James C. Davis

School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47906

davisjam@purdue.edu
765-494-3133
https://davisjam.github.io

RESEARCH INTERESTS

My research improves computing systems, grounded in empirical measurements of how and why they are used. My projects have flavors of software engineering, cybersecurity, and systems.

PROFESSIONAL EXPERIENCE

Assistant Professor — Fall 2020-present
Purdue University — Electrical and Computer Engineering

Intern, Microsoft Research (RiSE group: Cloud Security) — Summer 2019
Microsoft Research, Redmond, WA — Mentored by Patrice Godefroid

Intern, IBM Research (Storage) — Summer 2018
IBM Research, Almaden, CA — Mentored by Deepavali Bhagwat

Graduate Research Assistant — 2016-2020
Virginia Tech — Advised by Dongyoon Lee

Software Engineer, IBM (GPFS) — 2012-2015, Summer 2016, Summer 2017
IBM, Poughkeepsie, NY

EDUCATION

Ph.D, Computer Science and Applications — 2015-2020
Virginia Tech, Blacksburg, VA

B.Sc. Computer Science, B.Sc. Mathematics — 2008-2012
Clarkson University, Potsdam, NY

RESEARCH GRANTS

[1] **Google, LLC**

Principal Investigator (Co-PI: Yung-Hsiang Lu)

Unrestricted Gift

2020. $80,000

[2] **Google, LLC**

Co-Principal Investigator (PI: Yung-Hsiang Lu)

Unrestricted Gift

2020. $20,000

[3] **Purdue University VEIL Program**

Principal Investigator (Co-PI: Kirsten Davis)

Intercultural Engineering Education

2020. $5,000

[4] **Under review: NSF OAC Core Small**

Co-PI (PI: Yung-Hsiang Lu + 2 collaborators at Loyola University Chicago)

Low-power machine learning: Theory and engineering

2020. $500,000

[5] **In preparation: NSF SaTC Small**

PI (Co-PI: Dongyoon Lee

Correct and Secure Regular Expressions

2021. $500,000

REFEREED CONFERENCE PUBLICATIONS

[1] **Davis**, Servant, Lee. *Using Selective Memoization to Defeat Regular Expression Denial of Service (ReDoS)*. Proceedings of the 42nd IEEE Symposium on Security and Privacy (**IEEE S&P'21**).

[2] Cha, Wittern, Baudart, **Davis**, Mandel, Laredo. *A Principled Approach to GraphQL Query Cost Analysis*. Proceedings of the 28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'20**). ***ACM Distinguished Paper Award***.

[3] Rupprecht, **Davis**, Arnold, Gur, Bhagwat. *Improving Reproducibility of Data Science Pipelines through Transparent Provenance Capture*. Proceedings of the 46th International Conference on Very Large Data Bases (**VLDB'20 Industry track**).

[4] **Davis**, Moyer, Kazerouni, and Lee. *Testing Regex Generalizability And Its Implications: A Large-Scale Many-Language Measurement Study*. Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (**ASE'19**).

[5] Michael, Donohue, **Davis**, Lee, and Servant. *Regexes are Hard: Decision-making, Difficulties, and Risks in Programming Regular Expressions*. Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (**ASE'19**). ***ACM Distinguished Paper Award***.

[6] **Davis**, Michael, Coghlan, Servant, and Lee. *Are Regular Expressions a Lingua Franca? An Empirical Study on the Re-use and Portability of Regular Expressions*. Proceedings of the 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'19**).

[7] Wittern, Cha, **Davis**, Baudart, Mandel. *An Empirical Study of GraphQL Schemas*. Proceedings of the 17th International Conference on Service-Oriented Computing (**ICSOC'19**).

[8] Fu, Ghaffar, **Davis**, and Lee. *EdgeWise: A Better Stream Processing Engine for the Edge*. 2019 USENIX Annual Technical Conference (**USENIX ATC'19**).

[9] **Davis**, Coghlan, Servant, and Lee. *The Impact of Regular Expression Denial of Service (REDOS) in Practice: an Empirical Study at the Ecosystem Scale*. Proceedings of the 26th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'18**). ***ACM Distinguished Paper Award***.

[10] **Davis**, Williamson, and Lee. *A Sense of Time for JavaScript and Node.js: First-Class Timeouts as a Cure for Event Handler Poisoning*. Proceedings of the 27th USENIX Security Symposium (**USENIX Security'18**).

[11] **Davis**, Thekumparampil, and Lee. *Node.fz: Fuzzing the Server-Side Event-Driven Architecture*. Proceedings of the Twelfth European Conference on Computer Systems (**EuroSys'17**).

JOURNAL ARTICLES

[1] Kazerouni, **Davis**, Basak, Shaffer, Servant, Edwards. *Fast and Accurate Incremental Feedback for Students' Software Tests Using Selective Mutation Analysis*. Journal of Systems and Software (**JSS'21**).

[2] Ozkan, Davis, **Davis**, James, Murzi, Knight. *Expectations and Experiences of Short-Term Study Abroad Leadership Teams*. Journal of International Engineering Education (**JIEE'20**).

SHORT PAPERS

[1] **Davis**. *On the Impact and Defeat of Regex DoS*. **ACM SRC – Grand Finals**. ***Second place, graduate student division***.

[2] **Davis**. *Rethinking Regex Engines to Address ReDoS*. **ACM SRC – ESEC/FSE'19**. ***First place, graduate student division***.

[3] Rupprecht, **Davis**, Arnold, Lubbock, Tyson, and Bhagwat. *Ursprung: Provenance for Large-Scale Analytics Environments*. Proceedings of the 2019 International Conference on Management of Data (**SIGMOD'19 Demo**).

[4] **Davis**, Kildow, and Lee. *The Case of the Poisoned Event Handler: Weaknesses in the Node.js Event-Driven Architecture*. Proceedings of the 10th European Workshop on Systems Security (**EuroSec'17**).

POSTERS

[1] Vivek, Chinnakotla, Banna, Vegesana, Yan, **Davis**, Lu, Thiruvathukal. *Exemplars for Machine Learning: Towards Software Engineering & Reproducibility*. SIAM Conference on Computational Science and Engineering (**CSE'20**).

PATENTS

[1] Davis, **Davis**. *Injection of Simulated Hardware Failure(s) in a File System for Establishing File System Tolerance-to-Storage-Failure(s)*. IBM, U.S. patent application 20200264961. U.S. patent pending.

[2] Davis, **Davis**. *Verification of the integrity of data files stored in copy-on-write (CoW) based file system snapshots*. IBM, U.S. patent application 20200242075. U.S. patent pending.

[3] **Davis**, Davis. *File metadata verification in a distributed file system*. IBM, U.S. patent 10,678,755B2, granted Jun. 9, 2020..

[4] Davis, **Davis**. *Testing of lock managers in computing environments*. IBM, U.S patent 10,061,777 B1, granted Aug. 28, 2018.

[5] **Davis**, Davis, Knop. *Detection of file corruption in a distributed file system*. IBM, U.S. patent 10,025,788, granted Jul. 17, 2018.

COURSES TAUGHT

ECE 595 — Advanced Software Engineering Spring 2021
Purdue University

ECE 368 — Data Structures Fall 2020
Purdue University

VIP: Open-Source TensorFlow Software Fall 2020, Spring 2021

Purdue University

VIP: SafeRegex Fall 2020, Spring 2021
Purdue University

CS 3114 — Data Structures and Algorithms Fall 2019
Virginia Tech

CS 1064 — Introduction to Programming in Python Spring 2019
Virginia Tech

Rising Sophomore Abroad Program (Track Leader) Spring 2018, Spring 2019
Virginia Tech

INVITED TALKS

Regexes Awry: Characterizing and Defeating Regex-based Denial of Service 2020
Clemson University CS department colloquium

Regex-based Denial of Service 2020
Clarkson University CS department colloquium

Regexes are Hard: Qualitative and Quantitative Perspectives 2019
NC State CS department colloquium

The Dangers of Copy/Pasting Code 2019
Episode of the Podcast "The Secure Developer": `https://tinyurl.com/DavisResearchPodcast`

Regexes in the Wild 2019
Virginia Tech department seminar

Academic Perspectives on Node.js 2018
Node.js Collaborator Summit, Vancouver

International Engineering Annual, 2015-2019
Rising Sophomore Abroad Program, Virginia Tech

AWARDS AND RECOGNITION

ACM Distinguished Paper Award, ESEC/FSE 2020	2020
Outstanding Graduate Student Service Award, CS@VT	2020
Second place, Grand Finals of the ACM Graduate Student Research Competition	2020
First place, Graduate Student Research Competition, ESEC/FSE 2019	2019
ACM Distinguished Paper Award, ASE 2019	2019
ACM Distinguished Paper Award, ESEC/FSE 2018	2018
Microsoft Security Researcher Acknowledgments (Regex DoS)	2018
Pratt Fellowship, Virginia Tech College of Engineering	2017-2019
Davenport Fellowship, Virginia Tech College of Engineering	2019
Graduate Fellow, VT Academy for Global Engineering	2019-2020
IBM Significant Contributor Award (Node.js)	2018
IBM Poughkeepsie's New hire of the month	2014
Frederica Clarkson Award	2012
Clarkson University's Outstanding Senior (x2): Mathematics, Computer science	2012
Clarkson University Phalanx Commendable Leadership	2011

EXTERNAL SERVICE

Reviewer, ACM Transactions on Software Engineering (TSE)	2020-present
Member, ICSE Demonstrations Track	ICSE 2021
Member, ESEC/FSE Artifact Evaluation Committee	ESEC/FSE 2020
Member, CGO Artifact Evaluation Committee	CGO 2019
Sub-reviewer: ASPLOS'18, EuroSys'18, MASCOTS'18, HPCA'19, CGO'19, ISMM'19	2016-2019
Regional Judge, ACM ICPC	Fall 2015

DEPARTMENTAL SERVICE

Committee member, Purdue ECE Undergraduate Curriculum Committee	2020-present
President, VT CS Graduate Student Council	2018-2019
Organizer, VT Systems Reading Group	2017-2020

PROFESSIONAL MEMBERSHIPS

Member, Association for Computing Machinery