A survey on attacks faced by deep reinforcement learning

Author: Parvathy Viswanath

Guest Lecturer, NSS college Ottapalm

ABSTRACT

Deep reinforcement learning can be defined as a combined form of reinforcement learning(RL) and deep learning .This field introduced solutions for different complex decision making tasks. It has an ability to adapt to the surrounding environments. Despite of having many advantages it also faces some attacks when we use them in real life applications. This paper includes some of the attacks faced and survey on the countermeasures can be taken to resolve them.

Key terms: Deep learning, reinforcement learning, artificial intelligence

1.INTRODUCTION

In the stream of 1artificial intelligence , machine learning provides automated methods that can detect patterns in data and can be used to get positive results in tasks[1]. Machine learning tasks can be divided into three different streams- Supervised learning , Unsupervised learning and reinforcement learning. Reinforcement learning is a process of learning and finding the sequence of actions in an environment in order to increase the cumulative rewards .In supervised learning, data is trained with corresponding labels and this combined form helps for the decision making. When compared to other methods , it is found to be the best well suited branch of machine learning and it is applied in various fields like speech recognition , spam detection etc.

In unsupervised , there is no presence of labels and each decision making process is made without knowing the true labels of the input data .Reinforcement learning is a procedure where the best actions to be performed by the agents is calculated on the basis of information that they observe while interacting with the environment surrounding it. Thus the later process make use of online data and observations obtained through real time interaction with the environmental factors. Deep learning is a human epitome of knowledge given to the computer systems. This system works almost like human brain with a capability of knowledge assimilation.

Recently a new technique was introduced by combining deep learning and reinforcement learning . Deep learning has shown results in complex decision making processes such as in designated task completion in robotics[2], health care[3] ,wireless and data management[4] etc. Adversarial attacks are like optical illusions for machines that an attacker has intentionally designed to cause the model to make a mistake .Deep reinforcement learning is found to be vulnerable to these types of attacks. This paper discuss about the survey on these attacks and some countermeasures that can be taken.

2.ADVERSARIAL ATTACKS

Adversarial attacks can be divided in to four different categories

1.Targeting the reward

As rewards has high importance in DRL, it can be targeted to get the advantage of it and it may show a wrong result. The reward attacks can be performed in different ways.

A. Attacks directly perturbing the reward
This attack was discussed by Han etal[5] , and proposed two kinds of attacks: flipping reward signals and manipulating states. Adversary can manipulate the reward by

flipping it to certain number of times and on other hand the states can be manipulated by making some changes in the first few steps of the training. On the basis of these attacks we can classify the attackers to four categories. 1) Omniscient attacker- the one who knows all information , 2)Peer attacker-one who is unknown about transition probabilities ,3)ignorant attacker –who knows only cost signals, 4)blind attacker-one who has no information.

B. Attack by disturbing states

"Robust deep reinforcement learning with adversarial attacks"[6] proposes different attacks on this type . Here the attackers creates perturbations to the observations. Once the states are disturbed it may never generate the expected reward.

C. Attacks by disturbing action space

X. Yeow Lee[7] explains different attacks by affecting the action space. These types of attacks can be classified in to two. One is for minimizing the cumulative reward with decoupled constraint and the second one with temporarily coupled constraints called look ahead action space attack.

2.Targeting the policy

Policies can also get attacked and this can be done in various ways.

A. Attacks on policy by affecting the states

Z.-W. Hong[8] proposed two different approaches where the policies get attacked by disturbing states. They are strategically timed attacks and enchanting attacks. In strategically timed attacks , subset of time steps in an episode of the DRL operation is used to reduce the reward. Enchanting attacks can be covered by offering the agent a predefined target state by using generative model and a sophisticated planning algorithm.

B. Attack on policy by effecting the environment.

L. Hussenot[9] discussed pre-observation attack and constant attack. Both of these attacks formally affects the environment. In pre-observation attack ,every observation of the agent get tampered and adding that disturbance to the environment. In constant attack includes a perturbation which is created at the start of the attack and it is added to all observations.

3. Targeting the observations

Environment can be targeted by either destroying the sensory data or sensors.This can be achieved by-

A. Attacks affecting states

These types of attacks are carried out in two different phases. At first, the training of Deep Q-networks on reward function to generate adversarial policy id done and this phase is known as initialisation. Exploitation phase includes the creation of different inputs that can lead the agent to follow the actions governed by the adversarial policies. Whereas another type of attack was proposed by Behzadan[10], where Deep Q-network policy is attacked by exploiting the transferability of samples.

4. Targeting the environment

The environment around agent can be also targeted. A detailed study on online sequential attack on environment was introduced by Xiao[11]. That study is carried by exploiting the temporal consistency of states and even their study proved that there is no need of back propagation here. Authors provide two method for this , namely adaptive based finite difference method and optimal frame selection method.

3.DEFENSIVE MECHANISMS AGAINST ATTACKS ON DRL

We have discussed different types attacks faced by DRL .Now some of the defensive mechanisms are explained

A. Robust Learning

This procedure ensures the robustness against adversarial attacks. This can be achieved by adding noise to the parameter state and it was found that this technique

was effective in mitigating the effects of training and test time. This was proposed by Behzadan[12].Later Mandelkar [13]proposed a better resilience by introducing an algorithm called ARPL(robust policy learning)algorithm.

Wasserstein robust reinforcement learning was the method introduced by Pinto et al[14].The policy they formulated with an objective of zero-sum minimax. There were other proposals which were using novel minmax game with convergent solver.

B. Adversarial Detection

This technique involves the detection of adversarial samples by using a model trained set to separate the true samples and so that we can discard adversarial ones and this procedure makes sure that the original model is not modified. The same procedure can be done by leverging an action conditioned frame prediction formula and it was formatted by Lin.et.al[8].

Advanced Q- learning algorithm for this action was proposed by Xiang[14] and this model is used to predict the inputs based on five different factors .This proposed model were used in automatic path finding robots. Energy point gravitation ,key point gravitation , path gravitation ,included angle and placid point are the some of influential factors.

Another proposed method generate a different matrix which generates rewards that can help reinforced agents to learn even in case of noisy inputs and such rewards are called as perturbed rewards. With the help of perturbed rewards unbiased reward estimator aided robust framework can be generated and proposal was made by Wang.et.al[15]

C. Defensive Distillation

This is considered to be the accurate method where the output probabilities of another model is trained on baseline.Rusu.et.al [16] explained a method of extracting the policy of dense network to train a lesser dense network .Lesser dense network is found take expert level decisions. Defensive distillation can be also made used in both networks and its expected entropy regularized distillation make it more promising.

D. Game approach

Past payoff observations that are subject to attack can be learnt and players can adjust their actions according to that .This observation can effectively work against attackers[19].Even minimax dynamic game framework can also generate some robust policies.

4. CONCLUSION.

This paper discussed about some of the adversarial attacks and some of the solutions can be carried out. As reinforcement learning is considered as growing technology , the attacks faced and the solutions that can be taken is having high importance. As the technology is growing more issues and attacks can be expected and those can be studied and resolved in the future.

5.REFERNCES

1. Vincent Francois-Lavent ,Riashant Islam, Joelle Pineau, Peter Henderson ,Marc.G.Bellemare ," An introduction to deep learning".

2. V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G.Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovskiet al., "Human-level control through deep reinforcement learning,"Nature, vol. 518, no. 7540, p. 529, 2015.

3. A. Raghu, M. Komorowski, L. A. Celi, P. Szolovits, and M. Ghassemi, "Continuous state-space models for optimal sepsis treatment-a deep reinforcement learning approach," arXiv preprint arXiv:1705.08422,

4. N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang,and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," arXiv preprint arXiv:1810.07862, 2018.

5. Y. Han, B. I. Rubinstein, T. Abraham, T. Alpcan, O. De Vel, S. Erfani,
D. Hubczenko, C. Leckie, and P. Montague, "Reinforcement
learning for autonomous defence in software-defined networking," in

Robust deep reinforcement learning with adversarial attacks," in Proceedings
of the 17th International Conference on Autonomous Agents
and MultiAgent Systems. International Foundation for Autonomous
Agents and Multiagent Systems, 2018, pp. 2040–2042.

7. X. Yeow Lee, S. Ghadai, K. L. Tan, C. Hegde, and S. Sarkar, "Spatiotemporallyconstrained action space attacks on deep reinforcemen learning agents," arXiv preprint arXiv:1909.02583, 2019.

8. Y.-C. Lin, Z.-W. Hong, Y.-H. Liao, M.-L. Shih, M.-Y. Liu, and M. Sun,"Tactics of adversarial attack on deep reinforcement learning agents," arXiv preprint arXiv:1703.06748, 2017.

9. L. Hussenot, M. Geist, and O. Pietquin, "Targeted attacks on deep
reinforcement 7learning agents through adversarial observations," arXiv
preprint arXiv:1905.12282, 2019.

10. V. Behzadan and A. Munir, "Vulnerability of deep reinforcement learning to policy tinduction attacks," in International Conference on Machine Learning and Data Mining in Pattern Recognition. Springer,
2017, pp. 262–275.

11. C. Xiao, X. Pan, W. He, J. Peng, M. Sun, J. Yi, B. Li, and D. Song,
"Characterizing attacks on deep reinforcement learning," arXiv preprint
arXiv:1907.09470, 2019.

12. B. Vahid and A. Munir, "Mitigation of policy manipulation attack on deep Q-networks with parameter-space noise," in International Conference on Computer Safety, Reliability, and Security. Springer,
2018, pp. 406–417.

13. A. Mandlekar, Y. Zhu, A. Garg, L. Fei-Fei, and S. Savarese, "Adversarially robust policy learning: Active construction of physicallyplausible perturbations," in 2017 International Conference on Decision and Game Theory for Security.
Springer, 2018, pp. 145–165.

6. A. Pattanaik, Z. Tang, S. Liu, G. Bommannan, and G. Chowdhary,
"
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2017, pp. 3932–3939

14. Y. Xiang, W. Niu, J. Liu, T. Chen, and Z. Han, "A PCA-based model
to predict adversarial examples on Q-learning of path finding," in 2018IEEE Third International Conference on Data Science in Cyberspace
(DSC). IEEE, 2018, pp. 773–780.

15. J. Wang, Y. Liu, and B. Li, "Reinforcement learning with perturbed rewards," arXiv preprint arXiv:1810.01032, 2018

16. A. Rusu, S. G. Colmenarejo, C. Gulcehre, G. Desjardins, J. Kirkpatrick,
R. Pascanu, V. Mnih, K. Kavukcuoglu, and R. Hadsell, "Policy
distillation," arXiv preprint arXiv:1511.06295, 2015

17. W. M. Czarnecki, R. Pascanu, S. Osindero, S. Jayakumar, G. Swirszcz,and M. Jaderberg, "Distilling policy distillation," in Proceedings ofMachine Learning Research, 2019, pp. 1331–1340.

18. M. Bravo and P. Mertikopoulos, "On the robustness of learning in
games with stochastically perturbed payoff observations," Games and
Economic Behavior, vol. 103, pp. 41–66, 2017

19. Inaam Ilahi_, Muhammad Usama_, Junaid Qadir, Muhammad Umar Janjua, Ala Al-Fuqaha, Dinh Thai Hoang, and Dusit Niyat," Challenges and Countermeasures for Adversarial Attacks on Deep Reinforcement Learning"