

Tse-Hsun (Peter) Chen

PhD Candidate, Queen's University
196 Westmount Rd N. Waterloo, ON, Canada

tsehsun@cs.queensu.ca +1 6135392423
<http://petertsehsun.github.io/>

Highlights of Qualifications

- Research already used on a daily basis in practice (i.e., BlackBerry) [H1–H3, R2]
- Proven student mentorship record (Papers with my mentees [C2, R1])
- Research Interests: software performance engineering, software log analysis, program analysis, mining software repositories, and empirical software engineering.
- Canadian Citizen

High Impact Publications (Top cited, Used in practice)

- H1** Tse-Hsun Chen, Weiyi Shang, Zhen Ming Jiang, Ahmed E. Hassan, Mohamed Nasser and Parminder Flora. *Finding and Evaluating the Performance Impact of Redundant Data Access for Applications that are Developed Using Object-Relational Mapping Frameworks*. IEEE Transactions on Software Engineering (TSE). In Press. **Impact factor 2.292 (2013)**.
- H2** Tse-Hsun Chen, Weiyi Shang, Zhen Ming Jiang, Ahmed E. Hassan, Mohamed Nasser and Parminder Flora. *Detecting Performance Anti-patterns for Applications Developed Using Object-Relational Mapping*. The 36th International Conference on Software Engineering (ICSE 2014). 12 pages. **Acceptance rate 20% (99/495)**. [Cited 20+ Times]
- H3** Tse-Hsun Chen, Weiyi Shang, Ahmed E. Hassan, Mohamed Nasser and Parminder Flora. *Detecting Problems in Database Access Code of Large Scale Systems - An Industrial Experience Report*. The 38th International Conference on Software Engineering, Software Engineering in Practice (ICSE 2016). 10 pages. **Acceptance rate 26% (28/108)**.
- The techniques and developed frameworks presented in the above-mentioned papers are used by BlackBerry to verify the quality of their large-scale software systems on a daily basis.
- H4** Tse-Hsun Chen, Stephen W. Thomas, Meiyappan Nagappan, and Ahmed E. Hassan. *Explaining Software Defects Using Topic Models*. The 9th Working Conference on Mining Software Repositories (MSR 2012). 10 pages. **Acceptance rate 28% (18/64)**. [Cited 30+ Times]
- This paper was cited over 30 times.

Other Journal Publications (see High Impact Publications for additional publications)

- J1** Tse-Hsun Chen, Weiyi Shang, Meiyappan Nagappan, Weiyi Shang, Ahmed E. Hassan, Stephen W. Thomas. *Topic-based Software Defect Explanation*. Journal of Systems and Software. In Press. (Impact Factor 1.485)
- J2** Tse-Hsun Chen, Stephen W. Thomas, and Ahmed E. Hassan. *A Survey on the Use of Topic Models when Mining Software Repositories*. Empirical Software Engineering. In Press. (Impact Factor 2.161)

Other Conference Publications (see High Impact Publications for additional publications)

- C1** Tse-Hsun Chen, Weiyi Shang, Jinqiu Yang, Ahmed E. Hassan, Michael W. Godfrey, Mohamed Nasser, and Parminder Flora. *An Empirical Study on the Practice of Maintaining Object-Relational Mapping Code in Java Systems*. 12 pages. The 13th International Conference on Mining Software Repositories (MSR 2016). 12 pages. (Acceptance rate 35%)

- C2** Tarek M. Ahmed, Cor-Paul Bezemer, Tse-Hsun Chen, Ahmed E. Hassan and Weiyi Shang. *Studying the Effectiveness of Application Performance Management (APM) Tools for Detecting Performance Regressions for Web Applications: An Experience Report*. The 13th International Conference on Mining Software Repositories (MSR 2016). 12 pages. (Acceptance rate 35%)
- C3** Tse-Hsun Chen, Meiyappan Nagappan, Emad Shihab, Ahmed E. Hassan. *An Empirical Study of Dormant Bugs*. The 11th Working Conference on Mining Software Repositories (MSR 2014). 10 pages. (Acceptance rate 33%)

Other Posters and Technical Demonstrations

- D1** Big Data Analytics for Software Performance in Ultra-large Scale Systems. IBM Centre for Advanced Studies Conference (CASCON) Exhibit Session. November, 2014. Toronto, Canada.
- D2** Big Data Analytics and Software Engineering. IBM Centre for Advanced Studies Conference (CASCON) Exhibit Session. November, 2013. Toronto, Canada.
- D3** Detecting Performance Anti-patterns for Applications Developed Using Object-Relational Mapping. Consortium for Software Engineering Research (CSER) Exhibit Session. November, 2013. Toronto, Canada.
- D5** Explaining Software Defects Using Topic Models. Consortium for Software Engineering Research (CSER) Exhibit Session. November, 2012. Toronto, Canada.

Industrial Experience

Software Performance Associate Jan. 2013 – Present
BlackBerry, Waterloo, Ontario, Canada

- Designed and led the development of a static bug detection framework
- Designed and led the development of a real-time log analysis framework using Spark and Kafka
- Designed and led the development of a workload-aware real-time system configuration optimization framework
- Designed and led the development of test case prioritization tools using static code analysis
- Designed and executed performance tests

Software Developer Aprl. 2010 – Oct. 2010
Scalable Analytics Inc., Vancouver, BC, Canada

- Developed data visualization tools for big data analytics
- Developed tools that utilize graph structures to store large-scale data for doing data analysis

Part-time Computer Repair Technician May. 2007 – Sept. 2007
Shang-Yi Ltd., Tainan, Taiwan

- Helped customers setup computers and install software
- Troubleshooted software configuration problems

Research Experience

Graduate Research Assistant at Queen's University Sept. 2011 – Present

- Conducted research on applying program analysis techniques to detect problems in database access code
- Conducted research on an infrastructure real-time system analysis using software logs
- Conducted research on mining large-scale execution logs
- Collaborated on research topics of software performance, software maintenance and evolution, software quality, software testing, and mining software repositories
- Mentored students in the research lab
- Published research findings in top software engineering conferences and journals
- Peer-reviewed papers for software engineering conferences and journals

Mentoring and Teaching Experience

Guest Lecturer at Queen's University

Jan. 2015, Jan. 2016

CISC 880 (graduate-level course): Engineering Ultra-large-scale systems

Tutor for first year computer science course at the University of British Columbia (volunteer)

Sept. 2009 – Dec. 2009

CISC 111: Introduction to Computation

Other Refereed Publications

- O1** Tse-Hsun Chen. *Improving the Quality of Large-Scale Database-Centric Software Systems by Analyzing Database Access Code*. Doctoral Symposium Workshop of the 31st International Conference on Data Engineering (ICDE 2015). 5 Pages.

Papers under Review

- R1** Tse-Hsun Chen, Weiyi Shang, Ahmed E. Hassan. *CacheOptimizer: Helping Developers Configure Caching Frameworks for Hibernate-based Database-centric Web Applications*. The 24th ACM SIGSOFT International Symposium on the Foundations of Software Engineering (FSE 2016).
- R2** Heng Li, Tse-Hsun Chen, Weiyi Shang, Ahmed E. Hassan. *Studying the Relationships between Logs and Topics*. The 31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016).

Tools Developed

- T1 DB Checker:** DB Checker is a static analysis tool for detecting data-access-layer related problems. The tool is designed and implemented by myself, and is integrated to the BlackBerry build system. The tool is executed on a daily basis to detect new problems. So far the tool has detected many performance-critical as well as functional bugs, and has helped improve system memory performance by over 70%.
- T2 AutoTuner:** AutoTuner is an automated system configuration tuning tool for dynamic workloads. AutoTuner uses Kafka and Spark to do real-time data stream analysis, and learns system behaviours using Machine Learning (ML) models. Then, AutoTuner uses the ML models to find the optimal server configuration to improve system performance and user experience. AutoTuner also leverages ML models to detect data anomalies in real time.
- T3 Cache Optimizer:** Cache Optimizer is an automated tool to configure the second-level cache of Hibernate-based systems for different workloads. Cache Optimizer is workload-independent, and can generate the optimal cache configuration when workloads change. Cache Optimizer uses log and static code analysis to recover the database accesses of a workload, and uses a coloured Petri net to find the optimal cache configuration.

Professional Services

Journal Reviewer

- Empirical Software Engineering Journal (EMSE)
- Journal of Systems and Software (JSS)

Co-Reviewer

- Empirical Software Engineering Journal (EMSE)
- Journal of Software: Evolution and Process (JSEP)
- Information and Software Technology
- The 36th International Conference on Software Engineering (ICSE), 2014.

- The 30th IEEE International Conference on Software Maintenance and Evolution (ICSME), 2014.
- The 11th IEEE Working Conference on Mining Software Repositories (MSR), 2014.
- The 10th IEEE Working Conference on Mining Software Repositories (MSR), 2013.

Education

Ph.D., Computer Science Sept. 2012 – Sept. 2016 (expected)
 Queen's University, Kingston, Ontario, Canada
 Dissertation Topic: Improving the Performance of Large-Scale Database-Centric Applications Through Source Code Analysis
 Supervisor: Dr. Ahmed E. Hassan

- Queen Elizabeth II Scholarship Science Technology, 2015
- ACM SIGSOFT CAPS Award, 2014
- Queen's Graduate Award, Queen's University, 2013, 2016
- Conference Travel Award, Queen's University, 2013, 2016

M.Sc., Computer Science Sept. 2011 – Nov. 2012
 Queen's University, Kingston, Ontario, Canada
 Dissertation Topic: Studying Software Quality Using Topic Models
 Supervisor: Dr. Ahmed E. Hassan

- Queen's Graduate Award, Queen's University, 2011
- Conference Travel Award, Queen's University, 2011

B.Sc. Computer Science Sept. 2005 – June 2010
 University of British Columbia, Vancouver, BC, Canada