Int. J Rec. Adv. Sci. Tech., 2020; 7(2):1-6 **e-ISSN: 2395-2318**

DOI: 10.30750/ijrast.721

Network Security in the Age of Reinforcement Learning

Mohit Dayal[1], Rupender Duggal[2], Siddharth Aggarwal[3]
[1]Ambedkar Institute of Advanced Communication Technologies & Research
New Delhi, India
[2]Ambedkar Institute of Advanced Communication Technologies & Research
New Delhi, India
[3]Maharaja Agrasen Institute of Technology, New Delhi, India

***Correspondence:** Mohit Dayal, Ambedkar Institute of Advanced Communication Technologies & Research New Delhi, India. **E-Mail:** mohitdayal.md@gmail.com

Abstract
In the internet age, networks are everywhere. The ability to connect on a call, to browse the web, to push your content for viewers online, to send someone a message and to do a billion of other things comes from the availability of networks. As the networks have diversified and have become omnipresent, they have also become vulnerable and exposed. Thus, network security has become a huge priority for commercial and personal users. The concern of data leakage or stealing is not only a question of privacy but many also incur loss of money, property and sometimes even identity. Fortunately, along with the growth of computer networks, computational methods and practices such as machine learning, cryptography, block-chain etc. have also evolved and have become more efficient and applicable to solve real world problems. One such area of computer science, that is just breaking out of its shell is Reinforcement Learning. Reinforcement Learning is a field within Machine Learning which aims to make machine make intelligent and planned decisions. This paper discusses some of the recent works done in the field of network security using Deep Reinforcement Learning.

Keywords: Reinforcement Learning, Network Security.

Introduction

As the name itself suggests, a network is like a net which, similar to a spider's net, builds a path between two points. This path is the means to facilitate movement of data to and fro these points. A network has nodes which are the hosts that are a part of the network to share resources with other nodes. This sharing can be done through wired networks that use cables for the purpose of sharing data or through wireless networks using the electro-magnetic waves as the medium. The story of network started with a network which initially had just four nodes for the purpose of sharing educational information (Roberts, 1988). From there, technology has come a long road. In the twentieth century, computer networks are an indispensable part of society. Not only are they important for work and educational purposes but also vital for social and personal life. Hence, now, a major issue in designing a network is to ensure that it is safe, reliable and hack-proof.

Many methods have been developed to solve the network design issues to make it more secure and resilient to attacker. One such recent method makes use of Reinforcement Learning algorithms. Reinforcement Learning deals with an agent acting in an environment and performing actions in that environment (Sutton & Barto, 2018). These actions can either be good or bad for the agent which is decided by the reward that is received once the action has been taken. In general, the agent strives to maximize the rewards that it can get from that state onwards from the environment.

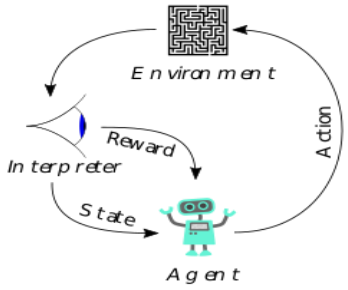

Figure1: Working of a Reinforcement Learning agent

In the past decade, there have been many advances in the field of Reinforcement Learning, especially with the combination of Reinforcement Learning with Deep Learning. This confluence between two important and

complex fields is called Deep Reinforcement Learning. Using deep reinforcement learning in applications like game play, robot navigation, autonomous driving and many other applications have produced breakthrough results that have further spurred the research of using Deep RL in other domains as well. Usage of Deep RL methods in Computer Networks have been widely studied in many fields of computer networks (Luong, et al., 2018). Many researchers have implemented Deep RL algorithms for improving Network Access, Data Offloading, Traffic Routing, Resource Scheduling, Rate Control and even Network Security. This paper reviews the recent algorithms that have been studied in the context of improving security of computer networks.

Interruptions to network security

A network can be jeopardized if an attack takes place to break the integrity of the network. Such an attack can be made to misuse the resources being shared on the network, to modify the information on network, to cause denial of service or other malicious intent. There are two major categories of attacks on the network. They are classified on the basis of how they affect the network. One category comprises of Active attacks which include all the attacks that directly disrupt the operations of the network. For instance, an SQL injection can harm the normal functioning of a network. Another category is of Passive attacks that aim to get a hold of the data that is being shared on the network. For example, an attacker might analyze the traffic of a network to gain hazardous information. In any case, they can be harmful to both the service providers of these networks and the users of these services. As the technology of networks and communication continues to evolve with advancements in speed, efficiency, low latency and increased security, new applications have arisen that make use of these improvements. From building smart devices using Internet of Things (IoT) to using networks to build Intelligent Transportation Systems(ITS), there have been a large variety of new use cases which have diversified the field of networks and communication, especially wireless communication and have led to new developments in network security.

In this paper we will conduct our analysis on two major types of attacks that undermine the security of these new technologies- jamming attacks and cyber physical attacks.

Jamming attacks

Jamming attacks are more of an issue in wireless communication than in wired networks. This is because the medium of communication is radio wave which can be easily intruded. Jamming is defined as the emission of radio signals aiming at disturbing the transceivers' operation(Adamy & Adamy, 2004). This type of attack is done intentionally against a particular target device to cause denial of service (DoS). Since radio frequency is an accessible medium that can be accessed by anyone, jamming becomes an easy and harmful attack. Jammers are the devices used for this purpose and they emit radio signals in the wireless medium. Emission of radio signals that have the same frequency as of the channel which the attacker wants to disrupt, will cause in loss, modification or corruption of data on that channel leading to loss in security of the network.

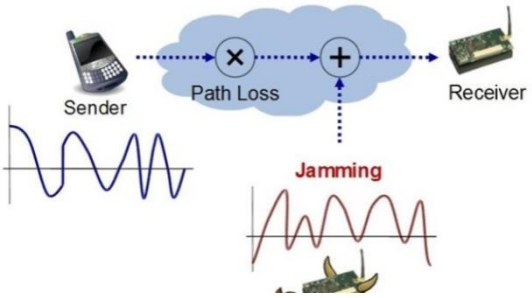

Figure 2: Network depicting Jamming (Krazytech, 2017)

The equation that helps in deciphering the influence of a jammer on a network is the jamming to signal ratio (J/S) equation.(Berg, 2008)

$$J/S = \left(ERP_J\right)\left(G_{RJ}\right)\left(d_S^2\right) / \left(ERP_S\right)\left(G_R\right)\left(d_J^2\right)$$

Equation 1. Jamming to Signal(J/S) Ratio

where,
ERPJ = the effective radiated power of the jammer (in any units);
ERPS = the effective radiated power of the desired transmitter (in the same units);
dJ = the distance from the jammer to the receiver (in any units);
dS = the distance from the desired transmitter to the receiver (in the same units);

GRJ = the gain of the receiving antenna toward the jammer (not in decibels);
GR = the gain of the receiving antenna toward the desired transmitter (not in decibels).

Cyber Physical Attacks

A Cyber Physical Systems (CPSs) are systems that engages computer software to perform physical tasks. They are hence the conflation of computation and physical processes (Lee, 2008). Consider for example, a modern car. It has a lot of embedded computers and networks that control and monitor the various aspects of a car such as the locks of the doors or automatic rain sensing wipers that automatically start to work when it begins to rain. It is easy to see that such automations are a result of communication between the hardware and software components that happen through computer networks that felicitate the control and communication in CPSs.

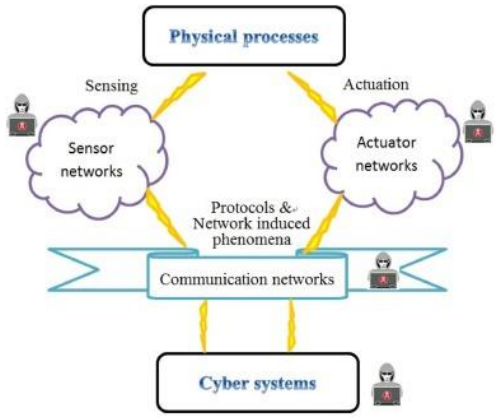

Figure 2: Architecture of CPSs (Ding, et al., 2018)

In certain cars, such as those with partial auto driving, speed of the car is regulated by the sensor readings such as from pre-collision radar sensors and many others. With the advent of Internet of Things (IoT), there are a large number of such devices that work by combining both sophisticated mechanical components and computer based algorithms. Thus, because of their popularity, security of CPSs has been a hot research area in recent years.

Attacks on CPSs

Attacks on Cyber Physical Systems amount to the intentional attacks that the attackers can make on the systems to cause damage. A system can be disturbed by an attacker by modifying the measurements of the sensor and injecting external control inputs (Mo & Sinopoli, 2012). Although CPSs are equipped with Failure detection systems(Han, et al., 2014) that sense abnormality and prevent attacks that intend to cause reading changes and other alterations, an attack can be planned in a way so as to fool the failure detection system. A smart grid – an electrical grid that manages electronic power distribution and production merged with smart meters and resources – which is a CPS, can suffer expensive damages if even the lowest intensity attack is made on the system. Security measures have to be therefore kept in place to ensure safe working of CPSs.

Deep Reinforcement Learning

Before we dive into some exciting applications and results of using Deep RL methods, let us look at what Deep RL actually is. In 2015, (Mnih, et al.), used Convolutional Neural Networks along with the most popular reinforcement learning algorithm Q- Learning and presented the Deep Q-Network (DQN) algorithm and showed the results of the algorithm in game playing. In their experiment, they showed that DQN's performance was superhuman when it came to playing Atari games from the Arcade Learning Environment. This prompted further research in the field and has led to a large number of organizations and researchers working on using Deep Reinforcement Learning on optimization problems where the task is to select the best action for a given state, one that leads to the maximum expected reward in the future.

Deep Reinforcement Learning for Jamming Attacks

Han, et al., in (2017), applied a Deep Q Networks (DQN) on cognitive radio networks and found that using DQN led to an increase in the signal-to-interference plus-noise ratio. Just like in any Reinforcement Learning problem, here there is an agent that can perform actions based on its current state and get rewarded or penalized for the same by the environment. In the Cognitive Radio Network, there is one Secondary User (SU) and more than one Primary Users (PU). There are also jammers or attackers trying to jam the network from SU to the Base station (BS). The SU can choose to stay in the geographical area and transmit signal to the same BS or to move to another geographical location if there is a chance of being intercepted in the first area. The agent in this case is the SU which can take any of these two actions. The SU has to take an action so as to avoid interfering with the PU.

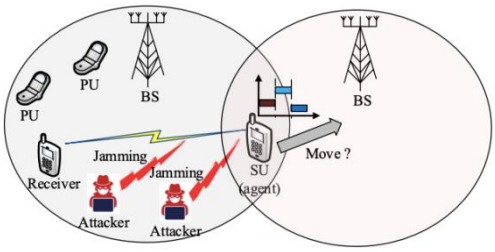

Figure 3: Architecture of the network in the work of Han, et al.

Applying DQN here requires a Convolutional Neural Network to be trained. The DQN algorithm updates a quality-function or Q-function for each state-action pair, which is the expected discounted long-term reward for state s and action x.

$$Q(\mathbf{s}, x) = \mathbb{E}_{\mathbf{s}'}\left[u_s + \gamma \max_{x'} Q(\mathbf{s}', x') | \mathbf{s}, x\right]$$

Equation 2: Q-Learning

A major caveat in this approach is that it discusses the approach only when there are two jammers. An improvement to this was found by (Xiao, et al., 2018). This approach used the same DQN algorithm but now in an underwater setting. They used Underwater Acoustic Networks (UANs) along with jammers to develop anti-jamming techniques to increase the jamming and intrusion endurance of underwater machines like robots and vehicles. Here again like in the previous work, an optimal policy is calculated using the algorithm to find the power with which the transmitter must send its signal and whether or not to change the location of the receiver based on the jamming. (Chen, et al., 2018) proposed a DQN based approach to tackle jamming in IoT devices. Like in the works of (Xiao, et al., 2018) and (Han, et al., 2017), they aimed to improve the signal-to-interference plus-noise ratio for IoT signals without being concerned of the architecture of the IoT device or jammer. Against jamming, they found that using DQN for IoT devices increased the SINR of the signals by 15.2% higher than the previous benchmarks which is a major improvement.

Deep Reinforcement Learning for Cyber Physical System Attacks

As mentioned above, Cyber Physical Systems can be attacked by sending forge sensor readings to the system that eventually make it take the wrong decision. Consider a self-driving car with airbags installed. Although these airbags inflate at the time of a collision, but an attacker can forge in data and the computer can be made to believe that there has been a collision and inflate the airbags when in fact that is not the case. Some of the starting work in CPSs security started with autonomous vehicles as this technology is at the forefront of research and innovation in this decade. (Ferdowsi, et al., 2018) formulated a game to study how deep reinforcement learning can be used to find the optimal actions for the autonomous vehicles. In this game, the authors developed a number of autonomous vehicles, with d_i being the space between vehicle *i* and *i-1*. For the system of autonomous vehicles to be successful, this distance should be enough to avoid collision and also be small enough so that there is good amount of flow of traffic at any time. Thus, the attacker's aim would be to manipulate the distance d_i so as to obstruct optimal flow of traffic in the system whereas the vehicle would try to minimize this change in distance which is also called as its regret function. In essence, both the Attacker and the Autonomous Vehicle(AV) want to optimize their policy and in doing so they use a Long Short Term Memory(LSTM) Neural network to find the Q values for different state action pairs in their state and action space.

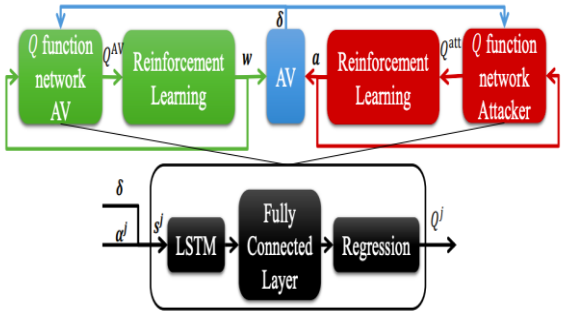

Figure 4. Model of network in Ferowsi, et al.

The results of their experiments showed that, using the deep RL algorithm, the AV was able to mitigate the effect of data injection attacks on the sensor data and thus stay robust to such attacks.

Another work that makes use of Long Short Term Memory (LSTM) to obstruct attacks in CPSs was done by (Ferdowsi & Saad, 2018). Similar to many researches, they used the approach of a game to study there results of deep reinforcement learning algorithm on an Internet of Things (IoT) platform. In a large scale IoT system, the devices are connected to the cloud for the purpose of connecting, processing, storing and analyzing data. Thus, extensive device-to-cloud communications take place for every device in the system. In such an IoT system, a gateway, which is a hardware that acts as the intermediary between device-

to-cloud communications, is used for management and also security purposes. When the devices are large in number, the gateway cannot authenticate all of the communications that take place in the IoT system because of the high computation required for this task. Thus it has to solve the problem of choosing which IoT device it should authenticate and those which it should not. This again is similar to finding the optimal policy where RL algorithms tend to perform well. This work aimed to dynamically predict the state of unauthenticated IoT devices and allow the gateway to decide on which device to authenticate using Deep RL. In their results they showed that using Deep RL approach reduced the number of devices whose security have to compromised by up to 30% which is a major improvement from the previous benchmarks.

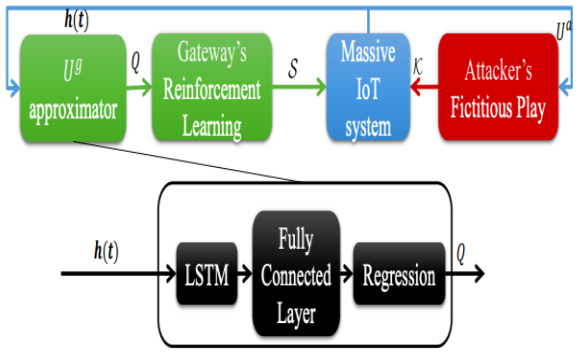

Figure6: Model of network in Ferowsi and Saad

Future Scope

Although only the use cases of network security have been considered in this paper, Deep RL has been applied into many areas of Networks and Communication. This has further prompted the research in many fields of Communication to try Deep RL algorithms wherever there is an optimal policy to be found. Some future research in the field of network security that seems to be not far away is the case of using multiple RL agents which has proven to be more efficient than using a single agent in many cases (Tan, 1993). Using Distributional RL algorithms such as C51(Bellemare, et al., 2017) and IQN(Dabney, et al., 2018) which have shown better results in almost all cases against the standard DQN algorithm can also be tried in the case of network security problems.

References

1. Adamy, D. L. & Adamy, D., 2004. EW 102: A Second Course in Electronic Warfare. s.l.: Artech House Publishers.
2. Bellemare, M., Dabney, W. & Munos, R., 2017. A distributional perspective on reinforcement learning. Ar Xiv.
3. Berg, J. S., 2008. Broadcasting on the Short Waves, 1945 to Today. s.l.:s.n.
4. Chen, Y., Li, Y., Xu, D. & Xiao, L., 2018. DQN-based Power Control for IoT Transmission Against Jamming. IEEE 87th Vehicular Technology Conference.
5. Dabney, W., Ostrovski, G., Silver, D. & Munos, R.́., 2018. Implicit Quantile Networks for Distributional Reinforcement Learning. arxiv.
6. Ding, D. et al., 2018. A survey on security control and attack detection for industrial cyber-physical systems. Neurocomputing.
7. Ferdowsi, A., Challita, U., Saad, W. & Mandayam, N. B., 2018. Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems. arxiv.
8. Ferdowsi, A. & Saad, W., 2018. Deep learning for signal authentication and security in massive internet-of-things systems. IEEE Transactions on Communications.
9. Han, G., Xiao, L. & Poor, H. V., 2017. Two-Dimensional Anti-Jamming Communication Based On Deep Reinforcement Learning. IEEE.
10. Han, S., Xie, M., Chen, H.-H. & Ling, Y., 2014. Intrusion Detection in Cyber-Physical Systems: Techniques and Challenges. IEEE Systems Journal .
11. Krazytech, 2017. Jamming and anti-Jamming Techniques. [Online] Available at: https://krazytech.com/technical-papers/jamming-and-anti-jamming-techniques
12. Lee, E. A., 2008. Cyber Physical Systems: Design Challenges.
13. Luong, N. C. et al., 2018. Applications of Deep Reinforcement Learning in Communications and Networking: A Survey. IEEE.
14. Mnih, V. et al., 2015. Human-level control through deep reinforcement learning. nature, pp. 529-533.
15. Mo, Y. & Sinopoli, B., 2012. Integrity attacks on cyber-physical systems. International conference on High Confidence Networked Systems, April.pp. 47-54.
16. Roberts, L., 1988. The arpanet and computer networks.

17. Sutton, R. & Barto, A., 2018. Reinforcement learning: An introduction. Vancouver: MIT press .
18. Tan, M., 1993. Multi-agent reinforcement learning: Independent vs. cooperative agents. international conference on machine learning.
19. Xiao, L. et al., 2018. Anti-Jamming Underwater Transmission With Mobility and Learning. IEEE.

Conflict of Interest: None
Source of Support: Nil