네트워크 침투 테스트 자동화를 위한 시뮬레이션 환경 개발 연구

김정윤, 김혜지, 박종열*
서울과학기술대학교
kjy97426@naver.com, kimhyejee923@seoultech.ac.kr, jongyoul@seoultech.ac.kr*

A Study on the Development of Simulation Environment for Automating Network Penetration Test

Kim Jeong Yoon, Kim Hye Ji, Park Jong Youl*
Seoul National Univ. of Science and Technology

요 약

ICT 기술의 발전에 따라 사이버 위협도 진화하고 있어, 보다 정교한 사이버 보안 대책이 필요해지고 있다. 기존의 보안 메커니즘은 정적 규칙과 과거 데이터에 기반하여 작동하기 때문에 새로운 유형의 복잡한 사이버 공격을 탐지하고 대응하는 데 한계를 보인다. 이에 따라, 미지의 위협에 대한 적응적이고 선제적인 방어를 가능하게 하는 강화학습이 사이버 보안 분야에서 주목받고 있다. 본 논문은 강화학습 기반 보안 에이전트 개발을 위한 낮은 추상화 수준의 시뮬레이션 환경을 구축하여, 현실과의 간극을 최소화하는 연구를 제시한다. 이를 통해 개발된 에이전트는 실제 네트워크 환경에서도 효과적으로 작동할 수 있음을 확인하였다.

Ⅰ. 서 론

ICT기술이 점점 발전함에 따라 사이버 위협도 같이 진화하고 있다. 지금까지는 많은 보안 전문가들은 규칙 기반의 사이버 보안 프로그램을 만들어 이러한 위협에 대비했다. 하지만 네트워크 시스템의 규모가 방대해진 현재의 사이버 위협은 매우 복잡하고 막기 어려워졌다[1]. 더욱이 각 분야에서 인공지능의 성능이 인간의 실력을 앞서면서, 사이버 위협 및 보안 분야에도 인공지능의 도입이 불가피해졌다. 기존의 방어 매커니즘이 정교해진 사이버 공격을 탐지하기 힘들어지게 되자 많은 보안 전문가들은 머신러닝 방식의 사이버 보안을 개발했다.

하지만 다양한 방법의 머신러닝(SVM, LR, KNN 등), 딥러닝과 같이 과거의 데이터로부터 얻어진 인사이트를 기반으로 만들어진 보안 매커니즘은 새로운 유형의 공격을 탐지해내는데 어려움을 겪는다[2]. 이에 실시간으로 발전하는 사이버 위협에 대처하기 위해 이전에 알려지지 않던 공격 패턴을 탐지하고 대응할 수 있는 알고리즘을 만들 수 있는 강화학습이 최근 사이버 보안 분야에서 각광받고 있다. 머신러닝 알고리즘 중 하나인 강화학습은 에이전트가 환경과 상호작용하여 보상의 누적값을 최대화하는 방향으로 행동하게끔 학습하는 알고리즘이다. 기존의 데이터에 의존하지 않고, 시행착오를 통해 학습하기에, 새로운 유형의 공격에도 대응이 가능하다.

강화학습으로 보안 매커니즘을 개발하기 위해서는 환경이 필요하다. 기존의 연구들은 높은 추상화 수준으로 구현된 환경을 사용해 보안 에이전트를 학습시켰다[3]. 시뮬레이터와 실제 환경은 큰 간극이 존재하는데, 추상화 수준이 높을수록 그 간극은 훨씬 커진다. 본 연구는 낮은 추상화 수준으로 시뮬레이터를 구현해, 현실과의 간극을 최소화시켰다. 본 연구는 인공지능을 기반으로 일반적인 네트워크망에서 공격 시나리오 탐지 및 자가학습 기술을 개발하는 것으로 최근 마이크로소프트사에서 개발한 사이버배틀심 기술 조사를 바탕으로 환경을 구축하고 강화학습 기법을 적용하여 네트워크 보안 시뮬레이션을 가능하게 하는 환경을 개발함에 목적이 있다.

본 연구의 기여는 다음과 같다.

- 낮은 추상화 수준으로 시뮬레이터를 구현해 현실과의 간극을 최소화하여 본 시뮬레이터에서 학습된 에이전트는 현실과 유사하게 작동될 수 있다.

논문의 구성은 다음과 같다. 본론에서 시뮬레이터의 구현, 학습에 쓰인 강화학습 알고리즘, 실험 및 실험 결과를 보여준다. 결론에서는 이 논문의 기여도와 향후 계획 등을 설명한다.

Ⅱ. 본론

2.1 Mitre Att&ck 제공 시나리오

본 연구 과정에서는 실제 네트워크 보안 위협 사례를 테스트해 보고자 Mitre Att&ck (https://attack.mitre.org/)에서 제공되는 시나리오 중 한 가지를 선택하여 공격 방법을 재현하였다. 사용한 시나리오는 “APT28”이며 해당 그룹은 2016년 힐러리 클린턴 캠페인, 민주당 전국위원회(DNC), 민주당 의회 선거운동위원회(DOC)를 손상시켜 미국 대통령 선거에 개입하려는 시도를 한 것으로 알려져 있다.

2.2. 컴퓨터 네트워크 구성

컴퓨터 네트워크의 구조와 데이터 통신 과정은 다음과 같다. TCP/IP 모델의 네 가지 계층(Application, TCP, IP, Physical)을 중심으로 구성되어 있다. 각 계층은 특정 역할을 담당하며, 데이터가 네트워크를 통해 전송될 때 각 계층을 거치며 포맷이 변환된다. Application 계층에서는 애플리케이션이 네트워크를 통해 데이터를 송수신하기 위해 소켓을 할당받는다. 소켓은 데이터 전송의 출발점이며, send 및 receive buffer를 통해 데이터를 처리한다. TCP 계층은 데이터를 TCP 세그먼트로 변환하며, 포트 정보를 포함한 TCP 헤더를 붙인다. 애플리케이션 간 통신은 TCP 세션을 통해 관리되며, TCP 3 way handshake를 통해 세션이 설정된다. IP 계층에서는 TCP 세그먼트에 IP 헤더가 추가되어 IP 패킷이 생성된다. 이 패킷은 각 컴퓨터의 네트워크 인터페이스 카드(NIC)를 통해 라우터로 전송되며,

이 과정에서 링크 헤더가 추가되어 프레임으로 변환된다. NIC와 Router는 패킷을 목적지로 전달하는 역할을 한다. 라우터는 라우팅 테이블을 사용하여 프레임을 다음 라우터 또는 클라이언트로 전달하며, DHCP를 통해 IP 주소를 할당한다. 방화벽은 규칙에 따라 패킷을 필터링하여 보안 기능을 제공한다. 각 단계에서의 PDU (Protocol Data Unit)는 해당 계층의 헤더를 포함하여 다음 계층으로 전달되며, 최종적으로 데이터를 전송하고 수신하는 데 필요한 모든 정보를 포함하게 된다.

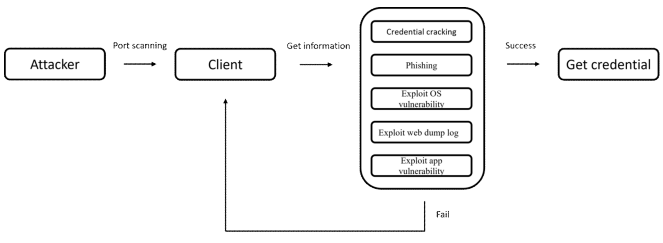
2.3. 컴퓨터 시스템 구성

컴퓨터의 커널은 여러 관리 모듈로 구성되어 있으며, 이들 각각이 시스템 운영에 중요한 역할을 한다. 주요 구성 요소로는 파일 시스템 관리자, 프로세스 관리자, 메모리 관리자, 그리고 네트워크 관리자가 있다. 장치 관리자와 I/O 관리자는 시뮬레이터의 성능을 고려하여 이 구성에 포함되지 않았다. 파일 시스템 관리자는 파일 및 디렉토리의 구조를 관리하며, 접근 권한에 따라 파일의 읽기, 쓰기, 실행 권한을 제어한다. 또한, 저널링 기능을 통해 데이터의 무결성을 보장하고, 변경된 파일이나 디렉토리를 복구할 수 있도록 지원한다. 프로세스 관리자는 시스템 내에서 프로세스의 생성과 종료를 관리한다. 파일 시스템 관리자로부터 애플리케이션 실행 요청을 받아 메모리 관리자에게 요청을 전달해 애플리케이션이 메모리에 로드되도록 한다. 메모리 관리자는 프로세스 관리자로부터 전달받은 요청에 따라 애플리케이션을 메모리에 로드하고, 필요한 메모리 공간을 효율적으로 할당한다. 이를 통해 애플리케이션이 안정적으로 실행될 수 있도록 지원한다. 네트워크 관리자는 커널의 네트워크 통신을 총괄하며, TCP/IP 계층의 작동을 관리한다. 네트워크 관리자는 소켓을 통해 애플리케이션 간의 데이터 전송을 처리하며, 전송 데이터를 프레임으로 변환하여 네트워크를 통해 전달한다. 이들 관리자는 시스템의 각기 다른 측면을 담당하며, 함께 협력하여 컴퓨터 시스템의 안정성과 효율성을 유지하는 데 기여한다.

2.4 시뮬레이션 환경 구성

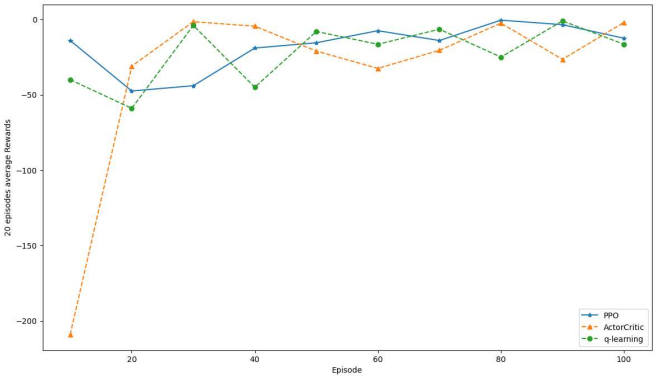
본 시뮬레이션 네트워크 환경은 4대의 컴퓨터로 구성되어 있으며, 각 컴퓨터에는 고유한 IP 주소, 포트 번호, 권한 정보가 부여되어있다. 네트워크 모델링은 사이버배틀심에서 활용된 내부 이동 모델링 기법을 기반으로 설계되었다. 이 시뮬레이션의 목표는 5가지 점령 기법을 통해 컴퓨터 노드들을 점령하는 과정을 분석하는 것이다.

각 컴퓨터의 상태(state)는 비밀번호, 설치된 운영체제, 웹 로그, 방화벽 설정, 열려 있는 포트, 닫혀 있는 포트 등의 정보로 정의된다. 에이전트(공격자)는 이러한 상태 정보를 입력받아 해당 상황에 가장 적합한 공격(action)을 선택하여 수행한다. 가능한 공격 유형으로는 자격 증명 크래킹, 피싱, 운영체제 취약점 악용, 웹 로그 분석, 애플리케이션 취약점 공격이 포함된다. 공격이 성공할 경우 에이전트는 +1의 보상을, 실패할 경우 -1의 보상을 받도록 설정되어 있다. 적용한 강화학습 알고리즘으로는 Q-learning, PPO, AC를 사용하였다.



[그림 1] 공격 시뮬레이션 흐름도

2.5. 실험 결과



[그림 2] Network 환경에 대한 강화학습 알고리즘에 따른 누적 보상 추이

그림 2은 구성된 네트워크 환경에서 강화학습을 적용한 결과를 나타낸다. 학습이 진행될수록 공격자가 획득하는 보상이 점진적으로 증가하는 경향을 보인다. 공격자는 네트워크 노드의 상태에 따라 제공된 5가지 공격 방식을 모두 시도하게 되며, 강화학습을 통해 점차 각 상태에 최적화된 공격 방법을 선택하는 능력을 학습한다. 6차원의 상태 공간과 5차원의 행동 공간을 가진 이 환경에서, Q-learning 알고리즘의 학습 성능은 상대적으로 저조한 것으로 나타났다. 반면, 신경망을 활용한 Proximal Policy Optimization(PPO) 및 Actor-Critic(AC) 알고리즘은 해당 환경에서 우수한 학습 성능을 보였다.

Ⅲ. 결론

ICT와 인공지능의 발전으로 사이버 위협이 커진 현재, 네트워크 보안을 위해 인공지능을 사용하는 방안이 각광받고 있다. 이에 본 연구는 강화학습을 통해 새로운 유형의 공격에도 대응이 가능한 보안 에이전트를 성공적으로 개발하였다. 향후 사회공학적 공격에도 대응이 가능한 에이전트를 개발할 예정이다.

ACKNOWLEDGMENT

본 연구는 정부(교육부, 과학기술정보통신부)의 재원으로 일부 한국연구재단의 지원을 받아 수행(No.2024-0762 ,기여율: 50%)하고 일부 정보통신기획평가원의 지원을 받아 수행된 IITP-2024-RS-2023-00262158, 기여율:50%) 연구임

참 고 문 헌

[1] Sangho Oh Jeongyoon Kim, Jongyoul Park. A study on the development of adversarial simulator for network vulnerability analysis based on reinforcement learning. Journal of The Korea Institute of Information Security and Cryptology, 34(1):21 - 29, 2024.

[2] Kamran Shaukat, Suhuai Luo, Vijay Varadharajan, Ibrahim A Hameed, Shan Chen, Dongxi Liu, and Jiaming Li. Performance comparison and current challenges of using machine learning techniques in cybersecurity. Energies, 13(10):2509, 2020

[3] Guo, Xiaotong, et al. "Automated penetration testing with fine-grained control through deep reinforcement learning." Journal of Communications and Information Networks 8.3 (2023): 212-220.