



# Cloud Computing and Big Data

## 云计算与大数据 研究室

孔超

kongchao@ahpu.edu.cn

# 提纲



研究方向



研究团队

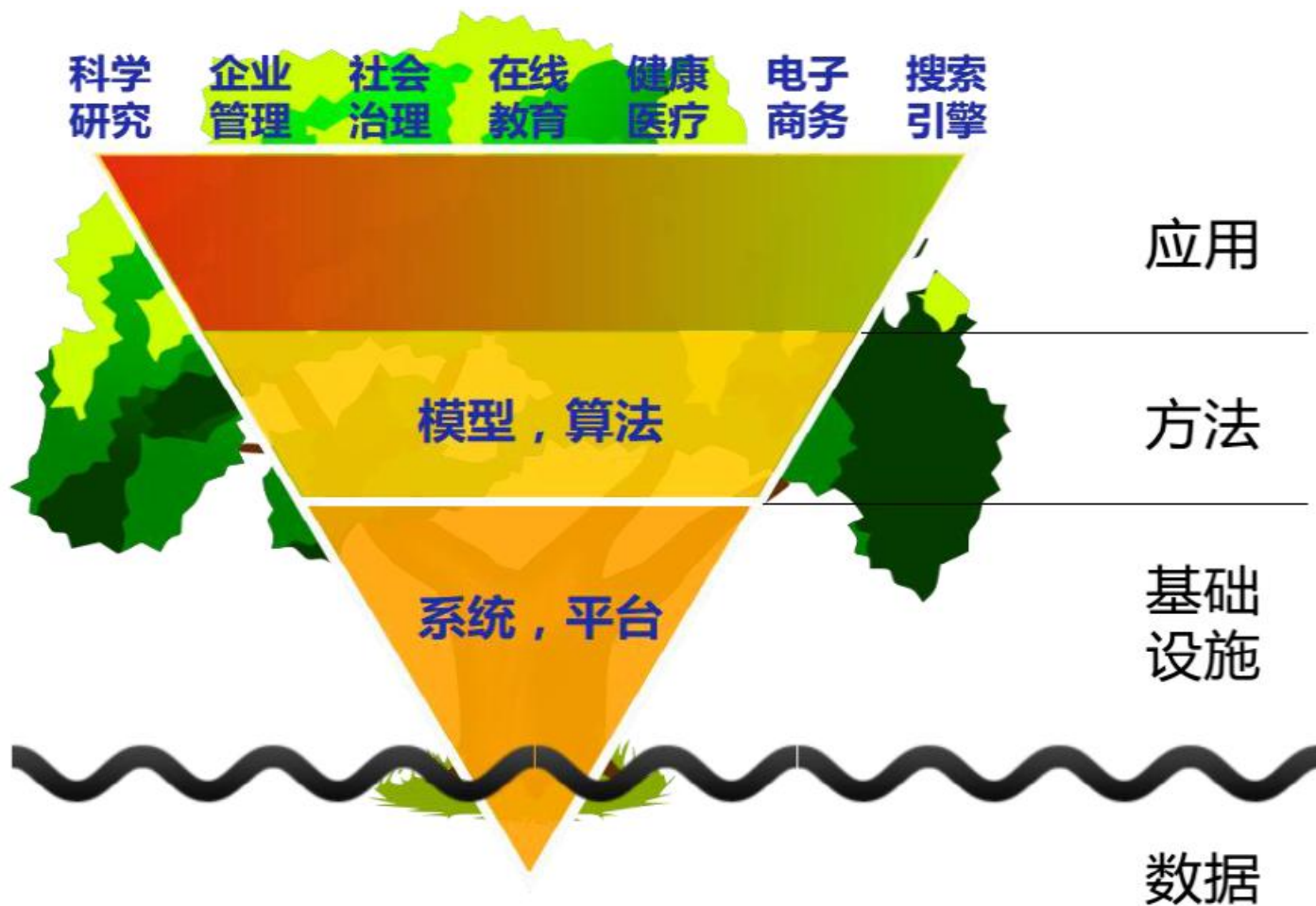


合作交流



人才培养

# 大数据全景图



# 研究方向

- 面向互联网开放数据的实体匹配
  - 支撑项目：
    - 国家自然科学基金青年项目
- 图神经网络算法
  - 支撑项目：
    - 安徽工程大学引进人才科研启动基金(已结题)
- 动态多智能体博弈策略
  - 支撑项目：
    - 安徽工程大学国家自然科学基金预研项目(已结题)

# 部分研究成果

## Selected Publications<sup>[DBLP]</sup>

### >> 2021

- >> Privacy Attack in Multiplex Network Embedding Model. Chao Kong, Baoxiang Chen, Liping Zhang. Journal of Combinatorial Optimization (JOCO, Research Article)
- >> Adaptive High-order Relations Modeling for Social Recommendation. Chao Kong, Shaoying Li et al. (Apweb-Waim 2021, Full Paper)
- >> Spatiotemporal Data Forecasting for Biological Invasion Detection. Shaoying Li, Chao Kong et al. (NDBC 2021, Full Paper)

### >> 2020

- >> DEM: Deep Entity Matching across Heterogeneous Information Networks. Chao Kong, Baoxiang Chen, Liping Zhang. Journal of Computer Science and Technology (JCST, Research Article)
- >> An Advanced Q-learning Model for Multi-agent Negotiation in Real-time Bidding. Chao Kong, Baoxiang Chen et al. (WISA 2020, Full Paper)
- >> GNE: Generic Heterogeneous Information Network Embedding. Chao Kong, Baoxiang Chen et al. (WISA 2020, Short Paper)
- >> D2NE: Dual Attention Network Embedding for Dynamic Link Prediction. Chao Kong, Baoxiang Chen et al. (ADMA 2020, Short Paper)
- >> Privacy Attack and Defense for Network Embedding. Chao Kong, Baoxiang Chen et al. (CSoNet 2020, Full Paper)
- >> Empowering Knowledge Graph with Neural Networks for Question Answering System. Chao Kong, Shaoying Li et al. (BigDIA 2020, Full Paper)

### >> 2019

- >> EnAli: entity alignment across multiple heterogeneous data sources. Chao Kong, Ming Gao, Chen Xu, Yunbin Fu, Weining Qian, Aoying Zhou. Frontiers of Computer Science (FCS, Research Article)
- >> Drug Abuse Detection via Broad Learning. Chao Kong, Jianye Liu, Hao Li, Ying Liu, Haibei Zhu, Tao Liu. Web Information Systems and Applications - 16th International Conference (WISA 2019, Short Paper)
- >> Link Prediction on Dynamic Heterogeneous Information Networks. Chao Kong, Hao Li, Liping Zhang, Haibei Zhu, Tao Liu. Computational Data and Social Networks - 8th International Conference (CSoNet 2019, Full Paper)

More details  
:  
<https://kongchao315.github.io/>

# 提纲



研究方向



研究团队



合作交流



人才培养

# 研究团队

- 李绍莹（2020级硕士）
- 陈嘉慧（2018级本科）
- 高兴（2018级本科）
- 陈云洋（2018级本科）
- 崔柏婷（2019级本科）
- 江涛（2019级本科）
- 朱旭（2019级本科）
- 陶文豪（2019级本科）
- 李梦飞（2019级本科）
- 王舒婷（2019级本科）



# 实验室成员去向（持续更新）

- 陈宝祥（2018级硕士，海康威视）
- 陈艺凡（2017级本科，军事科学院军事医学研究院）
- 韦振昆（2017级本科，网易互娱）
- 孙逸帆（2017级本科，上海理工大学硕士在读）
- 朱海蓓（2016级本科，字节跳动、腾讯实习，GeorgiaTech硕士在读）
- 刘建业（2016级本科，西北大学硕士在读）
- 刘颖（2016级本科，国防科大硕士在读）



# 提纲



研究方向



研究团队



合作交流



人才培养

# 国内外学术交流

- 2019年5月，赴台湾宜兰大学参加ICCE-TW 2019国际学术会议，并做现场汇报



# 国内外学术交流

- 2019年8月，赴日本福岡九州工业大学参加SCDS 2019国际学术会议，并做现场汇报



# 国内外学术交流

- 2019年11月，赴越南胡志明文朗大学参加CSoNet 2019国际学术会议，并做现场汇报



# 提纲



研究方向



研究团队



合作交流

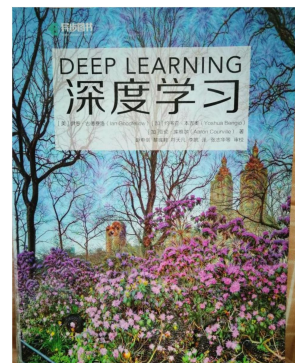


人才培养

# 期望

## ■ 对未来学生的期望：

- **编程基础**：扎实的Python/JAVA等编程基础，具有编程的sense，即利用计算机程序解决问题的能力；熟练使用GnuPlot、Tableau、MATLAB等可视化工具；熟练在集群环境下跑实验；熟练使用Latex进行科研论文写作
- **理论基础**：理解机器学习、深度学习方法和框架。推荐阅读《机器学习》和《深度学习》；



- **兴趣志向**：对社交网络分析和海量数据挖掘技术感兴趣，而不仅限于使用现成框架，有志于从事上游算法设计，勇于挑战自我。

# 个性化培养

## ■ 学校、学院培养流程：

- 课程学习
- 科研实践
- 资格考试、中期考核
- 毕业论文、答辩

## ■ 实验室（个性化）培养：

- 熟悉主流的机器学习和深度学习框架，如MF、GCN、Hadoop、TensorFlow、Pytorch等
- 结合项目需求和个人兴趣确定研究主题/项目任务
- Seminar：实验室集体研讨
  - 论文阅读
  - 研究工作进展
- Meeting：单独面谈30分钟–1小时
  - 研究/学习中的问题



# 学习过程管理

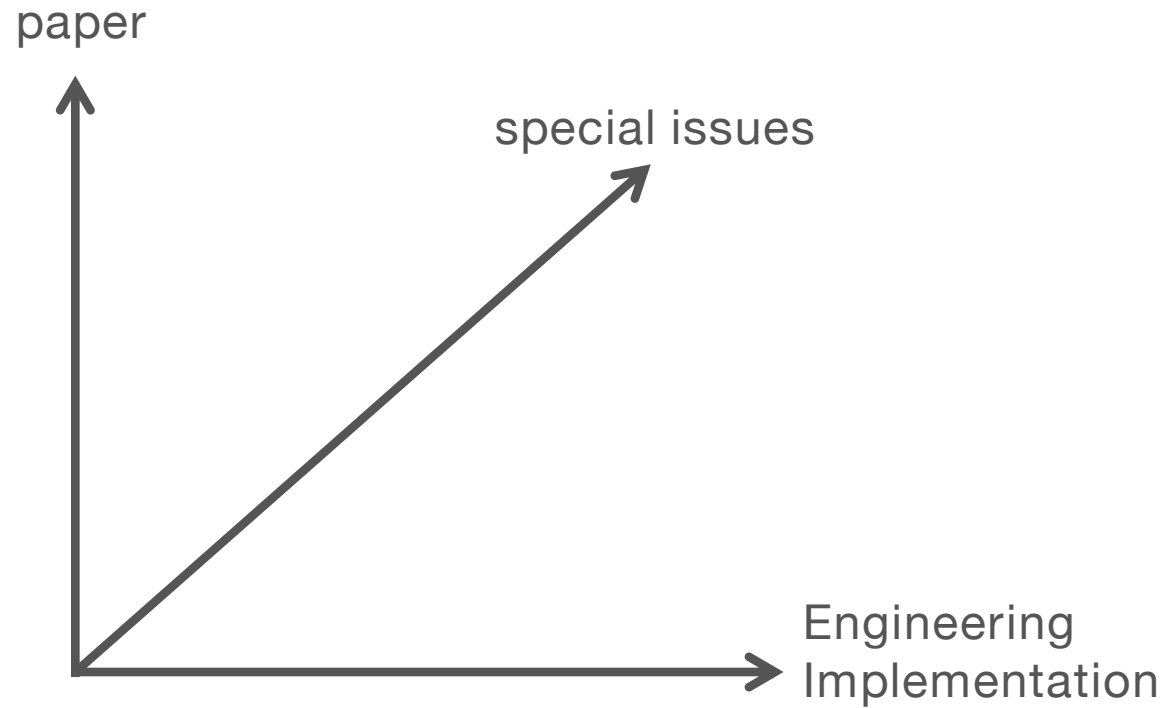
## ■ 严格的“711”工作时间及周报制度：

Kong's Seminar 工作周报（第5周）

序号	姓名	本周工作内容	状态	下周工作计划
1	陈宝祥			
2	李邵莹	1.阅读论文。 2.统计学习方法。 3.看论文代码。	1.论文算法部分，还未理解。 2.代码还要继续理解。	1.形成整体框架，写paper。 2.继续阅读相关论文。
3	陈嘉慧	1.项目申请+多智能体博弈类论文和知识学习 2.统计学习方法第二章	本周学习时间不多，统计学习方法需要加快进度	1.科研立项周三前完成！ 2.统计学习方法三四章 3.博弈大赛周四开始准备
4	陈艺凡			
5	陈云洋	1.学习《统计学习方法》代码 辅助陈宝祥运行代码，其中包括学习dockerfile编写	2. dockerfile由于没有一些固定模板编写出错，考虑转换国内镜像源来提升docker生成速度	继续辅助陈宝祥进行代码在集群上的运行和修改
6	高兴	1. linux 系统编程 2. 学习数学分析新讲	1.学习数学分析新讲 2.配置实验室服务器环境	1.搭建实验室服务器GPU环境 2.继续学习数学分析
7	陶文豪	1.统计学习方法代码+课件 博弈大赛报名及相关资料学习	2.统计学习方法部分代码还在理解当中 2.报名已完成，下载了源代码和相关资料	1.统计学习方法 2.阅读材料和源代码
8	王舒婷	1.博弈大赛报名及相关学习 2.算法相关学习	运行博弈大赛相关代码并找相关论文	1.博弈大赛相关算法学习并进行代码优化 2.程序设计算法学习
9	甘宇恒	统计学方法PPT的阅读以及对博弈大赛相关资料的查找。	了解了博弈大赛的整个过程以及确定了棋种和相关棋类的下法	深入了解与爱恩斯坦棋的相关算法。
10	朱旭	1、运行openne代码 2、配置linuxpython环境 3、看统计学习代码	1、linux的python环境配置出错 2、openne代码跑通 3、统计学习到第三章	1、阅读完计算机博弈大赛代码 2、继续看统计学习
11	崔佰婷	1.复习对应数据结构的c语言知识 2.英语文章阅读 3.博弈大赛相关资料阅读	1.英语六级熟悉过程 2.熟悉博弈大赛 3.查找linux系统安装方法	1.英语六级学习 2.学习统计学方法 3.博弈大赛代码 4.安装linux系统 5.科研立项
12	李梦飞	1、结合代码学习《概率导论》； 2、理解计算机博弈大赛代码； 3、学习C++；	1、对不围棋代码不够熟悉 2、时间利用不够充分	1、找出不围棋程序代码的优缺点，寻找更理想的数学模型； 2、辅助崔佰婷完成科研立项； 3、继续结合代码学习概率导论
13	江涛	1.学习统计学方法 2.查找一些关于计算机博	统计学还在继续学习当中，代码也在理解当中	继续学习统计学方法，继续优化代码



# 期末工作总结



# 期末工作总结

## 2019-2020 学年第二学期期末工作总结

姓名:                  入学时间:                  总结时间:

说明: 请按照以下提纲**简明扼要**回答下列问题, 对本学期工作进行总结。  
注意: 请勿删除以下提纲的内容, 包括括号中的文字。

1. 过去一学期的时间投入情况 (平均每周在课程学习、科研中的时间分配情况, 每周投入学习的情况, 在家是否能够按照工作日规定时间工作, 以及非工作日的学习情况等)
2. 过去一学期取得的工作成绩 (包括课程学习、完成的实验、项目系统实现、发表或提交的论文等)
  - (1) xxx
  - (2) xxx
  - (3) xxx
3. 目标完成情况 (针对上一学期末或学期初拟定的目标, 在本学期的完成情况如何? 例如, 达到、超过或未能达到的目标等)
  - (1) Xx
  - (2) Xx
4. 自我综合评价 (从学习态度、工作成绩、能力提升、个人感悟等方面对本学期工作进行综合评价, 是否满意? 如不满意, 打算如何改进)
5. 下一个学期的研究工作计划与目标 (包括拟研究的问题、项目系统实现、实验验证、发表的论文等)
  - (1) 拟研究 xx
  - (2) 拟在 x 月前实现 xx

(3) 拟在 x 月完成论文 xx, 投稿至 xx

6. 下一阶段的自我规划 (包括拟毕业时间、如临近毕业时间保障按期毕业的措施、毕业后的就业倾向、学术型硕士进行硕博连读的意向等)
7. 其它希望说明的问题 (例如, 承担了某项工作占用过多时间导致未能达到某一目标, 希望得到的某些帮助等)

# 愿景

- 经过3年实验室学习后的愿景：
  - 具有扎实的机器学习**算法设计与实现的理论基础**，能够熟练阅读英文论文、准确理解并综述前沿技术；
  - 吃透若干开源框架的源码，具备**修改源码**或依据基础理论**开发原型系统的能力**；
  - **硕士生有高水平论文**（CCF-C类及以上）的发表和良好的presentation技能，以及带领小团队的经验；
  - 到**大型企业或海外高水平实验室工作**，从事机器学习、大数据等相关的研究与开发。



# 欢迎感兴趣的同学加盟

## 谢谢！



将简历、成绩单等材料发送至  
孔超 [kongchao@ahpu.edu.cn](mailto:kongchao@ahpu.edu.cn)