

흥행영화의 온라인 구전패턴과 관객수의 관계에 대한 실증연구 An Empirical Study on the Relationship between the Pnline WOMs and the Number of Audience of Successful Films

황예나, 남윤재

경희대학교 문화관광콘텐츠학과

Yena Hwang(vlel91@nate.com), Yoonjae Nam(ynam@khu.ac.kr)

요약

본 연구는 흥행영화의 온라인 구전패턴을 확인하고, 온라인 구전과 관객수 간의 관계를 실증분석 했다. 분석 대상으로 삼은 흥행영화는 2017년 한국에서 개봉한 영화 중 누적 관객수 기준으로 상위 20위까지의 영화이며, 본 영화에서 온라인 구전으로 도출한 변인은 각 20개 영화에 대한 블로그 포스팅, 기사, 리뷰, 그리고 검색의 양(volume)이다. 본 연구의 분석결과는 다음과 같다. 첫째, 온라인 구전의 패턴을 확인하기 위해 SPSS의 곡선추정으로 한 결과, 모든 변인이 3차 모형으로 가장 잘 설명된다고 나타났다. 둘째, 온라인 구전과 관객수와 의 관계를 보기 위해 패널 분석을 한 결과, 블로그양, 리뷰양, 검색량은 관객수와 모두 정의 관계에 있는 것으로 나타났으나 기사량은 부의 관계에 있는 것으로 도출되었다.

■ 중심어 : | 흥행영화 | 온라인 구전 | 관객수 | 곡선추정 | 패널분석 |

Abstract

This study investigates the relationship between the online WOMs(such as volume of blogs, articles, reviews, searches) and the number of audience of successful film.The results are as follow: Frist, using a curve-estimation method, the results show that the longitudinal trends of the online WOMs can be best described by a cubic indicating. Second, using panel analysis in model(t) the volume of blogs, reviews, and searches is positively associated with the number of audience. All of the variables' coefficient are significant. However the volume of articles is negatively related to the number of audience with a significant coefficient.

■ keyword : | Film | Online WOM | Curve Estimation | Audiences | Panel Analysis |

1. 서론

영화는 대표적인 영상문화콘텐츠이며 콘텐츠 산업은 창의성과 같이 무형자산을 생산요소로 하는 고부가가치 산업이다[1]. 콘텐츠산업의 개념은 1998년 OECD의 '신(新)성장 산업으로서 콘텐츠'라는 언급을 통해 처음 소개되었고, 이후 하위 분야로 디지털콘텐츠의 중요성이 부각되었다[2]. 특히 사회경제구조가 산업경제와 지식경제의 형태에서 창조경제의 중심으로 변화됨에

따라 문화콘텐츠의 중요성이 커지고 있고, 영화는 문화콘텐츠의 대표적인 영역으로 자리매김했다. 미국영화협회(MPAA)의 자료에 의하면 2016년 한국 영화시장은 세계 6위를 기록하였으며, 2017년도 기준 국민 연평균 영화관람 횟수는 4.25회로 최고 수준을 유지하고 있다 [3].

영화는 인기 있는 소수 영화로 상업적인 성공이 집중되며 긴 생산단계에 비해 짧은 판매기간을 갖는다. 때문에 영화 마케팅은 짧은 판매 기간 안에 많은 관객이

영화를 통해 만족을 얻어 수익을 극대화하는 것에 초점을 맞추게 된다[4]. 특히 영화는 관람하기 전까지 그 특성을 알 수 없고 관람하고 나서야 평가가 가능하다는 경험재적 특성을 갖기 때문에 소비자의 광고나 홍보 등의 정보의존도를 높이는데, 이때 소비자는 선택의 불확실성을 줄이기 위해 영화에 대한 사전 정보 탐색 과정을 거친다[5]. 이 사전정보탐색의 과정에서 소비자는 영화의 구전 정보를 활용하게 된다.

고정민(2002)은 한국영화 성장의 주요 요인 중 하나로 인터넷을 통한 영화정보의 노출과 구전에 의한 정보유통을 강조했는데, 포털 사이트나 영화 관련 커뮤니티에서 공유되는 리뷰와 영화 평점의 영향력은 갈수록 커진다고 하였다[6]. 영화진흥위원회(이하 영진위)의 2016년 영화 소비자 조사 보고서에 따르면¹⁾ 소비자들의 영화 선정 시 주 참조 정보원은 인터넷이 59.6%로 1위를 차지하였으며, 그 다음이 주변인의 의견(58.6%)으로 나타나 영화 선택시 구전정보의 중요성을 뒷받침했다[7]. 최근에는 소셜네트워크서비스(SNS)를 통한 온라인 구전 커뮤니케이션이 활발하게 이루어지고 있기 때문에 온라인 구전이 영화에 관객수에 어떤 영향을 미치는지를 연구하는 것은 영화마케팅 측면에서나 영화의 산업적 성공이라는 측면에서도 중요한 시사점을 제공할 것이라고 본다.

이러한 맥락에서 본 연구는 흥행한 영화를 대상으로 온라인에서 수집할 수 있는 영화에 대한 데이터들을 영화의 구전으로 측정하고, 이들이 관객수와 어떤 관계가 있는지 알아보려 한다. 온라인 구전과 관객수의 관계에 대한 분석에 앞서 본 연구에서는 영화의 온라인 구전이 어떤 패턴을 가지고 있는지를 확인하고자 한다. 영화는 음악, 게임 등 타 콘텐츠들과 다르게 유통에 있어 상영기간이라는 기한이 정해져 있으며, 보통 개봉 첫 주 관객수에 의해 영화의 성패가 결정되는 특징을 가지고 있기 때문에 온라인 구전변수가 개봉일을 기준으로 어떤 패턴으로 나타나는지 확인해 보는 것은 관객수와의 관계를 해석하는 데 필요한 과정으로 보았다.

따라서 본 연구에서는 흥행영화의 온라인 구전을 선정하고 그들의 패턴을 확인하며, 온라인 구전과 관객수

의 관계를 실증분석하여 영화 마케팅 측면 혹은 영화선택의 정보원으로서 구전활용의 방향성에 대한 시사점을 제공하고자 한다.

II. 이론적 배경

1. 영화흥행에 관한 논의

영화는 문학적, 예술적, 상품적인 특성뿐만 아니라 사회적으로 복합적인 특성을 지닌 창작물이다[8]. 산업적인 측면에서 영화는 다른 예술분야와 비교했을 때 제작에 오랜 시간이 걸리고, 영화흥행을 매개로 하는 하나의 단일 프로젝트로서, 투자-제작-배급이라는 구조를 가진 사업이다.

영화의 성공에 대해 한마디로 정의하기란 어렵지만, 일반적으로 작품성과 흥행성 두 가지로 나눠 설명된다. 전자는 예술영화 혹은 독립영화에서 관심을 두고 있으며, 후자는 상업영화에서 초점을 두고 있다. 영화흥행의 측정으로는 크게 매출액과 관객수를 기준으로 삼고 있는데, 영화 산업의 세분화를 반영한 매출액을 기준으로 하는 것이 분석에 더 합리적이라는 의견이 있지만, 한국에서는 통상적으로 영진위에서 운영하는 영화관 입장권 통합 전산망 시스템의 관객수를 흥행의 기준으로 삼는다.

미디어 경제학의 분야로서 영화의 흥행에 대한 연구는 1970년대부터 진행되어 왔지만, 경제적 접근방법의 이론적 틀을 만들고 영화흥행 요소에 대한 연구를 본격화시킨 것은 Litman(1983)의 연구부터라 할 수 있다. Litman(1983)은 영화의 성공을 결정짓는 중요 세 가지 영역으로 창조적 영역, 편성과 배급 그리고 마케팅 노력을 뽑았다[9]. 창조적 영역에는 감독, 배우, 등급, 장르, 제작비 등이 포함되고, 편성과 배급은 영화의 개봉 시기 및 배급사의 규모, 그리고 스크린 수가, 마지막 마케팅 영역에는 영화의 평점 및 수상여부가 포함된다. 국내에서는 1990년대 후반부터 영화흥행에 관련된 연구가 본격화되기 시작했는데, 대부분 Litman(1983)의 연구모형을 기반으로 어떤 요인이 흥행에 영향을 미치는지에 관련한 연구가 주를 이룬다. 다음 [표 1]은 Litman의 연구를 비롯해 영화흥행 요소에 대한 국내외 선행연구를 정리한 것이다[9-20].

1) 전국 17개 시도(2015년부터 세종시 포함)에 거주하는 만 15세 이상 59세 이하 소비자들을 대상으로 조사.

표 1. 영화흥행 요소에 대한 국내·외 선행연구

국외 연구			
연구자	분석영화 편수	흥행에착요인	
Litman 1983 [9]	1970년대 개봉한 155편의 영화	제작비, 개봉스크린 수, 전문가 평점, 배급사, 장르, 수상실적	
Litman 1989 [10]	1981~1986년에 개봉한 697편의 영화	제작비, 개봉스크린 수, 전문가 평점, 배급사, 수상실적, 장르, 개봉시기, 배우파워	
Wyatt Badger 1990 [11]	1983~1986년에 개봉한 영화 중 1백만 불 이상의 매출액을 기록한 512편의 영화	제작비, 전문가 평점, 배우 파워	
Sochay 1994 [12]	1980년대 후반 개봉한 263편의 영화	전문가 평점, 수상경력, 배우 파워, 스크린 수, 개봉시기, 장르	
Ravid 1999 [13]	1990년대 초반 개봉한 영화 180편	제작비, 전문가 평점, 등급	
Chang Ki 2005 [14]	2000~2002년에 미국에서 개봉한 영화 중 1백만 불 이상의 매출액을 기록한 463편의 영화	속편, 배우파워, 제작비, 장르, 등급, 상영기간, 개봉스크린 수	
국내 연구			
연구자	분석영화 편수	흥행에착요인	
김홍중 1998 [15]	1988년부터 10년간 개봉한 637편의 한국영화	개봉스크린 수, 수상경력, 배우파워, 감독파워	
유현석 2002 [16]	1988~1999년에 개봉한 732편의 한국영화	감독파워, 배우파워, 제작사 파워, 장르, 등급	
김은미 2003 [17]	1995~2000년에 개봉한 한국영화 318편, 1996~1997년에 상영된 350편의 할리우드 영화	제작비, 개봉스크린 수, 장르, 등급, 감독파워	제작비 고려
최은희 외 3명 2009 [18]	2005년 말~2008년에 상영된 영화 중 리뷰수 100건 이상인 229편의 영화	개봉스크린 수, 배급사 파워	
김소영 임승희 정예슬 2010 [19]	2008년 국내 개봉된 379편의 영화	스크린 수, 관객 평가, 장르	사업 영화
박승현 송현주 정완규 2011 [20]	2009~2010년에 개봉된 한국영화	감독파워, 배우파워, 네티즌 리뷰 양	예술 영화
		스크린 수, 감독파워, 등급, 관객 평가, 전문가 평가, 배급사 파워, 제작국, 개봉 시기, 장르	예술 영화
		온라인 리뷰 양, 제작비, 배급사 파워, 상영등급	고예산
		배우파워, 개봉스크린 수, 배급사 파워	저예산

이상의 선행연구를 보면 연구마다 약간의 차이는 있지만 스크린 수, 제작비, 감독 혹은 배우의 파워, 전문가 평점, 개봉시기, 장르, 등급 등이 흥행에 영향을 주는 요인으로 나타났다.

2. 구전에 관한 논의

구전(Word-of-Mouth)은 마케팅 연구에서 처음 사용된 이래로 연구 분야와 연구자마다 차이는 있지만, 사람들의 입에서 입으로 전달되는 정보라는 뜻으로 사용된다

[21]. 최근에는 인터넷, 모바일과 같은 새로운 매체의 등장으로 인해 온라인을 통한 구전의 중요성이 대두되고 있는데, 온라인 구전이란 인터넷과 같은 가상공간에서 이루어지는 커뮤니케이션을 의미한다. 각종 커뮤니티, 웹사이트 게시판, 온라인 쇼핑몰 리뷰, 쇼핑 후기, 블로그 포스팅, 평점 등이 온라인 구전의 예라고 할 수 있다.

Henning-Thurau 외 3명(2004)은 온라인 구전에 대해 제품과 사회에 대한 잠재적, 실재적, 혹은 과거의 고객에 의한 긍정적이나 부정적인 언급이라고 정의하였으며[22][23], 온라인 구전은 그 익명성과 시간 및 공간의 비제약성 때문에 오프라인 구전에 비해 자유롭고 편리한 커뮤니케이션이라고 했다. Bickart, Schindler (2001)은 소비자 사이에서 발생하는 제품에 대한 정보, 사용 후기, 추천과 같은 정보들의 교환이 인터넷이라는 공간을 통해 커뮤니케이션 되는 것이 온라인 구전이라고 정의했다[24]. Senecal 과 Nantel(2004)은 다른 제품에 비해 경험재를 구매하고자 할 때 소비자는 온라인 구전에 더 의존한다고 하였다[25].

구전은 영화흥행에 영향을 미치는 주요 요인으로 여러 연구에서 언급되어 왔는데, Henning-Thurau 외 3명(2004)은 관람객의 리뷰가 영화흥행에 중요 결정요인이며, 특히 평점은 첫 주의 관객수와 총 관객수에 영향을 준다고 밝혔다[22]. 정영호(2013)는 영화 선택 시 소비자들은 온라인 구전을 중요하게 고려하고, 다수의 의견이 제시된 경우 영화 선택에 긍정적 영향을 미친다고 말했다[26].

영화와 같이 관여도가 높은 상품일수록 구전의 수용과 확산이 활발하다는 연구결과도 존재한다[27][28]. 정재학과 김영찬(2004)의 연구에 따르면 영화는 소비자들이 소비를 선택하기까지의 과정을 가장 잘 기억하는 상품이며, 매스 커뮤니케이션보다 구전 커뮤니케이션이 영화관람 선택 시 더 영향을 미친다고 밝혔다[29].

영화와 구전의 관계에서도 구전 정보원의 전문성과 관련된 연구들이 자주 연구되었는데, Kim, Park, Park(2013)은 영화는 유형의 제품과 달리 경험을 통해 소비되고 쾌락적 소비의 특성을 보이기 때문에 소비자들은 전문가보다 일반인 후기에 주목한다고 하였으며 [30], 김세준과 이세진(2013)의 연구에서도 소비자들은 평론가의 리뷰보다는 일반인 리뷰에 더 긍정적인 반응을 보였다[31]. Zufryden(1996)은 구전이 개봉 초기의

영화 인지도에 유의미한 영향을 미치며 영화가 상영되는 내내 지속적인 영향력을 갖는다고 밝혔다[32].

영화흥행과 구전의 연구에서는 구전의 방향성(valence)과 구전의 양(volume)을 나누어 분석한 연구들이 있는데, 최근 연구에서는 구전의 방향성은 영화흥행에 유의미한 영향을 미치지 못하고 구전의 양만이 유의미한 영향을 미친다는 결과가 많았다. Wyatt와 Badger(1984)는 실험을 통해 영화 리뷰의 긍/부정의 방향성에 따라 관객의 흥미와 평가가 달라진다는 것을 실증적으로 밝혔는데[33], 후속 연구[11]에서는 대중매체에 실린 비평의 방향성보다는 정보의 양이 관객에게 더 큰 영향을 미친다고 하였다. Duan, Gu, Whiston(Jun. 2008)은 2003년과 2004년의 개봉 첫 1, 2주간의 일일 구전효과 측정치와 일일 흥행 수입 데이터로 상관관계를 분석하였는데, 분석결과 영화흥행 수입과 구전의 방향성은 구전의 양에 유의미한 영향을 미치며, 구전의 양에 의해 영화흥행 수입이 증가함을 밝혔다[34]. Liu(2006)은 포털사이트 Yahoo Movie의 구전 정보를 사용하여 구전 정보가 영화수입에 영향을 주는지를 실증 분석하였는데, 영화에 대한 구전은 개봉 전과 개봉 첫 주에 가장 활발하며 그 이후에는 줄어드는 경향이 있음을 밝혔다. 특히 영화흥행에 영향을 주는 것은 구전의 양이지 방향성에 의한 것은 아니라고 밝혔다[35]. Duan, Gu, Whiston(Nov. 2008)은 영화 일일 흥행성적에 대한 온라인 리뷰의 설득력과 인지효과를 분석하였는데, 리뷰의 방향성은 영화흥행에 영향을 미치지 않았으나 리뷰의 양은 영향을 미친다고 나타났다[36]. 김연형과 홍정환(2011)의 연구에서는 영화 포털사이트의 평점은 영화 흥행에 유의미한 영향을 주지 않았으나, 블로그 포스팅 같은 소셜 미디어의 버즈량은 흥행에 유의미한 영향을 미치는 것으로 나타났다[37]. Jeon과 Son(2016)도 영화 흥행에 미치는 온라인 구전 변수들의 영향을 분석했는데, 온라인 구전의 방향을 나타내는 평점 보다 온라인 구전의 크기를 나타내는 포털의 평가자 수, 블로그 수, 뉴스 수가 영화 흥행과 더욱 연관성이 있다는 결과가 나타났다[38].

이 밖에 영화관람의 시기별로 집단을 구분하여 구전이 영화선택에 미치는 영향에 관해 검증한 연구에서는 영화 초기 관객보다는 후기 관객들이 구전에 더 많은

영향을 받으며, 구전을 통해 영화를 선택한 관객의 만족도는 높은 것으로 나타났다[39-41].

III. 연구방법

1. 연구설계

1.1 연구대상

본 연구에서는 온라인 영화정보의 구전패턴을 확인하고, 구전패턴이 영화의 관객수와 어떤 관계가 있는지 알아보려고 한다. 분석을 위해 사용한 흥행영화는 작년(2017년) 한 해 동안 한국에서 개봉한 영화 중 누적 관객수 상위 20위까지의 영화로 선정하였으며, 그 대상은 다음 [표 2]와 같다.

관객수 상위 20위까지의 영화를 분석 대상으로 선택한 이유는 한국에서는 관객수를 흥행지표의 기준으로 삼고 있고, 기대치가 낮은 영화는 기대치가 높은 영화보다 구전을 활용하기 힘들다[42]는 Moul(2007)의 연구를 바탕으로 기대치가 높은, 다시 말해 흥행이 기대되는 영화에서 구전효과의 측정이 더 쉬울 것으로 예상되기 때문이다.

표 2. 2017년 관객수 상위 20위 영화

순위	영화명	개봉일	국적	관객수
1	택시운전사	08-02	한국	12,099,614
2	신과함께-죄와 벌	12-20	한국	8,539,600
3	공조	01-18	한국	7,817,459
4	스파이더맨: 홈 커밍	07-05	미국	7,257,973
5	범죄도시	10-03	한국	6,851,300
6	군함도	07-26	한국	6,592,151
7	청년경찰	08-09	한국	5,636,259
8	더 킹	01-18	한국	5,316,015
9	미녀와 야수	03-16	미국	5,137,995
10	킹스맨: 골든 서클	09-27	미국	4,945,484
11	토르: 라그나로크	10-25	미국	4,853,778
12	꾼	11-22	한국	4,018,035
13	강철비	12-14	한국	4,012,891
14	남한산성	10-03	한국	3,829,891
15	미이라	06-06	미국	3,689,325
16	분노의 질주: 더 익스트림	04-12	미국	3,653,238
17	너의 이름은.	01-04	일본	3,637,113
18	슈퍼배드 3	07-26	미국	3,324,861
19	아이 캔 스피크	09-21	한국	3,279,296
20	캐리비안의 해적: 죽은 자는 말이 없다	05-24	미국	3,049,894

1.2 연구모형

본 연구에서 연구모형은 온라인 구전이 관객수와 어떤 관계가 있는지를 확인하기 위한 패널 분석에서 사용하였다. 흥행영화의 온라인 구전과 관객수의 관계를 알아보기 위해 다음 식(1)과 같은 연구모형을 설정하였다.

$$Audience_{it} = \alpha + \beta_1 Blog_{it} + \beta_2 News_{it} + \beta_3 Reply_{it} + \beta_4 Searchvolume_{it} + \epsilon_{it} + u_i \quad (1)$$

(i = 영화; t = 개봉일, ..., $t + 30$)

식(1)은 본 연구에서 알아보고자 하는 온라인 구전변수들이 관객수에 미치는 영향을 알아보기 위한 기본 모형으로 모델(t)라 하며, 종속변수 $Audience_{it}$ 는 i 영화에 대한 t 시점의 일일 관객수를 의미한다. 첫 번째 변수인 $Blog_{it}$ 는 블로그양으로 t 시점에 i 영화에 대한 블로그 포스팅의 개수로 측정하였다. 두 번째 변수인 $News_{it}$ 는 기사량으로 t 시점에 i 영화의 인터넷 기사수를 의미하며, 세 번째 변수인 $Reply_{it}$ 는 리뷰양으로 t 시점에 i 영화의 리뷰 개수를, 마지막 변수인 $Searchvolume_{it}$ 는 검색량으로 t 시점에 i 영화에 대한 검색의 상대적 수치를 의미한다. ϵ_{it} 는 시간에 따라 변화하는 오차항을, u_i 는 시간에 따라 변하지 않는 오차항을 의미하며, α 는 절편을 나타낸다.

패널 모형에서 t 는 관측 시점을 의미하며 일반적으로 연도나 분기를 의미한다. 하지만 본 연구의 분석 대상은 흥행영화이고, 영화는 각기 다른 개봉일을 갖기 때문에 여기에서 t 는 동일한 관측 시점을 의미하는 것이 아니라 i 영화의 개봉 일부터 개봉 후 30일까지의 총 31일이라는 기간을 의미한다고 가정한다.

1.3 연구문제 및 가설설정

본 연구에서 알아보고자 하는 연구문제는 크게 두 가지다. 우선 흥행영화의 온라인 구전 변수들을 추출하고, 구전패턴을 알아보고자 한다. 두 번째로 온라인 구전패턴이 관객수와 어떤 관계가 있는지 알아보고자 한다.

따라서 본 연구에서는 다음과 같은 연구문제와 가설을 설정하였다.

연구문제 1. 온라인 영화정보의 구전패턴은 어떠한가?

연구문제 2. 온라인 영화정보는 관객수와 어떤 관계가 있는가?

가설2-1: 블로그양은 관객수와 정(+)의 관계에 있을 것이다.

가설2-2: 기사량은 관객수와 정(+)의 관계에 있을 것이다.

가설2-3: 리뷰양은 관객수와 정(+)의 관계에 있을 것이다.

가설2-4: 검색량은 관객수와 정(+)의 관계에 있을 것이다.

2. 데이터

2.1 변수의 조작적 정의

본 연구에서 분석을 위해 사용한 흥행영화의 온라인 구전 변수들은 각 영화에 대한 블로그, 기사, 리뷰, 그리고 검색 데이터이다.

임현정, 양희동, 백현미(2014)는 소셜 미디어 매체로서 블로그 게시글과 온라인 뉴스를 선택하고 이들과 영화 매출액과의 관계에 대한 실증연구를 통해 온라인 구전효과를 밝혔다. 분석결과 개봉 후에는 블로그가 온라인 뉴스 보다 영화의 매출액에 더 영향력이 큰 것으로 나타났다[43]. 이수범과 강은희(2010)는 블로그를 이용하는 20대와 30대를 대상으로 온라인과 오프라인 설문 조사를 하여 블로그를 이용한 영화 마케팅 수용자 인식을 조사하였다. 분석결과 과반수 이상의 응답자들이 영화 후기를 가장 즐겨보는 블로그 포스팅으로 꼽았고, 이는 영화관람 의도에 유의한 영향을 주는 것으로 나타났다[44]. 이상의 선행연구에 따라 본 연구에서도 영화에 대한 블로그 포스팅이 영화흥행과 관련된 구전이라고 보고, 각 영화에 대한 블로그 포스팅 수를 블로그양(volume)으로 집계하였다.

박조원, 이귀옥(2004)은 신문 보도와 영화흥행 간의 실증적 관계를 검증하였는데, 분석결과 개봉 영화의 신문 보도와 흥행 실적은 높은 상관관계가 있었다[45]. 특히 보도의 양은 관객수에 긍정적인 영향을 미치는 것으로 나타났다. Basuroy, Chatterjee, Ravid(2003)도 영화 관련 기사의 방향성과 스타 영향력, 제작비가 매출에 미치는 영향에 대해 분석하였는데, 기사의 보도 방향성에 따라 그 효과가 다르게 나타났음을 밝혔다[46]. 온라인 뉴스 혹은 기사와 영화 흥행에 관련해서는 기사의 양과 방향성을 동시에 보는 연구들도 존재하였지만, 본 연구에서는 블로그 포스팅과 마찬가지로 기사의 보도 방향성을 제외한 기사의 양(volume)을 측정하

여 분석에 사용하였다.

구전으로서 리뷰는 많은 선행연구에서 영화홍행 예측요인으로 언급되었기 때문에 본 연구에서도 온라인 구전의 변인으로 채택하였는데, 분석에서 사용한 리뷰는 전문가 리뷰가 아닌 네티즌 리뷰만을 수집하여 사용하였다. 영화홍행에 관한 기존 선행연구에서 리뷰는 전문가의 평론적 리뷰와 일반인이 작성한 대중적 리뷰가 나뉘어 사용되는데, 본 연구에서는 전문가의 리뷰의 영향력은 감소하고 있고 개봉 후에는 일반인의 평가가 흥행에 더 영향을 준다는 박승현과 장정현(2012)의 연구를 참고하여[47], 네이버 영화에서 영화를 관람한 네티즌들이 140자 이내로 작성하는 관람평 수집하여 리뷰의 양(volume)으로서 분석에 사용하였다.

검색량 데이터는 네이버 데이터 랩의 '검색어 트렌드'를 통해 수집하였는데, 이는 네티즌 정보탐색의 상대적 수치를 알려주는 서비스이며 구글에서 제공하는 구글 트렌드와 비슷하다. Goel 외 4명(2010)은 인터넷 검색량이 영화 개봉 첫 주의 흥행, 비디오 게임의 첫 달 판매량, 빌보드 HOT100 차트의 순위를 예측함을 밝혔으며, 검색 횟수는 모델의 설명력을 높인다고 밝혔다[48]. 이유석, 차경천, 김상훈(2016)은 소비자의 검색 행동과 영화홍행 사이의 상관관계를 알아보기 위해 구글 트렌드로 정보를 수집하여 실증분석한 결과, 영화의 일별 검색량은 일별 박스오피스의 선행 변수이자 결과변수로서 작용함을 밝혔다[49]. 검색량은 바이럴 마케팅 효과를 측정하는 결과변수로서 사용되기도 하지만 본 연구에서는 위와 같은 선행연구와 함께, 최근 소비자들의 검색량과 상품구매량 간의 상관관계에 대한 실증적 연구들[50-52]이 검색량 데이터는 온라인상에서 상호작용을 통하여 확산되는 구전을 의미하여, 이러한 적극적인 소비자 행동이 예측변수로서 작용할 수 있음을 시사하기 때문에 검색량(volume)을 구전에 영향을 주는 독립변수로서 분석에 사용하였다.

이상의 네 가지가 본 연구에서 수집한 온라인 구전 변수이며, 독립변수로서 분석에 사용한다. 영화홍행성과의 지표로 가장 많이 사용되는 변수는 관객수 또는 박스오피스로 불리는 매출액이다. 미국의 경우 박스오피스를 흥행의 기준으로 많이 사용하지만, 한국에서는 통상적으로 관객수를 흥행의 지표로 사용하므로, 본 연구

에서도 관객수를 종속변수로 사용하여 온라인 구전과 관객수의 관계에 대해 분석하고자 한다.

2.2 데이터 수집

영화의 일일 관객수는 영진위에서 운영하는 영화관 입장권 통합전산망(<http://www.kobis.or.kr>)에서 수집하였고, 관객수를 제외한 나머지 변수들은 포털 사이트인 네이버(www.naver.com)를 통해 수집하였다. 네이버를 수집 대상으로 선택한 이유는 웹 사이트 이용 순위 서비스를 제공하는 전문기관인 랭키닷컴(www.rankey.com)과 코리아 클릭(www.koreanclick.com)이 제공하는 방문자 수 순위에서 1위를 기록하고 있고(2018년 5월 기준), 2017년 기준 하루 순 방문자 2,700만 명, 회원 수 4,200만 명을 보유하며 국내 검색 엔진 시장에서 73%에 달하는 점유율을 갖고 있어) 네티즌의 활용성이 높고 구전량이 많을 것으로 예상되기 때문이다. 데이터 크롤링은 통계프로그램인 R을 이용하여 수집하였다. 다음 [표 3]는 변수의 조작적 정의 및 출처를 정리한 것이다.

표 3. 변수의 조작적 정의

변수명	변수 단위	변수 정의	출처
$Blog_{it}$ 블로그양	독립 개	영화 일별 블로그 포스팅 개수	네이버 블로그
$News_{it}$ 기사량	독립 개	영화 일별 온라인 기사 전체 개수	네이버 기사
$Reply_{it}$ 리뷰양	독립 개	영화 140자 평의 개수	네이버 영화
$Searchvolume_{it}$ 검색량	독립 %	분석 기간 내 영화에 대한 일별 검색량을 0부터 100사이의 값으로 정규화 한 데이터	네이버 데이터랩
$Audience_{it}$ 관객수	종속 명	영화 일일 관객수	영화진흥위원회 통합전산망

3. 연구방법

3.1 곡선추정

본 연구에서는 시간에 따른 온라인 영화정보의 구전 패턴을 알아보기 위해서 비선형회귀(non-linear)의 곡선추정을 한다. 일반적인 회귀 분석은 독립변수와 종속

2) 매경이코노미.제1951호.별책부록 (2018.03.28.). 검색일 2018년 4월 19일.
<http://news.mk.co.kr/v2/economy/view.php?year=2018&no=197771>

변수 사이에 선형의 관계(linear relation)가 있다고 가정하는데, 현실에서는 독립변수의 증감에 따라 종속변수가 비선형적으로 변화하는 경우가 더 많이 발생한다. 회귀식이 선형성을 만족하지 않으면 연구자는 두 가지 방법을 사용하여 분석할 수 있다. 첫째는 회귀식을 변수변환(transformation of variable)을 활용하여 선형성이 만족하도록 데이터를 수정하는 방법이고 둘째는 선형성 가정이 충족되지 않아도 되는 비선형 회귀모형에 적용하는 방법이다[53].

본 연구에서는 비선형 회귀모형에 직접 적용하는 방법인 곡선추정을 통해 흥행영화의 온라인 구전패턴을 가장 잘 설명하는 최적 곡선모형을 도출하고자 한다. 곡선추정 회귀모형은 독립변수가 1개일 경우에 적용할 수 있으며, 여러 가지 형태의 곡선을 적합 시키는 분석을 의미한다[54]. 이때 선택할 수 있는 모형에는 선형 모형, 대수, 역모형, 2차 모형, 3차 모형, 제곱값, 복합 모형, S모형, 성장 모형, 지수 모형, 그리고 로지스틱 곡선이 있다.

Nam·Barnett(2010)의 연구에서는 한국의 커뮤니케이션 미디어의 확산과 대체과정에서 미디어의 확산의 패턴과 추세를 알아보기 위해 곡선추정을 사용했는데 [55], 이 연구를 바탕으로 본 연구에서도 시간에 따른 온라인 구전의 패턴을 확인하기 위해 SPSS를 사용하여 곡선추정을 수행한다. 독립변수는 시간이고 종속변수는 각 영화의 블로그양, 기사량, 리뷰양, 검색량이다.

3.2 패널분석

온라인 구전과 관객수의 관계 분석을 위해 패널 분석 방법을 선택하였다. 본 연구에서 패널 분석을 하기 위해 20개 영화의 온라인 구전 데이터를 관측한 횡단면 데이터(cross-sectional data)와 구전 변수들이 개봉일부터 개봉 후 30일까지의 시간의 흐름에 따라 어떻게 변화했는지에 대한 시계열 데이터(time-series data)를 하나로 합친 패널 데이터구조가 필요하다. 패널 데이터는 분석 대상이 되는 개체들을 여러 기간 동안 관측한 자료로, 특정 시점의 여러 개의 개체의 현상 혹은 특성을 모아 놓은 데이터를 의미한다. 본 연구에서 사용한 패널 데이터 예시는 다음과 같고, 이를 long-form data 구조라 한다.

표 4. 패널데이터 예시

제목	t	블로그양	기사량	리뷰양	검색량	관객수
1 택시운전사	t+0	636	227	4961	100	698088
1 택시운전사	t+1	992	287	4473	96	651779
...						
2 신과 함께	t+0	329	216	2612	40	406365
2 신과 함께	t+1	431	218	2493	45	401449
...						
20 캐리비안의 해적	t+0	280	24	740	19	206541
...						
20 캐리비안의 해적	t+60	2	0	8	0	0

패널 자료는 하나 혹은 여러 개의 개체가 반복 관찰되기 때문에 동태적인(dynamic) 관계를 추정할 수 있고 다양한 변수의 변동성을 제공해준다는 장점이 있다. 또한, 시간에 따라 변하지 않는 개별특성효과와 시간에 따라 변화하는 시간특성효과를 모두 통제할 수 있다는 점이 장점이다. 반면, 데이터 수집 및 설계에 많은 시간이 소모된다는 점과 국가나 지역이 아닌 개인을 패널 개체로 설정할 경우 시간 변수의 길이가 짧아지고 분석 결과의 신뢰성이 떨어진다는 단점이 있다[56][57].

다음 식(2)는 일반적인 패널 회귀모형이다.

$$y_{it} = \alpha + \beta x_{it} + \epsilon_{it}, \quad i = 1, 2, \dots, n \quad t = 1, 2, \dots, T \quad (2)$$

n 은 패널 개체의 수, T 는 개체 i 의 데이터 포괄 기간이다. 본 연구에서 n 은 영화이므로 1부터 20이 되며, t 는 개봉 일을 기준으로 개봉 후 30일까지의 데이터를 의미하므로 1부터 31이 된다.

식(2)에서 각 패널 개체의 특성을 고려할 수 있는 모형을 설정하기 위해 두 개의 오차항을 가진 식으로 변형하면 다음 식(2.1)과 같다.

$$y_{it} = \alpha + \beta x_{it} + u_i + e_{it} \quad (2.1)$$

식(5.1)의 오차항들이 가진 의미는 다음과 같다.

u_i	패널 개체 간 이질성(heterogeneity), 시간에 따라 변하지 않는 고유 변수 ex) 성별, IQ
e_{it}	패널 개체와 시간에 따라 변하는 순수 오차항(idiosyncratic error term)

패널모형은 오차항의 형태와 특성에 따라 고정효과모형과 확률효과모형의 두 가지로 나눌 수 있는데 이해를 위해 식(2.1)을 다르게 표현하면 다음 식(2.2)와 같다.

$$y_{it} = (\alpha + u_i) + \beta x_{it} + e_{it} \quad (2.2)$$

고정효과모형에서는 위 식의 $(\alpha + u_i)$ 를 패널 개체별로 특정하게 고정된 값을 갖는 모수(parameter)로 간주하고, 확률효과모형에서는 이를 확률변수로 가정한다.

고정효과모형을 추정하는 데는 두 가지 방법이 있는데 첫 번째는 u_i 를 모형에서 제외하는 변환을 사용하는 것과 u_i 를 직접 추정하는 방법이다. u_i 를 모형에서 제외하는 방법은 일반 선형회귀식 식(2.1)에서 패널 그룹별 평균으로 이루어진 between 모형의 식인 식(3)을 빼주는 방법을 사용하며, 이를 식으로 나타내면 다음과 같다.

$$y_{it} = \alpha + \beta x_{it} + u_i + e_{it} \quad (2.1)$$

$$\bar{y}_i = \alpha + \beta \bar{x}_i + u_i + e_{it} \quad (3)$$

$$\bar{y}_{..} = \alpha + \beta \bar{x}_{..} + e_{..} \quad (4)$$

식(4)는 식(2.1)에서 식(3)을 빼서 수학적으로 u_i 를 없앤 모형이다. 본 연구에서 고정효과 모형은 각 개체인 영화들의 특성에 따른 영향력을 통제하여 설명변수인 영화정보 변인들이 종속변수에 미치는 영향력을 정확하게 추정할 수 있는 모형이라고 할 수 있다[58].

한편, 본 연구에서의 확률효과모형은 각 개체인 영화의 특성에 따른 효과를 확률변수로 보는 것으로, 모든 오차항이 독립적이고 등질적으로 분포되어 있다고 가정한다. 즉, 시간 불변적인 개체특성 효과는 독립변수와 상관성이 없다는 점이 고정효과모델과의 차이점이다 [59].

3.3 패널모형 검정

패널 데이터 분석에서는 관찰되지 않은 이질성의 효과를 통제하고, 모수를 추정하기 위해 어떤 패널 모형을 선택할지 결정하는 것이 중요한 문제이다. 실증분석에 있어서 어떤 모형이 더 적합한지는 하우스만 검정(Hausman test)을 통해 판단한다. 하우스만 검정은 패

널 데이터에서 어느 모형의 추정량이 더 효율적인가에 대한 검정방법으로, 독립변수와 개체특성효과 간의 상관성의 여부로 구분한다. 하우스만 검정의 귀무가설과 대립가설은 다음과 같다.

$$H_0 : cov(x_{it}, u_i) = 0$$

$$H_1 : cov(x_{it}, u_i) \neq 0$$

귀무가설은 독립변수와 개체특성효과 간에 상관관계가 없는 확률효과 모형을 채택하는 것이다. 검정을 통해 귀무가설이 맞다면 확률효과모형이 더 적합하고, 귀무가설이 틀리다면 일차추정량을 얻을 수 있는 고정효과 모형을 선택한다.

하우스만 검정에 앞서 본 연구에 적합한 모형은 고정효과모형을 가정하며, 검정을 통해 고정효과모형이 적합하다면 오차항 u_i 를 제외하여 추정하는 방법으로 가설을 검정하고자 한다. 그 이유는 영화가 가지고 있는 개별적 특성(국적, 감독 영향력, 스타 영향력, 배급사, 제작비, 계절적 요인, 등급, 장르 등)을 수학적으로 제거하여 본 연구에서 알아보고자 하는 구전과 관객수의 관계만을 효율적으로 추정할 수 있기 때문이다. 본 연구에서 패널 분석을 위해 사용하는 통계프로그램은 R version 3.3.1이다.

IV. 분석결과

1. 영화의 온라인 구전패턴 분석결과

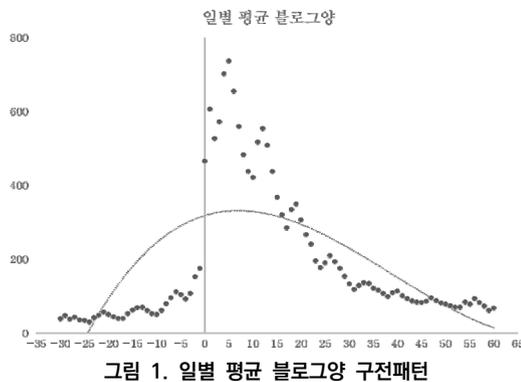
1.1 블로그양의 최적곡선 및 구전패턴

연구문제 1번을 일별 블로그 포스팅을 종속변수로, 개봉 전 30일부터 개봉 후 60일까지의 시간(91일)을 독립변수로 하는 회귀-곡선추정을 하였다. 분석결과 20개 영화 모두 삼차모형으로 가장 잘 묘사되었으며, 그중 모형의 설명력을 나타내는 결정계수인 R^2 이 가장 높은 영화는 범죄도시($R^2 = .510, p = .000$)로 나타났다. 다음은 분석 대상이 되는 영화 20개의 일별 블로그양의 최적곡선을 요약한 것이다.

표 5. 일별 블로그양 최적곡선

영화	식	모형 요약			모수 추정값			
		R ²	F	p	상수항	b1	b2	b3
택시 운전사	삼차	.460	24.67	.000	-348.42	52.47	-.918	.004
신과함께	삼차	.422	21.20	.000	-153.86	24.62	-.394	.002
공조	삼차	.484	27.21	.000	-162.41	23.15	-.295	.001
스파이더맨	삼차	.422	21.18	.000	-261.75	42.77	-.861	.005
범죄도시	삼차	.510	30.18	.000	-80.13	17.87	-.127	.000
군함도	삼차	.398	19.14	.000	-335.90	64.39	-1.37	.008
청년경찰	삼차	.474	26.10	.000	-125.59	27.27	-.474	.002
더 킹	삼차	.453	24.05	.000	-257.44	39.61	-.685	.003
미녀와 야수	삼차	.457	24.37	.000	-257.28	38.94	-.655	.003
킹스맨	삼차	.418	20.85	.000	-142.55	25.53	-.548	.003
토르	삼차	.399	19.24	.000	-145.94	25.03	-.488	.003
꾼	삼차	.449	23.65	.000	-86.48	15.79	-.282	.001
강철비	삼차	.439	22.73	.000	-160.04	24.52	-.397	.002
남한산성	삼차	.358	16.18	.000	-111.52	34.71	-.736	.004
미이라	삼차	.369	16.98	.000	-142.42	25.18	-.514	.003
분노의 질주	삼차	.404	19.69	.000	-77.03	12.74	-.247	.001
너의 이름은	삼차	.509	30.10	.000	-427.35	65.51	-1.04	.004
슈퍼배드 3	삼차	.493	28.18	.000	-95.02	19.09	-.391	.002
아이 캔 스피크	삼차	.361	16.36	.000	-149.84	22.16	-.422	.002
캐리비안의 해적	삼차	.388	18.40	.000	-94.63	15.47	-.314	.002
평균	삼차	.471	25.36	.000	-180.78	30.84	-.558	.003

다음 [그림 1]은 20개 영화의 일별 평균 블로그양의 구전패턴(R²= .471, p=.000)의 그래프이다.



1.2 기사량의 최적곡선 및 구전패턴

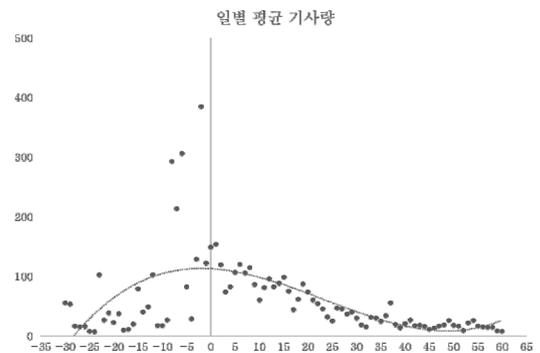
일별 인터넷 기사를 종속변수로, 시간(91일)을 독립 변수로 하는 회귀-곡선추정을 하였다. 분석결과 8개의 영화(스파이더맨: 홈 커밍, 범죄도시, 청년경찰, 킹스맨: 골든 서클, 쿤, 강철비, 남한산성)를 제외한 나머지 13

개의 영화는 삼차모형으로 가장 잘 묘사되었으며, 그중 모형의 설명력을 나타내는 결정계수인 R²이 가장 높은 영화는 분노의 질주: 더 익스트림(R²= .547, p=.000)으로 나타났다. 다음 [표 6]은 일별 기사량의 최적곡선을 요약한 것이다.

표 6. 일별 기사량 최적곡선

영화	식	모형 요약			모수 추정값			
		R ²	F	p	상수항	b1	b2	b3
택시운전사	삼차	.148	5.04	.003	-22.22	16.95	-.362	.002
신과함께	삼차	.235	8.90	.000	-96.29	17.91	-.375	.002
공조	삼차	.136	4.56	.005	-86.87	20.40	-.498	.003
스파이더맨								
범죄도시	삼차	.160	5.53	.002	4.19	25.66	-.674	.004
청년경찰								
더 킹	삼차	.182	6.45	.001	-120.02	23.31	-.548	.003
미녀와 야수	삼차	.396	18.97	.000	-11.82	3.14	-.059	.000
킹스맨								
토르	삼차	.483	27.14	.000	-8.60	2	-.032	.000
꾼								
강철비								
남한산성								
미이라	삼차	.285	11.58	.000	-15.47	3.10	-.065	.000
분노의 질주	삼차	.547	35.02	.000	-8.67	1.65	-.031	.000
너의 이름은	삼차	.160	5.54	.002	13.02	5.13	-.079	.000
슈퍼배드 3	삼차	.413	20.37	.000	-5.89	1.82	-.039	.000
아이 캔 스피크								
캐리비안의 해적	삼차	.429	21.76	.000	-6.29	1.60	-.030	.000
평균	삼차	.372	17.15	.000	-28.09	11.19	-.262	.002

다음 [그림 2]는 일별 평균 기사량의 구전패턴(R²= .372, p=.000) 그래프이다.



1.3 리뷰량의 최적곡선 및 구전패턴

일별 영화 리뷰를 종속변수로, 시간(91일)을 독립변수로 하는 회귀-곡선추정을 하였다. 분석결과 20개 영화 모두 삼차모형으로 가장 잘 묘사되었으며, 그중 모형의 설명력을 나타내는 결정계수인 R²이 가장 높은 영화는 아이 캔 스피크(R²= .398, p= .000)로 나타났다. 다음 [표 7]은 분석 대상이 되는 영화 20개의 일별 리뷰량의 최적곡선을 요약한 것이다.

표 7. 일별 리뷰량 최적곡선

영화	식	모형 요약			모수 추정값			
		R ²	F	p	상수항	b1	b2	b3
택시 운전사	삼차	.299	12.39	.000	-941.56	145.21	-3.017	.017
신과함께	삼차	.331	14.33	.000	-1014.5	140.73	-2.665	.014
공조	삼차	.331	14.36	.000	-359.68	44.72	-.649	.002
스파이더맨	삼차	.326	14.05	.000	-312.74	43.890	-.825	.004
범죄도시	삼차	.376	17.46	.000	-353.55	49.368	-.808	.003
군함도	삼차	.170	5.96	.001	-1404.87	223.878	-5.113	.031
청년경찰	삼차	.359	16.25	.000	-304.22	46.002	-.906	.005
더 킹	삼차	.337	14.71	.000	-392.9	54.929	-1.038	.005
미녀와 야수	삼차	.284	11.48	.000	-303.54	41.932	-.778	.004
킹스맨	삼차	.321	13.69	.000	-435.81	63.144	-1.303	.007
토르	삼차	.264	10.42	.000	-321.55	46.566	-.934	.005
꾼	삼차	.293	12.02	.000	-242.57	34.923	-.695	.004
강철비	삼차	.306	12.81	.000	-332.01	48.645	-.940	.005
남한산성	삼차	.213	7.83	.000	-368.85	56.724	-1.244	.007
미이라	삼차	.204	7.43	.000	-294.12	45.199	-.975	.006
분노의 질주	삼차	.318	13.49	.000	-192.04	27.185	-.513	.003
너의 이름은.	삼차	.367	16.78	.000	-556.2	81.601	-1.623	.009
슈퍼배드 3	삼차	.327	14.06	.000	-174.34	24.208	-.466	.002
아이 캔 스피크	삼차	.398	19.19	.000	-201.55	27.623	-.477	.002
캐리비안의 해적	삼차	.312	13.17	.000	-173.49	25.133	-.502	.003
평균	삼차	.323	13.84	.000	-.434	63.580	-1.273	.007

다음 [그림 3]은 20개 영화 일별 평균 리뷰량의 구전 패턴(R²= .323, p= .000)그래프이다.

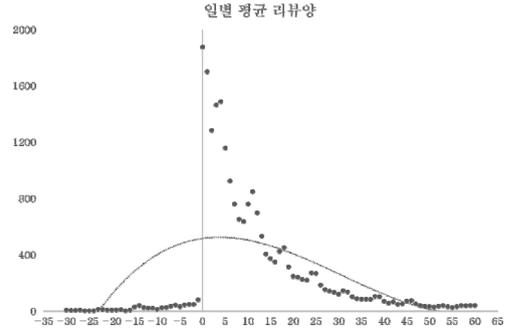


그림 3. 일별 평균 리뷰량 구전패턴

1.4 검색량의 최적곡선 및 구전패턴

일별 검색량을 종속변수로, 시간(91일)을 독립변수로 하는 회귀-곡선추정을 하였다. 분석결과 20개 영화 모두 삼차모형으로 가장 잘 묘사되었으며, 그중 모형의 설명력을 나타내는 결정계수인 R²이 가장 높은 영화는 킹스맨: 골든 서클(R²= .622, p= .000)로 나타났다. 다음 [표 8]은 영화 20개의 일별 검색량 최적곡선을 요약한 것이다.

표 8. 일별 검색량 최적곡선

영화	식	모형 요약			모수 추정값			
		R ²	F	p	상수항	b1	b2	b3
택시 운전사	삼차	.483	27.14	.000	-19.95	4.04	-.083	.000
신과함께	삼차	.340	14.97	.000	-6.21	2.34	-.033	.000
공조	삼차	.509	30.11	.000	-18.02	2.46	-.031	5.486 E-05
스파이더맨	삼차	.592	42.14	.000	-18.37	3.46	-.052	.000
범죄도시	삼차	.306	12.79	.000	-20.68	3.34	-.064	.000
군함도	삼차	.193	6.94	.000	-6.44	1.57	-.038	.000
청년경찰	삼차	.145	4.92	.003	-8.86	1.63	-.033	.000
더 킹	삼차	.261	10.24	.000	-10.05	1.87	-.039	.000
미녀와 야수	삼차	.244	9.37	.000	-12.44	2.04	-.044	.000
킹스맨	삼차	.622	47.65	.000	-4.02	3.83	-.085	.000
토르	삼차	.607	44.73	.000	-16.88	3.99	-.061	.000
꾼	삼차	.440	22.76	.000	-21.24	3.57	-.061	.000
강철비	삼차	.212	7.82	.000	-6.88	1.75	-.034	.000
남한산성	삼차	.226	8.46	.000	-15.21	2.86	-.067	.000
미이라	삼차	.174	6.12	.001	-9.86	1.83	-.043	.000
분노의 질주	삼차	.526	32.16	.000	-7.97	2.42	-.028	5.748 E-06
너의 이름은.	삼차	.429	21.82	.000	-21.93	4.08	-.080	.000
슈퍼배드 3	삼차	.244	9.36	.000	-10.56	1.96	-.042	.000
아이 캔 스피크	삼차	.279	11.24	.000	-13.16	2.20	-.045	.000
캐리비안의 해적	삼차	.421	21.10	.000	-13.17	2.76	-.053	.000
평균	삼차	.524	31.96	.000	-13.09	2.70	-.051	.000

다음 [그림 4]은 20개 영화의 일별 평균 검색량의 구전패턴($R^2 = .524, p = .000$) 그래프이다.

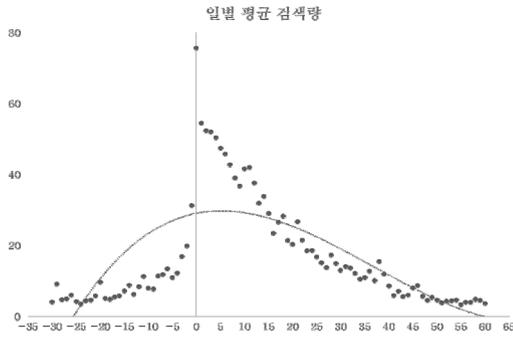


그림 4. 일별 평균 검색량 구전패턴

2. 온라인 구전과 관객수의 패널 분석결과

2.1 기술 통계량

온라인 영화의 구전과 관객수의 관계를 알아보기 위해 사용한 패널 데이터 자료는 영화의 개봉 당일부터 개봉 후 30일까지의 시계열 데이터와, 블로그양, 기사량, 리뷰양, 검색량, 그리고 관객수의 5개 변인에 대한 횡단면 데이터를 포함한다. 본 연구에서 사용한 데이터의 관측값은 2017년 한국에서 개봉한 영화 중 누적 관객수 상위 20위의 영화를 패널 그룹으로 가지고 있고, 31일간의 시계열 데이터를 포함하므로 변수별 620개이다. 패널 데이터의 기술통계량을 요약하면 다음과 같다.

표 9. 기술통계량 요약

변수	관측	평균	표준 편차	최소 값	최대값
블로그양 (개)	620	389.58	292.320	0	1751
기사량 (개)	620	76.95	77.716	0	595
리뷰양 (개)	620	623.93	866.220	36	10171
검색량 (%)	620	32.58	26.558	1	100
관객수 (명)	620	177023	207044	219	1268537

변수별 기술통계량 값을 보면, 블로그 포스팅 수의 평균은 389.58개로 하루 최소 0개에서 최대 1751까지

로 나타났다. 인터넷 기사 수의 평균은 약 77개, 최소 0개에서 최대 595개, 리뷰는 평균 약 624개, 하루 최소 36개에서 최대 10171개로 분포하고 있다. 검색량의 경우 평균 32.58%, 최소 1에서 최대 100%로 나타났다.

패널 분석에 앞서 20개 영화 변수들의 일별 평균값으로 이변량 상관관계 분석을 하였다. 분석결과 리뷰양과 검색량이 가장 높은 상관관계(.956^{**})를 보였고 기사량과 관객수가 가장 낮은 상관관계(.519^{**})를 보였다.

표 10. 일별 평균 구전량의 상관관계 분석결과

	블로그양	기사량	리뷰양	검색량
기사량	.789 ^{**}			
리뷰양	.805 ^{**}	.767 ^{**}		
검색량	.834 ^{**}	.834 ^{**}	.956 ^{**}	
관객수	.792 ^{**}	.519 ^{**}	.883 ^{**}	.847 ^{**}

^{**}p<0.05
상관관계가 0.01 수준에서 유의합니다(양측).

2.2 하우스만 검정결과

패널 데이터 분석에서 가장 중요한 것은 어떤 패널 모형을 설정하고 모수를 추정하는가에 대한 문제이다 [58]. 본 연구에서는 각각의 영화들이 가진 영화 고유의 특성들(국적, 장르, 등급, 계열 등)을 제거하고 설명변수인 온라인 영화정보 변인들이 종속변수인 관객수에 미치는 순수한 영향관계를 보기 위해 고정효과모형이 적합할 것이라는 가정하에 고정효과모형과 확률효과모형에 대해 하우스만 검정을 하였다. 연구문제 2번을 위해 모델(t)의 하우스만 검정결과 결과는 다음과 같다.

표 11. 모델(t) 하우스만 검정결과

>	phtest(Model_1.FE, Model_1.RE)
	Hausman Test
	chisq = 35.924, df = 4,
	p-value = 0.0000002999

검정결과 카이스퀘어 값은 35.924이며 p값은 1% 유의 수준에서 고정효과모형이 더 적절한 것으로 나타났다.

2.3 연구문제 및 가설검정

하우스만 검정 결과에 따른 적합모형 채택 결과 모델(t)는 고정효과모형이 더 적합하다고 나타났다. 다음은 모델(t)의 고정효과모형 분석 결과이다.

표 12. 연구문제 2번 가설 검정결과

변수	가설	Coef.	SE	가설채택 여부
블로그양	2-1	275.925***	30.387	채택
기사량	2-2	-926.408***	100.237	기각
리뷰양	2-3	135.106***	8.831	채택
검색량	2-4	1,940.931***	271.815	채택
Observations	620			
R ²	0.644			
Adj R ²	0.631			
F Statistic(df=4;596)	269.833***			

*p<0.1, **p<0.05, ***p<0.01

분석결과 R²=0.644로 모델(t)의 전체 모형은 64.4%의 설명력을 가지는 것으로 나타났다. 블로그양은 관객수와 정(+)의 관계가 있을 것이라는 가설2-1은 1% 유의수준에서 채택되었다. 이는 다른 모든 조건이 같을 때 블로그양이 하나 증가하면 관객수는 약276명 증가한다는 의미이다. 또한, 이러한 블로그양의 한계효과는 모든 패널 개체, 즉 영화 20개에 동일하게 적용된다.

기사량은 관객수와 정(+)의 관계가 있을 것이라는 가설 2-2는 분석결과 부(-)적인 관계에 있으므로 기각된다. 기사량이 하나 증가할 때 관객수는 926명 감소한다고 해석할 수 있다.

리뷰양은 관객수와 정(+)의 관계가 있을 것이라는 가설 2-3은 1% 유의 수준에서 채택되었다. 따라서 리뷰양이 하나 증가하면 관객수는 135명씩 증가한다는 결과를 얻을 수 있다.

마지막으로 검색량은 관객수와 정(+)의 관계가 있을 것이라는 가설 2-4 또한 채택되었다. 다른 모든 조건이 같을 경우 검색량이 한 단위 증가하면 관객수는 약 1941명이 증가하며, 검색량의 경우 모델(t)에서 가장 큰 회귀계수 값을 가지는 구전변수로 도출되었다. 이는 이우석, 차경천, 김상훈(2016)의 연구에서처럼[49] 검색량 데이터와 소비자 행동이 높은 상관관계가 있다는 분석과 일치하는 결과라고 할 수 있다.

V. 결론 및 제언

본 연구는 흥행영화의 온라인 구전 변수들을 수집하여 온라인 구전과 관객수가 어떤 관계가 있는지에 대해 실증적인 분석을 수행하였다. 본 연구에서 분석을 위해 수집한 온라인 구전은 블로그, 기사, 리뷰 그리고 검색

의 양(volume)이며 분석을 위해서 작년(2017) 한국에서 개봉한 영화 중 누적 관객수 상위 20개 영화를 선정하였다. 분석 대상을 바탕으로 본 연구에서는 온라인 구전 변수들이 시간에 따라 변화하는 패턴을 확인하고자 곡선추정을 하였고, 온라인 구전이 관객수와 어떤 관계가 있는지 알아보기 위해 패널 데이터 분석을 했다.

본 연구의 실증분석 결과는 다음과 같다.

첫째, 흥행영화의 온라인 구전 변수들의 패턴 분석을 위해 곡선추정을 이용하여 분석한 결과 블로그양, 기사량, 리뷰양, 검색량의 네 변인 모두 3차 모형이 구전의 패턴을 가장 잘 설명하는 것으로 나타났다. 온라인 구전의 관측값을 시간에 따라 산점도로 나타낸 그래프를 통해 구전패턴을 확인한 결과, 변수별로 차이는 있으나 개봉 일을 기점으로 종 모양의 곡선을 그리는 것을 확인할 수 있었다. 개봉 전까지는 구전이 점차 증가하면서 개봉 첫 주 가장 많은 구전들이 발생하고, 개봉 이후부터 점차 감소했기 때문이다. 단순히 구전량의 수치로만 봤을 때는 리뷰가 가장 많은 구전량을 보였고 기사가 가장 적은 구전량을 보였다.

둘째, 연구문제 2번을 위해 모델(t)를 설정하고 패널 분석을 위한 적합 모형 채택을 위해 하우스만 검정을 실시한 결과 고정효과모형이 적합한 것으로 나타났다. 따라서 연구문제 2번의 가설검정을 위해 모델(t)을 고정효과모형으로 분석한 결과 블로그양, 리뷰양, 검색량은 모두 관객수와 정(+)의 관계가 나타났고 검색량과 관객수의 관계가 가장 큰 회귀계수 값을 갖는 것으로 도출되었다.

반면, 기사량은 관객수와 부(-)의 관계가 있는 것으로 나타났다. 기사량이 블로그양, 리뷰양, 검색량과 다른 결과를 보이는 이유는 다음과 같이 추론할 수 있다. 기사는 본 연구에서 수집한 다른 구전 변인들과는 달리 일반인들에 의해 발생한 구전이 아니며, 영화에 대한 정보성을 띠기보다는 광고나 홍보의 일환으로 배포되는 경우가 다수이다. 때문에 관객이 기사를 접하는 것은 블로그 포스팅이나 영화 리뷰를 직접 찾아보고 영화를 검색해보는 행위보다 수동적 행위에 해당한다. 이러한 맥락에서 기사보다는 관객이 직접 찾아서 접하는 능동적 행위의 구전이 관객수에는 더 긍정적인 영향을 미친다고 해석할 수 있다.

실증분석 결과를 토대로 본 연구는 다음 몇 가지 의의를 갖는다.

우선 기존의 구전 연구에서 중요하게 다루지 않았던 구전의 패턴을 확인했다는 점이다. 곡선추정을 통해 구전의 패턴을 확인하고 이를 가장 잘 설명하는 최적 곡선을 발견했다는 점에서 의의가 있다.

두 번째, 본 연구는 기존의 영화흥행 관련 논문들이 영화와 관련된 여러 요소를 미시적 관점에서 분석하고 설명하려 한 것과는 다르게, 횡단면 데이터와 시계열 데이터를 합친 패널 데이터를 이용하여 구전과 관객수의 관계에 대한 거시적인 분석했다는 점에 의의가 있다. 특히 고정효과모형이 적합모형으로 도출되어, 영화마다 시간 불변하는 특성들의 오차항 효과, 다시 말해 선행연구들에서 영화흥행 요소로 밝혀진 바 있는 요소들(장르, 국적, 배우나 감독의 파워, 배급사 등)을 수학적으로 제거하여 오로지 구전과 관객수의 관계만을 파악하려고 시도했다는 데 의의가 있다. 영화의 개별적 특성들은 관객수 변화를 설명하는 중요 외생변수이지만, 현실적으로 모든 변수를 측정하여 데이터에 포함하는데 한계가 있다. 따라서 고정효과모형을 통해 영화의 개별적 특성을 사전에 제거하여 구전과 관객수의 관계에만 초점을 맞춰 분석했다는 점은 중요한 시사점이 될 수 있다.

마지막으로 본 연구결과의 한계점 및 향후 연구방향을 제시하면 다음과 같다.

첫째, 최근 소셜 미디어의 발달로 소셜 미디어 상에서 정보 탐색 및 공유까지 이뤄지고 있는데 본 연구에서 사용한 소셜 미디어 데이터는 블로그 포스팅이 유일했고 다른 소셜 미디어 데이터는 활용하지 못했다는 한계가 있다. 특히, 최근 많이 이용되는 인스타그램, 유튜브, 트위터 등의 SNS를 통한 구전효과 또한 새로 개봉하는 영화에 대한 정보제공과 선택에 많은 영향을 줄 수 있지만 본 연구에서는 이들 데이터에 대한 분석을 시도하지 못했다. 또한, 일반 영화관람객들의 리뷰양만을 변인으로 분석했지만, 주요 영화 전문 사이트에서 제공하는 전문가들 또는 파워 블로거들의 리뷰 또는 평점도 구전효과와 중요 요인으로 별도로 고려되고 분석되어야 할 필요가 있다.

두 번째로는 독립변인으로 사용된 기사량을 해당 영

화에 대한 보도 기사의 수가 아니라, 실제 기사를 읽은 양 또는 기사를 각종 SNS에 공유한 구전량을 변인으로 보았을 때는 결과가 다르게 나타났을 수 있다.

세 번째로 패널 분석의 장점은 개체의 수 혹은 시간의 단위가 커져도 분석이 쉽다는 점인데 본 연구는 영화 20개, 관측일 수는 31일로 분석 단위가 적어 패널 분석의 장점을 최대한 활용하지 못했다.

네 번째, 기존의 선행연구에서 구전효과와 매출액 혹은 구전과 관객수의 관계는 내생성(endogeneity)으로 인해 상호영향을 준다는 결과가 있었는데, 본 연구에서는 구전을 독립변수로, 관객수를 종속변수로 하여 분석하였기 때문에 상호영향의 관계는 확인하지 못했다는 점이다.

이상의 한계점을 보완하기 위해서 향후 연구에서는 분석 대상인 영화 수를 늘려 관측 데이터의 양을 증가시키고, 관객과 구전의 상호작용을 함께 분석할 수 있는 모형을 설정하고 소셜 미디어 데이터 변수를 더 추가한다면 더 신뢰성 있는 결과를 얻을 수 있을 것이라 예상된다.

참 고 문 헌

- [1] 최은영, “한국영화산업의 발전방향 분석,” 한국콘텐츠학회논문지, 제8권, 제11호, pp.134-143, 2008.
- [2] 한국콘텐츠진흥원, 2016 해외콘텐츠시장 동향조사, 2017.
- [3] MPAA, *Theatrical Market Statistics 2016*, Motion Picture Association of America, Inc, 2017.
- [4] 김기태, *영화흥행을 위한 마케팅*, 서울: 삶과 꿈, 1997.
- [5] 전범수, “영화 소비 결정 요인,” 한국콘텐츠학회논문지, 제13권, 제10호, pp.226-233, 2013.
- [6] 고정민, *한국 영화산업의 선순환구조와 발전전략*, 삼성경제연구소, Issue paper, 2002.
- [7] 영화진흥위원회, 2016 온라인 영화소비자 조사 결과 보고서, 2017.
- [8] 최건용, *비즈니스로 보는 한국영화산업*, 형설출판사, 2012
- [9] B. R. Litman, “Predicting success of theatrical

- movies: An empirical study,” *The Journal of Popular Culture*, Vol.16, No.4, pp.159-175, 1983.
- [10] B. R. Litman and L. S. Kohl, “Predicting financial success of motion pictures: The '80s experience,” *Journal of Media Economics*, Vol.2, No.2, pp.35-50, 1989.
- [11] R. O. Wyatt and D. P. Badger, “Effects of information and evaluation in film criticism,” *Journalism Quarterly*, Vol.67, No.2, pp.359-368, 1990.
- [12] S. Sochay, “Predicting the performance of motion pictures,” *Journal of Media Economics*, Vol.7, No.4, pp.1-20, 1994.
- [13] S. A. Ravid, “Information, blockbusters, and stars: A study of the film industry,” *The Journal of Business*, Vol.72, No.4, pp.463-492, 1999.
- [14] B. H. Chang and E. J. Ki, “Devising a practical model for predicting theatrical movie success: Focusing on the experience good property,” *Journal of Media Economics*, Vol.18, No.4, pp.247-269, 2005.
- [15] 김휴중, “한국 영화스타의 스타파워 분석,” *문화경제 연구*, 제1권, 제1호, pp.165-200, 1998.
- [16] 유현석, “한국영화의 흥행 요인에 관한 연구,” *한국언론학보*, 제46권, 제3호, pp.183-213, 2002.
- [17] 김은미, “한국 영화의 흥행 결정 요인에 관한 연구,” *한국언론학보*, 제47호, 제2권, pp.190-220, 2003.
- [18] 최은희, 표티나, 박윤진, 염준근, “한국영화의 흥행 결정요인에 대한 통합적 연구 -2005년 말부터 2008년까지 한국 영화산업을 중심으로-,” *한국자료분석학회*, 제11호, 제5호, pp.2773-2784, 2009.
- [19] 김소영, 임승희, 정예슬, “영화 유형별 영화 흥행 성과 예측 요인의 비교 연구,” *한국콘텐츠학회논문지*, 제10권, 제2호, pp.381-393, 2010.
- [20] 박승현, 송현주, 정완규, “한국영화의 흥행성과 결정 요인에 관한 연구,” *언론과학연구*, 제11권, 제4호, pp.231-258, 2011.
- [21] 배정호, 심범준, 김병도, “온라인 구전과 영화 매출 간 상호영향에 관한 연구,” *ASIA MARKETING JOURNAL*, Vol.12, No.2, pp.1-25, 2010.
- [22] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, “Electronic word of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?,” *Journal of interactive marketing*, Vol.18, No.1, pp.38-52, 2004.
- [23] 홍성태, 이은영, “온라인 구전마케팅에 대한 이해와 그 활용전략에 대한 연구,” *사회과학연구*, 제19권, pp.1-15, 2004.
- [24] B. Bickart and R. M. Schindler, “Internet forums as influential sources of consumer information,” *Journal of interactive marketing*, Vol.15, No.3, pp.31-40, 2001.
- [25] S. Senecal and J. Nantel, “The influence of online product recommendations on consumers' online choices,” *Journal of retailing*, Vol.80, No.2, pp.159-169, 2004.
- [26] 정영호, “온라인 구전의 정보적 특성 (구전량, 일치도, 정보원) 이 소비자 선택에 미치는 영향,” *방송과 커뮤니케이션*, 제14권, 제1호, pp.47-76, 2013.
- [27] G. T. Lau and S. Ng, “Individual and situational factors influencing negative word-of-mouth behaviour,” *Canadian Journal of Administrative Sciences*, Vol.18, No.3, pp.163-178, 2001.
- [28] 김성훈, “제품 관여도 및 제품 지식에 따른 온라인 구전정보 활용 연구,” *광고학연구*, 제14권, 제1호, pp.257-280, 2003.
- [29] 정재학, 김영찬, “구전 효과가 소비자 제품 선택에 미치는 영향에 대한 연구,” *마케팅연구*, 제19권, 제3호, pp.1-20, 2004.
- [30] S. H. Kim, N. Park, and S. H. Park, “Exploring the effects of online word of mouth and expert reviews on theatrical movies' box office success,” *Journal of Media Economics*, Vol.26, No.2, pp.98-114, 2013.
- [31] 김세준, 이세진, “정보원, 메시지의 유형 및 수신자 성향에 따른 구전효과 연구,” *광고연구*, 제96권, pp.100-136, 2013.
- [32] F. Zufryden, “Linking advertising to box office performance of new film releases-A marketing planning model,” *Journal of Advertising Research*, Vol.36, No.4, pp.29-42, 1996.
- [33] R. O. Wyatt and D. P. Badger, “How reviews affect interest in and evaluation of films,”

- Journalism Quarterly, Vol.61, No.4, pp.874-878, 1984.
- [34] W. Duan, B. Gu, and A. B. Whinston, "The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry," *Journal of retailing*, Vol.84, No.2, pp.233-242, 2008(6).
- [35] Y. Liu, "Word of mouth for movies: Its dynamics and impact on box office revenue," *Journal of marketing*, Vol.70, No.3, pp.74-89, 2006.
- [36] W. Duan, B. Gu, and A. B. Whinston, "Do online reviews matter?—An empirical investigation of panel data," *Decision support systems*, Vol.45, No.4, pp.1007-1016, 2008(11).
- [37] 김연형, 홍정환, "영화 흥행 결정 요인과 흥행 성과 예측 연구," *Communications for Statistical Applications and Methods*, 제18권, 제6호, pp.859-869, 2011.
- [38] S. Jeon and Y. S. Son, "Effect of online word-of-mouth variables as predictors of box office," *Korean Journal of Applied Statistics*, Vol.29, No.4, pp.657-678, 2016.
- [39] 전범수, "국내 영화관람객의 영화 소비 행동," *한국방송학보*, 제17권, 제2호, pp.297-326, 2003.
- [40] 조재영, "영화홍보에 있어 구전 커뮤니케이션의 효과," *한국광고홍보학보*, 제7권, 제4호, pp.193-217, 2005.
- [41] A. Chakravarty, Y. Liu, and T. Mazumdar, "The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation," *Journal of Interactive Marketing*, Vol.24, No.3, pp.185-197, 2010.
- [42] C. C. Moul, "Measuring word of mouth's impact on theatrical movie admissions," *Journal of Economics & Management Strategy*, Vol.16, No.4, pp.859-892, 2007.
- [43] 임현정, 양희동, 백현미, "블로그와 온라인 뉴스가 영화흥행에 미치는 영향에 대한 실증연구," *Journal of Information Technology Applications & Management*, Vol.21, No.4, pp.157-171, 2014.
- [44] 이수범, 강은희, "블로그를 이용한 영화 마케팅의 사용자 인식에 관한 연구," *광고학연구*, 제21권, 제1호, pp.329-342, 2010.
- [45] 박조원, 이귀옥, "영화 관련 신문 보도가 영화 흥행에 미치는 영향에 관한 연구," *한국언론학보*, 제48권, 제6호, pp.62-83, 2004.
- [46] S. Basuroy, S. Chatterjee, and S. A. Ravid, "How critical are critical reviews? The box office effects of film critics, star power, and budgets," *Journal of marketing*, Vol.67, No.4, pp.103-117, 2003.
- [47] 박승현, 장경현, "온라인 영화 리뷰의 내용과 품질에 관한 탐색적 연구," *언론과학연구*, 제12권, 제4호, pp.221-256, 2012
- [48] S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts, "Predicting consumer behavior with Web search," *Proceedings of the National Academy of Sciences of the United States of America*, Vol.107, No.4, pp.17486-17490, 2010.
- [49] 이유석, 차경천, 김상훈, "인터넷 정보 검색 행동과 영화 흥행의 상관관계에 대한 연구," *경영학연구*, 제45권, 제5호, pp.1501-1526, 2016.
- [50] H. Choi and H. Varian, "Predicting the present with Google Trends," *Economic Record*, Vol.88, No.1, pp.2-9, 2012.
- [51] S. P. Jun, J. H. Yeom, and J. K. Son, "A study of the method using search traffic to analyze new technology adoption," *Technological Forecasting and Social Change*, Vol.81, pp.82-95, 2014.
- [52] G. Kulkarni, P. K. Kannan, and W. Moe, "Using online search data to forecast new product sales," *Decision Support Systems*, Vol.52, No.3, pp.604-611, 2012.
- [53] 한광중, *SPSS활용 미래예측과 시계열분석*, 백산출판사, 2015.
- [54] 데이터솔루션 컨설팅팀, *SPSS Statistics 회귀분석*, SPSS 데이터솔루션, 2013.
- [55] Y. J. Nam and G. A. Barnett, "Communication media diffusion and substitutions: Longitudinal trends from 1980 to 2005 in Korea," *new media & society*, Vol.12, No.7, pp.1137-1155, 2010.
- [56] 민인식, 최필선, *STATA 패널데이터 분석*, 지필미디어

어, 2012.

[57] 최충익, “패널모형: 시계열 분석과 횡단면 분석을 한 번에,” 국토, 통권320호, pp.120-127, 2008.

[58] 이희연, 노승철, *고급통계분석론*, 문우사, 2012.

[59] 배은송, *패널데이터분석을 활용한 인바운드 관광수 요에 미치는 한류의 영향*, 한양대학교 대학원, 박사학위논문, 2017.

저 자 소 개

황 예 나(Yena Hwang)

정회원



- 2018년 8월 : 경희대학교 문화관광콘텐츠학과(석사)

〈관심분야〉 : 영화, 미디어, 소셜네트워크

남 윤 재(Yoonjae Nam)

정회원



- 1998년 2월 : 연세대학교 신문방송학과(학사)
- 2004년 1월 : 럿거스대학교, 커뮤니케이션, 정보학(석사)
- 2011년 9월 : 뉴욕주립대학교 버펄로, 커뮤니케이션(박사)
- 2012년 9월 ~ 2018년 8월 : 경희대학교 문화관광콘텐츠학과 조교수

- 2018년 9월 ~ 현재 : 경희대학교 문화관광콘텐츠학과 부교수

〈관심분야〉 : 미디어, 관광콘텐츠, 사회관계망