

논문 2024-3-5 <http://dx.doi.org/10.29056/jsav.2024.09.05>

광고사 식별 코드를 활용한 저작권 침해 대체 사이트 식별 방안 연구

오상원*, 박명찬*, 김영모**†

Utilizing Advertiser Unique Identification Numbers for Illegal Copyright Infringement Site Feature Information Collection Strategy

Sang-Won Oh*, Myung-Chan Park*, Young-Mo Kim**†

요 약

한국 콘텐츠 산업의 지속적인 성장과 비례하여 불법 저작물 유통 또한 꾸준히 증가하고 있다. 저작권 침해 사이트는 해외 클라우드 서비스를 이용하고, 주기적으로 도메인을 변경하여 법적 조치를 회피하고 있다. 이를 해결하기 위해 정보는 보다 신속하고 효과적인 차단 방법을 모색하고 있다. 본 논문은 저작권 침해 사이트의 광고배너에 포함된 광고사 식별 코드를 활용하여 이러한 사이트의 특징을 정의하고, 이를 바탕으로 도메인 변경을 통해 운영되는 대체 사이트 식별 방안을 제한한다. 51개의 저작권 침해 사이트를 6개월간 분석한 결과, 99.3%의 광고 배너에서 동일한 광고사 식별 코드를 확인할 수 있었다. 이를 통해 도메인이 변경된 대체 사이트를 효과적으로 식별하고, 신속히 차단할 수 있는 정당성을 확보할 수 있었다. 광고배너 내 식별 코드를 기반으로 저작권 침해 사이트를 식별하는 방법론을 제시하여, 주기적인 도메인 변경에도 불구하고 저작권 침해 사이트를 효과적으로 차단할 수 있는 가능성을 입증하였다. 이는 향후 저작권 보호 및 불법 저작물 유통 방지에 있어 중요한 기여를 할 것으로 기대된다.

Abstract

The steady growth of the Korean content industry has been accompanied by a rise in the distribution of illegal copyrighted materials. Copyright-infringing sites often use overseas cloud services and frequently change their domains to evade legal action. To address this, more rapid and effective blocking methods are needed. This paper proposes identifying these sites by utilizing advertiser identification codes found in ad banners, providing a method to detect alternative sites that operate through domain changes. An analysis of 51 infringing sites over six months revealed that 99.3% of ad banners contained identical advertiser codes. This validates the effectiveness of identifying and quickly blocking sites with changed domains. The proposed method demonstrates the potential to effectively block copyright-infringing sites despite frequent domain changes, contributing significantly to copyright protection and the prevention of illegal content distribution.

한글키워드 : 광고사 코드, 저작권 침해 사이트, 특징 정보, 대체 사이트, 웹툰, 스트리밍

keywords : AD code, Infringement sites, Feature information, Alternative sites, Webtoon, Streaming

* 에이치엠컴퍼니(주)

접수일자: 2024.09.13. 심사완료: 2024.09.16.

** 숭실대학교 컴퓨터학과

게재확정: 2024.09.20.

† 교신저자: 김영모(email: ymkim828@ssu.ac.kr)

1. 서론

한국 콘텐츠 산업은 최근 몇 년간 눈부신 성장을 이루고 있다. 한국콘텐츠진흥원의 보고서에 따르면, 2023년 상반기 한국 콘텐츠 산업의 매출액은 전년 대비 2.5% 증가하였고, 수출액 또한 1.3% 상승하였다[1]. 그러나 이러한 성장에도 불구하고, 불법 저작물 유통은 꾸준히 증가하고 있다. 경찰대학 치안정책연구소의 '치안전망 2024' 보고서에 따르면, 2023년 9월까지 사이버 저작권 침해 범죄가 전년 대비 116.4% 증가한 5,078건이 발생했다[2].

특히 저작권 침해 사이트는 국내법의 적용을 회피하기 위해 해외 클라우드 서비스를 적극적으로 활용하고 있다. 해외 클라우드 서비스는 국내의 법적 규제를 통한 적용이 어려워 불법 사이트들이 이를 이용하여 지속적으로 서비스를 제공하고 있다. 이에 대한 대응방안으로 대한민국의 저작권법 제133조의3에 따른 한국저작권보호원의 '복제·전송자에 대한 경고', '삭제 또는 전송 중단', '복제·전송자의 계정 정지' 등의 시정 조치만으로는 이러한 불법 활동을 근절하는 데 한계가 있어 불법 콘텐츠로 인한 피해를 최소화할 수 있는 효과적인 대응 방안으로 국내 유입을 차단하는 조치를 시행하고 있다[1].

기업에서는 카카오엔터테인먼트의 P.CoK 팀이 불법 저작권 침해 사이트에 대해 집중 단속과 접속 차단 업무를 수행하고 있으며, 이를 통해 2016년부터 2023년까지 약 2,400만 건의 불법 콘텐츠를 삭제하는 성과를 거두었다[3]. 또한, 문화체육관광부, 방송통신위원회, 경찰청 등 관계 기관이 합동으로 침해 대응 특별 전담팀을 구성하여 집중 단속을 시행하고 주요 불법 사이트를 폐쇄하며 운영자를 사법 처리하는 등 의미 있는 성과를 거두었다. 그러나 특정 사이트를 폐쇄하면 다른 사이트로 사용자들이 이동하고 이름을 변경

한 사이트가 새로 생기는 등 대체사이트를 통한 서비스가 지속되고 있다[4].

최근 대부분의 저작권 침해 사이트는 주기적으로(주 1회 이상) 도메인을 변경하거나 외형을 조금씩 바꾸는 “대체 사이트”를 운영한다. ‘대체 사이트’는 차단을 회피할 목적으로 기존 사이트와 동일한 형태로 URL 주소만 일부 변경하여 서비스하는 것을 말한다.

2024년 2월, 방송통신심의위원회(이하 방심위)는 K-콘텐츠를 불법 유통하는 6천여 개의 사이트를 차단했으며, 대부분은 기존 차단된 사이트가 URL을 변경해 운영하는 ‘대체 사이트’였다. 방심위는 저작권 침해 정보를 적발하기 위해 방송 사업자, OTT, 웹툰, 음원 플랫폼 등 33개 권리와 협력하고 있으며, 대체 사이트는 별도 확인 절차 없이 즉시 차단하고 있다[5]. 그러나 절차 간소화로 행정력 남용과 검열 우려가 제기되고 있다. 방심위의 접속 차단 결정은 방송통신위원회 설치법과 시행령에 근거한 독립적 절차로, 정부의 임의 개입은 없다. 차단 절차의 신속성은 저작권 침해 대응에 필수적이지만, 검열 우려를 해결할 새로운 접근이 필요하다. 본 논문은 방심위의 차단 시스템에 광고 배너를 통한 대체 사이트 식별 방안을 제안한다. 광고 배너의 식별 코드를 분석해 저작권 침해 사이트를 특정하는 방법으로, 이를 통해 방심위는 절차의 간소화를 유지하면서도 대체 사이트를 더 정확히 식별할 수 있다. 이 접근은 행정력 남용 및 검열 논란을 줄이고 법적 투명성을 높이는 데 기여할 수 있다.

본 논문의 구성은 제2장에서 저작권 침해 관련 연구를 소개하며, 제3장에서 광고 식별 코드를 활용한 대체 사이트 식별 방안을 제안한다. 제4장에서는 제안된 방안을 기반으로 검증을 수행하고, 제5장 결론에서는 연구의 유효성과 방심위의 차단 시스템 개선에 미치는 영향을 논의하며 마무리한다.

2. 관련 연구

저작권 침해 사이트와 관련된 연구는 주로 불법 콘텐츠 유통 방식과 그 대응 방안에 초점을 맞추고 있다. 이 장에서는 불법 저작권 침해 사이트 분석, 광고사 추적을 통한 연구, 불법 침해 사이트 광고배너 내 특징정보에 대한 기존연구를 살펴본다.

2.1 저작권 침해 사이트 분석

저작권 침해 사이트는 일반적으로 메뉴, 광고, 카테고리 등으로 구성되며, 웹 페이지 상단에 로고와 메뉴바, 검색창이 위치하고, 가장 눈에 띄는 곳에 광고배너가 자리 잡는다(그림 1 참고).

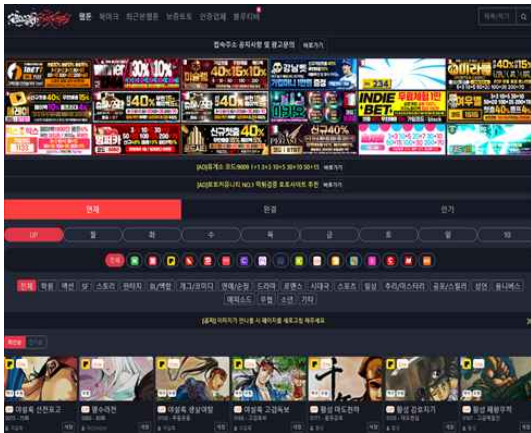


그림 1. 저작권 침해 사이트 예시
Fig. 1. Example of Copyright-Infringing Sites

메뉴바 및 로고는 사이트 상단에 위치하여 사용자가 쉽게 탐색할 수 있도록 도와주고, 로고는 특정 사이트를 상징하는 것으로 이미지 또는 텍스트 형태로 구성되어 있으며, 독자적인 모양을 가지고 있어 각 사이트를 구분할 수 있다. 검색창은 사용자가 원하는 불법 콘텐츠 검색을 지원하고, 광고배너는 저작권 침해 사이트의 주 수익원으로 광고배너를 통해 광고사를 홍보하고 수익

을 발생시키는 핵심 구성요소이다.

해외의 경우 저작권 침해사이트 내 광고 배너가 존재하지 않는 것이 다수이며, 콘텐츠 이용시 무작위로 광고 사이트로 리다이렉트 되어야 콘텐츠를 볼 수 있었다. 이러한 방식은 사용자에게 불편을 주면서도 광고 노출을 극대화하기 위한 수단으로 사용된다. 한편, 중국의 일부 저작권 침해 사이트(예: 91MJW.org)는 국내 사이트와 유사하게 광고 배너에 특정 혜택 코드를 제공하여 사용자 참여를 유도하기도 한다. 이처럼 저작권 침해 사이트들은 수익을 극대화하기 위해 다양한 광고 배치 방식을 활용하고 있다.

2.2 저작권 침해 사이트 식별 방안의 선행 사례

저작권 침해 사이트를 식별하고 차단하는 다양한 방안이 연구되어 왔으며, 그 중 도메인 변경 패턴 분석, IP 추적, 광고 배너 추적 등의 방법이 주목받고 있다.

HTML 태그 간 유사도를 분석하는 LCS 유사도 측정 알고리즘을 활용하여, 저작권 침해 사이트의 도메인 변경 패턴을 파악하고 이를 기반으로 사이트를 차단하는 방안을 제시하였다[6]. 이 방법은 도메인 주소를 빈번하게 변경하는 사이트를 빠르게 식별할 수 있다는 장점이 있다. 그러나 사이트가 HTML 구조를 변형하거나, 동적 콘텐츠로 구성을 바꾸는 경우 유사도 측정의 정확도가 떨어질 수 있다는 한계가 있다.

또한, Censys, Shodan, Criminal IP 등의 사이버 위협 정보 검색 서비스를 이용하여 저작권 침해 사이트의 실제 IP 주소를 추적하는 기법을 연구하였다[7]. 이를 통해 단순히 도메인 차단에 그치지 않고, 실 운영 서버를 대상으로 직접적인 대응이 가능해졌다. 이러한 기술적 접근은 불법 사이트의 실체를 더욱 명확히 파악할 수 있는 장점이 있으나, IP 우회 및 프록시 사용과 같은 방어 기법에 취약한 면이 존재한다.

2.3 광고사를 이용한 연구

저작권 침해 사이트의 주요 수익원인 광고배너를 추적하는 연구가 다수 진행되었다. 신명섭 등[8]은 불법 저작권 침해 사이트의 자금 추적 기반 방식을 제안하여 광고를 차단하는 방안을 연구하였으나, 국내 환경 적용에 한계가 있었다. 송봉규[9]는 광고배너를 중심으로 저작권 침해 사이트의 연결 실태와 대응 방안을 구글 웹 검색을 통해 조사하였으나, 모든 사이트를 분석하는데 어려움이 있었다. 강하연 등[10]은 유해 사이트 배너 광고 추적을 통한 광고주 분석을 통해 최종 광고주를 차단하는 연구를 제안하였으나, 대부분의 불법 광고는 저작권 침해 사이트와 동일한 형태로 운영되어 추적이 어려운 점이 있었다. 기존 연구에서는 광고주를 식별하여 차단함으로써 불법 저작권침해 사이트의 수익원을 차단하는 연구를 중심으로 진행되었으나, 광고사 또한 불법 저작권 침해 사이트와 유사한 형태로 운영되어 식별 및 차단에 한계가 있었다.

2.4 저작권 침해사이트에 적용된 광고배너 식별 정보

합법적인 웹툰 및 영화 유통 사이트는 구독 및 유료 결제를 통해 수익을 창출한다. 그러나 저작권 침해 사이트는 주로 광고를 게재하여 수익을 창출한다. 광고주는 사이트 이용자가 많고 활발한 사이트에 광고를 게재하기 때문에, 저작권 침해 사이트는 불법 콘텐츠를 이용하여 더 많은 이용자를 유치하고, 이를 기반으로 광고 수익을 창출한다[11].

저작권 침해 사이트에서 배너 광고는 주로 직접 계약 방식으로 이루어진다. 이는 광고주가 매체(웹 사이트)를 소유한 사업자와 직접 계약하여, 사이트의 특정 부분에 광고주가 제공한 광고 표현물을 담아 사용자에게 노출하는 방식이다. 이러한 배너 광고는 검색 광고와 함께 가장 많이

사용되는 광고 형태다[12, 13].

저작권 침해 사이트에 게재되는 광고배너에는 광고사를 식별할 수 있는 식별코드가 포함되어 있다. 이는 저작권 침해 사이트를 소유한 사업자에게 부여되는 식별 수단으로, 광고사가 광고를 의뢰하면 클라이언트는 계약을 통해 광고사로부터 광고배너와 식별코드를 받는다(그림 2 참고).

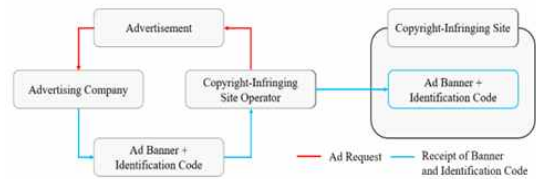


그림 2. 광고 배너 흐름
Fig. 2. The Flow of Ad Banners

저작권 침해 사이트를 이용하는 이용자가 광고배너를 통해 신규 가입하는 경우, 식별코드를 이용하여 추천인이 된다. 즉, 식별코드는 광고 수익을 계산하는 주요 수단이 된다. 광고배너에는 식별코드가 포함된 경우와 포함되지 않는 경우로 표 1과 같다.

표 1. 광고 배너 식별 코드 예시
Table 1. Example of AD Identification code

Includes Identification Code	Does Not Include Identification Code

식별코드가 없는 경우에는 HTML 코드에 그림 3과 같이 'join=AZZT' 값이 존재하고, 이 값은 광고 사이트 접속 후 가입 시 추천인 코드로 활용된다.

광고배너에 포함된 식별코드는 광고주로부터 부여받은 고유 식별 값으로 사이트마다 상이하여 광고주가 대상 업체를 식별하는 용도로 활용된다.



그림 3. 식별 코드가 없는 광고 이미지
Fig. 3. Advertising image without identification code

3. 광고배너의 식별 정보를 이용한 대체 사이트 식별 방안

3.1 광고배너에 식별코드가 있는 경우

관련연구에서 광고배너에 포함된 식별코드가 광고사에서 부여한 식별 값을 확인할 수 있었다. 이는 특정 사이트의 광고배너 식별코드를 분석하여 저작권 침해 사이트를 효과적으로 식별할 수 있는 중요한 수단이 될 수 있다. 이를 기반으로 대체 사이트를 명확히 특정할 수 있어, 동일한 사이트임을 증명할 수 있다. 이는 즉시 차단 of 정당성을 확보할 수 있게 한다. 따라서 본 장에서는 저작권 침해 사이트의 광고배너 속 식별 코드를 활용하여 대체 사이트를 식별하는 방안을 제시한다.

불법침해사이트는 광고배너에 식별코드가 있는 경우와 코드가 없는 경우가 있다. 식별코드가 없는 광고배너의 경우, 일부에서 HTML 코드에 식별코드가 있는 경우도 있으며, 식별코드가 없는 경우도 존재한다. 이는 광고사의 운영 정책에 따라 상이하다. 다음 그림 4는 본 연구에서 제안 하는 대체사이트 식별 방안이다.

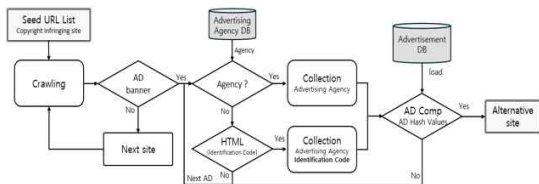


그림 4. 대체 사이트 식별 방안
Fig. 4. Alternative Site Identification Methods

3.2 광고배너에 식별코드가 있는 경우

광고배너에 식별코드가 있는 경우, 광고배너의 식별코드를 획득하기 위해 OCR(광학 문자 인식)을 이용하는 방법이 제안되었으나, 이미지 색이 복잡하거나 애니메이션 효과가 있는 경우 인식이 매우 낮다. 또한, 문자가 많은 경우 식별코드를 인식하는 데 어려움이 있다. 따라서, OCR을 이용하지 않고 식별코드를 식별할 수 있는 방안 연구가 필요하다.

본 연구에서는 광고배너의 식별코드를 획득하기 위한 방안으로 광고배너 이미지 자체를 특정하고자 한다. 즉, 특정 광고사에서 제공하는 광고배너는 고객사마다 식별코드가 다른 고유의 광고배너를 사용하고 있어, 해당 광고배너의 해시 값을 이용하여 고객사를 식별하고자 한다.

따라서, 식별코드가 있는 광고사를 특정할 수 있다면 광고배너의 해시 값을 이용해 기존 사이트와 대체 사이트가 동일한 식별코드를 사용하는 지 검증할 수 있다. 그림 5는 식별코드가 있는 광고사를 특정하기 위한 절차로 다음과 같다.

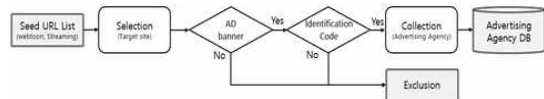


그림 5. 광고사 DB 생성
Fig. 5. Creation of an Advertising Agency Database

- 1) 광고사 DB 생성: 저작권 침해 사이트에서 광고배너에 식별코드가 존재하는 광고사를 선정하여 광고사 정보 등록(그림 5 참조)
- 2) 광고 DB 생성: 광고사 DB에 등록된 광고사를 기준으로 광고배너 수집(그림 4)
- 3) 대체 사이트 식별: 광고 DB에 수집된 광고사별 해시값과 비교

먼저, 수동으로 광고배너에 식별번호가 있는 광고사를 식별하고, 광고사 DB에 광고사 메타정

보를 수집하여 등록한다. 수집 정보는 표 2와 같이 광고배너 해시값, 광고배너 파일명, 광고사 주소이다.

표 2. 식별 코드가 포함된 이미지 수집
Table 2. Collection of images with identification codes

Collection target	
Image Hash (SHA256)	886434A425E86A30EC4982C540F4213B148794AAC5021D20A8ED7EC5AA456511
File Name	202303300153452954519.jpg
Advertising company URL	http://1bet1.vip/
Advertising Domain	1bet1

저작권 침해 사이트에서 생성된 광고 DB는 “대체 사이트”로 의심되는 사이트에서 수집된 광고배너의 해시 값과 비교하여 동일한 해시값이 존재하면 해당 사이트를 대체 사이트로 인식한다.

3.3. 광고배너가 아닌 HTML에 식별코드가 있는 경우

광고배너에 식별코드가 존재하지 않고, HTML에 식별코드가 존재하는 경우는 표 3 과 같이 광고 배너에는 광고문구만 존재하고, 식별코드는 표 3과 같이 HTML 코드에서 확인할 수 있는 유형이다.

표 3. 식별 코드가 없는 수집 대상의 예시와 수집된 데이터
Table 3. Example of Collected Targets without Identification Codes and the Obtained Data

Collection target (Advertising Banner and HTML Code)	
Image Hash (SHA256)	4B9D32A71F08E2D36C82FB80FBC1BAB6108B4406A0DA99223978A328418E7763
File Name	202402191922285803703.gif
Advertising company URL	https://xn--o39a72x5kxyxg.com
Advertising Domain	xn--o39a725kxyxg

표 3의 수집 대상의 HTML 코드를 확인하면 ‘<a>’ 태그에 “join=AZZT”를 확인 할 수 있으며, 이 경우 해당 광고사로 이동하면 추천인에 “AZZT”이 자동으로 입력되는 것을 확인 할 수 있다. 본 연구에서는 HTML에 식별코드가 있는 경우, 광고사와 관계없이 수집 대상으로 간주하여 광고배너 및 식별코드를 획득하여 광고DB에 저장하고, 이후 그림 4와 같이 대체 사이트 점검에서 동일한 해시값이 발견되면 식별코드를 확인하여 식별코드가 동일한 경우 대체 사이트로 판단한다.

4. 광고배너를 이용한 대체 사이트 식별 방안 실험

4.1. 실험검증 개요

광고배너에 포함된 광고사 식별코드를 활용하여 대체 사이트를 효과적으로 식별하는 방안을 실험을 통해 검증하였다.

4.1.1. 저작권 침해 사이트 선정

- 1) 불법 저작권 침해 사이트 선정: 주소모음 사이트를 대상으로 저작권을 침해하는 웹툰과 스트리밍 사이트 등 84개 사이트를 수집하고, 수집된 사이트에 대해 중복제거 후 광고배너가 존재하는 51개 사이트 최종 선정
- 2) 광고사 DB 생성: 최종 선정된 51개 사이트의

광고배너를 분석하여 식별코드가 존재하는 광고사를 216개를 수동으로 식별하여 광고사 DB 생성

- 3) 침해사이트 광고 배너 DB 생성: 저작권 침해 사이트의 광고배너를 대상으로 광고사 DB에 등록된 광고사 광고배너가 있는 경우 해시 값을 계산하여 광고 DB에 저장하고, 그 외 HTML에 식별코드가 있는 경우에도 이미지 해시 값을 계산하여 광고 DB에 저장(예: A광고사 식별번호=1234, B광고사=TRTR 등)
- 4) 대체 사이트 식별: 대체 사이트로 의심되는 사이트에 대해서도 식별을 위해 광고사 DB에 저장된 광고사가 있는 경우 해시 값을 계산하여 수집하고, 그외 HTML에 식별코드가 있는 경우에도 이미지 해시 값을 계산하여 광고 DB에 등록된 해시 값과 비교(예: B광고사=TRTR, C광고사=HHHH 등), 동일한 식별코드가 발견되면 대체사이트로 판단(예: B광고사=TRTR)

4.1.2. 광고사 DB의 적절성

저작권 침해 사이트에 존재하는 광고배너를 수집하기 위해 주소 모음 사이트를 대상으로 웹툰 및 스트리밍 사이트 등 총 84개 사이트를 수집하였고, 중복제거 및 광고배너가 존재하는 51개를 대상으로 식별코드를 수집 및 분석하였다. 광고사 216개 광고배너 614개가 수집하였으며, 이를 분석하여 광고사 Top 10을 표 4와 같이 선정하였다.

표 4. 상위 10개 광고사
Table 4. Top 10 Advertising companies

No	Advertising Company (Top 10)	Copyright infringement sites							Identification Code Verification	
		1	2	3	49	50	51			
1	one*one	O	O	O	O	O	O	49	96.4%
2	wj*	O	O	O	O	X	O	47	92.9%
3	win*	O	O	O	O	O	O	38	82.1%
4	sun*ti	X	X	X	O	X	X	28	39.3%
5	si*bat	O	O	X	X	O	X	20	32.1%
6	sk*all	O	O	X	X	O	O	15	28.6%
7	cal*	O	O	O	X	X	X	14	28.6%
8	Hz*	O	O	O	O	X	X	14	28.6%
9	we*ch	X	X	O	X	O	X	12	21.4%
10	aa*ba	O	O	X	X	X	O	7	25.0%

분석 결과, 원*원 광고는 51개 사이트 중 대부분에 49개(96.4%) 걸쳐 광고를 게재하고 있었다. 이를 통해 광고사를 중심으로 한 식별코드 수집이 효과적임을 확인하였다. 광고사 DB를 기반으로 대체 사이트를 점검한 결과 높은 일치율을 보였으며, 이를 통해 동일한 사이트임을 명확히 증명할 수 있었다.

4.2. 대체 사이트 동일성 실험

기존 차단된 저작권 침해 사이트의 URL을 수정하여 서비스하는 “대체 사이트”의 동일성 검증을 위해 웹툰 및 스트리밍 관련 저작권 침해 사이트 51개를 대상으로 URL을 등록하였다. 저작권 침해 사이트는 대체 사이트 변경 주기가 각 사이트별로 상이하어 2023년 6월부터 2023년 12월까지 6개월간 주 2회, 총 48회 점검을 수행하였다. 대체 사이트 점검 절차는 다음과 같다.

- 1) 초기 URL 등록: 최초 입력된 URL에 대해 ‘URL State, Current URL, Before URL, Advertisement Image URL, Advertising Agency, Advertisement HashValue’ 정보를 ‘광고 DB’에 등록
- 2) 주기적 점검: 주기적으로 대체 사이트 유무를 확인하고, 대체 사이트로 의심되는 사이트가 발견되면 ‘광고사 DB’에 등록된 광고사 대상으로 검색
- 3) 해시값 비교: 광고사 배너가 발견되면 해시값을 계산하여 기존 ‘광고 DB’에 저장된 결과와 비교하고, 같으면 대체 사이트로 판단
- 4) 광고 DB 갱신: 최종 발견된 대체 사이트 정보를 기존 ‘광고 DB’ 업데이트

4.3. 대체 사이트 동일성 실험

6개월간 48회 점검한 결과, 총 51개 사이트 중 중간에 차단되거나 폐쇄된 사이트 22개를 제외한 29개 사이트에서 대체 사이트가 발견되었다. 초

기 수집된 641개의 광고배너는 표 5와 같이 최종 382개로 40% 감소하였으며, 이중 최종 비교 대상 광고배너는 298개가 되었다. 그 중 296개(99.3%)는 일치하였다.

일치하지 않은 광고배너 2개는 1개 사이트에서 확인되었으며, 광고사 식별코드가 'BT' → 'BTBT'로 수정되었다. 확인 결과 광고 내용이 기존 내용과 다르게 새로운 내용으로 변경된 것을 확인할 수 있었다. 즉, 식별코드가 있는 광고배너의 경우 광고 내용이 바뀌지 않는 한 변경되지 않았고, 광고사는 수시로 변경되었지만, 한 번에 모두 변경되지는 않았다. 즉, 주기적인 점검을 통해 광고배너 정보를 갱신하면 사이트 식별이 가능한 것으로 분석되었다.

표 5. 수집된 광고 배너의 비교 및 인식률
Table 5. Comparison and Recognition Rate of Collected Advertisement Banners

Category	Result
Initial Registered Sites	51 sites
Final Target Sites	29 sites(57%)
Blocked & Closed Sites	22 sites(43%)
Initially Collected Ad Banners	641 banners
Finally Collected Ad Banners	382 banners
Ad Banners for Comparison	298 banners
Matching Ad Banners	296 banners
Ad Banner Matching Rate	99.3%

5. 결론

본 연구에서는 저작권 침해 사이트의 광고배너 속 식별코드를 활용하여 대체 사이트를 식별하는 방안을 제시하였다. 저작권 침해 사이트는

도메인 변경과 같은 다양한 방법으로 차단을 회피하고 있으며, 이러한 사이트들을 효과적으로 식별하고 차단하기 위해서는 광고배너 내 식별코드의 활용이 중요하다.

저작권 침해 사이트의 광고배너를 수집하고 분석한 결과, 주요 광고사들의 식별코드를 통해 대체 사이트를 특정할 수 있었다. 특히, 대형 광고사가 차지하는 비중이 높아 이들 광고사를 중심으로 한 식별코드 수집이 효과적이었다. 광고사 DB를 기반으로 대체 사이트를 점검한 결과, 높은 일치율을 보였으며, 이를 통해 동일한 사이트임을 명확히 증명할 수 있었다.

광고배너의 식별코드를 활용한 대체 사이트 식별 방법은 즉시 차단의 정당성을 확보할 수 있게 하여, 국내로 유입되는 저작권 침해 사이트를 효과적으로 차단하고 저작권 보호에 기여할 수 있다. 이러한 방법은 차단 회피를 위한 대체 사이트의 빈번한 도메인 변경에도 불구하고, 지속적인 감시와 차단이 가능하게 한다.

본 연구는 저작권 보호를 위한 정책 수립과 기술적 대응 방안에 중요한 기초 자료를 제공하며, 향후 더 많은 데이터와 다양한 유형의 사이트를 대상으로 한 추가 연구가 필요하다.

This research project supported by Ministry of Culture, Sport and Tourism(MCST) and Korea Creative Content agency(KOCCA) in 2022(Project Name: "Cloud-based remote storage evidence collection and copyright Infringement data DB development", Project Number: R2022020109, Contribution Rate: 100%)

참고 문헌

- [1] J. Park, "Annual report on copyright protection 2024", Korea Copyright Protection Agency, Report No. KP 2023-03, 2024. Available: https://www.kcopa.or.kr/lay1/bbs/S1T283C290/F/25/view.do?article_seq=5216&cpage=1&rows=9&condition=&keyword=
- [2] D. Kwon, "Public safety outlook 2024: Cyber crime projections in the field of criminal investigation", Police University, Institute of Public Safety Policy Research, Dec. 2023. Available: https://www.police.go.kr/user/bbs/BD_selectBbs.do?q_bbsCode=1002&q_bbscttSn=20231229153717729
- [3] Kakao Entertainment, "Illegal Distribution Response White Paper Vol. 3", 2023. Available: https://bit.ly/pcok_wp3
- [4] S. H. Park, "A Study on Characteristics and Problems of Illegal Distribution of Comics and Webtoons", *Cartoon and Animation Studies*, pp. 341-374, 2019. DOI: 10.7230/KOSCAS.2019.57.341.
- [5] Korea Communications Standards Commission, "Detection and Blocking of Over 6,000 Illegal 'Alternative Sites' for Distributing K-Content", Press release, 2024. Available: <https://www.kocsc.or.kr/cop/bbs/selectBoardArticle.do>
- [6] J. Y. Jang, K. D. Lim, and S. J. Lee, "Collection System of Harmful Sites Using HTML and URL Features", *Journal of Digital Forensics*, vol. 16, no. 1, pp. 54-63, 2022. DOI: 10.22798/KDFS.2022.16.1.54.
- [7] Y. S. Hwang, J. H. Han, and S. J. Lee, "Real IP Address Tracking Techniques for Illegal Sites Using Cyber Threat Intelligence Search Services", *Journal of Digital Forensics*, vol. 16, no. 2, pp. 116-125, 2022. DOI: 10.22798/KDFS.2022.16.2.116.
- [8] M. S. Shin, M. R. Yong, and Y. J. Lee, "Study on Preventing Copyrights Infringement through Blocking Advertisements of Illegal Copyrighted Websites", *Journal of The Korea Contents Association*, vol. 20, no. 7, pp. 331-341, 2020. DOI: 10.5392/JKCA.2020.20.07.331.
- [9] B. G. Song, "The Status and Countermeasures of Connecting Illegal and Harmful Websites: Focusing on Banner Advertising", *The Korean Association of Police Science Review*, vol. 22, no. 1, pp. 121-149, 2020.
- [10] H. Kang, Y. Choi, and S. Lee, "Analysis of advertisers by tracking banner ads on piracy websites", *Journal of Digital Forensics*, vol. 15, no. 3, pp. 15-26, 2021. DOI: 10.22798/kdfs.2021.15.3.15.
- [11] E. S. Choi, Y. Kim, and M. C. Park, "Research on Methods of Feature Information Gathering for Identifying Illegal Copyright Infringement Sites", *Journal of Software Assessment and Valuation*, vol. 19, no. 3, pp. 1-10, 2023. DOI: 10.29056/jsav.2023.09.01.
- [12] E. H. Park, and J. H. Kim, "Advertising Psychology", Communicationbooks Co., 2009. ISBN: 8964060008.
- [13] W. H. Kim, H. J. Ryu, and W. B. Ji, "The Impact of Attention and Content Involvement toward Network Advertisement on Perceived Intrusiveness and Ad Attitude", *Journal of Digital Contents Society*, vol. 20, no. 2, pp. 377-385, 2019. DOI: <http://dx.doi.org/10.9728/dcs.2019.20.2.377>.

— 저 자 소 개 —



오상원(Sang-Won Oh)

2024.8 성공회대학교 정보통신공학과 졸업
2024.4-현재 : 에이치엠컴퍼니(주) 매니저
<주관심분야> 디지털 포렌식



박명찬(Myung-Chan Park)

2005.8 대전대학교 컴퓨터공학과 박사
2012.11-현재 : 에이치엠컴퍼니(주) 이사
<주관심분야> 저작권 보호 및 이용 활성화



김영모(Young-Mo Kim)

2003.2 대전대학교 컴퓨터공학과 졸업
2005.2 대전대학교 컴퓨터공학과 석사
2011.2 대전대학교 컴퓨터공학과 박사
2012-현재 : 숭실대학교 교수
<주관심분야> 저작권 보호 및 이용 활성화