Privacy and Evolutionary Cooperation in Neural-network-based Game Theory

Zishuo Cheng, Dayong Ye, Tianqing Zhu, Wanlei Zhou
University of Technology Sydney, School of Computer Science

Abstract

Privacy preserving is an essential requirement in cooperative multi-agent systems. However, it is challenging to balance the trade-off between preserving privacy and promoting cooperation of agents, especially in the domain of evolutionary games.

Existing methods have two common drawbacks:
1. Most existing methods did not take the privacy of agents into account. Due to this ignorance, adversarial agents can utilize other agents' private information to maximize their personal payoffs by minimizing other agents' payoff. However, this could lead to a decrease on the overall level of cooperation.
2. Most existing methods have a limited capability to be generalized in various domains, such as different types of networks. The performance of most existing methods strongly depends on some specific factors, such as network structures and the initial proportion of cooperative agents.

To overcome these, we propose a novel model which jointly adopts differential privacy mechanisms and neural networks. Differentially private (DP) mechanisms can be used to prevent agents' private information from being utilized by adversaries. Neural networks optimize agents' decision making in different types of networks. In this way, the proposed model can promote the evolution of cooperation in different domains even if adversarial agents exist.

Introduction

In artificial intelligence (AI), many tasks, such as evolution of cooperation in multi-agent social dilemma, multi-agent game and multi-agent control, also require agents to cooperate. However, as the raising concern of agents' privacy in AI, protecting the privacy of agents while promoting the level of cooperation has been a challenge in game theoretical multi-agent systems.

Two issues of existing methods in evolutionary games:

1. ***The protection of sensitive information***: On the one hand, each agent' actions and neighborhood information in the network can be considered as sensitive information which should be hidden from adversarial agents. On the other hand, the agents with limited information regarding other agents have the potential to lead to a dramatically low level of cooperation due to the partially exploration of the system state. Therefore, how to balance the trade-off between preserving the privacy of agents and promoting the level of cooperation is challenging.
2. ***The level of cooperation***: Previous mechanisms can achieve a high level of cooperation in some situations, while exhibiting a lower level in other situations. A situation consists of various factors, such as the proportion of cooperators, the structure of networks, and the state updating rule. For example, using Win-stay lose-shift rule, cooperation evolves in scale-free network when the initial fraction of cooperators is larger than 0.5, while not in random networks. Hence, the evolution of cooperation is easily affected by different situations.

Application Scenario: Iterated Prisoner Dilemma Games

Payoffs satisfy two conditions:
1. $T > R > P > S$, and
2. $2R > T + S$.

A higher value of T states more strict conditions for cooperation among agents in the long run.

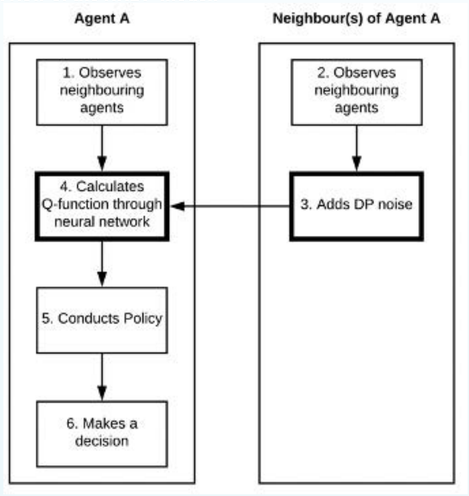
	Cooperate	***Defect***
Cooperate	***R***	***S***
Defect	***T***	***P***

Contributions of our research:

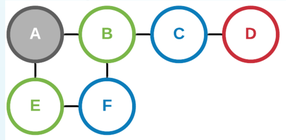
1) We are the first to propose a model that protects the privacy of agents in game theoretical multi-agent systems. By adopting DP mechanisms, the neural-network-based game with differential privacy (***NNDP***) model protects agents' sensitive information while sacrificing less performance on promoting the level of cooperation. Compared to encryption method, the *NNDP* saves a great volume of computational resources and time while provably guaranteeing agents' privacy.
2) We are the first to utilize a neural network to promote the level of cooperation in evolutionary games. By adopting deep multi-agent reinforcement learning algorithms, the *NNDP* improves the adaptivity and stability to resist the changes of conditions.
3) We theoretically proof that our proposed method satisfies the definition of differential privacy and implemented extensive experiments to examine the performance.

Methodology

Overview of the *NNDP* model:



An example of a network consisting of 6 agents:



True actions taken in last turn:
Agent A: Cooperate
Agent B: Cooperate
Agent C: Cooperate
Agent D: Cooperate
Agent E: Cooperate
Agent F: Defect

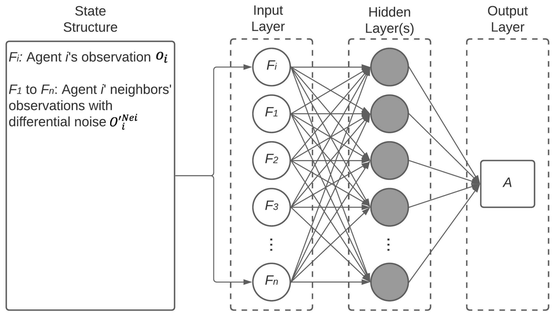
Sensitive information:

An observer's double-hop neighbors' information is sensitive that needs to be protected by DP mechanisms because this information is indirectly reported from the single-hop neighbors to the observer.

For example, **taking Agent A (grey) as the observer**, then:
1. Single-hop neighbors **(green)** : B and E
2. Multi-hop neighbors:
   1) Double-hop neighbors **(blue)**: C and F
   2) Other multi-hop neighbor **(red)**: D
3. **Green ones** transfer their observations on **blue ones** to help **Agent A's** decision making.
4. **Blue ones'** information is sensitive to **Agent A**.
5. **Red one** cannot be observed by **Agent A** anyway.

***NNDP* in static networks: DQN + Exponential DP**

1. Deep learning model: deep Q-learning model



2. Input state example:

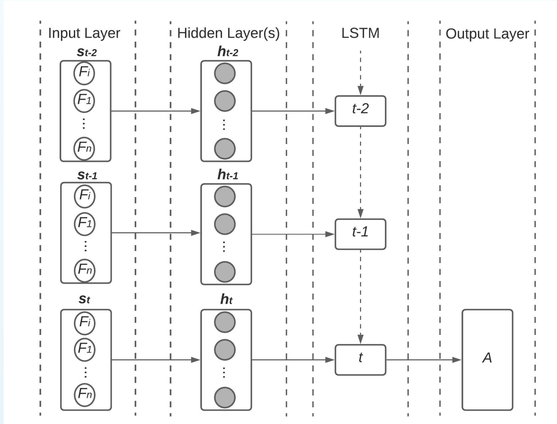
$S_A = \{O_A, O_B, O_C\}$ where
$O_A = \{a_A = Cooperate, a_B = Cooperate, a_C = Unknown, a_D = Unknown, a_E = Cooperate, a_F = Unknown\}$,
$O_B = \{a_A = Cooperate, a_B = Cooperate, a_C = Cooperate, a_D = Unknown, a_E = Unknown, a_F = Defect\}$,
$O_C = \{a_A = Cooperate, a_B = Unknown, a_C = Unknown, a_D = Unknown, a_E = Cooperate, a_F = Defect\}$.

3. Differential privacy model: exponential noise

Equation: $a'_z \leftarrow \exp(\frac{\epsilon \cdot Q(o_z, a_z)}{2\Delta S \cdot t})$ where a_z is the true action taken by the agent observed, ϵ is the whole privacy budget, ΔS is the sensitivity, and t is the time step.

***NNDP* in dynamic networks: DRQN + Laplacian DP**

1. Deep learning model: deep recurrent Q-learning model



2. Input state example:

$S_A = \{O_A, O_B, O_C\}$ where
$O_A = \{N_{Cooperate} = 3, N_{Defect} = 0\}$,
$O_B = \{N_{Cooperate} = 4, N_{Defect} = 1\}$,
$O_C = \{N_{Cooperate} = 2, N_{Defect} = 1\}$.

3. Differential privacy model: Laplacian noise

Equation: $L'_{a_k} \leftarrow L_{a_k} + Lap(\frac{\Delta S \cdot \ln t}{\epsilon})$ where L is the number of the single-hop neighbors taking the action a_k at each time step t.

Experiments

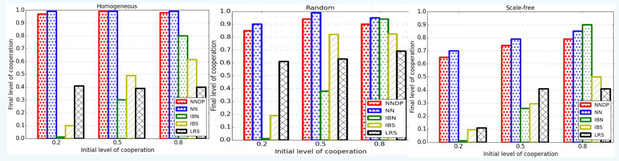
Benchmarks: (*NNDP* - **red**)
1. Neural-network-based (**blue**) reinforcement learning, where agents makes decision with only the proposed neural network.
2. Imitate-best-neighbor (**green**), where each agent imitates the action of the wealthiest agent (including itself) in the next round.
3. Imitate-best-strategy (**yellow**), where each agent adopts the strategy that accumulates the highest payoff in its neighborhood.
4. Local-redistribution-strategy (**black**), where wealthy agents in multi-agent systems share a fraction of their income with neighbors.

Network structure:
1. Homogeneous network (**left**): In a homogeneous network, each node has the same number of connections n.
2. Random network (**center**): In a random network, the connection between nodes is set with a connected probability p.
3. Scale-free network (**right**): A scale-free network has the property that a minority of nodes have many connections, while a majority have few connections.

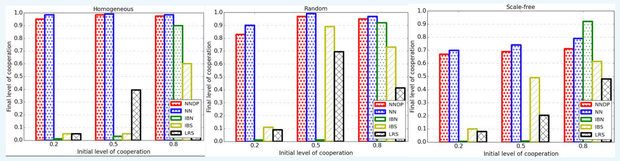
Result discussion:

According to the experimental results, the proposed *NNDP* method can achieve a desirable stability and adaptivity for the evolution of cooperation in static and dynamic networks. In terms of adaptivity, the evolution of cooperation can adapt to different initial proportion of cooperators and different types of static and dynamic networks; the level of cooperation is significantly higher than other the three mechanisms in most situations.

1. Adaptivity in static networks:



2. Adaptivity in dynamic networks:



Conclusion

This paper proposes a novel deep reinforcement learning model with differentially private mechanisms and relevance weights for game theoretical multi-agent systems. By adopting a customised neural network, the agents' adpativity to cooperative have been significantly improved. Moreover, by implementing differentially private mechanisms with a proper privacy budget, normal agents' privacy can be provably guaranteed and different but similar observation sent to neighboring agents can remain the level of cooperation to a large extent. Therefore, the proposed model is able to preserve agents' privacy while promoting the level of cooperation in game theoretical multi-agent systems.