

Bihe Zhao

PhD candidate, CISPA
Saarbrücken, Germany

+49-681870832632

bihe.zhao@cispa.de

Personal Website

Google Scholar

EDUCATION

- **CISPA Helmholtz Center for Information Security & Saarland University** 06/2024–present
Ph.D. candidate in Computer Science, supervised by Prof. Adam Dziedzic and Prof. Franziska Boenisch
– Research Topic: Data Provenance for Generative Models
- **Beihang University (BUAA)** 09/2021–01/2024
Master in Cyber Science and Technology Score: 93.05/100 Rank: 6/49
– Major Courses: Matrix Theory (99/100), Cyber Security (97/100), Algorithm Design and Analysis (96/100)
- **Beihang University (BUAA)** 09/2017–06/2021
Bachelor in Cyber Science and Technology Score: 91.20/100 Rank: 5/50
– Major Courses: Information Theory and Encoding (99/100), Natural Language Processing (95/100)
- **University of Illinois at Urbana-Champaign (UIUC)** 07/2018–08/2018
Visiting Student at Global Education and Training Program for Accounting and Finance GPA: 4.0/4.0

PUBLICATIONS

1. **Zhao B.**, Maini P., Boenisch F., Dziedzic A. Unlocking Post-hoc Dataset Inference with Synthetic Data. International Conference on Machine Learning (**ICML**) 2025.
2. **Zhao B.**, Kerner L., Meintz M., Boenisch F., Dziedzic A. Data Provenance for Image Auto-Regressive Generation. International Conference on Learning Representations (**ICLR**) 2026.
3. **Zhao B.**, Meintz M., Xu J., Boenisch F., Dziedzic A. Data Circuit Breaker for Image Generative Models. European Conference on Computer Vision (**ECCV**) 2026.
4. **Zhao B.**, Xu J., Boenisch F., Dziedzic A. PRISM: Post-hoc Recognition of Image Source Models. *Under review at AAAI*. 2026.
5. Wang X., **Zhao B.**, Backes M., Boenisch F., Dziedzic A. AgentProv: Auditing Agentic LLM API Providers via Tool-use Policy Probes. Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2026.
6. Kerner L., Meintz M., **Zhao B.**, Boenisch F., Dziedzic A. BitMark: Watermarking Bitwise Autoregressive Image Generative Models. Annual Conference on Neural Information Processing Systems (**NeurIPS**) 2025.
7. Deng X.*, **Zhao B.***, Guan Z., Xu M. A New Finding and Unified Framework for Fake Image Detection. IEEE Signal Processing Letters, 2023. (* *Equal contribution*)
8. Zhang, L., Gao, X., **Zhao, B.**, Guan, Z., Bian, S. GRAMSSAT: An efficient label inference attack against two-party split learning based on gradient matching and semi-supervised learning. Journal of Information Security and Applications. 2025.
9. Guan Z., Zhang L., Huang B., **Zhao B.**, Bian S. Adaptive Hyperparameter Optimization for Black-box Adversarial Attack. International Journal of Information Security. 2025.
10. Zhang Y., Liu J., Guan Z., **Zhao B.**, Leng X., Bian S. ARMOR: Differential Model Distribution for Adversarially Robust Federated Learning. Electronics. 2025.

PROFESSIONAL EXPERIENCE

- **Research Intern at SenseTime Technology** 01/2022–04/2023
– Proposed a query-efficient model extraction attack based on public datasets that outperforms state-of-the-art model extraction attacks by a large margin.
– Revealed an observation for face forgery detection and designed a unified detection framework based on the finding.
– Implemented both projects with Pytorch.
- **Software Engineer Intern at ByteDance Technology** 08/2020–02/2021
– Assisted in the development of data annotation and management platform.
– Developed and improved an alarm center that has more than 20,000 rules to detect unusual data traffic.
– Wrote more than 5,000 lines of code with Go.

AWARDS

- Excellent Graduate of Beijing (top 3%) 11/2023
- Excellent Graduate of Beihang University (top 5%) 06/2021
- First Prize, Academic Excellence Award (top 5%) 10/2019
- First Prize, 12th National College Student Information Security Competition (top 3%) 08/2019
- Excellent Student of Beihang University (top 5%) 06/2019

SERVICE AND VOLUNTEERING

- **Reviewer** of ICML 2026 (**Gold Reviewer**), ECCV 2026, NeurIPS 2026, TIFS 08/2025–present
- **Mentor** for European Championship in AI & Cybersecurity 2025/2026 08/2025–present
- **Teaching Assistant** of The Secret of Cryptology, Beihang University 09/2021–01/2023
- **Mentor** for 13th National College Student Information Security Competition, First Prize 03/2022–08/2022