模型的估计融合方法难以满足复杂竞争环境的要求，因此我们需要研究基于不确定模型的鲁棒估计融合方法。在本报告中，我们结合优化理论和非线性集合隶属度估计，推导出多传感器集合融合的一些解析公式，为保证跟踪系统的高精度，稳定性和有效性提供了理论依据。

嘉宾介绍：

王治国博士于 2018 年 6 月在四川大学获得博士学位。现在他是罗智泉教授的博士后研究员。他的研究兴趣集中在目标跟踪、信息融合、非凸优化、机器学习等方面。

十一、高效有效的机器阅读理解模型

时间：2018/10/30 星期二 上午 10:30–11:30

地点：诚道楼 207

嘉宾：余韻

讲座摘要：

机器阅读理解在机器学习和自然语言处理领域中引起了广泛的关注。在本次演讲中，他将介绍两种高效且有效的模型来完成这项任务。

首先，他将提出一个模型 LSTM-Jump，它可以跳过序列数据中不重要的信息，模仿人类阅读时的略读行为。通过有效的强化学习算法训练，该模型在推理时间上会比 vanilla LSTM 快几倍。



随后，他将介绍一种避免循环网络的序列编码方法，从而完全支持并行训练和推理的进行。基于这种技术，一种新的问答模型 QANet 被提出。与通过反向翻译的数据增强方法结合，该模型获得比寻常模型更快的速度，在 2018 年 8 月斯坦福问答数据集 (SQuAD) 中达到了最佳性能。值得注意的是，QANet 的精确匹配得分大大超过了人类的表现。

嘉宾介绍：

余韡是卡内基梅隆大学机器学习系的博士候选人，导师是 Jaime Carbonell 教授和 Alex Smola 教授。他的研究兴趣是人工智能，包括深度学习、大规模优化和自然语言处理。他研究的重心是通过设计有效的模型和算法来加速 AI。他的研究成果已发表在各种顶尖会议和期刊上，包括 ICML、NIPSICLR、ACL、COLT、JMLR、AISTATS、AAAI 和 VLDB。他的论文入围 NFORMS 2014 数据挖掘最佳学生论文，他的合著论文也被提名为 2011 年 ICML 最佳论文。他是 Nvidia 博士研究员、Snap 博士研究员、Siebel 学者和 CMU 院长研究员。此外，他还曾担任 2017 年 AISTATS 的工作流主席。

十二、能量收集无线系统的多样性及其开发利用

时间：2018/11/2 星期五 下午 2:30–3:30

地点：诚道楼 207

嘉宾：李航，深圳市大数据研究院访问学者

讲座摘要：

本演讲主要讨论我们如何看待能量收集 (EH) 无线系统中的多样性——一种重要的无线通信技术，以及我们如何利用多样性。首先，回顾一些无线通信中的基本概念，例如信道、错误率、容量、和多样性。此外，还将介绍有关 EH 设置的基础知识。其次，EH 设置下的多样性基本原理将分别在点对点通道和多用户场景中呈现。第三，分布式机会调度问题也会被讨论，其中会简要介绍了最优停止理论，因为它是调度问题中利用多用户、多样性的主要技术框架。

嘉宾介绍：

李航博士分别于 2008 年和 2011 年在北京航空航天大学获得 B.E. 和 M.S. 学位。2016 年，他获得美国德克萨斯 A&M 大学博士学位。2016 年 9 月至 2017 年 8 月，他曾担任德克萨斯 A&M 大学的博士后研究员，2017 年 9 月至 2018 年 3 月他又于加州大学戴维斯分校担任博士后



研究员。自 2018 年 4 月起，他担任深圳市大数据研究院的访问研究学者。他目前的研究兴趣包括无线网络、随机优化、和机器学习的应用。

十三、一种解决大型跳跃约束 Steiner 树问题的双标准方法及其在无人机定位通信中的应用

时间：2018/11/16

星期五 下午 4:00–5:00

地点：诚道楼 207

嘉宾：Oleg Burdakov 教授，瑞典林雪平大学

讲座摘要：

我们考虑定向的 Steiner 树问题 (DSTP)，其中约束树中的弧（跳）总数。这个问题已被证明是 NP hard。对于大小超出现有精确算法容量的实例，只能应用启发式算法。跳跃约束的 DSTP 被视为双标准问题，其中树成本和跳数被最小化。我们推导出最优化条件，并将它们用于开发一种旨在近似解决跳跃约束 DSTP 的方法。该方法还可用于改进由其他启发式算法产生的、或者作为精确算法的一部分的近似解。我么将介绍基于该方法产生的特定标签校正类型算法，并报告它们在一组测试问题上表现性能的初步结果。测试实例源自用于多目标监视的无人驾驶飞行器的 3D 放置。它们的特征在于相对较少数量的终端节点和非常多的节点以及大量的弧（超过十亿）。

嘉宾介绍：

Oleg Burdakov 自 1999 年来于瑞士林雪平大学任职教授，同时兼任美国盖恩斯维尔佛罗里达大学应用优化中心的附属学院的教职员。1998 年，他担任巴西坎皮纳斯大学的客座教授。他还曾担任俄罗斯科学院计算中心 (1980–1994) 和法国图卢兹 CERFACS (1995–1997) 的研究科学家。他分别于 1977 年和 1980 年获得了莫斯科物理科学与技术学院的硕士学位与博士学位。他是期刊《优化方法和软件》的主编。2014 年他参加由美国离散数学和理论计算机学中心

(DIMACS) 组织的解决最小生成树问题国际挑战赛并获奖。2017 年，他获得了中国科学院院长国际奖学金计划下的访问科学家奖。他的研究方向包括：解决优化问题和非线性方程系统的数值方法，尤其是牛顿型、稳定割线和插值方法、全球化策略、基数约束优化、反问题、多线性最小二乘、非光滑优化和方程、线性和非线性鞍问题和单调方程、单调回归、数据拟合和插值、跳限制最短路径和最小生成树问题。



十四、为听力设备解决鸡尾酒会问题

时间：2018/11/20 星期二 上午 11:00–12:00

地点：道远楼会议室

嘉宾：张涛博士，Starkey Hearing Technologies

讲座摘要：

即使经过数十年的研究，鸡尾酒会问题仍然是听力设备面临的最具挑战性的问题之一。在本次演讲中，我们将回顾基于深度学习、针对单麦克风的语音增强技术。我们将介绍并讨论我们对多麦克风语音增强技术的研究。此外，我们还将展示我们关于多模式语音增强技术的最新研究。在实践中部署这些算法的挑战也将会被讨论。最后，我们将分享我们对未来研究方向的看法，尤其是在使用大数据和机器学习的个性化和定制方面。

嘉宾介绍：

张涛于 1986 年获得了南京大学物理学学士学位，1989 年于北京大学获得了电气工程硕士学位，1995 年又于美国俄亥俄州立大学获得了语言和听力科学博士学位。他于 2001 年加入 Starkey Laboratories 的高级研究部并担任高级研究科学家。自 2004 年至 2008 年他管理 DSP 部门，从 2008 年至 2014 年又管理信号处理研究部门。自 2014 年起，他担任 Starkey Hearing Technologies 信号处理研究部主任，该部门是提供创新型听力技术的全球领导者。他曾获得许多著名奖项，包括年度发明家奖、雷尼尔山最佳研究团队奖、最有价值创意奖、杰出技术领导奖、和 Starkey 工程服务奖。

他是 IEEE、信号处理学会、以及医学和生物学工程学会的高级成员，同时也是 IEEE AASP 技术委员会和工业关系委员会的成员。他是 IEEE SPS 杰出工业演讲者、IEEE SPS 美国工业大使（1–6 区）、和 IEEE Twin-cities 信号处理与通信分会主席。

他目前的研究兴趣包括音频、声学、语音和多模态信号处理以及用于听力增强的机器学习、健康和健康监测、心理声学、房间和耳道声学、超低功耗实时嵌入式系统设计、以及设备电话云生态系统设计。他撰写及合著的演讲和出版物多达 100 余篇、获得了 20 多项批准专利、同时还有 30 多项专利正在申请。



十五、解构区块链以接近物理极限

时间：2018/11/29 星期四 上午 10:00–11:00

地点：知新楼 110

嘉宾：David Tse 教授，斯坦福大学

讲座摘要：

区块链的概念是由 Satoshi Nakamoto 提出的，用于维护电子支付系统：比特币的分布式账本。除了安全性之外，区块链协议还具有交易吞吐量、确认延迟性、和确认可靠性等重要性能指标。这些指标受到两个基础物理网络的限制：通信容量和光速传播延迟。现有系统的运作还远远达不到这些物理的极限。在这项工作中，我们介绍了一种新的区块链协议 Prism，它可以证明：
1) 防御高达 50% 的对抗哈希能力；2) 高达网络容量 C 的最大吞吐量；3) 与传播延迟 D 成比例的诚实交易的确认等待时间，确认错误概率是带宽延迟乘积 CD 的指数级小；4) 所有交易的最终总排序。我们设计此协议的方法是基于将区块链解构为其基本功能，并系统地扩展这些功能以接近其物理极限。

此项工作为与 Vivek Bagaria、Sreeram Kannan、Giulia Fanti 和 Pramod Viswanath 的合作研究。详细论文请参考 <https://arxiv.org/abs/1810.08092>

嘉宾介绍：

David Tse 于 1989 年获得滑铁卢大学系统设计工程专业学士学位，又分别于 1991 和 1994 年获得美国麻省理工学院的电气工程硕士和博士学位。1995 年到 2014 年，他在加州大学伯克利分校任教，目前担任斯坦福大学的 Thomas Kailath 和 Guanghan Xu 教授。他于 2017 年获得 Claude E. Shannon 奖，并于 2018 年当选为美国国家工程院院士。此前，他于 1998 年获得 NSF CAREER 奖，2000 年获得 INFORMS 应用概率学会颁发的 Erlang 奖和 Frederick Emmons 奖，2009 年获得美国工程教育学会颁发的特曼奖。他与 Pramod Viswanath 合著了无线通信基础，被全球 60 多个机构使用。他获得了 IEEE 信息理论、通信和信号处理协会颁发的最佳论文奖，并且是所有第三代和第四代蜂窝系统中使用的比例公平调度算法的发明者。



编辑：陈怿、王文琪

指导：罗智泉

深圳市大数据研究院
中国广东省深圳市龙岗区龙翔大道 2001 号道远楼 225 室
邮编：518172
电话：(86) 755-84273615
<http://sribd.cn/index.php/cn/>