International Journal of Science, Engineering and Technology
An Open Access Journal

A Study on Applying Reinforcement Learning in Intrusion Detection and Prevention System

[1]R. Arunraj, [2]C. Umarani

Abstract

Attacking vectors are increasing day by day with various dimensions. Attacks use to happen against vulnerable services, data driven attack on application, host based attacks like unauthorized login, stealing sensitive information through malware like viruses, worms and Trojans. By this we forcibly loose our CIA (Confidentiality, Integrity, and Availability) mantra.

In order to prevent that we use Intrusion Detection and Prevention System (IDPS) as one of the solution. But our efforts are always tend to be placed behind an attacker. Applying artificial intelligence in IDPS could resolve this issue by having automated system to resolve threats in real time. This paper presents a clear study to proactively identify and mitigate new threat and also reduce false positive, false negative effectively by using Reinforcement learning in IDPS.

Introduction

Automation of intrusion system for an organization is essential and became most wanted in recent years. Reinforcement learning is a part of Artificial Neural Networks (ANN). ANN is a technique which used in machine learning to achieve Artificial Intelligence. The inventor of the first neurocomputer, Dr. Robert Hecht-Nielsen, defines a neural network as “a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.”

Intrusion detection and prevention system is a proactive approach to network security used to identify threat and respond to them in quickly manner. When an attacker gains access, IDPS has the ability to take immediate action. It has various detection methodologies such as signature based, anomaly based, stateful protocol analysis and network based. All together to this IDPS are also broadly classified into four sector such as network based which used to sniff packet and analyse network traffic, host based which used to monitor events occurring within that host, network behaviour analysis used to identify threat that generate unusual traffic flows, wireless based used to monitors wireless network traffic.

This paper focuses on finding out how to use IDPS effectively and reduce high false positives, false negative and zero day attack. Applying Reinforcement learning seems to be quite easy but implementation is not as easy as we think.

Reinforcement learning

Reinforcement learning(RL) is an art of machine learning which helps an agent to take best action by absorbing the environment so that it maximizes the cumulative reward.RL is formulated to solve the so-called Markov decision problem (MDP). In other words, the problem that we are trying to solve with RL should be an MDP or its variant. The theory of RL relies on artificial intelligence (AI) and dynamic programming (DP). With the help of this technique we are going to successfully implement this in IDPS. It’s important to understand the players who make this as possible.

(1) Agent - who interact with the packet
(2) State - Absorb the current state
(3) Action - Based on the state, apply action
(4) Reward- Give some value for the action

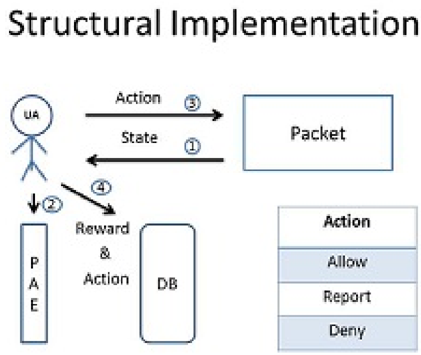

Figure 1

As shown in figure 1, first our IPS will analyse the state of the packet and then send to the PAE (Packet Analysis Engine).Once it get entered into the PAE it analyse each and every nook of the packet by separating each information of the packet and assign state for each information and then giving reward value for each state. At last our PAE will decide that what action it should take based on cumulative reward and state of the packet. Here comes the next issue that how this PAEanalyse the state and perform the action. To answer that we choosing Q-Learning algorithm as a solution.

Q-Learning Algorithm

The Q-Learning algorithm falls under the umbrella of active reinforcement learning whose main job role is to take an action by absorbing the environment.It can be used to find an optimal action-selection policy for any given (finite) Markov decision process (MDP). Q-learning effectively provides an output by working under the principle of learning an action-value function that predominantly gives the expected utility of taking a mentioned action in a given state and thereafter following the optimal policy. To define policy we can say that it is a rule helps to choose the best action from the given state it is in built by normally choosing the action with the maximum reward value in each state. The Q-learning algorithm learns an action-utility representation instead of learning utilities. The value of doing action 'a' in state's' is given by the Q-function, Q(s, a).

The relation between Q-values and utility values is given by: [7]

$U(s) = max_a Q(s, a)$[7]

Q-learning is a model-free method and hence it does not use transition function. The update equation for Q-learning is given by: [7]

$Q(s, a) = Q(s, a) + \alpha(R(s) + \gamma(max_{a'} Q(s', a') - Q(s, a)))$[7]

where a' is the action actually taken in state s' . The Q-learning algorithm is given in figure 2.

The agent absorbs the current state s' and reward r' and returns an action based on the Q-table that stores action values indexed by state and action. The Q-value updating is done according to the above mentioned update equation. This updating process happens continuously until we reach an equilibrium i.e. we should not see change in the Q-values for two consecutive iterations. With the help of above processed Q-values, we select the action and therefore making user agent to reach its final stage. Through this way Q value tells the agent that what action it should take on time. According to our scenario we limiting our action in three categories such as to (1) Allow, (2) Report and (3) Deny as shown in figure1.

As shown in figure 2, we can clearly see that our packet analysis engine effectively inspect each and every corner of the packet as like state full inspection. With the help of Q-learning algorithm our PAE analyse the state of the packet as well as with whom it need to communicate in the network. By doing so we can able to handle the packet effectively.

```
function Q-LEARNING-AGENT(percept)
  inputs: percept, a percept indicating the current state s' and reward signal r'

  persistent: Q, a table of action values indexed by state and action, initially zero
              N_{s,a}, a table of frequencies for state-action pairs, initially zero
              s, a, r, the previous state, action and reward, initially null

  if TERMINAL?(s) then
     Q[s, None] <- r'
  end if
  if s is not null then
     increment N_{s,a}[s, a]
     Q[s, a] <- Q[s, a] + α(r + γ.max_{a'} Q[s', a'] − Q[s, a])
  end if
  s, a, r <- s', argmax_{a'} f(Q[s', a'], N_{s,a}[s', a']), r'
return a
end function
```

Figure 2: The Q-Learning algorithm that learns the value Q(s, a) of each action in each situation[7].

Conclusion

In this paper, I mentioned some common study techniques of Reinforcement learning and Q-learning algorithm with respect to the proposed IDPS which helps to see the best automated service on relate to subject topic. By using Q-learning algorithm we can identify the packet state as well as what it tries to do in the network. The user agent will take the decision according to the situation prevailing in the environment. In further studies I am trying to explore that how to refine the Packet Analysis Engine (PAE) effectively.

References

1. R Shekar, K perumal, Department of Computer Applications, Madurai Kamaraj University, Madurai,S Vijaya Rani, Assistant Professor, MCA Department, Brindavan College, Dwaraka Nagar, "Analysis of Next Generation Intrusion Prevention System Using Sensor Fussion and Fuzzy Logic".

2. Muhammad Imran Shafi, Muhammad Akram, Sikanadar Hayat and Imran Sohail, "Effectiveness of Intrusion Systems (IPS) in Fast Networks".

3. Why Artificial Neural Networks (ANN) Technology offers a promising future in IDS/IPS

http://resources.infosecinstitute.com/why-artificial-neural-networks-ann-technology-offers-a-promising-future-in-idsips/#gref

4. Artificial Intelligence – Neural Networks

https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_neural_networks.htm

5. Arturo Servin and Daniel Kudenko, "Multi-Agent Reinforecement Learning for Intrusion Detection".

6. Vaibhav Mohan, June 2014, "A comparison of various reinforcement learning algorithms to solve racetrack problem".

7. Abhijit Gosavi, Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology, February 11, 2017, "A Tutorial for Reinforcement Learning".

8. Markov Decision Processes and Reinforcement Learning

https://danieltakeshi.github.io/2015-08-02-markov-decision-processes-and-reinforcement-learning/

.9.Network Security: Prevalence Of Intrusion Detection System-C.Umarani, ISBN (13 Digits) : 978-81-910024-3-0, ISBN (10 Digits) : 81-910024-3-4

10.Application of Genetic Algorithm in Network Intrusion Detection-C.UMARANI, ISBN: 978-81-207-8818-3

11.Genetic Algorithm and its implementation in Network Intrusion Detection with MATLAB. - C.UMARANI

12. Fuzzy Neural Networks with GA for Network Intrusion Detection System- C.UMARANI

Author's details

[1]MCA-ISMS Student, Jain University, Bangalore

[2]Professor,Jain University,Bangalore