

2020

한국정보보호학회 하계학술대회

CISC- S'20

Conference on Information Security
and Cryptography - Summer 2020

07.15^{WED}

온라인 컨퍼런스

<http://www.cisc.or.kr/>

주최
주관



한국정보보호학회
Korea Institute of Information Security & Cryptology

후원



국가정보원
National Intelligence Service Korea



과학기술정보통신부



행정안전부



한국인터넷진흥원
KISA KOREA INTERNET & SECURITY AGENCY

ETRI

한국전자통신연구원
Electronics and Telecommunications
Research Institute

NSR

국가보안기술연구소
National Security Research Institute



아이티센
ITCen



적대적 공격에 견고한 Perceptual Ad-Blocking 디자인	김민재 · 김보민 · 허준범 / 381
--	-----------------------

AI보안

DNN 모델 취약점 적대적 예제 생성 기법과 방어기법에 대한 분석	서성관 · 문현준 · 윤주범 / 387
딥러닝 시스템 커버리지 기반 퍼징 프레임워크 최적화 연구	유지현 · 문현준 · 윤주범 / 392
Variant AutoEncoder 기반 모바일 맬웨어 클러스터링 기법	고은별 · 김진성 · 반영훈 · 이정현 / 396
딥러닝을 활용한 CCTV 내 폭력 탐지 시스템 개발	김태옥 · 박현아 · 임다연 · 최은정 / 400
머신러닝 기반 네트워크 트래픽 분석 기법 성능 비교	황종배 · 하재철 / 404
금융데이터를 활용한 연합학습 평가	장진혁 · 최대선 / 409
CNN을 이용한 스마트 폰 부채널 공격	김동현 · 김지은 · 유건우 · 이승대 · 백유진 / 413
K-means Clustering을 이용한 kNN 속도개선에 관한 연구	김도연 · 이용수 · 윤재현 · 이태진 / 417

신원/인증시스템

분산 ID 기반 비대면 데이터 공유 접근 시스템	이정륜 · 윤태연 · 이우기 / 423
차세대 주민등록정보시스템 구축에 관한 연구	이현철 · 김대영 · 김현수 / 427
탈 중앙화된 분산 ID 시스템의 특징 및 한계에 대한 동향 조사	이길희 · 김석현 · 김은수 · 전소원 · 김봉곤 · 조영섭 / 431
열 기반 슬라이싱을 이용한 폐기 가능한 홍채 템플릿 생성 방법	신영아 · 정재열 · 정익래 / 436
안면인식을 통한 USB 파일 복호화 시스템	송경주 · 강혜진 · 백민진 · 김현지 · 서화정 / 440
스마트폰과 스마트워치를 활용한 2차 사용자 인증 방식 제안	심민주 · 엄시우 · 김현지 · 서화정 / 444

암호화폐/양자암호

암호화폐 금전가치 분석 동향	박재훈 · 최승주 · 서화정 / 451
암호화폐 거래 믹싱을 통한 익명화기술 동향조사	장철아 · 이경현 / 455
스마트 컨트랙트의 동적 모니터링을 지원하는 이더리움 가상머신	김지수 · 이수연 · 조은선 / 459
NIST 양자내성암호 공모전 부호 기반 암호 구현 동향	최승주 · 김현지 · 심민주 · 서화정 / 463
양자 컴퓨팅 환경에서 안전한 다중 서명 기법	이영경 · 박종환 · 이광수 · 이동훈 / 467
μ -Hope: 오류 정정 부호를 사용한 RLWE 기반 KEM	이주엽 · 김수리 · 홍석희 / 471

데이터식별/추적

저작권 침해 의심 사이트의 생애주기 단계 자동 탐지 기법	정해선 · 곽진 / 477
링크분석 기반 영상 저작권 침해 스트리밍 사이트 탐지 방안	이희승 · 곽진 / 481
스팸 리뷰 판별을 위한 텍스트 분석 관련 기술 동향	임세진 · 김현지 · 서화정 / 485
음란물 유포 경로 추적을 위한 Tumblr 애플리케이션 메모리 포렌식 분석	박진성 · 공성현 · 이창훈 / 489
안드로이드 어플리케이션의 DRM 기술 현황 분석 :	
카카오톡, 카카오톡	이나연 · 김도연 · 양승욱 · 김희석 / 493
메시지 시그니처 분석을 통한 송신자 식별 방안	김동현 · 곽진 / 497

링크분석 기반 영상 저작권 침해 스트리밍 사이트 탐지 방안

이희승¹,곽진²

¹아주대학교 컴퓨터공학과 정보보호응용및보증연구실

²아주대학교 사이버보안학과

Video Copyright Infringement Streaming Site Detection Method based on Link Analysis

Hee-Seung Lee¹, Jin Kwak²

¹ISAA Lab., Department of Computer Engineering, Ajou University

²Department of Cyber Security, Ajou University

요 약

최근 한류열풍이 불면서, 한국의 음악, 드라마, 영화 등이 세계 각지에서 한국의 콘텐츠에 대한 주목이 높아지고 있다. 하지만 주목이 높아지면서 정당한 방법으로 금전을 지불하고 이용하는 것이 아닌, 스트리밍이나 다운로드 등 불법적인 경로를 통해 이용하는 방식이 많아지고 있다. 이에 따라 본 논문에서는 불법적으로 운영되는 스트리밍 사이트와 스트리밍 콘텐츠의 링크에 대한 특성을 분석하고, 불법적으로 영상 스트리밍을 제공하는 사이트에 대해 탐지하는 방안을 제시한다.

I. 서론

시대가 발전하면서 자신의 생각이나 감정을 여러 가지 방법을 통해 보여줄 수 있는 매체들이 증가하면서 저작권과 저작물에 대한 관심도가 높아지고 있다. 이에 따라 원저작자의 허가나 승인 없이 저작물을 이용하는 경우 저작권 침해로 표현을 하고 있으며, 이러한 저작물이나 저작권 침해에 대해 민감하게 받아들이고 있다.

최근 세계 각지에 한류열풍이 불면서 K-Pop, K-Drama 등 한국의 음악과 드라마, 영화들에 대한 소비가 증가하고 있다. 하지만 한국의 콘텐츠에 대한 소비가 증가함에 따라 불법적인 경로를 통해 콘텐츠를 이용하는 사용자들도 증가하고 있다. 즉, 저작물을

정당한 방법으로 금전을 통해 구매하여 이용하는 것이 아닌, 불법적인 경로로 저작물을 무료로 이용하는 소비가 증가하고 있다[1].

특히 영상 저작물의 경우, 원저작자의 허가 없이 스트리밍 형태로 영상을 무료로 배포하거나, 이를 이용하는 소비자가 크게 늘고 있으며, 2018년 10,505억 원에 달하는 생산 감소가 발생했다[2].

이에 따라 본 논문에서는 원저작자의 허가나 승인 없이 저작물을 사용, 배포하는 영상 스트리밍 사이트에 대해 특성을 분석하고 탐지하는 방안을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서 저작권에 대한 개요와 영상 저작물 침해에 대한 유형을 설명하고, 3장에서 저작권을 침해하는 영상 스트리밍 사이트의 특징 분석 및 탐지하기 위한 제안 사항을 제시한 뒤 4장에서 실험결과에 대해 설명하고 5장에서 결론을 맺는다.

본 연구는 문화체육관광부 및 한국저작권위원회의 2020년도 저작권연구개발사업의 연구결과로 수행되었음(2019-PF-9500).

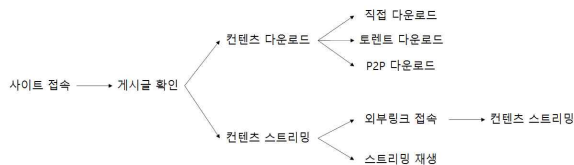
II. 관련 연구

2.1 저작권 개요

저작권위원회에서는 저작권이란 사람의 생각이나 감정을 표현한 결과물에 대하여 그 표현한 사람에게 주는 권리를 뜻하고 있다[3]. 또한 저작권법 제 2조 1항에서는 저작물을 인간의 사상 또는 감정을 표현한 창작물로 정의하고 있으며, 원저작자에게 일정 기간 그 창작물을 독점적으로 사용하게 하고 다른 사람이 무단으로 복제, 공연, 공중 송신, 전시, 배포 등의 행위를 하거나 그 창작물에 대한 원저작자의 인격권을 침해하는 행위를 금지하는 권리로 표현하고 있다[4].

2.2 영상 저작물 침해 유형

인정받는 저작물의 유형이 증가함에 따라 저작권을 침해하는 유형도 증가하고 있다. 영상 저작물의 경우 (Fig. 1)과 같이 다운로드 혹은 스트리밍 형태로 저작권이 침해된다.



(Fig. 1) Infringement Types of Video Literary Work

III. 영상 스트리밍 사이트 특징 분석 및 제언 사항

3.1 링크 특징 분석

본 논문에서는 링크들의 특징을 분석한다. 단순 링크와 딥 링크의 경우 저작권 침해에 해당되지 않지만 임베디드 링크와 프레이밍 링크의 경우 저작권 침해에 해당이 되며, 2심 법원에서는 공중송신권 침해를 용이 하게 해주는 행위로 방조에 해당한다고 판결하고 있다[5].

□ 단순 링크(Simple Link)

웹사이트의 이름과 URL만 게시

□ 딥 링크(Deep Link)

저작물의 간단한 정보 제시와 저작물이 존재하는 세부 페이지로 연결시키거나 어플리케이션 내부의 콘텐츠로 직접 연결

□ 임베디드 링크(Embedded Link)

홈페이지를 열거나 링크를 클릭하면 링크된 자료가 링크를 건 웹사이트의 자료인 것처럼 페이지 이동 없이 실행

□ 프레이밍 링크(Framing Link)

임베디드 링크와 유사하지만 <iframe> 태그 내에서 페이지 이동 없이 실행된다는 차이가 있음

3.2 링크 내에서의 스트리밍 특징

107개의 불법 스트리밍 사이트 샘플 중 접속이 가능하며, 영상 재생이 가능한 17개의 사이트를 대상으로 연구를 진행하였다. 17개 사이트의 스트리밍 페이지를 분석한 결과 크게 3개의 유형으로 나뉜다. 또한 영상의 소스에 직접 접근 시 2개의 유형으로 나뉜다.

3.2.1 스트리밍 제공 유형

나뉜 3개의 유형은 각각 사용되는 것이 아닌, 혼합되어 사용되기도 한다.

□ 외부 링크 접속

스트리밍 재생 시 외부 링크로 연결되는 유형으로 supervid, jetload 등 불법 스트리밍 사이트에서 해당 사이트의 링크를 제공한다.

□ jw player 사용

jw player라는 별도의 플러그인을 사용하는 유형으로 <Table 1>과 같은 형태를 가지고 있다.

<Table 1> Form of jw player

```

video class="jw-video jw-reset"
options src="url-link"
  
```

□ html5_api 사용

html5의 api를 사용하는 유형으로 html5의 특징을 이용하여 별도의 플러그인 없이도 영상 재생을 지원한다.

3.2.2 소스 접근 시 유형

페이지 내의 스트리밍 소스에 직접 접근할 경우 영상을 다운로드하는 경우와 영상이 스트리밍 되는 경우로 나뉜다.

□ 영상 다운로드

소스 접근 시 avi, mp4의 확장자를 갖는 영상 파일들이 다운로드

□ 영상 스트리밍

소스 접근 시 영상파일이 다운로드 되지 않으며, 스트리밍을 제공하는 링크와 blob:url 형태의 링크를 갖는 유형이 있다.

blob:url 링크의 경우 데이터의 송수신을 위한 의사링크로 해당 링크로 직접 접근할 수 없다.

3.3 제안 사항

저작권을 침해하는 콘텐츠에 대해 스트리밍을 제공하는 사이트에 대한 탐지를 위한 것으로, 링크의 특징과 링크 내에서 스트리밍의 특징을 이용하여 제안 사항을 제시한다.

Step 1. 키워드 검색

각 방송사의 홈페이지에는 해당 방송사에서 제작 또는 방영하는 드라마, 예능만이 검색 가능하다. 이러한 특징 때문에 임의의 사이트에 접속하여 각 방송사에서 방영하는 콘텐츠들을 검색했을 때, 검색 결과가 모두 나온다면 저작권 침해 사이트로 의심해볼 수 있다.

각 방송사에서 방영하는 콘텐츠들의 제목을 키워드로 하여 사이트 내에서 검색이 되는지 확인한다.

Step 2. 검색된 키워드 접근 & 소스 분석

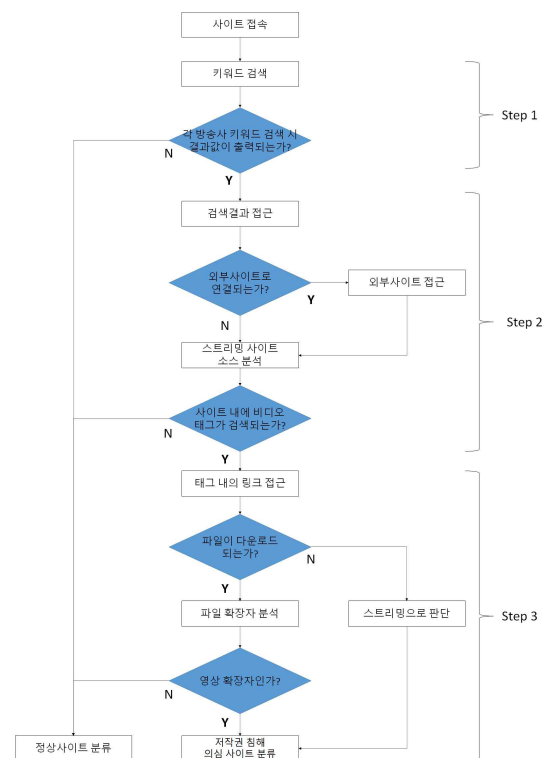
키워드 검색 시 결과가 나온다면 해당 콘텐츠의 링크로 접속하여 페이지 내의 소스를

분석한다. 이때 소스 분석 시에는 임베디드 링크를 위한 <embed>, <object>, <video> 태그와 프레임 링크를 위한 <iframe> 태그 존재 여부를 확인한다.

Step 3. 분석한 소스 기반 영상 유형 파악

분석한 페이지 소스를 기반으로 영상의 소스에 접근 시 파일 다운로드와 스트리밍 유형에 대해 파악한다.

파악한 영상의 유형을 기반으로 다운로드 되는 경우 파일의 확장자를 검사한다. 영상파일 확장자이거나, 파일이 다운로드 되지 않는 경우 스트리밍으로 판단하여 저작권 침해 의심 사이트로 분류한다.



(Fig. 2) Video Copyright Infringement Streaming Site Detection Process

IV. 실험 결과

실험은 17개의 저작권 침해 사이트와 18개의 정상 사이트로 총 35개의 데이터 샘플로 진행한다.

제시한 제안 사항으로 탐지 진행 시 <Table 2>와 같은 결과를 얻을 수 있었으며, 수식을 통해 약 83%의 정확도가 산출되었다.

<Table 2> Result of Detection Experiment

		Condition	
		Positive	Negative
Prediction	Positive	13(TP)	2(FP)
	Negative	4(FN)	16(TN)

$$\text{확도 (Accuracy)}: \frac{TP + TN}{P + N} \quad (1)$$

애니메이션만을 스트리밍하는 사이트의 경우 키워드화한 국내 방송사의 콘텐츠들이 검색되지 않아 저작권 침해 의심사이트로 분류되지 않은 것으로 판단된다. 또한 국내 콘텐츠가 아닌 해외 콘텐츠만 스트리밍하는 경우도 검색이 되지 않아 저작권 침해 의심사이트로 분류되지 않았다. 정상사이트 샘플 중 포털사이트의 경우 모든 검색결과에 대해 검색이 되고, 원본영상보다 짧은 클립영상을 스트리밍하여 저작권 침해 의심사이트로 분류되었다.

V. 결론

본 논문에서는 링크 분석을 이용한 저작권 침해 스트리밍 콘텐츠에 대한 탐지 방법을 제안하였다. 불법으로 스트리밍을 제공하는 사이트들은 타 사이트의 영상을 링크를 이용해 제공하는 임베디드 링크, 프레임 링크 방식을 사용하여 스트리밍을 제공하고 있다. 이러한 스트리밍 사이트의 특징을 이용하여 불법 스트리밍을 탐지할 수 있는 프로세스를 제안하였다.

향후, 불법 스트리밍 사이트들을 더욱 정확하게 탐지할 수 있는 방안에 대한 연구를 진행할 예정이다.

[참고문헌]

- [1] 한국저작권보호원, “저작권 보호기술 기반 사업 전략 로드맵”, Sep. 2017.
- [2] 한국저작권위원회, “불법복제에 따른 연간 경제적 손실 규모”,

http://kosis.kr/statHtml/statHtml.do?orgId=443&tblId=DT_443001_073, Feb. 2020.

- [3] 한국저작권위원회, “저작권이란”, <https://www.copyright.or.kr/education/educlass/learning/what-the-copyright/definition/index.do>
- [4] 전문예술법인단체, “저작권이란”, https://www.gokams.or.kr:442/artsdb/05_report/copyright01.asp
- [5] 권순일, 고영한, 조희대, 조재연, “대법원 2017. 9. 7. 선고 2017다222757 판결”, Sep. 2017.