Scott Alfeld

salfeld@amherst.edu • http://www.scottalfeld.net

INTERESTS The security of artificial intelligence and machine learning.

POSITION Assistant Professor of Computer Science, Amherst College 2017 – Present

EDUCATION

University of Wisconsin–Madison, Department of Computer Sciences 2011 – 2017

- Ph.D. in Computer Science (Minor: Mathematics)
  - Co-Advisers: Paul Barford and Xiaojin (Jerry) Zhu
- Master's Degree in Computer Science

University of Southern California, Department of Computer Science 2009 – 2011

- Ph.D. Program, no degree (Transferred to UW–Madison)

University of Utah, School of Computing 2004 – 2008

- Bachelor's of Science in Computer Science

PUBLICATIONS

PREPRINTS

- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld
  "Defense Against Shortest Paths"
  arXiv:2305.19083
- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld
  "Optimal Edge Weight Perturbations to Attack Shortest Paths"
  arxiv:2107.03347

JOURNAL PAPERS

- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld **TKDD '23**
  "Attacking Shortest Paths by Cutting Edges"
  Transactions on Knowledge Discovery from Data

CONFERENCE PAPERS

- A. Sarkar, M. Lanier, S. Alfeld, R. Garnett, N. Jacobs, Y. Vorobeychik **WACV '23**
  "A Visual Active Search Framework for Geospatial Exploration"
  in *Proceedings of the Winter Conference on Applications of Computer Vision*
- Z. Kong, S. Alfeld **ECAI '23**
  "Approximate Data Deletion in Generative Models"
  in *Proceedings of the European Conference on Artificial Intelligence*
- A. Vartanian, W. Rosenbaum, S. Alfeld **AAAI '23**
  "Training-Time Attacks Against k-Nearest Neighbors"
  in *Proceedings of the AAAI Conference on Artificial Intelligence*
- N. Marchant, B. I. P. Rubinstein, S. Alfeld **AAAI '22**
  "Hard to Forget: Poisoning Attacks on Certified Machine Unlearning"
  in *Proceedings of the AAAI Conference on Artificial Intelligence*
- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld **ECML-PKDD '21**
  "PATHATTACK: Attacking Shortest Paths in Complex Networks"
  in *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*
- D. Liu, Z. Shafi, W. Fleisher, T. Eliassi-Rad, S. Alfeld **AIES '21**
  "RAWLSNET: Altering Bayesian Networks to Encode Rawlsian Fair Equality of Opportunity."
  in *Proceedings of the ACM conference on Artificial Intelligence, Ethics, and Society*
- S. Yu, L. Torres, S. Alfeld, T. Eliassi-Rad, Y. Vorobeychik **SDM '21**
  "Optimizing Graph Structure for Targeted Diffusion"
  in *Proceedings of the Siam International Conference on Data Mining*
- S. Alfeld, A. Vartanian, L. Newman-Johnson, B. I. P. Rubinstein **AAAI '19**
  "Attacking Data Transforming Learners at Training Time"
  in *Proceedings of the AAAI Conference on Artificial Intelligence*
- A. Sen, S. Alfeld, X. Zhang, A. Vartanian, Y. Ma, X. Zhu **GameSec '18**
  "Training Set Camouflage"
  in *Proceedings of the Conference on Decision and Game Theory for Security*
- L. Tong, S. Yu, S. Alfeld, Y. Vorobeychik **ICML '18**
  "Adversarial Regression with Multiple Learners"
  in *Proceedings of the International Conference on Machine Learning*
- S. Yu, Y. Vorobeychik, S. Alfeld **AAMAS '18**
  "Adversarial Classification on Social Networks"
  in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*
- S. Alfeld, X. Zhu, P. Barford **AAAI '17**
  "Explicit Defense Actions Against Test-Set Attacks"
  in *Proceedings of the AAAI Conference on Artificial Intelligence*

PUBLICATIONS CONT.

- A. Cahn, S. Alfeld, P. Barford, S. Muthukrishnan **ASONAM '16**
  "What's in the Community Cookie Jar?"
  in *Proceedings of the IEEE/ACM Conference on Advances in Social Network Analysis and Mining*
- S. Alfeld, X. Zhu, P. Barford **SoCS '16**
  "Machine Teaching as Search" (Short Paper)
  in *Proceedings of the Symposium on Combinatorial Search*
- S. Alfeld, X. Zhu, P. Barford **AAAI '16**
  "Data Poisoning Attacks Against Autoregressive Models"
  in *Proceedings of the AAAI Conference on Artificial Intelligence*
- A. Cahn, S. Alfeld, P. Barford, S. Muthukrishnan **WWW '16**
  "An Empirical Study of Web Cookies"
  in *Proceedings of the World Wide Web Conference*
- M. Malloy, S. Alfeld, P. Barford **ISIT '15**
  "Contamination Estimation via Convex Relaxations"
  in *Proceedings of IEEE International Symposium on Information Theory*
- S. Alfeld, P. Barford **Energycon '14**
  "Targeted Residual Analysis for Improving Electric Load Forecasting"
  in *Proceedings of IEEE Energy Conference*
- S. Alfeld, C. Barford, P. Barford **e-Energy '12**
  "Toward an Analytic Framework for the Electrical Power Grid"
  in *Proceedings of the Third International Conference on Future Energy Systems*

ABSTRACTS AND WORKSHOP PAPERS

- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld
  " Defense Against Shortest Path Attacks"
  in *NETWORKS – A Joint Sunbelt and NetSci Conference '22*
- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld
  "Attacking the Shortest Path by Perturbing Edge Weights"
  in *NETWORKS – A Joint Sunbelt and NetSci Conference '21*
- M. Stein, S. Alfeld,
  "GARFD: Gradient-based Autoregressive Forecaster Defense"
  in *Workshop on Dataset Curation and Security at NeurIPS '20* 2020.
- X. Zhang, H. Ohannessian, A. Sen, S. Alfeld, X. Zhu
  "Optimal Teaching for Online Perceptrons"
  in *Constructive Machine Learning at NIPS '16*
- S. Alfeld, P. Barford, X. Zhu
  "Optimal Defense Actions Against Test Set Attacks"
  in *ICML Workshop on Reliable Machine Learning in the Wild '16*
- S. Alfeld, K. Berkele, S. DeSalvo, T. Pham, D. Russo, L.J. Yan, M.E. Taylor
  "Reducing the Team Uncertainty Penalty: Empirical and Theoretical Approaches"
  in *AAMAS workshop on Multiagent Sequential Decision Making in Uncertain Domains '11*
- S. Alfeld, M.E. Taylor, P. Tandon, M. Tambe
  "Towards a Theoretic Understanding of DCEE"
  in *AAMAS Distributed Constraint Reasoning Workshop '10*

PATENT

- M. Malloy, S. Alfeld and P. Barford, 2020
  Fraudulent Traffic Detection and Estimation
  US10832280B2

ADDITIONAL POSTERS

In addition to those for publications, above:

- *PATHATTACK: Attacking Shortest Paths in Complex Networks* 2021
  GraphEx
- *POTION: Optimizing Graph Structure for Targeted Diffusion* 2021
  NETWORKS – A Joint Sunbelt and NetSci Conference ’21
- *Machine Learning in the Presence of an Adversary* 2016
  Greater Chicago Area Systems Research Workshop (GCASR)
- *Improving Energy Efficiency: A Data-Driven Approach* 2013
  ACM SIGKDD Conference on Knowledge Discovery and Data Mining
- *Understanding and Improving the Electric Grid* 2012
  Wisconsin Institute on Software-Defined Datacenters Of Madison (WISDOM)

AWARDS

- Fund for Applied Research in Artificial Intelligence Security at Amherst 2023
  External gift GFT0118079, $20,000
- Amherst College Provost’s Research Fellowship 2020-2021
  100% of 9-month salary
- Google Cloud Platform Education Grant 2018
  Funds for Google Compute Engine use, $3,200
- Cisco Distinguished Graduate Fellowship 2016-2017
  One of two annual awards for graduate students in UW’s CS department.
  Provides tuition and a stipend for 9 months.
- Student-voted Favorite Talk for UW’s *Estimating Functions From Data*, 2012
  Talk Title: *Maximum Covariance Unfolding*
  $200 Prize
- Center for Risk and Economic Analysis of Terrorism Events (CREATE) Fellowship 2010–2011
  Part of the DHS Career Development Student Fellowship Program.
  Provides tuition and a stipend for 12 months.
- University of Utah *School of Computing Outstanding Teaching Assistant* Award 2008 – 2009
  An annual award to up to two teaching assistants in the School of Computing.

PREVIOUS WORK EXPERIENCE

- University of Wisconsin-Madison, Graduate Student Researcher 2011 - 2017
- *comScore* Intern 2014
  I worked with their data scientists toward designing, implementing, and deploying an anomaly detection system.
- *mDotLabs* (acquired by comScore) Contractor 2014
  I designed and implemented a statistical model evaluation framework.
- *The Eric and Wendy Schmidt Data Science For Social Good Fellowship* 2013
  I led the Energy Team. We built a tool to profile a building’s electricity usage, and audited several facilities for the Illinois Department of Corrections to improve their energy efficiency.
- University of Southern California, Graduate Student Researcher 2009-2011
- University of Utah, Teaching Assistant 2005-2008

SERVICE

- Member of Advisory Board, Cranium.ai 2023 – Present
- PhD Dissertation Comitteee, *Northeastern University* 2023
- Invited Participant, *Center for Advancing Safety of Machine Intelligence Workshop* 2023
- Advised teams in KPMG's Enterprise Innovation organization as a subject matter expert on AI 2022
- Invited Participant, *Northwestern University Machine Learning Impact Initiative Summit* 2021
- Invited Participant, *Northwestern University Machine Learning Impact Initiative Workshop* 2020
- Honors Examiner, *Swarthmore College* 2019, 2020
- Invited Panelist, *Panel on Adversarial Learning* at GameSec 2018 2018
- Book Reviewer for *Adversarial Machine Learning*, Morgan & Claypool 2018

CONFERENCES AND JOURNALS

- Regularly review for: AAAI, EAAI, ICML, NeurIPS 2016 – Present
- EAAI Mentored Undergraduate Research Challenge Reviewer 2022 – Present
- EAAI Model AI Assignment Reviewer 2019 – Present
- AAAI Undergraduate Consortium Reviewer 2021, 2022
- AAAI 2020 SPC (Meta-Reviewer) 2020
- AIRE Reviewer 2019
- AISTATS 2017 Workflow Chair 2017
- Energycon 2014 Reviewer 2014

AT AMHERST COLLEGE

- *Task Force on Guidelines for the Use of Generative AI Tools* 2023
  Task Force Member
- *AI and War* 2023
  Invited Panelist
- *The Impact of Artificial Intelligence on Art* 2022
  Invited Panelist
- *Lavender Mentor* 2022
  Program hosted by the Amherst College Queer Resource Center for LBGTQ+ students
- *College Council* 2022 – 2023
  Committee Member
- *The Constant Conundrum: Academic Freedom and Civil Discourse* 2022
  Invited Panelist
- *Ad Hoc Faculty Committee on Academic Structures during COVID-19* 2020
  Committee Member
- *Faculty Computer Committee* 2019-2020
  Committee Member
- Instruction Weekend Science Center Panel 2019
  Invited Panelist
- Amherst Promise Campaign Data Science Panel, NYC 2019
  Invited Panelist
- Summer Undergradate Research Fellowship Ethics Luncheon 2018
  Invited Panelist
- Amherst College Marker Faire Table 2018
- Five-College Data Science Initiative Working Group 2017 – Present

INVITED TALKS

- *What is Gen AI & How is it Impacting Education?* — Sep 2023
  Inaugural talk for the AI Learning Lab - Lunch Series, Amherst College
- *Attacks and Defenses for Networked Systems* — Aug 2023
  Talk at "The Brain Drain - Keeping up with the Cutting Edge of AI Security", Zoom
- *The Future of AI and Amherst's Role in it* — Jun 2023
  Talk at Amherst College Reunion
- *On the Offensive: Attacking AI Systems* — Apr 2022
  Talk at KPMG
- *Manipulating Learners at Training Time* — Feb 2021
  Talk at MassMutual as part of the ECS:edu series
- *Attacking at Training Time: Complicated Attacks Against Simple Learners* — Aug 2020
  Talk at Robustness of AI Systems Against Adversarial Attacks (RAISA3) 2020
- *Attack, Defend, Steal:* — Jul 2020
  Student-Led Research Projects in Adversarial Machine Learning at Amherst College
  Talk at Amherst College Reunion, Zoom
- *When Hackers Meet Data* — Jun 2020
  Talk as part of the STEM Incubator Colloquium Series
- *The Intersection of Machine Learning and Security* — Feb 2020
  Talk at "An Evening with Professor Scott Alfeld", hosted by Alex Ginsburg NYC
- *The Role of Hackers in Data Analysis* — Nov 2019
  Talk at Hampshire College, Amherst, Massachusetts
- *Adversarial Machine Learning* — May 2019
  Talk at Amherst College Reunion
- *Manipulating Learners: Machine Learning in the Presence of Adversarial Input* — Mar 2019
  Talk at MIT Lincoln Labs, Lexington, Massachusetts
- *Hacking Machine Learning* — Aug 2018
  Talk at University of Melbourne, Melbourne
- *Attacking and Defending Forecasters* — Jul 2017
  Talk at Vanderbilt University, Nashville, Tennessee
- *Deep Security – How to Pick a Lock* — Jan 2017
  Lecture on Physical Security at Google, Madison
- *Time Series Forecasting in the Presence of an Adversary* — Nov 2015
  Lecture for the Human, Animal, and Machine Learning: Experiment and Theory (HAMLET) organization
- *Improving Load Forecasting by Augmenting the MISO Model* — Sep 2012
  Presentation to the Load Forecasting Team of Midwest ISO (MISO)
- *Analyzing the Grid via Wholesale Electricity Markets* — Feb 2012
  Presentation to the BACTER Institute at UW–Madison

VOLUNTEER WORK

- *ASA Five College DataFest, Amherst* — 2023
- *Physical Security Workshop, All Campus Makerspace, UMass, Amherst* — 2020
- *UMass ECE Department's Circuits and Code* — 2018, 2019
- *Recreational Lockpicking Workshop, UMass, Amherst* — 2018
- *LockDown* IT Security event — 2017
- *Wisconsin Science Festival* — 2016
- *LockDown* IT Security event — 2015
- 2014 Milwaukee Maker Faire — 2014
- American Player's Society, Madison, WI — 2011
- Hi-GEAR High School Girl's Outreach Program, Salt Lake City — 2008
- University of Utah School of Computing High School Programming Competition — 2007
- National Forensics League National Tournament, Salt Lake City — 2004