

미디어 작품 캐릭터 가치 측정 연구: 네트워크 중심성 척도와 검색 데이터를 활용하여

A Study on the Estimation of Character Value in Media Works:
Based on Network Centralities and Web-Search Data

조성현 (Seonghyun Cho)	한국과학기술원 경영대학 ¹⁾
이민형 (Minhyung Lee)	한국과학기술원 경영대학 ²⁾
최한별 (HanByeol Stella Choi)	한국과학기술원 경영대학 ³⁾
이희석 (Heeseok Lee)	한국과학기술원 경영대학 ⁴⁾

〈 국문초록 〉

무형자산의 가치에 대한 중요성이 대두되면서 이를 측정하는 것에 관한 다양한 연구가 진행되었다. 그러나 미디어 산업의 빠른 성장에도 불구하고 해당 산업 내 캐릭터 가치를 정량적으로 평가하는 데 많은 어려움이 존재한다. 최근에는 소셜 네트워크 분석 (Social Network Analysis) 방법론이 미디어 사용자의 행태를 분석하는 데 유용하게 활용되고 있다. 본 연구는 SNS 데이터를 통하여 미디어 작품의 캐릭터 네트워크 특징과 인간의 검색 행위 사이의 상관 관계를 분석하였다. 분석 결과 미디어 작품의 캐릭터 네트워크 중심성 척도와 검색 데이터 간 유의미한 상관 관계 및 인과성이 확인되었다. 본 연구 결과는 캐릭터 네트워크가 캐릭터 자산의 가치평가를 위한 단서로서 활용될 수 있음을 시사한다.

주제어: 무형자산, 지식경영, 미디어 산업, 캐릭터 가치, 캐릭터 네트워크, 소셜 네트워크 분석, 검색 데이터

1) 제1저자, seonghyun.cho@kaist.ac.kr

2) 제2저자, white_wishes@kaist.ac.kr

3) 교신저자, chb3@kaist.ac.kr

4) 제4저자, hsl@kaist.ac.kr

1. 서론

‘무형자산의 가치를 어떻게 측정할 것인가’에 대한 논의는 오랜 시간 다양한 방식으로 이어져 왔다 (Molloy et al. 2011, 박영수 and 최성호 2020). 기업의 브랜드, 상표권, 지적재산권 등 무형자산의 가치는 경영성과에 많은 영향을 미치지만, 그 가치를 정량적으로 측정하기 어렵기 때문에 가치평가 방식에 대해 지속적으로 연구가 진행되고 있다 (Allee 2000, Castilla-Polo and Gallardo-Vázquez 2016). 대표적으로, 비정형 요소들을 고려한 가치 네트워크 (Value Network)를 활용하여 무형 자산의 가치를 측정하는 방법이 여러 측면에서 대안적으로 연구되고 있다 (Allee 2008).

특히 미디어 산업은 개별 기업이 보유하는 자산 중 무형자산이 차지하는 비중이 높아, 그 중요성이 더욱 대두되는 분야이다. 예를 들어 출판사의 문학 작품이 지니는 가치, 미디어 콘텐츠 기업의 캐릭터 라이선스 가치 (고영희 and 이서현 2016)는 각 기업체가 가지는 자산 중 상당 부분을 차지한다. 이러한 콘텐츠 내 무형자산의 가치를 고려하여 문학작품, 연극, 영화 등 미디어 산업 부문에서 소셜 네트워크 분석 (Social Network Analysis, SNA) 방법론을 활용하여 작품을 분석하는 연구가 진행되었다 (Sparavigna 2013). 기존의 연구들은 작품 간 캐릭터들의 관계를 더 잘 이해하는 것에 분석의 초점을 맞추거나 (Masías et al. 2017) 작품의 이해도를 높이기 위해 새로운 SNA 분석 기법을 발굴하고 그 효용성을 연구하는 데 중점을 두었다 (Weng et al. 2009). 이러한 연구들은 미디어 작품에 대한 이해도를 높였다는 의의를 지니지만, 그 효용성이 개별 작품 분석에 그친다는 한계를 지닌다.

본 연구는 미디어 산업 분야에서 활발히 적용되고 있는 SNA 방법론을 활용하여 영화로 대표되는 미디어

어 콘텐츠 작품 내 특징을 분석하고자 한다. 인물, 대사 등의 비정형 텍스트 데이터를 활용하여 캐릭터 네트워크 (Character Network)를 구성하여 다층적으로 미디어 작품을 분석하였다. 더불어 작품 내 캐릭터들에 중심성 척도 (Centrality) 를 부여하여 주요 인물들 간의 관계를 정량적, 시각적으로 도식화하였다. 이야기가 인간의 행동을 유발하는 힘을 지니고 있다는 점을 고려하면 (Hineline 2018), 미디어 작품에 내재된 특징이 인간의 행동에 영향을 미칠 수 있음을 가정할 수 있다. 예를 들어, 미디어 작품은 개인에게 감동을 주어 특정 행동을 유발하거나 (Shrum 2012, Blasco-Arcas, Hernandez-Ortega et al. 2013, An 2018) 작품 속 여행지에 대해 방문할 동기를 부여하기도 한다 (O'Connor and Kim 2016). 특히 작품 내 등장인물 혹은 캐릭터는 작품에 대한 소비자의 인식 및 행동에 영향을 주는 중요한 특징으로 언급된다. 구체적으로, 작품 내 캐릭터를 표현하는 상품을 구매할 유인을 제공하기도 한다 (Kwak 2010). 소비자의 구매 의도는 작품 자체 혹은 작품 내 특징의 가치를 의미하며 (Wang et al. 2013), 미디어 콘텐츠의 중요 특징 중 하나인 캐릭터의 가치는 콘텐츠 내 캐릭터 간의 네트워크를 통해 형성된다 (Ding and Yilmaz 2010). 나아가 작품의 캐릭터는 소비자의 공감을 이끌어 작품에 대한 가치 평가에 유의미한 영향을 미친다. 따라서 캐릭터 네트워크의 수치적 특징은 캐릭터의 가치 평가에 대한 정량적 도구로 활용될 수 있으며, 이는 소비자의 구매 행동 예측에 중요한 요인으로 작용할 수 있다.

온라인 포털 검색량은 대중의 관심을 판단하는 가장 효과적인 지표이다. 특히 미디어 산업은 소비자의 평판 및 인지도가 매출에 큰 영향을 주는 산업이기에

(장리 et al. 2017, Andina-Díaz and García-Martínez 2020) 미디어 산업에서 검색량은 제품 및 서비스의 성과를 예측하는 중요한 수단으로 활용된다. 영화 대본 내의 캐릭터 네트워크가 영화의 가치를 보여주는 핵심 속성이라는 점과 검색량이 미디어 제품의 성과를 직접적으로 드러낸다는 사실을 종합하여 본 연구는 아래의 연구 질문을 제안한다.

RQ: “미디어 작품에서 추출한 캐릭터 네트워크 특징이 검색량에 영향을 미치는가?”

본 연구에서는 영화 대본 데이터를 활용하여 캐릭터 네트워크를 구축하고, 이로부터 추출한 각 캐릭터들의 네트워크 중심성 척도를 주요 요소로 활용하고자 한다. 개인의 검색 행위는 실제 검색 대상 관련 행동으로 이어진다 (Choi and Varian 2012). 이에 영화 캐릭터 네트워크 중심성 척도 값들과 각 캐릭터에 대한 검색량 데이터 간 상관관계 및 회귀 분석을 통하여, 캐릭터 네트워크 중심성 척도가 개인의 검색 행위에 영향을 미치는지 파악한다.

2. 문헌 연구

2.1. 미디어 산업에서의 SNA 활용 연구

SNA 방법론은 개인들이 이루는 사회적 관계망을 중심으로 다양한 현상을 분석하는 방법론이다 (Wasserman and Faust 1994). SNA 방법론에서는 개인을 노드로 (nodes, or vertices), 개인간의 사회적 관계를 링크로 (links, or edges) 표현하여 구성한다. SNA 방법론의 적용 분야가 확장됨에 따라, 인간의 커뮤니케이션 연구와 같은 사회학 연구 (Bavelas 1950) 및 기업의 협력 네트워크와 같은 경영학 연구 (박지혜 2017) 등 다양한 분야에서 SNA 방법론이 사용되고 있다.

미디어 분야 연구에서 또한 SNA 방법론을 활용하여 작품을 분석하는 등 여러 연구가 활발히 진행되어 왔다 (Sparavigna 2013). 연극, 소설 작품, 드라마, 만화책, 영화에 이르기까지 모든 장르에서 적용되고 있으며, SNA 방법론을 통한 작품 내 핵심 인물 탐색, 인물 간 상호작용 분석, 주요 인물 커뮤니티 확인 및 작품의 캐릭터 네트워크 위상 체계 (Topology) 특징을 분석하는 연구가 주를 이루고 있다 (Labatut and Bost 2019).

연극 분야에서는 1990년대부터 셰익스피어 작품을 중심으로 활발하게 문학 분석 연구가 진행되어 왔다 (Stiller et al. 2003). 해당 연구들은 작품 내 인물들의 구성, 관계 그리고 인물들 간 내부 커뮤니티 구조를 중심으로 작품 분석을 진행하였다. 소설 작품 분야에 있어서는 영국의 대표적 소설 작품인 ‘해리포터’ 시리즈를 바탕으로 주인공들의 청소년 시기의 캐릭터 관계 분석을 통하여 청소년 시기 피어그룹 (Peer Group)의 역할에 대한 연구하기도 하였다 (Bossart and Meidert 2013). 이 외에도 드라마 (Bazzan 2020), 각종 영화 (Lafhel et al. 2020), 캐릭터 작품 (Labatut and Bost 2019), 마블 만화책 (Alberich et al. 2014) 등의 모든 장르 연구에서 SNA 방법론이 사용되고 있다.

미디어 작품을 연구하는 기본이자 핵심 SNA 기법은 바로 특정 인물을 대상으로 하는 중심성 척도 (Bonacich 1987)와 네트워크 시각화 분석이라고 할 수 있다 (John et al. 2019). 작품 내 인물들 간 관계를 기반으로 캐릭터 네트워크를 구축한 뒤, 각종 중심성 척도를 활용하여 핵심 노드를 파악하며 (Bonacich 1972, Freeman 1977) 네트워크 시각화를 통해 작품을 직관적으로 이해할 수 있다. 실제 셰익스피어 작품 ‘로미오와 줄리엣’ 분석에 있어서 다양한 중심성 척도를 활용하여 주요 캐릭터들의 중요도 특징에 대해 연구한 바 있으며 (Weng et al. 2009, Tsai et al. 2013, Masías et al. 2017), 미국에서 10년간 방영된 드라마 시리즈

<표 1> 미디어 작품 관련 SNA 활용 선행연구

출처	분석 작품	적용 방법론	시사점
Carolina 2013	연극 '햄릿', 소설 '해리포터'	• 작품 내 등장인물들의 Degree 분포도 비교 분석	• 작품 내 캐릭터 네트워크가 인간 세계의 네트워크 구성과 유사함 (Scale-Free Network)
Masias et al., 2017	연극 '로미오와 줄리엣'	• 주요 캐릭터들의 중심성 척도 분석	• 여러 중심성 척도를 활용한 주요 캐릭터 분석의 유용성 제시
Masias et al., 2015	셰익스피어 연극 및 그 외 연극, 영화 작품 총 28편	• 개별 작품 내 커뮤니티 그룹 사이즈 비교 분석	• 셰익스피어 연극 작품과 주요 영화 작품의 그룹 사이즈 유사함
Stiller et al., 2003	셰익스피어 연극 작품 총 10편	• 작품 내 캐릭터 네트워크 구조 비교 분석	• 셰익스피어 작품 간 네트워크 구조의 유사함
Waumans et al., 2015	소설 '해리포터' 포함 다수의 문학 작품 총 35편	• 중심성 척도, 커뮤니티 분석 등	• 문학 작품 분석에서 다양한 SNA 기법의 유용성 및 의의 제시
Bossaert & Meidert, 2013	소설 '해리포터' 시리즈 총 6편	• 시리즈 별 (시계열) 청소년 캐릭터들의 피어그룹 변화 분석	• 문학 작품을 통해 청소년기 피어그룹 변화에 영향을 미치는 요인 간접적 확인
John et al., 2019	소설 '해리포터' 1편	• 캐릭터 네트워크 시각화 분석	• 네트워크 시각화를 통한 직관적 분석의 유용성 제시
Bazzan, 2018	드라마 시리즈 'Friends'	• 에피소드 별 네트워크 구조 변화 분석	• 드라마 내용 다층적 분석 제시
Weng et al., 2009	영화 'You've Got Mail'	• 영화 Scene 별 네트워크 구성 및 커뮤니티 분석	• Audio-Visual 특성을 지닌 영화 작품의 SNA 적용 방법론 제시
Tsai et al., 2013	할리우드 영화 총 12편	• 영화 작품 내 캐릭터 그룹 별 역할 분석	• 캐릭터 그룹 별 역할 분석을 통한 영화 작품 요약 방법론 제시
Majda Lafhel et al, 2020	영화 'Star Wars' 시리즈 총 6편	• 캐릭터 네트워크의 다층적 구조를 시각적으로 분석	• 캐릭터 네트워크의 층위 구조 시각적 분석의 유용성 제시 (Network Portrait Divergence)
Alberich et al., 2002	만화 시리즈 'Marvel Comics'	• 캐릭터 네트워크 Topology 분석	• 마블 캐릭터 네트워크와 인간 세계의 네트워크 구성과 유사함 발견

'Friends'를 대상으로 핵심 인물 중요도 분석 및 인물들 간 관계 구성 탐색 연구가 이루어진 바 있다 (Bazzan 2020).

<표 1>에 정리된 바와 같이, 기존의 SNA 방법론을 활용한 미디어 작품 연구는 대부분 작품 내 등장 인물 간 관계 분석에 그치고 있다. 이는 작품의 캐릭터 분석 및 캐릭터 네트워크 위상 체계 (Topology) 분석을 통해 작품 자체에 대한 이해를 높여주지만, 이러한 특성이 외부 환경에 미치는 효과를 파악하지 못하고 있다. 미디어 산업이 고부가가치 산업으로 부상하는 오늘날 미디어 상품을 소비하는 개인들의 행동을 파악하는 것은 경영적으로 중요한 함의를 지닌다 (Yadav

and Srivastava 2020). 이에, 본 연구는 영화 내 등장인물 간의 네트워크 분석을 기반으로 영화로 대표되는 미디어 상품 내의 등장인물 간 관계라는 상품 특성이 작품의 성과에 어떠한 영향을 주는지 분석하고자 한다.

2.2. 미디어 산업에서의 검색 데이터

검색 데이터는 대중의 관심을 보여주는 가장 효과적인 척도이다. 검색 데이터 중 가장 대중적으로 활용되는 데이터는 구글의 '구글 트렌드'이다 (Choi and Varian 2012). 많은 경영학 연구에서 구글 트렌드 데이터를 활용하여 고객의 기술 수용 수준 (Jun et al. 2014),

제품의 판매량 (Hu et al. 2012, Kulkarni et al. 2012, Hu et al. 2014)을 예측하였다. 미디어 산업에서도 검색 데이터를 활용하여 소비자의 행태를 예측한 연구가 진행되었다. 예를 들어 구글 트렌드 데이터를 활용하여 미술관 및 박물관 일자별 검색량 데이터와 실제 미술관, 박물관의 일자별 방문객 수의 높은 상관성이 확인된 바 있으며 (Botta et al. 2020), 일자별 특정 영화 검색량이 실제 일간 영화 박스오피스 매출액과 높은 상관관계를 지닌다는 사실을 발견하였다 (이유석 et al. 2016). 이처럼 검색 데이터는 단순히 이용자의 관심 및 정보 탐색 수준을 나타내는 것에서 멈추지 않고, 실제 소비자의 구매와 같은 의사결정 행위로 이어지는 강력한 지표로서 사용될 수 있음을 알 수 있다 (Choi and Varian 2012).

미디어 산업은 인지도 및 평판이 미치는 영향이 매우 중요한 산업이다 (Andina-Díaz and García-Martínez 2020). 따라서 미디어 산업에서 대중의 관심도가 제품 및 서비스의 성과에 미치는 영향을 파악함과 동시에 대중의 인지도에 영향을 주는 선행 요인을 찾는 것이 매우 중요하다. 검색 데이터가 소비자의 관심을 확인하는 가장 효과적인 척도라는 점을 감안했을 때, 미디어 산업에서 미디어 제품의 검색량에 미치는 요인을 찾는 시도가 필요하다.

Grant et al. (2007)는 소비자의 온라인 검색 행동에 관한 연구에서 온라인 검색 행위에는 소비자의 개인적 특성과 제품 및 서비스의 속성이 영향을 미친다 하였다. 특히 무형의 경험을 제공하는 제품 및 서비스의 경우 제품의 질적 속성이 소비자의 검색 행동에 유의한 영향을 미친다는 사실을 보였다. 영화의 질을 판단하는 가장 중요한 속성은 영화의 줄거리, 즉 대본이다 (Eliashberg et al. 2014). 선행 연구에서 대본의 감정 (sentiment) 지수가 영화의 평점을 예측하는 효과적인 수단이라는 (Frangidis et al. 2020) 사실이 밝혀졌고, 대

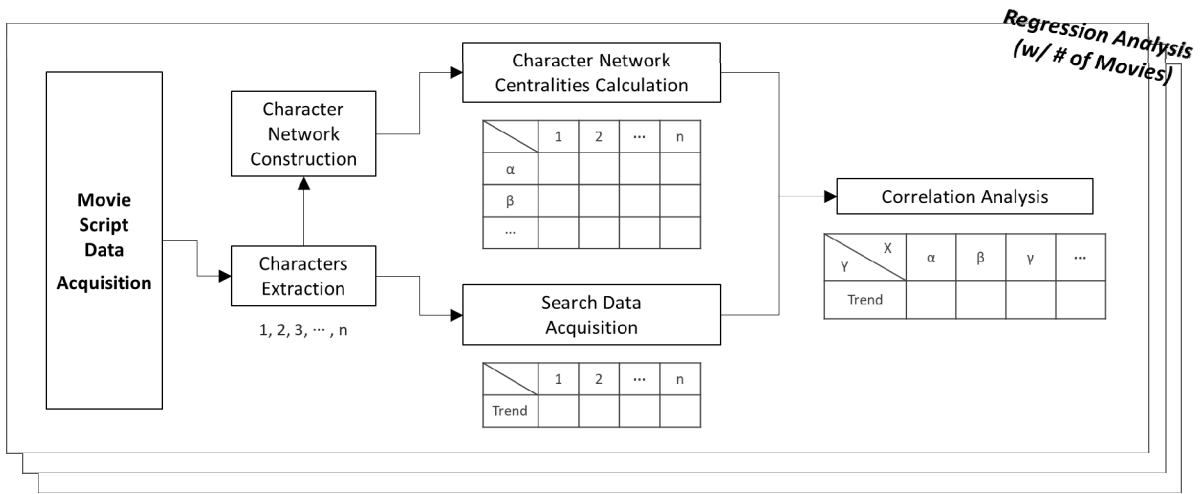
본 텍스트 분석을 통해 영화의 매출을 예측할 수 있다는 (Eliashberg et al. 2007) 사실이 입증되었다. 이러한 결과를 통해 영화 대본이 영화의 성과에 영향을 주는 중요 지표라는 사실을 확인할 수 있다. 이에 본 연구는 영화 대본 데이터를 활용하여 파악한 작품 내 캐릭터 네트워크가 구글 트렌드 검색량에 미치는 영향을 분석하여 미디어 산업에서 상품의 콘텐츠 특성이 소비자의 행동에 미치는 영향을 파악하고자 한다.

3. 연구절차 및 데이터 분석

3.1. 연구절차

본 연구에서는 <그림 1>에서 표현된 연구절차에 따라 데이터를 획득 및 전처리하고, ‘캐릭터 네트워크 중심성 척도 - 검색량 데이터 상관관계 및 회귀분석’을 진행하였다. SNA 방법론에서 네트워크 중심성 척도는 캐릭터의 작품 내 중요도를 나타내게 된다 (Freeman et al. 1979). 작품 스토리 상 중요한 캐릭터일 수록 그것을 취하는 관중의 더 높은 관심을 받을 수밖에 없다. 따라서 이와 같은 연구 방법론을 통해 캐릭터 네트워크 중심성 척도가 개별 캐릭터들의 검색량 데이터, 즉 사람들의 개별 캐릭터에 대한 관심 수준에 예측력을 지닐 수 있는지 탐색하고자 한다. 동시에 어떠한 네트워크 중심성 척도가 캐릭터의 검색량 데이터에 가장 큰 영향을 주는지 확인하고자 한다.

분석을 진행하기 위해 우선 한 편의 영화 대본을 획득하여 영화에 출연하는 모든 캐릭터 (1,2,3 ~ n)를 추출하였다. 이후, 영화 대본 내 캐릭터들 사이의 대화를 기반으로 캐릭터 네트워크를 구성한 뒤, 각 캐릭터에 네트워크 중심성 척도 (Centrality) 값들을 부여하였다. 이와 동시에 추출된 모든 캐릭터를 대상으로 해당 영화의 개봉시점부터 3개월 간의 검색량 데이터를 수



〈그림 1〉 개괄적 연구절차 도식화

집하였다. 각 캐릭터에 대한 검색량 데이터는 종속변수로서 역할을 하며, 해당 캐릭터에게 부여된 네트워크 중심성 척도 값들은 독립변수로서 역할을 한다. 우선적으로 개별 영화 단위에서 상관관계 분석을 진행하여 종속변수인 각 캐릭터의 검색량 데이터에 가장 높은 상관성을 보이는 독립변수는 어떠한 네트워크 중심성 척도인지 살펴보고자 한다. 이에 더하여 전체 영화 단위에서 분석 대상이 되는 모든 영화의 캐릭터 중심성 척도와 그에 따른 검색량 데이터를 활용하여 회귀분석을 진행하여 어떠한 중심성 척도가 검색량 데이터에 높은 예측력을 지니는지 확인한다.

분석에는 2021년 상반기까지 개봉된 모든 마블 영화 (총 23편)의 데이터가 활용되었다. 상기 기재된 과정을 영화 한 편이 아닌 모든 마블 영화를 대상으로 진행한 뒤 결과를 확인하여 분석의 예측력 수준을 신뢰성 있게 확인하였다. 선행 연구에서 마블 만화를 대상으로 SNA 분석을 진행한 바 있지만 (Alberich et al. 2014), 마블 영화를 대상으로 진행된 연구는 없었다는 점에서 추가적인 연구 의의를 지닌다. 또한, 캐릭터 산업에서의 마블 캐릭터들이 차지하는 영향력 (Harrison et al. 2019) 을 고려할 때 실제 경영환경에서

의 실무적 시사점도 지닐 것으로 보인다. 종속 변수를 획득함에 있어서는 구글 트렌드 서비스를 활용하여 (Choi and Varian 2012) 영화에 출연하는 캐릭터들의 상대적 검색량 수준을 확인한다.

3.1.1. 네트워크 중심성 척도

본 연구에서는 영화 대본에 등장하는 각 캐릭터들의 중요도를 확인하기 위한 지표로 여섯 가지 주요 네트워크 중심성 척도 (Centralities)를 사용하였다. Edwards et al. (2020)은 문학 작품 연구에서 연결중심성 (Degree), 매개중심성 (Betweenness), 근접중심성 (Closeness), 고유벡터중심성 (Eigenvector), 군집계수 (Clustering Coefficient) 라는 다섯 가지 중심성 척도를 사용하여 각 캐릭터의 작품 내 상대적 중요성을 좀 더 명확히 밝힐 수 있음을 증명하였다. 본 연구에서는 이에 더하여 네트워크 상에서 노드의 내부자 수준을 확인할 수 있는 K-Shell Factor (Csermely et al. 2013)를 추가하여 총 여섯 가지 네트워크 중심성 척도를 활용한다.

연결중심성 (Degree) 척도는 노드에 연결되어 있는 링크 (Link)들의 총합 값을 나타낸다. 특정 노드 (N_i)를 대상으로 하는 연결중심성 척도 ($C_D(N_i)$)는 아래와

같은 산식으로 산출된다 (Diestel 2005).

$$C_D(N_i) = \sum_j x_{ij}$$

x_{ji} 는 이항 변수(Binary Variable)로 노드 i (N_i)와 노드 j (N_j)간 Link로 연결되어 있으면 1, 연결되어 있지 않으면 0으로 표시된다. 연결중심성 척도는 노드 수의 총합 ($\sum N_i$)으로 나누어 정규화(Normalize)하여 사용할 수 있다. 캐릭터 네트워크 상에서 연결중심성 척도는 해당 캐릭터가 얼마나 많은 다른 캐릭터들과 대화를 주고 받았는지를 의미한다.

매개중심성 (Betweenness) 척도는 네트워크 상에서 얼마나 많은 최단 경로 (Shortest Path)들이 해당 노드를 경유하는지를 의미한다. 특정 노드 (N_i)를 대상으로 하는 매개중심성 척도 $B(N_i)$ 는 아래와 같은 산식으로 산출된다 (Freeman, 1979).

$$B(N_i) = \sum_j g_{jk}(i) / g_{jk}$$

g_{jk} 는 노드 j 와 노드 k 를 지나는 최단 경로의 수를 의미하며, $g_{jk}(i)$ 는 노드 i (N_i)를 경유하는 최단 경로의 수를 의미한다. 매개중심성 척도는 네트워크 상의 연결가능한 모든 링크 수 ($\frac{N(N-1)}{2}$)로 나누어 정규화하여 사용할 수 있다. 캐릭터 네트워크 상에서 매개중심성 척도는 해당 캐릭터가 얼마나 다른 캐릭터들을 연결하는데 얼마나 핵심적인 역할을 하는지를 의미한다.

근접중심성 (Closeness) 척도는 특정 노드로부터 얼마나 네트워크 내에 있는 모든 노드들에게 빠르게 도달할 수 있는지를 의미한다. 해당 노드로부터 몇 차례의 링크를 경유했을 때 모든 노드들에게 도달할 수 있는지를 측정한다. 특정 노드 (N_i)를 대상으로 하는 근접중심성 척도 $C(N_i)$ 는 아래와 같은 산식으로 산출된다 (Murray A. Beauchamp State, 1965).

$$C(N_i) = \frac{1}{\sum d(i,j)}$$

$d(i,j)$ 는 노드 i 에서 노드 j 까지의 최단 경로 길이를 의미한다. 근접중심성 척도는 네트워크 상의 모든 노드 수에서 특정 노드를 제외한 값 ($N - 1$)으로 나누어 정규화하여 사용할 수 있다. 캐릭터 네트워크 상에서 근접중심성 척도는 해당 캐릭터가 얼마나 네트워크 내에서 중심에 위치하여 다른 모든 캐릭터들 가까운 관계를 유지하고 있는지를 나타낸다.

고유벡터중심성 (Eigenvector) 척도는 특정 노드가 네트워크 상의 중요 노드들과 얼마나 많이 연결되어 있는지를 의미한다. 고유벡터중심성 척도를 산출하기 위해서는 사전에 네트워크 상에 있는 모든 노드들에 연결중심성 (Degree) 척도를 선행적으로 부여한 후, 높은 연결중심성 (Degree) 척도들을 나타내는 노드들과 연결이 되어 있는 수준을 확인하게 된다. 특정 노드 (N_i)를 대상으로 하는 고유벡터중심성 척도 X_i 는 아래와 같은 산식으로 산출된다 (Bonacich 2007).

$$\lambda \cdot X_i = A_{ij} \cdot X_j \quad \text{thus,} \quad X_i = 1/\lambda * \sum_j A_{ij} \cdot X_j$$

A_{ij} 는 주어진 네트워크로부터의 인접 행렬을 의미하며, 선형대수론에 기반하여 X_i 는 A_{ij} 의 고유벡터값, λ 는 고윳값을 의미한다. 캐릭터 네트워크 상에서 고유벡터중심성 척도는 특정 캐릭터가 네트워크 내 중요 캐릭터들과 얼마나 많이 직접적인 관계를 맺고 있는지를 나타낸다.

군집계수 (Local Clustering Coefficient) 척도는 특정 노드가 네트워크 상에서 다른 노드들과 얼마나 많은 삼각형 관계를 이루면서 위치하고 있는지를 의미한다. 특정 노드를 포함하여 직접적인 링크 관계로 삼각형을 만들어지는 삼각형 수의 최대값을 확인하게 된

다. 특정 노드 (N_i)를 대상으로 하는 군집계수 척도 C_i 는 아래와 같은 산식으로 산출된다 (Opsahl 2013).

$$C_i = \text{Number of closed triangles} / \left(\frac{k_i(k_i - 1)}{2} \right)$$

$\frac{k_i(k_i-1)}{2}$ 은 특정 노드 (N_i)를 포함하여 네트워크 상에 만들어질 수 있는 삼각형 수의 최대값을 의미한다. 캐릭터 네트워크 상에서 군집계수 척도는 특정 캐릭터가 단순히 다른 하나의 단일 캐릭터들과 관계를 맺는 것이 아닌 복수의 캐릭터와 삼각관계로 대사를 주고받으며, 얼마나 많은 삼각 커뮤니티에 포함되어 있는지를 나타낸다.

K-Shell Factor 척도는 특정 노드가 네트워크 상에서 어느 수준의 내부자 집단 (Inner Circle)에 속해 있는지를 나타낸다. 행렬구조로 이루어진 네트워크에서 각각의 노드는 특정 클러스터에 위치하게 된다. $G = (V, E)$, $|V| = n \text{ vertices and } |E| = e \text{ edges}$ 의 행렬 구조로 이루어진 전체 네트워크 안에 하위 네트워크 $H = (C, E|C)$, $C \subseteq V$ 로 이루어진 행렬이 존재하게 된다. 해당 하위 네트워크 내에는 추가적인 하위 네트워크가 존재할 수 있다. 몇 번째 하위 네트워크에 노드가 위치하고 있는지를 기준으로 K-Shell을 부여하게 된

다. 즉, K-Shell Factor는 특정 노드 입장에서 몇 번째 하위 네트워크에까지 위치해 있는 지로 산출된다 (Alvarez-Hamelin et al. 2005). 캐릭터 네트워크 상에서 K-Shell Factor는 특정 캐릭터가 얼마나 내부자 집단에 위치하고 있는지를 나타낸다.

아래 <표 2>를 통해 본 연구에서 활용하는 여섯 가지 중심성 척도의 정의와 각 척도가 캐릭터 네트워크 상에서 나타내는 의미를 요약 기술하였다.

3.1.2. 상관관계 분석

본 연구에서는 상관관계 분석 모형을 활용하여 두 종류의 변수간 상관성을 확인하고자 한다. 영화 대본에 등장하는 캐릭터의 중심성 척도 값을 하나의 변수, 캐릭터의 검색량 데이터를 다른 하나의 변수로 두고서 두 변수간 상관성을 확인하였다. 두 변수 간 높은 상관성이 존재한다면, 캐릭터의 중심성 척도 값이 높을수록 사람들의 해당 캐릭터에 대한 검색량 수준이 높다는 것을 의미하게 된다.

두 변수 간 상관관계를 분석함에 있어 피어슨 상관계수 (Pearson Correlation Coefficient; PCC)를 활용하였다. 피어슨 상관계수 (PCC)는 연속적으로 (continuous) 분포되어 있는 데이터를 대상으로, 스피어만 순위 상

<표 2> 중심성 척도의 캐릭터 네트워크 상에서의 의미

중심성 척도 (Centrality)	용어 정의 (Definition)	캐릭터 네트워크에서의 의미
연결중심성 (Degree)	노드에 연결되어 있는 링크 수의 총합이 어떻게 되는지	얼마나 많은 다른 캐릭터들과 직접적 관계를 맺고 있는 수준
매개중심성 (Betweenness)	네트워크 상에서 얼마나 많은 최단 경로들이 해당 노드를 경유하는지	다른 캐릭터들을 연결하는 데 핵심적인 역할을 하는 수준
근접중심성 (Closeness)	얼마나 네트워크 내에 있는 모든 노드들에게 빠르게 도달할 수 있는지	다른 모든 캐릭터들에 도달하는 데 걸리는 거리 수준
고유벡터중심성 (Eigenvector)	네트워크 상의 중요 노드들과 얼마나 많이 연결되어 있는지	중요 캐릭터들과 직접적인 관계를 맺고 있는 정도
군집계수 (Local Clustering Coefficient)	다른 노드들과 얼마나 많은 삼각형 관계를 이루면서 지역 단위에서 조밀하게 뭉쳐 있는지	지역 커뮤니티 단위에서 다른 캐릭터들과 밀접한 관계를 나타내는 수준
K-Shell Factor	네트워크 상에서 어느 깊이 수준의 내부자 집단에 속해 있는지	캐릭터가 커뮤니티 내에서 위치하고 있는 내부자 수준

관계수 (Spearman Rank correlation)는 서수적으로 (ordinal) 하게 분포되어 있는 데이터를 대상으로 하는 상관관계에서 주로 사용된다 (Schober et al., 2018). 본 연구 분석을 진행함에 있어서는, 사람들의 영화 내 개별 캐릭터에 대한 검색량 순위 확인보다 캐릭터들 간 검색량 값의 차이가 절대적으로 어느정도 수준인지를 확인하기 위함이 목표이기에 피어슨 상관관계수 (PCC) 방식을 선택하였다.

피어슨 상관관계수는 (PCC)는 두 변수 X와 Y의 데이터들의 차이 값이 최소화되는 선형상관 관계를 계량화하게 된다. 두 변수의 공분산을 분자로, 각 변수 X와 Y의 표준편차를 곱한 값을 분모로 하여 결과 값을 산출한다 (Benesty et al. 2009). 그 결과값은 +1과 -1 사이에 위치한다. 결과값이 +1에 가까울수록 두 변수간 상관성이 높으며, 0에 가까우면 상관관계가 없으며, -1에 가까울수록 음의 상관성을 지님을 의미한다.

3.1.3. 회귀분석

상관관계 분석이 두 변수간 선형적 상호 관계를 통계적으로 측정하는 것을 목표로 하는 반면, 회귀분석은 인과적으로 영향을 미치는 두 변수 간 수치적 예측 수준을 확인하는 것을 목표로 한다 (Darlington 1990). 이에 따라 두 변수는 영향을 주는 독립변수 x 와 영향을 받는 종속변수 y 로 구분된다. 선형 회귀분석에서 가장 보편적으로 사용되는 모형은 최소자승법 (OLS; Ordinary Least Squares)이다.

회귀방정식은 $\hat{y}_i = b_0 + b_1 \cdot x_i$ 식으로 구성된다. n 개의 관찰치 $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$ 로 구성된 자료에서 \hat{y}_i 는 독립변수 x 의 값이 x_i 일 때 종속변수 y 의 예측치를 나타내고, y_i 는 그에 상응하는 y 의 관찰치를 나타낸다. 회귀분석의 주 목적은 예측이기 때문에 \hat{y}_i 와 y_i 간 편차를 최소화하는 직선을 찾아내

는 것을 목적으로 하며, 최소자승법에서는 이를 위해 잔차의 자승합 (RSS: Residual Sum of Squares) 값을 최소로 만드는 계수인 b_0 와 b_1 를 구하게 된다. 잔차의 자승합은 아래와 같은 식을 통해 산출된다.

$$RSS = \sum_i^n (y_i - \hat{y}_i)^2 = \sum_i^n (y_i - b_0 - b_1 \cdot x_i)^2$$

산출의 대상이 되는 계수 b_0 와 b_1 값은 아래 식을 통해 구하게 된다. 아래 식에서 \bar{x} 와 \bar{y} 는 표본평균을 의미한다.

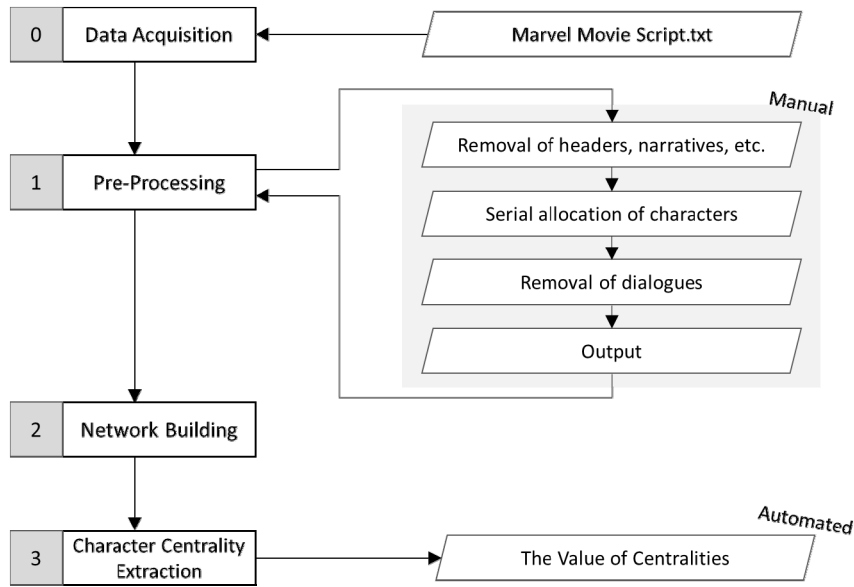
$$b_1 = \sum_i^n (x_i - \bar{x})(y_i - \bar{y}) / \sum_i^n (x_i - \bar{x})^2, b_0 = \bar{y} - b_1 \cdot \bar{x}$$

3.2. 데이터

3.2.1. 캐릭터 네트워크 구축

캐릭터 네트워크를 산출하기 위해서는 영화 대본 데이터를 획득하고 전처리하여 네트워크를 사전에 구축해야 한다. 본 연구에서는 Waumans et al. (2015) 의 SNA 분석 기법 전처리 방식을 차용하여 <그림 2>와 같은 프로세스에 따라 캐릭터 중심성 척도를 산출하기 위한 사전 작업을 진행하였다.

우선 모든 마블 스튜디오의 영화 (MCU Films; Marvel Cinematic Universe)를 대상으로 데이터를 획득하였다. 마블 영화는 2008년 ‘아이언맨 (Iron Man)’ 개봉을 시작으로 2021년 상반기까지 총 23편의 영화가 제작 및 배급되었다. 마블 영화의 대본은 ‘FANDOM (팬덤, fandom.com)’사이트를 통해 확보하였다. 팬덤은 무료 위키 호스팅 서비스를 제공하고 있으며, 위키백과와 유사하게 팬들이 TV Series, 음악, 영화와 같은 각종 미디어 관련 내용을 공유하고 자발적으로 등록하는 사이트이다. 마블 영화는 전세계 수많은 팬 층을 보유한만큼 영화 개봉 직후, 많은 팬들에 의해 해당 영화의 대본이 빠르게 업로드 된다. 팬덤 사이트에 올라와 있는 텍스트 데이



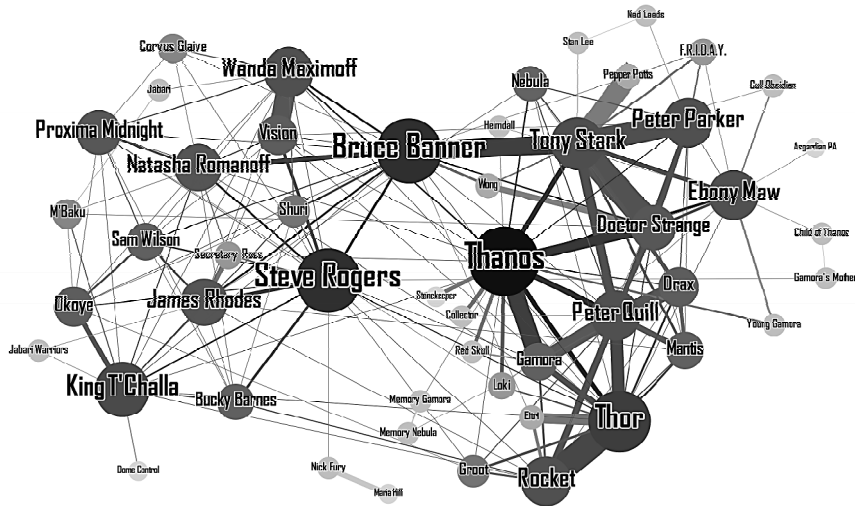
〈그림 2〉 대본 기반 캐릭터 네트워크 구축 및 중심성 척도 산출 프로세스

터를 엑셀 파일로 옮겨 영화 데이터를 획득하였으며, 그 이후 데이터 전처리 작업이 수반되었다.

영화 대본 데이터를 기반으로 캐릭터 네트워크를 구축하기 위해서는 사전 전처리 작업이 필수적으로 수반된다. 대본의 텍스트 데이터에는 캐릭터의 대사 이외에도 상황을 설명하는 안내 글과 캐릭터의 동작 및 감정 상태를 나타내는 다양한 문구들이 포함되어 있다. 캐릭터 네트워크를 구축하는데 있어 해당 텍스트 데이터는 불필요하므로 모두 제거하였다. 이후, 캐릭터들을 대사 순서에 맞게 모두 차례대로 배치하였으며 (who-talked-after-whom heuristic) (Masías et al. 2017), 이후 캐릭터의 대사들은 네트워크 구축에 불필요하여 제거하였다. 대본 텍스트 데이터에는 캐릭터명 오기입 등 각종 오류 값들이 존재하여 자동화가 어려워 전처리 과정은 모두 수작업으로 진행되었다. 이와 같은 전처리 과정을 거쳐 캐릭터 네트워크 구축을 위한 데이터셋을 재구성하였다. 네트워크 상에서 하나의 캐릭터는 하나의 노드 (Node)로 표현된다. 또한, 하나의 대사를 기준으로 앞의 캐릭터와 뒤에 배치된

캐릭터의 연결 관계는 링크 (Link)로 표현된다.

캐릭터 네트워크를 구축하는 과정은 Python의 Network X 모듈을 활용하여 자동화하여 진행하였다. 전처리한 데이터셋을 바탕으로 <그림 3> 과 같이 영화 캐릭터 네트워크를 구성하였다. 네트워크 시각화를 구현하기 위해서는 Gephi 프로그램을 활용하였다 (John et al. 2019). 캐릭터 네트워크 시각화를 통해서는 어떤 캐릭터가 작품 상에서 핵심 인물인지 탐색적으로 확인할 수 있다. 예시적으로 표현한 <그림 3>의 ‘Avengers: Infinity War’ 영화 캐릭터 네트워크는 연결중심성 (Degree) 척도의 높고 낮음을 기준으로 노드 (Node)의 크기와 색이 다르게 표시되어 있으며, 링크 (Link)의 굵기를 통해 노드 간 관계의 강도를 나타내고 있다. 이를 통해 타노스 (Thanos), 브루스 배너 (Bruce Banner, 헐크), 스티브 로저스 (Steve Rogers, 캡틴 아메리카), 토르 (Thor) 등과 같이 주요 캐릭터를 중심으로 이야기가 구성되어 있으며, 그 외 수많은 주변 캐릭터들이 이야기에 동참하여 영화가 전개된다는 점을 직관적으로 파악할 수 있다. 본 연구에서는 상기 과정을 분석의 대상이



〈그림 3〉 영화 ‘Avengers: Infinity War’ 캐릭터 네트워크 시각화

되는 총 23편의 영화를 대상으로 반복하였으며, 영화별로 각각 캐릭터 네트워크를 구축하였다.

3.2.2. 캐릭터 검색량 확보

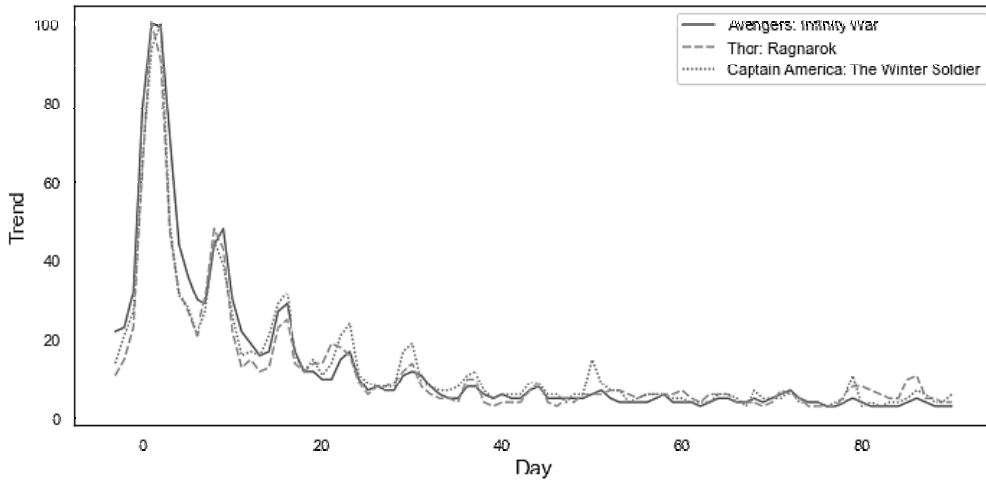
마블 영화의 캐릭터들에 대한 사람들의 관심 지표를 확인하기 위해 구글의 검색 데이터를 활용하였다. 구글은 구글 트렌드라는 공공 서비스 (trends.google.com)를 통해서 모든 사람들에게 특정 검색 키워드에 대한 정규화된 (Normalized) 시계열 검색량 데이터를 무료로 제공하고 있다. 이용자는 특정 검색 키워드에 대한 검색량을 국가별, 기간별, 카테고리별, 검색종류별로 확인하고 다운로드 할 수 있다. 검색량은 절대값이 아니며 최대값은 100으로 고정되어 있다. 이용자가 설정한 기간 내에서 최대값 대비 검색량 수준을 0에서 100 사이의 수준으로 환산하여 표시해준다 (이유석 et al. 2016).

본 연구에서는 마블 영화의 미국 개봉일을 기준으로 하여, 개봉일로부터 3일 전부터 개봉일 이후 90일 (3개월)까지의 구간을 설정하여 영화 내 등장하는 모든 개별 캐릭터에 대한 미국 지역에서의 검색량 데이

터를 확보하였다. 대부분의 영화들은 일반적으로 개봉 이후 약 4주간 상영된다. 영화별로 편차는 있으나 비인기 영화는 조금 더 상영관에서 빨리 내리기도 하며, 블록버스터 급 영화들은 1년 이상 상영이 되기도 한다. 그렇지만 대부분의 영화는 개봉 직후 1주일 이내에 가장 높은 관심을 받고 뒤이어 개봉하는 다른 영화들에 밀려 관심 수준이 낮아지는 것이 일반적이다 (Elberse and Eliashberg 2003).

〈그림 4〉는 세 편의 대표적인 마블 영화에 대하여 개봉일 3일 전부터 이후 개봉일 90일까지의 영화 제목 검색량 데이터를 겹쳐서 선그래프(Lineplot)으로 보여주고 있다. 대부분의 영화들은 상영관 개봉 당일 시점부터 1~2일 내 검색량이 최고치를 기록하고, 개봉일 이후 30일 시점부터는 검색량이 현저하게 낮아지는 것을 확인 할 수 있다. 이에 따라 각 영화의 상영일 전체에 대한 데이터를 굳이 확인하지 않더라도 사람들의 영화에 대한 관심도 수준의 최대값은 영화 개봉 후 초기 시점에 나타난다는 것을 알 수 있다.

영화에 등장하는 특정 개별 캐릭터는 개봉 이후 어느 정도 시점이 지나서 사람들의 관심을 뒤늦게 받고



〈그림 4〉 영화 개봉일 전후 영화제목 구글 검색량

서 검색량이 높아질 가능성이 있기에 개봉일 30일 시점 이후로도 최대한 검색량 데이터를 확보하고자 하였다. 하지만, 마블 영화들이 1년에 최대 3편까지 개봉하는 점을 고려하면 직전 영화에 등장한 캐릭터가 다음 영화에서 다시 등장하면서 후속 작품으로 인해 특정 캐릭터에 대한 검색량 데이터가 다시 증가하는 왜곡 현상이 발생할 수 있다. 본 연구에서는 하나의 영화 작품에 등장하는 캐릭터에 대한 사람들의 검색량 수준을 확인하는 것을 목적으로 하기에, 후속 영화 개봉 시점과 겹치지 않도록 개봉일 3일 전부터 90일 이후 시점까지 검색량 데이터 수집 구간을 설정하였다.

구글 트렌드에서는 복수의 검색 키워드의 검색량 데이터를 확인함에 있어 최대 다섯 개의 검색 키워드까지만 동시에 데이터를 제공하고 있다. 따라서 모든 캐릭터의 기간 내 검색량 데이터를 확보하기 위해, 영화 내 핵심 캐릭터들의 검색량 수준을 확인하여 최대 값을 나타내는 캐릭터를 먼저 확인하고, 해당 캐릭터를 고정시킨 후 나머지 캐릭터들을 네 개씩 차례대로 입력하여 각 캐릭터의 표준화된 검색량 데이터를 획득하였다.

개별 캐릭터에 대한 관심 즉, 인기 수준은 해당 기

간 내 데이터의 최대값을 기준으로 활용하였다. 상기 과정을 분석의 대상이 되는 총 23편의 영화를 대상으로 반복하여, 영화별로 모든 개별 캐릭터들의 검색량 데이터를 확보하였다.

4. 연구 결과

영화 캐릭터 네트워크 중심성 척도가 인간의 검색량에 영향을 미치는 지를 확인하기 위해 다음과 같은 순서로 분석을 진행하게 되었다. 우선, 한 편의 영화 캐릭터 네트워크 특징 결과를 예시적으로 먼저 살펴봄으로써 개별 영화 내에서 캐릭터 네트워크가 어떻게 구성 되어있는지, 캐릭터 별 네트워크 중심성 척도의 분포가 어떠한 지를 확인하였다. 이후, 분석의 대상이 되는 총 23편에 대하여 개별 영화 단위로 캐릭터 네트워크 중심성 척도와 캐릭터의 검색량 데이터 간 상관관계 분석을 진행하였다. 이에 더하여 전체 영화 단위 데이터셋을 활용하여 회귀분석을 진행하여 검색량 데이터에 가장 높은 예측력을 지니는 캐릭터 네트워크 중심성 척도를 확인하였다.

4.1. 네트워크 중심성 척도 분석

영화 ‘Avengers: Infinity War’의 캐릭터 네트워크 특징을 <표 3>을 통해 예시적으로 살펴보면, 총 51명의 캐릭터가 영화에 등장하였으며, 캐릭터들 간 179개의 대사 관계를 기반으로 캐릭터 네트워크가 구성되어 있음을 알 수 있다. 노드들의 평균 연결중심성 척도 (Degree)는 7.02로 하나의 캐릭터가 평균적으로 영화 내에서 다른 7명의 캐릭터와 대사 관계 (Links)를 맺고 있음을 확인할 수 있다. 캐릭터들은 평균적으로 2.32 개의 링크를 통해 다른 모든 캐릭터와 연결되어 있으며, 네트워크 전체 단위에서 0.45 군집계수 (Global Clustering Coefficient) 수준으로 밀집되어 있다.

네트워크 중심성 척도에 있어서 가장 기본이 되는 연결중심성 (Degree) 척도의 분포도를 <그림 5>와 같이 확인하였다. 대부분의 노드들은 소수의 이웃 노드들과 링크를 통해 연결되어 있으며 (Most nodes with low degree), 일부 소수의 핵심 노드들이 많은 수의 다른 노드들과 연결관계를 나타내고 있음이 (Core nodes

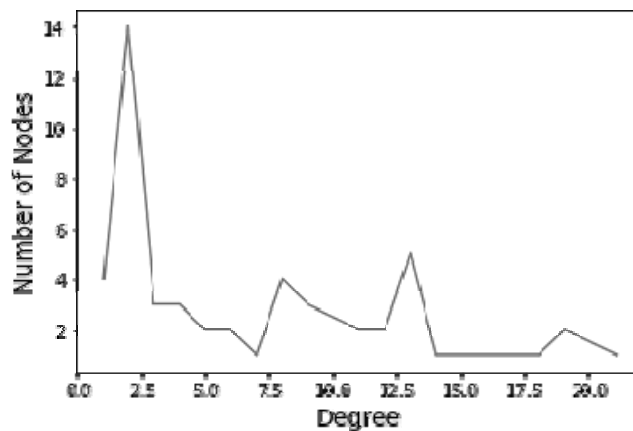
with high degree) 관찰된다. 실제 영화를 떠올려보면, 소수의 핵심 주인공들이 영화 내에서 많은 조연 및 단역 캐릭터들과 대사를 주고받으며 영화 스토리가 진행되는 것과 일치하는 결과이다.

<표 4>를 통해서 중심성 척도 별 상위 5개의 캐릭터 노드를 확인할 수 있다. 그 결과를 살펴보면, <그림 3>의 네트워크 시각화에서 탐색적으로 확인된 주요 핵심 캐릭터들이 연결중심성 (Degree) 척도의 상위 순위에 위치하는 것이 확인된다. ‘타노스 (Thanos)’가 0.42로 가장 높았으며, ‘브루스 배너 (Bruce Banner, 헐크)’, ‘스티브 로저스 (Steve Rogers, 캡틴 아메리카)’가 0.38으로 뒤이었다. ‘토르 (Thor)’가 0.36으로 네 번째 순위, ‘티찰라 (King T’Challa, 블랙 팬서)’가 0.30으로 다섯 번째 순위에 위치하였다.

매개중심성 (Betweenness) 척도에서는 상위 노드 캐릭터들의 명단이 일부 달라졌으며, 가장 높은 값을 기록한 ‘타노스 (Thanos)’ 대비 그 다음 순위 캐릭터 노드들의 값 격차가 벌어진 게 확인 되었다. 근접중심성 (Closeness) 척도에서는 상위 노드들의 명단이 연결중

<표 3> 영화 ‘Avengers: Infinity War’ 캐릭터 네트워크 특징

노드 (Nodes)	링크 (Links)	Average Degree	Average Path Length	Clustering Coefficient
51	179	7.02	2.32	0.45



<그림 5> 영화 ‘Avengers: Infinity War’ 네트워크 연결중심성 척도 분포도

<표 4> 영화 ‘Avengers: Infinity War’ 상위 5개 캐릭터 네트워크 중심성 척도

	Node	Degree	Node	Between-ness	Node	Close-ness	Node	Eigen-vector	Node	K-Shell	Node	Local Clustering
1	Thanos	0.42	Thanos	0.23	Thanos	0.59	Bruce Banner	0.29	Thanos	7	Corvus Glaive	1.0
2	Bruce Banner	0.38	Steve Rogers	0.15	Thor	0.59	Thor	0.29	Thor	7	Cull Obsidian	1.0
3	Steve Rogers	0.38	Ebony Maw	0.14	Bruce Banner	0.56	Steve Rogers	0.28	Bruce Banner	7	Heimdall	1.0
4	Thor	0.36	Bruce Banner	0.12	Steve Rogers	0.55	Thanos	0.24	Steve Rogers	7	Jabari Warriors	1.0
5	King T'Challa	0.30	King T'Challa	0.10	King T'Challa	0.54	King T'Challa	0.23	King T'Challa	7	Collector	1.0

심성 (Degree)과 아주 유사하게 확인되었으나 값들이 좀 더 커지고 밀접하게 붙어있는 점을 발견하였다. 고유벡터중심성 (Eigenvector) 척도에서는 연결중심성 (Degree)과 상위 노드들의 명단은 큰 차이는 없으나 값들이 아래로 좀 더 치우치게 된 점을 확인하였다. K-Shell Factor를 통해서 가장 핵심 노드들이 일곱 번째 하위 네트워크 (k-Shell)에 위치하고 있음을 확인하였다. 군집계수 (Local Clustering Coefficient)를 통해서 기존에 언급되지 않았던 주변 캐릭터들이 상위 권에 포진되어 있음을 발견하였으며, 이는 다수의 비 핵심 노드들이 네트워크 외곽에서 소규모로 일부 노드들끼리만 밀접하게 삼각관계를 맺고 있으며 네트워크 상에 다른 여러 노드들과의 많은 관계를 맺지 못하고 있기 때문으로 추론된다.

4.2. 모형 예측력 분석

4.2.1. 상관관계 분석

앞의 한 편의 영화 캐릭터 네트워크를 예시적으로 살펴본 것에서 나아가, 총 23편의 개별 마블 영화에 대한 캐릭터 네트워크를 각각 구축하였으며, 이를 통해 개별 영화에 등장하는 각 캐릭터들의 네트워크 중심성 척도 값들을 산출하였다. 구글 트렌드 서비스를

활용하여, 각 영화의 개봉일 3일 전부터 90일까지 기간을 설정하여 영화 내 개별 캐릭터들에 대한 검색량 데이터를 추가로 확보하였다. 캐릭터 네트워크 중심성 척도 값과 검색량 데이터를 두 가지 변수로 하여 개별 영화 단위에서 상관관계 분석을 진행하였다. <표 5>를 통해 개별 영화들을 기준으로 등장하는 캐릭터들의 네트워크 중심성 척도 값과 검색 데이터 사이의 상관관계 분석 결과를 확인할 수 있다.

이를 통해 여섯 가지 중심성 척도 중 가장 높은 상관계수를 보인 척도는 매개중심성 (Betweenness) 척도로 확인하였다. <표 5>에서 요약 기술된 바와 같이 매개중심성 척도는 평균 수치 0.68을 기록하며 중심성 척도 중 가장 높은 값을 기록하였다. 영화 편수에 있어서도 총 23편의 영화 중 13편의 영화에서 가장 높은 상관성을 보였다. 그 다음으로 연결중심성 (Degree) 척도는 평균 수치 0.66을 기록하며, 6편의 영화에서 가장 높은 상관성을 보였다. 이 두가지 중심성 척도는 별도로 확인한 p-value 값에서도 모든 영화의 결과값에서 0.01 이하를 기록하며 통계적으로 99% 수준 이상으로 유의미함을 나타냈다. 뒤이어 고유벡터중심성 (Eigenvector) 척도에서 3편, 근접중심성 (Closeness) 척도에서 1편의 영화 순으로 네트워크 중심성 척도 중

<표 5> 캐릭터 네트워크 중심성 척도와 구글 검색량 간 상관계수

No.	영화제목	Degree	Betweenness	Closeness	Eigenvector	K-Shell	Local clustering
1	Iron Man	0.65	0.61	0.60	0.62	0.25	-0.06
2	The Incredible Hulk	0.56	0.69	0.48	0.47	0.31	-0.24
3	Iron Man 2	0.87	0.88	0.73	0.77	0.57	-0.13
4	Thor	0.76	0.84	0.65	0.63	0.35	-0.22
5	Captain America: The First Avenger	0.83	0.84	0.57	0.73	0.52	-0.15
6	The Avengers	0.69	0.49	0.64	0.72	0.69	0.04
7	Iron Man 3	0.58	0.55	0.38	0.51	0.36	-0.14
8	Thor: The Dark World	0.71	0.73	0.58	0.60	0.43	-0.19
9	Captain America: The Winter Soldier	0.49	0.37	0.53	0.57	0.54	-0.09
10	Guardians of the Galaxy	0.56	0.63	0.52	0.51	0.40	-0.13
11	Avengers: Age of Ultron	0.77	0.63	0.73	0.77	0.63	-0.20
12	Ant-Man	0.63	0.66	0.50	0.51	0.32	-0.25
13	Captain America: Civil War	0.73	0.71	0.68	0.68	0.54	-0.20
14	Doctor Strange	0.86	0.96	0.84	0.72	0.31	-0.26
15	Guardians of the Galaxy Vol. 2	0.33	0.22	0.37	0.38	0.38	0.09
16	Spider-Man: Homecoming	0.62	0.54	0.50	0.57	0.41	0.33
17	Thor: Ragnarök	0.70	0.87	0.63	0.54	0.30	-0.23
18	Black Panther	0.67	0.72	0.54	0.58	0.44	-0.13
19	Avengers: Infinity War	0.70	0.79	0.62	0.61	0.44	-0.13
20	Ant-Man and the Wasp	0.61	0.62	0.62	0.62	0.37	-0.02
21	Captain Marvel	0.54	0.57	0.46	0.50	0.31	0.08
22	Avengers: Endgame	0.70	0.70	0.59	0.63	0.53	-0.16
23	Spider-Man: Far From Home	0.72	0.93	0.57	0.49	0.26	-0.14
평균 값 (Mean Value)		0.66	0.68	0.58	0.60	0.42	-0.11

높은 상관성을 보였다.

K-Shell Factor와 군집계수 (Local Clustering Coefficient) 척도는 단 한 편의 영화에서도 가장 높은 상관성을 보이지 못하는 낮은 상관성을 보여주었다. 특히, 군집계

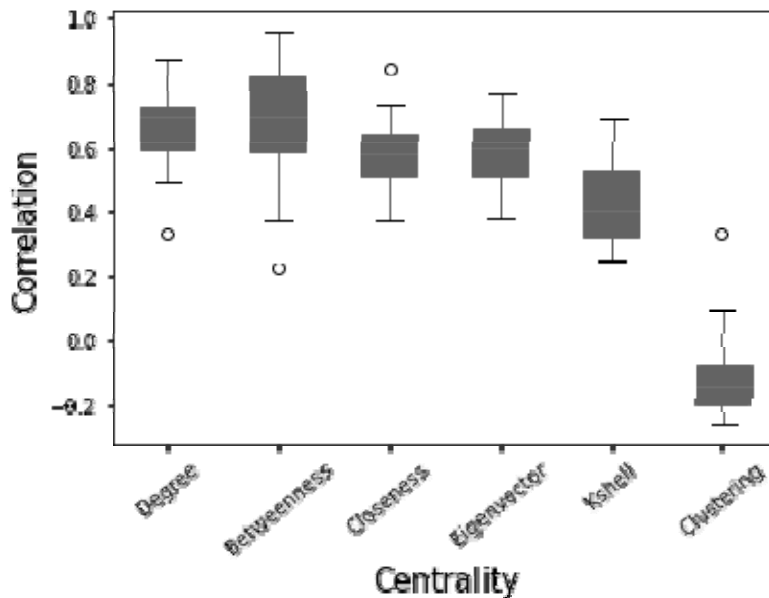
수는 대부분의 영화에서 음의 상관성을 나타내며 검색 데이터와 전혀 상관성이 없음을 나타냈다. K-Shell Factor의 경우, <표 4>에서 선행적으로 확인한 바와 같이 개별 캐릭터 노드의 전체 네트워크에서의 내부자

집단 수준에 따라 정수 단위로 값을 나타내게 된다. 이에 따라 다수의 캐릭터들간 중심성 척도 스케일 차이가 크게 발생하지 않아 검색량 데이터와 낮은 상관성을 나타냈다. 군집계수 척도는 영화 전체 캐릭터 네트워크 단위가 아닌 노드의 지역 단위 인접 노드와의 결합 수준을 확인하는 척도라는 특징으로 인하여 사람들의 관심 수준, 즉 검색 행위와는 상관성이 없는 것으로 확인되었다.

위와 같은 결과를 통해 단순 영화 편수를 통해서 네트워크 중심성 척도의 상관계수 순위를 확인하였을 때, 매개중심성 (Betweenness) 척도가 사람들의 검색량 수준과 가장 높은 상관성을 보인다는 점을 확인할 수 있다. 연결중심성 (Degree) 척도와 연결중심성 (Degree) 척도와 고유벡터중심성 (Eigenvector) 척도, 근접중심성 (Closeness) 척도 역시 상관계수 평균값에서 0.66, 0.60, 0.58을 각각 기록하며 높은 상관성을 보임을 확인하였다. 또한, 개별 영화 단에 있는 매개중심성 (Betweenness) 척도보다 일부 높은 수치를 보이거나 유사 값을 나타내는 경우도 존재하였다.

이에 따라 중심성 척도들의 전체적인 수준을 <그림 6>의 Boxplot을 통해 시각적으로 추가 확인하게 되었다. 연결중심성 (Degree) 척도와 매개중심성 (Betweenness) 척도의 상관계수는 거의 동등한 수준의 중위값 (Median)을 나타내고 있으며, 연결중심성 (Degree) 척도의 1~3사분위 값의 범위가 모두 매개중심성 (Betweenness)의 1~3사분위 값의 범위 내에 있음을 확인할 수 있다. 즉, 연결중심성 (Degree) 척도와 매개중심성 (Betweenness) 척도는 거의 동등한 수준의 상관계수를 나타내지만, 연결중심성 (Degree) 척도가 좀 더 낮은 분산 수준을 나타냄을 확인할 수 있다. 근접중심성 (Closeness) 척도와 고유벡터중심성 (Eigenvector) 척도는 매개중심성 (Betweenness) 척도와 일부 겹치는 구간이 존재하지만 그 중위값의 위치가 확연히 시각적으로 차이가 나는 것을 확인할 수 있었다. K-Shell Factor와 군집계수 (Local Clustering Coefficient)에서는 그 차이가 더욱 벌어졌다.

각 중심성 척도들의 평균 차이를 통계적으로 확인하기 위하여 T-test 차이 검정을 진행하였다. 가장 높은 예측력을 나타내는 매개중심성 (Betweenness) 척도를



<그림 6> 캐릭터 네트워크 중심성 척도 상관계수 분포도

〈표 6〉 매개중심성 척도와 다른 중심성 척도 값들 T-test 차이 검증 결과

Centrality	Degree	Closeness	Eigenvector	K-Shell	Local Clustering
statistic	-0.1936	2.1207	1.7671	5.6201	16.5903
p-value	0.8475	0.0407**	0.0857*	0.0000***	0.0000***

〈표 7〉 캐릭터 네트워크 중심성 척도와 검색량 데이터 간 회귀분석 결과 값

Centrality	Degree	Between-ness	Closeness	Eigen-vector	K-Shell	Local Clustering
Coefficient	304.95***	594.04***	387.50***	344.03***	9.61***	-19.51**
R-squared	0.134	0.188	0.091	0.098	0.028	0.005

기준으로 하여 다른 다섯개의 네트워크 중심성 척도 값들을 차례대로 T-test 검정을 진행하였다. <표 6>의 결과를 살펴보면, <그림 6>을 통해 시각적으로 확인한 바와 유사한 결과를 확인할 수 있었다.

매개중심성 (Betweenness) 척도와 연결중심성 (Degree) 척도의 평균값은 p-value가 0.84라는 높은 수치를 기록하여 통계적으로 유의미한 차이를 보이지 않음을 확인하였다. 고유벡터중심성 (Eigenvector) 척도와는 p-value 값이 0.1 밑을 기록하여 90% 유의수준에서 통계적으로 차이가 존재하였으며, 그 외 중심성 척도들과는 p-value 값이 0.05 밑을 기록하여 95% 유의수준 이상에서 확연하게 통계적으로 중심성 척도 평균값 차이가 존재함을 발견할 수 있었다.

4.2.2. 회귀분석

캐릭터 네트워크 중심성 척도 값이 검색 데이터에 인과적으로 영향을 미치는 지를 확인하기 위하여 전체 영화 단위에서 OLS 회귀분석을 추가적으로 진행하였다. 사전에 상관관계 분석을 진행하며 확보한 총 23편 영화의 개별 영화 단위 캐릭터 중심성 척도 값을 하나의 데이터셋으로 결합하는 과정을 거쳤다. 개봉일을 기준으로 확보한 개별 영화 단위에서 캐릭터들의 검색량 데이터 역시 전체 통합 데이터셋에 결합하였다. 총 23편의 영화 통합 데이터셋은 1,083명의 캐릭터와 개별 캐릭터들의 중심성 척도값들 및 검색량

데이터로 구성되었다. 시리즈물로 제작되는 마블 영화 특성 상 여러 영화에 중복되는 캐릭터들이 등장하지만, 시간 차이를 두고 상영되는 개별 영화 내에서 캐릭터들의 상대적 인기 수준 확인을 목표로 하기에 중복되는 캐릭터들이 다른 영화에 출연하는 경우 별도의 인덱스 값으로 구분하여 별개의 캐릭터로 처리하였다.

캐릭터의 검색량 데이터를 종속변수로 하여, 여섯가지 네트워크 중심성 척도를 개별 독립변수로 하여 단일 변수 최소자승법 (OLS) 회귀분석을 여섯 차례에 걸쳐 각각 진행하였다. <표 7>의 분석 결과를 통해, 매개중심성 (Betweenness) 척도가 캐릭터의 검색량 데이터에 가장 높은 예측력을 지닌다는 점을 확인할 수 있었다. 네트워크 중심성 척도 중 가장 높은 계수 (Coefficient)를 기록하였으며, p-value 값이 0.01 이하를 나타내며 통계적으로도 99% 수준에서 유의미함을 나타냈다. 설명력 (R-squared) 측면에서도 매개중심성 척도는 가장 높은 수준을 보였다.

5. 시사점 및 한계

5.1. 학문적 시사점

본 연구는 미디어 작품 내 캐릭터 네트워크 특징과 인간의 행동 사이의 관계를 분석한 최초의 연구라는

점에서 의의를 지닌다. 지금까지 SNA 방법론을 활용하여 미디어 작품의 특징을 바탕으로 작품을 깊게 이해하려는 수많은 선행 연구들이 존재하였다 (Aggarwal and Subbian 2014). 많은 선행연구들이 작품 내 캐릭터 네트워크 특징이 실제 세계의 사람들의 네트워크 형태와 유사함을 밝혀낸 바 있지만, 캐릭터 네트워크 특징이 인간의 행동에 미치는 영향을 분석하는 데까지 나아가지 못하였다. 본 연구는 문학 작품이 사람의 행동에 영향을 줄 수 있다는 사실을 주목하여 (Hineline 2018), 미디어 작품 상의 캐릭터 네트워크 특징이 인간의 행동에 영향을 줄 수 있는지를 분석하였다.

연구 분석을 진행함에 있어 다음과 같은 방법론적 기여를 하였다. 우선, ‘캐릭터 네트워크 중심성 척도 - 검색량 데이터 상관관계분석 및 회귀분석’ 연구 방법론을 제시하였다. 네트워크 중심성 척도는 캐릭터의 작품 내 중요성을 의미하게 되며, 작품 스토리에서 더 중요한 캐릭터일수록 영화 관람객의 더 높은 관심을 받을 수밖에 없다. 본 연구에서는 영화 대본을 활용하여 캐릭터 네트워크를 구축하였으며, 캐릭터의 주요 네트워크 중심성 척도와 캐릭터의 검색량 간 상관성 및 인과 관계를 실증적으로 증명하였다.

또한, 인간의 검색 행위에 가장 높은 영향을 미치는 미디어 작품의 캐릭터 네트워크 중심성 척도는 매개 중심성 (Betweenness) 척도임을 발견하였다. 단순히 여러 캐릭터와 많은 대사 관계를 맺고 있는지를 의미하는 연결중심성 (Degree) 척도보다 캐릭터 네트워크 상에서 중요도를 지니는 캐릭터들과 직접적 관계를 맺으며, 영화 스토리 전환에 핵심이 되는 캐릭터를 나타내는 매개중심성 (Betweenness) 척도가 캐릭터 검색량 데이터와 더 높은 상관성을 지닌다는 사실을 밝혔다. 또한, 회귀분석을 통해 매개중심성 (Betweenness) 척도라는 캐릭터 네트워크 특징이 인간의 검색 행위에 대한 예측력을 지닌다는 점을 발견함으로써 미디어

어 작품 내 캐릭터 네트워크를 활용하여 인간의 행동을 예측하는 (Vespignani 2009) 새로운 연구 기반을 마련하게 되었다.

이에 더하여 본 연구는 검색량 데이터를 종속변수로 활용하여 분석한 새로운 시도를 한 연구라고 할 수 있다. 검색량 데이터를 활용한 대부분의 선행 연구는 검색량 데이터가 인간의 행동을 예측하는 지표로서 역할을 할 수 있음을 증명하는데 초점을 두었다 (Choi and Varian 2012). 본 연구는 수많은 선행 연구들을 기반으로 검색량 데이터가 인간의 관심 및 행동 수준을 나타내는 지표로서 역할을 할 수 있다는 점을 전제로 하여, 특정 캐릭터에 대한 검색량 데이터 값을 소비자의 캐릭터에 대한 관심 수준으로 활용하였다. 이후, 영화 대본에서 추출한 캐릭터 네트워크 중심성 척도 값과 검색 데이터 사이의 상관성을 확인하여 네트워크 중심성 척도 값이 인과적으로 사람의 검색량, 즉 소비자의 관심 수준에 영향을 미친다는 점을 발견하였다.

5.2. 실무적 시사점

영화 캐릭터 네트워크를 통해 산출한 캐릭터의 중심성 척도가 개별 캐릭터의 인기 지표로 사용될 수 있는 가능성을 제시하였다. 기존의 캐릭터 라이선스 협상에 있어서는 그 로열티 수준을 산정할 때 정성적인 양측의 주장에 기반하여 그 수준이 정해지는 것이 일반적이다. 캐릭터 라이선스 로열티는 그 캐릭터에 대한 사람들의 관심과 인기 수준에 의해 정해지지만 (이선영 and 이승진 2015), 그 수준을 수치적으로 그리고 선행적으로 판단하기는 용이하지 않다.

본 연구 결과는 첫번째로 라이선스 협상에 있어서 캐릭터의 네트워크 매개중심성 (Betweenness) 척도가 영화 내 캐릭터들의 개별 인기 수준을 판단하는 정량

적인 참고 수치로서 활용될 수 있음을 제시한다. 업계의 일반적인 현황에 따르면 기존에는 수많은 캐릭터들에 대한 사람들의 인기 수준을 개별적으로 판단하기 어려웠기에, 모든 캐릭터들에게 동일한 수준으로 매출의 적정 비율의 로열티를 산정하는 게 일반적이다. 하지만, 영화 내 핵심 주인공 캐릭터는 사람들의 더 많은 관심을 받으며 해당 캐릭터 상품이 더 많은 매출을 기록한다는 것은 상식적으로 예측이 가능하다. 그렇지만 모든 캐릭터에게 동일 비율로 로열티를 산정한다면 라이선스 권한 구매사 (Licensee)는 핵심 주인공 캐릭터를 위주로만 캐릭터 제품을 생산하여 판매하여 수익을 창출하고자 할 것이다. 조연급 주인공 캐릭터를 활용하여 제품을 생산하였다가 해당 제품이 팔리지 않을 경우, 제품의 보관 및 폐기 비용 리스크와 불확실성을 홀로 떠안아야 하기 때문이다. 본 연구의 결과를 근거로 하여 핵심 주인공 캐릭터 대비 다른 조연급 캐릭터의 중심성 척도 수준을 참고하여 로열티 수준을 더 낮게 할인하여 개별적으로 책정한다면, 캐릭터 라이선스 권한 구매사 (Licensee)는 조연급 캐릭터 제품의 생산 판매에 대한 리스크가 줄어들어 해당 상품을 개발하여 수익을 창출할 유인이 추가로 생기게 된다. 권리 보유사 (Licensor) 입장에서도 캐릭터 상품의 매출에 대한 일정 비율로 로열티 수익을 창출하기에, 조연급 캐릭터 상품을 기반으로 추가 매출을 기록하여 이익을 증대시킬 사업적 기회로 작용할 수 있다.

두번째로, 본 연구 결과를 활용하여 영화 캐릭터 상품을 적시에 출시하여 이익 창출의 극대화를 꾀할 수 있을 것이다. <그림 4>에서 보인 바와 같이 영화는 개봉 직후 단기간에만 높은 인기를 누리기 때문에 (Elberse and Eliashberg 2003), 캐릭터 상품 또한 영화 개봉 시점에 맞추어 빠른 출시가 이루어져야만 높은 이익을 창출할 수 있다. 개봉 이후 2~4주의 시간이 흘

러서 개별 캐릭터의 인기 수준을 판단하고 캐릭터를 활용한 제품을 판매하거나 광고를 기획한다면 그 캐릭터의 인기를 상업적으로 온전히 누리기 어렵다. 영화에 등장하는 각 캐릭터의 네트워크 중심성 척도는 영화 대부분만 있으면 빠르게 산출이 가능하다. 그렇기 때문에 영화 시사회 직후 혹은 영화 개봉 당일에도 각 캐릭터에 대한 인기 수준을 예측할 수 있다. 해당 수치를 참고하여 적시에 라이선싱 계약을 진행한다면 계약의 양 당사자 모두에게 더 높은 이익을 누리게 할 기회를 제공할 것이다. 세번째로, 본 연구는 OTT 산업의 확산에 따라 지속적으로 성장하고 있는 콘텐츠 시장과 연계하여 영화 캐릭터 산업을 확산시키는 데 기여할 수 있을 것으로 보인다. 최근 코로나 상황으로 인하여, 집에서 머무는 시간이 늘어나면서 극장이 아닌 Netflix, IPTV 플랫폼 등 OTT 플랫폼을 통한 영화 콘텐츠 소비가 증가하고 있다. OTT 플랫폼은 시간과 장소의 영향을 받지 않기 때문에, 더 많은 고객이 영화 콘텐츠를 소비할 수 있으며, 특히 시리즈로 연계되는 영화들은 플랫폼 상에서 더 쉽게 홍보되고 노출됨에 따라 해당 영화 캐릭터의 인기도가 올라갈 수 있다. 본 연구의 네트워크 분석과 더불어, OTT 플랫폼의 영화 소비 데이터를 분석하여, 어떤 영화 캐릭터가 인기를 끌고 있는 지 또는 인기를 끌 수 있는 지를 분석하여 그 수요를 정확하게 예측하는 데 기여할 수 있을 것이다.

마지막으로, 영화 산업뿐만 아니라, 다양한 인물이 존재하는 애니메이션 산업 또는 이모티콘 캐릭터 산업에도 적용될 수 있을 것으로 보인다. 본래 캐릭터 산업은 영화 이전에 애니메이션 분야에서 크게 두각을 드러내 왔으며, 최근에는 카카오톡 또는 라인과 같은 모바일 메신저의 확산과 함께 메신저 내 이모티콘에서도 그 시장성이 커지고 있다. 이모티콘 캐릭터 간 스토리나 세계관이 구축된다는 점에서 향후 이모티콘

캐릭터들의 인기도는 더욱 중요해질 수 있으며, 구축된 스토리 또는 출시 이모티콘들의 특징을 분석하여 캐릭터간 네트워크를 측정할 수 있을 것이다. 이와 같이 다른 산업에서도 본 연구를 적용하여 캐릭터 간의 네트워크를 분석하여, 캐릭터들의 인기를 예측하여 관련 상품들의 수요를 측정할 수 있을 것이다.

5.3. 미래 연구를 위한 제언

현재의 연구 결과를 기반으로 아래의 추가 연구가 기대 된다. 첫째, 적은 편수의 마블 영화만을 대상으로 하여 본 연구의 결과를 모든 영화에 일반화하여 적용하기 어렵다. 총 23편의 마블 영화 대본 데이터만으로 분석이 진행되었기에 다른 영화들에서는 비슷한 캐릭터 네트워크 특징이 발견되지 않을 수 있다. 대부분의 영화들이 핵심 주인공들과 그 외 캐릭터들로 스토리가 구성된다는 것을 고려하면 영화들의 캐릭터 네트워크 특징 차이는 크지 않을 수 있다. 그러나, 마블 영화들은 시리즈 형태로 제작되며, 짧은 주기로 다량의 작품이 개봉되기 때문에 영화의 흥행도, 캐릭터 종류 및 캐릭터 간 관계에 있어 충분한 다양성을 보여준다. 더불어 마블 스튜디오는 두터운 팬 층을 보유하여 일반적으로 개봉 전부터 많은 관심을 받는다. 이에, 흥행을 거두지 못하여 대중의 충분한 관심을 받지 못하는 (Kumar et al. 2014) 타 영화 대비 높은 검색량 데이터를 확보할 수 있어 캐릭터 네트워크를 파악하기에 적합한 분석 대상이 될 수 있다.

다음으로, 본 연구는 공개 데이터를 기반으로 영화 캐릭터 간 네트워크를 분석하였다. 이에 공개된 정보인 캐릭터들의 대사관계만을 통하여 캐릭터 네트워크를 구성하였다. 하지만 캐릭터 간 네트워크를 파악하는 변수에는 캐릭터 특징, 개별 캐릭터의 검색량에 영향을 미치는 요인 등 영화 내/외부적 요인이 존재한다

(Lazer et al. 2014). 예를 들어, 영화 대본의 특징 외에도 출연 배우의 특징, 영화 내 다양한 특수 효과 및 장치, 광고, 감독의 인지도 등 영화 흥행에 영향을 미치는 요소 등이 검색량 데이터에 편차를 줄 수 있다. 이에 후행 연구에서는 내부 데이터 등 다양한 데이터셋을 기반으로 유기적인 캐릭터 네트워크를 분석할 수 있을 것이다. 이로써 문학 작품 연구에서 시도되는 벡터값을 도입한 SNA 방법론 (Lee and Jung 2020) 및 기계학습 (Machine Learning) 알고리즘을 적용하여 캐릭터 중심성 척도와 검색량 간 예측력 수준을 좀 더 높일 수 있을 것이다.

마지막으로, 작품 및 캐릭터의 실질적인 인기를 측정할 수 있는 다른 변수들을 고려하여 네트워크 특징의 영향을 살펴볼 수 있을 것이다. 예를 들어, 실질적인 인기도는 검색량 자체 보다는 해당 작품 또는 캐릭터 상품의 실제 판매량을 통해 측정할 수 있다. 검색량은 실제 구매량을 예측하기 위한 중간 척도이기에, 실제로 캐릭터 네트워크 특징에 따른 사람들의 행동 양상을 살피기 위해서는 캐릭터 관련 상품의 매출 및 판매량을 살펴보아야 할 것이다. 어떤 캐릭터의 이모티콘이 더 많이 판매되었고 어떤 영화 캐릭터 인형 또는 코스튬이 많이 판매되고 있는 지를 살펴보면서 네트워크 특징이 사람들의 실질적인 소비 행위에 미치는 영향을 측정한다면, 본 연구의 경영학적인 기여도는 더욱 높아질 것이다.

6. 결론

본 연구는 영화 대본에서 추출한 캐릭터의 네트워크 특징이 인간의 검색 행위에 영향을 준다는 것을 발견하였다. 데이터셋으로는 현 시점까지 개봉한 마블 스튜디오의 모든 영화 (총 23편) 대본을 활용하였다. SNA 방법론을 적용하여 각 영화별 캐릭터 네트워크를 구축하였으며, 이후 개별 캐릭터들에 주요 네트워크 중심성 척도 값들을 부여하였다. 해당 중심성 척도 값들과 각 캐릭터의 영화 개봉 이후 구글 검색량 간의 상관관계를 확인하여 캐릭터의 매개중심성 (Betweenness) 척도가 캐릭터 검색량과 높은 상관성을 지닌다는 점을 확인하였다. 이에 더하여 매개중심성 척도가 검색량 데이터에 수치적으로 높은 예측력을 지닌다는 점을 회귀 분석을 통해 확인하였다.

기존의 선행연구들이 SNA 방법론을 사용하여 미디어 작품 자체를 깊이 있게 분석했던 것에서 나아가, 미디어 작품의 네트워크 특징이 인간의 행동을 유발할 수 있다는 점 (Hineline 2018)에 착안하여 진행된 본 연구가 유효성을 거둔 것이다. 캐릭터 중심성 척도와 같은 미디어 작품의 네트워크 특징이 인간 행동에 영향을 줄 수 있다는 결과를 통해, 미디어 연구와 인간 행동 연구 사이에 새로운 연구의 발판을 마련하는 학문적 기여를 제공하였다.

또한, 인간의 검색 행위가 실제 단순히 사람의 탐색 행위 및 관심 수준을 넘어서 실제 인간의 행동 혹은 구매 행위로 이어질 수 있다는 점에서 실무적 시사점이 존재한다. 캐릭터 네트워크 중심성 척도가 해당 캐릭터에 대한 인기 수준으로 활용될 수 있으며, 그 점에 착안하여 캐릭터 네트워크 중심성 척도는 캐릭터의 가치 수준을 판단하는 참고 지표로 사용될 수 있다. 기존의 캐릭터 라이선스 협상이 라이선스 권리 보유사 (Licensor)와 권리 활용사 (Licensee)의 정성적 협

상에 근거하여 모든 캐릭터에 대해 일괄적으로 동일한 수준의 로열티를 부여하였다는 것을 고려하면, 본 연구 방법론을 활용하여 수치적으로 개별 캐릭터에 적정 수준의 로열티를 책정할 수 있을 것이다.

〈참고문헌〉

[국내 문헌]

1. 고영희, 이서현 (2016). 미디어 콘텐츠 기업의 무형자산 중심 지식재산 가치 연결 전략: 아이코닉스 애니메이션 뽀로로에 대한 탐색적 사례연구. **지식경영연구**, 17(3), 181-206.
2. 박영수, 최성호 (2020). 지식재산평가의 적정성에 대한 연구: 기술보증기금의 특허가치평가를 중심으로. **지식경영연구**, 21(4), 195-210.
3. 박지혜 (2017). 초연결시대의 협력: IT 기업 간 협력 네트워크와 성과에 관한 연구. **Information Systems Review**, 19(2), 21-35.
4. 이선영, 이승진 (2015). 애니메이션의 게이미피케이션과 캐릭터 라이선싱 연계 연구: 〈타닝메카드〉를 중심으로. **만화애니메이션연구**, 41, 357-378.
5. 이유석, 차경전, 김상훈 (2016). 인터넷 정보 검색 행동과 영화 흥행의 상관관계에 대한 연구. **경영학연구**, 45(5), 1501-1526.
6. 장리, 최강준, 이재영 (2017). 온라인 구전량 및 평점과 시기별 영화 흥행과의 관계. **지식경영연구**, 18(2), 65-83.

[국외 문헌]

7. Aggarwal, C., & Subbian, K. (2014). Evolutionary network analysis: A survey. **ACM Computing Surveys (CSUR)**, 47(1), 1-36.
8. Alberich, R., et al. (2014). *Marvel universe looks almost like a real social network (2002)*. arXiv preprint cond-mat/0202174.
9. Allee, V. (2000). The value evolution: Addressing larger implications of an intellectual capital and intangibles perspective. **Journal of Intellectual Capital**, 1(1), 17-32.
10. Allee, V. (2008). Value network analysis and value conversion of tangible and intangible assets. **Journal of Intellectual Capital**, 9(1), 5-24.
11. Alvarez-Hamelin, J. I., et al. (2005). *K-core decomposition: A tool for the visualization of large scale networks*. arXiv preprint cs/0504107.
12. An, H. S. (2018). *Essays on consumer behavior in*

contemporary music market in the US: Millennial's perception. Rutgers University-Graduate School-Newark.

13. Andina-Díaz, A., & García-Martínez, J. A. (2020). Reputation and news suppression in the media industry. **Games and Economic Behavior**, 123, 240-271.
14. Bavelas, A. (1950). Communication patterns in task-oriented groups. **The Journal of the Acoustical Society of America**, 22(6), 725-730.
15. Bazzan, A. L. (2020). I will be there for you: Clique, character centrality, and community detection in friends. **Computational and Applied Mathematics**, 39(3), 1-25.
16. Benesty, J., et al. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4). Springer.
17. Blasco-Arcas, L., et al. (2013). Adopting television as a new channel for e-commerce. The influence of interactive technologies on consumer behavior. **Electronic Commerce Research**, 13(4), 457-475.
18. Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. **Journal of Mathematical Sociology**, 2(1), 113-120.
19. Bonacich, P. (1987). Power and centrality: A family of measures. **American Journal of Sociology**, 92(5), 1170-1182.
20. Bossaert, G., & Meidert, N. (2013). "We are only as strong as we are united, as weak as we are divided" A dynamic analysis of the peer support networks in the Harry Potter books. **Open Journal of Applied Sciences**, 3(02), 174-185.
21. Botta, F., et al. (2020). In search of art: Rapid estimates of gallery and museum visits using Google Trends. **EPJ Data Science**, 9(1), 14.
22. Castilla-Polo, F., & Gallardo-Vázquez, D. (2016). The main topics of research on disclosures of intangible assets: A critical review. **Accounting, Auditing & Accountability Journal**, 29(2), 323-356.

23. Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*, 2–9.
24. Csermely, P., et al. (2013). Structure and dynamics of core/periphery networks. *Journal of Complex Networks*, *1*(2), 93–123.
25. Darlington, R. (1990). Multiple tests. In *Regression and linear models* (pp. 249–276). McGraw–Hill.
26. Ding, L., & Yilmaz, A. (2010). Learning relations among movie characters: A social network perspective. In *European Conference on Computer Vision*, Springer.
27. Edwards, M., et al. (2020). The one comparing narrative social network extraction techniques. *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, IEEE.
28. Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, *22*(3), 329–354.
29. Eliashberg, J., et al. (2007). From story line to box office: A new approach for green–lighting movie scripts. *Management Science*, *53*(6), 881–893.
30. Eliashberg, J., et al. (2014). Assessing box office performance using movie scripts: A kernel–based approach. *IEEE Transactions on Knowledge and Data Engineering*, *26*(11), 2639–2648.
31. Frangidis, P., et al. (2020). Sentiment analysis on movie scripts and reviews. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, Springer.
32. Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*(1), 35–41.
33. Freeman, L. C., et al. (1979). Centrality in social networks: II. Experimental results. *Social Networks*, *2*(2), 119–141.
34. Grant, R., et al. (2007). A review of factors affecting online consumer search behaviour from an information value perspective. *Journal of Marketing Management*, *23*(5–6), 519–533.
35. Harrison, S., et al. (2019). MARVEL'S BLOCKBUSTER MACHINE How the studio balances continuity and renewal. *Harvard Business Review*, *97*(4), 136–+.
36. Hineline, P. N. (2018). Narrative: Why it's important, and how it works. *Perspectives on Behavior Science*, *4*(2), 471–501.
37. Hu, Y., et al. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research*, *51*(3), 300–319.
38. John, M., et al. (2019). A visual approach for the comparative analysis of character networks in narrative texts. *2019 IEEE Pacific Visualization Symposium (PacificVis)*, IEEE.
39. Jun, S. P., et al. (2014). A study of the method using search traffic to analyze new technology adoption. *Technological Forecasting and Social Change*, *81*, 82–95.
40. Kulkarni, G., et al. (2012). Using online search data to forecast new product sales. *Decision Support Systems*, *52*(3), 604–611.
41. Kumar, A., et al. (2014). Information discovery and the long tail of motion picture content. *MIS Quarterly*, *38*(4), 1057–1078.
42. Kwak, D. (2010). A study on the purchase intention of character product and the attribute factors of character design. *The Journal of Korean Society of Design Culture*, *16*(3), 11–23.
43. Labatut, V., & Bost, X. (2019). Extraction and analysis of fictional character networks: A survey. *ACM Computing Surveys(CSUR)*, *52*(5), 1–40.
44. Laffel, M., et al. (2020). Movie script similarity using multilayer network portrait divergence. In *International Conference on Complex Networks and Their Applications*, Springer.
45. Lazer, D., et al. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, *343*(6176), 1203–1205.
46. Lee, O. J., & Jung, J. J. (2020). Story embedding: Learning distributed representations of stories based on character networks. *Artificial Intelligence*, *281*, 103235.
47. Masías, V. H., et al. (2017). Exploring the prominence of Romeo and Juliet's characters using weighted centrality measures. *Digital Scholarship in the Humanities*, *32*(4), 837–858.

48. Molloy, J. C., et al. (2011). Making intangibles "tangible" in tests of resource-based theory: A multidisciplinary construct validation approach. *Journal of Management*, *37*(5), 1496–1518.
49. O'Connor, N., & Kim, S. (2016). Media-related tourism phenomena: A review of the key issues. In *Mediating the tourist experience* (pp. 13–32). Routledge.
50. Opsahl, T. (2013). Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks*, *35*(2), 159–167.
51. Shrum, L. (2012). *The psychology of entertainment media: Blurring the lines between entertainment and persuasion*. Taylor & Francis.
52. Sparavigna, A. C. (2013). On social networks in plays and novels. *International Journal of Sciences*, *2*(10).
53. Stiller, J., et al. (2003). The small world of Shakespeare's plays. *Human Nature*, *14*(4), 397–408.
54. Tsai, C. M., et al. (2013). Scene-based movie summarization via role-community networks. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(11), 1927–1940.
55. Vespignani, A. (2009). Predicting the behavior of techno-social systems. *Science*, *325*(5939), 425–428.
56. Wang, Y. S., et al. (2013). What drives purchase intention in the context of online content services? The moderating role of ethical self-efficacy for online piracy. *International Journal of Information Management*, *33*(1), 199–208.
57. Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press
58. Waumans, M. C., et al. (2015). Topology analysis of social networks extracted from literature. *PloS One*, *10*(6), e0126470.
59. Weng, C. Y., et al. (2009). Rolenet: Movie analysis from the perspective of social networks. *IEEE Transactions on Multimedia*, *11*(2), 256–271.
60. Yadav, M., & Srivastava, D. M. K. (2020). A study of changing consumer trends in the entertainment industry. *Iconic Research and Engineering Journals*, *4*(4), 9–16.

● 저 자 소 개 ●



조 성 현 (Seonghyun Cho)

KAIST 경영대학 석사과정을 졸업했으며, SK 이노베이션 배터리 사업부에 재직 중이다. 현재 사업기획본부 운영최적화실에서 투자사업 경제성 평가, 합작법인 (JV) 추진, 운영 프로세스 및 원가 혁신 업무를 담당하고 있다. 주요 관심 분야는 프로젝트 가치평가, 사회 구조망 네트워크 분석 등이다.



이 민 형 (Minhyung Lee)

KAIST 경영대학에서 경영공학 박사 학위를 취득하였다. 주요 관심분야는 디지털 콘텐츠, 공유경제, 에너지 IT, 정보 시스템의 사회적 효과, 비즈니스 애널리틱스 및 빅데이터 기반 최적화 전략 등이다. 지금까지 Decision Support System 등 주요 학술지에 논문을 발표하였다.



최 한 별 (HanByeol Stella Choi)

KAIST 경영대학 디지털혁신연구센터 연구조교수로 재직 중이다. 주요 관심분야는 프라이버시, 정보 보안, 공유 경제, 정보 시스템의 사회적 효과 및 비즈니스 애널리틱스 등이다. 지금까지 Journal of Management Information Systems, Decision Support System, Security Journal 등 주요 학술지에 논문을 발표하였다.



이 희 석 (Heeseok Lee)

KAIST 경영대학 교수로 재직 중이다. 아리조나대학 경영학 박사 취득 후 네브라스카대학 교수를 역임하였다. 주요 관심분야는 IT Strategy와 Digital Innovation이다. MIS Quarterly, Journal of Management Information Systems, Information and Management 등 주요 학술지에 논문을 발표하였다.

〈 Abstract 〉

A Study on the Estimation of Character Value in Media Works: Based on Network Centralities and Web-Search Data

Seonghyun Cho^{*}, Minhyung Lee^{**}, HanByeol Stella Choi^{***}, Heeseok Lee^{****}

Measuring the intangible asset has been vigorously studied for its importance. Especially, the value of character in media industry is difficult to quantitatively evaluate in spite of the industry's rapid growth. Recently, the Social Network Analysis (i.e., SNA) has been actively applied to understand human usage patterns in a media field. By using SNA methodology, this study attempts to investigate how the character network characteristics of media works are linked to human search behaviors. Our analysis reveals the positive correlation and causality between character network centralities and character search data. This result implies that the character network can be used as a clue for the valuation of character assets.

Key Words: Intangible Asset, Knowledge Management, Media Industry, Character Value, Character Network, Social Network Analysis (SNA), Search Data

* KAIST College of Business

** KAIST College of Business

*** KAIST College of Business

**** KAIST College of Business