

宋学志

手机：13127702205
邮箱：sxz4se@outlook.com



个人简介

我当前任职于字节跳动质量智能部门，担任大模型算法工程师。我的核心工作是利用程序分析与大语言模型技术研发关键算法，以支撑服务端系统实现从自动化测试、缺陷定位到智能修复的闭环解决方案。我的研究兴趣聚焦于 AI4SE，例如代码生成、缺陷分析与自动修复等方向。

教育背景

复旦大学，计算机科学技术学院，计算机软件与理论，博士 2021.09 - 2025.06

- 导师：彭鑫教授、吴毅坚副教授、陈碧欢副教授

东华大学，计算机科学与技术学院，计算机技术，硕士 2018.09 - 2021.06

- 新加坡国立大学 Research Intern，由林云副教授指导

青岛大学，计算机科学与技术学院，软件工程，学士 2013.09 - 2017.06

工作经历

字节跳动—质量 Intelligence 实验室，大模型算法工程师 2025.08 – 至今

服务端测试用例生成平台 TestGPT

- 项目简介：TestGPT 是字节跳动内部面向服务端测试的全链路解决方案，涵盖测试场景挖掘、用例自动生成、结果分析、缺陷定位与智能修复多个环节。

- 个人职责：

- 调研并评估 Go 语言程序分析工具链，设计与实现支撑大模型算法的 Go 语言程序分析方法；
- 调研前沿 AI 缺陷修复技术，设计适用于服务端测试场景的缺陷修复算法，在标准评测集上分析效果与瓶颈，提出优化策略；
- 调研基于大语言模型的 Mock 数据生成技术，设计可用于服务端测试的 Mock 生成算法方案。

- 项目成果：程序分析模块成为部门代码大模型的核心工具与上下文支撑组件；缺陷修复成功率提升至 70%。

科研项目

高质量缺陷数据集/评测集的自动化构建，项目负责人 2020.01 – 2022.05

- 项目描述：高质量缺陷数据集可用于自动缺陷检测、定位与修复的效果评测，也可在累积到一定规模后用于大语言模型（LLMs）数据增强，从而提升相关算法性能。本项目结合测试迁移、代码演化分析与深度学习技术，实现缺陷数据集的自动化高质量构建。

- 个人职责：

- 设计半自动化缺陷收集与验证方法论；
- 开发缺陷案例的自动收集、验证与分类算法。

- 项目成果：

- 收集缺陷样本 3.3 万余个，其中包含 6000 个回归缺陷，构建当前最大且具备测试用例的缺陷库；
- 相关论文被 ISSTA'22 (CCF-A)、ASE'23 (CCF-A) 及 FSE'22 demo (CCF-A) 录用；
- 获中国软件原型赛一等奖、复旦大学原型赛二等奖；
- 工具在汇丰银行成功落地应用；
- 半自动化构建的深度学习缺陷库 gDefect4DL 被 ICSE'22 demo (CCF-A) 录用。

缺陷自动化检测、定位与修复，项目负责人 2022.05 – 2025.06

- 项目描述：自动缺陷检测、定位与修复是软件工程领域的研究热点之一，本项目旨在解决这些任务的关键技术问题（特别是基于 LLMs 和 Agent 的技术），提升任务性能。

- 个人职责：

- 开发并优化早期传统缺陷定位（增量调试、基于深度学习的方法）算法；

2. 设计与实现基于 Agent 架构的缺陷定位与检测算法,;
3. 构建缺陷知识库，并设计基于 RAG 的缺陷修复算法优化方案。

- **项目成果：**

1. 基于 Agent 的缺陷检测与定位方法分别被 ICSE'26 (CCF-A) 和《软件学报》'26 (CCF-A) 录用;
2. 传统缺陷定位方法被 ISSTA'24 (CCF-A) 和 ICSME'23 (CCF-B) 录用。

论文发表

当前发表会议/期刊论文 9 篇，包括 CCF-A 类长文 5 篇（第一作者 4 篇）、工具短文 2 篇，CCF-B 类长文两篇（一作者 1 篇）。

Characterizing Regression Bug Inducing Changes and Improving LLM Based Regression Bug Detection, ICSE 2026 (CCF-A)

Xuezhi Song, Yijian Wu, Zhengjie Lu, Bihuan Chen, Shuning Liu, Xin Peng

Defect Isolation via Semantic Analysis of Code Changes, Journal of Software (China) 2025 (CCF-A)

Shuning Liu, Yijian Wu, Xuezhi Song, Bihuan Chen, Xin Peng

C2D2: Extracting Critical Changes for Real-World Bugs with Dependency-Sensitive Delta Debugging, ISSTA 2024 (CCF-A)

Xuezhi Song, Yijian Wu, Shuning Liu, Bihuan Chen, Yun Lin, Xin Peng

BugMiner: Automating Precise Bug Dataset Construction by Code Evolution History Mining, ASE 2023 (CCF-A)

Xuezhi Song, Yijian Wu, Junming Cao, Bihuan Chen, Yun Lin, Zhengjie Lu, Dingji Wang, Xin Peng

RegMiner: Towards Constructing Ultra-Large Regression Dataset from Code Evolution History, ISSTA 2022 (CCF-A)

Xuezhi Song, Yun Lin, Siang Hwee Ng, Yijian Wu, Xin Peng, Jin Song Dong, Hong Mei

An Empirical Study on Fault Diagnosis in Robotic Systems, ICSME 2023 (CCF-B)

Xuezhi Song, Yi Li, Zhen Dong, Shuning Liu, Junming Cao, Xin Peng

RegMiner: Mining Replicable Regression Dataset from Code Repositories, FSE (Demonstrations Track) 2022 (CCF-A)

Xuezhi Song, Yun Lin, Yijian Wu, Yifan Zhang, Siang Hwee Ng, Xin Peng, Jin Song Dong, Hong Mei

gDefect4DL: A Dataset of General Real-World Deep Learning Program Defects, ICSE (Demonstrations Track) 2022 (CCF-A)

Yunkai Liang, Yun Lin, Xuezhi Song, Jun Sun, Zhiyong Feng, Jin Song Dong

Characterizing the Complexity and Its Impact on Testing in ML-Enabled Systems - A Case Study on Rasa, ICSME 2023 (CCF-B)

Junming Cao, Bihuan Chen, Longjie Hu, Jie Gao, Kaifeng Huang, Xuezhi Song, Xin Peng

荣誉奖项

奖学金

冠名奖学金 x1, 华为智能基座奖学金 x1, 一等奖学金 x5, 二等奖学金 x1

竞赛

中国开源软件大会自由赛道一等奖 (2022)、优胜奖 (2023)

中国开源软件大会命题赛道二等奖 (鲲鹏软件迁移, 2022)

华为智能基座未来之星 (2022)

技能 / 证书及其他

编程语言：Java / Python / Kotlin / Go

程序分析工具：JavaParser / Spoon / JDT / ASM / LLVM

语言能力：CET-6 (英语六级)