更新于 2019 年 5 月 26 日

马凌霄

xysmlx@pku.edu.cn (邮件)
xysmlx@gmail.com (邮件)
博士四年级 · 北京大学分布式系统组　　北京市海淀区颐和园路 5 号北京大学 (地址)

研究方向

机器学习系统， 大规模图计算， GPU 并行计算： 博士研究方向为利用现代高性能计算设备（例如：GPU）为大规模数据分析场景构建高性能并行计算系统，例如：深度学习、机器学习、大规模图计算

教育经历

- **北京大学**　信息科学技术学院 · 计算机系统结构 · 导师：代亚非 教授　2015.09 – 今　理学博士
- **北京师范大学**　信息科学与技术学院 · 计算机科学与技术　2011.09 – 2015.06　理学学士

实习经历

- **微软亚洲研究院（北京）– 系统组**　全职研究实习生　导师：薛继龙、伍鸣　2017.11 – 今

项目经历

- **NeuGraph (NGra)**　图神经网络计算系统 · 发表于 ***USENIX ATC'19*** · 第一作者　2017.12 – 今
  - **简介**　NeuGraph (NGra) 是面向大规模图神经网络 (Graph Neural Network, GNN) 的计算框架，它构建于现有基于数据流图的深度学习系统之上，解决了 GNN 的图结构带来的 scalability 和 efficiency 的问题
  - **技术**　结合顶点编程和数据流编程的 SAGA-NN 编程抽象; SAGA-NN 模型到数据流图的翻译; 单 GPU/多 GPU 的流式执行机制; GPU 的高性能图传播操作和内核融合机制
  - **成果**　USENIX ATC'19 论文 [1]; 被新智元、机器之心等人工智能媒体报导; 项目在 2018 微软人工智能大会 (Microsoft AI Innovate) 和 2018 中国计算机大会 (CNCC) 进行了展示
- **Garaph**　*CPU-GPU* 异构图计算系统 · 发表于 ***USENIX ATC'17*** · 第一作者　2016.03 – 2017.03
  - **简介**　Garaph 是一个面向 CPU/GPU 异构环境的图计算系统，支持 CPU 和多 GPU 协同对大规模图数据进行高效处理
  - **技术**　解决 GPU 线程冲突的自动副本机制; 解决 CPU 负载均衡和线程冲突的副本机制; CPU/GPU 之间的动态任务调度
  - **成果**　USENIX ATC'17 论文 [2]; 专利 [1]; NASAC'17 受邀报告
- **DL Compiler**　深度学习 *Inference* 编译器 · 在投论文 · 第一作者　2018.10 – 今
  - **简介**　现有深度学习框架（如 TensorFlow）在 inference 场景中会导致很高且不稳定的 latency。我们针对该问题，设计了一个编译器
- **SeerNet**　稀疏卷积计算 · 发表于 ***CVPR'19*** · 第二作者　2018.10 – 2018.11
  - **简介**　SeerNet 关注卷积神经网络中输出特征图的稀疏性，例如，经过 ReLU 或 Max-pooling 层后，卷积层的大部分输出被置为零或丢弃。如果跳过这部分对应的先导卷积计算，则可以大大减少卷积层的计算量
  - **技术**　利用低精度模型以极低的代价预测输出特征的稀疏性，通过稀疏的原精度计算加速原神经网络计算；可以直接应用于预训练好的模型中而无需对原始模型做任何修改或重训练
  - **成果**　CVPR'19 论文 [3];

发表论文

[1] NeuGraph: Parallel Deep Neural Network Computation on Large Graphs
Lingxiao Ma, Zhi Yang, Youshan Miao, Jilong Xue, Ming Wu, Lidong Zhou, Yafei Dai
2019 USENIX Annual Technical Conference (**USENIX ATC'19**) (CCF A)　(北大第 6 篇，组内第 2 篇)

[2] Garaph: Efficient GPU-accelerated Graph Processing on a Single Machine with Balanced Replication
Lingxiao Ma, Zhi Yang, Han Chen, Jilong Xue, Yafei Dai
2017 USENIX Annual Technical Conference (**USENIX ATC'17**) (CCF A) (北大第 4 篇，组内首篇)

[3] SeerNet: Predicting Convolutional Neural Network Feature-Map Sparsity through Low-Bit Quantization
Shijie Cao, **Lingxiao Ma**, Wencong Xiao, Chen Zhang, Yunxin Liu, Lintao Zhang, Lanshun Nie, Zhi Yang
30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR'19**) (CCF A)

[4] CuWide: Towards Efficient GPU Training for Large-Scale Wide Model
Xupeng Miao, **Lingxiao Ma**, Yingxia Shao, Bin Cui, Zhi Yang, Jiawei Jiang, Lele Yu
Submitted to **VLDB'20** (single-blind peer-review, under one shot revision) (CCF A)

申请专利

[1] 一种基于自动选择副本因子模型的图计算方法. 201710533444.5

主要奖励

- **北京大学优秀科研奖** 2018.12, 2017.12
- **智唐奖学金** 2017.12
- **北京大学博士研究生校长奖学金** 2017.06
- **秒针奖学金** 2016.12
- **北京师范大学优秀毕业生** 2015.05
- **北京师范大学第 12 届励耘优秀本科生奖学金** 全校 *6* 人 2014.12
- **国家奖学金** 入选教育部主编《希望—*2014* 年国家奖学金获奖学生风采录》，全校唯一 2014.10

竞赛获奖

- **第 39 届 ACM/ICPC 国际大学生程序设计竞赛亚洲区域赛鞍山站** 银牌 2014.10
- **美国数学建模竞赛** 一等奖 *(Meritorious Winner)* 2014.02
- **全国大学生数学建模竞赛北京赛区** 一等奖 2013.10

专业能力

- **编程语言**: C, C++, CUDA, Python, LaTeX, Markdown, Java, Shell
- **熟悉领域**: 面向 GPU 和 Multi-Core 环境的并行编程, 大规模图计算, 机器学习, 分布式系统, 数据结构与算法

其他经历

- **北京大学网络与信息系统研究所** 系统管理员 2016.12 – 今
- **北京大学信息科学技术学院本科生专业必修课程《计算概论 A》** 助教 2016.09 – 2017.01

更多个人信息请访问: https://xysmlx.github.io/