Yang Zhang | CV

✉ zhang@cispa.de • yangzhangalmo.github.com

last update: June 15, 2022

Employment

CISPA Helmholtz Center for Information Security — **Saarbrücken, Germany**
Faculty — *2/2020 -*

CISPA Helmholtz Center for Information Security — **Saarbrücken, Germany**
Research Group Leader — *1/2019 - 1/2020*

CISPA, Saarland University — **Saarbrücken, Germany**
Postdoctoral Researcher — *1/2017 - 12/2018*
Host: Michael Backes

Education

University of Luxembourg — **Luxembourg, Luxembourg**
Ph.D. in Computer Science, highest honor — *12/2012 - 11/2016*
Supervisor: Sjouke Mauw and Jun Pang

Shandong University — **Jinan, China**
Master in Computer Science — *9/2009 - 62012*

University of Luxembourg — **Luxembourg, Luxembourg**
Master in Informatics, exchange student — *9/2010 - 10/2011*

Shandong University — **Jinan, China**
Bachelor in Software Engineering — *9/2005 - 6/2009*

Research Projects

- Understanding the individual host response against Hepatitis D Virus to develop a personalized approach for the management of hepatitis D; Funding agency: Horizon Europe; Role: Co-PI; Time: 10/2022 - 9/2026; Budget: 318,375€
- The Norton Labs Graduate Fellowship;[1] Funding agency: Norton Lab; Role: Lead PI; Time: 5/2022 - 4/2023; Budget 20,000$
- Dual-level Specification Inference; Funding agency: NGI Assure; Role: Lead PI; Time: 10/2022 - 9/2023; Budget: 34,391€
- Developing AI/Cloud based User Data Security Technologies; Funding agency: Ministry of Science and ICT, Korea; Role: Lead PI; Time: 9/2021 - 9/2022; Budget: 7,275€
- Trustworthy Federated Data Analytics; Funding agency: Helmholtz Association; Role: Co-PI; Time: 1/2020 - 1/2023; Budget: 110,925€

[1] This fellowship is given to my PhD student Xinlei He.

Research Interests

- Trustworthy Machine Learning
- Social Network Analysis
- Misinformation, Hate Speech, and Memes

Service

- PC member
  - 2023: IEEE S&P, NDSS
  - 2022: USENIX Security, CCS, NeurIPS, ICLR, KDD, WWW, AAAI, PETS, ASIACCS
  - 2021: USENIX Security, CCS, WWW, AAAI, Euro S&P, PETS, ASIACCS
  - 2020: CCS, WWW, ICWSM, RAID, PETS
  - 2019: CCS, ISMB/ECCB
- PC chair
  - The 1st International Workshop on Ethics in Computer Security (EthiCS 2022)
- Organizer
  - Privacy and Security in ML Seminars[2]

Awards

- Busy Beaver teaching award for seminar "Privacy of Machine Learning" at Saarland University (2021 Winter)
- Distinguished paper award, NDSS 2019
- Distinguished reviewer award, TrustML Workshop 2020 (co-located with ICLR 2020)
- Best paper award, ARES 2014

Publication

I have published more than 50 peer-reviewed papers so far. My publication list can also be found at DBLP[3] and Google Scholar,[4], however, they may not be up to date. Note that in the domain of information security, the most prestigious conferences are IEEE S&P, CCS, USENIX Security, and NDSS.

Conference

[1] Zheng Li and Yiyong Liu and Xinlei He and Ning Yu and Michael Backes and **Yang Zhang**. Auditing Membership Leakages of Multi-Exit Networks. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2022.

[2] Min Chen and Zhikun Zhang and Tianhao Wang and Michael Backes and Mathias Humbert and **Yang Zhang**. Graph Unlearning. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2022.

[3] Tianshuo Cong and Xinlei He and **Yang Zhang**. SSLGuard: A Watermarking Scheme for Self-supervised Learning Pre-trained Encoders. In *ACM SIGSAC Conference on Computer and*

[2] https://prisec-ml.github.io/
[3] https://dblp.org/pid/06/6785-16.html
[4] https://scholar.google.com/citations?user=Xeb2888AAAAJ

Communications Security (CCS). ACM, 2022.

[4] Yun Shen and Yufei Han and Zhikun Zhang and Min Chen and Ting Yu and Michael Backes and **Yang Zhang** and Gianluca Stringhini. Finding MNEMON: Reviving Memories of Node Embeddings. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2022.

[5] Yugeng Liu and Rui Wen and Xinlei He and Ahmed Salem and Zhikun Zhang and Michael Backes and Emiliano De Cristofaro and Mario Fritz and **Yang Zhang**. ML-Doctor: Holistic Risk Assessment of Inference Attacks Against Machine Learning Models. In *USENIX Security Symposium (USENIX Security)*. USENIX, 2022.

[6] Yufei Chen and Chao Shen and Cong Wang and **Yang Zhang**. Teacher Model Fingerprinting Attacks Against Transfer Learning. In *USENIX Security Symposium (USENIX Security)*. USENIX, 2022.

[7] Zhikun Zhang and Min Chen and Michael Backes and Yun Shen and **Yang Zhang**. Inference Attacks Against Graph Neural Networks. In *USENIX Security Symposium (USENIX Security)*. USENIX, 2022.

[8] Xinyue Shen and Xinlei He and Michael Backes and Jeremy Blackburn and Savvas Zannettou and **Yang Zhang**. On Xing Tian and the Perseverance of Anti-China Sentiment Online. In *International Conference on Weblogs and Social Media (ICWSM)*. AAAI, 2022.

[9] Yun Shen and Xinlei He and Yufei Han and **Yang Zhang**. Model Stealing Attacks Against Inductive Graph Neural Networks. In *IEEE Symposium on Security and Privacy (S&P)*. IEEE, 2022.

[10] Ahmed Salem and Rui Wen and Michael Backes and Shiqing Ma and **Yang Zhang**. Dynamic Backdoor Attacks Against Machine Learning Models. In *IEEE European Symposium on Security and Privacy (Euro S&P)*. IEEE, 2022.

[11] Ahmed Salem and Michael Backes and **Yang Zhang**. Get a Model! Model Hijacking Attack Against Machine Learning Models. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2022.

[12] Junhao Zhou and Yufei Chen and Chao Shen and **Yang Zhang**. Property Inference Attacks Against GANs. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2022.

[13] Xinlei He and **Yang Zhang**. Quantifying and Mitigating Privacy Risks of Contrastive Learning. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 845–863. ACM, 2021.

[14] Min Chen and Zhikun Zhang and Tianhao Wang and Michael Backes and Mathias Humbert and **Yang Zhang**. When Machine Unlearning Jeopardizes Privacy. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 896–911. ACM, 2021.

[15] Minxing Zhang and Zhaochun Ren and Zihan Wang and Pengjie Ren and Zhumin Chen and Pengfei Hu and **Yang Zhang**. Membership Inference Attacks Against Recommender Systems. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 864–879. ACM, 2021.

[16] Zheng Li and **Yang Zhang**. Membership Leakage in Label-Only Exposures. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 880–895. ACM, 2021.

[17] Xiaoyi Chen and Ahmed Salem and Michael Backes and Shiqing Ma and Qingni Shen and Zhonghai Wu and **Yang Zhang**. BadNL: Backdoor Attacks Against NLP Models with Semantic-preserving Improvements. In *Annual Computer Security Applications Conference (ACSAC)*, pages 554–569. ACSAC, 2021.

[18] Xinlei He and Jinyuan Jia and Michael Backes and Neil Zhenqiang Gong and **Yang Zhang**. Stealing Links from Graph Neural Networks. In *USENIX Security Symposium (USENIX Security)*, pages 2669–2686. USENIX, 2021.

[19] Zhikun Zhang and Tianhao Wang and Jean Honorio and Ninghui Li and Michael Backes and Shibo He and Jiming Chen and **Yang Zhang**. PrivSyn: Differentially Private Data Synthesis. In *USENIX Security Symposium (USENIX Security)*, pages 929–946. USENIX, 2021.

[20] Fatemeh Tahmasbi and Leonard Schild and Chen Ling and Jeremy Blackburn and Gianluca Stringhini and **Yang Zhang** and Savvas Zannettou. "Go eat a bat, Chang!": On the Emergence of Sinophobic Behavior on Web Communities in the Face of COVID-19. In *The Web Conference (WWW)*. ACM, 2021.

[21] Rui Wen and Yu Yu and Xiang Xie and **Yang Zhang**. LEAF: A Faster Secure Search Algorithm via Localization, Extraction, and Reconstruction. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 1219–1232. ACM, 2020.

[22] Dingfan Chen and Ning Yu and **Yang Zhang** and Mario Fritz. GAN-Leaks: A Taxonomy of Membership Inference Attacks against Generative Models. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 343–362. ACM, 2020.

[23] Ahmed Salem and Apratim Bhattacharya and Michael Backes and Mario Fritz and **Yang Zhang**. Updates-Leak: Data Set Inference and Reconstruction Attacks in Online Learning. In *USENIX Security Symposium (USENIX Security)*, pages 1291–1308. USENIX, 2020.

[24] Inken Hagestedt and Mathias Humbert and Pascal Berrang and Irina Lehmann and Roland Eils and Michael Backes and **Yang Zhang**. Membership Inference Against DNA Methylation Databases. In *IEEE European Symposium on Security and Privacy (Euro S&P)*, pages 509–520. IEEE, 2020.

[25] **Yang Zhang** and Mathias Humbert and Bartlomiej Surma and Praveen Manoharan and Jilles Vreeken and Michael Backes. Towards Plausible Graph Anonymization. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2020.

[26] Jinyuan Jia and Ahmed Salem and Michael Backes and **Yang Zhang** and Neil Zhenqiang Gong. MemGuard: Defending against Black-Box Membership Inference Attacks via Adversarial Examples. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 259–274. ACM, 2019.

[27] Zheng Li and Chengyu Hu and **Yang Zhang** and Shanqing Guo. How to Prove Your Model Belongs to You: A Blind-Watermark based Framework to Protect Intellectual Property of DNN. In *Annual Computer Security Applications Conference (ACSAC)*, pages 126–137. ACSAC, 2019.

[28] Zhiqiang Zhong and **Yang Zhang** and Jun Pang. A Graph-Based Approach to Explore Relationship Between Hashtags and Images. In *International Conference Web Information Systems Engineering (WISE)*, pages 473–488. Springer, 2019.

[29] Tahleen Rahman and Bartlomiej Surma and Michael Backes and **Yang Zhang**. Fairwalk: Towards Fair Graph Embedding. In *International Joint Conferences on Artifical Intelligence (IJCAI)*, pages 3289–3295. IJCAI, 2019.

[30] **Yang Zhang**. Language in Our Time: An Empirical Analysis of Hashtags. In *The Web Conference (WWW)*, pages 2378–2389. ACM, 2019.

[31] Ahmed Salem and **Yang Zhang** and Mathias Humbert and Pascal Berrang and Mario Fritz and Michael Backes. ML-Leaks: Model and Data Independent Membership Inference Attacks and Defenses on Machine Learning Models. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2019.

[32] Inken Hagestedt and **Yang Zhang** and Mathias Humbert and Pascal Berrang and Haixu Tang and XiaoFeng Wang and Michael Backes. MBeacon: Privacy-Preserving Beacons for DNA Methylation Data. In *Network and Distributed System Security Symposium (NDSS)*. Internet Society, 2019.

[33] Fanghua Zhao and Linan Gao and **Yang Zhang** and Zeyu Wang and Bo Wang and Shanqing Guo. You Are Where You App: An Assessment on Location Privacy of Social Applications. In *International Symposium on Software Reliability Engineering (ISSRE)*, pages 236–247. IEEE, 2018.

[34] **Yang Zhang** and Mathias Humbert and Tahleen Rahman and Cheng-Te Li and Jun Pang and Michael Backes. Tagvisor: A Privacy Advisor for Sharing Hashtags. In *The Web Conference (WWW)*, pages 287–296. ACM, 2018.

[35] Pascal Berrang and Mathias Humbert and **Yang Zhang** and Irina Lehmann and Roland Eils and Michael Backes. Dissecting Privacy Risks in Biomedical Data. In *IEEE European Symposium on Security and Privacy (Euro S&P)*, pages 62–76. IEEE, 2018.

[36] Michael Backes and Mathias Humbert and Jun Pang and **Yang Zhang**. walk2friends: Inferring Social Links from Mobility Profiles. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 1943–1957. ACM, 2017.

[37] Jun Pang and **Yang Zhang**. Quantifying Location Sociality. In *ACM Conference on Hypertext and Social Media (HT)*, pages 145–154. ACM, 2017.

[38] Jun Pang and **Yang Zhang**. DeepCity: A Feature Learning Framework for Mining Location Check-Ins. In *International Conference on Weblogs and Social Media (ICWSM)*, pages 652–655. AAAI, 2017.

[39] Yan Wang and Zongxu Qin and Jun Pang and **Yang Zhang** and Xin Jin. Semantic Annotation for Places in LBSN Using Graph Embedding. In *ACM International Conference on Information and Knowledge Management (CIKM)*, page 2343–2346. ACM, 2017.

[40] **Yang Zhang** and Minyue Ni and Weili Han and Jun Pang. Does #like4like Indeed Provoke More Likes? In *International Conference on Web Intelligence (WI)*, pages 179–186. ACM, 2017.

[41] Minyue Ni and **Yang Zhang** and Weili Han and Jun Pang. An Empirical Study on User Access Control in Online Social Networks. In *ACM Symposium on Access Control Models and Technologies (SACMAT)*, pages 12–23. ACM, 2016.

[42] Jun Pang and Polina Zablotskaia and **Yang Zhang**. On Impact of Weather on Human Mobility in Cities. In *International Conference Web Information Systems Engineering (WISE)*, pages 247–256. Springer, 2016.

[43] Jun Pang and **Yang Zhang**. Location Prediction: Communities Speak Louder than Friends. In *ACM Conference on Online Social Networks (COSN)*, pages 161–171. ACM, 2015.

[44] **Yang Zhang** and Jun Pang. Distance and Friendship: A Distance-based Model for Link Prediction in Social Networks. In *Asia-Pacific Web Conference (APWeb)*, pages 55–66. Springer, 2015.

[45] Jun Pang and **Yang Zhang**. Event Prediction with Community Leaders. In *Conference on Availability, Reliability and Security (ARES)*, pages 238–243. IEEE, 2015.

[46] Marcos Cramer and Jun Pang and **Yang Zhang**. A Logical Approach to Restricting Access in Online Social Networks. In *ACM Symposium on Access Control Models and Technologies (SACMAT)*, pages 75–86. ACM, 2015.

[47] Jun Pang and **Yang Zhang**. Cryptographic Protocols for Enforcing Relationship-based Access Control Policies. In *Annual IEEE Computers, Software and Applications Conference (COMPSAC)*, pages 484–493. IEEE, 2015.

[48] Jun Pang and **Yang Zhang**. Exploring Communities for Effective Location Prediction. In *International Conference on World Wide Web (WWW)*, pages 87–88. ACM, 2015.

[49] **Yang Zhang** and Jun Pang. Community-driven Social Influence Analysis and Applications. In *International Conference on Web Engineering (ICWE)*. Springer, 2015.

[50] Jun Pang and **Yang Zhang**. A New Access Control Scheme for Facebook-style Social Networks. In *Conference on Availability, Reliability and Security (ARES)*, pages 1–10. IEEE, 2014.

Journal

[51] Cheng-Te Li and Cheng Hsu and **Yang Zhang**. FairSR: Fairness-aware Sequential Recommendation through Multi-Task Learning with Preference Graph Embeddings. *ACM Transactions on Intelligent Systems and Technology*, 2022.

[52] Xinlei He and Qingyuan Gong and Yang Chen and **Yang Zhang** and Xin Wang and Xiaoming Fu. DatingSec: Detecting Malicious Accounts in Dating Apps Using a Content-Based Attention Network. *IEEE Transactions on Dependable and Secure Computing*, 2021.

[53] Bo-Heng Chen and Cheng-Te Li and Kun-Ta Chuang and Jun Pang and **Yang Zhang**. An Active Learning-based Approach for Location-aware Acquaintance Inference. *Knowledge and Information Systems*, 2018.

[54] Jun Pang and **Yang Zhang**. A New Access Control Scheme for Facebook-style Social Networks. *Computers & Security*, 2015.

Teaching

- Advanced Lecture: Machine Learning Privacy (2022 Summer)
- Seminar: Data-driven Understanding of the Disinformation Epidemic (2022 Summer)
- Seminar: Privacy of Machine Learning (2021 Winter)
- Advanced Lecture: Privacy Enhancing Technologies (2021 Summer)
- Seminar: Data-driven Understanding of the Disinformation Epidemic (2021 Summer)
- Seminar: Data Privacy (2020 Winter)
- Advanced Lecture: Privacy Enhancing Technologies (2020 Summer)
- Seminar: Data-driven Approaches on Understanding Disinformation (2020 Summer)
- Seminar: Data Privacy (2019 Winter)
- Advanced Lecture: Privacy Enhancing Technologies (2019 Summer)
- Seminar: Biomedical Privacy (2019 Summer)
- Seminar: Data Privacy (2018 Winter)
- Advanced Lecture: Privacy Enhancing Technologies (2018 Summer)
- Seminar: Adversarial Machine Learning (2018 Summer)

Students

Ph.D. Students

Boyang Zhang — **CISPA Helmholtz Center for Information Security**, *12/2021 -*

Zheng Li — **CISPA Helmholtz Center for Information Security**, *2/2021 -*

Xinlei He — **CISPA Helmholtz Center for Information Security**, *2/2020 -*

Co-supervised Ph.D. Students

Yuan Xin — **CISPA Helmholtz Center for Information Security**, *3/2022 -*
with Michael Backes

Yugeng Liu — **CISPA Helmholtz Center for Information Security**, *1/2022 -*
with Michael Backes

Wai Man Si — **CISPA Helmholtz Center for Information Security**, *11/2021 -*
with Michael Backes

Hai Huang — **CISPA Helmholtz Center for Information Security**, *11/2021 -*
with Michael Backes

Yiting Qu — **CISPA Helmholtz Center for Information Security**, *11/2021 -*
with Michael Backes

Rui Wen — **CISPA Helmholtz Center for Information Security**, *10/2021 -*
with Michael Backes

Ph.D. Preparatory Phase

Zeyang Sha — **CISPA Helmholtz Center for Information Security**, *10/2021 -*

Yixin Wu — **CISPA Helmholtz Center for Information Security**, *10/2021 -*

Ziqing Yang — **CISPA Helmholtz Center for Information Security**, *10/2021 -*

Yiyong Liu — **CISPA Helmholtz Center for Information Security**, *5/2021 -*

Xinyue Shen — **CISPA Helmholtz Center for Information Security**, *5/2021 -*

Minxing Zhang — **CISPA Helmholtz Center for Information Security**, *5/2021 -*

Visitors

Tianshuo Cong — **Tsinghua University**
visiting Ph.D. student — *8/2021 -*

Alumni

Shannon Pierson — **Saarland University**
visiting researcher — *3/2022 - 6/2022*

Ahmed Salem — **CISPA Helmholtz Center for Information Security**
co-supervised Ph.D. student, now at MSR — *2/2017 - 1/2022*

Fan Zhang — **Tsinghua University**
intern — *7/2021 - 10/2021*

Bartlomiej Surma — **CISPA Helmholtz Center for Information Security**
co-supervised Ph.D. student, now at Google — *6/2016 - 9/2021*

Joann Chen — **UC Irvine**
intern — *7/2021 - 9/2021*

Ge Han — **Shandong University**
visiting Ph.D. student — *10/2019 - 8/2021*

Suliya — **Chinese Academy of Science**
visiting Ph.D. student, now at JD — *1/2020 - 1/2021*

Xiaoyi Chen — **Peking University**
visiting Ph.D. student — *10/2019 - 10/2020*

Yuhao Mao — **Zhejiang University**
intern, now at ETH — *6/2020 - 9/2020*

Leonard Schild — **Saarland University**
research assistant, now at KU Leuven — *3/2017 - 7/2020*

Zeyu Yang — **Zhejiang University**
visiting Ph.D. student — *10/2019 - 3/2020*