



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

경영학 석사학위 논문

데이터 분석을 통한 온라인
음원차트의 어뷰징 영향요인 탐색

An empirical study on the effecting factors
of online music chart abusing

2017년 2월

서울시립대학교 대학원
경영학과 MIS 전공
곽 아 영

데이터 분석을 통한 온라인 음원차트의 어뷰징 영향요인 탐색

An empirical study on the effecting factors
of online music chart abusing

지도교수 장 남 식

이 논문을 경영학석사학위 논문으로 제출함

2016년 12월

서울시립대학교 대학원
경영학과 MIS 전공
곽 아 영

곽아영의 경영학 석사학위 논문을 인준함.

심사위원장 김 충 영 인

심 사 위 원 조 우 제 인

심 사 위 원 장 남 식 인

2016년 12월

서울시립대학교 대학원

국문 초록

온라인 음악 서비스 시장이 확대되고 음악 소비 채널이 오프라인 매체에서 디지털 매체로 전환됨에 따라 음원차트의 중요성은 점차 강화되었다. 이에 음원차트의 순위 조작 또는 저작권 사용료 수입을 목적으로 특정 음원을 부당하게 구입하는 음원사재기 현상이 나타났다. 음원사재기는 소비자에게 왜곡된 정보를, 저작권자에게는 사재기가 아니었으면 얻었을 저작권 수입의 감소와 나아가 방송출연 기회의 박탈로 이어질 수 있으므로 반드시 근절되어야 할 행위이다. 과거 일부 기획사의 음원 사재기 논란과 더불어 대중성으로 승부해온 아티스트보다는 강력한 팬덤을 가진 아이돌에게 유리한 차트라는 평가 등 음원차트의 공정성에 대한 논란은 예전부터 지속되어왔다.

이에 대한 방안으로 주요 음악 사이트들은 차트 집계에 ‘1일, 1계정, 1카운트 정책’을 반영하였지만 다량의 계정을 통해 이를 무력화시키고 스트리밍 서비스가 차트에 반영되는 재생시간을 고려한 자동 매크로 개발 등 그 방법이 더욱 지능화되고 고도화되는 실정이다.

따라서 본 논문에서는 국내 음원 유통 사이트의 실시간 음원차트 집계에 반영된 데이터를 이용해 어뷰징 여부 판단 패턴을 도출하였다. 구체적으로 어뷰징 의심군 유형을 구분하였고, 이를 기반으로 고객별 프로파일을 통해 어뷰징을 판단하기 위한 속성들을 제시하였다.

주제어 : 음악시장, 스트리밍 서비스, 음원차트 어뷰징

경영학 석사학위 논문

데이터 분석을 통한 온라인
음원차트의 어뷰징 영향요인 탐색

An empirical study on the effecting factors
of online music chart abusing

2017년 2월

서울시립대학교 대학원
경영학과 MIS 전공
곽 아 영

[목 차]

1. 서론	1
1.1 연구의 필요성 및 목적	1
1.2 연구 방법 및 구성	4
2. 온라인 음악시장 환경	6
2.1 해외 음악시장 동향	6
2.2 국내 음악시장 동향	10
3. 문헌 연구	14
3.1 어뷰징의 개념	14
3.2 이론적 배경	15
3.3 어뷰징 관련 연구	17
3.4 이상거래 탐지 사례	19
4. 분석	22
4.1 데이터 수집	22
4.2 트래픽 로그 1차 분석	24
4.3 계정별 프로파일 2차 분석	29
5. 결론	36
5.1 요약 및 결론	36
5.2 연구의 한계점 및 향후 연구방향	38

참고문헌	39
Abstract	42



[표 목차]

<표 2-1> 세계 음원시장 규모 및 전망	6
<표 2-2> 국가별 음악 산업 규모 Top15	7
<표 2-3> 2012~2015 국내 음반 발매 형태	11
<표 2-4> 국내 주요 서비스 이용 현황	12
<표 4-1> 시간대별 전체 재생건수 대비 특정곡의 점유율	23
<표 4-2> 특정곡의 시간대별 전체 재생건수/이용자 비교	25
<표 4-3> 시간대별 5회 이상 그룹별 회원수 및 재생건수	26
<표 4-4> 샘플 그룹별 일간(0~24시) 재생이력 추적	27
<표 4-5> 발매 당일 특정곡 순도 100% 재생 이용자	30
<표 4-6> Top1 이용자의 발매 당일 재생이력	31
<표 4-7> 계정별 프로파일 : 2차 분석 - 유형 1 해당 계정	32
<표 4-8> 계정별 프로파일 : 2차 분석 - 유형 2 해당 계정	33
<표 4-9> 계정별 프로파일 : 2차 분석 - 유형 4 해당 계정	35
<표 4-10> 2차 의심군의 누적 활동이력	35
<표 5-1> 어뷰징 여부 판단 패턴	37

[그림 목차]

<그림 2-1> 국내 주요 서비스 이용 현황	12
<그림 2-2> 국내 온라인 음원 유통 사이트의 실시간 음원차트	13



1. 서론

1.1 연구의 필요성 및 목적

인터넷과 정보통신기술의 발달에 따라 지식정보사회가 본격화되면서 정보화는 경제사회 전반의 구조와 개인·사회집단의 생활양식에 변화와 혁신을 불러오고 있으며 나아가 국가 발전의 원동력으로써 그 잠재력과 파급력에 지대한 관심이 쏠리고 있다.

네트워크로 사람, 사물, 데이터 등을 포함한 모든 것이 연결된 사회 즉 초연결사회가 도래하면서 거대한 융합의 시대, 인터넷을 넘어선 모바일 중심의 시대가 시작되었으며 클라우드, 사물인터넷(IoT), 빅데이터, 인공지능(AI) 등과 같은 서비스로 이어져 정치, 사회, 경제, 문화 등 전반에 걸쳐 혁신을 일으키고 있다. 특히 모바일 기기와 통신 속도의 발전으로 모든 산업 분야에서 모바일이 비즈니스의 중심이 되는 모바일 퍼스트 시대와 전자상거래, 콘텐츠 소비, 비즈니스 등 대부분의 경제사회 활동을 모바일에서만 처리하는 모바일 중심 시대가 선언되기도 하였다. 실제로 게임, 방송 등 다양한 콘텐츠 산업 부문에서 모바일 퍼스트 전략에 초점을 맞추고 있으며 음악 서비스 또한 모바일 시장이 활성화되고 있는 영역이다.

최근 수년간 온라인-모바일 음악 서비스 내에서 LTE 등의 네트워크와 고사양 스마트폰, 태블릿 등의 IT 발달에 따라 음원 다운로드 시장은 실시간 스트리밍¹⁾ 서비스로 점차 전환되어왔다. 과거 개인소장 중심의 다운로드에서 데이터 기반의 저렴한 스트리밍 서비스가 확산되면서 ‘장기적

1) 인터넷 상에서 음성이나 동영상 등을 실시간으로 재생하는 기술

재생'이 아닌 '단발적 소비'형태로의 이용자 패턴 변화에 따라 음원 콘텐츠 간의 경쟁은 날로 치열해지고 있는 상황이다.

음악시장이 음반에서 디지털 다운로드, 스트리밍으로 개편되면서 각 음원 사이트들은 이를 반영하여 음원 발매 직후 한 시간 단위로 집계되는 '실시간 차트'를 발표하고 있다. 대부분의 실시간 차트는 매시간 서비스 이용량 중 스트리밍40%, 다운로드 60%를 반영하고 있으며, 실시간 차트에 반영되는 디지털 음원 소비 행위(다운로드, 스트리밍) 건수는 일반 소비자의 의도적인 행위여야 하며 어떠한 인위적인 조작도 없어야 한다.

그러나 음원 문화의 변화에 따라 음원차트의 상위권에 노출되지 못하면 사장되어버린다는 위기의식이 콘텐츠 제작자와 극성팬들 사이에 팽배하면서 일부 이해관계자들의 불안 심리를 이용해 음원차트의 공정성에 악의적인 위해를 가하고자 행해지는 음원 사재기, 어뷰징(Abusing)성 트래픽의 유입 등이 문제시 되고 있다. 2015년 9월 JTBC '음원 사재기 의혹, 실체 추적' 보도 후 음원 사재기 의혹이 불어졌으며 이후 '뉴스룸'을 통해 JYP 엔터테인먼트 대표 박진영, 가수 이승환 등이 직접 출연하여 실태를 증언, YG엔터테인먼트는 보도에 환영한다는 입장을 표명하여 논란이 증폭되었다. 또한 2015년 9월 말 국내 최대 음원 사이트인 멜론이 각 음반 기획사를 대상으로 '[로엔] 음원사재기(어뷰징) 행위 근절을 위한 협조요청의 건'이라는 제목의 공문을 보낸 사실이 드러났으며, 공문에는 음원 사재기를 이용한 순위 차트 조작에 유감을 표하며 외부 업체와의 부당한 사례가 있을 경우 법률적 조치를 취하겠다는 내용이 담겨있었다.

이와 관련하여 문화체육관광부 주관으로 관련 업계와의 공청회를 통해 2016년 3월 3일 '음악 산업 진흥에 관한 법률'을 통과시켰다. 온라인 음악 서비스 제공자가 발표하는 음원차트의 순위가 소비자의 구매에 큰 영

향을 미침에도 불구하고 조직적이고 인위적인 개입에 따른 음원 사재기가 유통질서를 왜곡시킨다는 지적은 과거부터 계속 되어왔으나, 현행법상 처벌의 근거가 없어 한계가 있었던 점으로 개정안의 시행에 따라 음반제작자와 관련 업체들의 저작권료 수입 등을 얻기 위해 음원을 대량으로 구매하는 식의 행위를 규제하게 되는 것이다.

악의적인 목적을 가지고 음원차트의 공정성에 해를 가하여 온라인-모바일 음악시장을 교란하는 행위는 근절되어야 하며 이와 관련하여 법안이 통과되었으나 이를 근절할 기술적 방안은 관련 업계에 위임하는 수준이라 실효성이 매우 떨어지는 실정이다. 음원 사이트 관계자들은 인위적인 조작을 막기 위한 일환으로 실시간 차트는 특정 곡을 반복 재생하더라도 한 아이디 당 1시간 1회만을 카운트하여 집계하는 방식과 일간 차트는 하루에 한 아이디 당 1회만 반영되는 방식으로 음원차트를 발표하고 있으나, 어뷰징 방법은 끊임없이 진화하고 있으며 팬덤²⁾의 활동과 악의적인 어뷰징 행위를 구분하는 것 또한 현실적으로 어려운 부분이 있다.

따라서 본 연구에서는 음원차트의 공정성 확보를 위한 탐색방안으로 음원차트의 어뷰징성 트래픽을 분석하고 이를 방지하기 위한 영향요인을 식별하고자 한다.

2) 대중적인 특정 인물이나 분야를 열성적으로 좋아하는 사람들을 하나의 큰 틀로 묶어 정의한 개념

1.2 연구 방법 및 구성

본 연구에서 대상이 되는 음원차트 어뷰징의 대표적인 유형은 두 가지이다.

- 1) 이해관계자 혹은 대행사 등 소수에 의한 특정 음원 집중화 유형
- 2) 팬덤 등 다수에 의한 특정 음원 집중화 유형

본 연구에서는 어뷰징 여부를 판단한 수 있는 변수들을 식별하기 위해

1) 차트 순위 집계에 영향을 주는 트래픽 로그를 분석, 2) 어뷰징으로 추정되는 대상의 유형을 1차적으로 식별, 3) 해당 트래픽이 실제 어뷰징에 해당하는지에 대한 검증을 2차적으로 수행하여 온라인 음원차트 어뷰징을 탐지할 수 있는 요인들을 구별하고자 한다.

먼저, 국내 온라인 디지털 음원 사이트 한 곳의 음원차트 집계에 반영된 원천 트래픽 로그를 수집하여 특정 곡을 대상으로 팬덤 규모가 큰 아티스트 혹은 기획사 기대주에 의한 어뷰징 가능성이 높은 시점의 트래픽 로그 데이터 샘플을 분석한다. 이 과정을 통해 1차적으로 식별한 의심패턴을 이용하여 2차 검증을 수행한 이후 의심 계정에 대한 프로파일, 로그인 이력, 활동이력(타 음원 감상 이력, 팬 맺기/좋아요 이력) 등을 분석하여 어뷰징 여부에 대한 검증을 시행한다.

본 논문의 구성은 다음과 같다. 1장은 논문의 서론으로 본 연구의 필요성과 목적, 연구방법에 대해서 서술하였다. 2장에서는 국내외 음악시장의 동향에 대해 조사하였고 3장에서는 관련 선행 연구 조사를 통해 온라인 음원차트의 효과, 어뷰징 관련 연구, 이상거래 탐지 기법 사례를 살펴보

았다. 4장에서는 국내 음원사이트의 트래픽 데이터를 통해 어뷰징 의심 유형을 구분, 이를 활용하여 1차 어뷰징 의심군을 식별한다. 이후 어뷰징 의심계정에 대해 앨범 발매 전후 3개월간의 재생 이력을 추적하고, 재생 활동외의 사용자 활동이력 분석을 통해 어뷰징 여부 판단 속성을 제시한다. 5장에서는 위의 내용들을 정리하고, 본 연구의 결론과 한계점 및 향후 연구방향을 제시하였다.



2. 온라인 음악시장 환경

2.1 해외 음악시장 동향

음악시장이 실물(CD) 음반시장에서 디지털 음원시장으로 개편되면서 전 세계적으로 디지털 음원의 비중은 매년 증가하고 있으며 디지털 음원 시장 중에서도 음원 스트리밍 서비스는 두드러진 증가 추세를 보이는 부분이다. <표 2-1>에서 볼 수 있듯이 2018년 세계 온라인 음원 스트리밍 시장은 약 4조원에 이를 것으로 전망되며, 음원에 접근하는 방식 또한 점차 다운로드 서비스 중심에서 스트리밍 서비스 중심으로 변화할 것으로 보인다(PwC. 2014).

<표 1-1> 세계 음원시장 규모 및 전망(2013년 기준)

(단위 : 백만달러, %)

구분	2010	2011	2012	2013p	2014	2015	2016	2017	2018
공연음악	24,583	25,540	26,170	26,696	27,262	27,900	28,630	29,487	30,521
후원	5,793	5,848	5,930	6,008	6,097	6,197	6,305	6,423	6,548
티켓 판매	18,790	19,692	20,240	20,688	21,165	21,703	22,324	23,064	23,973
음반	22,892	22,098	21,435	20,720	20,351	20,175	20,095	19,951	19,925
디지털	7,260	7,757	8,719	9,432	10,180	10,905	11,544	12,020	12,491
다운로드	4,677	5,055	5,743	5,986	6,258	6,520	6,755	6,891	7,057
스트리밍	657	1,038	1,530	2,148	2,692	3,174	3,557	3,837	4,035
모바일	1,925	1,664	1,447	1,298	1,230	1,212	1,233	1,292	1,399
오프라인 음반	15,633	14,342	12,716	11,287	9,270	9,270	8,551	7,391	7,434
합계	47,476	47,638	47,605	47,415	48,075	48,075	48,725	49,438	50,446

*2013년 이후는 전망치

출처 : PwC(2014)

<표 2-2>를 살펴보면 2014년 세계 음악 산업은 미국이 33%의 점유율로 2위인 일본과는 3배 이상의 차이를 보이며 압도적인 영향력으로 선두자리를 지키고 있으며 독일, 영국, 프랑스, 호주 등이 그 뒤를 이었다. 향후 국가별 순위를 예상하면 고령화, 소비지출 감소 등의 문제를 겪고 있는 일본을 제치고 독일이 2위에 오를 것으로 예상되며 성장세가 높은 중국, 스웨덴이 10위권 내에 진입할 것으로 보인다. 다음은 주요 국가인 미국, 영국, 일본, 중국의 최근 음악시장 동향에 대해 구체적으로 살펴보겠다.

<표 2-2> 국가별 음악 산업 규모 Top15(2014년 기준)

(단위 : 백만달러, %)

순위	국가	2010	2011	2012	2013	2014p	2015	2016	2017	2018
1	미국	14,938	15,083	15,075	15,086	15,061	1,5015	15,062	15,240	15,562
2	일본	4,968	4,880	4,965	4,606	4,483	4,367	4,291	4,246	4,220
3	독일	4,304	4,358	4,301	4,379	4,423	4,455	4,491	4,527	4,553
4	영국	4,653	4,610	4,338	4,380	4,275	4,298	4,327	4,359	4,396
5	프랑스	1,759	1,776	1,797	1,831	1,789	1,767	1,752	1,740	1,730
6	호주	1,130	1,194	1,259	1,312	1,232	1,211	1,201	1,199	1,202
7	캐나다	1,161	1,179	1,208	1,239	1,225	1,222	1,225	1,234	1,250
8	러시아	1,013	998	1,053	1,110	1,045	1,004	1,064	1,081	1,110
9	이탈리아	990	986	955	922	915	906	904	904	907
10	중국	619	645	682	725	790	860	917	970	1,015
11	스웨덴	641	628	682	761	771	796	830	871	921
12	한국	692	720	753	744	743	741	740	744	759
13	스페인	692	642	626	571	606	605	607	610	616
14	브라질	432	449	465	481	500	506	516	527	538
15	멕시코	395	437	424	429	435	435	438	444	452

*2014년 이후는 전망치

출처 : PwC(2015)

1) 미국

디지털 음원 시장의 규모가 가장 큰 미국은 애플(Apple)을 비롯하여 판도라(Pandora), 구글 플레이(Google Play) 등 세계적인 디지털 음원 공급사들이 있어 디지털 음원 시장을 장악하고 있다. 스포티파이(Spotify), 판도라(Pandora), 디저(Deezer) 등 주요 업체들의 광고 기반 유·무료 스트리밍 서비스 증가로 2012년 오프라인 음반을 추월하였으나, 이후 디지털 음원 다운로드의 감소로 2014년 25%이던 산업 규모가 2019년 21%로 축소될 것으로 예측된다(PwC, 2015). 미국의 디지털 음원 시장은 스포티파이가 86%의 점유율을 차지하고 있으며 유튜브 뮤직(Youtube Music), 애플뮤직(Apple Music), 타이달(Tidal)의 시장 진출로 기존 스트리밍 업체들과의 경쟁이 치열할 것으로 보인다.

2) 영국

PwC(2015)에 의하면 2014년 영국의 음악 산업은 공연 산업이 성장하였으나 오프라인 음반과 디지털 음원 산업의 하락세로 인해 전년대비 2.4% 감소한 약 42억 달러로 집계되었다. 전 세계적인 추세와 유사하게 오프라인 음반 산업은 지속적으로 감소하고 있으며 디지털 스트리밍 산업이 성장하고 있다.

2014년 영국의 음악 판매 41%는 오프라인 음반으로 디지털 음원의 45%보다는 낮지만 스마트폰 사용의 확대로 음악을 휴대하기 위해 리핑³⁾ 서비스를 이용하는 사람들이 많을 것으로 예상되었다. 그러나 최근 영국에서는 저작권 관련 법 개정으로 아이튠즈(iTunes)의 사용이 모두 불법

3) CD나 DVD에 담겨 있는 디지털 오디오 파일 또는 비디오 파일 등을 PC 하드디스크로 복사하는 작업

이 되었다. 2015년 8월 5일 영국 고등법원은 저작권으로 보호된 음악 및 기타 콘텐츠를 다른 형태로 복제하는 행위를 불법으로 규정한 것이다. 애플의 큰 장점인 아이튠즈 서비스를 이용하지 못하게 되면서 영국 음악 산업에 큰 혼란이 예상된다.

3) 일본

세계 음악시장 2위인 일본은 엄격한 저작권법으로 음원 라이선싱 확보가 까다롭고 스트리밍 음원을 통한 불법 복제에 대한 우려로 일부 음반사들은 콘텐츠 제공을 꺼려하는 등 그간 글로벌 음악 스트리밍 서비스 업체들이 진출하는데 어려움을 겪어왔다. 그러나 10~20대를 대상으로 유튜브를 통한 스트리밍 음악 서비스를 이용하는 소비자들이 증가하고 있으며, LINE에서 '소셜'과 '음악'을 결합해 만든 서비스인 라인 뮤직(LINE Music)이 2015년 6월 일본에서 음악 스트리밍 서비스를 개시하였다. 일본 내에서 사용률이 가장 높은 라인 메신저와 연계해 음악을 첨부하여 메시지를 전송할 수 있게 하는 서비스를 제공하고 있다. 스포티파이, 애플 뮤직 등 글로벌 스트리밍 서비스 업체들 또한 일본 시장에 진입하고 있으며 일본의 디지털 음악시장에서 주도권을 잡으려는 경쟁이 치열할 것으로 예상된다.

또한 일본의 음악 산업은 여전히 오프라인 음반이 음악 산업을 주도하고 있으나 오프라인 음반의 지속적인 감소세와 디지털 음원의 급격한 성장이 이어지면서 향후 일본의 음악 산업은 디지털 음원의 수익이 증가할 것으로 전망된다.

4) 중국

중국에서는 온라인을 통해 음원을 구입하는 소비 행태가 일반적이며, 전체 음악 산업 중 디지털 음원 산업이 차지하는 비중이 크다. 오프라인 음반 산업의 경우, 불법 복제로 수입원 창출이 어려운 문제가 있는데다 음반 소매가격 또한 인상되어 그 비중이 더욱 축소될 것으로 전망된다. 최근 인터넷 업체인 ‘텐센트’, ‘바이두’, ‘알리바바’ 등의 엔터테인먼트 시장 투자로 디지털 음원 시장 또한 높은 성장세를 나타내고 있으며 향후 중국의 디지털 음원 시장은 매우 긍정적인 것으로 전망된다.

또한 사회가 안정되고 대중의 소비 수준이 향상되면서 지방중소도시를 중심으로 다양한 장르의 콘서트 수요가 증가하는 추세이다. 이런 추세에 따라 중국 온라인 스트리밍 서비스 업체들은 음원 저작권 확보에 집중하여 음악 감상 서비스에서 나아가 ‘온라인 콘서트’시장에 진입하기 위한 박차를 가하고 있다. 온라인 콘서트는 공간과 시간의 제약이 없다는 장점을 갖고 있으며 다수의 소비자를 확보할 수 있다는 점에서 여러모로 온라인 콘서트에 대한 수요가 꾸준히 증가할 것으로 보인다.

2.2 국내 음악시장 동향

국내 음악시장 또한 2003년을 기점으로 오프라인 음반 시장에서 온라인 음원시장으로 개편되었으며, 스마트 기기 사용의 확대와 네트워크 환경 개선과 같은 기술적 발전은 소비자들의 음악 소비형태 변화에 큰 영향을 미쳤다고 볼 수 있다. 국내 소비자들은 과거 음악을 소유하던 것에

서 소비하는 것으로 바뀌고 있으며 이에 따라 음원 다운로드가 아닌 스트리밍 서비스를 선택하는 비중이 커지고 있다. 이러한 소비 양상의 변화는 음반의 발매 형태에도 영향을 미치고 있다. <표 2-3>을 보면 2012년부터 2015년 상반기까지 국내 음반시장의 앨범 발매는 꾸준히 증가하고 있으며, 디지털 싱글 발매 형태는 전체 앨범 발매 형태의 약 70% 수준으로 디지털 싱글 발매 형태의 강세를 볼 수 있다.

<표 2-3> 2012~2015 국내 음반 발매 형태

구분	2012년	2013년	2014년	2015 (상반기)	누적
디지털 싱글	1,566	2,162	2,676	1,427	7,831
오프라인 CD	949	805	874	532	3,160
합계	2,515	2,967	3,550	1,959	10,991

출처 : (사)한국음반산업협회 방송심의 음반 관련 통계

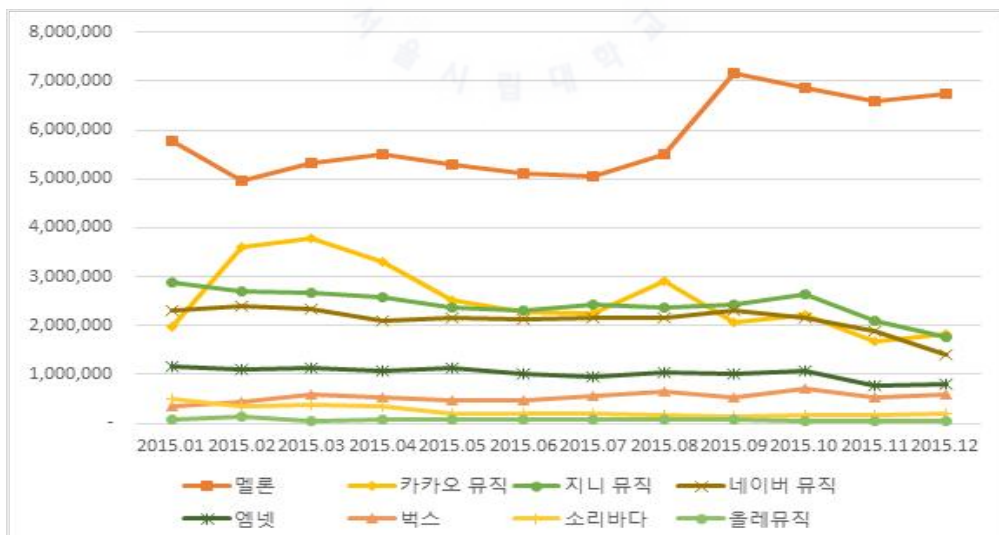
국내 음원의 유통과정을 살펴보면 가수나 연예기획사는 각 음원 사이트에 음반, 음원을 판매하기 위해 ‘중간 상인’에 해당하는 유통사와 배급계약을 맺은 음원사이트에 음원을 제공하고 소비자는 이를 스트리밍, 다운로드 등을 통해 구매하게 된다. 하루에도 수백에서 수천에 이르는 음원 및 음반이 출시되는 상황에서 음원 사이트들은 중간상인인 유통사를 통해 신곡들을 제공하며 SM, YG, JYP를 비롯한 유명 기획사, 소속사가 없는 무명 가수들 모두 유통사를 거쳐야만 신곡을 발표할 수 있다. 주요 음원 서비스 제공 사이트 중 로엔, 벅스, KT뮤직, Mnet 등은 유통사 업무 또한 병행하고 있으며 인터파크나 윈드밀 등과 같은 중견 업체들이 음원 유통 업무를 하고 있다. 반면 기획사나 가수가 의뢰한 곡들에 대해 일정 비율이나 액수를 받고 각 음원 사이트에 노출을 담당해주며 각 유통사에 음원들을 전달하는 대리인 역할을 하는 유통대리업도 존재한다.

2015년 6월 말 기준, 국내 디지털 음원 서비스를 제공하는 사이트는 총 17개이며, 주요 5대 사이트인 로엔엔터테인먼트의 Melon과 KT뮤직의 geine, CJ E&M의 Mnet, 네오위즈의 Bugs, 소리바다 등이 확인되었다.

<표 2-4> 국내 주요 서비스 이용 현황(Android 운영 체제)

기간	멜론	카카오 뮤직	지니 뮤직	네이버 뮤직	엠넷	벅스	소리바다	올레뮤직
2015.01	5,780,296	1,970,643	2,878,693	2,319,516	1,155,532	352,329	510,297	76,432
2015.02	4,956,411	3,616,258	2,711,671	2,388,709	1,101,733	442,673	350,179	132,756
2015.03	5,328,449	3,799,653	2,665,230	2,350,238	1,125,058	594,014	364,771	58,649
2015.04	5,491,307	3,311,790	2,588,526	2,108,759	1,062,649	520,588	353,596	71,021
2015.05	5,288,800	2,532,444	2,374,789	2,145,997	1,122,388	459,767	199,271	73,351
2015.06	5,098,015	2,263,726	2,307,432	2,119,731	1,019,160	469,974	200,035	63,722
2015.07	5,060,482	2,256,274	2,442,977	2,164,784	952,280	551,350	184,868	73,203
2015.08	5,499,793	2,904,004	2,354,602	2,162,865	1,033,704	663,265	179,305	87,470
2015.09	7,160,493	2,068,280	2,441,353	2,293,348	1,014,477	539,823	139,318	79,759
2015.10	6,855,846	2,230,100	2,635,705	2,163,011	1,067,579	698,084	173,576	61,333
2015.11	6,586,746	1,671,267	2,093,511	1,891,290	779,729	522,988	169,688	37,302
2015.12	6,734,142	1,820,914	1,750,518	1,410,830	788,803	580,147	190,781	40,432

출처 : Korean Click, 2015



<그림 2-1> 국내 주요 서비스 이용 현황(Android 운영 체제)

출처 : Korean Click, 2015

위의 표<2-4>과 그림<2-1>은 국내 이용자들 중 가장 많은 비중을 차지하는 안드로이드(Android)기반 서비스 ‘이용자 지수(UV, Unique Visitors)’를 나타낸 것으로 1위 사업자의 비중이 월등히 크며 향후 온라인 음악 서비스 시장의 경쟁이 더욱 치열해질 것으로 보인다. 2015년 상반기 다날엔터테인먼트의 달콤뮤직 서비스와 현대카드의 음악 서비스 종료, 다음과 카카오의 합병 소식 등으로 볼 때 다른 서비스들의 통·폐합 가능성 또한 예상된다.

국내 온라인 음원차트는 멜론, 지니, 엠넷, 벅스 등과 같은 음원 사이트와 한국음악콘텐츠산업협회가 출범시킨 가온차트를 포함해 총 12개의 음원차트가 존재한다. 국내 시장 점유율이 가장 높은 온라인 음원 유통 사이트 5곳의 경우를 살펴보았을 때 ‘실시간 차트’, ‘일간 차트’, ‘주간 차트’, ‘월간 차트’를 기본적으로 제공하고 있었으며 그 외에도 ‘장르별 차트’, ‘아티스트 차트’ 등 다양한 차트를 제공하는 것을 <그림 2-2>와 같이 확인할 수 있었다.

순위	곡명	작가	곡명	작가	순위	곡명	작가
1	It Ain't Love Carnaval Jackson Bnet		1	에라 모르겠다 BIGBANG / MADE	1	에라 모르겠다 BIGBANG / MADE	
2	바라 보르겠다 BIGBANG / MADE		2	LAST DANCE BIGBANG / MADE	2	LAST DANCE BIGBANG / MADE	
3	LAST DANCE BIGBANG / MADE		3	GIRLFRIEND BIGBANG / MADE	3	GIRLFRIEND BIGBANG / MADE	
4	GIRLFRIEND BIGBANG / MADE		4	Hush Lasse Lindh / 도깨비	4	Stay With Me 한영 & 펀치 (Punch)	
5	저 별 헤이즈 (HEIZE) / 저 별		5	내 눈에만 보여 10cm / 도깨비 OST	5	Hush Lasse Lindh / 도깨비	
6	Stay With Me 한영 (CHANYEOL) / 펀치						

순위	곡명	작가	순위	곡명	작가
1	에라 모르겠다	BIGBANG	1	에라 모르겠다	BIGBANG
2	LAST DANCE	BIGBANG	2	LAST DANCE	BIGBANG
3	GIRLFRIEND	BIGBANG	3	GIRLFRIEND	BIGBANG
4	Hush	Lasse Lindh	4	Stay With Me	한영 & 펀치
5	Stay With Me	한영 & 펀치	5	저 별	헤이즈 (HEIZE)

<그림 2-2> 국내 온라인 음원 유통 사이트의 실시간 음원차트

3. 문헌 연구

3.1 어뷰징의 개념

어뷰징이란 ‘남용’과 ‘오용’을 뜻하는 영어 어뷰스(abuse)에서 파생된 어휘로 인터넷 포털 사이트에서 조회수를 올리기 위해 같은 기사의 제목이나 내용을 살짝 바꾼 뒤, 반복해서 전송하는 행위이다(김병희, 2015). 한편 온라인 게임 용어로도 자주 사용되는데 여기서 어뷰징은 게임의 시스템을 이용해 불법적인 이익을 취하는 행위를 뜻한다. 즉 게임에서 허용하지 않는 방법이나 게임의 허점을 이용해 대결 결과를 조작하고, 이를 통해 게임 내에서 부당한 이득을 누리는 것을 가리킨다.

음악시장에서 발생하는 음원 사재기(Abusing)란 음악 차트 순위 조작 또는 저작권 사용료를 목적으로 저작권자 또는 저작인접권자⁴⁾가 음원을 부당하게 구입하거나, 전문 위탁 업체 및 기타 관련업체를 통해 음원을 부당하게 구입하도록 하는 행위이다. 음원 사재기의 원인은 국내 온라인 음원 서비스 업체들이 제공하는 실시간 차트와 월정액 무제한 스트리밍 서비스에서 찾을 수 있다. 과도한 트래픽 유발과 사재기를 통해 실시간 차트에 진입하게 되면 방송 및 공연 섭외 시 이점을 얻을 수 있기 때문이다. 게다가 월정액 무제한 스트리밍 서비스의 경우 일반적으로 6,000원 내외의 금액으로 이용할 수 있으며 다양한 할인 이벤트를 통해 소액만 지불하고도 해당 서비스를 이용할 수 있다. 음악 1곡이 1회 스트리밍 될

4) 저작물의 창작자는 아니지만 저작물의 배포에 기여하는 자로 실연자, 음반 제작자, 방송 사업자가 이에 해당

때, 월정액 무제한 스트리밍 상품을 기준으로 저작자와 제작자, 실연자⁵⁾에게 총 3.6원의 사용료가 지급되는 것이다. 이렇듯 1곡을 스트리밍 하는데 많지 않은 비용이 드는 구조로 인해 음원 사재기가 끊임없이 발생하는 것으로 파악된다.

3.2 이론적 배경

1) 편승효과

편승효과 또는 밴드웨건 효과는 어떤 선택이 대중적으로 유행하고 있다는 정보로 인하여 그 선택에 더욱 힘을 실어주게 되는 효과를 뜻하는 것으로, 다양한 상품 중 음원에 대한 정보 또한 편승효과가 큰 상품임이 입증되어 있다(Leibenstein, 1950; Towse, 1992).

음반의 품질은 소비자가 직접 들어보기 전까지 알 수 없는 특징을 가진 경험재로서 이러한 경험재들은 소비되기 이전에 생산자와 소비자 간 정보 비대칭(Information Asymmetry)의 문제가 발생하게 된다. 정보비대칭을 극복하기 위해서 소비자는 그 상품의 품질에 대한 정보를 수집하거나 과거의 소비 경험을 바탕으로 소비를 결정한다(임성준 외, 2004). 음원 소비를 결정할 때 음원차트의 순위를 고려하는 것도 이에 해당하는 것으로 정보비대칭을 방지하기 위해 대중의 선호를 참고하는 것이다. 이는 음원을 특정 소비자의 선호도가 다른 소비자의 선호도에 영향을 받는 편승효과가 작용하는 상품으로 만들며(Leibenstein, 1950; Towse, 1992), 음악

5) 저작물을 연주나 가창 등으로 표현하거나 저작물이 아닌 것을 이와 유사한 방법으로 표현하는 사람

산업에서 소수의 슈퍼스타들이 시장을 지배하는 현상을 설명한다(Rosen, 1981; Alder, 1985).

이처럼 소비자들은 많은 사람들이 구매하는 음원을 구매하는 성향이 있으며, 음원 발매 초기의 성과는 편승효과로 인해 향후 지속적인 성과로 이어지게 된다(Strobl & Tucker, 2000). 이러한 소비자 심리를 바탕으로 음원 판매자들이 편승효과를 이용한 음원판매 마케팅 활동을 펼친다면 음반의 성적에 큰 영향을 미칠 수 있을 것이다.

2) 랭킹효과

사람은 합리적인 선택을 위해서 개인 스스로가 수집한 정보를 무시하고 타인이 판단한 결정을 따르는 무리효과(Information Cascade)에 의존하는 경향이 있다(Leibenstein, 1950). 무리효과는 불확실한 정보를 가지고 있는 상황에서 합리적인 의사결정을 위해 타인의 정보를 참고하는 과정에서 발생하는 것으로 정보를 가지고 있지 않는 상태에서 오로지 다수의 사람들이 하는 선택을 따르는 군중심리와는 차이가 있다.

소비자는 온라인 디지털 콘텐츠를 이용하기 위해 온라인 디지털 콘텐츠를 유통하는 사이트에 접속하게 되고 이는 종종 콘텐츠 판매 순위 정보가 포함된 화면에 노출되게 한다(Chevalier & Goolsbee, 2003). 디지털 음원 소비자는 디지털 음원 유통 사이트의 음원차트를 통해 음원 선택에 필요한 정보를 획득할 수 있고, 그 정보를 통해 소비할 음원을 결정할 수 있다.

이러한 성향에 따라 소비자가 순위차트의 상위에 위치한 제품을 선택하고자 하는 현상을 랭킹효과(Ranking Effect)라 한다(Spoerri, 2008). 즉,

상위권 상품일수록 소비가 증가하고, 하위권 상품일수록 소비는 감소한다는 사실을 도출할 수 있다.

음원차트 순위 또한 랭킹효과가 발생할 수 있다. 음원 유통 사이트 이용자들은 음원차트에 노출되면서 얻는 음원 정보로 인해 상위권 음원을 선택할 가능성이 높게 되고, 이 과정이 반복되면서 상위권 음원의 경우 꾸준히 상위 순위를 유지할 수 있게 되는 것이다.

위의 편승효과, 랭킹효과 두 가지 이론을 온라인 음원차트의 순위가 대중들과 음악 산업 관계자들에게 막대한 영향력을 미치는 근간으로 설명할 수 있다.

3.3 어뷰징 관련 연구

다양한 분야에서 발생하고 있는 어뷰징 행위와 이를 방지하기 위하여 진행된 연구들에 대해 살펴보고자 한다. 먼저 포털 사이트에서 기사 어뷰징을 방지하기 위한 검색 알고리즘 개발 연구로 김병희(2015)는 뉴스 유통사(포털사), 뉴스 제공자(인터넷 신문사), 뉴스 효과 활용자(광고주) 등 정보 제공자를 대상으로 질적 연구방법의 하나인 초점집단법(FGI: Focus Group Interview)을 적용하여 포털 검색 알고리즘 개발에서 기사 어뷰징 방지를 위한 연구를 하였다. 기사 어뷰징이란 인터넷 언론사가 검색 클릭 수를 늘리기 위해 처음 보도된 기사를 유사한 내용과 제목으로 변형해 반복 전송하는 행위로 온라인 언론사의 증가로 인터넷 언론사의 경쟁에 의한 어뷰징이 급증하고 있으며 이러한 어뷰징 행위는 저널리즘의

윤리의식과 언론 본연의 책무를 상실하는 등 사회문제로 직결되고 있다. 이 연구에서는 현재 포털에서의 기사 어뷰징 실태를 분석하고 기사 어뷰징 방지를 위한 검색 알고리즘 개발 가이드라인을 제시하였다.

다른 분야의 연구로 박요한(2014)은 데이터마이닝 기법인 의사결정나무를 활용하여 언론매체에 보도된 사회복지시설의 정부보조금 부정수급 경로를 분석하고 관리 우선순위를 결정하는 부정수급 주요 요인 모형을 제안하였다. 이선아 외(2015)는 의사결정나무를 이용하여 농산물 도매시장에서의 허위거래를 탐지하는 모형을 제시하였다. 허위거래란 실제 발생하지 않은 거래를 실물 없이 거래한 것처럼 조작하여 대금을 정산하는 것으로, 변수 간 상관관계 및 분포도 등을 통해 데이터 특성을 파악하고 예측모형을 구축하여 허위거래와 정상거래를 분류하는 패턴을 도출해냈다.

이처럼 어뷰징 행위는 포털, 정부 보조금, 농산물 유통 등 다양한 분야에서 부정수급, 허위거래라는 명칭으로 일어나고 있으며 단순히 금액에 대한 피해뿐만이 아니라 신뢰와 공정성에 대한 우려로 이어질 수 있는 문제이다. 이러한 어뷰징을 방지하기 위한 목적으로 여러 연구가 진행되고 있었으며 각 산업군에서 이를 예측하고 적발하기 위해 대응방안을 구상 및 수립하고 있었다.

3.4 이상거래 탐지 사례

본 연구에서는 데이터의 기초 통계 분석을 통해 의심군을 설정하고 어뷰징 여부 판단에 영향을 미치는 속성을 추정하는 것에서 그치나, 이 속성들이 실제 어뷰징 판별에 유의한가를 판단하기 위해서는 추후 데이터 마이닝과 같은 모형 선정과정을 거쳐야 한다.

네트워크, 시스템, 응용서비스 등에서 발생하는 데이터 및 보안 관련 이벤트 간의 연관성을 분석하여 보안 지능을 향상시키는 차세대 정보 분석 기술이 필요하며, 이를 빅데이터 분석기술을 활용한 시큐리티 인텔리전스라고 한다. 이를 위해서는 데이터마이닝, 웹마이닝, 텍스트마이닝, 소셜 네트워크 분석 기법을 이용하여 빅데이터를 이해하고 추론 및 학습하는 과정이 요구된다(우지영, 2016). 시큐리티 인텔리전스 중 하나인 이상거래 탐지 시스템(FDS : Fraud Detection System)은 데이터 분석 기반의 보안시스템으로 이용자의 행동, 환경 등에 대한 데이터를 분석하고 모니터링하여 평소 패턴과 다른 거래임이 탐지될 경우 금융기관과 이용자에게 탐지사실을 알리고 나아가 임의로 거래를 중단시키는 것에 활용될 수 있다. 대표적으로 간편 결제 또는 금융결제를 수행하는 산업에서 이상거래 탐지 시스템이 요구되고 있으며 은행, 신용카드, 증권사에서 이상거래 탐지 시스템을 구축하고 있다.

관련 연구들을 살펴보면, 전병국 외(2014)는 사기를 탐지하기 위해 데이터마이닝 기법을 적용한 문헌들을 연구하고 이를 포괄적으로 분석할 수 있는 분류 체계를 제시하였다. 1997년부터 2014년 사이에 발매된 논문을 중심으로 연구를 진행하였으며 주로 은행사기, 보험사기, 금융사기, 보이

스 피싱, 이상치 탐지, 네트워크 침입 탐지, 모바일 뱅킹 사기 등을 예로 들 수 있다. 사기 탐지에 사용된 주된 알고리즘을 조사한 결과, 국내의 사기 탐지에 이용된 데이터마이닝 기법은 인공신경망(Neural Network), LOGIT이 가장 많이 사용되었으며 주로 기업사기와 거래사기, 은행사기를 적발하는 데에 많이 사용되었다. 그리고 국외의 경우에는 LOGIT이 가장 많이 사용되었으며 주로 보험사기, 기업사기 문제를 해결하는데 사용되었다는 것을 제시하였다. 정성훈 외(2015)는 전자결제 로그 데이터에 대한 분석과 데이터마이닝을 이용한 이상거래 탐지 기법에 대해 조사하였다. 이상거래 탐지 연구를 통해 데이터 셋, 알고리즘, 연구 관점으로 분류하여 연도별 연구 분포를 통해 최신 연구 경향을 제시하였다.

김승현 외(2015)는 현재 운영 중인 오프라인 거래 기반의 FDS에 온라인 신용카드 거래의 특성 및 부정사용 발생 패턴 등을 반영하여 온라인 전자상거래에 특화된 신용카드 전자상거래 FDS 탐지 기능 개선방안을 제안하였다. 박재훈 외(2015)는 국내 모 은행의 전자금융 사고 데이터 분석을 통해 이상 징후 패턴을 파악하고 탐지 룰을 설정하였으며, 설정된 룰을 기반으로 고객 개인별 거래 패턴과 고객 프로파일을 비교해 이상 거래 여부를 판단했다. 장기만 외(2015)는 연구개발비의 오남용 및 부정 집행을 방지하고자 연구개발비 집행계획 단계에서부터 정보를 수집하여 이상거래를 탐지할 뿐만 아니라 그 결과를 주관 연구기관, 전문기관, 신용카드사 간의 상호 실시간 연동으로 공유하여 활용할 수 있도록 하였다. 품질진단 기법 중 연구개발 관련 규정 및 매뉴얼, Q&A, 담당자 인터뷰 결과 등과 같은 정보를 수집하여 그 정보로부터 업무규칙을 도출하는 아웃사이드 인(Outside-in) 분석방법을 이용하였다. 도출된 업무규칙을 바탕으로 사전에 작성된 업무규칙에 적용, 이상거래를 탐지하도록 하였다.

우지영 외(2016)는 온라인 게임의 소액결제 사기를 방지하기 위해 게임 사용자의 결제 패턴을 분석하였다. 의사결정트리를 이용해 이상거래를 탐지할 수 있는 머신러닝 기반의 이상거래 탐지 모델을 제시하고, 이 모델을 특정 게임에 적용한 사례를 보여주었다. 신용카드나 은행에 도입되어온 이상거래 탐지 시스템을 게임 분야에 적용한 선두 연구라는 점에서 의의가 있었다.

이렇듯 실시간 로그 데이터를 반영한 온라인 음원차트의 경우에도 실시간으로 어뷰징성 트래픽의 유입을 방지할 수 있는 이상거래 탐지 시스템과 같은 시스템 구축이 필요하다고 사료된다.



4. 분석

4.1 데이터 수집

본 연구의 분석대상이 되는 A곡의 경우 발매 당일부터 이후 5일까지 국내 온라인 디지털 음원 유통 사이트에서 1위를 유지한 곡으로 어뷰징성 트래픽이 유입되었을 것으로 의심되는 곡이다. 1차 분석에 사용되는 데이터는 국내 온라인 디지털 음원 유통 사이트의 음원차트 집계에 반영된 신인 그룹의 A곡에 대한 앨범 발매 5일 후 일간 트래픽 로그 총 1,793,720개로 변수는 상품 구분, 접속 디바이스, 회원키(암호화), 재생시간, 접속 IP등으로 구성되어있다. 2차 분석에 사용되는 데이터는 A곡 발매 당일의 24시간 재생 로그 데이터 99,894,998건 중 전체 약 3.4%의 점유율을 차지하는 관련 로그 3,405,814개를 이용하였으며, 분석대상 앨범 발매일 전후 3개월간의 재생 로그 데이터를 추가하여 어뷰징 의심 계정을 추적하였다. 이후 의심 계정에 대한 사용자 활동이력 누적 데이터 분석을 통해 어뷰징 여부를 추정하였다.

<표 4-1>의 변수를 설명하면, 시간대의 경우 발매 당일 5일 후의 자정부터 1시간 간격으로 그 시간 범위 내의 각각 스트리밍과 다운로드, 종량⁶⁾에 해당하는 서비스 이용건수를 뜻한다. 총 재생건수는 스트리밍과 다운로드, 종량 세 값을 모두 더한 것으로 세 가지 서비스로 이용된 곡들의 총 건수를 나타낸다. 특정곡 재생건수는 본 연구의 분석에 사용되는

6) 음원의 가격을 매길 때 스트리밍 횟수나 내려 받은 수를 따르는 것

해당 특정곡의 재생건수를 나타내며 점유율은 총 재생건수 중 특정곡의 재생건수가 차지하는 비중을 나타낸다. 위 A곡의 경우를 살펴보면 시간대별 전체 재생건수 대비 점유율이 2%에 불과함에도 온라인 음원차트 순위에서 1위를 하는 것을 볼 수 있었다.

<표 4-1> 시간대별 전체 재생건수 대비 특정곡의 점유율(약 2%), 1위

시간대	스트리밍	다운로드	총량	총 재생건수	특정곡 재생건수	점유율
0:00~0:59	3,258,866	98,826	1,556	3,359,248	69,392	2.07%
1:00~1:59	2,827,273	73,564	1,745	2,902,582	58,422	2.01%
2:00~2:59	2,373,354	47,448	770	2,421,572	47,907	1.98%
3:00~3:59	1,975,503	34,107	192	2,009,802	39,587	1.97%
4:00~4:59	1,690,549	21,156	117	1,711,822	34,229	2.00%
5:00~5:59	1,536,311	16,273	100	1,552,684	32,158	2.07%
6:00~6:59	1,540,564	17,413	219	1,558,196	33,480	2.15%
7:00~7:59	1,736,045	32,155	240	1,768,440	40,065	2.27%
8:00~8:59	2,255,569	52,033	529	2,308,131	56,567	2.45%
9:00~9:59	2,858,981	72,690	774	2,932,445	72,738	2.48%
10:00~10:59	3,505,150	91,421	806	3,597,377	85,597	2.38%
11:00~11:59	4,071,401	90,026	773	4,162,200	93,304	2.24%
12:00~12:59	4,453,703	95,834	811	4,550,348	101,522	2.23%
13:00~13:59	4,699,148	103,435	1,085	4,803,668	101,818	2.12%
14:00~14:59	4,789,846	97,160	921	4,887,927	98,280	2.01%
15:00~15:59	4,869,516	93,757	990	4,964,263	97,600	1.97%
16:00~16:59	4,917,561	97,183	1,099	5,015,843	97,456	1.94%
17:00~17:59	4,782,694	92,921	897	4,876,512	91,793	1.88%
18:00~18:59	4,611,427	94,215	1,076	4,706,718	89,135	1.89%
19:00~19:59	4,628,657	108,380	1,551	4,738,588	91,048	1.92%
20:00~20:59	4,881,747	131,261	1,439	5,014,447	97,920	1.95%
21:00~21:59	4,698,817	148,153	2,135	4,849,105	95,484	1.97%
22:00~22:59	4,193,292	159,193	2,054	4,354,539	88,489	2.03%
23:00~23:59	3,693,167	160,591	1,936	3,855,694	79,729	2.07%
합계	84,849,141	2,029,195	23,815	86,902,151	1,793,720	2.06%

4.2 트래픽 로그 1차 분석

어뷰징을 파악하기 위해서 우선 이용자의 이용패턴을 구분할 필요가 있기 때문에 샘플 그룹을 설정하였다. 샘플 그룹을 설정하기 위해 먼저 특정 곡에 대한 시간대 별 재생건수와 회원수를 비교하였다. <표 4-2>는 본 연구에서 분석 대상이 되는 특정곡의 시간대별 전체 재생건수/이용자 대비 5회 이상 재생건수/이용자를 비교한 것이다. 특히 새벽시간대에 집중하여 시간당 5회 이상 이용자 비율 대비 재생건수 비율이 월등히 높음을 확인할 수 있었다.

시간당 5회 이상 재생한 이용자들 중 최대치는 시간당 19회 재생이었으며 <표 4-3>은 5~10회, 11~15회, 16회 이상 그룹별로 회원수 및 재생건수를 비교하였다. 16회 이상 재생한 회원수 및 재생건수 또한 마찬가지로 새벽시간대에 집중되어 있었으며, 05시 전체 이용자의 7.4%가 전체 재생건수의 45.7%를 차지함을 살펴볼 수 있었다. 따라서 05시 16회 이상 재생한 회원 869명을 샘플로 식별하여 일간(0~24시) 재생 이력을 추적하였다.

<표 4-4> 일간 재생 이력은 회원별로 최소 21건에서 최대 410건으로 이 또한 100건 이하, 101~200건, 201~300건, 301~400건, 401건 이상 다섯 그룹으로 나누어 회원수 및 재생건수를 비교하였다. 시간당 전체 평균 회원수의 1.9%인 869명의 재생건수가 전체 11%의 비중을 차지하였고, 이 중 200회 이상을 재생한 상위 52%(453명)가 재생건수의 73%를 차지하였다.⁷⁾

7) 곡 시간 3분 30초로 24시간 재생 시 약 410회를 재생할 수 있으므로 약 200회부터는 12시간 이상을 재생한 것으로 200회 이상 재생을 기준으로 설정

<표 4-2> 특정곡의 시간대별 전체 재생건수/이용자 비교

시간대	전체 재생건수	시간당 5회 이상 재생건수	5회 이상 재생비율	전체 회원수	시간당 5회 이상 회원수	5회 이상 회원비율
0:00~0:59	69,392	23,118	33.32%	39,709	2,121	5.34%
1:00~1:59	58,422	22,278	38.13%	31,383	1,876	5.98%
2:00~2:59	47,907	21,761	45.42%	23,127	1,737	7.51%
3:00~3:59	39,587	20,809	52.57%	16,933	1,554	9.18%
4:00~4:59	34,229	20,047	58.57%	13,150	1,458	11.09%
5:00~5:59	32,158	19,383	60.27%	11,830	1,395	11.79%
6:00~6:59	33,480	18,573	55.47%	13,591	1,351	9.94%
7:00~7:59	40,065	18,654	46.56%	19,024	1,434	7.54%
8:00~8:59	56,567	20,416	36.09%	31,523	1,755	5.57%
9:00~9:59	72,738	21,369	29.38%	44,067	2,004	4.55%
10:00~10:59	85,597	22,574	26.37%	54,249	2,184	4.03%
11:00~11:59	93,304	23,061	24.72%	60,203	2,262	3.76%
12:00~12:59	101,522	24,512	24.14%	65,566	2,463	3.76%
13:00~13:59	101,818	24,161	23.73%	66,337	2,426	3.66%
14:00~14:59	98,280	23,594	24.01%	64,175	2,354	3.67%
15:00~15:59	97,600	23,435	24.01%	63,638	2,296	3.61%
16:00~16:59	97,456	23,382	23.99%	63,737	2,270	3.56%
17:00~17:59	91,793	21,762	23.71%	60,297	2,080	3.45%
18:00~18:59	89,135	21,424	24.04%	58,614	2,059	3.51%
19:00~19:59	91,048	21,422	23.53%	59,888	2,066	3.45%
20:00~20:59	97,920	23,221	23.71%	64,022	2,277	3.56%
21:00~21:59	95,484	23,363	24.47%	61,568	2,298	3.73%
22:00~22:59	88,489	23,594	26.66%	55,502	2,280	4.11%
23:00~23:59	79,729	23,555	29.54%	48,340	2,265	4.69%
	1,793,720		33.43%			5.46%

<표 4-3> 시간대별 5회 이상 그룹별 회원수 및 재생건수

시간대	그룹별 회원수			그룹별 재생건수			전체 재생건수	16회 이상 재생건수 비율	전체 회원	16회 이상 재생 회원 비율
	5~10회	11~15회	16회 이상	5~10회	11~15회	16회 이상				
0	1,115	381	625	7,625	4,979	10,514	69,392	15.15%	39,709	1.57%
1	806	339	731	5,540	4,417	12,321	58,422	21.09%	31,383	2.33%
2	645	290	802	4,414	3,798	13,549	47,907	28.28%	23,127	3.47%
3	479	190	885	3,383	2,470	14,956	39,587	37.78%	16,933	5.23%
4	401	158	899	2,765	2,086	15,196	34,229	44.40%	13,150	6.84%
5	356	170	869	2,481	2,208	14,694	32,158	45.69%	11,830	7.35%
6	361	160	830	2,494	2,061	14,018	33,480	41.87%	13,591	6.11%
7	490	171	773	3,367	2,232	13,055	40,065	32.58%	19,024	4.06%
8	806	257	692	5,454	3,290	11,672	56,567	20.63%	31,523	2.20%

<표 4-4> 샘플로 선정한 869명의 그룹별 일간(0~24시) 재생 이력 추적

구분	100이하	101~200	201~300	301~400	401이상	샘플합계	전체 데이터
회원수	106	310	192	194	67	869	45,436
비율	12.20%	35.67%	22.09%	22.32%	7.71%		1.91%
재생 건수	7,613	45,854	48,279	68,507	27,179	197,432	1,793,720
비율	3.86%	23.23%	24.45%	34.70%	13.77%		11.01%
			200회 이상 = 일 12시간 이상 (73%)				

1차 분석에서는 어뷰징성 트래픽의 유입이 의심되는 특정곡을 대상으로 앨범 발매 직후 시간대별 재생건수/이용자를 비교하여 869명의 샘플을 추출하였다. 이중 일간 200회 이상 재생자의 이용패턴을 분석하여 의심 군으로 추정되는 유형을 1차로 식별하였다.

869명 중 200회 이상 재생자(453명)의 이용패턴을 분석하면 크게 네 유형으로 구분할 수 있다.

- 유형 1 : 24시간 무한 재생(스트리밍), 동일 디바이스, 동일 IP → 어뷰징 의심(방치)
- 유형 2 : 특정 시간대 집중하여 무한 재생(주로 새벽), 동일 디바이스, 동일 IP → 어뷰징 의심(방치)
- 유형 3 : 24시간 무한 재생, 디바이스 변경, IP동일 또는 변경 → 팬덤 추정(중간에 이용자 개입)

■ 유형 4 : 특정 시간대 집중하여 무한 재생, 디바이스 변경, IP 동일 또는 변경 → 팬덤 추정(중간에 이용자 개입)

유형을 예를 들어 설명하면, 유형 1에 해당하는 이용자의 경우 24시간 내내 동일 IP설정 하에서 동일 디바이스를 사용하여 특정 곡 스트리밍을 무한 재생한 것이다. 이 유형의 경우 대부분이 핸드폰 단말기를 사용한 것으로 확인되었으며 디바이스 특성상 24시간 내내 동일 IP설정 하에 있는 것은 어려우므로 디바이스 방치의 가능성을 유추하였다.

전체 869명 중 425명(49%)이 해당되는 유형 1과 2의 이용패턴을 분석해 보면 전체 재생건수 50%의 비중을 차지하였다. 주로 안드로이드에 집중해 있으며 이어서 아이폰, PC 플레이어, 아이패드, 맥킨토시 순이었다. 안드로이드, 아이폰을 제외하고 PC플레이어의 경우 대부분 고정된 장소라는 환경 특성을 보았을 때 팬덤일 소지가 있으며 나머지 디바이스 또한 미미한 수준이라 개인 팬덤으로 추정하였다.

4.3 계정별 프로파일 2차 분석

1차 분석 후 추가 데이터 셋 분석을 통해 의심스러운 이용패턴을 추가적으로 식별할 필요가 있었으며, 식별된 패턴에 해당하는 이용자 그룹을 대상으로 실제 어뷰징 여부를 판단하기 위해 2차 분석을 시행하게 되었다.

2차 분석에서는 연구 대상 A곡의 발매 당일('15. 9월 중순) 24시간 재생 로그 데이터 전체 99,894,998건 중 전체 약 3.4% 점유율에 해당하는 관련 로그 3,405,814건을 대상으로 곡, 앨범, 아티스트와 같은 콘텐츠 데이터와 계정, 접속 디바이스, IP, 재생 시간과 같은 사용자 재생 데이터로 구성되어있다.

<표 4-5>는 2015년 9월 00일 특정곡을 순도 100% 재생한 사용자 총 135명으로 이중 56명이 일 400건 이상, 24시간 동일 곡만 재생한 사용자에게 해당되었다. 이들 중 1차 분석의 유형 1에 해당하는 총 27명을 어뷰징 1차 의심군으로 식별하였다. 다음 <표 4-6>은 대표적인 예로 Top1 사용자의 발매 당일 재생이력을 보여주는 것으로 동일 디바이스, 동일 IP로 시간당 최대 18회를 재생하며 총 410건 재생하는 것을 확인할 수 있다.

<표 4-5> 발매 당일 특정곡 순도 100% 재생 이용자

재생건수	회원수	계정	앨범	건수	IP	디바이스ID
410	1	1	A 앨범	410	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
409	1	2	A 앨범	409	210.178.93.177	FE188C262B9F45BD872C6ECDBEA9F48D
408	4	3	A 앨범	408	121.142.202.120	개인정보 삭제
407	6	4	A 앨범	407	112.151.247.106	개인정보 삭제
406	7	5	A 앨범	407	125.178.162.5	개인정보 삭제
405	6	6	A 앨범	404	182.215.124.195	개인정보 삭제
404	10	7	A 앨범	403	59.13.47.170	개인정보 삭제
403	9	8	A 앨범	403	61.99.178.176	888888888888&&c114137b&&LGF100L&&
402	3	9	A 앨범	402	125.182.220.119	888888888888&&0b7322161f4d25fd&&LGV507L&&
401	4	10	A 앨범	402	222.101.45.17	개인정보 삭제
400	5	11	A 앨범	400	115.136.92.217	개인정보 삭제
399	2	12	A 앨범	400	221.153.66.173	개인정보 삭제
398	2	13	A 앨범	398	116.34.54.106	78DAF7C266FB43E58B6A1F491116655A
396	3	14	A 앨범	396	128.134.1.1	888888888888&&c1600a414b3d8af&&YPGP1&&
395	1	15	A 앨범	394	58.225.97.12	BAF7B571C012411BA10749565EDFCECB
• • •		16	A 앨범	392	1.253.196.189	개인정보 삭제
173	1	17	A 앨범	392	222.106.121.130	888888888888&&32041569038ac06f&&SHWM440S&&
57	1	18	A 앨범	390	211.247.98.237	개인정보 삭제
합계	135	19	A 앨범	389	123.254.226.191	개인정보 삭제
		20	A 앨범	387	116.36.92.167	개인정보 삭제
		21	A 앨범	387	14.51.55.127	개인정보 삭제
		22	A 앨범	385	183.105.222.92	개인정보 삭제
		23	A 앨범	378	113.130.177.194	개인정보 삭제
		24	A 앨범	374	115.161.1.109	개인정보 삭제
		25	A 앨범	372	49.174.86.212	SCSI\DISK&VEN_TOSHIBA&PROD_MQ01ABF050\4&37AC1C0D&0&0
		26	A 앨범	369	211.36.134.228	개인정보 삭제
		27	A 앨범	362	14.51.53.115	개인정보 삭제

<표 4-6> Top 1 이용자의 발매 당일 재생이력

계정	디바이스	곡	앨범	건수	시간	IP	디바이스 ID
1	IT40	A곡	A앨범	17	2015091500	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	18	2015091501	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091502	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091503	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091504	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091505	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091506	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091507	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091508	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091509	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091510	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091511	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091512	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091513	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	18	2015091514	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091515	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091516	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091517	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091518	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091519	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091520	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091521	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091522	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38
1	IT40	A곡	A앨범	17	2015091523	220.87.10.55	932CBDFD1B2E4A01944115F5ACC39D38

1차 의심군 27개 계정에 대해 연구 대상 앨범 발매 전후 3개월간의 재생 이력에 대한 추적 및 분석을 시행한다. 월간 재생 비율 100% 근접 여부, 월간 재생건수, 월간 기타 앨범 재생 여부(앨범수, 건수), 재생 지속성, 회원가입 시점, 타겟 앨범 발매 전 이력 등 이러한 속성에 따라 2차 분석에서 정의한 4개의 유형 중 3개 유형에 속하는 9개 계정을 2차 어뷰징 의심군으로 식별한다.

■ 유형 1 : 어뷰징 → 5개 계정(1, 5, 8, 11, 15)

- A 아티스트 앨범 발매 시점에 신규가입(이전 기록 없음)
- 4개월 간 순도 100% 해당 앨범만 재생(타 앨범 재생 이력 미미한 수준)
- 9~10월간 24시간 무한 재생 한도치에 육박한 수준(3분 30초 곡 기준 월 1.2만 건이 한도)

<표 4-7> 계정별 프로파일 : 2차 분석 - 유형 1에 해당하는 대표 계정 1의 예

9월			10월			11월			12월		
앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율
A아티스트	6136	100%	A아티스트	10588	100%	A아티스트	6217	100%	A아티스트	2324	62%
									A아티스트	1423	38%

■ 유형 2 : 어뷰징 → 3개 계정(4, 6, 21)

- 특정 타겟 앨범 발매 전 특정 아티스트 앨범 집중 현상(기타 앨범 재생 미미한 수준)
- 9월 타겟 앨범 발매 후 지속적으로 해당 앨범만 재생(월 24시간 재생 수준으로 타 아티스트 어뷰징 후 타겟 앨범 발매일에 어뷰징 대상을 전환 가능성)

<표 4-8> 계정별 프로파일 : 2차 분석 - 유형 2에 해당하는 대표 계정 4의 예

6월			7월			8월			9월			10월			11월			12월		
앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율
B아티스트	433	61%	B아티스트	9346	97%	B아티스트	691	33%	A아티스트	6501	81%	A아티스트	11965	100%	A아티스트	11341	100%	A아티스트	3981	67%
B아티스트	32	4%	B아티스트	77	1%	B아티스트	503	24%	A아티스트	831	10%							A아티스트	1976	33%
B아티스트	28	4%	B아티스트	76	1%	B아티스트	490	23%	A아티스트	627	8%									
B아티스트	26	4%	B아티스트	76	1%	B아티스트	110	5%	B아티스트	104	1%									
B아티스트	25	4%	B아티스트	76	1%	B아티스트	110	5%												
B아티스트	17	2%				B아티스트	109	5%												
B아티스트	11	2%				B아티스트	109	5%												

■ 유형 3 : 어뷰징 → 1개 계정(17)

- A 아티스트 앨범 발매 당일 24시간 재생(399회)외 전후 기록 없음(당일 순위 상승만을 목적으로 재생 후 종적을 감춘 것으로 보여짐)

■ 유형 4 : 팬덤 → 16개 계정(2, 3, 7, 9, 12, 13, 16, 18, 19, 22, 23, 24, 25, 26, 27)

- 특정 타겟 앨범 발매 전인 6~8월 특정 아티스트 앨범에 집중되거나 기타 앨범도 다양하게 재생하거나 전반적으로 편중 없이 재생
- 9월 타겟 앨범 발매 후 지속적으로 해당 앨범의 재생 비율이 높으나, 기타 앨범도 다양하게 재생
- 타 아티스트 팬덤이 타겟 아티스트 앨범 발매 후 집중 재생하는 것으로 보여짐

다음 <표 4-10>는 2차 의심군 9개 계정을 대상으로 해당 아티스트에 대한 사용자 누적 활동이력 데이터를 분석한 것으로, 일반적인 팬덤이라면 재생 활동외 사용자 활동이력이 높을 것이라는 전제 하에 팬덤으로서 재생 외의 일반적인 활동을 했는지 확인하였다. 재생 외 활동으로 플레이리스트 담기, 좋아요, 공유, 댓글 등을 고려하였으며 포토, 영상 조회이력은 아티스트의 순위에 반영되는 항목으로 고려대상에서 제외하였다.

결과적으로, 3개 계정의 어뷰징이 확실히 되었고 나머지 5개 계정은 어뷰징이 유력한 것으로 판단하였다.

<표 4-9> 계정별 프로파일 : 2차 분석 - 유형 4에 해당하는 대표 계정 2의 예

6월			7월			8월			9월			10월			11월			12월		
앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율	앨범	건수	비율
B아티스트	7585	87%	B아티스트	7386	80%	B아티스트	3372	69%	A아티스트	2146	83%	A아티스트	727	63%	A아티스트	4845	93%	A아티스트	1851	63%
B아티스트	860	10%	B아티스트	1141	12%	B아티스트	539	11%	A아티스트	46	2%	A아티스트	55	5%	기 타 76개	390	7%	A아티스트	575	20%
기 타 76개	231	3%	B아티스트	617	7%	B아티스트	354	7%	기 타 45개	400	15%	A아티스트	27	2%				기 타 84개	495	17%
			기 타 30개	70	2%	B아티스트	293	6%				기 타 135개	344	30%						
						기 타 17개	309	7%												

<표 4-10> 2차 의심군의 누적 활동이력

계정	총재생수	좋아요	포토	영상	playlist	공유	댓글	기타	팬맺기	어뷰징
1	35,128	15	38	13	0	0	1	26	Y	어뷰징 유력 : 재생수 대비 활동 저조
4	62,281	54	46	228	0	3	3	4	Y	어뷰징 유력 : 재생수 대비 활동 저조
5	-	0	0	0	0	0	0	0	-	어뷰징 확실 : 활동 없음
6	48,002	64	8,878	18	0	85	0	0	Y	팬덤으로 추정 : 활동 높은 편(공유)
8	27,388	66	151	18	0	0	0	0	Y	어뷰징 유력 : 재생수 대비 활동 저조
11	45,743	74	1,500	450	0	0	1	9	Y	어뷰징 유력 : 재생수 대비 활동 저조
15	45,569	61	38	27	0	0	0	0	Y	어뷰징 유력 : 재생수 대비 활동 저조
17	399	2	0	0	0	0	0	0	Y	어뷰징 확실 : 활동 거의 없음
21	3,620	0	0	0	0	0	0	0	N	어뷰징 확실 : 활동 없음

5. 결론

5.1 요약 및 결론

본 논문에서는 국내 온라인 음원차트의 어뷰징 실태에 대해 조사하였고, 어뷰징 여부를 판단할 수 있는 근거 속성들을 구분하였다. 먼저 어뷰징성 트랙의 유입이 의심되는 곡을 대상으로 국내 온라인 음원 유통 사이트의 트랙 로그 데이터를 이용하여 1차적으로 의심패턴 유형을 식별하였다. 이를 기반으로 2차 분석에서는 계정별 프로파일을 통해 어뷰징 의심군 유형을 2차로 식별하였다. 이 과정에서 일반적인 이용자의 재생패턴과 어뷰징 의심군의 재생패턴을 비교하여 일부 속성들에서 특이점을 발견하였다.

본 연구 결과, 어뷰징 여부를 판단하기 위해서는 다음 <표 5-1>과 같은 다양한 속성을 활용해야 할 것으로 제시하였다.

<표 5-1> 어뷰징 여부 판단 패턴

번호	평가 속성	기준	계정
1	발매일 특정 앨범 재생 비율 100% 근접 여부	특정 곡 몰아주기 의도로 추정	계정 1, 5, 8, 11, 15
2	특정일 24시간 동일 IP, 동일 디바이스 재생	사용자 개입 없는 것으로 추정	모든 계정 해당
3	발매 후 3개월 일정기간 특정 앨범 재생 비율	특정 곡 몰아주기 의도로 추정	모든 계정 해당
4	재생 횟수의 수준	월 한도치 근접 여부(1.2만여건)	2차 분석 : 유형 1, 2, 3
5	같은 기간 기타 앨범 재생 여부(앨범수, 건수 기준)	기타 앨범도 재생 → 팬덤으로 추정	2차 분석 유형 4 제외
6	재생 지속성	발매 후 2~3개월 지속되거나 당일 등 짧은 기간 집중된 경우	모든 계정 해당
7	계정가입시점	타겟 앨범 발매 시점	2차 분석 : 유형1, 계정 17
8	타겟 앨범 발매 전 이력	이전 이력이 존재하지 않거나 타 앨범 집중 여부 타 앨범 어뷰징 후 해당 앨범으로 타겟 전환	계정 4, 6, 21
9	재생 횟수 대비 기타 활동이력 (좋아요, 팬 맺기, 공유, 댓글, 플레이리스트 담기 등)	없거나 저조한 수준	계정 6 제외
비고	의심 계정의 포토, 영상 조회 이력은 고려대상에서 제외	해당 속성도 아티스트 순위를 올리기 위한 어뷰징 가능성 있음	

5.2 연구의 한계점 및 향후 연구방향

본 연구에서는 국내 음원 유통 사이트의 트래픽 로그 데이터를 통해 어뷰징 의심군 패턴을 도출해냈지만 이들이 정말 어뷰징에 해당하는지에 대한 적합성 확인이 필요하다. 음원차트 어뷰징의 경우 실시간 집계되어 반영된다는 점과 그 수법이 점차 고도화되고 있다는 점에서 적시에 어뷰징성 트래픽을 구분하는 것이 어려울 수 있다. 앞서 도출된 결과를 다른 곡에 적용하여 어뷰징 의심군을 추출하고 관련 실무자의 확인을 통해 의심군의 어뷰징 여부를 확인하는 과정이 필요할 것이다.

또한 앞서 확인한 10가지 속성에 대한 추가 데이터 검증을 통해 정규화된 어뷰징 모델 개발이 필요할 것이다. 어뷰징 탐지 룰의 효과적인 적용을 위해 추가적으로 어뷰징 데이터를 수집하여 실험 데이터를 생성하고 분석 결과의 검증을 위해 데이터 셋을 모형 구축 데이터와 모형 검증 데이터로 분류한다. 이를 바탕으로 Decision Tree 등과 같은 데이터마이닝 기법을 통해 분석 결과를 도출하고, 모형 검증 데이터를 통해 특정 계정의 어뷰징 여부가 일치하는지를 확인해야 한다.

그러나 어뷰징과 팬덤 간의 식별은 매우 어려우며 어뷰징 수법은 지속적으로 고도화되고 있는 실정이라 어뷰징 모델이 제시하는 확률을 기반으로 최종적으로는 관련 실무자의 판단이 필요하다. 해킹과 유사하게 어뷰징 수법 또한 갈수록 진화하고 있어 지속적인 모니터링을 통한 어뷰징 모델의 업데이트가 필요할 것이다.

참 고 문 헌

Pricewaterhouse Coopers(이하 PwC). (2014). <Global entertainment and media outlook 2009-2018>

Pricewaterhouse Coopers(이하 PwC). (2015). <Global entertainment and media outlook 2010-2019>

김병희. (2015). 포털의 검색 알고리즘 개발에서 기사 어뷰징 방지를 위한 영향요인 탐색. *커뮤니케이션 이론*, 11(3), 47-89.

김승현, 김휘강, & 김은진. (2015). 전자상거래 신용카드 부정사용 사례 분석을 통한 FDS 효과성 향상에 관한 연구. [(A)Study on the Improvement of FDS Effectiveness over the Case Analysis on E-commerce Credit-card Fraud-to-sales]. *한국지식정보기술학회 논문지*, 10(6), 723-734.

박세희. (2016.3.3). 팬들의 음원 사재기를 법으로 막는게 효과가 있을까. <허핑턴포스트코리아>. URL : http://www.huffingtonpost.kr/2016/03/03/story_n_9371092.html

박요한. (2014). 데이터 마이닝을 이용한 정부보조금 부정수급 예측 사례 연구. *2014 한국정보과학회 제 41 회 정기총회 및 동계학술발표회*, 144-146.

박재훈, 김휘강, & 김은진. (2015). 의사결정나무를 이용한 이상금융거래

- 탐지 정규화 방법에 관한 연구. *정보보호학회논문지*, 25(1), 133-146.
- 우지영, 김하나, 박병일, & 김휘강. (2016). 온라인 게임 결제 데이터 분석 기반의 이상거래 탐지 모델. *정보보호학회지*, 26(3), 38-44.
- 윤샘이나. (2015.10.12). JTBC ‘음원 사재기 의혹’...이달의 기자상 수상. <JTBC>. URL : <http://bit.ly/2d7pZFY>
- 이선아, & 장남식. (2015). 데이터마이닝을 이용한 허위거래 예측 모형. *지능정보연구*, 21(1), 161-177.
- 이종완. (2015.10.2). 음원사재기 논란, 멜론이 음반기획사에 보낸 공문. <위키�트리>. URL : <http://i.wik.im/234558>
- 임성준, & 백운선. (2004). 한국 음반제작시스템에 관한 탐색적 연구. *전략경영연구*, 7(1), 1-32.
- 장기만, 김창수, & 정희경. (2015). 데이터 품질진단 기법을 이용한 연구 개발비 이상거래 실시간 탐지. *한국정보통신학회논문지*, 19(11), 2609-2614.
- 전병국, 김태훈, 최소윤, & 안현철. (2014). 사기 탐지에 대한 데이터 마이닝 문헌 연구와 분류 프레임워크. *한국지능정보시스템학회 2014 년 추계학술대회*, 277-283.
- 정성훈, 김하나, 신영상, 이태진, & 김휘강. (2015). 결제로그 분석 및 데이터 마이닝을 이용한 이상거래 탐지 연구 조사. *정보보호학회 논문지*, 25(6), 1525-1540.

- Adler, M. (1985). Stardom and talent. *The American Economic Review*, 75(1), 208–212.
- Chevalier, J., & Goolsbee, A. (2003). Measuring prices and price competition online: Amazon. com and BarnesandNoble. com. *Quantitative marketing and Economics*, 1(2), 203–222.
- Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *The quarterly journal of economics*, 183–207.
- Rosen, S. (1981). The economics of superstars. *The American Economic Review*, 71(5), 845–858.
- Spoerri, A. (2008). Authority and ranking effects in data fusion. *Journal of the American society for information science and technology*, 59(3), 450–460.
- Strobl, E. A., & Tucker, C. (2000). The dynamics of chart success in the UK pre-recorded popular music industry. *Journal of Cultural Economics*, 24(2), 113–134.
- Towse, R. (1992). The earnings of singers: An economic analysis *Cultural economics* (pp. 209–217): Springer.

Abstract

An empirical study on the effecting factors of online music chart abusing

Gwak, A Yeong

Major in Management Information System

Dept. of Business Administration

The Graduate School

The University of Seoul

As the online music service market has expanded and music consumption channels have shifted from offline media to digital media, the importance of music charts has been increasing. Thus, there was a music chart abusing operation, which is an act of illegally purchasing a specific sound source for manipulating the music chart rank or earning the royalty fee. It is an act that should be eradicated, because it can lead to distorted information to the consumer and to the reduction of copyright revenue and further deprivation of broadcasting opportunities if not a music chart abusing.

As a measure, major music sites reflected '1 day, 1 account, 1

count policy' in the chart aggregation, but it is impossible to disable this through a large number of accounts, and automatic macro development considering streaming service is becoming more intelligent and advanced.

Therefore, in this paper, I used the data reflected in the real-time sound source chart aggregation of the domestic sound source distribution site to derive the abusing judgment pattern. Specifically, the types of the abusing group are classified and based on this, I offer attributes for judging abusing by the customer profile.

Keyword : Music market, Streaming service, Music chart abusing