AI 기반 Web Agent를 활용한 불법 데이터 자동 탐지 및 수집 연구

이수정, 이요셉*, 조용성**, 송경우
연세대학교, *숭실대학교, **한국전자통신연구원

drlisa@yonsei.ac.kr, *dytpq0916@soongsil.ac.kr, **yscho73@etri.re.kr,
kyungwoo.song@yonsei.ac.kr

Web Agent Approach to Automated Monitoring and Collection of Illegal Data

Lee Soojeong, Lee Joseph*, Cho Yongseong**, Song Kyungwoo
Yonsei Univ., *Soongsil Univ., **Electronics and Telecommunication Research Institute (ETRI)

요 약

사이버 범죄는 불법 웹사이트와 다크웹을 중심으로 빠르게 확산되고 있으나, 기존의 데이터 수집 기법은 변화하는 환경에 대응하기 어렵다. 본 연구는 이러한 한계를 극복하기 위해 AI 기반 Web agent 를 활용한 자동 탐지 및 데이터 수집 방법을 제안한다. Web agent 는 웹 환경을 인식하고 계획하며 실행하는 과정을 통해 자율적으로 탐색과 수집을 수행하며, 텍스트, 이미지, 영상 등 멀티모달 데이터를 실시간으로 처리할 수 있다. 제안된 체계는 불법 데이터 탐지, 수집, 불법 유형별 분류의 단계로 구성되어 사이버 범죄 단서 확보를 자동화한다. 이를 통해 수사 기관은 불법 행위의 패턴을 효과적으로 파악하고, 정책적 대응과 집행 전략을 수립하는 데 활용할 수 있다.

Ⅰ. 서 론

사이버 범죄의 종류가 날로 다양해지고, 범죄 발생 건수 또한 급격히 증가함에 따라 사이버 범죄에 대한 단서 수집을 사람이 수동적으로 수행하기에 많은 인력을 필요로 하고 있다. 특히 불법 웹사이트와 다크웹 환경은 수사망을 회피하기 위해 구조와 운영 방식을 지속적으로 변화시키고 있어, 기존의 수작업 기반 단서 수집이나 단순 크롤링 방식으로는 효율적인 대응이 어렵다. 이러한 문제를 해결하기 위해 불법 데이터를 자동으로 점검하고, 구조적으로 수집 할 수 있는 체계가 필요하다.

이에 본 연구에서는 변화하는 환경에 동적으로 대응할 수 있는 AI agent 기반 방법론을 제안한다. Web agent[1]는 다크웹을 포함한 웹 환경에서 자율적으로 탐색, 수집, 분석을 수행하는 소프트웨어 에이전트로, 텍스트, 이미지, 영상 등 멀티모달 데이터를 다루면서도 크롤링과 같은 기존 방식보다 빠른 속도로 데이터를 수집하여 실시간으로 사용 가능하다는 장점을 가진다[2]. 이러한 특성을 활용하면 불법 데이터를 효과적으로 탐지하고 수집할 수 있으며, 더 나아가 수집된 데이터를 정리하고 구조화함으로써 사이버 범죄 수사를 위한 단서 제공과 분석 지원이 가능하다. 본 연구는 Web agent 를 기반으로 한 자동 탐지 및 데이터 수집 방법을 활용하며, 수사 기관이 효율적으로 불법 행위의 패턴을 파악하고 대응 전략을 마련할 수 있도록 돕는다.

Ⅱ. 본론

가. Web agent 정의 및 특성

Web agent 는 웹 환경에서 사용자의 명령을 독립적으로 수행하는 지능형 소프트웨어 에이전트로 정의된다. 기존의 단순한 매크로나 스크립트 기반 자동화와 달리, Web agent 는 복잡하고 동적인 웹 페이지 구조를 이해하고, 주어진 목표를 달성하기 위한 합리적인 행동을 스스로 선택하고 실행할 수 있다는 점에서 차별화된다[3]. 예를 들어, 사용자가 “회의 일정을 잡아줘”라고 지시하면, Web agent 는 이메일 애플리케이션을 열고, 수신자 주소를 탐색하며, 메일을 작성하고 전송하는 일련의 과정을 자동으로 처리할 수 있다. 이러한 과정은 기존 챗봇처럼 지속적인 사용자 개입에 의존하지 않고 독립적으로 진행된다는 점에서 의의가 크다.

Web agent 의 주요 특성은 크게 세 가지로 요약할 수 있다. 첫째, 사용자 편의성이다. 반복적이고 시간이 많이 소요되는 온라인 예약, 정보 검색과 같은 작업을 자동화하여 사용자의 부담을 경감한다. 둘째, 복잡한 업무 수행 능력이다. 단순히 단일 명령만 처리하는 것이 아니라, 다단계 상호작용과 장기적 목표 달성이 가능하다는 점에서 기존 매크로 기반 자동화 도구와 구별된다. 셋째, 확장성과 일반화 가능성이다. 특정 웹사이트나 도메인에 종속되지 않고 다양한 웹 환경에서 일반화된 행동을 수행할 수 있어, 보다 광범위한 응용 가능성을 지닌다. 이러한 특성은 Web agent 가 지능형 개인 비서, 자동 데이터 수집, 온라인 협업 지원 등 다양한 분야에서 활용될 수 있는 근거가 되며, 궁극적으로 사용자의 웹 활용 경험을 향상시킨다.

나. Web agent 방법론

Web agent 는 웹 환경에서 명령을 수행하기 위해 “인식 – 계획 및 추론 – 실행”의 세 단계를 따른다[3]. 이 절에서는 각 단계의 방법론적 특징을 기술한다.

먼저, 인식 단계에서 Web agent 는 웹 환경을 이해하기 위해 정보를 수집한다. 가장 흔히 활용되는 방식은 HTML 기반의 DOM(Document Object Model) 분석과 웹페이지의 시각적 스크린샷이다. DOM 은 웹 요소의 구조적 관계를 명확히 제공하지만, 사용자가 실제로 경험하는 시각적 맥락은 충분히 반영하지 못한다. 반대로 스크린샷은 인터페이스의 실제 모습을 반영하나, 세부 구조 정보를 담기에는 한계가 있다. 최근 연구들은 이 두 가지 방식을 융합한 멀티모달 인식 기법을 활용하여 보다 정교한 웹 환경 표현을 구축하고 있으며, 이는 복잡한 사용자 상호작용을 처리하는 데 효과적이다[4].

다음으로, 계획 및 추론 단계에서는 사용자의 목표를 달성하기 위해 필요한 행동을 합리적으로 결정한다. 이 과정에는 작업을 단계적으로 분해하는 작업 분해(task decomposition), 다음 동작을 예측하는 행동 예측(action prediction), 그리고 장기적 맥락을 반영하기 위한 메모리 기반 추론(memory-based reasoning)이 포함된다. 예컨대, 특정 상품을 구매하는 경우, Web agent 는 ‘검색창 클릭 → 키워드 입력 → 결과 확인 → 상품 선택 → 결제 진행’과 같은 일련의 행위를 계획한다.

마지막으로, 실행 단계에서는 계획된 행동을 실제 웹 환경에서 수행한다. 이 과정에서 agent 는 DOM 요소를 탐색하여 클릭, 스크롤 등의 상호작용을 재현하거나, 특정 API 와의 연동을 통해 효율적으로 작업을 처리한다. Browser-use 프레임워크[5]는 이러한 실행을 지원하기 위해 Python 기반 API 호출 및 헤드리스 브라우저 제어 기능을 제공하며, 이는 웹 자동화의 대표적인 구현 방식으로 자리 잡아가고 있다.

[Figure 1] Web Agent 를 활용한 불법 데이터 탐색 및 수집 과정

다. Web agent 기반 불법 데이터 탐지 및 수집

앞 절에서 Web Agent 의 구조와 방법론을 고찰하였다. 본 절에서는 이를 토대로 불법 데이터 대응 문제에 적용 가능한 방안을 제시하며, 해당 과정을 데이터 탐지, 수집, 분류의 세 단계로 구분하였다.

첫째, 데이터 탐지 단계에서는 단순 키워드 검색을 넘어, Web Agent 가 프롬프트 기반 BFS 탐색 기법을 활용하여 다크웹을 포함한 웹 환경을 체계적으로 계획하여 추론한다. 예를 들어 ‘스포츠 베팅’과 ‘최신 영화’ 등 불법 콘텐츠 관련 키워드를 기반으로, BFS 를 사용하여 한 페이지 내 하위 링크를 최대 depth 2 까지 구조적으로 탐색하고, 추가 콘텐츠를 위해 다음 페이지 이동도 자동으로 처리한다. 이를 통해 단일 URL 에 국한되지 않고, 웹 전반의 잠재적 불법 데이터 확산 경로를 식별할 수 있다는 점에서 기존 단순 크롤링 기법과 차별성을 가진다.

둘째, 데이터 수집 단계에서는 텍스트, 이미지, 영상 등 다양한 멀티모달 데이터를 확보한다. 본 연구에서는 프롬프트에 결과 예시 형식을 직접 제공하여, URL 단위로 수집된 데이터를 마크다운 형식으로 구조화하였다. 나아가 다운로드 자동화 모듈을 추가하여 원본 파일(이미지, 영상 등)을 직접 저장할 수 있도록 구현하였다. 이를 통해 단순한 텍스트 크롤링을 넘어, 실제 수사기관에서 활용 가능한 원본 증거 데이터 확보가 가능해진다.

셋째, 데이터 분류 단계에서는 수집된 데이터를 불법 행위의 유형(불법 링크 공유, 불법 영상물 유통, 불법 도박 거래 등)과 불법 여부에 따라 정밀하게 분류한다. 단순 식별을 넘어, 정책적 대응 및 집행 전략 수립에 필요한 실질적 근거 데이터를 제공한다는 점에서, 본 연구는 기존 연구 대비 실무적 활용 가능성을 높인다.

Ⅲ. 결론

본 연구는 기존의 수작업 기반 데이터 수집과 단순 크롤링 기법이 가지는 한계를 극복하기 위해, AI 기반 Web agent 를 활용한 자동 탐지 및 데이터 수집 체계를 제안하였다. 제안된 방법은 웹 환경을 인식, 계획, 실행하는 과정을 통해 멀티모달 데이터를 실시간으로 확보하고, 이를 불법 데이터 탐지와 분류에 적용할 수 있음을 보였다. 이러한 체계는 변화하는 불법 웹사이트 및 다크웹 환경에서 안정적으로 동작할 수 있어, 수사기관이 효율적으로 단서를 확보하고 분석할 수 있는 기반을 마련한다는 점에서 의의가 크다.

ACKNOWLEDGEMENT

이 논문은 2025 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행되었으며 (No. RS-2023-00224740, 디지털성범죄 피해 예방을 위한 불법촬영물 이미지 유포 차단 및 추적 기술 개발), 또한 2025 년도 정부(경찰청)*의 재원으로 과학치안진흥센터 사이버범죄 수사단서 통합분석 및 추론시스템 개발 사업의 지원을 받아 수행된 연구임(RS-2025-02218280).

참 고 문 헌

[1] Gur, Izzeddin, et al. "A real-world webagent with planning, long context understanding, and program synthesis." arXiv preprint arXiv:2307.12856 (2023).

[2] Huang, Wenhao, et al. "AutoScraper: A progressive understanding web agent for web scraper generation." arXiv preprint arXiv:2404.12753 (2024).

[3] Ning, Liangbo, et al. "A survey of webagents: Towards next-generation ai agents for web automation with large foundation models." Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2. 2025.

[4] Deng, Xiang, et al. "Mind2web: Towards a generalist agent for the web." Advances in Neural Information Processing Systems 36 (2023): 28091-28114.

[5] Browser-use Documentation, 2025, (https://docs.browser-use.com/introduction)