

# Yinuo Du

✉ yinuod@andrew.cmu.edu     yinuo-du  
 <https://aduyinuo.github.io/>  
 <https://scholar.google.com/citations?user=XdY3VB0AAAAJ>




## Summary

---

My research is to study **human and AI decision-making in cybersecurity**. Toward this goal, I create human-like agents, design autonomous adaptive defense strategies, and complementary human-AI teaming paradigms. My work involves **game theory**, **reinforcement learning**, **cognitive modeling**, and **behavioral experiments**.






## Education

---

- 2021 – Present     **Ph.D. Societal Computing**, School of Computer Science, Carnegie Mellon University.  
Advisor: *Cleotilde Gonzalez, Fei Fang*  
Committee: *Cleotilde Gonzalez, Fei Fang, Christian Lebiere, Prashanth Rajivan, Tiffany Bao*  
Thesis topic: *Human and AI Decision-Making in Cybersecurity: A Multiagent Modeling Perspective*.
- 2019 – 2021     **M.Sc. Information Technology**, Information Networking Institute, Carnegie Mellon University.
- 2015 – 2019     **B.S. Software Engineering**, Xi'an Jiaotong University.

## Honors and Awards

---

- 2025     **SCS Presidential Fellowship**, Carnegie Mellon University
- 2023-2024     **Women in Cybersecurity Student Scholarship**
- 2023     **Accelerating Foundation Models Research**, PI: Cleotilde Gonzalez,  
To advance the development and application of foundation models in AI
- 2018     **Foho Technical Innovation Grants**,  
For research excellence and academic performance, 1 of 77 awarded
- 2016-2018     **National Encouragement Scholarship**, Xi'an Jiaotong University.  
For strong academic performance, 3 of 77 awarded

## Research Publications

---

### Working Journal Articles

- 1    **Yinuo Du**, Palvi Aggarwal, Kuldeep Singh, Fei Fang, and Cleotilde Gonzalez (2024). *A Model of Human Behavior in Group Prisoner's Dilemma Games*. Under preparation.

### Journal Articles

- 1    **Yinuo Du**, Palvi Aggarwal, Kuldeep Singh, Fei Fang, and Cleotilde Gonzalez (2025). "Emergent Cooperative Behavior in Group Prisoner's Dilemma Games". In: *Acta Psychologica*. Under Review.
- 2    **Yinuo Du**, Baptiste Prebot, Tyler Malloy, Fei Fang, and Cleotilde Gonzalez (2024). "**Experimental Evaluation of Cognitive Agents for Collaboration in Human-Autonomy Cyber Defense Teams**". In: *Computers in Human Behavior: Artificial Intelligence*. Under review,  
[https://drive.google.com/file/d/1J2xKi8-GrwBYztlHVGacevL\\_DmDSpjnQ/view?usp=drive\\_link](https://drive.google.com/file/d/1J2xKi8-GrwBYztlHVGacevL_DmDSpjnQ/view?usp=drive_link).

- 3 **Yinuo Du**, Baptiste Prebot, Tyler Malloy, and Cleotilde Gonzalez (2024). “A Cyber-War Between Bots: Cognitive Attackers are More Challenging for Defenders than Strategic Attackers”. In: *ACM Transactions of Social Computing*. Just Accepted, [https://drive.google.com/file/d/19J2JxneqbtXRzwJKzMW6j3L8wIVdEEvn/view?usp=drive\\_link](https://drive.google.com/file/d/19J2JxneqbtXRzwJKzMW6j3L8wIVdEEvn/view?usp=drive_link).
- 4 Prebot, Baptiste, **Yinuo Du**, and Cleotilde Gonzalez (2023). “Learning about simulated adversaries from human defenders using interactive cyber-defense games”. In: *Journal of Cybersecurity* 9.1.

## Books and Chapters

- 1 Miah, Mohammad Sujan, Palvi Aggarwal, Marcus Gutierrez, Omkar Thakoor, **Yinuo Du**, Oscar Veliz, Kuldeep Singh, Christopher Kiekintveld, and Cleotilde Gonzalez (2022). *Diversifying Deception: Game-Theoretic Models for Two-Sided Deception and Initial Human Studies*.

## Conference Papers

- 1 Aggarwal, Palvi, Saeefa Rubaiyet Nowmi, **Yinuo Du**, and Cleotilde Gonzalez (2024). “Evidence of Cognitive Biases in Cyber Attackers from An Empirical Study”. In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*.
- 2 **Yinuo Du**, Prashanth Rajivan, and Cleotilde Gonzalez (2024). “Large Language Models for Collective Problem-Solving: Insights into Group Consensus”. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46 (o).
- 3 Malloy, Tyler, **Yinuo Du**, Fei Fang, and Cleotilde Gonzalez (2023a). “Accounting for Transfer of Learning Using Human Behavior Models”. In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*. Vol. 11. 1, pp. 115–126.
- 4 **Yinuo Du**, Palvi Aggarwal, Kuldeep Singh, and Cleotilde Gonzalez (2022). “Modeling of Multi-Defender Collaboration in a Cyber-Security Scenario”. In: *Proceedings of the Annual Joint Meeting of the Society for Mathematical Psychology and the International Conference on Cognitive Modeling*. URL: [https://drive.google.com/file/d/19J2JxneqbtXRzwJKzMW6j3L8wIVdEEvn/view?usp=drive\\_link](https://drive.google.com/file/d/19J2JxneqbtXRzwJKzMW6j3L8wIVdEEvn/view?usp=drive_link).
- 5 **Yinuo Du**, Baptiste Prébot, Xiaoli Xi, and Cleotilde Gonzalez (2022). “Towards autonomous cyber defense: predictions from a cognitive model”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 66. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1121–1125.

## Workshop Papers and Extended Abstracts

- 1 **Yinuo Du**, Baptiste Prebot, and Cleotilde Gonzalez (2024). “Turing-like Experiment in a Cyber Defense Game”. In: *Proceedings of the AAAI Symposium Series*. Vol. 3. 1, pp. 547–550.
- 2 Malloy, Tyler, **Yinuo Du**, Fei Fang, and Cleotilde Gonzalez (2023b). “Generative environment-representation instance-based learning: a cognitive model”. In: *Proceedings of the AAAI Symposium Series*. Vol. 2. 1, pp. 326–333.
- 3 **Yinuo Du**, Zimeng Song, Stephanie Milani, Cleotilde Gonzales, and Fei Fang (2022). “Learning to play an adaptive cyber deception game”. In: *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems. Auckland, New Zealand*. Vol. 6.
- 4 Aggarwal, Palvi, **Yinuo Du**, Kuldeep Singh, and Cleotilde Gonzalez (2021). “Decoys in cybersecurity: an exploratory study to test the effectiveness of 2-sided deception”. In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*.

## Invited Talks and Events

---

- 2024     **Invited Lightning Talk** *CMU Industry-Academia Partnership (IAP) Workshop*  
          **Invited Participant** *Human-AI Teaming for Decision-Making Workshop*  
          **Selected Speaker**, Turing-like Experiment in a Cyber Defense Game, *AAAI Spring Symposium on Human-Like Learning*  
          **Selected Speaker**, Human-AI Team Defense Game, *Women in Cybersecurity (WiCyS)*
- 2023     **Invited Panelist**, Multi-defender collaboration for Threat Intelligence Sharing, *The Future of Cyber Deception Workshop*  
          **Invited Speaker**, Using Cognitive Agents to Collaborate with Cyber Defenders: Current Work and Major Challenges, *INFORMS*  
          **Invited Speaker**, Human-AI Teaming for Cyber Defense, *Ellis-DDMLab Workshop*  
          **Selected Speaker**, Cognitive Modeling of Attackers, *Women in Cybersecurity (WiCyS)*
- 2022     **Invited Speaker**, Learning about Attackers through Interactive Cyber Defense Game, *Cylab Partners Conference*
- 2021     **Invited Speaker**, Cognitive Modeling of Attackers, *Cylab Partners Conference*

## Teaching Experience and Preparation

---

- 2024 - Present     **Eberly Future Faculty Program Participant**  
                          *Learn course design principles and pedagogy through seminars, receive feedback on teaching through teaching feedback consultations, and complete a course & syllabus design project.*
- 2024 Spring     **Teaching Assistant**, 88-312 Decision Models and Games
- 2023 Fall     **Teaching Assistant**, 17-759/17-599 Advanced topics in Machine Learning and Game Theory  
                          *Designed programming assignment on strategic language agent*

## Mentoring

---

### Masters Students

- 2024-present     Saeefa Rubaiyet Nowmi, *University of Texas El Paso*  
                          *Thesis project in Attacker Modeling*
- 2024     Adam Hunt, *Carnegie Mellon University*  
                          *Course project in AI for Social Good*
- Pau Balcells Sanchez, *Carnegie Mellon University*  
                          *Course project in AI for Social Good*

### Undergraduate Students

- 2024-present     Rony Dahnal, *Elon University*  
                          *Research Experiences for Undergraduates in Software Engineering (REUSE) Program*  
                          *Reinforcement learning for Cyber Deception*
- Neil Ramen, *Carnegie Mellon University*  
                          *CMU Path to AI Program*
- 2023-2024     Mason Kim, *Carnegie Mellon University*  
                          *Independent study on Interactive Cyber Defense Games*

## Service and Leadership

---

### Community Service

2024 Student Secretary of Human Factors and Ergonomics Society's Cyber Technical Group

### Conference Reviewing

2025 AAAI Conference on Artificial Intelligence (AAAI) Demo Track

2024 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)

2023-2024 Annual Meeting of the Cognitive Science Society (CogSci)

2022 Conference on Decision and Game Theory for Security (GameSec)

### Departmental Service

2024 SCS Teaching Awards committee

Societal Computing PhD Admission Committee

REUSE Admission Committee

AI-SDM Student Leadership Council

2023-2024 S3D Distinguished Speaker Selection Committee

2022-2023 S3D Women & Non-Binary & Trans Lunch

### Others

2023-2024 CMU Paths to AI Research Mentoring Program

2022-2024 CMU Graduate Application Support Program

2022 Workshop for Undergraduates in Computer Science (OurCS)

## Other Research and Industry Experience

---

2020 **Independent Study**, Dynamic Decision Making Lab, Carnegie Mellon University

*Advised by Palvi Aggarwal*

**Research Student**, Mobile, Embedded, & Wireless Security Lab, Carnegie Mellon University

*Advised by Patrick Tague*

**Software Engineer Intern**, BlockApps Inc.

2018 **Research Intern**, Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University

*Advised by Jing Tao*

## References

---

### Cleotilde Gonzalez

Full Research Professor  
Carnegie Mellon University,  
✉ coty@cmu.edu

### Prashanth Rajivan

Assistant Professor  
University of Washington,  
✉ prajivan@uw.edu

### Fei Fang

Associate Professor  
Carnegie Mellon University,  
✉ feifang@cmu.edu

### Palvi Aggarwal

Assistant Professor  
University of Texas El Paso,  
✉ paggarwal@utep.edu

### Christian Lebiere

Research Scientist  
Carnegie Mellon University,  
✉ cl@cmu.edu