李沐晟 Musheng Li

清华大学电子工程系·直博生 | 导师：谷源涛教授

lms21@mails.tsinghua.edu.cn | GitHub: Liusuthu | Google Scholar | 北京，中国

教育背景

清华大学·电子工程系　博士研究生（直博） 2025.09 – 至今

研究方向：大语言模型、时间序列建模、多模态理解与推理；导师：谷源涛教授

清华大学·电子工程系　工学学士 2021.09 – 2025.07

本科毕业设计：基于多模态融合的时间序列文本问答大模型（成绩 A）

研究方向

大语言模型（LLMs）、时间序列理解与推理、多模态表示学习、测试时计算扩展（Test-Time Scaling）、循环深度与权重共享 Transformer（Looped Transformers）。

学术论文

[1] **Musheng Li**, Jiawei Zhang, Ziyuan Liu, Yuantao Gu.
Rectified Looped Transformer.
Manuscript, 2026. （第一作者；arXiv 提交处理中，尚未公开）

[2] **Musheng Li**, Ziying Zhang, Cheng Jin, Yuantao Gu.
Continuity and Ordinality Matter: Constraining Time Series Tokens for Effective Time Series Analysis with Large Language Models.
arXiv preprint, arXiv:2605.28866, 2026. （第一作者）

[3] Yao Yin, Zhenyu Xiao, **Musheng Li**, Yiwen Liu, Sutong Nan, Yiting He, Ruiqi Wang, Zhenwei Zhang, Qingmin Liao, Yuantao Gu.
MMTS-BENCH: A Comprehensive Benchmark for Time Series Understanding and Reasoning.
arXiv preprint, arXiv:2602.08588, 2026. （第三作者）

科研项目

COM：面向时间序列大模型的连续性与序关系建模 2026

- 针对离散时间序列 Token 忽视数值连续性和顺序结构的问题，研究具有几何先验的嵌入表示。
- 提出结合**硬初始化约束**与**软正则化**的框架，在 TSQA 等任务上提升时间序列理解能力及泛化表现。

MMTS-BENCH：多模态时间序列理解与推理基准 2025 – 2026

- 参与构建由 **2,424 个问答样本**组成的评测基准，覆盖 Base、InWild、Match、Align 四类子集。
- 从结构感知、特征分析、时间推理、序列匹配、跨模态对齐五个层级评估通用与时间序列专用大模型。

AdaTTS：大语言模型自适应测试时计算策略 2026

- 建立统一计算预算下 Self-Consistency、Best-of-N、Beam Search 的理论与实验比较框架。
- 研究按任务难度自适应分配推理策略；实验中相较最优固定策略平均提高 **2.38 个百分点**。

Rectified Looped Transformer（RLT）：稳定的循环反馈机制 2026 – 至今

- 提出**流式插值反馈**，将循环块输出视为终点预测，通过剩余步数调度平滑更新隐状态；无需额外可训练参数或循环块前向计算。
- 从理论上分析状态有界性与扰动放大上界；实验验证相较直接反馈基线具有更低困惑度、更好的噪声鲁棒性及推理深度泛化能力。

其他科研与实践经历

南加州大学（USC）·暑期访问研究学生 2024.07 – 2024.08

指导教师：Prof. Meisam Razaviyayn；清华–南加大暑期学术交流项目

- 研究差分隐私（Differential Privacy）机器学习，比较隐私保护与非隐私训练在图像分类任务中的表现。

多模态抑郁症检测系统·清华大学语音与音频技术实验室 本科期间

- 参与多模态特征融合与抑郁症检测系统设计；项目获**清华大学第 42 届“挑战杯”校级二等奖**。

荣誉与学生工作

- **2026.05**：清华大学四星级紫荆志愿者。
- **本科期间**：清华大学第 42 届“挑战杯”校级二等奖（团队项目）。
- **学生组织**：清华大学电子工程系志愿服务学生骨干（曾任志愿部书记/负责人）；学生电子音乐协会副会长。

专业技能

- **编程与工具**：Python、PyTorch、C/C++、MATLAB、Git、Linux、LaTeX。
- **大模型研发**：Hugging Face Transformers、SFT、数据合成与评测、DeepSpeed 分布式训练、GPU 集群实验。
- **其他**：英文科研写作与学术交流；兴趣包括乒乓球、游泳、电子音乐与摄影。