Xueyuan Han Vanbastelaer

Assistant Professor

Department of Computer Science
227 Manchester Hall, 1834 Wake Forest Road
Wake Forest University

Email: vanbasm@wfu.edu
Website: www.vanbastelaer.com
Phone: +1 (336) 758-5472

Area of Interest and Specialization

My research interests are in **systems** and **security**, and applying **machine learning** in these contexts. Topics include audit systems for capturing endpoint (e.g., workstations and mobile devices) data provenance, analysis of graph-structured data, operating system security (and more specifically, intrusion detection and investigation), and interpretable and robust machine learning application for systems security.

Academic Appointments

July 2022 - *present* **Assistant Professor**, Wake Forest University
Sept. 2016 - May 2022 **Research Assistant**, Harvard University
Aug. 2018 - Feb. 2019 **Research Assistant**, University of Cambridge
Oct. 2015 – June 2016 **Undergraduate Research Assistant**, University of California at Los Angeles

Education

Sept. 2016 - May 2022 **Ph.D., Computer Science**, Harvard University, Cambridge, MA, USA
Thesis: *Detecting System Anomalies Using Kernel-level Data Provenance*
🏆 Harvard Computer Science Dissertation Award Nominee
Advisors: Dr. Margo Seltzer, Dr. James Mickens

Sept. 2016 - May 2022 **S.M., Computer Science**, Harvard University, Cambridge, MA, USA

Sept. 2011 - Dec. 2015 **B.S., Computer Science**, University of California at Los Angeles, Los Angeles, CA, USA
Summa cum Laude 4.0/4.0
🏆 Outstanding Bachelor of Science in Computer Science
Advisor: Dr. Miryung Kim

Research Support

PI, National Science Foundation
CRII: SaTC: Robust Explainable Provenance-based Intrusion Detection
NSF CNS-2245442, $174,999, 03/15/2023 - 02/28/2025

Honors & Awards

Provost's Fund for Faculty Travel, Wake Forest University, 2023
Siebel Scholar, Thomas and Stacey Siebel Foundation, 2022
The Siebel Scholars program was founded in 2000 by the Siebel Foundation to recognize the most talented students at the world's leading graduate schools of business, computer science, bioengineering, and energy science.
Student Travel Award, ACM Conference on Computer and Communications Security (CCS), 2018
Student Travel Award, ACM European Conference on Computer Systems (EuroSys), 2018
Distinction in Teaching, Derek Bok Center, Harvard University, 2018
Dean's Honors List for Superior Academic Achievement, UCLA, 2011 - 2015
Rose Gilbert Honors Scholarship, UCLA, 2012

Industry Experience

May 2020 – July 2020 **Research Intern**, Microsoft Research, New York, NY, USA
May 2019 – Aug. 2019 **Summer Research Intern**, NEC Laboratories America, Princeton, NJ, USA

Service

Broadening Participation in Computing

- **Faculty Mentor**, *Wake Forest LEAP (Lab Experiences: Academics and Professions).* A six-week, lab-based summer STEM internship program that targets underrepresented minority high-school junior and senior students who attend one of the Title I Winston-Salem/Forsyth County schools.

  Summer 2023 Justin Nguyen (Parkland Magnet High School)

- **Faculty Mentor**, *Wake Forest Undergraduate Research and Creative Activities Center*. Provide undergraduate students an opportunity to engage in mentored scholarship.

  Summer 2023 Yukun Michael Xi *and* Shiyu Andrea Jiang (Wake Forest Research Fellows)

Panelist

- National Science Foundation (NSF), 2023

Program Committees and Journal Reviews

- ACM Asia Conference on Computer and Communications Security (AsiaCCS), 2024
- Workshop on Cybersecurity Experimentation and Test (CSET), 2023
- International Provenance and Annotation Workshop (IPAW), 2023
- USENIX Workshop on the Theory and Practice of Provenance (TaPP), 2020, 2023
- IEEE Transactions on Information Forensics and Security (TIFS), 2023
- IEEE Transactions on Dependable and Secure Computing (TDSC), 2023
- PeerJ Computer Science, 2023
- ACM Computing Surveys (CSUR), 2022

Poster Jury and Shadow Program Committees

- IEEE Symposium on Security and Privacy (S&P), 2020, 2022
- ACM European Conference on Computer Systems (EuroSys), 2018

University Service

- Member, *Graduate Committee*, Department of Computer Science, Wake Forest University, 2022 - 2023
- Member, *Curriculum Committee*, Department of Computer Science, Wake Forest University, 2023 - *present*

Publicly Released Software

Kairos. A provenance-based anomaly detection system that not only detects intrusions but reconstructs succinct attack footprints. `https://github.com/ProvenanceAnalytics/kairos`.

Unicorn. A provenance-based intrusion detection system. Unicorn has been used as an evaluation benchmark in numerous academic publications. `https://github.com/crimson-unicorn`.

FRAPpucino. An intrusion detection system, Unicorn's precursor. `https://github.com/michael-hahn/frap`.

Xanthus. A tool to reproducibly generate system execution trace datasets. Many have used it to generate audit datasets to evaluate host intrusion detection systems. `https://github.com/tfjmp/xanthus`.

saBPF. An extension to the Linux Kernel to allow users to attach eBPF programs at the intersection of namespaces and security hooks. It enables deployment of security and audit mechanisms within the scope of a container without affecting the rest of the operating system. Output of this work was re-implemented by Google engineers and submitted as patches to the Linux kernel. `https://github.com/saBPF-project`.

CamFlow. A state-of-the-art audit system used in a number of research institutions, including Harvard University, University of Illinois in Urbana-Champaign, University of North Carolina at Charlotte, University of Cambridge, Universität Regensburg, Huazhong University of Science and Technology, University of British Columbia, University of Edinburgh, SRI International, and IBM. `https://github.com/CamFlow`.

Journal Publications

1. Thomas F. J.-M. Pasquier, Matthew K. Lau, Xueyuan Han, Elizabeth Fong, Barbara Staudt Lerner, Emery R. Boose, Mercè Crosas, Aaron M. Ellison, Margo I. Seltzer. "Sharing and Preserving Computational Analyses for Posterity with Encapsulator." IEEE Computing in Science & Engineering (CiSE) 20(4): 111-124 (2018).

Conference Publications

2. Zijun Cheng, Qiujian Lv, Jinyuan Liang, Yan Wang, Degang Sun, Thomas F. J.-M. Pasquier, Xueyuan Han. "KAIROS: Practical Intrusion Detection and Investigation using Whole-system Provenance." 45^{th} IEEE Symposium on Security and Privacy (S&P'24). San Francisco, CA, USA. May 20, 2024. *To Appear.*

3. Xueyuan Han, James Mickens, Siddhartha Sen. "Splice: Efficiently Removing a User's Data from In-memory Application State." 30^{th} ACM Conference on Computer and Communications Security (CCS'23). Copenhagen, Denmark. November 26, 2023, *To Appear.*

4. Akul Goyal, Xueyuan Han, Gang Wang, Adam Bates. "Sometimes, You Aren't What You Do: Mimicry Attacks against Provenance Graph Host Intrusion Detection Systems." 30^{th} ISOC Network and Distributed System Security Symposium (NDSS'23). San Diego, CA, USA. February 27, 2023. *Acceptance rate=14.6%.*

5. Soo-Yee Lim, Bogdan Stelea, Xueyuan Han, Thomas F. J.-M. Pasquier. "Secure Namespaced Kernel Audit for Containers." 12^{th} ACM Symposium on Cloud Computing (SoCC'21). Seattle, WA, USA. November 1, 2021.

6. Xueyuan Han, Xiao Yu, Thomas F. J.-M. Pasquier, Ding Li, Junghwan Rhee, James W. Mickens, Margo I. Seltzer, Haifeng Chen. "SIGL: Securing Software Installations Through Deep Graph Learning." 30^{th} USENIX Security Symposium (Security'21). August 11, 2021. *Acceptance rate=18.7%.*

7. Xueyuan Han, Thomas F. J.-M. Pasquier, Adam Bates, James Mickens, Margo I. Seltzer. "Unicorn: Runtime Provenance-Based Detector for Advanced Persistent Threats." 27^{th} ISOC Network and Distributed System Security Symposium (NDSS'20). San Diego, CA, USA, February 23, 2020. *Acceptance rate=17.4%.*

8. Thomas F. J.-M. Pasquier, Xueyuan Han, Thomas Moyer, Adam Bates, Olivier Hermant, David M. Eyers, Jean Bacon, Margo I. Seltzer. "Runtime Analysis of Whole-System Provenance." 25^{th} ACM Conference on Computer and Communications Security (CCS'18). Toronto, ON, Canada. October 15, 2018. *Acceptance rate=16.6%.*

9. Thomas F. J.-M. Pasquier, Xueyuan Han, Mark Goldstein, Thomas Moyer, David M. Eyers, Margo I. Seltzer, Jean Bacon. "Practical Whole-System Provenance Capture." 8^{th} ACM Symposium on Cloud Computing (SoCC'17). Santa Clara, CA, USA. September 24, 2017.

10. Muhammad Ali Gulzar, Matteo Interlandi, Xueyuan Han, Mingda Li, Tyson Condie, Miryung Kim. "Automated Debugging in Data-Intensive Scalable Computing." 8^{th} ACM Symposium on Cloud Computing (SoCC'17). Santa Clara, CA, USA. September 24, 2017.

Workshop Publications

11. Soo Yee Lim, Xueyuan Han, Thomas Pasquier. "Unleashing Unprivileged eBPF Potential with Dynamic Sandboxing." 1^{st} SIGCOMM 2023 Workshop on eBPF and Kernel Extensions (eBPF'23). New York City, NY. September 10, 2023.

12. Xueyuan Han, James Mickens, Ashish Gehani, Margo I. Seltzer, Thomas F. J.-M. Pasquier. "Xanthus: Push-button Orchestration of Host Provenance Data Collection." 3^{rd} ACM International Workshop on Practical Reproducible Evaluation of Computer Systems (P-RECS'20). Stockholm, Sweden. June 23, 2020.

13. Xueyuan Han, Thomas F. J.-M. Pasquier, Margo I. Seltzer. "Provenance-Based Intrusion Detection: Opportunities and Challenges." 10^{th} USENIX Workshop on the Theory and Practice of Provenance (TaPP'18). London, UK, July 11, 2018.

14. Xueyuan Han, Thomas F. J.-M. Pasquier, Tanvi Ranjan, Mark Goldstein, Margo I. Seltzer. "FRAPpuccino: Fault-Detection Through Runtime Analysis of Provenance." 9^{th} USENIX Workshop on Hot Topics in Cloud Computing (HotCloud'17). Santa Clara, CA, USA. July 10, 2017.

15. Muhammad Ali Gulzar, Xueyuan Han, Matteo Interlandi, Shaghayegh Mardani, Sai Deep Tetali, Todd D. Millstein, Miryung Kim. "Interactive Debugging for Big Data Analytics." 8^{th} USENIX Workshop on Hot Topics in Cloud Computing (HotCloud'16). Denver, CO, USA. June 20, 2016.

Talks

1. SIGL: Securing Software Installations Through Deep Graph Learning. 30^{th} *USENIX Security Symposium (Security'21)*. Virtual. August 2021.

2. Leveraging System-Level Data Provenance for Intrusion Detection. *IBM Research*. December 2020.

3. Runtime Provenance-Based Detector for Advanced Persistent Threats. *SRI International*. June 2020.

4. Unicorn: Runtime Provenance-Based Detector for Advanced Persistent Threats. 27^{th} *ISOC Network and Distributed System Security Symposium (NDSS'20)*. San Diego, CA, USA, February 2020.

5. Detecting Advanced Persistent Threats at Runtime Using Whole-System Data Provenance. *Nokia Bell Labs*. October 2019.

6. Using Provenance for Security and Interpretability. 12^{th} *EuroSys Doctoral Workshop (EuroDW'18)*. Porto, Portugal. April 2018.

7. FRAPpuccino: Fault-Detection Through Runtime Analysis of Provenance. 9^{th} *USENIX Workshop on Hot Topics in Cloud Computing (HotCloud'17)*. Santa Clara, CA, USA. July 2017.

Teaching

CSC 791: Thesis Research I, Instructor, *Wake Forest University*. Fall 2023
CSC 393: Individual Study, Instructor, *Wake Forest University*. Spring 2023
CSC 250: Computer Systems I, Instructor, *Wake Forest University*. Fall 2023
CSC 193: Independent Study, Instructor, *Wake Forest University*. Fall 2023
CSC 111: Introduction to Computer Science, Instructor, *Wake Forest University*. Fall 2022, Spring 2023, Fall 2023
COMPSCI 61: Systems Programming and Machine Organization, Head Teaching Fellow, *Harvard University*. Fall 2017
COM SCI 130: Software Engineering, Academic Grader, *UCLA*. 2015

Student Advising

Ph.D. Research Advising (from Other PhD-Granting Institutions)*

**Department of Computer Science at Wake Forest University does not grant doctorate degrees.*

- Dongqi Han, *Tsinghua University*, 2023 - *present*
- Ziyang Yu, *University of Chinese Academy of Sciences*, 2023 - *present*
- Md. Monowar Anjum, *University of British Columbia*, 2022 - 2023
- Zijun Cheng, *University of Chinese Academy of Sciences*, 2020 - 2023
- Akul Goyal, *University of Illinois at Urbana-Champaign*, 2019 - 2022
- Soo Yee Lim, *University of British Columbia*, 2020 – 2021

M.S. Research Advising

- Debashis Gupta, *Wake Forest University*, 2023 - *present*
- Raniery Mendes, *Wake Forest University*, 2022 - *present*
- Jinyuan Liang, *University of British Columbia*, 2021 - *present*

Undergraduate Research Advising (Wake Forest University): Dalal Ahmidouch (2023 - *present*), Olivia Caulfield (2022 - 2023, received the Canadian **Fulbright Scholarship** in 2022 under my mentorship), Yukun Michael Xi (2022 - 2023), Jason Zhang (2022 - *present*)

Undergraduate Research Advising (Other Institutions): Japraj Sandhu (*University of British Columbia*, 2023 - *present*), Jude Shamsi (*University of British Columbia*, 2020), Dennis Li (*Carnegie Mellon University*, 2017)

Dissertation Committees: Md Asifur Rahman (M.S., expected Fall 2023), Raniery Mendes (M.S., expected Spring 2024)

Patents

1. Xiao Yu, Xueyuan Han, Ding Li, Junghwan Rhee, Haifeng Chen. *Securing Software Installation Through Deep Graph Learning*. U.S. Patent 11321066. May 3, 2022.