Xia Li

Research Assistant
Department of Computer Science
Erik Jonsson School of Engineering & Computer Science
The University of Texas at Dallas

+1 (469) 769-2402
Xia.Li3@utdallas.edu
lx0704.github.io/
McKinney, TX 75071, US

Research Interests

Software Engineering, in particular: software testing and debugging involving dynamic/static program analysis, machine learning, deep learning and big-code mining.

Education

2014 – 2020 (Expected) **Ph.D. in Computer Science**
The University of Texas at Dallas, Richardson, US
GPA: 3.67/4.0, Advisor: Lingming Zhang (lingming.zhang@utdallas.edu)

2012 – 2014 **M.S. in Information Technology and Management**
The University of Texas at Dallas, Richardson, US

2009 – 2012 **M.S. in Management Science and Engineering**
Shandong Jianzhu University, Jinan, China

2004 – 2008 **B.S. in Mathematics and Applied Mathematics**
Jiangxi University of Science and Technology, Ganzhou, China

Publications

- Yiling Lou, Ali Ghanbari, **Xia Li**, Lingming Zhang, Dan Hao, Lu Zhang. **Can Automated Program Repair Refine Fault Localization?** *arXiv preprint arXiv:1910.01270,* October 2019.
- **Xia Li**, Wei Li, Yuqun Zhang, and Lingming Zhang. **DeepFL: Integrating Multiple Fault Diagnosis Dimensions for Deep Fault Localization.** In *Proceedings of the 28th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2019)*, July 2019. ACM SIGSOFT Distinguished Paper Award
- Mengshi Zhang, Yaoxian Li, **Xia Li**, Lingchao Chen, Yuqun Zhang, Lingming Zhang, Sarfraz Khurshid. **An Empirical Study of Boosting Spectrum-based Fault Localization via PageRank.** *IEEE Transactions on Software Engineering (TSE)*, April 2019.
- **Xia Li** and Lingming Zhang. **Transforming Programs and Tests in Tandem for Fault Localization.** In *Proceedings of the ACM SIGPLAN Conference on Object-Oriented Programming System, Languages, and Applications (SPLASH/OOPSLA 2017)*, October 2017.
- Mengshi Zhang, **Xia Li**, Lingming Zhang and Sarfraz Khurshid. **Boosting Spectrum-based Fault Localization using PageRank.** In *Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2017)*, July 2017. Nomination for ACM SIGSOFT Distinguished Paper Award

- **Xia Li**, Jiajun Jiang, Yingfei Xiong and Lingming Zhang. **A Massive Study on API Misuses in the Wild and Its Implications.** In *Proceedings of the 29th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2020)*, July 2020. *Manuscript in preparation*.
- Junjie Chen, **Xia Li**, Lingming Zhang, Dan Hao, Lu Zhang and Bing Xie. **Boosting Automated Debugging Supports via Delta Debugging: An Extensive Study.** *Manuscript in preparation*.
- **Xia Li**, Wei Li, Yuqun Zhang, and Lingming Zhang. **An Empirical Study of Integrating Multiple Fault Diagnosis Dimensions for Boosting Fault Localization via Deep Learning.** *IEEE Transactions on Software Engineering (TSE). Manuscript in preparation*.

Research Experience

2018 – 2019 **Detecting GitHub Bugs via Big Code Mining and Static Analysis**.

- Mined millions of historical bug-fixing commits from GitHub and automatically extracted various bug-fixing patterns via static program analysis.
- Implemented a bug detection tool according to the patterns to detect bugs in the latest Apache projects.
- Detected 400+ bugs and 55 of them are confirmed and fixed by developers to date.

2017 – 2019 **Deep-Learning-Based Fault Localization**.

- Extracted suspiciousness-based features, fault-proneness-based features and textual-similarity-based features via dynamic analysis, static analysis and information retrieval.
- Implemented various Deep Learning techniques via TensorFlow such as Multiple Layer Perceptron (MLP), Recurrent Neural Networks (RNN) and a tailored hierarchical MLP for fault localization by combining these features.
- Ranked 213 bugs (out of 395 studied real bugs) within Top-1, the best result compared with other state-of-the-art techniques.
- Published one paper in ISSTA 2019.

2016 – 2017 **Localizing Bugs by Transforming Programs and Tests via Learning-to-Rank**.

- Transformed test cases to capture more detailed failure messages and assertion outcomes.
- Used LIBSVM and XGBoost to implement the Learning-to-Rank algorithm for localizing bugs by combining spectrum-based and mutation-based fault localization via various failure messages.
- Localized 142 bugs (out of 357 real bugs) within Top-1 by LIBSVM.
- Published one paper in OOPSLA 2017.

2016 – 2017 **Boosting Spectrum-Based Fault Localization via PageRank**.

- Collected spectrum information between tests and source code entities as well as the static call graph information among source code entities via bytecode instrumentation and analysis.
- Used PageRank to re-compute the spectrum information by considering the contributions of different tests.
- Found 104 bugs (out of 357 studied real bugs) within Top-1 by PageRank.
- Published one paper in ISSTA 2017.

Industry Experience

Summer 2018 **R&D Software Support Engineer Intern**, *FutureWei Technologies, Inc*, Plano, TX, US.

- Worked as an R&D intern to work on an automated program repair project of the company.
- Helped set up a state-of-the-art repair tool and mined bug-fixing patterns from GitHub for improving the tool.

Teaching Experience

Summer 2019 **Teaching Assistant**, Organization of Programming Languages (CS4337), The University of Texas at Dallas, Richardson, US

Spring 2016 **Teaching Assistant**, Big Data Management and Analytics (CS6350), The University of Texas at Dallas, Richardson, US

Expertise and Skills

Languages Python, Java, R

Systems Windows, Linux

Tools TensorFlow, PyTorch, Spark, Scikit-learn, Eclipse, JUnit, ASM Bytecode Manipulation Framework, Eclipse JDT, Git

Professional Service

2020 **Co-Reviewer**:ICSE

2019 **Co-Reviewer**:ICST,ISSTA,QRS,ICSME,ASE

2018 **Reviewer**:Journal of Systems and Software (JSS)
Co-Reviewer:QRS,COMPSAC,ASE,SPE

2017 **Co-Reviewer**:ICST,QRS,COMPSAC,ASE

References

Advisor **Dr. Lingming Zhang**, Assistant Professor
Department of Computer Science, The University of Texas at Dallas
lingming.zhang@utdallas.edu
ECSS 4.205, UT Dallas, Richardson, TX 75080, US

Co-Author **Dr. Sarfraz Khurshid**, Professor
Department of Electrical and Computer Engineering, The University of Texas at Austin
khurshid@ece.utexas.edu
EER 7.880, 1 University Station C0803, Austin, TX 78712, US

Co-Author **Dr. Yingfei Xiong**, Associate Professor
Department of Computer Science and Technology, Peking University
xiongyf@pku.edu.cn
Room 1431, Science Building 1, Peking University, Beijing, China

Co-Author **Dr. Yuqun Zhang**, Assistant Professor
Department of Computer Science and Engineering, Southern University of Science and Technology
zhangyq@sustech.edu.cn
Room 604, Building 10, Innovation Park, Southern University of Science and Technology, Shenzhen, China