https://doi.org/10.7840/kics.2022.47.11.1776

강화학습 기반 자율형 사이버 위협 공격 COA 탐색 기술 동향

손 석 빈◆, 이 해 민*, 박 수 현*, 김 동 화**, 김 중 헌°

Trends in Autonomous Course of Action for Cyber Attacks Using Reinforcement Learning

Seok Bin Son◆, Haemin Lee*, Soohyun Park*, Dong Hwa Kim**, Joongheon Kim°

요 약

네트워크의 규모가 커지고 복잡해짐에 따라 악의적인 공격자에 의한 보안 위협은 지속적으로 증가하고 있다. 잠재적인 사이버 공격에 대처하기 위해 다양한 테스트 방법이 개발되어 타깃 네트워크의 보안을 평가하고 실제 적대자를 생성하여 일련의 전략적 공격 행동(COA, Course of Action)을 탐색한다. 그러나 현재의 공격 COA 기술에는 특히 크고 복잡한 네트워크의 경우 비용이 많이 들고 광범위한 인적 노력이 필요하다는 점이 있다. 또한, 불확실한 네트워크의 상태에 대한 자율적 공격의 필요성이 대두되고 있다. 이러한 맥락에서, 강화학습 기반 접근법은 유망한 해결책으로, 본 논문에서는 강화학습 기반 공격 COA 전략의 최신 연구 동향과 그 한계점을 소개한다.

키워드 : 사이버 공격, 공격 COA, 강화학습, 공격자동화, 침투테스팅
Key Words : Cyber Attacks, Attack COA, Reinforcement Learning, Penetration Testing

ABSTRACT

As the network is getting large and complex with massively connected endpoints, the security threats from malicious adversaries have continuously increased. Various testing methods have been developed for the prior security evaluation of a target network to derive a series of strategic attack actions that we call Course of Action (COA) by emulating real active adversaries. However, current attack COA techniques require extensive human effort and cost, especially for large and complex networks, and the need for autonomous attack COA on the uncertain state of the target network has emerged. In this context, the reinforcement learning-based approach is regarded as a promising solution. This paper introduces an overview of the state-of-the-art research in RL-based attack COA strategies with the remaining limitations.

※ 이 논문은 국방과학연구소의 지원으로 수행된 연구임(UI210009XD).
◆ First Author : Korea University School of Electrical Engineering, lydiasb@korea.ac.kr, 학생회원
° Corresponding Author : Korea University School of Electrical Engineering, joongheon@korea.ac.kr, 종신회원
* Korea University School of Electrical Engineering, haemin2@korea.ac.kr, 학생회원
* Korea University School of Electrical Engineering, soohyun828@korea.ac.kr, 학생회원
** Agency for Defense Development(ADD), dhkim@add.re.kr
논문번호 : 202209-219-B-RN, Received September 19, 2022; Revised October 14, 2022; Accepted October 18, 2022

Ⅰ. 서 론

다양한 산업 기술이 발전함에 따라 네트워크 환경은 더욱 복잡해지고 대규모화되고 있으며 동적으로 변화하는 추세이다. 이에 따라 사이버 위협의 가능성 및 다양성은 더욱 커져갔다. 해킹으로 개인 정보 유출 사고가 발생하는 경우가 빈번해졌으며, 시스템 장악을 통해 대규모로 기업의 서버나 국가 시설망을 장악하여 악성코드 유포 및 기밀문서를 빼돌리는 경우도 많아졌다. 이렇듯 사이버 보안의 중요성은 갈수록 증대되었으며, 이를 방지하기 위해 다양한 사이버 보안 방어 기법 및 보안 환경을 테스트하는 방법들이 개발되었다. 취약점 분석이나 보안 평가를 위해 보편적으로 사용되고 있는 침투 테스팅(Penetration Testing) 기술[1]은 식별되거나 악용되는 보안 취약점에 대한 정보를 취합하여 시스템 관리자에게 제공하여 본 논문에서 공격 COA라 정의하고 있는 일련의 전략적 의사 결정을 내리거나 교정조치 작업의 우선순위를 정할 수 있다.

그러나 기존의 사이버 보안 위협 공격 COA 탐색 기법을 수동으로 설정하는 것에는 많은 한계가 있다. 전통적인 공격 COA 탐색 기법에는 사이버 보안에 대한 전문적인 지식을 가진 보안 전문가가 참여해야 하며, 그 전문성에 따라 활용 가치가 달라진다. 또한, 직접 사이버 보안을 점검해야 하므로 시간과 비용이 많이 드는 작업이다. 따라서 사이버 보안 분석의 효율성을 높이고 효과적인 방어 전략을 수립하기 위해 자동으로 공격 COA 탐색을 할 수 있는 기술에 대한 필요성이 존재한다.

이에 따라 국내외적으로 다양한 공격 COA 탐색 기술이 개발되었으며, 공격 COA 탐색 및 모델링 기법의 세부기술로는 공격 트리(Attack Tree) [2], 공격 그래프(Attack Graph) [3], BAS (Breach and Attack Simulation) [4], 침투 테스팅 기법이 있다. 최근에는 다양한 인공지능 알고리즘을 사용하여 공격 COA 탐색을 하는 경우가 다수 있었다. 특히 다양한 딥러닝 알고리즘 중에서 강화학습 알고리즘을 사용하여 자동으로 공격 COA 탐색하는 연구가 존재하였다. 이에 본 논문은 공격 COA 사이버 탐색을 위해 국내외 공격 COA 기술 동향을 조사하고, 강화학습을 사용한 다양한 공격 COA 탐색 기법 동향을 조사하고 분석하였다. 본 논문에서 2장은 사이버 공격 COA 탐색 기술의 정의 및 동향에 대해 서술하고 강화학습 기반 자동 탐색 기술의 필요성에 대해 논한다. 3장에서는 강화학습 기반 공격 COA 탐색 기법과 한계에 대해 서술하며, 4장에서는 결론을 맺는다.

Ⅱ. 사이버 위협 공격 COA 탐색 기술

2.1 기술 개요

미국 국방부(Department of Defense)에서는 COA를 (1) 개인 또는 단위가 따를 수 있는 활동의 순서, (2) 임무를 수행하거나 임무 수행과 관련된 개인 또는 지휘관에게 공개될 수 있는 계획, (3) 업무나 임무를 완수하기 위해 채택된 제도, (4) 교전에서의 행동 방식, (5) 공동 운영 계획 혹은 개발단계에서의 실행 시스템과 같이 정의하였다[5]. 즉, COA는 넓은 의미로 문제에 대한 잠재적인 해결책을 의미한다.

공격 COA는 사이버 공격 시뮬레이션 기술 분야에

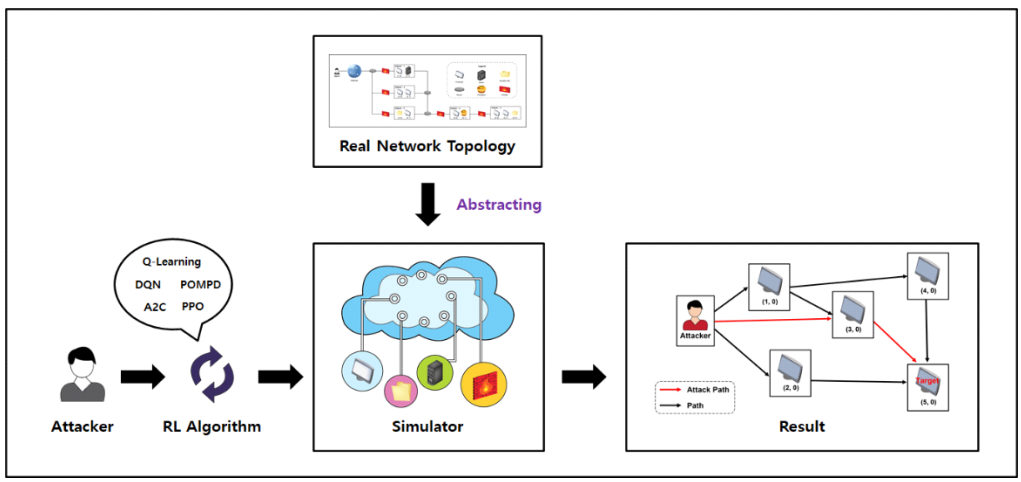

그림 1. 강화학습 기반 자율 공격 COA 기법
Fig. 1. RL based autonomous attack COA techniques

서 사용되며, 조직의 정책 및 상태를 점검하여 잠재적인 공격 위험 요소를 사전에 파악해 보안 강화 전략을 수립하는데 주로 사용되었다. 따라서 공격 COA는 사이버 분석을 수행하고 보안에 위협이 되는 일련의 공격을 탐색 및 분석하는 사이버 보안에서 필수적인 기법이 되었다. 기존의 공격 COA 탐색 기법은 공격 트리, 공격 그래프, 침투 테스팅, BAS 등의 기술들을 포함하며, 자동화를 통해 조직에서 보안성 평가를 위해 반복적으로 사용할 수 있는 기술로 발전해나가고 있다[4]. 본 논문은 여러 공격 COA 기법 중에서도 대표적으로 사용되는 기술인 침투 테스팅에 집중하였다.

침투 테스팅 기술은 사이버 공격에서 악용될 수 있는 컴퓨터 시스템의 보안 취약점을 찾아내는 방법을 제공해주는 기술이다. 특히 공격자가 악의적인 목적으로 시스템을 공격할 수 있는 경우를 고려하여 취약점을 찾아내며, 이를 바탕으로 통제된 공격을 수행한다. 따라서 침투 테스팅은 실제 환경에서의 사이버 공격과 유사한 공격을 하는 것이므로, 조직 내의 시스템에서의 보안을 평가하고, 이를 바탕으로 보안을 강화하는 것에 주로 사용되는 효과적인 방법 중 하나이다[1].

전통적으로 사용되는 침투 테스팅은 수동적으로 전문적인 지식을 가진 보안 전문가가 참여해야 한다. 또한, 다양한 데이터 분석이 필요하므로, 시간과 비용이 많이 드는 작업이었다. 산업 기술이 발전함에 따라 네트워크 환경은 점점 대규모화 되고 복잡화되었으므로, 여전히 수동적으로 인력을 사용하여 침투 테스팅을 하는 것은 효율성 면에서 많은 한계가 발생하였다. 따라서 인공지능 알고리즘을 사용하여 침투 테스팅을 할 필요성이 대두되었으며, 이에 따른 다양한 자동화 침투 테스팅 기술이 개발되었다.

2.2 사이버 위협 공격 COA 자동 탐색 기술 동향

침투 테스팅을 자동화하기 위해 다양한 연구들이 진행되어 왔는데, 그 중에서 대표적으로 공격 트리, 공격 그래프, 게임 이론(Game Theory) [6]을 사용한 연구들이 진행되어왔다.

공격 트리는 공격자가 공격하고자 하는 타깃 시스템을 표현해주는 수학적이고 트리 구조화된 다이어그램 또는 모델을 의미한다. 공격 트리는 공격자가 공격할 수 있는 선택과 공격하고자 하는 목표로 구성되어 있다. 공격 트리에서의 최상위 루트 노드는 공격자의 전반적인 목표로 나타나게 된다. 높은 수준의 타깃 시스템까지의 공격 경로가 다양할수록 공격 트리의 다이어그램은 확장된다[2]. 공격 트리는 일반적으로 공격 절차를 구성하기 위해 분석하는 기술로 주로 사용된다[7].

표 1. 침투 테스팅 자동화 알고리즘
Table 1. Autonomous algorithm of penetration testing

알고리즘	내용
공격 트리	공격 절차를 구성하기 위해 분석하는 알고리즘
공격 그래프	다양한 취약점을 찾아 모델링 및 다양한 공격 시나리오를 구상해주는 알고리즘
게임 이론	가장 합리적인 결정을 할 수 있도록 도와주는 의사 결정 알고리즘
강화학습	누적 보상의 합을 최대화하는 방법으로 의사 결정 알고리즘

공격 그래프는 공격을 모델링 할 수 있는 또 다른 방법 중에 하나이다. 공격 그래프는 공격자가 네트워크의 취약점을 분석하여, 공격자가 타깃 시스템까지의 가능한 모든 공격 경로를 단순화하여 표현할 수 있는 모델을 의미한다. 이때, 타깃 시스템에 도달하지 못하는 공격까지 포함이 된다. 공격 그래프에서의 노드는 공격자가 다양한 공격 단계에서 사용하는 취약점을 의미한다. 그리고 엣지는 공격이 성공적으로 실행되거나 공격이 수행된 후에 얻게 되는 결과를 위해 요구되는 보안 조건들로 여겨진다. 즉, 공격 그래프는 다양한 취약점을 다양한 호스트에서 찾아서 모델링하고, 이러한 취약점을 연결하여 다양한 공격 시나리오를 구상하는 것을 말한다[3]

게임 이론은 두 명 이상의 경쟁자가 상호 연관 관계를 맺음으로써 각자 개인의 이익을 추구하고 있으나, 경쟁자 중 누구도 결과를 결정할 수 없는 경쟁적인 상황을 가정한다. 이 상황에서 가장 합리적인 결정을 할 수 있도록 도와주는 의사 결정 알고리즘 중에 하나인 것이다[6]. 게임 이론을 사용하는 경우, 방어자와 공격자 모두의 관점에서 모델링할 수 있으므로, 자율적 침투 테스팅은 주로 게임 이론을 많이 사용하였다[8]. 특히 Stackelberg Security Game과 같은 게임 이론 또한 침투 테스팅에 적용한 사례가 있다[9,10].

2.3 기술의 한계

자율 공격 탐색 기법은 공격 그래프와 공격 트리가 대표적인 방법으로, 두 방법 모두 시스템보안을 평가하고 잠재적인 공격 경로를 검출하기 위한 설명 가능한 모델을 제공한다. 그러나 두 방법은 타깃 네트워크의 토폴로지와 모든 호스트의 구성에 대한 완전한 정보가 필요하기 때문에 실제 공격자의 관점에서 적용하는 것은 비현실적이다. 또한, 대규모 네트워크 시나리오에서는 상태 공간의 증가와 복잡한 모델링 과정으로 인해 적용하기 어려운 문제가 있다[11].

또한, 게임 이론을 사용하여 침투 테스팅을 할 경우에도 한계점이 발생한다. 게임 이론을 사용할 경우, 모든 네트워크의 구성을 관찰할 수 없다는 가능성과 신뢰할 수 없는 공격 툴을 사용할 가능성을 고려하지 못한다. 게임 이론으로 침투 테스팅을 수행할 경우, 완벽하게 통제할 수 있는 확률적인 게임을 수행해야 하는데 현실 네트워크 환경에서는 이는 충족하기 어렵다[8].

정리하면, 기존의 침투 테스팅은 실제 공격자의 활동을 시뮬레이션 하여 사이버 시스템의 취약성을 평가하는 효과적인 방법이지만, 주로 이용자의 전문성에 의존하며 공격 탐색방법의 반복성(Repeatability)을 감소시킨다. 이에, 탐색방법의 효율성을 높이고 사람의 개입을 줄이는 자동 공격 탐색 기법의 필요성이 증가한다. 자동 공격 탐색은 타깃(Target) 네트워크 환경의 불확실한 상태를 바탕으로 공격 의사 결정을 자율적으로 생성할 수 있어야 한다.

Ⅲ. 강화학습 기반 공격 COA 탐색 기술

3.1 기술 개요

3.1.1 강화학습 개요

강화학습은 기계학습의 한 분야로, 그림 2와 같이 마르코프 의사 결정 과정(Markov Decision Process, MDP)을 따른다. MDP는 의사결정자(Agent), 상태 정보(State), 정책(Policy), 행동(Action), 보상(Reward)으로 구성되어 있다. 강화학습은 의사 결정자가 의사결정자가 존재하는 환경으로부터 상태 정보를 관찰하고, 정책을 통해 확률적으로 행동을 결정하는데, 이러한 행동을 환경에 가하면 의사결정자는 행동에 대한 보상을 얻고, 다음 상태 정보를 받게 되는 과정으로 진행된다. 즉, 상태 관찰, 행동 결정, 상태 천이, 다음 상태 및 즉각적 보상 부여를 반복하면서 의사 결정을 하는 것이다. 강화학습은 시스템이 종료될 때까지 받을 수 있는 누적 보상의 합을 최대화하는 방법으로 의사 결정을 한다. 이러한 장점을 바탕으로, 강화학습 알고리즘을 다양한 분야에 적용한 연구들이 진행되어 왔다[12-19].

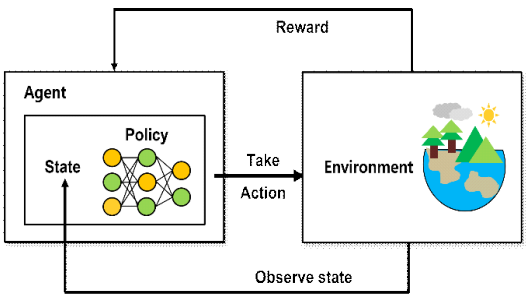

그림 2. 강화학습 알고리즘
Fig. 2. Reinforcement Learning algorithm

3.1.2 강화학습 기반 공격 COA 탐색 기술

다양한 분야에서 AI 알고리즘이 광범위하게 사용되고 있다. 고도화되는 사이버 공격 기술에 대응을 정확히 하고, 보안 위협에 효과적으로 대처하기 위해 인공지능 융합 보안 기술의 중요성이 강조되고 있다. 그러나 기존 AI 알고리즘을 사이버 보안 분야에 사용하기에는 한계가 있다. 데이터 부족으로 인한 변동성 문제가 발생할 수 있으며, 데이터 과적합 문제로 인해 잘못된 결과 도출의 위험성이 존재한다. 이러한 한계를 극복하기 위하여 강화학습 알고리즘을 보안 분야에서 사용할 것이 대두되었다. 그림 1과 같이 강화학습을 침투 테스팅에 사용할 경우, 일반적으로 데이터로 인해서 발생하는 AI 알고리즘의 문제를 해결할 수 있으며, 동적인 환경에서도 환경과 행동에 대한 보상 정보를 바탕으로 항상 최적의 솔루션을 학습하고, 이를 바탕으로 최종적으로 최적의 행동 정책을 도출할 수 있다. 이러한 강화학습의 특성을 고려하여, 강화학습 기반 공격 COA 탐색을 통해 발생 가능한 다양한 보안 문제에 효과적으로 대응할 수 있다. 잘못 학습된 인공지능에서의 오류로 인해서 문제가 발생하거나, 데이터 보안 등의 문제가 발생하였을 때, 강화학습을 통해 공격 COA 위협 탐색을 할 수 있다. 특히, 침투 테스팅 기법에 강화학습을 적용하는 연구들이 계속 진행되고 있는데, 심층 강화학습 기반 침투 테스팅은 인력을 줄이고 신뢰성을 향상시킴으로써 공격 탐색 COA를 자동화할 수 있는 유망한 해결책으로 평가되고 있다.

강화학습 알고리즘은 상태 공간(State Space), 행동 공간(Action Space), 보상(Reward)으로 구성될 수 있다. 여기서 상태 공간은 호스트의 구성 정보와 취약점 정보 등을 포함한 네트워크 관찰 정보이다. 이때, 상태 공간의 크기는 네트워크에서 존재하는 호스트의 수에 따라 달라진다. 행동 공간은 스캐닝(Scanning), 취약성 공격(Vulnerability Exploitation), 권한 상승(Privilege Escalation Operations)으로 구성되어 있다. 여기서 의사 결정자(Agent)가 호스트에서 행동을 수행하게 된다. 보상은 의사결정자가 공격한 모든 호스트의 가치에서 수행한 모든 행동에서 나오는 비용으로 산출되는 값이다. 이때 의사결정자는 최대한 비용

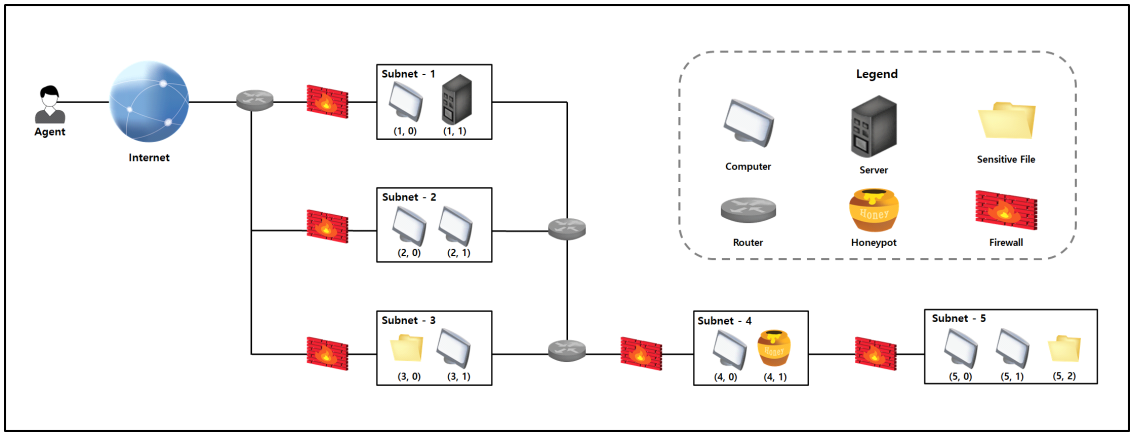

그림 3. 네트워크 공격 시나리오
Fig. 3. Network attack scenario

이 적게 드는 행동으로 가장 중요한 호스트를 손상시킬 수 있도록 구성된다. 즉, 보상을 최대화하는 과정에서 의사결정자는 최적의 공격 경로를 학습하게 되는 것이다. 이러한 구성 요소를 사용하여, 그림 3과 같이 네트워크 구성에서의 최적의 공격 경로를 찾는 시나리오를 구상할 수 있게 된다.

침투 테스팅 기법 또한 환경에서 주어지는 상태에 따라 동적으로 의사 결정을 하는 과정이라고 할 수 있다. 강화학습 기반 침투 테스팅 연구에서, 에이전트는 탐색과 시행착오(Trial-and-Error)를 통해 네트워크 환경에서 최적의 정책을 학습하도록 훈련한다. 이때, 에이전트가 주어진 네트워크 시나리오에 대해 사전 지식 및 가정에 관계없이 학습할 수 있다는 장점이 있다.

표 2. 강화학습 기반 공격 COA 탐색 기술 동향
Table 2. Trends of RL based attack COA techniques

알고리즘	내용
POMDP	실제 해커의 불완전한 관찰(Observation) 특성을 모델링하고 침투 테스팅 프로세스의 불확실성을 설명 가능
Q-Learning	사전 지식이 부족한 경우라도, 네트워크 구성 범위에 대해서 최적의 공격 경로를 탐색 가능
DQN	학습 데이터를 효율적으로 사용하면서, 학습 시에 발생하는 오차를 줄이는 효과를 달성 가능
Improved DQN	희소한 보상 문제에서의 최적의 공격 경로 탐색 가능
A2C	각각의 상태가 주어졌을 때 행동을 결정하고, 해당 상태의 공격 경로의 가치를 평가 가능
PPO	정책 신경망 근사를 통해서 최적의 공격 경로의 확률 도출

3.2 강화학습 기반 공격 COA 탐색 기술 동향

이번 장에서는 강화학습 알고리즘 기반 침투 테스팅 연구에 대한 동향을 정리한다. 먼저 최근 침투 테스팅 연구에서 사용되는 부분적으로 관측 가능한 마르코프 의사 결정 과정(POMDP, Partially Observable Markov Decision Processes) [20], Q-Learning[21], DQN[5]과 같은 강화학습 알고리즘에 대해 간단히 서술한 다음 연구 동을 서술하고, 마지막으로 기술의 한계에 대해 서술한다.

자율 공격 탐색 기법에 대한 한가지 접근법은 POMDP 모델이다. POMDP 알고리즘은 불확실한 환경과 소통하는 의사결정자의 불확실성을 고려하여 의사 결정하는 강화학습 기반 확률적 의사 결정 모델이다[20]. POMDP 방법을 사용하면 실제 해커의 불완전한 관찰(Observation) 특성을 모델링하고 침투 테스팅 프로세스의 불확실성을 설명할 수 있다. POMDP 모델을 사용하여 자동 침투 테스팅을 수행한 연구[22-23]에서는 네트워크 구성 및 방화벽 의존도와 같은 정보를 자연스럽게 표현할 수 있어, 침투 테스팅을 성공적으로 수행할 수 있다. 또한, 이 접근법을 사용하는 경우, 네트워크의 정보 수집을 중요한 부분으로 인식하여, 스캐닝과 공격을 지능적으로 융합할 수 있다. 그러나 대규모 네트워크 시나리오에서 POMDP 기반의 방법은 계산 비용 문제 등 적합하지 않다.

POMDP를 사용하여 지능형 자동 침투 테스팅 시스템인 IAPTS (Intelligent Automated Penetration Testing System)를 제안한 연구[24-25]에서 해당 시스템을 사용할 경우, 시간 소비와 인적 자원을 절약하여

자율적이고 지능적으로 침투 테스팅을 수행할 수 있으므로 정확도와 효율성 향상을 보인다. 특히 IAPTS는 대부분의 산업용 침투 테스팅 프레임워크와 통합할 수 있어서, 중대형 네트워크의 상황에서도 좋은 성능을 보일 수 있다.

Q-Learning 알고리즘은 각 행동과 보상에 따른 Q 값(Q-value)의 최적값을 예측하고, 현재 상태에서 가장 높은 Q 값을 가지는 동작을 선택하는 방식으로 동작하는 Model-Free 강화학습 알고리즘이다[21]. 강화학습 알고리즘이 침투 테스팅에 사용될 수 있는지 조사한 연구[26]에서는 Q-Learning 알고리즘을 사용하여 침투 테스팅을 할 경우, 네트워크 구성 정보와 같은 사전 지식만 주어진 경우라도, 네트워크 구성 범위에 대해서 최적의 공격 경로를 찾을 수 있다. 침투 테스팅을 자동화하기 위해 새로운 온톨로지(Ontology) 기반 BDI (Belief Desire Intention) 에이전트와 강화학습 기반의 Q-Learning을 결합한 프레임워크를 제안한 연구[27]에서는 해당 프레임워크를 사용함으로써 침투 테스팅을 수행할 때, 사전 지식이 알려지지 않은 환경에서 몇 번의 경험을 시도한 후에, 최적이고 효율적인 계획을 수립한 침투 테스팅을 시도하였다. 해당 프레임워크는 다양한 환경에서 적응력 있고 유연하다는 장점이 있으며, 정확성과 속도 측면에서도 성능이 향상되었다. 또한, AgentPen이라는 도구로 침투 테스팅을 수행한 연구[28]에서는 AgentPen 툴을 사용할 경우 성공적인 공격 훈련으로, 학습된 사전 지식 없이 최적의 동작 순서로 수행되는 침투 테스팅을 보인다. 특히, Q-Learning을 적용하여 침투 테스팅에서의 자율성을 달성하였다. AgentPen을 사용함으로써, 실제 공격 단계에서의 인간적 요소를 제거하고, 새로운 공격 시퀀스 및 공격 호스트 환경을 학습 및 탐색하고, 자체 내부 침투 테스팅의 로직을 개선할 수 있으므로, 공격 전략 자체를 도출할 수 있다는 장점이 있다. 그리고 고정 전략 기반 알고리즘인 Random, Greedy, Finite State Machine과, 학습 기반 알고리즘인 Q-Learning, DQN, Extended Classifier System을 침투 테스팅에 사용한 연구[29-30]에서는 해당 알고리즘의 침투 테스팅 성능을 측정하여 비교를 수행하였다. 이때, Q-Learning을 사용할 경우, 사람이 수동적으로 침투 테스팅을 수행하는 경우를 능가하는 최고의 성능을 보였다.

DQN은 Q-Learning 알고리즘에 심층 신경망인 DNN (Deep Neural Network)을 적용한 것으로, 학습 데이터를 효율적으로 사용하면서, 분리된 Neural Network (NN) 구조로 되어 있어 학습 시에 발생하는 오차를 줄이는 효과를 달성할 수 있다[5]. DQN을 기반으로 하는 자동화된 침투 테스팅 프레임워크를 제안한 연구[31]에서는 DQN을 사용하는 경우 주어진 환경에서의 최적의 공격 경로를 발견하였다. 또한 그들은 CVSS (Common Vulnerability Scoring System) 점수 정보를 사용하여 각 노드에 할당된 보상 점수를 활용하여 가장 실현 가능한 공격 경로를 결정한다. 사이버 지형 개념인 IPB (Intelligence Preparation of the Battlefiled)를 침투 테스팅에 도입하여 현실을 반영한 연구[32]에서는 IPB와 DQN을 함께 사용할 경우 자동 침투 테스팅을 수행할 때 보상과 공격 행위에 있어서 뚜렷한 성과를 보였다. IPB를 침투 테스팅에 도입함으로써 강화학습 모델을 유지하는 것에 도움이 되고, 강화학습 의사결정자의 행동 현실화에 도움을 주었다.

또한, 최근에는 DQN 알고리즘을 더욱 발전시켜 침투 테스팅을 수행하기도 한다. NDSP-IDQN이라는 발전된 DQN 알고리즘을 제안한 연구[11]에서는 해당 알고리즘은 AQN의 개선 방법을 합리적으로 선택하고, 이를 DQN 알고리즘에 통합하여 탐색 능력을 향상시킴으로써 희소한 보상 문제를 더 잘 해결할 수 있다. 또한, 공간 백터 분리를 통해 행동 공간을 줄임으로써 의사 결정자의 탐색 과정에서의 시행착오 비용을 줄일 수 있다는 장점이 있다. HA-DRL이라는 계층적 구조 기반의 심층 강화학습 아키텍처를 제시한 연구[33]에서는 해당 알고리즘을 사용한 침투 테스팅 자동화 기법을 제안하였다. HA-DRL은 대수적 행동 분해 전략을 사용하여 자동 침투 테스팅을 수행할 때 발생하는 큰 규모의 이산 동작 공간(Discrete Action Space)을 처리할 수 있으며, 기존의 DQN보다 우수한 성능을 보였다.

A2C (Advantage Actor Critic) 알고리즘[34]은 정책 자체를 학습시키는 policy-based 방법이다. 정책의 파라미터를 업데이트시키는 actor 신경망과 가치 함수를 업데이트 해나가는 critic 신경망으로 구성되어, 각각의 상태가 주어졌을 때 행동을 결정하고, 해당 상태의 가치를 평가한다. 침투 테스팅을 위한 최적의 데이터 유출 경로를 파악하기 위해 A2C 모델 및 Double 에이전트환경을 사용한 연구[35]에서는 A2C 모델을 사용함으로써 데이터 추출하기 위해 사용될 가능성이 가장 높은 서비스와 호스트 및 네트워크 위험 평가에 사용되는 평가 지표를 식별하기에 용이함을 보였다. 추가로, 규모가 상당한 네트워크 환경에서도 잘 동작할 수 있음을 성능 평가로 보였다.

PPO (Proximal Policy Optimization) [36] 역시 정책 자체를 학습시키는 policy-based 방법으로 정책 신

표 3. 공격 시뮬레이션 프레임워크
Table 3. Attack Simulation Framework

Framework	내용	강화학습 알고리즘 적용 여부
CALDERA	Scan, Exploit 외 다양한 액션 라이브러리 제공	X
HARMer	그래픽 보안 모델(HARM) 사용하여 확장성 및 적응력 향상	X
NASim	네트워크 추상화를 통해 현실 네트워크 토폴로지 반영	POMDP, Q-learning, DQN, Improved DQN
IAPTS	중대형 네트워크 상황에서의 정확성과 효율성 향상	POMDP
CybORG	공동 인터페이스가 있는 시뮬레이션 및 환경을 제공	Improved DQN

경망 근사를 통해서 행동 확률을 도출하는 알고리즘이다. PPO 알고리즘과 RND (Random Network Distillation)를 결합하여 침투 테스팅을 위한 에이전트를 학습한 연구[37]에서는 MOMDP (Multi-Objective Markov Decision Process)를 기반으로 침투 테스팅을 수행하여, 보다 효율적이고 안정적으로 자동 침투 테스팅을 수행하였다.

3.3 기술의 한계

강화학습 에이전트는 자율 침투 테스팅을 통해 보상을 극대화하는 최적의 공격 경로를 학습해야 한다. 공격 경로, 즉 공격 COA는 일련의 액션 또는 일련의 순서가 지정된 공격 벡터이다. 그런데, 이 공격 경로의 탐색 공간은 네트워크 규모의 영향을 받게 되어, 네트워크 사이즈가 증가하면 다음 이유로 에이전트의 학습 효율 낮아지게 된다[11].

먼저, 대부분의 경우에 네트워크 기밀 호스트 중 양의 가치(value) 값을 갖는 호스트는 일부이기 때문에, 네트워크 사이즈가 증가하면 보상이 드물게 발생하는 문제가 생겨 알고리즘의 수렴이 어려워지는 문제가 발생한다[38]. 두 번째로, 네트워크 크기가 커짐에 따라 행동 공간의 크기가 커지게 되어, 탐색(exploration) 효율이 떨어지게 되는 문제가 발생한다.

Ⅳ. 공격 COA 시뮬레이션 프레임워크

이번 장에서는 공격 COA 시뮬레이션 프레임워크를 정리한다. 공격 COA 시뮬레이션 프레임워크는 표 1처럼 CALDERA[39], HARMer[40], NASim[26], IAPTS[24], 적용한 시뮬레이터는 NASim, IAPTS, CybORG가 있다.

4.1 CALDERA

CALDERA (Cyber Adversary Language and Decision Engine for Red Team Automation) [39] 는 MITRE에서 사이버 공격에 대응하기 위해, 자체적으로 공격을 모델링하고 실험하는 시스템을 개발한 프레임워크이다. CALDERA는 새로운 기술에도 잘 적응하여 확장성이 뛰어나며, 자율적으로 실행되기 때문에 비용과 시간적인 측면에서도 효율적이다. CALDERA는 적을 모방하고 인스턴스화하는 소프트웨어 기반 시설인 ViRTS (Virtual Red Team System)와 CALDERA가 수행할 작업 선택에 사용되는 논리적인 모델인 LAVA (Logic for Advanced Virtual Adversaries)로 구성되어 있다. CALDERA는 스캔과 공격 이외에 더 강력한 기능을 갖춘 액션 라이브러리를 제공함으로써 다른 침투 테스트 프레임워크와는 차별점이 있다. 이 시스템의 내부에서는 고전적인 계획, 마르코프 의사 결정 프로세스, Monte Carlo 시뮬레이션 기술을 사용하여 솔루션을 개발하여 사용자 정의를 하였으며, 이를 통해 사이버 환경 및 적 프로파일의 논리적 인코딩을 맞춤으로 개발되었다. CALDERA는 공격자의 행동을 바탕으로 평가를 수행하고, 공격자 탐지를 수행할 수 있다. 또한 방어자는 이를 바탕으로 위협을 감지하고 대응하는 방법을 추론 할 수 있다는 장점이 있다.

4.2 HARMer

HARMer[40]은 사이버 공격 생성을 위해 HARM (Hierarchical Attack Representation Model)이라는 그래픽 보안 모델(GSM; Graphic Security Model)을 기반으로 하는 그 새로운 자동화 침투 테스팅 공격 프레임워크이다. 데이터는 NMAP및 OpenVAS와 같은 기존 도구를 사용하여, 네트워크 정보(호스트, 호스트 구성, 취약성, PORT, 서비스)와 위협 및 공격자 정보를 수집한다. 또한 CVSS 점수를 사용하여, 각 취약점의 심각도를 바탕으로 공격 경로를 설정하는 방법을 사용하였다. HARMer은 AWS (Amazone Web Service)에서 실험을 수행하여, 성능을 검증하였다. HARMer은 확장성이 좋고 적응력이 좋아, 합리적인 공격자의 가능한 공격 동작을 체계적으로 계획할 수 있으며, 계산 복잡성을 줄여주어 확장 가능하다는 장점이 있다.

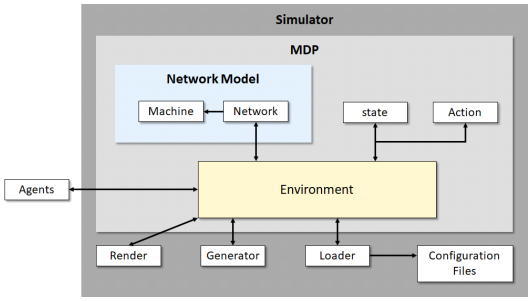

그림 4. NASim 프레임워크
Fig. 4. Framework of NASim

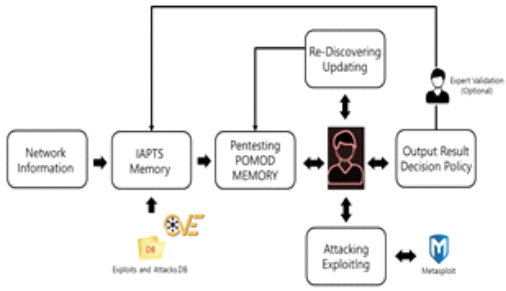

그림 5. IAPTS 프레임워크
Fig. 5. Framework of IAPTS

4.3 NASim

NASim (Network Attack Simulator)[26]은 그림 4와 같이 침투 테스팅을 위한 시뮬레이션을 제공하는 오픈 소스 연구 플랫폼이다. 네트워크 시나리오를 생성할 수 있으며, 강화학습 알고리즘을 사용하여 에이전트의 자동 침투 테스팅을 수행할 수 있도록 해주는 환경이다. 특히, NASim은 경량 네트워크 공격 시뮬레이터로 네트워크 추상화를 통해 현실 네트워크 토폴로지의 특성들을 최대한 반영하였다. 또한 네트워크의 규모 및 네트워크 구성 유형에 따라 다양한 벤치마크 시나리오를 제공하고 있다. 규모에 따라서 Tiny, Small, Medium 등으로 시나리오가 제공되며, 종류에 따라서Hard, Linear, Honeypot, Single, Multi 등으로 다양한 벤치마크 시나리오가 제공된다. 그리고 사용자는 Generated 버전을 사용하여, 호스트 구성, 방화벽 구성 등과 같은 구성 요소를 일부 변경할 수 있다. 이를 통해 새로운 네트워크 토폴로지 기반의 시나리오를 생성하여 침투 테스팅을 수행할 수도 있다는 장점이 있다. NASim 시뮬레이터에POMDP, Improved DQN, Q-learning, DQN과 같은 강화학습 알고리즘을 사용한 다양한 연구들이 진행되어왔다.

4.4 IAPTS

IAPTS (Intelligent Automated Penetration Testing System) [24-25]도구는 강화학습 알고리즘을 사용하는 지능형 도구로, 복잡한 침투 테스팅 활동을 학습하고 재현할 수 있는 환경을 그림 5와 같이 제공한다. 이 시스템을 사용하여 침투 테스팅을 할 경우, 신뢰성 및 테스트 빈도 측면에서 더 좋은 결과를 산출하면서, 사용하는 시간 줄일 수 있어 인적 자원을 절약할 수 있으므로 효율적이라는 장점이 있다. 이를 통해 대부분의 산업용 침투 테스트 프레임 워크와 통합할 수 있어, 중대형 네트워크 상황에서의 정확성과 효율성을 향상시킬 수 있다. IAPTS 시뮬레이터에 POMDP와 같은 강화학습 알고리즘을 사용한 다양한 연구들이 진행되어왔다.

4.5 CybORG

CybORG (Cyber Operations Research Gym) [42]는 ACO (Autonomous Cyber Operation) 요구사항을 지원하기 위해 그림 6과 같이 설계된 환경으로, 자율 에이전트 및 보안 인력의 교육을 위한 사이버 보안 연구 환경이다. CybORG는 미리 생성해놓은 설명을 기반으로 시나리오를 만들고, 에이전트들은 반복적으로 시나리오의 환경과 상호작용하면서 시나리오 상태 값을 업데이트 시켜서 마지막에는 원하는 조건을 충족시킨다. CybORG는 실제 시스템에서 테스트를 수행할 수 있도록 공동 인터페이스가 있는 시뮬레이션 및 환경을 제공한다. 이를 통해 빠른 속도로 학습할 수 있다는 장점이 있다. CybORG는 한 에이전트를 같은 코드를 사용하여 각기 다른 레벨의 현실재현도 (Fidelity)에서 학습하고 시험을 함으로써 사이버 자율성을 위한 새로운 가능성을 구현하였다. CybORG 환경을 사용하여, 다양한 침투 테스팅 시나리오가 제작되었으며, 그 중에서 CAGE (Cyber Autonomy Gym

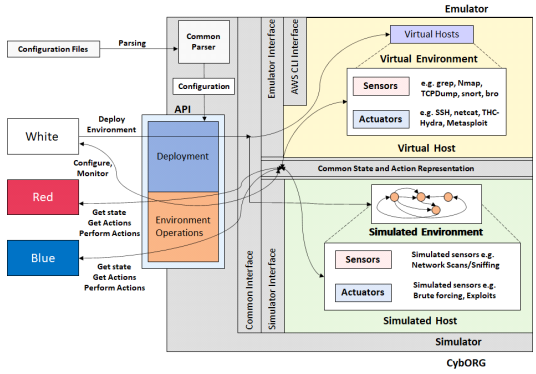

그림 6. CybORG 프레임워크
Fig. 6. Framework of CybORG

for Experimentation) [41] 시나리오 등이 개발되었다. CybORG 시뮬레이터에 HA-DRL이라는 발전된 DQN 알고리즘을 사용한 연구가 진행되기도 하였다.

V. 결 론

본 논문은 사이버 공격에 대한 선제적 대응 방법으로써 강화학습 기반 사이버 공격 시뮬레이션 기술의 동향에 대해 기술하였다. 현재 보편적으로 사용되는 침투 테스팅 기술로부터 공격 트리 기술, 공격 그래프 기술에서부터 강화학습 기반 자율 공격 기술까지 많은 관심을 받고 있다. 이처럼 사이버 공격 탐색 기술은 기존의 수동 분석 기술에서 점차 자동화된 기술로 발전하는 단계에 있다. 특히, 강화학습 알고리즘을 이용한 연구가 활발히 이루어지고 있다.

그러나 현 시점에서, 실제 네트워크 트래픽의 모델링이 어렵고, 실제 상황에서 적용하기 어려운 한계점들 때문에 강화학습을 이용한 자율 공격 테스트는 아직 시뮬레이션 실험 단계에 있다. 추후 다중 에이전트 강화학습 혹은 계층적 강화학습 기법 같은 보다 고급 알고리즘을 사용하여 더 많은 사이버 공격 시나리오에서의 성능을 개선하는 것 또한 도전적인 연구가 될 것으로 생각 된다.

References

[1] B. Arkin, S. Stender, and G. McGraw, "Software penetration testing," *IEEE Secur. & Privacy*, vol. 3, no. 1, pp. 84-87, 2005. (https://doi.org/10.1109/MSP.2005.23)

[2] T. R. Ingoldsby, "Attack tree-based threat risk analysis," *Amenaza Technologies Limited*, pp. 3-9, 2010.

[3] N. Ghosh and S. K. Ghosh, "A planner-based approach to generate and analyze minimal attack graph," *Appl. Intell.*, vol. 36, no. 2, pp. 369-390, 2012. (https://doi.org/10.1007/s10489-010-0266-8)

[4] J. Y. Lee, D. S. Moon, and I. K. Kim, "Technological trends in cyber attack simulations," *Electr. and Telecommun. Trend*, vol. 35, no. 1, pp. 34-48, 2020.

[5] P. P. Reese, "Military decision making process: Lessons and best practices," *Center for Army Lessons Learned Fort Leavenworth United States*, pp. 27-63, 2015.

[6] X. Liang and Y. Xiao, "Game theory for network security," *IEEE Commun. Surv. & Tuts.*, vol. 15, no. 1, pp. 472-486, 2012. (https://doi.org/10.1109/SURV.2012.062612.00056)

[7] J. H. Eom, S. H. Park, and T. M. Chung, "A study on an extended cyber attack tree for an analysis of network vulnerability," *J. Korea Soc. Digital Ind. and Inf. Manag.*, vol. 6, no. 3, pp. 49-58, 2010.

[8] J. Schwartz, H. Kurniawati, and E. El-Mahassni, "Pomdp+ information-decay: Incorporating defender's behaviour in autonomous penetration testing," in *Proc. ICAPS*, vol. 30, pp. 235-243, Jun. 2020. (https://ojs.aaai.org/index.php/ICAPS/article/view/6666)

[9] P. Speicher, M. Steinmetz, J. Hoffmann, M. Backes, and R. Künnemann, "Towards automated network mitigation analysis," in *Proc. 34th ACM/SIGAPP Symp. Appl. Comput.*, pp. 1971-1978, Limassol, Cyprus, Apr. 2019. (https://doi.org/10.1145/3297280.3297473)

[10] T. H. Nguyen, D. Kar, M. Brown, A. Sinha, A. X. Jiang, and M. Tambe, "Towards a science of security games," *Math. Sci. with Multidisciplinary Appl.*, pp. 347-381, 2016.

[11] S. Zhou, J. Liu, D. Hou, X. Zhong, and Y. Zhang, "Autonomous penetration testing based on improved deep Q-Network," *Applied Sci.*, vol. 11, no. 19, p. 8823, 2021. (https://doi.org/10.1007%2F978-3-319-31323-8_16)

[12] W. J. Yun, M. Shin, D. Mohaisen, K. Lee, and J. Kim, "Hierarchical deep reinforcement learning-based propofol infusion assistant framework in anesthesia," *IEEE Trans. Neural Netw. and Learn. Syst.*, 2022. (https://doi.org/10.1109/tnnls.2022.3190379)

[13] W. J. Yun, S. Park, J. Kim, M. Shin, S. Jung, D. A. Mohaisen, and J. H. Kim, "Cooperative multiagent deep reinforcement learning for reliable surveillance via autonomous Multi-

UAV control," *IEEE Trans. Ind. Informat.*, vol. 18, no. 10, pp. 7086-7096, 2022. (https://doi.org/10.1109/TII.2022.3143175)

[14] W. J. Yun, D. Kwon, M. Choi, J. Kim, G. Caire, and A. F. Molisch, "Quality-aware deep reinforcement learning for streaming in infrastructure-assisted connected vehicles," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 2002-2017, 2021. (https://doi.org/10.1109/tvt.2021.3134457)

[15] S. Jung, W. J. Yun, M. Shin, J. Kim, and J. H. Kim, "Orchestrated scheduling and multi-agent deep reinforcement learning for cloud-assisted multi-UAV charging systems," *IEEE Trans. Veh. Technol.*,vol. 70, no. 6, pp. 5362-5377, 2021. (https://doi.org/10.1109/TVT.2021.3062418)

[16] M. Choi, M. Shin, and J. Kim, "Dynamic video delivery using deep reinforcement learning for device-to-device underlaid cache-enabled Internet-of-Vehicle networks," *J. Commun. and Netw.*, vol. 23, no. 2, pp. 117-128, 2021. (https://doi.org/10.23919/JCN.2021.000006)

[17] W. J. Yun, S. Jung, J. Kim, and J. H. Kim, "Distributed deep reinforcement learning for autonomous aerial eVTOL mobility in drone taxi applications," *ICT Express*, vol. 7, no. 1, pp. 1-4, 2021. (https://doi.org/10.1016/j.icte.2021.01.005)

[18] D. Kwon, J. Kim, D. A. Mohaisen, and W. Lee, "Self-adaptive power control with deep reinforcement learning for millimeter-wave Internet-of-vehicles video caching," *J. Commun. and Netw.*, vol. 22, no. 4, pp. 326-337, 2020. (https://doi.org/10.1016/j.icte.2021.01.005)

[19] M. Shin, D. H. Choi, and J. Kim, "Cooperative management for PV/ESS-enabled electric vehicle charging stations: A multiagent deep reinforcement learning approach," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3493-3503, 2019. (https://doi.org/10.1109/TII.2019.2944183)

[20] G. E. Monahan, "State of the art—a survey of partially observable markov decision processes: Theory, models, and algorithms," *Manag. Sci.*, vol. 28, no. 1, pp. 1-16, 1982. (https://doi.org/10.1287/mnsc.28.1.1)

[21] C. J. C. H. Watkins and P. Dayan. "Q-learning," *Mach. Learn.*, vol. 8, no. 3, pp. 279-292, 1992. (https://doi.org/10.1007/bf00992698)

[22] C. Sarraute, O. Buffet, and J. Hoffmann, "POMDPs make better hackers: Accounting for uncertainty in penetration testing," in *Proc. 26th AAAI Conf. Artificial Intell.*, Toronto, Ontario, Canada, Jul. 2012. (http://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/view/4996)

[23] C. Sarraute, O. Buffet, and J. Hoffmann, "Penetration testing== POMDP solving?," in *Proc. 2011 IJCAI Wkshp. Intell. Secur. (SecArt)*, p. 66, Barcelona, Spain, Jul. 2011.

[24] M. C. Ghanem and T. M. Chen, "Reinforcement learning for efficient network penetration testing," *Information*, vol. 11, no. 1, p. 6, 2019. (https://doi.org/10.3390/info11010006)

[25] M. C. Ghanem and T. M. Chen, "Reinforcement learning for intelligent penetration testing," in *Proc. 2018 Second World Conf. Smart Trends in Syst., Secur. and Sustainability(WorldS4)*, pp. 185-192, London, UK, Oct. 2018. (https://doi.org/10.1109/worlds4.2018.8611595)

[26] J. Schwartz and H. Kurniawati, "Autonomous penetration testing using reinforcement learning," arXiv preprint arXiv:1905.05965, 2019.

[27] K. Qian, D. Zhang, P. Zhang, Z. Zhou, X. Chen, and S. Duan, "Ontology and reinforcement learning based intelligent agent automatic penetration test," in *Proc 2021 IEEE ICAICA*, pp. 556-561, Dalian, China, Jun. 2021. (https://doi.org/10.1109/icaica52286.2021.9497911)

[28] K. Pozdniakov, E. Alonso, V. Stankovic, K. Tam, and K. Jones, "Smart security audit:

Reinforcement learning with a deep neural network approximator," in *Proc. 2020 Int. Conf. Cyber Situational Awareness, Data Analytics and Assessment(CyberSA)*, Dublin, Ireland, pp. 1-8, Jun. 2020. (https://doi.org/10.1109/CyberSA49311.2020.9139683)

[29] S. Niculae, *Reinforcement Learning vs Genetic Algorithms in Game-Theoretic Cyber-Security* (2018), Retrieved Mar. 28, 2022, from https://thesiscommons.org/nxzep/ (https://doi.org/10.31237/osf.io/nxzep)

[30] S. Niculae, D. Dichiu, K. Yang, and T. Bäck, *Automating Penetration Testing using Reinforcement Learning*(2020), Retrieved Mar. 28, 2022, from https://stefann.eu/files/Automating%20Penetration%20Testing%20using%20Reinforcement%20Learning.pdf

[31] Z. Hu, R. Beuran, and Y. Tan, "Automated penetration testing using deep reinforcement learning," in *Proc 2020 IEEE EuroS&PW*, pp. 2-10, Genoa, Italy, Sep. 2020. (https://doi.org/10.1109/EuroSPW51379.2020.00010)

[32] R. Gangupantulu, T. Cody, P. Park, A. Rahman, L. Eisenbeiser, D. Radke, and R. Clark, "Using Cyber Terrain in Reinforcement Learning for Penetration Testing," arXiv preprint ar Xiv:2108.07124, 2021. (https://doi.org/10.1109/COINS54846.2022.9855011)

[33] K. Tran, A. Akella, M. Standen, J. Kim, D. Bowman, T. Richer, and C. T. Lin, "Deep hierarchical reinforcement agents for automated penetration testing," arXiv preprint arXiv:2109.06449, 2021.

[34] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous methods for deep reinforcement learning," in *Proc. 33rd Int. Conf. Mach. Learn.(PMLR)*, pp. 1928-1937, New York, USA, Jun. 2016.

[35] T. Cody, A. Rahman, C. Redino, L. Huang, R. Clark, A. Kakkar, and E. Bowen, "Discovering exfiltration paths using reinforcement learning with attack graphs," arXiv preprint arXiv:2201.12416, 2022. (https://doi.org/10.1109/DSC54232.2022.9888919)

[36] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv preprint arXiv:1707.06347, 2017.

[37] Y. Yang and X. Liu, "Behaviour-diverse automatic penetration testing: A curiosity-driven multi-objective deep reinforcement learning approach," arXiv preprint arXiv: 2202.10630, 2022.

[38] A. Nair, B. McGrew, M. Andrychowicz, W. Zaremba, and P. Abbeel, "Overcoming exploration in reinforcement learning with demonstrations," in *Proc 2018 IEEE ICRA*, pp. 6292-6299, May 2018. (https://doi.org/10.1109/ICRA.2018.8463162)

[39] MITRE, CALDERA, Retrieved Mar. 28, 2022, from https://www.mitre.org/research/technology-transfer/open-source-software/caldera%E2%84%A2

[40] S. Y. Enoch, Z. Huang, C. Y. Moon, D. Lee, M. K. Ahn, and D. S. Kim, "HARMer: Cyber-attacks automation and evaluation," *IEEE Access*, vol. 8, pp. 129397-129414, 2020. (https://doi.org/10.1109/ACCESS.2020.3009748)

[41] M. Standen, D. Bowman, S. Hoang, T. Richer, M. Lucas, and R. V. Tassel, *Cyber Autonomy Gym for Experimentation Challenge 1*(2021), Retrieved Mar. 28, 2022, from https://www.dst.defence.gov.au/partner-with-us/opportunities/call-submissions-cyber-autonomy-gym-experimentation

[42] C. Baillie, M. Standen, J. Schwartz, M. Docking, D. Bowman, and J. Kim, "Cyborg: An autonomous cyber operations research gym," arXiv preprint arXiv:2002.10667, 2020.

손 석 빈 (Seok Bin Son)

2022년 2월 : 서울여자대학교 정보보호학과 졸업
2022년 3월~현재 : 고려대학교 전기전자 공학과 석사과정
<관심분야> 인공지능, 정보보호, 빅데이터
[ORCID:0000-0002-3692-0752]

이 해 민 (Haemin Lee)

2020년 2월 : 숙명여자대학교 통계학과 졸업
2020년 9월~현재 : 고려대학교 전기전자 공학과 석박통합과정
<관심분야> 통신공학, 데이터 공학
[ORCID:0000-0002-7524-1672]

박 수 현 (Soohyun Park)

2019년 2월 : 중앙대학교 컴퓨터공학과 졸업
2020년 3월~현재 : 고려대학교 전기전자 공학과 석박통합과정
<관심분야> Connected mobility, Distributed computing, Stochastic optimization, Data science
[ORCID:0000-0002-6556-9746]

김 동 화 (Dong Hwa Kim)

2004년 2월 : 고려대학교 전기전자전파공학과 졸업
2007년 2월 : 고려대학교 전기공학과 석사 졸업
2007년~현재 : 국방과학연구소 선임연구원
<관심분야> 사이버 보안 M&S 및 훈련
[ORCID:0000-0002-0596-1199

김 중 헌 (Joongheon Kim)

2004년 2월 : 고려대학교 컴퓨터학과 졸업
2006년 2월 : 고려대학교 컴퓨터학과 석사 졸업
2014년 8월 : University of Southern California Computer Science 박사 졸업
2016년 3월 : 중앙대학교 소프트웨어학대학 조교수
2019년 9월~현재 : 고려대학교 전기전자공학부 조교수/부교수
<관심분야> Stochastic Optimization, Mobility, Reinforcement Learning, Quantum Deep Learning
[ORCID:0000-0003-2126-768X]