深層強化学習に基づくペネトレーションテスト手法の提案

A penetration testing method based on deep reinforcement learning

米田智紀・システム分科会・情報セキュリティ大学院大学

abstract: Penetration testing is a method of testing for security vulnerabilities by attempting to penetrate a device or system using a variety of techniques. The penetration tester selects and executes the necessary commands based on the system's response. This process is very similar to that used to solve text-based games such as TextWorld. In this study, we propose a penetration tester based on a neural agent that efficiently finds the goal of TextWorld using deep reinforcement learning based on partially observed Markov decision processes (POMDPs). In addition, while existing neural agents are based on GRUs, the proposed method proposes a system that introduces GTrXL, which has an internal attention mechanism, into the model that estimates the state of the neural agent. We have conducted experiments by running these systems in OS environments such as Linux and using exploit commands in the action set, and have succeeded in demonstrating the superiority of the proposed system over conventional models.

研究の背景 ・ 目的

■背景

近年、サイバー空間における驚異の増大に伴い、強化学習を基にしたペネトレーションテスト手法が注目されている。ASAPやDeepExploitなど様々な手法が提案されている。

既存の研究の問題点

・シミュレーション環境での実験
・脆弱性及び対応する攻撃方法が既知

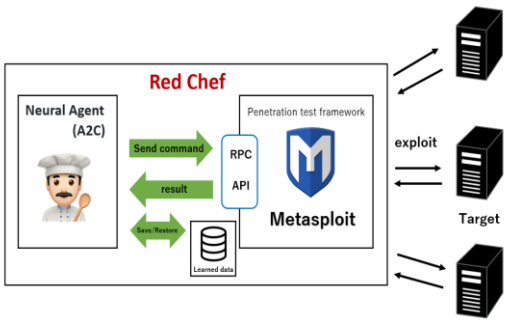
■目標

本研究では、POMDPを基にしたニューラルエージェントを使用して、コマンド実行時の結果から自然言語処理によって解釈し、状態を推定しながら次の最適な攻撃コマンドを選択するペネトレーションテストシステムを提案する

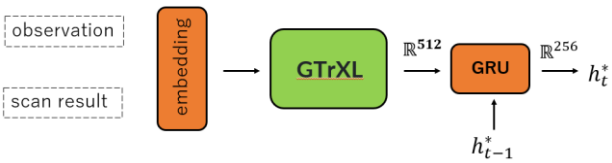
提案手法1

提案手法のシステムはニューラルエージェントとMetasploitやNmapといったツールが連携して、ターゲットサーバに攻撃を行う。 Nmapの実行結果、エージェントの行動の結果のテキスト情報を部分観測情報としてニューラルエージェントに入力する。行動は事前に実行されるNmapとディレクトリ探索によって獲得される。本提案手法ではMetasploit上で使用する以下の4種類の行動の最適化を行う。

・攻撃ポート番号の選択　（数ポート～数十ポート）
・攻撃ターゲットの選択　（1～数十）
・攻撃モジュールの選択　（数十～数千）
・攻撃ペイロードの選択　（数十～数百）



提案手法2

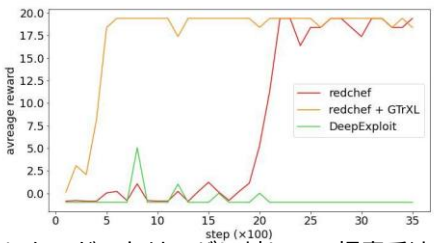


POMDPエージェントにとって状態推定の精度向上は極めて重要である。そこで、ニューラルエージェント状態推定の精度を向上させるために、注目すべき表現を絞り込むことのできるAttention機構を持つtransformerベースの手法であるGTrXLを導入する。既存のニューラルエージェントではテキスト情報をembeddingした後にGRUでエンコードを行っていたが本提案手法ではGTrXLエンコーダに変更している。

実験

実験1:提案手法であるRedchefとRedchef+GTrXL、既存手法であるDeepExploitとの性能評価を行った。対象サーバは脆弱なLinux仮想イメージであるMetasploitable2を用いた

実験1結果



実験2:マルチなターゲットサーバに対して、提案手法であるRedchefとRedchef+GTrXLにおける性能評価を行った。対象サーバはMetasploitable2を含む脆弱なサーバを6種類準備した。

実験2結果