

2019 한국정보보호학회 동계학술대회

# CISC-W'19

Conference on Information Security and Cryptography-Winter 2019

**일시** 2019. 11. 30(토)

**장소** 중앙대학교(흑석동)

**Proceedings**



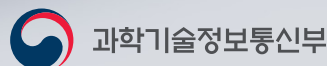
**주최**



**주관**



**후원**



## 구두발표 트랙 A

| 발표 세션                                                         | 발표논문                                                   |
|---------------------------------------------------------------|--------------------------------------------------------|
| <b>A4</b><br><b>콘텐츠 보안</b><br>13:10 ~ 14:10<br>좌장<br>곽 진(아주대) | 음악 스트리밍 프로파일링 기반의 불법 음악 스트리밍 사이트 탐지 방안<br>정혜선, 곽진(아주대) |
|                                                               | 공공 인프라 영상보안을 위한 보안 프레임워크 연구<br>유진용, 김영갑(세종대)           |
|                                                               | 클라우드 공유 기능의 저작권 침해 방지를 위한 연구<br>박소희, 곽진(아주대)           |
|                                                               | 공공장소 CCTV 프라이버시 확보 기법<br>권혁동, 장경배, 권용빈, 서화정(한성대)       |

| 튜토리얼                                            | 발표논문                           |
|-------------------------------------------------|--------------------------------|
| <b>차세대인증</b><br>14:20 ~ 15:20<br>좌장<br>최대선(공주대) | DID 개념과 영지식 증명<br>황정연 박사(ETRI) |

# 음악 스트리밍 프로파일링 기반의 불법 음악 스트리밍 사이트 탐지 방안

정해선\*, 곽진\*\*

\*아주대학교 컴퓨터공학과 정보보호응용및보증연구실

\*\*아주대학교 사이버보안학과

## Illegal Music Streaming Site Detection based on Music Streaming Profiling

Hae-Seon Jeong\*, Jin-Kwak\*\*

\*ISAA Lab., Department of Computer Engineering, Ajou University.

\*\*Department of Cyber Security, Ajou University.

### 요약

음악 저작물 복제는 간단하고 쉽게 할 수 있는 작업이기 때문에 불법 음악 저작물이 빈번하게 유포되고 있다. 이에 따라 내부 서버에 음악 복제물을 저장하고, 불법으로 음악 스트리밍을 제공하는 사이트가 운영되고 있다. 불법 음악 스트리밍 사이트들은 무료로 서비스를 제공하기 때문에 많은 사용자들이 사용하며, 해당 사이트들은 광고를 통해 수익을 얻기도 한다. 이에 따라 본 논문에서는 음악 저작물을 보호하기 위하여 합법적인 사이트와 불법 사이트를 비교 분석한 후 이에 기반한 불법 사이트 탐지 기법을 제시하였다.

### I. 서론

재생한 음악 녹음하기, 유튜브에서 mp3 파일 추출하기 등 음악을 복제하는 것은 쉽게 할 수 있다. 실제로 2018년 12월 기준으로 음악은 두 번째로 많이 복제된 콘텐츠로 분류되었다[1]. 불법 음악 스트리밍 사이트는 쉽게 복제한 음악 저작물을 파일의 형태로 무료 유포하며 광고로 수익을 얻기도 한다. 이에 따라 본 논문에서는 음악 저작물을 보호하기 위해 불법 음악 스트리밍 사이트 탐지 기법을 제시하였다.

### II. 관련 연구

#### 2.1 저작권 침해 기준

저작권법에 따르면 저작권자, 저작권 신탁

관리자의 이용 허가를 받지 않고 음악 저작물을 공유하는 경우 저작권법에서 규정한 전송권을 침해한다고 판단한다. 또한 권리자의 허가 없이 음악 파일을 통해 광고 수익을 보는 것 또한 저작권법을 위반하는 행위이다. 1분 이내로 편집하여도 명백한 전송권 침해 행위이며 면책 사유가 될 수 없다. “멜론” 등의 사이트에서 1분 미리듣기를 무료 제공하는 것은 이미 권리자의 허가를 받은 상태에서 서비스를 제공하는 것이므로 저작권 침해가 아니다. 이와 같은 저작권법 위반은 업로더의 책임이지만, 저작권 침해에 대해 온라인서비스제공자가 확인하여 방조한 경우에는 방조 책임이 있다.

본 연구는 문화체육관광부 및 한국저작권위원회의 2019년도 저작권기술개발사업의 연구결과로 수행되었음.

## 2.2 저작권 침해 사이트 사례

### □ A 사이트

A 사이트는 서버 내에 저장되어 있는 음악 파일을 무료로 제공하고 있다. 무료 음악 스트리밍, 다운로드뿐만 아니라 웹페이지에 음악을 넣기 위한 태그도 제공하고 있다. 또한, 저작권자의 허가 없이 복제된 음악 파일도 제공하고 있다. 이는 저작권자의 이용 허가를 받지 않고 음악 저작물을 공유하는 것이기 때문에 저작권법 위반 행위에 해당한다.

### □ B 사이트

B 사이트는 서버 내에 저장되어 있는 음악 파일에 대하여 스트리밍과 다운로드 기능을 제공하고 있다. B 사이트는 사용자에게 이와 같은 서비스를 제공하면서 광고를 통해 수익을 얻는다. 이는 저작권자의 이용 허가를 받지 않고 음악 저작물을 공유하며 저작물을 통해 수익을 얻는 저작권 위반 행위에 해당한다.

## III. 음악 스트리밍 사이트 비교·분석

본 논문에서는 합법적인 음악 스트리밍 사이트와 불법 음악 스트리밍 사이트를 비교·분석함으로써 불법 음악 스트리밍 사이트를 탐지할 수 있는 특징을 도출하였다. 합법적인 음악 스트리밍 사이트로 이용량이 가장 높은 “멜론”, “지니”, “벅스”를 대상으로 분석하였으며[2], 불법 음악 스트리밍 사이트로 2.2절의 A 사이트와 B 사이트를 대상으로 분석하였다.

### 3.1 합법적인 음악 스트리밍 사이트 분석

합법적인 음악 스트리밍 사이트는 사용자가 음악 파일에 대해 사용 권한이 없는 경우에는 1분 미리 듣기만 제공한다. 이를 위해 서버는 음악 재생 요청 시 음악 정보, 사용자 정보 및 가사를 JSON 형태로 먼저 전송한 후에 음악 파일을 전송한다.

음악 재생 요청부터 음악 파일 전송까지의 과정은 3개 사이트 모두 대략 30개의 세션으로 구성되어 있다. “벅스”와 “지니” 서버는

암호화되어 있지 않은 평문 정보를 JSON 형태로 전송한다. 음악 정보, 사용자 정보, 사용자의 이용권 결제 정보, 가사 등을 평문의 JSON 형태로 전송한다. 그 후에 음악 파일을 전송한다. 하지만 “멜론”의 경우 JSON 내 key, value 값이 모두 암호화되어 있어 정보를 식별할 수 없었다. 하지만 요청 URL을 통해 멜론 서버에 요청한 정보를 추정할 수 있었다.

음악 재생 요청 시 서버의 구체적인 정보 전송 과정은 다음과 같다.

### □ 벅스

#### Step 1. 음악 정보 및 사용자 정보 전송

“https://music.bugs.co.kr/player/track/‘노래 ID?’” 형태의 URL을 요청하면 서버는 [그림 1]과 같이 암호화되지 않은 평문의 JSON 형태로 음악 정보, 사용자 정보 및 사용자의 이용권 결제 정보를 전송한다. 해당 URL 내에서 음악 정보 요청에 해당하는 “player”, “track” 키워드를 볼 수 있다.

|                                                                                                                                                         |             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| auth=True<br>inventory=0<br>isAbuser=False<br>isLike=True<br>isLoggedIn=True                                                                            | → 사용자 정보    |
| track_id=31708907                                                                                                                                       | → 음악 ID     |
| track_main_artist_nm=AKMU (악동뮤지션)<br>track_no=3<br>track_status=OK<br>track_title=어떻게 이별까지 사랑하겠어, 널 사랑하는 거지                                             | → 음악 정보     |
| display_goods_end_dt=1573964367000<br>display_goods_exp_dt=1574002799000<br>display_goods_name=둘7(4스마트폰 저장(매월))<br>display_goods_start_dt=1571285967000 | → 이용권 결제 정보 |

[그림 1] 음악 정보 및 사용자 정보 일부

#### Step 2. 가사 전송

“https://music.bugs.co.kr/player/lyrics/T/‘음악 ID?’” 형태의 URL을 요청하면 서버는 [그림 2]와 같이 사용자가 요청한 음악의 가사를 암호화되지 않은 평문의 JSON 형태로 전송한다. 해당 URL 내에서 가사에 해당하는 “lyrics” 키워드를 볼 수 있다.

```
JSON
{
  "lyrics": "19.3]일부러 몇 발자국 물러나 #26.2]내가 없이 혼자 걷는 날 바라본다."
}
```

[그림 2] 서버가 전송한 음악 가사 일부

#### Step 3. 음악 파일 전송

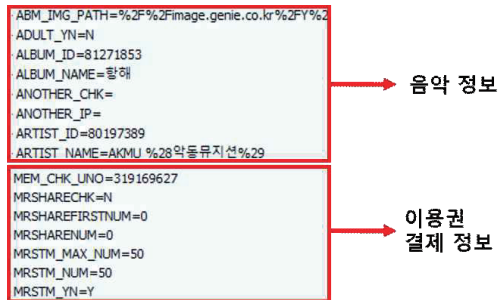
“https://w-aod.bugs.co.kr/raout/aac128/‘음악

ID'.m4a?mediatype=TRACK&mediaid='음악 ID'&" 형태의 URL을 요청하면 서버는 음악 파일을 전송한다. 해당 URL 내에서 요청한 음악을 명시하기 위해 사용되는 "mediaid" 키워드를 볼 수 있다.

□ 지니

#### Step 1. 음악 정보 및 사용자 정보 전송

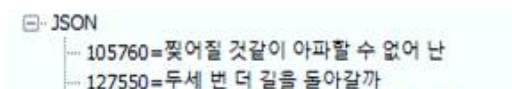
"https://www.genie.co.kr/player/playStmInfo.json음악 ID'?" 형태의 URL을 요청하면 서버는 [그림 3]과 같이 암호화되지 않은 평문의 JSON 형태로 음악 정보, 사용자의 이용권 결제 정보를 전송한다. 해당 URL 내에서 음악에 해당하는 "player"와 스트리밍의 약자 "stm" 키워드를 볼 수 있다.



[그림 3] 음악 정보 및 사용자 정보 일부

#### Step 2. 가사 전송

"https://dn.genie.co.kr/app/purchase/get\_msl.asp?path=a&songid='음악 ID'&..." 형태의 URL을 요청하면 서버는 [그림 4]와 같이 사용자가 요청한 음악의 가사를 암호화되지 않은 평문의 JSON 형태로 전송한다. 해당 URL 내에서 가사에 해당하는 "lyrics" 키워드를 볼 수 있다. 해당 URL 내에서 요청한 음악을 명시하는 "songid" 키워드를 볼 수 있다.



[그림 4] 서버가 전송한 음악 가사 일부

#### Step 3. 음악 파일 전송

"https://web-hls.genie.co.kr/web/\_definst/\_ND\_O/S ONG/.../mp3: '음악 파일 명'/playlist.m3u8?token=..." 형태의 URL을 요청하면 서버는 음악 파일을 전송한다.

□ 멜론

#### step 1. 사용자 인증 정보 전송

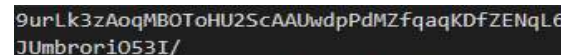
"http://appsvr.melon.com/cds/user/player/user\_authInfoUser5.json" 형태의 URL을 요청하면 서버는 [그림 5]와 같이 암호화된 JSON 데이터를 전송한다. 해당 JSON 데이터는 식별할 수 없지만 URL 내에서 사용자 및 인증과 관련된 "user", "player", "author" 키워드를 볼 수 있다.



[그림 5] user\_authInfoUser5.json의 일부 내용

#### Step 2. 음악 정보 전송

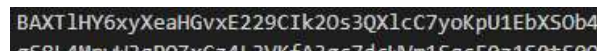
"http://appsvr.melon.com/cds/delivery/streaming/player/streaming\_getTgtInfo.json" 형태의 URL을 요청하면 서버는 [그림 6]과 같이 암호화된 JSON 데이터를 전송한다. 해당 JSON 데이터는 식별할 수 없지만 URL 내에서 음악과 관련된 "player", "stream" 키워드를 볼 수 있다.



[그림 6] streaming\_getTgtInfo.json 내용 일부

#### Step 3. 가사 전송

"http://appsvr.melon.com/cds/delivery/streaming/player/streaming\_getstnlyricsinfo2.json?\_cid=..." 형태의 URL을 요청하면 서버는 암호화된 JSON 데이터를 전송한다. 해당 JSON 데이터는 식별할 수 없지만 URL 내에서 가사와 관련된 "lyrics" 키워드를 볼 수 있다.



[그림 7] streaming\_getstnlyricsinfo2.json 내용 일부

#### Step 4. 음악 파일 전송

"http://sts.melon.com/p3/.../'음악 파일명'..." 형태의 URL을 요청하면 서버는 음악 파일을 전송한다.

### 3.2 불법 음악 스트리밍 사이트 분석

불법 음악 스트리밍 사이트의 경우 별도의 사용자 인증 없이 자체 서버 내에 있는 음악 파일을 바로 제공한다. 음악 재생 요청 시 음악 정보, 사용자 정보, 가사를 전송하지 않고 바로

음악 파일을 전송하기 때문에 전체 세션 수는 2, 3개로 매우 적다. [표 1]의 URL을 요청함으로써 단순히 서버 내부의 음악 파일을 요청한다.

[표 1] 불법 음악 스트리밍 사이트의 음악 파일 요청 URL

| 사이트 명 | 요청 URL                            |
|-------|-----------------------------------|
| A 사이트 | https://“A 사이트 도메인”/.../“음악 파일 명” |
| B 사이트 | https://“B 사이트 도메인”/.../음악 파일 명   |

#### IV. 불법 음악 스트리밍 사이트 탐지

3장에서 합법적인 음악 스트리밍 사이트와 불법 음악 스트리밍 사이트를 분석한 결과 도출한 불법 음악 스트리밍 사이트의 특징을 기반으로 제안하고자 하는 불법 음악 스트리밍 사이트 탐지 기법 동작 과정은 다음과 같다.

##### Step 1. 합법적인 음악 스트리밍 사이트 목록과 비교

불법 사이트를 탐지하기 위해 가장 먼저 합법 사이트 목록과 비교한다. 일치하지 않는 경우 다음 단계들을 통해 불법 사이트인지 탐지한다.

##### Step 2. 통신 세션 수 확인

음악 재생 시 불법 음악 스트리밍 사이트는 음악 정보, 사용자 정보 및 가사 전송 없이 음악 파일만 전송하기 때문에 세션 수가 5개 미만으로 적게 나타난다. 따라서 적은 세션 수를 통해 음악을 재생한다면 불법 음악 스트리밍 사이트로 탐지할 수 있다.

##### Step 3. 요청 URL 키워드 DB 비교

합법적인 음악 스트리밍 사이트의 경우 음악 파일 전송에 앞서 음악 정보, 사용자 정보 및 가사를 전송한다. 음악 정보 요청 URL에는 음악을 뜻하는 “track”, “player”와 스트리밍에 해당하는 “stream”, “stm”이 포함되어 있었다. 요청하는 음악을 명시하여 요청하는 경우에는 음악 ID를 “mediaid”, “songid” 등을 사용하여 URL 내에 명시하였다. 또한, 3개의 합법적인 음악 스트리밍 사이트 모두 가사 요청 URL에는 “lyrics”가 포함되어 있었다.

이에 따라 음악을 의미하는 키워드, 음악 ID를 의미하는 키워드, 가사를 의미하는 키워드 항목으로 구성된 DB 기반으로 불법 음악

스트리밍 사이트를 탐지하고자 한다. 해당 DB 내에는 3.1절에서 도출한 키워드 외에도 DB 키워드 항목에 대응 가능한 키워드가 포함되어 있다. 음악 재생 시 요청 URL을 [표 2]의 키워드 DB와 매칭하여 해당 사이트가 합법적인 스트리밍 사이트처럼 음악 파일 전송에 앞서 음악 정보, 사용자 정보 및 가사를 전송하는지 확인한다. [표 2]의 키워드 DB와 매칭되지 않으면 불법 음악 스트리밍 사이트로 탐지한다.

[표 2] 키워드 DB

| 키워드 항목 | 키워드                                                                                                                      |
|--------|--------------------------------------------------------------------------------------------------------------------------|
| 음악     | track, media, player, stream, stm, song                                                                                  |
| 음악 ID  | trackid(track_id)<br>mediaid(media_id)<br>playerid(player_id)<br>streamid(stream_id)<br>stmid(stm_id)<br>songid(song_id) |
| 가사     | lyrics                                                                                                                   |

#### V. 결론

본 논문에서는 음악 저작물을 보호하기 위하여 합법 음악 스트리밍 사이트와 불법 음악 스트리밍 사이트를 비교·분석하였고, 이를 통해 도출한 불법 음악 스트리밍 사이트의 특징 기반 탐지 기법을 제시하였다. 만약 불법 사이트가 의도적으로 합법 사이트와 비슷한 세션 수, 비슷한 키워드를 가지는 통신을 한다면 해당 방법은 사용할 수 없기 때문에 이를 보완하는 방법과 함께 사용한다면 더 높은 탐지율을 보일 수 있을 것이다.

#### [참고문헌]

- [1] 문화체육관광부(저작권정책과), “저작권법”, October, 2018.
- [2] 콘텐츠진흥원, “국내 음악이용자 실태 조사”, February, 2019.