

Shaowei Wang

Assistant professor, University of Manitoba
66 Chancellors Cir, Winnipeg, MB R3T 2N2

shaowei@cs.umanitoba.ca +1 6624970084
<https://sites.google.com/view/shaoweiwang>

Highlights of Qualifications

- Publications: 32 publications in software engineering, including 15 journal papers (e.g., 6 TSE, 1 TOSEM, and 5 EMSE) and 18 conference papers (e.g., 4 ASE). My H-index is 20 (1,000+ citations). My recent paper in the 22nd IEEE International Conference on Program Comprehension (ICPC) was nominated as the best paper.
- Proven student mentorship record: 5 PhD students, 3 master students, and 1 undergraduate student; 16 papers with my mentees.
- Industry-relevant research resulted in tools and techniques for software artifacts search and reuse, which are used in industry [C17, C11, J2], including companies such as Microsoft and Huawei.
- Research interests: data analytics for software engineering, machine learning applications in software engineering, mining software repositories, automated debugging, software maintenance; focus on using program analysis and data mining techniques to improve software quality and development productivity.

Journal Publications (Student trainees are underlined)

- J14** Haoxiang Zhang, **Shaowei Wang**, Tse-Hsun Chen, Ahmed E. Hassan. *Are comments on Stack Overflow well organized for easy retrieval by developers?* ACM Transactions on Software Engineering and Methodology (TOSEM 2020).
- J13** Aaditya Bhatia, **Shaowei Wang**, Muhammad Asaduzzaman, Ahmed E. Hassan. *A Study of Bug Management Using the Stack Exchange Question and Answering Platform*. IEEE Transactions on Software Engineering (TSE 2020). Impact factor 4.778.
- J12** Jiayuan Zhou, **Shaowei Wang**, Cor-Paul Bezemer, Ying Zou, Ahmed E. Hassan. *Studying the Association between Bountysource Bounties and the Issue-addressing Likelihood of GitHub Issue Reports*. IEEE Transactions on Software Engineering (TSE 2020). Impact factor 4.778.
- J11** Haoxiang Zhang, **Shaowei Wang**, Tse-Hsun Chen, Ahmed E. Hassan. *Reading Answers on Stack Overflow: Not Enough!* IEEE Transactions on Software Engineering (TSE 2020). Impact factor 4.778.
- J10** Gopi Krishnan Rajbahadur, **Shaowei Wang**, Yasutaka Kamei, Ahmed E. Hassan. *Impact of Discretization Noise on Machine Learning Classifiers when Studying Software Engineering Datasets*. IEEE Transactions on Software Engineering (TSE 2019) (to appear). Impact factor 4.778.
- J9** Jiayuan Zhou, **Shaowei Wang**, Cor-Paul Bezemer, Ahmed E. Hassan. *Bounties on Technical Q&A Sites: A Case Study of Stack Overflow Bounties*. Empirical Software Engineering Journal (EMSE 2019) (to appear). Impact factor 4.457.
- J8** Haoxiang Zhang, **Shaowei Wang**, Tse-Hsun Chen, Ying Zou, Ahmed E. Hassan. *An Empirical Study of Obsolete Answers on Stack Overflow*. IEEE Transactions on Software Engineering (TSE 2019). Impact factor 4.778.
- J7** **Shaowei Wang**, Tse-Hsun Chen, Ahmed E. Hassan. *How Do Developers Revise Answers on Technical Q&A Websites? A Case Study on Stack Overflow*. IEEE Transactions on Software Engineering (TSE 2018). **Invited to ICSE 2019 as part of the Journal First presentation**. Impact factor 4.778.
- J6** Yuhao Wu, **Shaowei Wang**, Cor-Paul Bezemer, Katsuro Inoue. *How Do Developers Utilize Source Code from Stack Overflow?* Empirical Software Engineering Journal (EMSE 2018), pages 1–37. Impact factor 4.457. **Invited to FSE 2018 as part of the Journal First presentation**.

- J5** Hanyang Hu, **Shaowei Wang**, Cor-Paul Bezemer, Ahmed E. Hassan. *Studying the consistency of star ratings and reviews of popular free hybrid android and iOS apps*. Empirical Software Engineering Journal (EMSE 2018), pages 1–27. Impact factor 4.457.
- J4** **Shaowei Wang**, Tse-Hsun Chen, Ahmed E. Hassan. *Understanding the Factors for Fast Answers in Technical Q&A Websites: An Empirical Study on Four Stack Exchange Websites*. Empirical Software Engineering Journal (EMSE 2017), pages 1552–1593. Impact factor 4.457. **Invited to ICSE 2018 as part of the Journal First presentation.**
- J3** **Shaowei Wang**, David Lo, Bogdan Vasilescu, Alexander Serebrenik. *EnTagRec++: An Enhanced Tag Recommendation System for Software Information Sites*. Empirical Software Engineering Journal (EMSE 2017), pages 800–832. Impact factor 4.457.
- J2** **Shaowei Wang**, David Lo. *AmaLgam+: Composing Rich Information Sources for Accurate Bug Localization*. Journal of Software: Evolution and Process (JSEP 2016), pages 921–942.
- J1** **Shaowei Wang**, David Lo, Lingxiao Jiang. *AutoQuery: Automatic Construction of Dependency Queries for Code Search*. Automated Software Engineering (ASE 2014), pages 393–425. Impact factor 2.625.

Conference Publications (Student trainees are underlined)

- C18** Gopi Krishnan Rajbahadur, **Shaowei Wang**, Yasutaka Kamei, Ahmed E Hassan. *The Impact of Using Regression Models to Build Defect Classifiers*. In Proceedings of the 14th International Conference on Mining Software Repositories (MSR 2017), pages 135–145. Acceptance ratio is 31%.
- C17** Fei Lv, Hongyu Zhang, Jian-guang Lou, **Shaowei Wang**, Dongmei Zhang, Jianjun Zhao. *CodeHow: Effective Code Search based on API Understanding and Extended Boolean Model*. In Proceedings of the 30th IEEE/ACM International Conference on Automated Software Engineering (ASE 2015), pages 260–270. Acceptance ratio is 19%. [Cited 40+ times]
- C16** Meili Lu, Xiaobing Sun, **Shaowei Wang**, David Lo, Yucong Duan. *Query expansion via WordNet for effective code search*. In Proceedings of the 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2015), pages 545–549. [Cited 40+ times]
- C15** Siqi Ma, **Shaowei Wang**, David Lo, Robert Huijie Deng, Cong Sun. *Active Semi-supervised Approach for Checking App Behavior against Its Description*. In Proceedings of the 39th IEEE Annual Computer Software and Applications Conference (COMPSAC 2015), pages 179–184.
- C14** **Shaowei Wang**, David Lo, Lingxiao Jiang. *Active Code Search: Incorporating User Feedback to Improve Code Search Relevance*. In Proceedings of the 29th IEEE/ACM International Conference on Automated Software Engineering (ASE 2014), pages 677–682.
- C13** **Shaowei Wang**, David Lo, Julia Lawall. *Compositional Vector Space Models for Improved Bug Localization*. In Proceedings of the 30th International Conference on Software Maintenance and Evaluation (ICSME 2014), pages 171–180. Acceptance ratio is 19%. [Cited 30+ times]
- C12** **Shaowei Wang**, David Lo, Bogdan Vasilescu, Alexander Serebrenik. *An Enhanced Tag Recommendation System for Software Information Sites*. In Proceedings of the 30th International Conference on Software Maintenance and Evaluation (ICSME 2014), pages 291–300. Acceptance ratio is 19%. [Cited 70+ times]
- C11** **Shaowei Wang**, David Lo. *Version History, Similar Report, and Structure: Putting Them Together for Improved Bug Localization*. In Proceedings of the 22nd IEEE International Conference on Program Comprehension (ICPC 2014), pages 53–63. **Nominated for Best Paper Award.** [Cited 100+ times]
- C10** **Shaowei Wang**, David Lo, Lingxiao Jiang. *An Empirical Study on Developer Interactions in StackOverflow*. In Proceedings of the 28th ACM SIGAPP Symposium On Applied Computing (SAC 2013), pages 1019–1024. [Cited 70+ times]

- C9 Tien-Duy B. Le, **Shaowei Wang**, and David Lo. *Multi-Abstract Concern Localization*. In Proceedings of the 29th IEEE International Conference on Software Maintenance (ICSM 2013), pages 364–367.
- C8 Ferdian Thung, **Shaowei Wang**, David Lo, and Julia Lawall. *Automatic Recommendation of API Methods from Feature Requests*. In Proceedings of the 28th IEEE/ACM International Conference on Automated Software Engineering (ASE 2013), pages 290–300. Acceptance ratio is 17%. [Cited 40+ times]
- C7 **Shaowei Wang**, David Lo, Lingxiao Jiang. *Understanding Widespread Changes: A Taxonomic Study*. In Proceedings of the 17th European Conference on Software Maintenance and Reengineering (CSMR 2013), pages 4–14 .
- C6 Tegawend F Bisseyand, Ferdian Thung, **Shaowei Wang**, Daniel Lo, Lingxiao Jiang, Laurent Reveillere *Empirical Evaluation of Bug Linking*. In Proceedings of the 17th European Conference on Software Maintenance and Reengineering (CSMR 2013), pages 89–98. [cited 40+ times]
- C5 Thung Ferdian, **Shaowei Wang**, David Lo, Lingxiao Jiang. *An Empirical Study of Bugs in Machine Learning Systems*. In Proceedings of the 23rd IEEE International Symposium on Software Reliability Engineering (ISSRE 2012), pages 271–280.
- C4 **Shaowei Wang**, David Lo, Lingxiao Jiang. *Semantically Related Software Terms and Their Taxonomy By Leveraging Collaborative Tagging*. In Proceedings of the 28th IEEE International Conference on Software Maintenance (ICSM 2012), pages 604–607. [Cited 40+ times]
- C3 **Shaowei Wang**, David Lo, Lingxiao Jiang. *Code Search via Topic-Enriched Dependency Graph Matching*. In Proceedings of the 18th IEEE Working Conference on Reverse Engineering (WCRE 2011), pages 119–123.
- C2 **Shaowei Wang**, David Lo, Lingxiao Jiang. *Concern Localization using Information Retrieval: An Empirical Study on Linux Kernel*. In Proceedings of the 18th IEEE Working Conference on Reverse Engineering (WCRE 2011), pages 92–96. [Cited 50+ times]
- C1 **Shaowei Wang**, David Lo, Lingxiao Jiang, Lucia, Hoong Chuin Lau. *Search-Based Fault Localization*. In Proceedings of the 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011), pages 556–559.

Book Chapter

- B1 **Shaowei Wang**, David Lo, Lingxiao Jiang. *Scalable Parallelization of Specification Mining Using Distributed Computing*. book chapter, The Art and Science of Analyzing Software Data, 2015

Papers under Review (Student trainees are underlined)

- R6 Gopi Krishnan Rajbahadur, **Shaowei Wang**, Gustavo A. Oliva, Yasutaka Kamei, Ahmed E. Hassan. *The impact of feature importance methods on the interpretation of defect classifiers*. IEEE Transactions on Software Engineering (TSE).
- R5 Haoxiang Zhang, **Shaowei Wang**, Tse-Hsun Chen, Heng Li, Ahmed E. Hassan. *Perils in Code Reuse: An Empirical Study of C/C++ Code Weaknesses on Stack Overflow* IEEE Transactions on Software Engineering.
- R4 Jiayuan Zhou, **Shaowei Wang**, Yasutaka Kamei, Ahmed E. Hassan, Naoyasu Ubayashi, *Studying Donations and their Expenses in Open Source Projects: A Case Study of GitHub Projects Collecting Donations through Open Collectives*. Empirical Software Engineering Journal (EMSE).
- R3 Anran Chen, Tse-Hsun (Peter) Chen, **Shaowei Wang**, *Demystify the Challenges and Benefits of Analyzing User-Reported Logs in Bug Reports*. Empirical Software Engineering Journal (EMSE).
- R2 Aaditya Bhatia, Gopi Krishnan Rajbahadur, Muhammad Asaduzzaman, **Shaowei Wang**, Ahmed E. Hassan. *An Exploratory Study of the Differences Between Bug Reports and Feature Requests Using the Stack Exchange Question and Answering Platform*. IEEE Transactions on Software Engineering (TSE).

Research Experience

Assistant Professor

Department of Computer Science, University of Manitoba, Manitoba, Canada

May 2020 – present

Assistant Professor

Department of Computer Science and Engineering, Mississippi State University, MS, USA

Aug. 2019 – May 2020

Postdoctoral Research Fellow

School of Computing, Queen's University, On, Canada.

Supervisor: Dr. Ahmed E. Hassan

Jan. 2016 – July 2019

Visiting Scholar

Heinz College, Carnegie Mellon University, PA, USA

Aug. 2013 – May 2014

Industrial Experience

Senior R&D Engineer

Huawei, Shenzhen, China.

Aug. 2015 – Jan. 2016

- Designed and led the development of a code search engine.

Research Intern

Microsoft Research Asia, Beijing, China.

May 2014 – Aug. 2014

- Designed and developed a code search plugin in Visual Studio.

Mentoring Experience

Current (co-)supervised students [publications]:

- Gopi Krishnan Rajbahadur (PhD) [C18, J10, R6]
- Haoxiang Zhang (PhD) [J8, J11, R1, R5]
- Jiayuan Zhou (PhD) [J9, J12, R4]
- Hengyi Fang (MSc)
- Anran Chen (PhD) [R3]

July 2016 - Present

Step. 2017 - Present

Step. 2017 - Present

Aug. 2019 - Present

Jan. 2019 - present

Graduated (co-)supervised students [publications]:

- Ferdian Thung (PhD) [C8]
- Hanyang Hu (MSc) [J5]
- Yuhao Wu (PhD) [J6]
- Aaditya Bhatia (MSc) [J13, R2]

Aug. 2013 - May 2015

Sep. 2016 - May. 2017

Sep. 2016 - May. 2018

Aug. 2018 - Aug. 2019

Undergraduated co-supervised students:

- Chuck Fu

Teaching Experience

Assistant professor at Mississippi State University

CSE 4833/6833: Introduction to Algorithms

CSE 4214/6214: Introduction to Software Engineering

Fall, 2019

Spring, 2020

Teaching assistant at Queen's University

CISC880: Mining Software Engineering Data (graduate course)

Fall, 2016

Teaching assistant at Singapore Management University
IS201: Object-Oriented Program Language (undergraduate course)

Spring, 2015

Teaching assistant at Singapore Management University
IS706: Software Mining and Analysis (graduate course)

Spring, 2015 and Spring, 2013

Awards

- Rapid Response Reliable Reviewer Award, ASE 2020.
- Distinguished Reviewer (4 out of 450) of Springer EMSE (SE's highest impact journal), 2018.
- Excellent Internship Award, Microsoft Research Asia, 2015.
- Nominated for Best Paper Award, ICPC, 2014.
- PhD Research Scholarship, Singapore Management University, 2010-2015.
- Honored Undergraduate Student, Zhejiang University, 2010.

Education

Ph.D., Software Engineering
Singapore Management University, Singapore
Dissertation Topic: Multimodal Code Search
Supervisor: Dr. David Lo

Sept. 2010 – July 2015

B.Sc., Software Engineering
Zhejiang University, Hangzhou, Zhejiang, China

Sept. 2006 – June 2010

Professional Services

Organizing Committee Member

- The 37th International Conference on Software Maintenance and Evolution (ICSME 2021), Organizing Committee.
- The 28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2020), Organizing Committee.
- 7th IEEE/ACM International Conference on Mobile Software Engineering and Systems (MOBILEsoft 2020), Organizing Committee.

Program Committee Member

- The 35th IEEE/ACM International Conference on Automated Software Engineering (ASE), Research track, 2020.
- The 27th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Research Track, 2020
- The 28th IEEE/ACM International Conference on Program Comprehension (ICPC), Research Track, 2020.
- The 34th IEEE/ACM International Conference on Automated Software Engineering (ASE), Demonstration track, 2019.
- The 26th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), Research Track, 2019
- The 27th IEEE/ACM International Conference on Program Comprehension (ICPC), Research Track, 2019.
- The 16th International Conference on Mining Software Repositories (MSR), Mining Challenge, 2019
- The 26th IEEE/ACM International Conference on Program Comprehension (ICPC), Research Track, 2018.
- The 25th IEEE International Conference on Program Comprehension (ICPC), Research Track, 2017.

Journal Reviewer

- Empirical Software Engineering Journal (EMSE) [on review board for 2016 and 2017].
- IEEE Transactions on Software Engineering (TSE).

- ACM Transactions on Software Engineering and Methodology (TOSEM).
- Science of Computer Programming (SCC).
- Information and Software Technology (IST).
- Journal of Software: Evaluation and Process (JSEP).
- Journal of Systems and Software (JSS)
- Software Testing, Verification & Reliability (STVR)

Co-Reviewer

- The 40th International Conference on Software Engineering (ICSE), 2018.