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
shawn870704@gmail.com
www.xialigithub.com
McKinney, TX 75071, US

Research Interests

Software Testing and Analysis, in particular: Automated Debugging and Dynamic/Static Progarm Analysis via machine learning, deep learning and big code mining.

Education

9/2014 – 5/2020 (Expected) **Ph.D. in Computer Science**
The University of Texas at Dallas, Richardson, USA
GPA: 3.67/4.0, Advisor: Lingming Zhang (lingming.zhang@utdallas.edu)

9/2012 – 5/2014 **M.S. in Information Technology and Management**
The University of Texas at Dallas, Richardson, USA

9/2009 – 6/2012 **M.S. in Management Science and Engineering**
Shandong Jianzhu University, Jinan, China

9/2004 – 7/2008 **B.S. in Mathematics and Applied Mathematics**
Jiangxi University of Science and Technology, Ganzhou, China

Conference Publications

[C3] **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**

[C2] **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.

[C1] 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.

Journal Publications

[J1] 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.

Research Experience

5/2018 – 8/2019 **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 latest Apache projects.
- Reported detected bugs to GitHub and 55 of them are confirmed and fixed by developers to date.

7/2017 – 1/2019 **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.

1/2016 – 3/2017 **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 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.

Industry Experience

6/2018 – 8/2018 **R&D Software Support Engineer Intern**, *FutureWei Technologies, Inc*, Plano,TX.

- 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

5/2019 – 8/2019 **Teaching Assistant**, Organization of Programming Languages (CS4337), The University of Texas at Dallas, Richardson, USA

1/2016 – 5/2016 **Teaching Assistant**, Big Data Management and Analytics (CS6350), The University of Texas at Dallas, Richardson, USA

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