**Reinforcement Learning approach in penetration testing.**

Research Questions:

* Can RL be used to automate the penetration testing?

Experiment:

* Create network model made up of subnetworks, machines and firewalls between each subnetwork. Fig 3.2.1
* Environment to be used: NAS simulator (<https://github.com/Jjschwartz>
* Construct custom scenarios for penetration testing.
* Train RL agents and measure the performance of attack policies in those scenarios.
* Use and compare different RL algorithm. Q learning, Epsilon greedy, etc

Results should explain:

1) Can RL be used to find an attack path through a network when one exists?

2) How optimal is the attack path generated by the RL agent, as defined by the reward

function?

3) How does RL scale with increased network size and number of exploits?

4) How general is the RL approach? Does it work on different network configurations?

5) How do the different RL approaches compare?

6) Assumptions

Possible problems:

* large combinatorial and changing action space - each exploit and scan program used for pentesting typically has lots of parameters, similarly the number of computers on a network can be large and vary a lot and may not be known beforehand. These properties can be a huge issue for existing RL and planning methods since models must deal with changing input/output sizes and also exploration becomes very difficult.
* partial information - typically in pentesting you only recieve partial information from each action, so need methods for dealing with this
* parsing observations into useful features - in general the observations received during pen testing are text in a terminal or maybe text in JSON format. To me it seems it would be extremely difficult (if not impossible without astronomical compute and data) to train a model on this raw data, so finding a way to extract useful features that an agent could then train is a real challenge.
* When I was doing research, this was always handled manually, but having a way to handle this automatically would be ideal.

Report: 15000 words with references.