자동화된 사이버 방어: 검토

SANYAM VYAS, 카디프 대학교, 영국
JOHN HANNAY, 카디프 대학교, 영국
ANDREW BOLTON, 영국 카디프 대학교
PETE BURNAP, 카디프 대학교, 영국

최근 사이버 범죄자들은 다양한 사이버 시스템 내에서 다양한 조직적이고 단호한 사이버 공격을 계획하여 민간 및 정부 기관에 결과적인 영향을 미치고 있습니다. 현재 보안 기반 자동화 및 오케스트레이션은 고정 목적 및 하드 코딩된 솔루션을 자동화하는 데 중점을 두고 있으며, 이는 현대의 사이버 공격으로 쉽게 능가됩니다. 자동화된 사이버 방어에 대한 연구를 통해 순차적인 의사 결정 에이전트를 통해 네트워크 시스템을 자율적으로 방어함으로써 인텔리전스 대응을 개발하고 활성화할 수 있습니다 . 이 기사에서는 자동화된 방어 및 공격 에이전트와 ACO(자율 사이버 운영) 체육관이라는 두 가지 하위 영역으로 구분된 요구 사항 분석을 통해 자동화된 사이버 방어 내 개발을 포괄적으로 설명합니다 . 요구 사항 분석을 통해 자동화된 에이전트를 비교할 수 있으며 지속적인 개발을 위한 ACO Gym의 중요성을 강조합니다. 요구 사항 분석은 실제 네트워크 시스템 내에 자동화된 에이전트를 배포하기 위한 전반적인 목표를 가지고 ACO 체육관을 비판하는 데에도 사용됩니다 . 자동화된 사이버 방어 분야의 개발을 가속화하기 위해 전반적인 분석에서 관련 미래 과제가 해결되었습니다 .

CCS 개념: • 컴퓨터 시스템 구성 → 인공 지능; 강화 학습; • 보안 및 개인 정보 보호 → 네트워크 보안; 악성 코드 완화.

추가 핵심 단어 및 문구: 자동화된 사이버 방어, 강화 학습, 침입 대응, 네트워크 보안

ACM 참조 형식: Sanyam
Vyas, John Hannay, Andrew Bolton 및 Pete Burnap. 2023. 자동화된 사이버 방어: 검토. 진행 ACM 측정 항문. 계산. 시스템. 37, 4, 제111조(2023년 2월), 32 쪽. https://doi.org/XXXXXXX.XXXXXXX

1. 소개

전 세계의 개인, 조직 및 정부는 일상 활동 및 운영에서 디지털 참여가 기하급수적으로 증가하는 상황에 직면해 있습니다. 이를 통해 정보 인식 및 통신 측면에서 전 세계가 효율적으로 연결되었지만 위에서 언급한 모든 개체는 기회주의자, 범죄자, 적대 국가 등 다양한 공격자의 사이버 공격에 지속적으로 직면하고 있습니다. 가정과 산업 내에서 이러한 정보 기술(IT) 및 운영 기술(OT) 장치가 기하급수적으로 증가하고 기술 부족이 증가함에 따라 발생하는 사이버 공격의 양이 증가하는 대신 사이버 보안 인프라와 조직이 압도당하게 되었습니다 . 숙련된 사이버 보안 실무자 가 절실히 필요한 반면 , 최근의 새롭고 자동화된 사이버 공격의 수준은 사이버 보안의 능력을 능가합니다.

저자 주소: Sanyam Vyas, vyass3@cardiff.ac.uk, Cardiff University, Cardiff, Wales, United Kingdom, CF24 4AG; John Hannay, hannayj1@cardiff.ac.uk, 카디프 대학교, 영국 웨일즈 카디프, CF24 4AG; Andrew Bolton, boltona2@cardiff.ac.uk, 카디프 대학교, 카디프, 웨일스, 영국, CF24 4AG; Pete Burnap, burnapp@cardiff.ac.uk, 카디프 대학교, 카디프, 웨일즈, 영국 , CF24 4AG.

2476-1249/2023/2-ART111 $15.00
https://doi.org/XXXXXXX.XXXXXXX

인간은 수동으로 방어합니다 [19]. 따라서 이제 이러한 시스템에 대한 위협을 관리하기 위해 IT 및 OT 인프라 내에 자동화된 방어 솔루션 [63, 77] 을 구현하는 것이 절실히 필요합니다 . 결과적으로 문헌에는 자동화된 방어 솔루션이 있었지만 최근 사이버 공격을 방어할 수 없기 때문에 금전적 및 지적 재산 기반 피해를 제한하기 위해 더 많은 연구가 수행되어야 합니다.

이 문제에 대한 해결책으로 매우 복잡한 사이버 공격을 완화하기 위해 네트워크로 연결된 시스템에 대한 자동화된 의사 결정 에이전트에 중점을 둔 영역인 자동화된 사이버 방어를 소개합니다. 이 문서에서는 ACD를 정의하고 자동화된 사이버 방어의 다양한 부서 내 문헌을 분석합니다. 분석은 네트워크 시스템 내에서 자동화된 의사 결정 에이전트 의 실제 배포에 초점을 맞춘 비전을 가지고 전반적인 요구 사항 분석을 통해 수행됩니다. 전반적으로 네트워크 시스템 내에서 사이버 공격을 방어하기 위한 자동화된 의사 결정 에이전트의 요구 사항을 강조한 출판물은 거의 없습니다 . 최근 간행물에는 이동 표적 방어, 사이버 방어 및 허니팟을 위한 강화 학습(RL) 솔루션에 대한 자세한 검토를 제공하는 [61] 이 포함되어 있습니다 . 또한 이 간행물은 최적의 제어 이론 원리와 최신 AI 개발을 통해 사이버 보안 내에서 RL 솔루션의 세부 개발을 제공합니다 . 그러나 검토에서는 네트워크 방어 및 공격을 위한 자동화된 의사 결정 에이전트의 신속한 개발을 위한 영역에 초점을 맞추지 않습니다. Wang 등 [123]은 또한 [61] 과 유사한 미래 과제를 해결하면서 네트워크 방어 및 공격을 위한 RL 솔루션 개발에 중점을 두고 있습니다 . 그러나 이 논문에서는 자동화된 사이버 방어의 필수적인 부분인 그러한 에이전트가 개발될 수 있는 지형을 철저하게 분석하지 않았습니다 . [22] 의 저자는 사이버 보안 내의 기계 학습(ML) 솔루션에 대한 리뷰를 제공하며 , 특히 침입 탐지 시스템 내에서 연구를 가속화하는 데이터 세트에 중점을 둡니다.

이 구현은 사이버 보안 내 자동화 접근 방식의 일부를 구성하지만 침입 탐지 접근 방식은 자동화된 사이버 대응을 포함하지 않으므로 이 백서에 정의된 자동화된 사이버 방어 용어에는 적용되지 않습니다. Burke 등 [23] 은 AcCD(Active Cyber Defense) 내에서 잠재적인 프로젝트 유형에 대한 심층적인 검토를 제공하며 , 그 중 일부는 이 백서에서 언급된 자동화된 사이버 방어라는 용어에도 적용됩니다. 그러나 보고서는 개발을 위한 ACO 체육관에 초점을 맞추지 않으며 자동화된 네트워크 공격 및 방어에 사용되는 자동화된 의사 결정 알고리즘에 대한 자세한 비교도 제공하지 않습니다 .

이 기사의 나머지 부분에는 이 논문에서 자주 사용되는 다양한 용어를 먼저 정의하는 섹션 2 가 포함되어 있습니다 . 그런 다음 섹션 1 에서는 관련 ACD 출판물을 찾는 데 사용되는 방법론을 설명합니다. 그런 다음 섹션 4 에서는 자동화된 사이버 방어의 선별된 용어와 최근 문헌에서 사용되는 유사한 용어와의 차이점을 자세히 설명합니다. 그런 다음 이 섹션에서는 국가 전략에서 해당 영역의 중요성을 제공합니다 . 마지막으로, 이 섹션 에서는 자동화된 사이버 방어의 일부로 인식된 선택된 간행물을 평가하는 데 사용되는 포괄적인 요구 사항 분석을 제공합니다 . 섹션 5 에서는 섹션 4 의 요구 사항 분석을 통해 맞춤형 ACO 체육관의 자동화된 방어 및 공격(파란색 및 빨간색) 에이전트에 대해 자세히 설명하고 비평 합니다. 섹션 6 에서는 오픈 소스 및 폐쇄 소스 ACO 체육관의 전체 목록을 자세히 설명하고 요구 사항을 사용하여 평가합니다. 섹션 4의 분석 . 섹션 5 에서는 ACO 체육관 내에 게시된 자동화 에이전트 목록을 자세히 설명 하고 섹션 4 의 요구 사항 분석을 사용하여 이를 평가합니다. 섹션 8 에서는 이전에서 수행된 평가를 사용하여 자동화된 사이버 방어 문헌의 과제와 격차를 식별하는 논의를 제공합니다. 섹션. 마지막으로 섹션 9 에서는 자동화된 사이버 방어 분야를 요약하여 기사를 마무리합니다 . 이 문서의 기여는 다음과 같습니다.

- 자동화된 사이버 방어라는 용어의 정의를 완료하고 자동화된 사이버 방어와 비교하여 해당 연구를 구별합니다. 기타 관련 용어.
- 자동 사이버 방어의 두 가지 중요한 영역, 즉 개발 기준의 요구 사항을 강조하는 정의된 자동 사이버 방어 분야에 대한 요구 사항 분석 개발

자동화된 블루 및 레드 에이전트와 ACO 체육관의 개발 기준을 통해 자동화된 사이버 방어 기능을 촉진합니다.

- 요구 사항을 통해 Automated Cyber Defense 문헌 내 출판물 평가
  분석.
- 미래 소설 연구를 강조하기 위해 문헌 내에서 새롭고 현실적인 과제를 식별합니다.
  지도.

2가지 주요 정의

이 문서는 사이버 보안 및 인공 지능 분야에서 일반적으로 사용되는 여러 기술 용어로 구성됩니다 . 이 섹션에서는 이 문서에서 사용되는 주요 용어를 정의합니다.

자동화된 레드팀 구성(Automated Red Teaming): 레드 팀 구성은 예상치 못한 상황을 줄이고, 네트워크 시스템의 견고성을 개선 및 보장한다는 전반적인 목표를 가지고 네트워크 시스템의 취약점을 찾아내거나 작전 개념에서 악용 가능한 격차를 찾기 위해 군사 및 산업 운영 내에서 사용되는 기술입니다 . [27]. 이 백서의 맥락에서 자동화된 레드 팀 구성은 일련의 작업(네트워크 시스템 내의 취약점과 악용을 발견하기 위한)을 작업 공간으로 보유하는 자율 에이전트를 의미합니다. 자동화된 레드팀 구성의 전반적인 목표는 알려진 취약점 및 악용으로부터 시스템을 보호하는 측면에서 자동화된 블루팀 에이전트(아래에 설명된 정의)의 견고성을 보장하는 것입니다 .

자동화된 블루 팀 구성: 블루 팀 구성은 네트워크 시스템 의 공백과 취약성을 악용하려는 모의 공격자 그룹에 대해 보안 태세를 유지함으로써 네트워크 시스템을 방어하는 기술입니다. 일반적으로 블루팀과 그 지지자들은 1) 상당한 기간에 걸쳐, 2) 대표적인 작전 상황(예: 작전 훈련의 일부)1 에서 실제 또는 시뮬레이션 공격을 방어해야 합니다 .

본 백서의 맥락에서 자동화된 블루 팀 구성은 노드/엔드포인트를 통해 네트워크 시스템에 진입하는 악성 프로세스를 파괴하기 위한 작업 공간으로 일련의 작업을 보유하는 자율 에이전트를 의미합니다.

자율 사이버 작전 체육관(Autonomous Cyber Operations Gym): 자율 사이버 작전(ACO) 은 자율적인 의사 결정과 행동을 통해 컴퓨터 시스템과 네트워크를 방어하는 것과 관련이 있습니다 . 모든 네트워크와 위치를 포괄하는 보안 전문가를 배치하는 것이 점점 더 어려워지고, 신뢰할 수 없는 통신 채널이나 적의 행동으로 인해 인간 방어자가 시스템에 안정적으로 액세스할 수 없는 경우에는 특히 필요합니다 . ACO 체육관은 끊임없이 진화하는 사이버 공격 으로부터 미래의 네트워크 시스템을 더욱 강화하기 위해 자율적인 레드 및 블루 팀 구성 에이전트 의 사용을 촉진하는 네트워크 시스템 환경입니다 [112]. ACO 체육관은 시뮬레이션 학습과 실제 환경 테스트를 결합하여 [114] 에서 사용되는 잠재적인 네트워크 시스템 의 '현실 격차'를 해결하고 줄이는 것을 목표로 합니다 .

순차 응답: 순차 응답 또는 순차적 의사 결정은 세계의 역학을 고려하여 문제가 해결될 때까지 문제의 일부를 지연시키는 알고리즘을 의미합니다 [42]. 환경 내에서 전반적인 목표로 작용하는 상태에 도달하기 위해 단기 및 장기 결정과 관련된 일련의 결정을 요구하는 것은 환경과의 확장된 상호 작용에서 모든 지능형 에이전트가 직면하는 근본적인 작업 입니다 .[73]. 이 백서의 맥락에서 자동화된 에이전트가 대상 작업을 수행하기 전에 탐색이 필요한 네트워크의 복잡성으로 인해 이 백서에서는 순차적 의사 결정 알고리즘을 자동화된 블루 및 레드 팀 에이전트로 간주합니다 (예: 다른 서브넷 내의 호스트) .

단일 단계 대응: 단일 단계 대응 알고리즘은 단기 결과 에만 초점을 맞춘 의사 결정 조치를 나타냅니다 . 예를 들어, 시간적 맥락에서 시간 t(n)의 알고리즘은 시간 t(n + 1)의 솔루션에 대해서만 계산을 수행합니다 .

1https://csrc.nist.gov/glossary/term/blue_team

시뮬레이션된 네트워크: 시뮬레이션된 네트워크는 유한 상태 기계로 설계된 ACO Gym(또는 ACO Gym의 훈련-테스트 전략의 일부)입니다 . 생성은 일반적 으로 시뮬레이션된 네트워크 내의 구성 요소, 에이전트 및 작업에 해당하는 개체를 포함하는 코드 형식으로 완료됩니다 . [83]
에뮬레이트된 네트워크: 에뮬레이트된 네트워크 는 컴퓨터 네트워크 시스템을 생성하는 데 사용되는 가상 머신 그룹을 통해 설계된 ACO Gym(또는 ACO Gym의 훈련-테스트 전략의 일부)입니다 [83].

3 검토 방법론

이 리뷰에 대한 모든 관련 기사를 찾기 위해 [67] 에서 영감을 받은 방법론이 구현되었습니다. 전반적인 ACD 정의와 이 기사의 연구 질문을 선별하기 위해 국내 및 국제 정부 기관과 민간 조직(섹션 4.1에서 언급) 의 논문을 활용했습니다. 이 문서에서는 다양한 영역 내 네트워크 시스템의 자율 대응 솔루션에 대한 필요성을 다루었습니다. 이를 통해 자동화된 사이버 방어 용어의 기존 영역 , 특히 자동화된 레드 및 블루 팀 내에서 자율 대응이 활용될 수 있는 영역을 분류할 수 있었습니다.

3.1 연구 질문 자동화된 사이버

방어에 대해 제시된 아이디어를 활용하기 위해 검색 전략을 파악하기 위한 연구 질문을 작성했습니다. 이를 통해 우리는 이 기사와 관련된 문헌을 찾을 수도 있었습니다. 세 번째 연구 질문은 검색 전략(다음 하위 섹션에서 자세히 설명)에 의해 식별되었으며 전체 검색 전략의 업데이트로 이어졌습니다.

연구 질문(RQ)에는 다음이 포함됩니다.

- RQ1: 자동화된 사이버 방어란 무엇입니까? • RQ2:
  자동화된 사이버 환경에서 사용된 가장 적합한 알고리즘 접근 방식은 무엇입니까?
  RQ1을 통해 정의된 국방 용어는 무엇입니까 ?
- RQ3: 가장 적합한 알고리즘 접근 방식이 가능한 최상의 환경은 무엇입니까?
  개발?

3.2 검색 용어 전략 모든 연구 질문
을 전체적으로 식별한 후 다음 단계는 관련 1차 연구를 검색하는 것입니다.

중요한 관련 작업을 간과하지 않기 위해 IEEE, ACM Digital Library, Springer 및 Science Direct를 포함한 인기 있는 디지털 라이브러리를 Google Scholar와 함께 활용합니다 . 자동화된 사이버 방어의 3가지 전체 주제 내에서 그룹화된 문자열 목록을 종합적으로 식별했습니다(표 1 참조). 그런 다음 디지털 도서관에서 다음과 같은 출판물을 식별하기 위한 목적으로 다양한 전체 테마의 문자열을 3가지 다른 순열 조합 그룹으로 그룹화합니다 .

- i: Automated Blue에서 식별된 알고리즘 계열의 성능을 탐색하고 순위를 매길 수 있습니다.
  팀 구성 (RQ1, RQ2).
- ii: Automated Red에서 식별된 알고리즘 제품군의 성능을 탐색하고 순위를 매길 수 있습니다.
  팀 구성 (RG1, RQ2).
- iii: 가장 적합한 알고리즘이 존재할 수 있는 최상의 환경을 발견할 수 있습니다.
  개발, 교육 및 테스트되었습니다 (RQ1, RQ3).

3.3 전반적인 관련 콘텐츠 추출

최근에야 자동화된 사이버 방어(Automated Cyber Defense) 분야가 인기를 얻고 있기 때문에 검색 전략에 나열되지 않은 자동화된 사이버 방어(Automated Cyber Defense)로 식별되는 출판물을 찾기 위해 여러 검색에 대한 역방향 눈덩이 뭉치 [65] 를 수행했습니다. 예를 들어, 최근 '자율사이버운영체육관' 등의 분야가

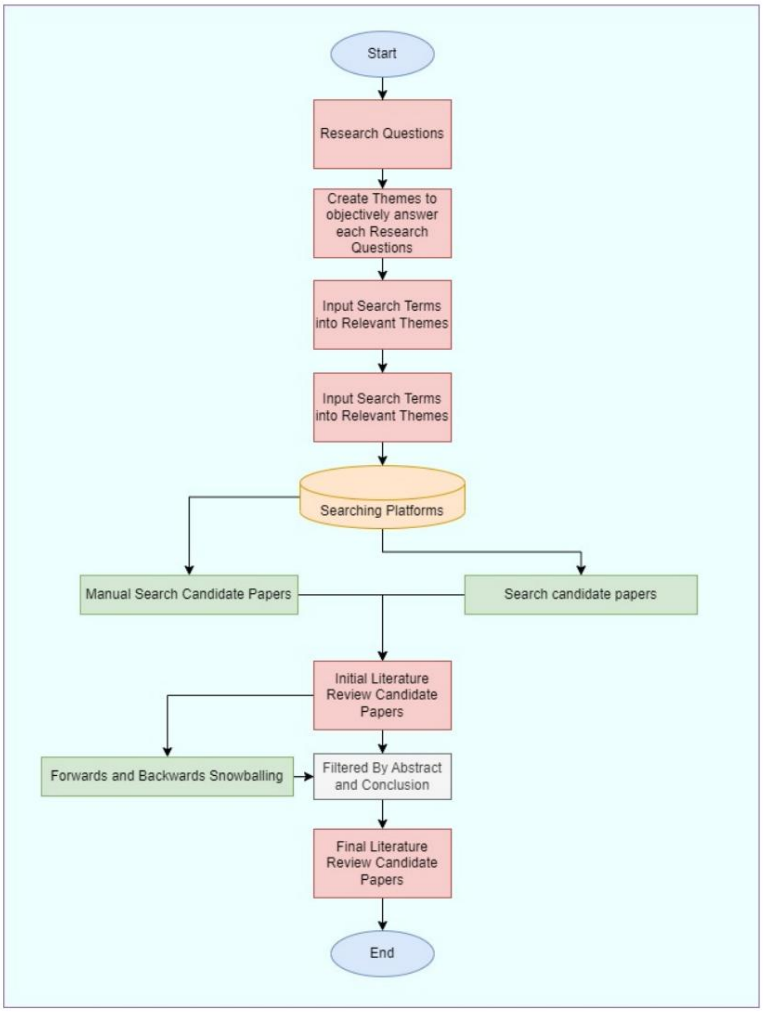

그림 1. 연구 방법론

이 영역 내에서 용어가 생성되었으므로 역방향 눈덩이를 사용하면 이 용어가 공식적으로 도입되기 전에 생성된 다른 인기 출판물(및 코드 저장소에서 발견된 구현)을 식별하는 데 도움이 되었습니다 . 또한 검색 전략을 통해 식별된 출판물을 인용 한 최신 Automated Cyber Defense 관련 논문(도메인 내의 추가 잠재적 영역을 강조하는 논문과 함께)을 식별하기 위해 수동 검색을 수행했습니다. 선정된 모든 논문은 자동화된 사이버 방어의 정의 전체적 정의를 바탕으로 논문의 범위를 정렬하기 위해 논문 초록을 바탕으로 또 다른 심사 과정을 거쳤습니다 . 마지막으로 나머지 논문을 완전히 읽고 추가 심사를 위해 분석했습니다. 그림 1 은 이 검색 방법론에 포함된 전체 단계를 보여줍니다 .

표 1. 검색 용어에 대한 중요한 주제를 제안합니다. ㅏ. 다음을 통합하는 알고리즘과 용어가 포함됩니다.
자동화된 대응(자동 사이버 방어 에이전트의 요구 사항) 비. 의 일부인 모든 용어를 포함합니다.
정의된 Automated Blue Teaming 용어. 씨. 정의된 자동화의 일부인 모든 용어를 포함합니다.
레드팀 용어.

ㅏ. 알고리즘 구혼	비. 자동화됨 블루팀	씨. 자동화됨 레드팀
- "인공지능" - "머신러닝" 또는 "딥 러닝" - "오픈 AI" AND "gym" - "강화 학 습" - "게임 이론" - "생성 모델링" - "자동화" 또는 "자 동"	- "자율사이버운영체육관" - "악성코드" - "프로세스 종료" - "사이버 방어" 또는 "사이버 방어" - "맬웨어" - "기만" - "대응"	- "프로세스" - "침투" AND "테스트" - "공격적인 사이버 보안" - "자율 악성코드" - "권한 승격"
또는 "자율" 또는 "오토메이션" - "응답"	- "워게임" 또는 "워게임" - "사이버 탄력성" - "고급 지속 위협" 또는 "아파트" - "블루 팀" 또는 "블루 팀" 또는 "블루 팀" - "사이버 위협 인텔리전스"	- "적대적 에뮬레이션" - "워게이밍" 또는 "워게임" - "레드팀" 또는 "레드팀" 또는 "레드팀 구성" - "정찰" - "자율사이버운영체육관" - "사이버 방어" 또는 "사이버 방어" - "기만"

4 자동화된 사이버 방어

ACD(자동 사이버 방어)는 최근 몇몇 출판물과 뉴스에서 언급된 주제입니다.
지난 몇 년간 발생한 사이버 공격의 증가에 비추어 지난 10년 동안의 기사를 살펴보았습니다. ~ 안에
이 용어를 정의하기 위해 간단한 검토가 완료되었습니다.

Rege et al [98] 은 전문가 수준의 의사결정 시스템으로서 ACD 알고리즘에 대한 높은 수준의 설명을 제공했습니다.
자동화된 블루 에이전트를 생성하는 출판물 [16] 을 인용하여 인간이 추론하고 학습하는 방식에서 영감을 얻은 능력
맞춤형 네트워크 시스템 내에서. Ko et al [68]은 목적을 정교화할 때 ACD에 대한 용어를 제공했습니다.
DARPA(Defense Advanced Research Projects Agency) 그랜드 챌린지에서 ACD를 다음과 같이 설명했습니다. [2],
사람의 개입 없이 실시간으로 소프트웨어 취약점을 자체 발견, 증명 및 수정할 수 있는 시스템
간섭. 2016년에 Baah 등 [15]은 ACD 시스템의 일반화된 개요를 제공했습니다. 설명된 논문
ACD는 현재 진행 중인 공격이나 네트워크의 기존 취약성을 감지하는 것으로 시작되는 대응입니다.
이 논문에서는 위험을 완화하기 위한 조치를 취하려면 탐지 속도와 정확성이 중요하다는 점을 강조했습니다.
네트워크 자산에 손상을 입히거나 임무를 방해하기 전에 위협을 차단합니다. 그것은 또한 기계의 해결책을 조명합니다
의심스러운 네트워크 활동과 양성 네트워크 활동을 구별할 수 있는 분석 학습 및 자동화된 퍼징
이전에 알려지지 않았던 소프트웨어의 취약점을 발견할 수 있는 기술. Benjamin 등 [16] 은 ACD를 정의합니다.

2https://www.darpa.mil/program/cyber-grand-challenge

CSISM(지능형 생존 가능성 관리를 위한 인지 지원)이라는 프로젝트를 통해 저자는 전문가 수준의 능력을 갖춘 자동화된 사이버 방어 의사 결정 메커니즘을 구현합니다. ACD 시스템은 경고와 관찰을 해석한 다음 네트워크 컴퓨팅 기능의 생존 가능성을 보장하기 위해 방어 조치를 취합니다. 저자들은 불확실하고 불완전한 정보를 바탕으로 전문가 수준의 응답을 실시간으로 생성하는 것이 어려운 목표라는 것을 알고 있습니다. 그러나 그들은 자동화된 추론 개발과 사이버 방어 작전을 위한 인지 아키텍처 사용을 통한 학습 사이에 디딤돌이 있다는 것을 알고 있습니다 .

Alan Turing Institute의 Burke 외 [23] 는 백서를 통해 AcCD(능동적 사이버 방어)에 초점을 맞춘 연구 이니셔티브를 소개했습니다. 이 연구 이니셔티브는 네트워크 방어자와 사이버 보안을 강화하기 위해 기업 내 자동화 향상을 추구하는 데 중점을 둡니다. AcCD와 ACD라는 용어 사이의 차이점을 해결하는 것이 중요합니다. AcCD에는 자동 보안 플래너가 포함되어 있다는 점입니다. 반면 ACD는 주로 자동화 된 블루 팀 구성 에이전트의 전반적인 개발을 위해 자동화된 레드 및 블루 팀 구성에 엄격하게 중점을 둡니다. 전반적으로, 이 문서에서는 시스템 방어자가 고도로 자동화된 미래 위협 및 공격으로 인한 위험을 관리 하고 사이버 관련 국가 규모에서 시스템을 방어할 수 있도록 하려면 지능형 자동화가 필수적이라고 설명합니다. 또한 백서는 자동화된 레드 및 블루 팀 구성의 필요성을 자세히 설명했습니다. 그러나 해당 분야의 연구 방향에 대한 높은 수준의 정보만 제공했습니다 . 최소한의 오류로 복잡한 사이버 공격을 탐지하고 대응하기 위해 지형(즉, 네트워크로 연결된 시스템)을 지능적으로 이해하는 방법으로 이러한 시스템 내에서 인공 지능의 사용이 제안되었습니다 .

Applebaum et al [13] 은 테이블 형식 Q-학습을 기반으로 하는 ACD(자율 사이버 방어) 에이전트를 활용하는 논문에서 자율 사이버 방어라는 용어를 강조합니다 . 이 용어는 자율 사이버 방어란 ML 기술을 활용하여 시스템을 자율적으로 방어할 수 있는 에이전트를 훈련하고 시끄러운 센서 데이터를 사용하는 응답으로 인한 자체 손상을 최소화하는 것임을 시사합니다. 이 정의는 ACD의 정의와 일치하지만 정의는 매우 간단 하며 자동화된 레드 및 블루 팀 에이전트와 함께 자율 사이버 운영 체육관의 병렬 개발을 포함하는 ACD의 전체 영역을 이해하기 위해 확장되어야 합니다 . 또한 정의된 용어에는 자동화된 레드팀 에이전트의 역할이 포함되어 있지 않습니다. 이는 이 백서에서 ACD의 필수적인 부분으로 다뤄집니다.

자율 사이버 작전(ACO)의 정의 는 ACO와 비교하여 ACD 내의 구체적인 연구 방향을 명확히 하기 위해 ACD와 관련하여 다루어질 필요가 있습니다 . [112] 는 ACO 를 게임 플레이 시나리오에서 서로 싸우는 네트워크 시스템 내에서 자동화된 빨간색(공격자) 에이전트와 자동화된 파란색(방어자) 에이전트의 병렬 개발로 정의합니다 . ACD는 자동화된 블루 에이전트의 전반적인 개발에 중점을 두고 있다는 점에서 ACO와 다릅니다. 자동화된 레드 에이전트는 특히 적의 훈련을 촉진하는 자동화된 침투 테스트 에이전트 로 설계되었습니다 . ACD의 관점에서 ACO 체육관 개발은 자동화된 블루 팀 에이전트 개발을 위해 특별히 설계되어야 한다는 점에서 AI ACO 체육관 개발과 다릅니다.

위에서 언급한 모든 문헌을 종합할 때 우리는 자동화된 사이버 방어를 매우 복잡한 사이버 공격을 완화하기 위해 사이버 시스템(예: 기업 네트워크, 산업 제어 시스템)에 대한 자동화된 의사 결정 에이전트에 초점을 맞춘 용어로 정의합니다. ACD 시스템의 개발은 다양한 유형의 작업을 조합하여 수행할 수 있습니다. 여기에는 자동화된 레드팀 에이전트를 사용하여 자동화된 블루팀 에이전트를 검증, 개발 및 강화하는 데 사용되는 지형 모드(실제 사이버 시스템을 복제하기 위해)로서 자율 사이버 운영 체육관 내에서 자동화된 블루팀 에이전트 의 개발이 포함됩니다. 실제 개발의 전반적인 목표.

4.1 국가 전략 문서 내 자동화된 사이버 방어의 중요성 민간 및 정부 기반 조직은 공격

탐지 및 대응 측면에서 AI가 곧 사이버 보안 내에서 최전선에 설 것이라는 점을 분명히 했습니다 . 표 2 는 여러 국가에서 ACD의 중요성을 자세히 설명합니다 .

표 2. ACD에 관한 국가 전략 보고서

국가/연맹부서/전략		ACD에 대한 참조는
호주	국방부 [111]	사이버 보안 기술을 확장하고 여기에 AI를 통합할 필요성을 제안합니다. 국방부는 정보 및 사이버 영역 전반에 걸쳐 역량을 강화하기 위해 AI 역량에 대한 연구와 투자를 조정하고 있다.
	의사결정 이니셔티브를 위한 AI 2022 [9]	사이버 공격에 즉각적으로 대응하기 위한 AI 기반 자동화 의사 결정 블루 팀 구성 알고리즘을 생성하는 TTCP CAGE 자율 사이버 방어 챌린지를 포함하여 연구원을 위한 AI 기반 챌린지 30개를 추가로 개발하는 것을 목표로 합니다.
	호주 왕립공군 [34]	기계가 결정을 내릴 수 있고 인간이 결정해야 하는 지속적인 평가를 조언합니다.
캐나다	국가 사이버 보안 전략 [33]	자율적 의사 결정 지원을 갖춘 국방 및 보안 애플리케이션의 중요성을 구체적으로 언급했습니다. 이 출판물에서는
	국방연구개발 [36]	진화하는 위협을 정확하게 식별하기 위해 딥 러닝과 RL 알고리즘의 조합을 제안하고 적절한 조치 과정을 권장하거나 실행합니다.
영국	국방인공지능전략 [120]	적시에 인간 운영자의 조치를 방해하는 속도와 규모로 작동하는 AI 강화 사이버 위협의 새로운 위험에 대해 논의합니다.
	정부 사이버 보안 전략 [90]	AI를 주목해야 할 새로운 기술로 설명했습니다. 악의적인 활동을 탐지하는 맥락에서 AI를 탐색하고 경우에 따라 "위협에 대한 자동화된 대응을 활성화"할 것을 제안합니다.
나토	협력적 사이버 방어 우수 센터 [85]	국가가 AI 기반 사이버 방어를 채택하고 탐색해야 할 필요성을 제안합니다.
	NATO AI 전략 [84]	전략에는 '사이버 방어를 위한 AI 기술 협력'이 포함된다.

4.2 자동화된 사이버 방어 요구 사항 NATO(북대서양 조

약 기구)는 참조 아키텍처 및 기술 로드맵인 AICA를 생성하여 자율 사이버 에이전트에 대한 요구 사항을 설명했습니다 [69]. 문서의 특정 부분은 자율 에이전트의 전장에서의 전략적 배치와 윤리적 우려에 중점을 둡니다. 본 문서와 관련된 AICA의 핵심 사항은 아래 표 3 의 ACD에 대한 요약된 요구 사항 분석에 추가 ACD 요구 사항과 함께 포함되었습니다 . 이 표에는 자동화된 빨간색 및 파란색 에이전트 의 사용을 허용할 수 있는 ACO Gyms 요구 사항 (G) 과 함께 자동화된 빨간색 및 파란색 에이전트 의 필수 요구 사항 (A)이 자세히 설명되어 있습니다. 이 표의 요구 사항은 ACD 내 연구자를 위한 체크리스트 역할을 하여 실제 네트워크 시스템 내에서 ACD 운영의 최종 배포를 개발할 수 있도록 해야 합니다.

표 3. ACD 요구 사항

요구 사항 요약	
일반화	- (G.1.1) ACO Gym은 새로운 설정으로 일반화하고 구성 요소를 원활하게 추가할 수 있는 능력을 갖추어야 합니다. - (G.1.2) ACO Gym은 다양한 유형의 에이전트를 추가할 수 있어야 합니다. - (G.1.3) 네트워크 시스템 교육-테스트는 네트워크 시스템 내에서 실제 네트워크 시스템 대기 시간 작업 지연을 일치시키는 것과 같은 측면을 포함하여 시뮬레이션에서 실제 설계로의 전환을 촉진해야 합니다. 예를 들면 교육-테스트 전략 내에서 시뮬레이션과 에뮬레이션의 하이브리드가 포함됩니다. - (G.1.4) ACO Gym은 구성 문제 없이 네트워크를 더 큰 크기로 확장할 수 있는 기능을 갖추고 있어야 합니다. - (A.1.1) 자동화된 에이전트는 다음과 관련된 결정을 일반화해야 합니다. 그것이 나타내는 에이전트 유형 - (A.1.2) 자동화된 에이전트는 ACO Gym 내 구조적 변화(서브넷 및 엔드포인트 추가 및 제거)에 일반화하고 적응해야 합니다. - (A.1.3) 자동화된 빨간색 및 파란색 에이전트는 시뮬레이션에서 실제까지 높은 성능을 유지하도록 설계되어야 합니다.
높은 레벨 의사결정	- (G.2.1) ACO 체육관은 네트워크 시스템 내에서 특정 이벤트가 발생한 후 상태를 설명할 수 있도록 설계되어야 합니다. - (G.2.2) ACO 체육관은 자동화된 결정이 내려질 수 있도록 MDP/POMDP 형식으로 구성되어야 합니다. - (A.2.1) 계획 및 집단 대응 계획을 위해서는 순차적 알고리즘을 고려해야 합니다. - (A.2.2) AICA 참조 아키텍처는 게임 이론과 인공 지능 모두 ACD 내 구현에 적합하다고 주장합니다. - (A.2.3) 설계된 자동화 에이전트에는 ACO 체육관의 복잡성을 유지하기 위한 "심층" 아키텍처가 필요합니다. - (A.2.4) 추가적으로 에이전트는 설명할 수 있어야 합니다 [21, 66, 97]. 즉, 실제 네트워크 시스템 내에서 작동할 수 있도록 실시간 결정을 내리는 것을 정당화합니다.
학습	- (A.3.1) AICA [69] 는 ACO 체육관 내에서 지속적인 학습을 가능하게 하는 가능성을 제시합니다. - (A.3.2) 또한 훈련-테스트 접근 방식의 중요성도 주장합니다. - (G.4.1) ACO 체육관은 다음
다중 에이전트 협동	과 같이 설계되어야 합니다. 다중 에이전트 강화 학습(MARL) 작동을 허용하는 방법 - (A.4.1) 다중 에이전트 시스템 표현은 자동화된 에이전트를 교육하고 작업/전략 협상을 위해 필요합니다. [8]. AICA는 [125] 에서 제작한 MARL 설문 조사와 결합하여 최소한의 커뮤니케이션 접근 방식과 중앙 집중식 교육 및 분산 실행 솔루션의 조합을 활용할 것을 제안합니다.
연구 협동	요구 사항은 ACD 와 일치하는 ACD [23] 내의 다른 연구자들과 설명하고 협력해야 한다는 것입니다. 따라서: - (G.5.1) ACO 체육관은 연구자들이 구현에 추가로 기여할 수 있도록 오픈 소스여야 합니다. - (G.5.2) ACO 체육관에 대한 문서는 체육관의 추가 개발과 체육관 내 자동화 에이전트의 연구 및 구현을 쉽게 할 수 있도록 제공되어야 합니다. AICA 참조 아키텍처는 다양한 악성 코드 샘플 및 기타 알고리즘 공격
회복력	에 대한 복원력의 필요성을 강조합니다. 따라서: - (G.6.1) ACO 체육관은 자동화된 레드 에이전트가 자동화된 블루 에이전트를 적대적으로 훈련하여 잘못된 작업 수를 줄일 수 있도록 설계되어야 합니다. - (A.6.1) 자동화된 블루 팀 에이전트의 성능을 향상하려면(단독 ACD 의 목적) 자동화된 레드 에이전트를 통한 적대적 훈련을 권장해야 합니다. - (A.6.2) 자동화된 레드 에이전트에는 다양한 사이버 공격이 제공되어야 합니다(MITRE ATT&CK 프레임워크 내에서 지정). - (A.6.3) 시스템 취약성을 해결하기 위한 다양한 알고리즘 공격 [59] 과 함께 . - (A.6.4) 자동화된 청색 및 적색 에이전트는 각각 기망 방어 및 공격을 실행할 수 있어야 합니다.

맞춤형 ACO 체육관에서 사용되는 5가지 ACD 알고리즘

섹션 4.2 에서 언급했듯이 일반적인 ACD 시스템 은 자동화된 레드 팀과 블루 팀 게임 플레이 시나리오를 허용하는 기능을 갖춘 지형 모드, 즉 네트워크 시스템으로 구성됩니다 . ACD 내의 최근 출판물에서는 맞춤형 ACO 체육관 내에서 자동화된 블루 및 레드 팀 구성을 위해 게임 이론(GT), 기계 학습(ML) 및 강화 학습(RL)과 같은 자동화된 의사 결정 알고리즘을 활용했습니다 .

ML 기반 솔루션(RL 기반 솔루션 [87, 99, 107] 과 함께) 도 수년 동안 신속한 사고 및 침입 대응을 위해서만 활용되었습니다 [50, 89, 118]. 특히 Zago 등 [126] 은 ML 기술을 활용하여 봇넷을 포함한 기존 및 향후 사이버 위협을 분석, 탐지 및 대응합니다. 제안된 접근 방식은 비지도 접근 방식과 감독 접근 방식을 결합하여 오류율을 줄이고 계산 리소스 측면에서 효율성을 높이는 확장 가능한 탐지 및 반응 프레임워크를 만듭니다 . 이 접근 방식은 차원 축소 알고리즘을 사용하고 구현을 위해 침입 탐지를 위해 공개적으로 사용 가능한 데이터 세트를 사용합니다.

이와 같은 유일한 ML 기반 구현은 특정 유형의 공격을 완화할 수 있지만 단일 단계 대응 특성으로 인해 제로데이 공격의 신속한 대응을 제공하지 못합니다. 또한 제로섬 GT 기반 솔루션과 마찬가지로 알고리즘이 작동 중인 시나리오에서 더 멀리 떨어진 상태 공간을 일반화할 만큼 복잡하지 않기 때문에 성능이 대규모 엔터프라이즈 네트워크로 확장되지 않습니다 . Cam et al [24] 은 또한 대부분의 ML 기반 솔루션(지도 학습 및 비지도 학습 알고리즘 포함)이 단일 단계 학습 문제에 대한 솔루션을 제공하는 방법을 강조합니다. 이는 알고리즘을 ACD 기반 솔루션으로 구현하는 것이 불가능하게 만드는 알고리즘의 기능입니다 . 네트워크로 연결된 시스템 내에서. 따라서 이 섹션에서 선택한 출판물은 자동화된 에이전트가 전체 네트워크 시스템 내에서 사이버 공격을 중지하는 데 필요한 순차적 대응에 중점을 둡니다.

이 섹션의 나머지 부분에서는 맞춤형 네트워크 시스템 내에서 각각 블루팀과 레드팀에 대한 자동 응답 내 최근 출판물에 대한 개요를 제공하고 섹션 4.2의 요구 사항 분석을 통해 자동화 에이전트 및 맞춤형 ACO Gym을 기반으로 출판물을 분석합니다 .

5.1 자동화된 블루팀 솔루션

네트워크 시스템 내의 자동화된 블루 에이전트는 전체 공격 표면을 실시간으로 방어하기 위해 끊임없이 경계해야 하며, 공격자는 단일 위치 내에서 한 번만 성공하면 됩니다. 사이버 공격자와 방어자 사이의 이러한 비대칭 시나리오로 인해 제한된 자원을 가진 방어자는 가능한 모든 공격에 대비할 여력이 없습니다.

이 하위 섹션에서 초점을 맞춘 문제 영역은 자세 관련 취약점(PrV)의 완화입니다. 즉, 방어자는 전체 공격 표면을 실시간으로 방어하기 위해 끊임없이 경계해야 하는 반면, 공격자는 내부 단일 위치 내에서 한 번만 성공하면 됩니다. 네트워크로 연결된 시스템. 이러한 보안 태세의 단점으로 인해 제한된 자원을 가진 방어자는 가능한 모든 공격에 대비할 여력이 없습니다.

아래 표 4 는 문헌에 게시된 맞춤형 ACO Gym과 함께 자동화된 블루 팀 구성 솔루션을 평가합니다 .

표 4. 맞춤형 네트워크 시스템 내의 자동화된 Blue Team 솔루션

	자동화된 Blue Team 맞춤형 네트워크 시스템 간행물															
요건	[131]	[18]	[60]	[88]	[78]	[46]	[37]	[25]	[29]	[24]	[122]	[47]	[121]	[110]	[101]	
A.1.1	+				+	+	+	+	+	+	+	+	+		+	+
A.1.2					+		+	+	+	+		+	+		+	+
A.1.3									+		+	+				
A.2.1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
A.2.2'	+				+	+	+	+	+	+	+	+	+	+	+	+
A.2.3				+						+	+		+			+
A.2.4	+	+														
A.3.1																
A.3.2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
A.4.1								+		+						
A.6.1					+		+			+						
A.6.2	+	+														
A.6.3																
A.6.4												+	+			
G.1.1	+	+	+	+		+	+		+			+	+		+	+
G.1.2				+	+		+	+		+					+	+
G.1.3			+			+						+	+			
G.1.4	+	+	+	+		+	+		+				+		+	+
G.2.1			+											+	+	+
G.2.2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
G.4.1						+		+		+						
G.5.1														+		
G.5.2														+		
G.6.1				+	+			+		+						

표 4 는 단독으로 설계된 네트워크 시스템 내의 관련 ACD 자동화 블루 팀 발행물을 보여줍니다.
각각의 자동화된 블루팀 에이전트 구현을 위해. 표에는 대부분의 출판물 회의가 강조되어 있습니다.
요구 사항 A.1.1, A.1.2, A.2.1, A.2.2. 이는 특히 대부분의 출판물이 다음의 필요성을 강조하기 때문입니다.
방어가 불가능한 단일 샷 블루 에이전트 응답과 반대되는 순차적 블루 에이전트 응답 [24]
현대의 사이버 공격에 대비하는 시스템. 이는 문제를 다음과 같이 구성하는 모든 출판물에서 추가로 나타납니다.
MDP/POMDP(G.2.2)는 자동화된 에이전트가 전환을 통해 순차적인 응답을 취할 수 있도록 합니다.

상태는 네트워크로 연결된 시스템의 특정 노드 내에서 수행되는 작업의 조합을 나타냅니다. 그러나 A.1.2의 요구 사항은 특정 간행물 내에서 충족되지만 이는 시뮬레이션 기반 네트워크 시스템 구현이므로 시스템이 실제 네트워크 시스템의 구성 변경의 복잡성을 완전히 나타내지는 않습니다. 이는 시뮬레이션된 네트워크 시스템 내에서만 알고리즘을 테스트하는 표 4 의 대부분의 간행물에서 충족되지 않는 A.1.3 요구 사항에서 특히 강조됩니다 .

대부분의 출판물은 에이전트에 대한 장기 조치의 적절한 일반화를 위해 복잡한 네트워크 환경에서 요구되는 A.2.3을 충족하지 못했습니다 . DRL(Deep RL) 구현만이 이 요구 사항을 충족할 수 있어 더 적합해졌습니다. Dhir 등 [35] 은 또한 ACO 체육관 내에서 성과를 유지할 수 있는 인과 추론 알고리즘 [48, 64, 94, 100, 127] 의 사용을 제안했습니다 . 표 4 의 대부분의 간행물은 또한 A.2.4의 설명 가능성 요구 사항을 충족하지 않습니다. 이는 SOC 환경 내에서 자동화된 에이전트를 활용하는 데 가장 중요한 요소이며, 이러한 에이전트는 작동하기 전에 인증을 받아야 합니다. 선택된 출판물 중 단 2개만이 A.4.1을 충족했습니다. 두 출판물 모두 특정 사이버 공격(다양한 사이버 공격을 탐지하고 대응할 수 있는 에이전트와 달리 DDoS)에 대해 자동화된 대응을 구현했습니다.

이러한 요구 사항은 A.6.1 및 A.6.2의 형태로 강조되며, 이는 다양한 사이버 공격에 대한 적대적 훈련을 통해 자동화된 블루 에이전트의 지식 기반을 지속적으로 개발할 필요성을 시사합니다.
또한 A.4.1 요구 사항을 충족하는 구현이 부족하여 A.4.3에 언급된 알고리즘 공격에 대한 자동화된 블루 에이전트의 개발이 방해를 받습니다 . 이 영역은 표 4 에서 강조된 출판물에서 집중적으로 다루지 않은 영역입니다 .

자동화된 블루 팀 구성의 요구 사항 A.6.4는 위협 탐지 기능의 증가 를 통해 네트워크로 연결된 시스템의 방어를 강화하는 사기성 요소를 전략적으로 실행할 수 있는 능력을 가진 방어자 에이전트를 나타냅니다. 문헌에서 사이버기만을 적용하는 방법은 적을 오도하거나 속도를 늦추고 궁극적으로 인지 프로세스를 방해할 목적으로 충실도가 높은 사기성 자산을 기존 인프라에 통합하는 것입니다. 이러한 자산은 일반적으로 실제 자산과 유사한 가상 환경 내에 캡슐화됩니다. 두 가지 전반적인 목표를 가지고 있습니다. 첫째, 미끼 및 미끼와 같은 위협 탐지 기능 강화를 통한 시스템 방어 , 둘째, 사이버 위협 인텔리전스(CTI) 수집을 지원하기 위해 공격자를 잘못된 방향으로 유도하고 격리하는 능력입니다 . 디셉션 기반 사이버 방어(DCD) 플랫폼은 상대방이 모르는 사이에 네트워크의 실제 보안 태세를 난독화하는 예방적 사이버 보안 도구를 실행하고 유지함으로써 전형적인 공격자-방어자 비대칭성에 대응합니다. 실제로, 공격자가 성공적인 사이버 공격을 실행하기 위해 예상되는 취약점의 바다를 '지뢰 제거'해야 한다는 점을 고려하면 PrV 완화에 DCD 사용이 점점 더 신중한 선택이 되고 있습니다. Wang 등 [122] 과 Ghao 등 [47]은 모두 적의 행동을 분석하기 위해 지능형 알고리즘의 사용과 동적 배포 전략을 결합하는 개념을 고려합니다. 두 솔루션 모두 최적의 배포 전략을 선택하기 위해 블루 에이전트를 훈련하는 데 성공했지만 공격자와 관련 환경이 결합 되어 많은 일반화 및 복원력 기반 요구 사항이 부족합니다 . 앞서 언급했듯이 DRL을 통합 한 [47] 과 같은 솔루션은 일반적으로 높은 수준의 의사 결정 요구 사항 A.2.3을 충족합니다. 저자는 일반적인 공격자-방어자 시나리오가 일반적인 RL 알고리즘의 공간 복잡성 에 미치는 영향을 알고 있기 때문에 이 경우 DRL을 사용하는 것이 합리적입니다. 이는 상태 공간을 수동으로 엔지니어링할 필요 없이 정책 기반 배포 결정을 내리기 위해 DNN(심층 신경망)이 도입되었기 때문입니다. ACD의 맥락에서 가능한 모든 상태의 시행착오를 통해 보상 경로를 결정하는 것은 네트워크 환경의 규모가 커짐에 따라 종종 계산 난치성으로 수렴될 수 있습니다. 따라서 DNN의 예측 요소를 활용하면 모든 개별 상태를 저장하고 조회하는 대신 각 Q 값을 근사화하여 지식이 일반화됩니다. [47] 의 저자는 온라인 학습을 활용하여 새로 수집된 공격 정보로 방어 모델을 업데이트합니다. 이는 '비연속적' 다양성이지만, 이는 치명적인 간섭에 대한 우려를 해결하기 위해 지속적인 학습 기술이 구현되지 않았음을 의미하므로 충족하지 못했습니다. 요구사항 A.3.1. Li et al [71]은 DRL의 근사치를 활용하여 시스템을 생성하여 최적의 방어기만 프레임워크를 제안합니다.

적대적인 행동을 모델링하는 위험 그래프(SRG). 그런 다음 공격 모델을 사용하여 DRL 에이전트를 교육하여 마이크로 서비스 아키텍처 내에서 최적의 배포 전략을 생성합니다. 일반적인 OT 네트워크의 다양성과 규모, 기술의 가상화 및 컨테이너 서비스의 역동성은 이미 압도적인 문제에 추가 공격 벡터를 노출시키기 때문에 컨테이너 기반 클라우드 환경에 방어적 속임수를 통합하는 것이 합리적입니다 . 기만적인 자산을 지능적으로 배치함으로써 확장되는 위협 표면을 유지하고 예방할 수 있습니다. 저자는 네트워크 환경과 위협 모델을 고차원 입력 공간으로 모델링하고 최대 60개 노드까지 확장되는 DRL 프레임워크를 구현할 때 확장성 문제를 강조합니다 . 다른 관점에서 Walter 등 [121]은 CyberBattleSim[115]이라는 기존 오픈 소스 ACO Gym의 소스 코드를 적용하여 방어적인 사이버기만 구성 요소로 ACD 환경을 강화할 수 있다는 전망에 주목했습니다 . 솔루션이 반드시 전용 블루 에이전트를 생성하지 않기 때문에 이 문서는 많은 요구 사항에 미치지 못합니다 . 대신, 이 문서의 목적은 자동화된 블루 팀 구성 에이전트에 궁극적으로 정보를 제공할 수 있는 공격자 행동에 대한 적극적인 사이버기만의 영향을 관찰하여 통찰력을 얻는 것이었습니다 .

표 4에 언급된 간행물 내 네트워크 시스템의 요구 사항 측면에서 G.1.1 및 G.1.2는 대부분의 시뮬레이션된 네트워크 시스템 간행물에서 충족되었습니다. 그러나 이전에 언급한 것처럼 시뮬레이션 된 시스템은 실제 시스템을 정확하게 나타내지 않으므로 언급된 구현 중 매우 적은 수만이 G.1.3 요구 사항을 충족할 수 있는 이유입니다. 요구사항 A.4.1과 유사하게, G.4.1은 자동화 에이전트의 포함을 촉진하기 위해 네트워크 시스템을 개발해야 하는 영역 입니다 . 연구 개발 영역에는 G.4.1 및 G.6.1도 포함되며, 여기서 네트워크 시스템은 이러한 요구 사항을 허용하도록 설계되어야 합니다.

5.2 자동화된 레드팀 솔루션

자동화된 레드팀 구성 솔루션에 대한 기존 문헌은 공격 계획, 침투 테스트 또는 레드팀 "자동화"를 통한 보안 분석가 지원 , 체육관 환경에서 수행되는 레드 에이전트 연구 등 세 가지 범주로 나눌 수 있습니다. 후자의 범주는 ACO 목표/목표와 밀접하게 관련되어 있는 반면, 전자는 이를 향한 중간 단계입니다.

공격 경로 계획 항목은 Nmap [1] , Nessus [2] 등 침투 테스트 도구에서 출력되는 스캐닝 정보를 활용하여 기업 네트워크를 대표하는 POMDP(G.2.2)를 설계한다. 그런 다음 취약성 스캔의 CVSS(Common Vulnerability Scoring System) 점수 [3] 를 활용하여 전환 확률을 정의합니다. [45] 또한 CVSS 점수를 활용하여 보상을 알렸습니다(예를 들어 관리자로 호스트에 착륙).

그런 다음 연구원은 이러한 환경에서 RL 알고리즘(A.2.2)을 활용하여 설정된 목표에 도달합니다(루프를 피하기 위해 각 단계에 부정적인 페널티를 추가함). 예를 들어 [45] 와 [28] 에서는 이 접근 방식을 활용하여 인간 전문가가 DQN 알고리즘(A.2.3)을 사용하여 테스트 목표를 달성하도록 지원하는 실행 계획을 생성했습니다. 마지막으로, Bloodhound [4] 와 같은 도구는 ML을 활용하지 않고 Active Directory 약점에 초점을 맞춘 공격 경로 계획을 제공한다는 점에 유의해야 합니다 .

침투 테스트를 자동화하려면 위 단락에 정의된 RL 게임을 확장하여 침투 테스트 또는 레드 팀 구성 도구(A.6.2)의 작업을 통합할 수 있습니다. 실제로 [129]는 Metasploit 프레임워크 [5]를 사용하여 침투 테스트를 자동화하기 위해 그렇게 한 반면, [76]은 PowerShell Empire 프레임워크 [6]를 활용하여 악용 후 활동을 자동화했습니다 . 또한 연구자들은 레드팀 구성의 특정 작업을 분석하고 이를 자동화하려고 시도했습니다. 예를 들어 [70] RL을 통한 자동 권한 에스컬레이션이 있습니다. [39] 에서 볼 수 있는 방어 회피와 같이 MITRE ATT&CK 매트릭스 [7] 의 여러 셀이 이러한 방식으로 자동화되는 것을 상상할 수 있습니다 .

자동화된 레드팀 솔루션을 위한 RL에 대한 연구가 시뮬레이션된 환경으로 추상화될 수 있다는 점을 고려하면

(ACO Gyms, G.1.3에 자세히 설명되어 있음), 문헌도 그러한 연구로 구성되어 있습니다. 예를 들어 [113] Network Attack Simulator Gym [106]에서 Deep RL 에이전트를 구축합니다. 저자는 PPO 및 DQN 알고리즘을 사용하여 구축된 다양한 크기와 복잡성의 5가지 시나리오에서 에이전트를 교육했습니다.
그들은 PPO가 약간 더 일반화되는 것처럼 보이는 테스트 시간에 더 큰 시나리오에서 어떻게 수행되는지 확인하기 위해 더 작은 시나리오에서 그들을 훈련시켰습니다. 액션 세트의 기하급수적인 증가를 고려하여 연구자들은 이 설정에서 계층적 RL의 사용을 분석하기 시작했으며 실제로 [117] 계층적 DQN 알고리즘을 제안한 CyBORG Gym 환경 [112] 에서 분석했습니다. 오픈 소스 체육관에서의 연구는 8 에 요약되어 있습니다.

마지막으로 GT 모델(A.2.2)도 연구되었지만(예는 [30]에서 제공됨) 이 경우 사이버 전쟁 게임과 같은 의사 결정자를 지원하는 데 활용된다는 점에 유의해야 합니다.

표 5. 맞춤형 네트워크 시스템 내의 자동화된 레드팀 솔루션

	자동화된 레드팀	맞춤형 네트워크 시스템	간행물
요구사항 [76] [70] [45]			
A.1.1			
A.1.2			
A.1.3		+	
A.2.1	+	+	+
A.2.2	+	+	+
A.2.3	+	+	+
A.2.4			
A.3.1			
A.3.2			
A.4.1			
A.6.1			
A.6.2			
A.6.3			
A.6.4			
G.1.1	+		+
G.1.2			
G.1.3	+	+	+
G.1.4	+		+
G.2.1			
G.2.2	+	+	+
G.4.1			
G.5.1			
G.5.2			
G.6.1			

6 자율 사이버 작전 체육관

이전 섹션에서 설명한 것처럼 공통 오픈 소스 ACO Gym이 부족하여 자동화된 블루 및 레드 에이전트(및 ACO Gym)의 별도 가속화 개발 가능성이 없습니다 . 이 섹션에서는 문헌 및 웹사이트 내에서 개발 및 출판된 자동화 에이전트와 함께 최근 ACO 체육관을 개발한 문헌에 대한 자세한 개요를 제공합니다. 이러한 ACO 체육관은 자동화된 블루 및 레드 팀 솔루션 개발을 위해 특별히 설계된 시뮬레이션 및/또는 에뮬레이트된 네트워크 시스템입니다 . 여러 리소스 의 가용성을 고려하여 다양한 출판물에서 교육 및 테스트 환경, 알고리즘 개발 유형 및 가능한 사이버 공격 유형에 대한 다양한 전략을 제시했습니다 .

6.1 훈련 전략 에이전트를 훈련

하고 테스트하는 가장 일반적인 접근 방식 은 훈련된 환경과 동일한 환경에서 에이전트의 정책을 검증하는 것입니다 . 이는 시뮬레이션된 환경과 에뮬레이트된 환경 모두에 적용됩니다. 불행하게도 이 전략은 자동화된 에이전트의 일반화 능력(즉, 요구사항 A.1.1, A.1.2 및 표 3의 충족) 에 대한 진술을 방해합니다 . 또한 자동화된 에이전트는 다양한 유형 의 환경을 사용함으로써 얻을 수 있는 이점(예: 현실성에 더 가까워지기 위한 확장성 및 에뮬레이션을 위한 시뮬레이션)을 완전히 활용하고 요구 사항을 충족할 수 없습니다.
G.1.3.

여러 연구 논문이 일반화 영역에서 진전을 이루기 위해 노력해 왔습니다. 예를 들어, [113]은 Network Attack Simulator Gym [106]에 Deep RL 에이전트를 구축했습니다. 자동화된 침투 테스트 연구를 수행하기 위한 시뮬레이션 환경입니다 . 자동화된 에이전트는 다양한 크기와 복잡성을 지닌 5가지 시나리오(서브넷, 호스트, 취약점 포함)에서 훈련되었으며, 여기서 작성자는 PPO와 DQN 알고리즘을 모두 채택했습니다. 복잡성이 낮은 시나리오에서 자동화된 에이전트를 교육한 후 PPO가 우수한 일반화를 제공하는 더 큰 복잡성 시나리오의 성능에 미치는 영향을 실험했습니다. [83], [112] 또는 [72] 에 의해 설계된 ACD 연구를 수행하기 위해 구축된 최첨단 플랫폼에는 모두 시간 효율적인 방식으로 에이전트를 교육하기 위한 시뮬레이션 환경이 포함됩니다 . 또한 실행 중인 서비스, 악의적인 작업을 수행하는 실제 악성 코드, 포트를 닫거나 감염을 제거하는 기능(시뮬레이션의 작업 공간에 매핑)을 갖춘 자동화된 블루 에이전트를 갖춘 클라우드 제공업체에서 환경 에뮬레이션을 실행할 수 있습니다. 또 다른 접근 방식은 시뮬레이션된 환경에서 훈련을 수행한 후 "실제" 테스트를 포함하는 것입니다. 언급할 가치가 있는 한 가지 예 는 작업별 에이전트입니다. 예를 들어 [70]는 MITRE ATT&CK 매트릭스 [7] 에서 가능한 모든 권한 상승 기술을 열거 하고 이 작업을 수행하기 위해 DQN을 사용하여 에이전트를 구축했습니다. 학습 프로세스 속도를 높이기 위해 Python으로 구축된 시뮬레이션 환경에서 에이전트를 교육한 다음 "실제 세계"(Windows 가상 머신)에서 테스트를 수행했습니다. 그들은 권한을 상승시키는 데 필요한 단계 수를 기준으로 성능을 측정했으며 일부 사례/취약성의 경우 자동화된 에이전트가 인간 전문가보다 성능이 뛰어났습니다.

6.2 기존 자율 사이버 운영 체육관 네트워크 시스템 내의 자동

화된 레드 및 블루 팀 구성 에이전트 영역 내에서 연구를 가속화하려면 오픈 소스 네트워크 시스템 또는 자율 사이버 운영 체육관(ACO 체육관)이 필요합니다.
ACO Gym을 제공하면 연구자는 표 3의 자동화 에이전트 기반 요구 사항을 충족하는 데 집중할 수 있습니다. 또한 이를 통해 연구자는 표 3 의 네트워크 시스템 요구 사항을 충족하는 더 많은 오픈 소스 ACO Gym을 개발하는 데 집중할 수 있습니다. 다음은 사이버 보안 연구를 위해 설계된 기존 환경에 대한 검토입니다 . 검토는 시뮬레이션인 기존 환경에 대한 개요를 제공하는 것으로 시작한 다음 게시된 다른 비공개 소스 에뮬레이션(및 기타 시뮬레이션) 환경을 자세히 살펴봅니다. 각 부분에서는 ACO Gym 에 대한 요구 사항 분석을 사용하여 ACO Gym 을 다른 오픈 소스/비공개 소스 ACO Gym과 비교합니다 .

6.2.1 오픈 소스 체육관. 첫째, Cyber Battle Sim [115] (CBS) 환경 은 구성된 취약점이 있는 고정 네트워크를 시뮬레이션하는 환경에서 사이버 공격의 측면 이동 단계에 초점을 맞춘 자동화된 레드 에이전트를 훈련하기 위해 만들어졌습니다 . 레드 에이전트는 측면 이동을 위해 익스플로잇(네트워크에 원격으로 액세스하여 높은 권한을 얻거나 네트워크 더 깊이 이동하는 특정 코드)을 활용하는 반면, 사전 정의된 블루 에이전트는 레드 에이전트를 탐지하고 액세스를 방해하는 것을 목표로 합니다. CBS 환경은 관련 노드와 함께 네트워크 레이아웃 과 취약점 목록을 정의할 수 있습니다. CBS에서 모델링된 사이버 자산은 최신 운영 체제와 최신 패치가 향상된 보호 기능을 제공할 수 있는 방법을 설명하는 데 중점을 두고 OS 버전을 캡처합니다.

"블루 에이전트" 교육을 위한 설계로 인해 구현을 확장할 수도 있습니다. 실제로 [121]는 블루 팀 속임수를 환경에 통합하기 위해 그렇게 했습니다 . 개발자는 취약점에 대한 MITRE ATT&CK 매트릭스 [7] 의 셀을 추상화하기 위해 환경 에 충분한 복잡성이 존재하는지 확인했습니다 (레드 에이전트가 보상을 받기 위해 악용). 전반적으로 문서는 새로운 시나리오/네트워크를 생성하고 보상 기능(손상된 서비스의 가치 및 악용 비용)을 조정하고 서비스에 취약점을 추가하는 데 충분합니다.

이를 통해 사용자는 환경을 광범위하게 실험할 수 있지만 코드는 시뮬레이션된 도메인 내 구현을 위해서만 존재하므로 환경의 현실성에 의문이 제기됩니다.

Gym IDS 게임 [54] 은 OpenAI 체육관 환경을 기반으로 구축된 단순한 Markov 게임입니다. 공격자는 두 가지 유형의 조치를 취할 수 있습니다.

- 정찰 활동
- 또는 유형 1...m의 공격

방어자는 또한 두 가지 유형의 행동을 취할 수 있습니다.

- 모니터링 조치 • 또는 유형 1...m
  의 방어 조치

파란색 에이전트나 빨간색 에이전트(또는 둘 다)를 훈련하는 데에는 다양한 시나리오가 있습니다. 불행히도 체육관 환경은 지나치게 단순하고 시뮬레이션된 환경만 제공하므로 CBS와 마찬가지로 현실감도 낮습니다.

위에서 설명한 Gym IDS 게임과 유사하게 Gym Threat Defense 체육관 [79] 도 POMDP 설정을 갖춘 시뮬레이션 기반 시스템입니다 . 그러나 이 경우 저자는 방어자가 네 가지 다른 행동을 할 수 있는 순전히 방어적인 게임으로 설계했습니다 .

- 조치 없음
- 서비스 차단 • 머신 연결 끊
  기 • 작업 2와 3을 동시에 수행

각 노드에 대한 탐지 확률, 공격 확률, 확산 확률 및 초기 상태를 정의할 수 있습니다.

네트워크 공격 시뮬레이터 환경 [106]은 침투 테스트 작업에서 AI 시스템을 테스트하기 위해 레드 에이전트(블루 에이전트가 없기 때문에)를 교육하기 위해 순수하게 구축되었습니다 . 이 환경은 OpenAI 체육관을 기반으로 구축되었으며 호스트 수, 서비스, 관찰 모드(예: 완전히 관찰됨) 및 문제의 호스트의 자산 중요도를 정의하여 시나리오를 생성할 수 있는 기능을 허용합니다. 마지막으로 네트워크에 존재하는 취약성을 결정 하고 조치 비용(예: 서브넷 스캔 비용)을 정의할 수 있습니다. 레드 에이전트는 악용, 권한 상승, 서비스 검색, 운영 체제 검색, 서브넷 검색, 프로세스 검색, 작업 없음 등 7가지 작업 유형 중에서 선택할 수 있습니다. 프로젝트의 목표는 블루 에이전트가 환경을 방해하지 않는 동안 시뮬레이션 시나리오에 대해 침투 테스트를 수행하도록 레드 에이전트를 교육하는 것입니다. 이 구현은 AI 기반 레드 에이전트 훈련 내에서 연구 진행을 강화할 수 있는 기능을 제공하지만 CBS, Gym IDS Game 및 Gym Threat Defense와 같은 시뮬레이션만 제공하여 구현의 현실성을 감소시킵니다.

언급된 환경과 유사하게 Optimal Intrusion Response Gym [55] 은 OpenAI Gym 라이브러리를 기반으로 구축된 Markov 게임입니다 . 환경은 각각 IDS로 구성된 여러 호스트가 있는 6개의 서브넷이 있는 시뮬레이션된 엔터프라이즈 네트워크로 구성됩니다. 불행하게도 방어자가 두 가지 행동 중에서만 선택할 수 있기 때문에 이 게임은 우리 사용 사례에 비해 지나치게 단순합니다 .

- "중지"는 게이트웨이를 차단합니다. 이로 인해 IT 서비스가 저하되고 그에 따른 비용도 발생합니다. 하지만, 또한 감염이 억제되도록 보장합니다.
- "계속"은 아무런 조치도 취하지 않습니다.

몇 가지 시뮬레이션/테스트를 수행한 후 [55]는 훈련된 파란색 에이전트가 시끄러운 공격자보다 은밀한 공격자를 상대할 때 더 일찍 "중지"할 가능성이 더 높다는 것을 발견했습니다.

CyBORG 환경 [112]은 블루 에이전트 교육을 위해 특별히 설계되었습니다. 그러나 CBS와 유사하게 레드팀 사용 사례를 위해 간단히 확장할 수 있습니다. 이 환경에서는 시뮬레이션된 환경과 에뮬레이트된 환경에서 각각 교육 및 테스트가 가능합니다. 시뮬레이션된 환경은 각 상태가 시스템과 네트워크를 나타내는 유한 상태 머신(FSM)에서 모델링된 시나리오와 상호 작용하는 에이전트로 구성됩니다.

한 상태에서 다른 상태로 이동하려면 해당 사전 조건을 만족하는 동작이 필요합니다. 또한 상태는 개별 파일의 생성 및 삭제, 네트워크 연결 설정 또는 중단과 같은 특정 세부 정보도 제공합니다. 이 모든 것이 결합되어 방어자와 적 에이전트 모두에게 이상적인 훈련 환경이 생성됩니다. 자동화 에이전트가 훈련되면 에뮬레이터에서 테스트가 완료됩니다. 에뮬레이터는 자동화 에이전트가 상호 작용하는 충실도가 높은 사이버 보안 환경을 생성하기 위한 AWS 가상 머신으로 구성됩니다 . 환경의 목적은 ACD 연구를 위한 플랫폼 역할을 하여 대중에게 도전 과제를 공개하는 것입니다. 즉, TTCP 케이지 챌린지 1, 2 및 3입니다. 챌린지는 복잡성이 증가하는 엔터프라이즈 네트워크 환경입니다(빨간색 및 파란색 에이전트의 관찰 및 작업 공간 측면에서).

TTCP CAGE 챌린지 2(가장 최근 챌린지)에서 파란색 에이전트의 작업 세트는 철저합니다.

- 제거 - 호스트에서 악성 코드를 제거합니다. • 복원 - 악성 코드에 높은 권한이 있는 경우 제거할 수 없으며 호스트를 다음에서 복원해야 합니다. 백업(관련 비용 포함).
- 분석 - 모니터링이 항상 감염을 탐지하는 것은 아니지만(5/100회) 호스트에 대한 분석을 수행합니다. 항상 그것을 감지합니다.
- 미끼 서비스 - 특정 호스트에 미끼 서비스를 설정하여 레드 에이전트 활동을 지연 및 감지합니다(7개 있음). 다양한 서비스 이용 가능) • 작업 없음
- 다른 작업과 관계없이 모니터링이 발생합니다.

시나리오는 YAML 파일에서 정의할 수 있습니다(예: 네트워크 토폴로지 및 자산 중요도). 또한 이 프로젝트에는 다양한 전략을 활용하는 다양한 레드 에이전트가 함께 제공됩니다. 마지막으로 문서는 철저합니다. 이 환경은 실험에 대한 모든 요구 사항에 맞는 것으로 보이며 자율적 블루 에이전트의 높은 수준의 원하는 작업을 자세히 설명합니다 . 이 시뮬레이션 환경 외에도 CybORG는 에뮬레이션(비공개 소스)으로 확장되며, 이는 훈련된 에이전트를 검증하기 위해 AWS에서 실행될 수 있습니다.

YAWNING TITAN [31]은 블루 에이전트 교육을 위한 고도로 추상화된 그래프 기반 체육관입니다. 파란색 에이전트와 빨간색 에이전트 모두의 행동 공간은 사이버 방어에 대해 예상되는 현실적인 공간과 매핑되지 않습니다. 대신 접근 방식/알고리즘을 효율적으로 테스트하고 검증하기 위해 체육관이 만들어진 것으로 보입니다. 또한 그래프 기반 설계는 네트워크가 YAML 파일이 동작과 공간을 결정하는 함수로 정의될 수 있으므로 일반화 A.1.2와 관련된 계산적으로 비용이 많이 드는 접근 방식을 탐색하는 것이 진정한 목적임을 시사합니다 . 표 6 은 실험할 수 있는 모든 오픈 소스 ACO 체육관을 요약하는 데 사용되었습니다 .

표 6. ACO 체육관(오픈 소스)

	자동화된 사이버 작전 체육관(오픈 소스)							
요구사항	CBS [115]	GIG	GTD	OIR	CybORG [112]	나심	YT [11]	
G.1.1								+
G.1.2					+	+		
G.1.3						+		
G.1.4								+
G.2.1	+					+	+	
G.2.2	+		+		+	+	+	+
G.4.1						+		
G.5.1	+		+	+	+	+	+	+
G.5.2	+		+		+	+	+	+
G.6.1								

6.2.2 비공개 소스 체육관. 나머지 ACO 체육관은 요구 사항을 통해 표 7 에서 분석되었습니다.
표 3에 표시된 분석. 강조 표시된 ACO 체육관은 오픈 소스는 아니지만 중요한 정보를 제공할 수 있습니다.
특히 디자인이나 디자인 시 영감을 얻을 수 있는 연구자를 위한 ACD 커뮤니티 내 통찰력
기존 ACO 체육관을 수정합니다.

표 7. ACO 체육관(비공개 소스)

	자동화된 사이버 작전 체육관(비공개 소스)								
요건	[44]	[81]	[43]	[20]	[102]	[103]	[72]	[83]	[38]
G.1.1	+		+	+	+			+	
G.1.2			+		+	+	+	+	+
G.1.3		+		+	+	+	+	+	+
G.1.4	+			+	+	+			+
G.2.1	+	+		+		+	+	+	
G.2.2						+	+	+	
G.4.1									
G.5.1									
G.5.2									
G.6.1					+	+		+	

6.3 모든 ACO 체육관에 대한 종합 분석

표 6에서 볼 수 있듯이 대부분의 저자는 원활한 추가 및 제거 요구 사항을 인식했습니다.
노드 및 구성요소(G.1.1). 저자는 또한 다음과 같은 자동화 에이전트 추가(G.1.2) 요구 사항을 충족합니다.
ACO 체육관 내의 구조적 변화를 이해하고 자신의 결정을 일반화할 수 있습니다(A.1.1
및 각각 A.1.2). 또한 모든 출판물은 AI 기반 순차의 요구 사항도 이해했습니다.
의사 결정을 자동화하는 빨간색 및 파란색 에이전트(각각 A.2.1 및 A.2.2)가 있으며 ACO를 구성했습니다.
이러한 에이전트를 활성화하기 위해 MDP로 Gym을 운영합니다. 그러나 이러한 ACO 체육관은 확장성이 뛰어나지만(G.1.4)
관련 자동화 에이전트의 개발을 허용하면 모든 구현에 활용되는 환경은 다음과 같습니다.
실제 네트워크 시스템에 대한 시뮬레이션으로 오픈 소스 에뮬레이션/실제 ACO 체육관의 부족을 강조
(G.1.3). 이로 인해 활용에 필수적인 자동화 에이전트의 "실제" 경험이 부족하게 됩니다.
현재 네트워크 시스템 내에서.

나머지 분석은 자동화 에이전트에 적용되지만 ACO 체육관의 현재 상태 설계는 다음과 같습니다.
ACO 체육관 내에서 설계할 수 있는 자동화 에이전트의 품질을 평가하는 데 사용할 수 있습니다. 전반적인,
단 하나의 ACO Gym(CybORG [112] Cage Challenge 3 [53])만이 자동화된 다중 에이전트의 필요성을 인식했습니다.
알고리즘(A.4.1)을 자동화된 블루팀 솔루션으로 사용합니다. [78] 및 [37] 출판물(특히 RL 사용에 중점을 두고 있음)
DDoS 공격 방어를 위한) 환경은 ACO 체육관을 구성하는 데 잠재적인 영감을 줄 수 있습니다.
다중 에이전트의 자동화된 레드 및 블루 팀 협업을 촉진합니다(요구 사항 G.4.1). ACO 체육관이 거의 없음
자동화된 블루를 강화하는 데 잠재적으로 활용될 수 있는 적대적 훈련(G.6.1 및 A.6.1)을 촉진합니다.
다양한 사이버 공격에 대한 에이전트(A.6.2). 현재 사용 가능한 오픈 소스 ACO 체육관은 인식되지 않았습니다.
자동화된 레드 에이전트의 행동 공간 내에 알고리즘 사이버 공격(A.6.3)을 통합할 필요성
자동화된 블루 에이전트에 대항합니다. 폐쇄 소스 ACO Gym [83] 에서 영감을 얻어 통합 할 수 있습니다.
DRL 알고리즘과 같은 자동화된 에이전트의 회피 및 포징과 같은 알고리즘 공격.

오픈 소스 ACO 체육관 내의 7가지 ACD 알고리즘

이전 섹션에서 언급한 오픈 소스 ACO Gym 중에서 여러 자동화된 의사 결정 알고리즘이 자동화 에이전트로 교육 및 테스트에 활용되었습니다. ACO Gym 제작자와 자동화된 블루 및 레드 팀 에이전트 개발자는 순차적 응답 특성으로 인해 도메인 내 DRL 기반 솔루션의 필요성을 인식했습니다 . DRL 기반 솔루션을 사용하여 많은 요구 사항이 충족되지만 이 섹션에서는 현재 게시된 구현을 통해 자동화 에이전트 설계 내에 여전히 존재하는 몇 가지 격차를 제시합니다 . 사이버 보안을 위해 알고리즘을 실제 네트워크 시스템에 배포하려면 먼저 이러한 격차를 해소해야 합니다 . 현재 ACO Gym 중 단 2개의 오픈 소스 ACO Gym만이 자동화된 빨간색 및 파란색 에이전트 게시에 활용되었습니다. 또한 도메인 내 연구 개발을 촉진하기 위해 많은 알고리즘이 개발되어 오픈 소스로 출시되었습니다.

CybORG [112] 는 작업 및 관찰 공간 측면에서 ACO Gym 복잡성이 다양한 시뮬레이션된 네트워크 시스템을 사용하여 세 가지 과제를 발표했습니다 . 과제는 자동화된 블루 에이전트의 개발에 초점을 맞추는 반면, 자동화된 레드 에이전트(두 가지 유형의 사이버 공격으로 구성됨)의 개발도 가능합니다. NaSim [106] 작성자는 자동화된 빨간색 에이전트 개발을 위해 코드를 오픈 소스로 만들었고 몇몇 간행물과 구현에서는 이러한 에이전트 개발을 위해 시뮬레이션된 네트워크를 활용했습니다.

7.1 자동화된 블루팀 솔루션

위에 언급된 두 ACO 체육관 중에서 CybORG는 공개된 챌린지 [8] 에 대한 결과를 발표했으며 Cage Challenge 1 [52] 및 Cage Challenge 2 [51] 에 사용된 RL 기반 알고리즘을 나열하고 순위를 매겼습니다 . Cage Challenge 3 결과는 작성자가 설정한 성능 지표를 통해 곧 공개될 예정입니다 [53] . 다양한 팀에서 취한 여러 가지 접근 방식과 자동화된 에이전트가 구현한 여러 고유한 전략이 있습니다 . 이 문서에서는 과제 전반에 걸쳐 가장 성과가 좋은 접근 방식을 선택했으며 표 3의 요구 사항 분석과 비교 했습니다 .

Cage Challenge 1에서 Team Mindrake [40]가 도전에서 승리하여 호기심을 가지고 Proximal Policy Optimization [104]을 포함하는 Hierarchical RL 알고리즘을 생성했습니다 . 알고리즘의 계층적 [56] 구성 요소 는 자동화된 에이전트(B_line 및 Meander APT 에이전트) 에 대해 배포된 공격자의 유형에 따라 관련 조치를 취하기 위해 컨트롤러를 통해 활용됩니다 . 모델은 훈련 단계와 별도로 두 적에 대해 사전 훈련된 다음 무작위 에피소드에서 동일한 적에 의해 테스트됩니다. 호기심 구성 요소는 내재적 보상을 통해 훈련 단계에서 환경 내 탐색을 허용하여 [92] 보상을 거의 두 배로 향상시킵니다. 자동 에이전트가 챌린지 내에서 승리했지만 A.1.3, A.2.4, A.3.1, A.4.1, A.6.3 및 A.6.4 요구 사항을 충족하지 않습니다 . 이는 주로 두 적 사이에서 수행할 수 있는 조치의 가용성 과 적들이 수행할 수 있는 공격의 다양성 때문입니다. 게다가 환경 [52] 은 A.4.1을 촉진할 수 없습니다. 마찬가지로, 다른 세 가지 제출물도 챌린지 우승자와 동일한 요구 사항을 충족했습니다.

Cage Challenge 2에서 Cardiff University의 팀(GitHub 코드를 사용하여 Cage Challenge [4]) 도전에서 승리했고 또한 1의 Team Mindrake와 유사한 계층적 PPO를 생성했습니다. 그러나 팀은 미끼 선택을 통해 두 번째 도전 내에서 속임수의 가용성을 활용했습니다. 요구사항 분석을 이용하여 자동화된 에이전트는 요구사항 A.1.3, A.2.4, A.3.1, A.4.1 및 A.6.3을 충족하지 못했으나 속임수를 사용하는 요구사항은 충족했습니다. Cage Challenge 2 내에서 가용성을 보장합니다 .

4https://github.com/john-cardiff/-cyborg-cage-2

전반적으로 두 가지 과제에서 볼 수 있듯이 계층적 PPO 에이전트의 변형은 다른 접근 방식과 비교하여 가장 최적의 성능을 보여주었습니다([124]에서도 제안되고 알고리즘으로 입증됨) . 자동화 에이전트는 두 적의 움직임을 일반화할 수 있지만 , 그들이 훈련받은 환경은 자동화 에이전트가 일반화할 수 있는 다양한 유형의 사이버 및 알고리즘 공격(A.6.2, A.6.3)으로 구성되지 않았습니다. 더 큰 알고리즘 공격 풀. 이 ACO Gym 내에서 이러한 요구 사항을 충족하기 위해 향후 구현에서는 ACO Gym을 수정하여 사이버 및 알고리즘 공격 기능을 향상시켜 더 큰 공격 풀에 대해 자동화된 에이전트의 일반화 품질을 평가할 수 있습니다. 대조적 으로 , 두 과제 모두에서 자동화된 에이전트 구현은 수행할 수신 작업에 대해 어떠한 형태의 설명 가능성(A.2.4)도 제공하지 않았습니다 .

7.2 자동화된 레드팀 솔루션

안타깝게도 Automated Blue Team 솔루션과 달리 공개 과제는 제안되지 않았습니다. 결과적으로 다양한 체육관과 다양한 구성에서 연구가 수행되었습니다. 따라서 공개 비교 벤치마크가 부족합니다.

자동화된 레드 팀 구성 솔루션은 지금까지 CyBORG [112], Network Attack Simulator [106] 및 CyberBattleSim [115] 과 같은 ACD 체육관 환경이나 IT 네트워크의 에뮬레이터 또는 사용자 정의 표현에서 강화 학습을 통해 주로 수행되었습니다. 문제가 강화 학습 게임(POMDP 탐색)에 대해 완벽하게 모델링되었으므로 이는 직관적으로 의미가 있습니다. Automated Blue Teaming 솔루션과 마찬가지로 Proximal Policy Optimization 알고리즘이 가장 성공적인 접근 방식인 것으로 나타났습니다.

주목할 만한 한 가지 예는 시뮬레이션된 빨간색 에이전트를 에뮬레이션 으로 전송하는 유일한 알려진 예를 제시하는 CyBORG 체육관에서 수행된 연구입니다 . 연구원은 CyBORG 시뮬레이터에서 DQN 에이전트를 구현했습니다 . 그런 다음 CyBORG 에뮬레이터(G.1.3)에서 자동화 에이전트를 검증했습니다.
대부분의 자동화 에이전트가 에뮬레이터로 성공적으로 전송되었습니다. 시뮬레이터의 관찰에 대한 과도한 맞춤으로 인해 실패할 가능성이 없는 것 (이산적 관찰에서 연속 시간 제한 관찰로 이동).
Nasim 체육관의 또 다른 예는 [113] 에 의해 수행된 확장 일반화(G.1.1)의 첫 번째 예를 제시합니다. 그들은 소규모 시나리오에서 교육을 받은 Deep RL 에이전트를 구현하고 테스트 시 더 큰 시나리오에서 검증했습니다 . 그들의 연구에 따르면 Proximal Policy Optimization 알고리즘은 다른 알고리즘보다 일반화가 약간 더 나은 것으로 나타났습니다.

그러나 이러한 알고리즘이 가장 적절한지는 아직 의문의 여지가 남아 있습니다. 실제로 최근 블루 팀 구성 측면에서 유망한 것으로 나타났음 에도 불구하고 자동화된 레드 팀 솔루션의 캐주얼한 접근 방식에 대한 연구가 부족한 것으로 보입니다 . 12].

표 8. 오픈 소스 체육관 내의 자동화된 레드팀 솔루션

	자동화된 레드팀		
논문 [117]	[112]	[86]	[113]
A.1.1			
A.1.2			
A.1.3	+		
A.2.1 +	+	+	+
A.2.2 +	+	+	+
A.2.3 +	+	+	+
A.2.4			
A.3.1			
A.3.2			
A.4.1			
A.6.1			
A.6.2			
A.6.3			
A.6.4			
체육관 CyBORG CyBORG Nasim Nasim			

8 토론

본 논문의 주요 목적은 사이버 보안 내에서 임박한 연구 분야인 ACD를 식별하는 것이었습니다.
미래의 사이버 공격을 완화하기 위해. 사이버 공격에 대한 자동화된 대응은 다음을 통해 해결되어야 합니다.
본질적으로 순차적인 자동화된 레드 및 블루 팀 구성 에이전트의 연구 개발
그들의 의사 결정. 이러한 알고리즘의 개발은 병행 연구를 통해 가속화될 수 있습니다.
ACO 체육관 지역 내 개발. 최근의 발전으로 연구 분야가 발전하면서
특정 방향에 대해서는 요구사항 분석(표 3) 을 통해 더 많은 과제가 확인되었습니다.
언급된 영역 내에서 향후 개발을 위한 논문입니다. 40개 이상의 출판물을 분석하고 비교했습니다.
Table 3의 요구사항 분석을 통해 ACO Gym과 자동화된 Red, Blue를 개발하면서
에이전트는 별도의 연구 및 개발 전략으로 구성되어 있으며, 한 영역의 진행은 크게 의존적입니다.
다른 한편으로는 공통된 연구 과제가 있다는 추론을 정당화합니다. 더 많은 어려움이 있을 수 있으므로
언급된 특정 요구 사항 내에서 연구자가 이 문서를 기반으로 추가 작업을 수행하는 것이 좋습니다.
산업 용도로의 발전을 더욱 촉진할 수 있는 ACD 내 영역을 다루고 개발합니다.

8.1 과제 및 중요성 표 3 의 요구 사항

분석을 ACD로 식별된 출판물에 직접 매핑하면 ACD 시스템을 실제 시스템에 구현하기 전에 추가로 개발하기 위해 채워야 할 분명한 과제가 있음을 알 수 있습니다 . 이 섹션 에서는 확인된 추가 연구 및 개발 영역의 개요를 설명 하고 해당 영역을 요구 사항 분석 내의 특정 요구 사항과 다시 연결합니다.

중요도에 따라 각 챌린지에 대한 요구 사항이 추가되었습니다.

8.1.1 자동화된 블루 에이전트의 AI 기반 공격 견고화(A.6.3, G.6.1, A.6.1). 여기에는 자동화 에이전트의 알고리즘 기능을 공격하는 것을 목표로 하는 중독 및 회피 공격에 대한 Deep RL 알고리즘의 견고성이 포함되는 영역에 중점을 둡니다 . 매우 적은 수의 출판물이 Deep RL 알고리즘에 대한 공격에 초점을 맞추고 있지만 미래의 사이버 공격자는 Deep RL 및 다른 영역 내의 신경망 기반 연구를 통해 미래에 이러한 공격을 구현할 것임이 분명합니다 [14, 26 , 108 , 130]. 이 문제가 해결되지 않으면 미래의 네트워크 시스템은 잠재적으로 자동화된 블루 에이전트를 제어하고 결국 전체 네트워크를 제어할 수 있는 알고리즘 공격에 취약할 수 있습니다 .

8.1.2 Automated Red Agents(A.3.1, G.6.1, A.6.1, A.6.2, A.6.3)에 대한 행동 공간의 지속적인 진화. 레드 에이전트의 행동 공간은 끊임없이 진화하고 있습니다. 실제로, 취약점이 결합된 새로운 서비스가 추가되는 경우가 많습니다 . 또한 "매년 소프트웨어에 대한 새로운 악용 사례가 발견되므로 자동화된 침투 테스트 에이전트가 유용하려면 점점 늘어나는 악용 데이터베이스를 처리할 수 있어야 합니다." 따라서 레드 에이전트와 이들이 훈련받은 체육관은 이 문제를 고려해야 합니다(G.1.1). 이 과제는 과제 7.1.1 및 7.1.5에 의존하지만 지속적인 학습 환경을 기반으로 하는 사이버 공격 자동화 레드 에이전트의 개발 및 추가는 아직 연구되지 않았습니다. 이러한 문제를 구현하지 못하면 자동화된 블루 에이전트 에이전트가 최신 사이버 및 알고리즘 공격으로부터 구식 상태로 유지됩니다.

8.1.3 설명 가능한 RL(A.2.4) 설명 가능한 RL은 XAI보다 더 복잡합니다. 실제로 "RL 에이전트에 대한 설명 가능성은 분명히 XAI의 하위 집합이고 IML(Interpretable ML)과 유사하지만 현재 XAI 및 IML 연구와 명시적으로 분리되어야 하는 뚜렷한 특성을 가지고 있습니다." [32]. 실제로 XRL의 첫 번째 어려움은 취해야 할 결정/조치를 결정하는 장기적인 관점 때문입니다. 두 번째는 레이블이 지정된 훈련 데이터를 기반으로 구축되지 않은 모델과 관련이 있습니다(설명 가능성을 단순화함). 관련 조사 문서 및 구현에서 추가 영감을 얻을 수 있습니다 [10, 49, 74, 75, 80 , 82 , 91, 95, 96, 109, 116]. 이 문제를 해결 하지 못하면 에이전트에 대한 신뢰도가 낮기 때문에 자동화된 블루 에이전트가 네트워크 시스템 내의 산업 직원에 의해 인증되지 않게 됩니다 .

8.1.4 다중 에이전트 RL(G.4.1). ACD를 위한 자동화된 블루 팀 구성의 또 다른 연구 영역 은 구현을 위해 단일 RL 알고리즘을 사용하는 것과 반대되는 다중 에이전트 RL 알고리즘을 활용하는 것입니다 . 이는 매우 복잡한 기업 네트워크 환경에서 특히 더 유용할 것입니다. [112] 저자는 세 번째 Cage Challenge에서 다중 에이전트 RL의 구현을 제안했지만 이 영역 내에서 더 많은 연구를 통해 더 많은 연구 영역이 나타날 수 있습니다 . 단일 자동화된 블루 팀 구성 에이전트를 사용하는 것이 유용할 것입니다. [5] 그러나 첫 번째 에이전트를 평가하고 잘못된 결정이 내려지면 경고하는 다른 에이전트가 없으면 근무 시간 외 시간에 에이전트가 저지른 실수는 해결되지 않습니다 .

8.1.5 자동화된 블루 에이전트의 사이버 보안 공격 견고성(A.6.2, A.6.1). 미래의 자동화된 블루 에이전트(및 ACO Gym)를 위한 개선 영역은 기업 네트워크 내에서 발생할 수 있는 더 많은 유형의 사이버 공격을 구현하는 것입니다 . 이를 위한 유용한 프레임워크는 MITRE ATT&CK 매트릭스에 나열된 다양한 사이버 공격을 사용하는 것입니다 . 소프트웨어 업데이트 및 패치와 마찬가지로 시스템은 더 많은 공격이 기술 자료에 나열되면 지식 기반에 추가될 수 있도록 설계 될 수 있습니다.

5https://github.com/cage-challenge/cage-challenge-3

MITRE ATT&CK 매트릭스와 같은 6 가지 프레임워크 . 교육 내에서 이 문제를 해결하지 못하면 자동화된 블루 팀 구성 에이전트가 특정 사이버 공격을 무시하고 결국 네트워크 내 침해로 이어질 것입니다.

8.1.6 자동화된 블루 에이전트의 속임수 기법의 견고화(A.6.4) ACD 목적으로 기만 기술을 활용하는 연구 분야의 필요성을 강조하는 것도 중요합니다 . ACO 체육관 에 포함시키면 사이버 킬 체인을 따라 적을 잘못된 방향으로 유도하고 방해하는 효과를 연구하기 위해 보다 복잡하고 적극적인 방어기만 기술을 도입할 수 있습니다 . 기존 문헌에서는 이 문제의 복잡성을 거의 고려하지 않고 ACD 도구로서 속임수의 초기 단계를 강조합니다. 이 범주에 속하는 연구 [47, 121, 122] 는 일반적으로 지능형 배포 전략을 통해 적의 행동을 분석하기 위해 honey-x 방법 [93] 또는 '미끼' 의 사용을 우선시합니다 . 기만적 자산 내에서 다양성을 장려하는 유용한 프레임워크 는 MITRE ENGAGE 매트릭스입니다. 이는 ACD의 다양한 영역에서 활용하여 적 교전을 최적화할 수 있는 수많은 기만 기술을 식별합니다. 이 문제를 해결하지 못하면 적들이 미끼의 동질성을 무기로 사용 하여 파란색 에이전트와 빨간색 에이전트 사이에 항상 존재하는 비대칭성을 확대할 수 있으므로 속임수의 주요 목적에서 벗어나게 됩니다 .
7

8.1.7 ACO 체육관의 현실성(G.1.3, A.3.1, A.3.2, G.1.4, G.1.1, G.1.2). ACO 체육관의 또 다른 과제는 현재 존재하는 대부분의 환경에 대한 현실성이 부족하다는 것입니다. 연구 분야로서 훈련-테스트(또는 지속적인 학습) 전략 의 품질을 분류하는 척도는 특히 중요합니다. 또한 연구자들은 일반적으로 시뮬레이션된 환경을 구축한 다음 학습된 정책을 실제 세계로 전송해야 합니다 (Sim-to-Real Transfer). 이는 [128] 에서 지적한 대로 로봇공학의 경우에 종종 수행됩니다 . CyBORG [112] 와 같은 환경은 시뮬레이션과 에뮬레이션을 모두 지원하여 이 문제를 해결하려고 시도하지만 두 구현 모두 실제 네트워크 시스템을 나타내지 않는 영역으로 구성됩니다(즉, 시뮬레이션의 대기 시간 지연 및 에뮬레이션의 네트워크 확장성). 또한, 전통적인 RL 작업과 달리 IT 및 OT 네트워크는 지속적이고 끊임없이 변화하는 환경입니다. 대부분의 RL 작업과 달리 기업 환경의 네트워크 및 호스트는 고정적이지 않은 반면, RL이 사용된 비디오 게임은 그렇지 않습니다. 에이전트가 완전히 새로운 지도에서 좋은 성능을 발휘할 것으로 기대합니다 [17, 62, 119]. 이러한 문제는 해결되어야 합니다. 그렇지 않으면 에이전트 가 작동하는 지형의 모드를 인식하지 못하여 에이전트가 잘못된 조치를 취하게 됩니다.

8.1.8 기만 기술의 현실성(A.6.4) 기만 충실도는 현재 문헌에서 제약이나 가정 의 일부로 간과되고 소개되는 경우가 많습니다 . 네트워크 에뮬레이션의 맥락 에서 물리적 자산의 가상화가 점점 보편화됨에 따라 DCD(기만 기반 사이버 방어) 플랫폼의 구현에는 시스템 충실도를 유지하고 공격자에게 그 사용을 알리지 않도록 물리적 프로세스를 모델링하고 시뮬레이션하는 기능이 있어야 합니다. 그러나 시스템 충실도와 상당한 공격 표면 사이의 균형을 맞추는 것은 어렵습니다. 특히 연구원이 설득력 있는 방식으로 장치를 에뮬레이트하는 방법을 찾아야 하는 운영 기술(OT) 환경과 같은 일부 네트워크 시스템의 복잡성과 규모를 고려할 때 더욱 그렇습니다. 네트워크 전체를 복제합니다 . 네트워크 구성 요소의 속성을 구현하는 자산에 대한 미끼 프로필을 생성하는 새로운 방법이 필요합니다 . 연구자들은 이미 통합 미끼의 충실도를 고려하거나 향상시키는 기만적인 기술을 고려할 수도 있습니다 . 'Honeyshills' [57]은 실제 구성 요소나 시스템을 사용 하고 미끼와 통신하도록 구성하여 더욱 사실적인 인상을 주는 예입니다. 이는 시뮬레이션 기반 네트워크 내에서 속임수 방법을 확장하고 궁극적으로 에뮬레이트된 도메인으로 이동하기 위한 제안을 장려합니다. 이러한 문제를 해결하지 못하면 공격자에게 속임수가 노출되어 공격자가 부주의하게 속임수를 사용하는 것보다 우선순위가 무효화될 수 있습니다. 이러한 모순은 속임수 기술을 올바르게 구현함으로써 제공되는 대칭적 이점을 상쇄합니다.

6https://attack.mitre.org/matrices/enterprise/
7https://engage.mitre.org/

8.1.9 잘못된 조치의 영향(G.6.1, G.1.3) [41] 위의 문제는 또한 자동화된 의사 결정 에이전트에 대한 ACD 문헌 내 격차를 초래합니다. 적절한 평가와 지표를 탐구해야 합니다. 또한 자동화된 에이전트에 대한 우수한 포렌식 평가를 위해서는 Hierarchical RL, Neurosymbolic RL 및 기타 설명 가능한 구현 [58, 82, 109] 과 같은 접근 방식을 탐구 해야 합니다 . 이 영역을 해결하지 않으면 자동화된 블루팀 RL 에이전트가 잠재적으로 네트워크 내의 중요한 프로세스를 제거하게 되어 높은 금전적 손실을 초래할 수도 있습니다.

8.1.10 행동 및 관찰 공간(G.2.1,G.2.2,A.2.3) 첫 번째 어려움은 거대한 행동 및 관찰 공간과 관련이 있습니다. ACD의 기존 연구는 행동 공간을 "실제 세계"에서 더 이상 사용할 수 없는 지점까지 추상화하여 행동 및 관찰 공간을 크게 줄입니다. 실제로, 에이전트가 수천 개의 호스트(단일 기업 네트워크 내)에 배포될 수 있는 사이버 보안 설정에서는 각각 거대한 작업 세트(프로세스 종료, 파일 제거/격리, 방화벽 설정 변경 등)가 있고 본질적으로 지속적인 관찰 공간이므로 훈련에서 공간을 충분히 탐색하는 것은 어려울 것입니다. 이 문제는 자동화된 레드 에이전트에도 적용됩니다. "침투 테스트를 자동화하기 위해 기존 DRL을 적용하는 것은 상대적으로 작은 시나리오에서도 작업 공간이 수천 개로 폭발할 수 있기 때문에 어렵고 불안정합니다. " 다른 서브넷에 있는 호스트 또는 다른 공격 방법” [117].

9 결론

이 기사는 연구 출판물, 정부 전략 보고서 및 사이버 보안 교육 기관을 통해 용어를 정의함으로써 자동화된 사이버 방어 분야에 대한 인식을 제공했습니다 . 그 후, 이 용어를 사용하면 해당 영역 내에 존재하는 연구 및 개발의 다양한 하위 주제, 즉 자동화 에이전트 및 자율 사이버 운영(ACO) 체육관을 세분화할 수 있었습니다 . 하위 주제를 인식함으로써 ACD(자동 사이버 방어) 문헌의 일부로 인식된 출판물을 평가하기 위한 측정 기준으로 사용되는 요구 사항 분석을 생성할 수 있었습니다. 맞춤형 ACO 체육관 내의 자동화된 블루 및 레드 팀 구성 알고리즘에 대한 기존 문헌을 광범위하게 검토한 결과 , 게임 이론 및 기계 학습 솔루션과 비교하여 심층 강화 학습(DRL) 솔루션이 자동화된 블루 및 레드 팀 구성에 가장 적합한 알고리즘이라는 사실이 밝혀졌습니다. . 이는 주로 장기 및 단기 목표에 대해 순차적으로 대응할 수 있는 능력 때문이었습니다 . 또한 검토에서는 두 영역 각각의 연구를 가속화하기 위해 자동화된(청색 및 빨간색) 에이전트와 ACO Gym의 병행 연구 및 개발이 필요하다고 제안했습니다. 자동화된 레드 및 블루 팀 구성 구현과 함께 기존 오픈 소스 및 비공개 소스 체육관에 대한 광범위한 검토도 수행되었습니다 . 요구사항 분석을 통해 출판물 과 구현을 평가하여 문헌 내에서 추가 개발 영역을 찾았습니다 . 자동화 에이전트 및 ACO Gym의 모든 출판물에 대한 요구 사항 분석을 통해 특정 과제와 격차를 발견하고 정교화했습니다. ACD 영역의 과제에는 자동화된 블루 에이전트의 사이버 보안 공격 강화, ACO 체육관의 현실성, 잘못된 조치의 영향 및 ACO 체육관 내의 조치/관찰 공간에 대한 연구가 포함되었습니다. 대조적으로, 자동화된 블루 에이전트에 대한 알고리즘 공격, 자동화된 블루 에이전트로 설명 가능한 RL 및 다중 에이전트 자동화 블루 에이전트에 대한 연구에는 격차가 있었습니다. 과제와 격차의 목표 는 자동화된 블루 에이전트를 ACO 체육관에서 실제 세계에 배포된 네트워크 시스템으로 전환하기 위해 ACD 내 미래 연구 및 개발 영역을 다룹니다 .

참고자료

[1] 2022. 엔맵(2022). https://nmap.org/ [2]
2022. 네소스(2022). https://www.tenable.com/products/nessus [3]
2022. CVSS(2022). https://nvd.nist.gov/vuln-metrics/cvss [4]
2022. 블러드하운드(2022). https://github.com/BloodHoundAD/BloodHound

[5] 2022. 메타스플로잇(2022). https://www.metasploit.com/
[6] 2022. 엠파이어(2022). https://github.com/EmpireProject/Empire
[7] 2022. 마이터(2022). https://attack.mitre.org/matrices/enterprise/
[8] 2022. 사이버작전연구실. https://github.com/cage-challenge/CybORG. Maxwell Standen, David Bowman, Son Hoang, Toby Richer, Martin Lucas, Richard Van Tassel, Phillip Vu, Mitchell Kiely, KC C., Natalie Konschnik, Joshua Collyer가 제작했습니다. https: //
[9] 퀸즈랜드 국방 과학 동맹. 2022. 의사결정 이니셔티브를 위한 인공지능(2022).
queenslanddefencesciencealliance.com.au/federal-and-state-defence-funding-opportunities-2/artificial-intelligence-for-decision-making-initiative-round-2022/
_ [10] Prithviraj Ammanabrolu 및 Mark Riedl. 2019. 그래프 기반 심층 강화 학습을 이용한 텍스트 어드벤처 게임 플레이.
3557-3565. https://doi.org/10.18653/v1/N19-1358
[11] Alex Andrew, Sam Spillard, Joshua Collyer 및 Neil Dhir. 2022. 사이버 보안을 통한 최적의 인과적 사이버 방어 에이전트 개발
시뮬레이션. 사이버 보안을 위한 머신러닝 워크숍(ML4Cyber)에서.
[12] Alex Andrew, Sam Spillard, Joshua Collyer 및 Neil Dhir. 2022. 사이버 보안을 통한 최적의 인과적 사이버 방어 에이전트 개발
시뮬레이션. arXiv 사전 인쇄 arXiv:2207.12355 (2022).
[13] Andy Applebaum, Camron Dennler, Patrick Dwyer, Marina Moskowitz, Harold Nguyen, Nicole Nichols, Nicole Park, Paul Rachwalski, Frank Rau, Adrian Webster 등 2022. 자동화와 자율 사이버 방어 연결: 테이블 형식 q-학습의 기초 분석.
제15회 인공지능 및 보안에 관한 ACM 워크숍 진행 중. 149-159.
[14] Giovanni Apruzzese, Mauro Andreolini, Mirco Marchetti, Andrea Venturi 및 Michele Colajanni. 2020. 봇넷 회피 공격에 대한 심층 강화 적대 학습. 네트워크 및 서비스 관리에 관한 IEEE 거래 17, 4(2020), 1975-1987. https: //doi.org/10.1109/TNSM.2020.3031843 [15] 조지 K. 바(George K. Baah), 토마스 홉슨(Thomas Hobson), 하마드 오크라비
(Hamad Okhravi), 섀넌 C. 로버츠(Shannon C. Roberts), 윌리엄 W. 스트라일린(William W. Streilein), 소피아 유디츠카야(Sophia Yuditskaya). 2015. 연구
사이버 방어 자동화의 격차.
[16] David Paul Benjamin, Partha Pal, Franklin Webber, Paul Rubel 및 Mike Atigetchi. 2008. 인지 아키텍처를 사용하여 사이버 방어 추론을 자동화합니다. 2008년 보안을 위한 생체 영감, 학습 및 지능형 시스템. 58-63. https://doi.org/10.1109/BLISS.2008.17 [17] Christopher Berner, Greg Brockman,
Brooke Chan, Vicki Cheung, Przemys aw D biak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse 등 2019. 대규모 심층 강화 학습을 갖춘 Dota 2. arXiv 사전 인쇄 arXiv:1912.06680 (2019).
[18] 라숀 부커(Lashon Booker)와 스콧 머스만(Scott Musman). 2022. 자동화된 사이버 대응에 대한 모델 기반 의사결정 이론적 관점. 자율 지능형 사이버 방어 요원에 관한 국제 회의.
[19] Robert A Bridges, Ashley E Rice, Sean Oesch, Jeff A Nichols, Cory Watson, Kevin Spakes, Savannah Norem, Mike Huettel, Brian Jewell, 브라이언 웨버, 그 외 여러분. 2022. 사용 중인 SOAR 도구 테스트. arXiv 사전 인쇄 arXiv:2208.06075 (2022).
[20] 스콧 브라운, 해롤드 브라운, 마이클 러셀, 브라이언 헨즈, 마이클 에드워즈, 프랭크 터너, 조르지오 베르톨리. 2016. 예제 악성코드를 이용한 네트워크 시뮬레이션 모델 검증 및 확장성 테스트. MILCOM 2016 - 2016 IEEE 군사 통신 컨퍼런스. 491-496. https://doi.org/10.1109/MILCOM.2016.7795375

[21] Miles Brundage, Shahar Avin, Jasmine Wang, Haydn Belfield, Gretchen Krueger, Gillian Hadfield, Heidy Khlaaf, Jingying Yang, Helen Toner, Ruth Fong 등. 2020. 신뢰할 수 있는 AI 개발을 향하여: 검증 가능한 주장을 지원하는 메커니즘. arXiv 사전 인쇄 arXiv:2004.07213 (2020).

[22] Ricardo Buettner, Daniel Sauter, Jonas Klopfer, Johannes Breitenbach 및 Hermann Baumgartl. 2021. 사이버 방어를 위한 머신러닝 접근 방식 의 최근 발전에 대한 검토 . 2021년 IEEE 빅데이터 국제컨퍼런스(빅데이터). IEEE, 3969-3974.
[23] A. 버크. 2017 [온라인]. 능동적 사이버 방어를 위한 강력한 인공 지능. 앨런 튜링 연구소. https://www.turing.ac.uk/sites/default/files/2020-08/publicaiacdtechreportfinal.pdf
[24] 하산 캠. 2020. 자율 에이전트와 강화 학습을 사용한 사이버 탄력성. 다중 도메인 운영 애플리케이션을 위한 인공 지능 및 기계 학습 II, Tien Pham, Latasha Solomon 및 Katie Rainey(Eds.), Vol. 11413. 국제 광학 및 포토닉스 협회, SPIE, 219 – 234. https://doi.org/10.1117/12.2559319

[25] Xinzhong Chai, Yasen Wang, Chuanxu Yan, Yuan Zhao, Wenlong Chen 및 Xiaolei Wang. 2020. DQ-MOTAG: DDoS 공격에 대한 심층 강화 학습 기반 이동 표적 방어. 375-379. https://doi.org/10.1109/DSC50466.2020.00065 [26] 첸 유잉, 첸 치아오팅, 추안윤상, 양야오춘, 황쓰하오.
2021. 강화 학습 기반 포트폴리오 관리 전략에 대한 적대적 공격 . IEEE 액세스 9(2021), 50667-50685. https://doi.org/10.1109/ACCESS.2021.3068768 [27] Chwee Seng Choo, Ching Lian Chua 및 Su-Han Victor Tay. 2007. 자동화된 레드팀 구성: 군사 적용을 위해 제안된 프레임워크. 유전 및 진화 계산에 관한 제9차 연례 회의 진행(영국 런던)(GECCO '07).

컴퓨팅 기계 협회, 뉴욕, 뉴욕, 미국, 1936-1942. https://doi.org/10.1145/1276958.1277345 [28] Ankur Chowdhary, Dijiang
Huang, Jayasurya Sevalur Mahendran, Daniel Romo, Yuli Deng 및 Abdulhakim Sabur. 2020. 자율적인 보안 분석 및 침투 테스트. 2020년 제16회 모빌리티, 센싱, 네트워킹(MSN)에 관한 국제 컨퍼런스. IEEE, 508-515.

[29] Ankur Chowdhary, Dijiang Huang, Abdulhakim Sabur, Neha Vadnere, Myong Kang 및 Bruce Montrose. 2021. 다중 에이전트 강화학습을 이용한 SDN 기반 이동 표적 방어. 자율지능형사이버방어요원회의.

[30] Edward JM Colbert, Alexander Kott, Lawrence P Knachel. 2020. 사이버워게임의 게임이론적 모델과 실험적 고찰 . 국방 모델링 및 시뮬레이션 저널 17, 1(2020), 21–38.
[31] 조쉬 콜리어(Josh Collyer), 알렉스 앤드류(Alex Andrew), 던컨 호지스(Duncan Hodges). 2022. ACD-G: 자율사이버방어 에이전트 일반화 강화 그래프 임베디드 네트워크 표현. 기계 학습에 관한 국제 회의.
[32] Richard Dazeley, Peter Vamplew 및 Francisco Cruz. 2021. broad-xai에 대한 설명 가능한 강화 학습: 개념적 프레임워크 그리고 설문조사. arXiv 사전 인쇄 arXiv:2108.09003 (2021).
[33] 국방. 2021. 캐나다 정부. https://www.canada.ca/en/department-national-defence/programs/defence- ideas/element/innovation-networks/challenge/autonomous-systems-defence-security-trust-barriers-adoption.html [34] 수잔나 케이트 데빗(Susannah Kate Devitt)
과 데미안 코플랜드(Damian Copeland). 2021. 보안 및 국방 분야의 AI 거버넌스에 대한 호주의 접근 방식. https: //doi.org/10.48550/ARXIV.2112.01252 [35] Neil Dhir, Henrique Hoeltgebaum, Niall
Adams, Mark Briers, Anthony Burke 및 Paul Jones. 2021. 미래의 인공지능 적극적인 사이버 방어를 위한 접근 방식. arXiv 사전 인쇄 arXiv:2104.09981 (2021).
[36] 맥스웰 돈도(Maxwell Dondo)와 나탈리아 나클라(Natalia Nakhla). 2021. 네트워크 운영에서 자율 방어 사이버 작전을 위한 프레임워크를 향하여 센터. (2021). https://cradpdf.drdc-rddc.gc.ca/PDFS/unc382/p814083_A1b.pdf
[37] Taha Eghtesad, Yevgeniy Vorobeychik 및 Aron Laszka. 2020. 적대적 심층 강화 학습 기반 적응형 이동 표적 방어. 보안을 위한 의사결정과 게임 이론: 제11차 국제 컨퍼런스(2020년 12월), 58–79. https://doi.org/10.1007/978-3- 030-64793-3_4

[38] Thomas C. Eskridge, Marco M. Carvalho, Evan Stoner, Troy Toggweiler 및 Adrian Granados. 2015. VINE: MTD 실험을 위한 사이버 에뮬레이션 환경 (MTD '15). 컴퓨팅 기계 협회, 뉴욕, 뉴욕, 미국, 43–47. https: //doi.org/10.1145/2808475.2808486

[39] Zhiyang Fang, Junfeng Wang, Boya Li, Siqi Wu, Yingjie Zhou 및 Haiying Huang. 2019. 심층 강화 학습을 통해 맬웨어 방지 엔진 회피 . IEEE 액세스 7(2019), 48867-48879.
[40] Myles Foley, Chris Hicks, Kate Highnam 및 Vasilios Mavroudis. 2022. 강화학습을 이용한 자율 네트워크 방어. 2022 컴퓨터 및 통신 보안에 관한 아시아 컨퍼런스 ACM 진행(일본 나가사키)(ASIA CCS '22). 컴퓨팅 기계 협회, 뉴욕, 뉴욕, 미국, 1252-1254. https://doi.org/10.1145/3488932.3527286 [41] 사이버 방어를 위한 AI 작동: 정확성-강
건성 트레이드오프. 2021. https://doi.org/10.51593/2021CA007 See More [42] 키스 프랭키시(Keith Frankish)와 윌리엄 램지(William
Ramsey). 2014. 캠브리지 인공지능 핸드북. 케임브리지 대학 출판부. https: //doi.org/10.1017/CBO9781139046855 [43] 안젤로 푸르파로, 안토니오 피콜로, 안드레아 파리세, 루치아노 아르젠토, 도메니코 사카.
2018. 복잡한 사이버 보안 시나리오를 에뮬레이션하기 위한 클라우드 기반 플랫폼입니다. 미래 세대 컴퓨터 시스템 89(2018), 791-803. https://doi.org/10.1016/j.future.2018.07.025 [44] Ariel Futoransky, Fernando Miranda, José Orlicki, Carlos Sarraute. 2009. 재미와 이익을 위해 사이버 공격을 시뮬레이션합니다. 컴퓨팅
연구 저장소 - CORR, 4. https://doi.org/10.1145/1537614.1537620
[45] Rohit Gangupantulu, Tyler Cody, Abdul Rahma, Christopher Redino, Ryan Clark 및 Paul Park. 2021. 공격 그래프를 이용한 강화 학습을 이용한 Crown Jewels 분석 . 2021년 IEEE 계산 지능 심포지엄 시리즈(SSCI). IEEE, 1-6.
[46] 가오 중앙강(Chungang Gao)과 왕용지에(Yongjie Wang). 2021. DDoS 공격에 대한 강화학습 기반의 자가 적응형 이동 표적 방어. 물리학 저널: 컨퍼런스 시리즈 1812(2021년 02월), 012039. https://doi.org/10.1088/1742-6596/1812/1/012039
[47] Yazhuo Gao, Guomin Zhang 및 Changyou Xing. 2021. 지능형 공격 경로 예측을 기반으로 하는 가상화된 허니팟의 다단계 동적 배포 메커니즘 . 보안 및 통신 네트워크 2021(2021년 10월). https://doi.org/10.1155/2021/6378218 [48] Maxime Gasse, Damien Grasset, Guillaume Gaudron 및 Pierre-Yves Oudeyer. 2021. 관찰 및 중재 데이터를 사용한 인과 강화 학습 . arXiv 사전 인쇄 arXiv:2106.14421 (2021).

[49] Claire Glanois, Paul Weng, Matthieu Zimmer, Dong Li, Tianpei Yang, Jianye Hao 및 Wulong Liu. 2021. 해석 가능한 강화 학습에 관한 설문조사 . arXiv 사전 인쇄 arXiv:2112.13112 (2021).
[50] Ross Gore, Saikou Diallo, Jose Padilla 및 Barry Ezell. 2018. 머신러닝을 활용한 사이버 사고 평가. 정보 및 컴퓨터 보안에 관한 국제 저널 10(2018년 1월), 341. https://doi.org/10.1504/IJICS.2018.095298 [51] TTCP Cage Working Grouo. 2022. TTCP CAGE 챌린지 2.
https://github.com/cage-challenge/cage-challenge-2.
[52] TTCP CAGE 워킹 그룹. 2021. CAGE 챌린지 1. arXiv.
[53] TTCP CAGE 실무 그룹. 2022. TTCP CAGE 챌린지 3. https://github.com/cage-challenge/cage-challenge-3.
[54] 킴 함 마르(Kim Hammar)와 롤프 슈타들러(Rolf Stadler). 2020. 강화 학습과 Self-Play를 통해 효과적인 보안 전략을 찾습니다. 2020년 16차 네트워크 및 서비스 관리에 관한 국제 컨퍼런스(CNSM). IEEE, 1–9.
[55] 킴 함 마르(Kim Hammar)와 롤프 슈타들러(Rolf Stadler). 2021. 최적의 차단을 통한 침입방지 정책을 학습합니다. 2021년 17차 인터내셔널 네트워크 및 서비스 관리 컨퍼런스(CNSM). IEEE, 509-517.
[56] 베른하르트 헝스트. 2010. 계층적 강화 학습. Springer 미국, 보스턴, 매사추세츠, 495–502. https://doi.org/10.1007/978-0-387-30164-8_363
[57] William Hofer, Thomas Edgar, Draguna Vrabie 및 Kathleen Nowak. 2019. 제어 시스템 환경을 위한 모델 기반 속임수. 2019년 IEEE 국토 안보 기술 국제 심포지엄(HST)에서. IEEE, 1-7.

[58] 로버트 R. 호프만, 셰인 T. 뮬러, 게리 클라인, 조던 리트먼. 2018. 설명 가능한 AI 측정항목: 과제 및 전망.
https://doi.org/10.48550/ARXIV.1812.04608
[59] 와이어트 호프만. 2021. 사이버 방어를 위한 AI 활용. (2021).
[60] Linan Huang과 Quanyan Zhu. 2019. 중요 인프라의 지능형 지속 위협에 대한 적응형 전략적 사이버 방어
네트워크. SIGMETRICS 수행. 평가. 개정판 46, 2(2019년 1월), 52–56. https://doi.org/10.1145/3305218.3305239 [61] 황윤한
(Yunhan Huang), 황리난(Linan Huang), 주취안얀(Quanyan Zhu). 2021. 피드백 지원 사이버 탄력성을 위한 강화 학습.
[62] 황윤한(Yunhan Huang)과 주취안옌(Quanyan Zhu). 2019. 비용 신호에 대한 적대적 조작 하의 기만적 강화 학습. ~ 안에
GameSec.
[63] 로렌스 아우 존슨 키뉴아. 2021. 보안 오케스트레이션, 자동화 및 대응 분야의 AI/ML: 향후 연구 방향.
지능형 자동화 및 소프트 컴퓨팅 28, 2(2021), 527–545. https://doi.org/10.32604/iasc.2021.016240
[64] Jean Kaddour, Aengus Lynch, Qi Liu, Matt J. Kusner 및 Ricardo Silva. 2022. 인과관계 머신러닝: 설문조사 및 미해결 문제.
https://doi.org/10.48550/ARXIV.2206.15475
[65] 직원 Keele et al. 2007. 소프트웨어 엔지니어링의 체계적인 문헌 검토 수행을 위한 지침. 기술 보고서. 인위적인
보고서, 버전 2.3 ebse 기술 보고서. ebse.
[66] 김미영, Shahin Atakishiyev, Housam Khalifa Bashier Babiker, Nawshad Farruque, Randy Goebel, Osmar R. Zaïane, Mohammad- Hossein
Motallebi, Juliano Rabelo, Talat Syed, Hengshuai Yao 및 Peter Chun. 2021. 설명 가능한 인공 지능의 분석 및 설계를 위한 다중 구성 요소 프레임워크 . 기계
학습 및 지식 추출 3, 4(2021), 900-921. https: //www.mdpi.com/2504-4990/3/4/45

[67] 바바라 키친햄. 2004. 체계적인 검토 수행 절차. Keele, 영국, Keele Univ. 33(2004년 8월).
[68] 라이언 KL 고. 2020. 사이버 자율성: 해커의 자가 치유, 자가 적응형 자동 사이버 방어 시스템과 이것이 산업, 사회 및 국가 안보에 미치는 영향을 자동화합니다. 신흥 기술
및 국제 안보 분야. 루트리지, 173-191.
[69] Alexander Kott, Paul Théron, Martin Dra ar, Edlira Dushku, Benoît LeBlanc, Paul Losiewicz, Alessandro Guarino, Luigi Mancini, Agostino
Panico, Mauno Pihelgas 등 2018. 자율 지능형 사이버 방어 에이전트(aica) 참조 아키텍처. 릴리스 2.0. arXiv 사전 인쇄 arXiv:1803.10664 (2018).

[70] Kalle Kujanpää, Willie Victor 및 Alexander Ilin. 2021. 심층 강화 학습을 통한 권한 상승 자동화. 제14차 ACM 인공지능 및 보안 워크숍 진행 중 . 157-168.

[71] Huanruo Li, Yunfei Guo, Penghao Sun, Yawen Wang 및 Shumin Huo. 2022. 심층 강화 학습을 갖춘 컨테이너 기반 클라우드를 위한 최적의 방어기만 프레임워
크 . IET 정보 보안 16, 3(2022), 178–192. https://doi.org/10.1049/ise2. 12050 arXiv:https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/
ise2.12050
[72] Li Li, Raed Fayad 및 Adrian Taylor. 2021. Cygil: 에뮬레이트된 네트워크 시스템을 통해 자율 에이전트를 교육하기 위한 사이버 체육관입니다. arXiv
사전 인쇄 arXiv:2109.03331(2021).
[73] 마이클 리트먼. 2009. 순차적 의사결정을 위한 알고리즘. (2009년 8월).
[74] Daoming Lyu, Fangkai Yang, Bo Liu 및 Steven Gustafson. 2019. SDRL: 상징적 계획을 활용하는 해석 가능하고 데이터 효율적인 심층 강화 학습. 인공 지능에 관
한 AAAI 회의 간행물 33(2019년 7월), 2970-2977. https://doi.org/10.1609/aaai.v33i01.33012970

[75] Prashan Madumal, Tim Miller, Liz Sonenberg 및 Frank Vetere. 2019. 인과적 렌즈를 통한 설명 가능한 강화 학습.
https://doi.org/10.48550/ARXIV.1905.10958
[76] 마에다 류세이와 미무라 마모루. 2021. 심층 강화 학습을 통해 공격 후 자동화. 컴퓨터 및 보안 100
(2021), 102108.
[77] Mohamad Imad Mahaini, Shujun Li 및 Rahime Belen Sa lam. 2019. 인간-기계 팀 구성을 기반으로 한 분류 체계 구축: 사이버 보안을 예로 들었습니다. 가용
성, 신뢰성 및 보안에 관한 제14차 국제 컨퍼런스 진행 중. 1~9.
[78] Kleanthis Malialis 및 Daniel Kudenko. 2013. 협력 강화 학습을 이용한 대규모 DDoS 대응.
[79] Erik Miehling, Mohammad Rasouli 및 Demosthenis Teneketzis. 2015. 베이지안 공격 그래프에서 부분적으로 관찰 가능한 확산 프로세스를 위한 최적의 방어 정
책 . 이동 표적 방어에 관한 두 번째 ACM 워크숍 진행 중. 67-76.
[80] Stephanie Milani, Nicholay Topin, Manuela Veloso 및 Fei Fang. 2022. 설명 가능한 강화 학습에 대한 조사. arXiv 사전 인쇄
arXiv:2202.08434 (2022).
[81] Jelena Mirkovic, Terry V. Benzel, Ted Faber, Robert Braden, John T. Wroclawski 및 Stephen Schwab. 2010. DETER 프로젝트: 사이버 보안 실험 및 테
스트 과학 발전. 2010년 IEEE 국토 안보 기술 국제 회의 (HST). 1~7. https://doi.org/10.1109/THS.2010.5655108

[82] Ludovico Mitchener, David Tuckey, Matthew Crosby 및 Alessandra Russo. 2022. 감지, 이해, 행동: 신경 상징
계층적 강화 학습 프레임워크. 기계 학습 111, 4(2022), 1523-1549.
[83] Andres Molina-Markham, Cory Miniter, Becky Powell 및 Ahmad Ridley. 2021. 자율주행을 위한 네트워크 환경 설계
사이버 방어. ArXiv ABS/2103.07583 (2021).
[84] 나토. 2021. 군의 인공지능과 자율성. https://ccdcoe.org/uploads/2021/12/Strategies_and_Deployment_
A4.pdf

[85] 나토. 2022. 협력적 사이버 방어 우수 센터. https://ccdcoe.org/library/publications/ [86] Hoang Viet Nguyen, Hai
Ngoc Nguyen, 우에하라 테츠타로. 2020. 침투 테스트를 위한 다중 레벨 액션 임베딩. 제4회 미래 네트워크 및 분산 시스템에 관한 국제 컨퍼런스(ICFNDS) 에서 . 1~9.

[87] Thanh Thi Nguyen 및 Vijay Janapa Reddi. 2021. 사이버 보안을 위한 심층 강화 학습. 신경망 및 학습 시스템 에 대한 IEEE 트랜잭션 (2021).

[88] Zhen Ni와 Shuva Paul. 2019. 스마트 그리드 보안의 다단계 게임: 강화 학습 솔루션. IEEE 거래
신경망 및 학습 시스템 30, 9(2019), 2684-2695. https://doi.org/10.1109/TNNLS.2018.2885530
[89] Constantin Nilă, Ioana Apostol 및 Victor Patriciu. 2020. 신속한 사고 대응을 위한 머신러닝 접근 방식. 2020년 제13차 국제통신학회(COMM)에서.
291-296. https://doi.org/10.1109/COMM48946.2020.9141989 [90] 내각실. 2022. 정부 사이버 보안 전략. https://www.gov.uk/
government/publications/government-cyber-
보안 전략-2022년부터 2030년까지
[91] Matthew L Olson, Roli Khanna, Lawrence Neal, Fuxin Li 및 Weng-Keen Wong. 2021. 에 대한 반사실적 상태 설명
생성적 딥러닝을 통한 강화학습 에이전트. 인공지능 295(2021), 103455.
[92] Deepak Pathak, Pulkit Agrawal, Alexei A. Efros 및 Trevor Darrell. 2017. 자기주도 예측을 통한 호기심 중심 탐색.
https://doi.org/10.48550/ARXIV.1705.05363
[93] Jeffrey Pawlick, Edward Colbert 및 Quanyan Zhu. 2017. 게임이론적 분류와 방어기만 조사
사이버 보안 및 개인 정보 보호. https://doi.org/10.48550/ARXIV.1712.05441
[94] Shaohui Peng, Xing Hu, Rui Zhang, Ke Tang, Jiaming Guo, Qi Yi, Ruizhi Chen, Xishan Zhang, Zidong Du, Ling Li, Qi Guo 및 Yunji Chen .
2022. 강화 학습을 위한 인과 중심 계층 구조 발견. https://doi.org/10.48550/ARXIV.2210.06964 [95] XIANGYU PENG, Mark Riedl 및 Prithviraj
Ammanabrolu. 2022. 자연어에서 본질적으로 설명 가능한 강화 학습.
신경 정보 처리 시스템의 발전, Alice H. Oh, Alekh Agarwal, Danielle Belgrave 및 조경현(Eds.). https://openreview.net/forum?id=DSEP9rCvZln

[96] Erika Puiutta 및 Eric MSP Veith. 2020. 설명 가능한 강화 학습: 설문조사. 기계 학습 및 지식 추출에서 Andreas Holzinger, Peter Kieseberg, A Min Tjoa
및 Edgar Weippl(Eds.). Springer International Publishing, Cham, 77-95.
[97] Hamon R, Junklewitz H, Sanchez Martin JI. 2020. 인공지능의 견고성과 설명 가능성. KJ-NA-30040-EN-N
(온라인) (2020). https://doi.org/10.2760/57493(온라인)
[98] Manjeet Rege 및 Raymond Blanch K Mbah. 2018. 사이버 방어 및 공격을 위한 머신러닝. 데이터 분석 2018(2018), 83.
[99] Kezhou Ren, Yifan Zeng, Zhiqin Cao 및 Yingchao Zhang. 2022. ID-RDRL: 심층 강화 학습 기반 특징 선택
침입 탐지 모델. 과학 보고서 12(2022년 9월). https://doi.org/10.1038/s41598-022-19366-3
[100] Danilo J. Rezende, Ivo Danihelka, George Papamakarios, Nan Rosemary Ke, Ray Jiang, Theophane Weber, Karol Gregor, Hamza Merzic,
Fabio Viola, Jane Wang, Jovana Mitrovic, Frederic Besse, Ioannis Antonoglou 및 Lars Buesing. 2020. 강화 학습을 위한 인과적으로 올바른 부분
모델 . https://arxiv.org/abs/2002.02836 [101] Ciaran Roberts, Sy-
Toan Ngo, Alexandre Milesi, Sean Peisert, Daniel Arnold, Shammya Saha, Anna Scaglione, Nathan Johnson, Anton Kocheturov 및 Dmitriy
Fradkin. 2020. 사이버 공격 완화를 위한 심층 강화 학습. 2020년 스마트 그리드를 위한 통신, 제어 및 컴퓨팅 기술에 관한 IEEE 국제 컨퍼런스(SmartGridComm).
IEEE, 1-7.
[102] 조지 러시, 다니엘 R. 타우리츠, 알렉산더 D. 켄트. 2015. 공진화 에이전트 기반 네트워크 방어 경량 이벤트 시스템(CANDLES)(GECCO Companion '15). 컴퓨팅
기계 협회, 뉴욕, 뉴욕, 미국, 859-866. https://doi.org/10.
1145/2739482.2768429
Kevin Schoonover , Eric Michalak, Sean Harris, Adam Gausmann, Hannah Reinbolt, Daniel Tauritz, Chris Rawlings 및 Aaron Pope.
2018. 갤럭시: 사이버 보안을 위한 네트워크 에뮬레이션 프레임워크.
[104] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford 및 Oleg Klimov. 2017. 근위 정책 최적화 알고리즘.
https://doi.org/10.48550/ARXIV.1707.06347
[105] Jonathon Schwartz와 Hanna Kurniawati. 2019. 강화학습을 이용한 자율 침투 테스트. ArXiv ABS/1905.05965
(2019).
[106] 조나단 슈워츠(Jonathon Schwartz)와 한나 쿠르니아와티(Hanna Kurniawatti). 2019. NASim: 네트워크 공격 시뮬레이터. (2019).
[107] Mohit Sewak, Sanjay K. Sahay 및 Hemant Rathore. 2022. 사이버 보안 위협 탐지 및 보호를 위한 심층 강화 학습: 검토. 인공지능 시대의 안전한 지식관리. 스
프링거 국제 출판, 51-72. https://doi.org/10.1007/978-3-030-97532-6_4

[108] Ali Shafahi, W. Ronny Huang, Mahyar Najibi, Octavian Suciu, Christoph Studer, Tudor Dumitras 및 Tom Goldstein. 2018. 독 개구리! 신경망
에 대한 표적화된 청정 라벨 중독 공격. 신경 정보 처리 시스템의 발전, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi 및 R. Garnett
(Eds.), Vol. 31. 커란 어소시에이츠(Curran Associates, Inc.) https://proceedings. neurips.cc/paper/2018/file/22722a343513ed45f14905eb07621686-
Paper.pdf
[109] Tianmin Shu, Caiming Xiong, Richard Socher. 2017. 다중작업 강화에서 계층적이고 해석 가능한 기술 습득
학습. arXiv 사전 인쇄 arXiv:1712.07294 (2017).

[110] Ryan Silva, Cameron Hickert, Nicolas Sarfaraz, Jeff Brush, Josh Silbermann 및 Tamim Sookoor. 2022. AlphaSOC: 사이버 물리 시스템을 위한 강화 학습 기반 사이버 보안 자동화. 2022년 ACM/IEEE 13차 사이버 물리시스템 국제 컨퍼런스(ICCCS). IEEE, 290-291.

벤 스펜서와 스티브 쿠퍼. [nd].
[112] Maxwell Standen, Martin Lucas, David Bowman, Toby J. Richer, 김준애, Damian Marriott. 2021. CybORG: 체육관 자율 사이버 에이전트의 개발. arXiv:2108.09118 [cs.CR]
[113] 마데나 술타나(Madeena Sultana), 아드리안 테일러(Adrian Taylor), 리 리(Li Li). 2021. 심층 강화 학습을 통한 자율 네트워크 사이버 공격 전략. 다중 도메인 운영 애플리케이션을 위한 인공 지능 및 기계 학습 III, Vol . 11746. SPIE, 490-502.
[114] Jie Tan, Zhaoming Xie, Byron Boots 및 C. Karen Liu. 2016. 휴머노이드 밸런싱을 위한 동적 컨트롤러의 시뮬레이션 기반 설계. 2016 년 IEEE/RSJ 지능형 로봇 및 시스템에 관한 국제 컨퍼런스(IROS). 2729-2736. https://doi.org/10.1109/IROS.2016.7759424
[115] 마이크로소프트 디펜더 연구팀. 2021. (2021).
[116] Ilaria Tiddi와 Stefan Schlobach. 2022. 설명 가능한 기계 학습을 위한 도구로서의 지식 그래프: 설문조사. 인공지능 302(2022), 103627. https://doi.org/10.1016/j.artint.2021.103627 [117] Khuong Tran, Ashlesha Akella, Maxwell
Standen, 김준애, David Bowman, Toby Richer 및 Chin-Teng Lin. 2021. 심층 계층 자동화된 침투 테스트를 위한 강화제. arXiv:2109.06449 [cs.AI]
[118] Vladislav D. Veksler, Norbou Buchler, Claire G. LaFleur, Michael S. Yu, Christian Lebiere 및 Cleotilde Gonzalez. 2020. 사이버 보안의 인지 모델: 전문 분석가로부터 학습하고 공격자 행동 예측. 심리학 11(2020)의 프론티어. https://doi.org/10.3389/fpsyg.2020.01049

[119] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, 정준영, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. 2019. 다중 에이전트 강화 학습을 사용한 스타크래프트 II의 그랜드마스터 레벨. 자연 575, 7782(2019), 350–354.

[120] 벤 월리스. 2022. 국방인공지능전략. 지능-전략/방어-인공지능-전략 https://www.gov.uk/government/publications/defence-artificial-

[121] 에리히 월터, 킴벌리 퍼거슨-월터, 아마드 리들리. 2021. 자율 방어를 위해 사이버 전투 시뮬레이션에 속임수를 통합합니다. arXiv 사전 인쇄 arXiv:2108.13980 (2021).

[122] Shuo Wang, Qingqi Pei, Jianhua Wang, Guangming Tang, Yuchen Zhang 및 Xiaohu Liu. 2020. 강화 학습 기반 속임수 자원 에 대한 지능형 배포 정책 . IEEE 액세스 8(2020), 35792-35804. https://doi.org/10.1109/ACCESS.2020.2974786
[123] Wenhao Wang, Dingyuanhao Sun, Feng Jiang, Xingguo Chen 및 Cheng Zhu. 2022. 강화에 대한 연구와 도전 인트라넷 보안을 위한 사이버 방어 의사결정 학습. 알고리즘 15, 4(2022), 134.
[124] Melody Wolk, Andy Applebaum, Camron Denver, Patrick Dwyer, Marina Moskowitz, Harold Nguyen, Nicole Nichols, Nicole Park, Paul Rachwalski, Frank Rau 등 2022. CAGE를 넘어: 학습된 자율 네트워크 방어 정책의 일반화 조사. arXiv 사전 인쇄 arXiv:2211.15557(2022).

[125] Annie Wong, Thomas Bäck, Anna V Kononova 및 Aske Plaat. 2021. 다중 에이전트 심층 강화 학습: 인간과 유사한 접근 방식에 대한 과제와 방향. arXiv 사전 인쇄 arXiv:2106.15691(2021).
[126] Mattia Zago, Víctor Sánchez, Manuel Pérez 및 Gregorio Martinez Perez. 2017. 자동 의사결정을 통해 사이버 위협에 대처하고 기계 학습 기술을 기반으로 한 반응.
[127] Matej Ze evi , Devendra Singh Dhami, Petar Veli kovi 및 Kristian Kersting. 2021. 그래프 신경망과 구조적 인과 모델의 관계. https://doi.org/10.48550/ARXIV.2109.04173 [128] Wenshuai Zhao, Jorge Peña
Queralta 및 Tomi Westerlund. 2020. 로봇공학을 위한 심층 강화 학습의 시뮬레이션에서 실제로의 전환: 조사. 2020년 IEEE 계산 지능 심포지엄 시리즈(SSCI). IEEE, 737-744.
[129] 주천양(Tian-yang Zhou), 장이차오(Yi-chao Zang), 주준후(Jun-hu Zhu), 왕칭셴(Qing-xian Wang). 2019. NIG-AP: 자동화된 침투 테스트를 위한 새로운 방법. 정보 기술 및 전자 공학의 개척자 20, 9(2019), 1277-1288.
[130] Chen Zhu, W. Ronny Huang, Hengduo Li, Gavin Taylor, Christoph Studer 및 Tom Goldstein. 2019. 심층 신경망에 대한 양도 가능한 클린 라벨 중독 공격. 기계 학습에 관한 제36차 국제 컨퍼런스 진행(기계 학습 연구 진행, Vol. 97), Kamalika Chaudhuri 및 Ruslan Salakhutdinov(Eds.). PMLR, 7614-7623. https://proceedings.mlr.press/v97/zhu19a. HTML

[131] Saman A. Zonouz, Himanshu Khurana, William H. Sanders 및 Timothy M. Yardley. 2009. RRE: 게임 이론상의 침입 대응 및 복구 엔진. 2009년 신뢰할 수 있는 시스템 네트워크에 관한 IEEE/IFIP 국제 컨퍼런스. 439-448. https://doi.org/10.1109/ DSN.2009.5270307