Scott Alfeld

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

INTERESTS The security of artificial intelligence and machine learning.

POSITION Associate Professor of Computer Science, Amherst College 2017 – Present (Tenured in 2024)

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
  "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

- S. Choi, S. Alfeld **GameSec '25**
  "Theoretical Bounds on the Adversarial Robustnessof Reduced-Rank Regression"
  in *Proceedings of the Conference on Decision and Game Theory for Security*
- B. A. Miller, Z. Shafi, W. Ruml, Y. Vorobeychik, T. Eliassi-Rad, S. Alfeld **SDM '25**
  "Defense Against Shortest Paths"
  in *Proceedings of the Siam International Conference on Data Mining*
- A. Sarkar, M. Lanier, S. Alfeld, R. Garnett, N. Jacobs, Y. Vorobeychik **WACV '24**
  "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

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