UNIVERSITY
OF TECHNOLOGY
SYDNEY

[Differential Privacy in Reinforcement Learning]

by [Sheng Shen]

Thesis submitted in fulfilment of the requirements for the degree of

[C02029: Doctor of Philosophy]

under the supervision of [Prof. Tianqing Zhu & Dr. Bo Liu]

University of Technology Sydney
Faculty of [Engineering & IT]

[22-Dec-2022]

Certificate of Original Authorship Template

Graduate research students are required to make a declaration of original authorship when they submit the thesis for examination and in the final bound copies. Please note, the Research Training Program (RTP) statement is for all students. The Certificate of Original Authorship must be placed within the thesis, immediately after the thesis title page.

Required wording for the certificate of original authorship

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, *Sheng Shen*, declare that this thesis is submitted in fulfilment of the requirements for the award of *C02029: Doctor of Philosophy*, in the *School of Computer Science, Faculty of Engineering & IT* at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.
**If applicable, the above statement must be replaced with the collaborative doctoral degree statement (see below).*

**If applicable, the Indigenous Cultural and Intellectual Property (ICIP) statement must be added (see below).*

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature: Signature removed prior to publication.

Date: 22/12/2022

Collaborative doctoral research degree statement

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with *[insert collaborative partner institution]*.

Indigenous Cultural and Intellectual Property (ICIP) statement

This thesis includes Indigenous Cultural and Intellectual Property (ICIP) belonging to *[insert relevant language, tribal or nation group(s) or communities]*, custodians or traditional owners. Where I have used ICIP, I have followed the relevant protocols and consulted with appropriate Indigenous people/communities about its inclusion in my thesis. ICIP rights are Indigenous heritage and will always remain with these groups. To use, adapt or reference the ICIP contained in this work, you will need to consult with the relevant Indigenous groups and follow cultural protocols.

C02029: Doctor of Philosophy
CRICOS Code: 00099F
PhD Thesis: Computer Science
December 2022

Differential Privacy in Reinforcement Learning

Sheng Shen

School of Computer Science
Faculty of Eng. & IT
University of Technology Sydney
NSW - 2007, Australia

Differential Privacy in Reinforcement Learning

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy
in
Computer Science

by

Sheng Shen

to

Center for Cyber Security and Privacy
School of Computer Science
Faculty of Engineering and Information Technology
University of Technology Sydney
NSW - 2007, Australia

December 2022

Abstract

Reinforcement learning is a principled AI framework for autonomously experience-driven learning. The primary goal of reinforcement learning is to train autonomous agents to learn the optimal behaviors for their interactive environments. Deep reinforcement learning promotes a higher-level understanding of the visual world in the field of reinforcement learning by combining deep learning models and reinforcement learning algorithms. Since reinforcement learning is achieving great success in an increasing number of application fields that may involve huge amounts of private information, the security of policies and privacy preservation in reinforcement learning have given rise to widespread concerns. In addition, deep reinforcement learning policies parameterized by neural networks have been demonstrated to be vulnerable to adversarial attacks in supervised learning settings. Privacy leakage also occurs in multi-agent reinforcement learning systems where agents' actions or behaviors are directly exposed to other agents.

To address these multiple privacy concerns in reinforcement learning, we apply differential privacy in variant scenarios of reinforcement learning. In this thesis, we introduce our differentially private methods in those diverse scenarios to preserve privacy, including the multi-agent advising framework, multi-agent planning framework, the deep reinforcement learning context, machine learning classifiers and multi-agent game theoretic framework, respectively. We have provided detailed theoretical analysis and comprehensive experimental results to demonstrate that our methods can guarantee privacy preservation as well as the utility of reinforcement learning in diverse scenario in different chapters.

DEDICATION

Dedicated to my love, Maiying, how bravely tolerated all my stubbornness, temper and craziness when things did not go as expected. I am so grateful for your understanding of my choice of research, and your constant love and support in my life. You are always my spiritual support. I love you.

Acknowledgments

I have received much support and assistance throughout the writing of this thesis. I would first like to acknowledge my principle supervisor Prof. Tianqing Zhu who provided much helpful and insightful advice at any time required. My fantastic journey of research began with your trust in my potential and the opportunities you offered. I am grateful to Prof. Wanlei Zhou who funds my scholarship during my PhD. Also thank Dr. Bo Liu to share with me your professional knowledge and research experience. Next, I wish to extend my thanks to Dr. Dayong Ye who directed, guided and co-authored with me on many works. I also acknowledge my colleagues at the Center for Cyber Security and Privacy, UTS, who showed enthusiasm for my research and contributed their talent in the teamwork.

Last but not least, thanks to my partner and mother who endured this long journey with me, always offering support and love. Thank you for always being my strongest support and offering me the courage to move forward.

List of Publications

1. Sheng SHEN, Tianqing ZHU, Dayong YE, Mengmeng YANG, Tingting LIAO & Wanlei ZHOU. (2019, December). Simultaneously advising via differential privacy in cloud servers environment. In International Conference on Algorithms and *Architectures for Parallel Processing* (pp. 550-563). Springer, Cham.

2. Sheng, SHEN, Tianqing ZHU, Dayong YE, Minghao WANG, Xuhan ZUO & Andi ZHOU. (2022). A novel differentially private advising framework in cloud server environment. *Concurrency and Computation: Practice and Experience*, 34(7), e5932.

3. Dayong YE, Tianqing ZHU, Sheng, SHEN, Wanlei ZHOU & Philip YU (2020). (2020). Differentially private multi-agent planning for logistic-like problems. *IEEE Transactions on Dependable and Secure Computing*.

4. Sheng SHEN, Dayong YE, Tianqing ZHU, & Wanlei ZHOU. (2022). Privacy Preservation in Deep Reinforcement Learning: a Training Perspective. Submitted to *IEEE Transactions on Cybernetics*.

5. Dayong YE, Sheng, SHEN, Tianqing ZHU, Bo LIU & Wanlei ZHOU. (2022). One Parameter Defense-Defending Against Data Inference Attacks via Differential Privacy. *IEEE Transactions on Information Forensics and Security*, 17, 1466-1480.

6. Dayong YE, Tianqing ZHU, Sheng SHEN, & Wanlei ZHOU. (2020). A differentially private game theoretic approach for deceiving cyber adversaries. *IEEE Transactions on Information Forensics and Security*, 16, 569-584.

7. Sheng SHEN, Tianqing ZHU, Di WU, Wei WANG, & Wanlei ZHOU. (2020). From distributed machine learning to federated learning: In the view of data privacy and security. *Concurrency and Computation: Practice and Experience*.

8. Xin CHEN, Tao ZHANG, Sheng SHEN, Tianqing ZHU, & Ping XIONG. (2021). An optimized differential privacy scheme with reinforcement learning in VANET. *Computers & Security*, 110, 102446.

9. Yuao WANG, Tianqing ZHU, Wenhan Chang, Sheng SHEN, & Wei REN. (2020, November). Model Poisoning Defense on Federated Learning: A Validation Based Approach. In *International Conference on Network and System Security* (pp. 207-223). Springer, Cham.

Table of Contents

LIST OF FIGURES

FIGURE **Page**

List of Tables