Suhaib Mujahid, PhD Candidate

suhaib.mujahid@concordia.ca
linkedin.com/in/suhaibmujahid
github.com/suhaibmujahid

Research Interests

- Mining Software Repositories
- Software Ecosystems
- Release Engineering
- Machine Learning on Code
- Defect Prediction and Avoidance
- Distributed Systems including Web Services
- Cloud Computing
- Software Architecture

Education

Jan 2018 – Present
- **PhD in Software Engineering**
  Concordia University, Montreal, QC, Canada
  Dissertation Topic: Managing Package Dependencies in Software Ecosystems
  GPA: 4.19

Sep 2016 – Dec 2017
- **MASc in Software Engineering**
  Concordia University, Montreal, QC, Canada
  Thesis Topic: Determining and Detecting Permission Issues of Wearable Apps
  Distinguish: Best Master's Thesis Award (F. A. Gerard Prize)
  GPA: 4.15

Feb 2011 – Dec 2015
- **BASc in Information Systems**
  Palestine Polytechnic University, Hebron, Palestine
  Distinguish: Highest graduation project score ever in the history of the faculty.

Research Publications

Journal Articles

[J1] Chen, X., Abdalkareem, R., **Mujahid**, **S.**, Shihab, E., & Xia, X. (2021). Helping or not helping? why and how trivial packages impact the npm ecosystem. *Empirical Software Engineering Journal*. EMSE'21, *26*(2), 27. doi:10.1007/s10664-020-09904-w

[J2] Hoyos, J., Abdalkareem, R., **Mujahid**, **S.**, Shihab, E., & Bedoya, A. E. (2021). On the removal of feature toggles. *Empirical Software Engineering Journal*. EMSE'21, *26*(2), 15. doi:10.1007/s10664-020-09902-y

[J3] Costa, D. E., **Mujahid**, **S.**, Abdalkareem, R., & Shihab, E. (2021). Breaking type-safety in go: An empirical study on the usage of the unsafe package. *IEEE Transactions on Software Engineering Journal*. TSE'21, 1–1. doi:10.1109/TSE.2021.3057720

[J4] Abdalkareem, R., Oda, V., **Mujahid**, **S.**, & Shihab, E. (2020). On the impact of using trivial packages: An empirical case study on npm and pypi. *Empirical Software Engineering Journal*. EMSE'20, *25*(2), 1168–1204. doi:10.1007/s10664-019-09792-9

[J5] Abdalkareem, R., **Mujahid**, **S.**, & Shihab, E. (2020). A machine learning approach to improve the detection of CI skip commits. *IEEE Transactions on Software Engineering Journal*. TSE'20, *47*(3), 448–463. doi:10.1109/TSE.2020.2967380

[J6] Abdalkareem, R., **Mujahid**, **S.**, Shihab, E., & Rilling, J. (2019). Which commits can be CI skipped? *IEEE Transactions on Software Engineering Journal*. TSE'19, *47*(3), 448–463. doi:10.1109/TSE.2019.2897300

[J7] **Mujahid**, **S.**, Sierra, G., Abdalkareem, R., Shihab, E., & Shang, W. (2018). An empirical study of android wear user complaints. *Empirical Software Engineering Journal*. EMSE'21, *23*(6), 3476–3502. doi:10.1007/s10664-018-9615-8

Conference Proceedings

[C1] **Mujahid**, **S.**, Abdalkareem, R., Shihab, E., & McIntosh, S. (2020). Using others' tests to identify breaking updates. In *Proceedings of the 17th international conference on mining software repositories* (pp. 466–476). MSR'20. doi:10.1145/3379597.3387476

[C2] Cabral, G. G., Minku, L. L., Shihab, E., & **Mujahid**, **S.** (2019). Class imbalance evolution and verification latency in just-in-time software defect prediction. In *Proceedings of the 41st IEEE/ACM international conference on software engineering* (pp. 666–676). ICSE'19. doi:10.1109/ICSE.2019.00076

[C3] **Mujahid**, **S.**, Abdalkareem, R., & Shihab, E. (2018). Studying permission related issues in android wearable apps. In *Proceedings of the 2018 ieee international conference on software maintenance and evolution* (pp. 345–356). ICSME'18. doi:10.1109/ICSME.2018.00043

[C4] Abdalkareem, R., Nourry, O., Wehaibi, S., **Mujahid**, **S.**, & Shihab, E. (2017). Why do developers use trivial packages? an empirical case study on npm. In *Proceedings of the 11th joint european software engineering conference and symposium on the foundations of software engineering* (pp. 385–395). FSE'17. doi:10.1145/3106237.3106267

[C5] **Mujahid**, **S.** (2017). Detecting wearable app permission mismatches: A case study on android wear. In *Proceedings of the 11th joint european software engineering conference and symposium on the foundations of software engineering* (pp. 1065–1067). FSE'17. doi:10.1145/3106237.3121279

[C6] **Mujahid**, **S.**, Sierra, G., Abdalkareem, R., Shihab, E., & Shang, W. (2017). Examining user complaints of wearable apps: A case study on android wear. In *Proceedings of the 4th IEEE/ACM international conference on mobile software engineering and systems* (pp. 96–99). MOBILESoft'17. doi:10.1109/MOBILESoft.2017.25

Theses

[T1] **Mujahid**, **S.** (2018). *Determining and detecting permission issues of wearable apps* (Master's thesis, Concordia University, Montreal, Quebec ,Canada). Retrieved from https://spectrum.library.concordia.ca/983405

Submitted Papers

[S1] **Mujahid**, **S.**, Costa, D. E., Abdalkareem, R., Shihab, E., Saied, M. A., & Adams, B. (2021). *Towards using package centrality trend to identify packages in decline*. Submited to IEEE Transactions on Engineering Management Journal.

Experience

Academic

Sep 2016 – Present
- **Research Assistant,** Data-driven Analysis of Software Lab
  Concordia University, Montreal, QC, Canada

Jan 2018 – Apr 2020
- **Teaching Assistant,** Dep. of Computer Science and Software Engineering
  Concordia University, Montreal, QC, Canada
  - *SOEN 490: Capstone Software Engineering Design Project:* Meet biweekly with student groups to advise on project challenges and evaluate progress. Grade project milestones and final projects.
  - *SOEN 691E: Software Re-engineering:* Held office hours to assist and tutor students with course assignments, projects and prepare for exams. Marked course assignments and final projects.

Experience (continued)

Research Projects

Jan 2020 – Present

- **Pitfalls in the Go programming language**
  - Identified problems and pitfalls in the usage of the Go type system.
  - Extracted evidence that helped the Go team at Google to dispel misconception about specific language features.
  - Proposed suggestions related to type safety that led to changes in the Go language.

  Key publication:
  - [J4] TSE'21: Breaking Type-Safety in Go: An Empirical Study on the Usage of the unsafe Package.

Apr 2018 – Present

- **Managing open-source dependencies**
  - Designed a technique to identify open-source dependencies that are likely to be deprecated.
  - Designed a technique to detect breakage-inducing updates of open-source dependencies by leveraging tests from "the crowd".
  - Identified the most important factors that developers take into consideration when selecting an `npm` package.
  - Developed a tool to measure community interest of `npm` packages (`github.com/centrality-checker`).

  Key publications:
  - [S2] TEM'21: Towards using package centrality trend to identify packages in decline.
  - [J5] EMSE'20: On the impact of using trivial packages: an empirical case study on npm and PyPI.
  - [C1] MSR'20: Using others' tests to identify breaking updates.

Mar 2018 – Present

- **Predicting risky software changes using ML**
  - Developed a scalable system that analyzes software repositories and uses ML to predict potential bug inducing commits.
  - Proposed a novel online class imbalance learning algorithm to improve the state-of-the-art bug prediction algorithms.

  Key publication:
  - [C2] ICSE'19: Class imbalance evolution and verification latency in just-in-time software defect prediction.

Feb 2018 – Jan 2020

- **Removing noise from CI pipelines**
  - Proposed a novel ML technique to detect commits that can be CI skipped.
  - Developed a tool that suggests skipping the CI for insignificant commits (Replicated by Google: `github.com/google/git-presubmit-linter`).

  Key publications:
  - [J6] TSE'20: A machine learning approach to improve the detection of CI skip commits.
  - [J0] TSE'19: Which commits can be CI skipped?

Experience (continued)

Industrial

Jul 2019 – Apr 2020
- **Software Engineering Intern**
  Société Générale, Montreal, QC, Canada
  - Developed a system to analyze software repositories and predict risky changes using ML and integrated the system with Société Générale's development environment.
  - Deployed the developed system on Kubernetes.

Aug 2018 – Mar 2019
- **Data Scientist - Part Time**
  Galilei Innovations Inc., Montreal, QC, Canada
  - Developed a pilot pipeline to analyze and link mall's traffic data and retailers' data to help them better understand their customers.
  - Collaborated directly with the top management of Ivanhoé Cambridge and major retailers (i.e., Tristan, BonLook, and La Vie en Rose) to improve the analysis pipeline.

Mar 2018 – Nov 2018
- **Software Engineering Intern**
  National Bank of Canada, Montreal, QC, Canada
  - Identified the best performing ML model in predicting risky commits at the National Bank of Canada.
  - Deployed an open-source ML system on Amazon VPC and implemented missed features.

Jun 2008 – Aug 2016
- **Software Developer**
  Freelance, Hebron, Palestine
  - Developed websites for local businesses and international groups; e.g., `www.al-ahliya.com` and `www.hebronapartheid.org`.
  - Developed simple accounting and management applications for non-professional users to help small businesses.

Honors and Awards

Achievements

2020
- **Arctic Code Vault Contributor**
  Contributed code to several projects in the GitHub Archive Program

2019
- **Best Master's Thesis Award**
  The F.A. Gerard Prize for the Master's thesis at Concordia University

2016
- **First Place in a Business Competition**
  The Business Simulation Competition at the College of Administrative Sciences and Informatics

2015
- **University President's Honor List**
  For the bachelor's degree at Palestine Polytechnic University

2014
- **Faculty Dean's Honor List**
  For the bachelor's degree at Palestine Polytechnic University

Honors and Awards (continued)

Publications

2017
- **Third Place in Microsoft Research Competition** for Graduate Students at FSE, Germany
  Paper [C5] FSE'17: Detecting wearable app permission mismatches: a case study on Android Wear.

- **Featured in a Concordia News Release**
  Paper [C4] FSE'17: Why Do Developers Use Trivial Packages? An Empirical Case Study on npm.

Financial

2020	• Concordia University Accelerator Award	$5,000
2018	• Concordia University Conference and Exposition Award	$1,000
	• GSA's Conference Funding Subsidy	$250
2017	• ACM FSE Student Research Competition Award, Germany	$700
	• ACM SIGSOFT CAPS Award	$533
	• Concordia University Tuition Award of Excellence	$35,949
	• Concordia University Conference and Exposition Award	$1,000
	• Concordia University 25th Anniversary Scholarship	$12,000
2015	• Palestine Polytechnic University Award of Excellence	JOD 550
2014	• Palestine Polytechnic University Award of Excellence	JOD 600

Talks and Posters

Conference Talks

May 2021
- Towards Using Package Centrality Trend to Identify Packages in Decline
  In the Consortium for Software Engineering Research (CSER), Ottawa, Canada (presented online)

May 2020
- Using Others' Tests to Identify Breaking Updates
  In the International Conference on Mining Software Repositories (MSR), Seoul, South Korea (presented online)

Sep 2018
- Studying Permission Related Issues in Android Wearable Apps
  In the IEEE International Conference on Software Maintenance and Evolution (ICSME), Madrid, Spain

Sep 2017
- Detecting Wearable App Permission Mismatches: A Case Study on Android Wear
  In the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), Paderborn, Germany

May 2017
- Examining User Complaints of Wearable Apps: A Case Study on Android Wear
  In the IEEE/ACM International Conference on Mobile Software Engineering and Systems (MOBILESoft), Buenos Aires, Argentina

Posters

Oct 2018
- Detecting Breakage Updates on npm Ecosystem
  In the Consortium for Software Engineering Research (CSER), Toronto, Canada

Oct 2017
- Manifest Mismatches In Wearable Apps: A Case Study On Android Wear
  In the Consortium for Software Engineering Research (CSER), Toronto, Canada

Sep 2017
- Detecting Wearable App Permission Mismatches: A Case Study on Android Wear
  In the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), Paderborn, Germany

May 2017
- Using Analytics for Effective Step Detection
  In the Software Engineering Research Centre (SERC), Montreal, Canada

Professional Service and Volunteers

External Reviewer

2021 • ACM Transactions on Software Engineering and Methodology Journal (TOSEM)
2019 • International Conference on Mining Software Repositories (MSR)
2018 • Springer's Journal of Empirical Software Engineering (EMSE)
2017 • IEEE International Conference on Program Comprehension (ICPC)

Student Volunteer

2019 • ACM/IEEE International Conference on Software Engineering (ICSE), Montreal, Canada
2018 • Concordia University Welcome International Students, Montreal, Canada
• ACM CHI Conference on Human Factors in Computing Systems, Montreal, Canada
2017 • ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), Paderborn, Germany
2015 • Palestine Polytechnic University Days Event, Hebron, Palestine

Technical Skills

Programming Languages	Strong experience in: **Go, Python, and JavaScript/TypeScript** Familiar with: Java, PHP, R, and C/C++
Data Analysis	Pandas, SQL, NumPy, SciPy, Matplotlib, and Seaborn
Machine Learning	Scikit-learn, ONNX, and ONNX Runtime
Databases	MongoDB, MySQL, PostgreSQL, Neo4J, and OrientDB
Web Development	React, Relay, GraphQL, and Redux
DevOps	Docker, Docker Swarm, and Kubernetes

Certificates

July 2018 • **Leadership Essentials Certificate**
Concordia University, Montreal, Canada

May 2017 • **School of Automated Software Testing**
University of Genoa, Genova, Italy

May 2015 • **Cisco Certified Network Associate (CCNA)**
Palestine Polytechnic University, Hebron, Palestine

Training Workshops

Technical Tutorials

May 2019 • Running Applications on Kubernetes — Google
• Train a model with TensorFlow and run it in the browser — Google

Leadership and Management

Oct 2017 • Leading Teams for Success — Concordia University
• Professional Negotiation and Persuasion Skills — Concordia University
• Managing People's Performance — Concordia University

Training Workshops (continued)

	• Data Analytics Essentials: Can we always trust numbers?	Concordia University
Sep 2017	• Solving Problems and Seeing the Big Picture	Concordia University
	• Foundations of Leadership	Concordia University
	• Navigating Emotional Intelligence	Concordia University

Other Workshops

Aug 2017	• Health and Safety Workshop for Teaching Assistants	Gina Cody School
Feb 2017	• Research Funding and Networking Strategies with Mitacs	Concordia University