



트위터 네트워크의 정보 전파과정 분석 - 유력자 및 하이퍼링크 효과

Analysis of the Information Diffusion Process on Twitter: Effects of Influentials and Hyperlinks

저자 (Authors) 이재현, 김찬균
Jae Hyun Lee, Chan Kyun Kim

출처 (Source) [한국언론학보 56\(3\)](#), 2012.6, 238-265 (29 pages)
[Korean Journal of Journalism & Communication Studies 56\(3\)](#), 2012.6, 238-265 (29 pages)

발행처 (Publisher) [한국언론학회](#)
Korean Society For Journalism And Communication Studies

URL <http://www.dbpia.co.kr/Article/NODE01896126>

APA Style 이재현, 김찬균 (2012). 트위터 네트워크의 정보 전파과정 분석. 한국언론학보, 56(3), 238-265.

이용정보 (Accessed) 아주대학교
202.30.23.141
2016/03/08 15:04 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다.

이 자료를 원저작자와의 협의 없이 무단게재 할 경우, 저작권법 및 관련법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

The copyright of all works provided by DBpia belongs to the original author(s). Nurimedia is not responsible for contents of each work. Nor does it guarantee the contents.

You might take civil and criminal liabilities according to copyright and other relevant laws if you publish the contents without consultation with the original author(s).

트위터 네트워크의 정보 전파과정 분석:

유력자 및 하이퍼링크 효과*

이재현**

(서울대학교 언론정보학과 교수)

김찬균***

(서울대학교 언론정보학과 석사)

트위터의 사회문화적 영향력과 관심을 반영하듯, 선행 연구에서는 정보 전파에 대한 다양한 연구가 진행되어왔다. 하지만 단순한 리트윗 받은 수만을 정보 전파를 분석하는 지수로 사용하여, 다차원적인 트위터 내 정보 전파의 과정을 분석하지 못하였다. 이 논문은 이러한 한계를 인식하고 선행 연구에서 사용된 정보 전파를 측정하는 다양한 지수를 활용하여 트위터 네트워크의 정보 전파과정을 분석하였다. 또한 기존 연구에서 정보 전파에 영향을 미치는 메시지 속성에 주목하고 있지 않다는 점을 한계로 지적하고, 트위터 내 정보 전파의 과정을 트윗을 작성한 이용자의 속성, 즉 유력자의 여부와 트윗의 메시지 속성, 즉 하이퍼링크 유무에 따라 그 차이를 검증하였다. 분석 결과, 유력자의 트윗은 비유력자의 트윗보다 더 많이, 더 깊게 전파되었으며, 더 오래 지속되었고, 비하이퍼링크 트윗은 하이퍼링크 트윗에 비하여 더 많이, 더 오래 지속되었다. 이용자 속성과 메시지 속성의 상호작용 효과를 분석한 결과, 유력자가 생산한 비하이퍼링크 트윗이 전파 정도가 가장 높은 것으로 나타났으며, 선형 회귀분석 결과, 이용자 속성은 메시지 속성보다 정보 전파에 미치는 영향이 더 큰 것으로 나타났다.

Key words: 트위터, 리트윗, RT, 정보 전파, 유력자, 하이퍼링크

* 이 연구는 2012년 서울대학교 언론정보연구소 연구기금의 지원을 받았습니다.

** leejh@snu.ac.kr

*** covi19xx@snu.ac.kr

1. 문제의 제기

트위터(Twitter)는 2006년 3월 미국에서 처음으로 서비스를 개시한 대표적인 소셜 네트워크 서비스(social network service; SNS)다. 트위터 이용자는 140자 이내로 자신의 생각, 느낌, 정보 등을 표현한다. 140자라는 제한은 휴대전화 문자 메시지의 형식을 차용한 것으로 다른 SNS들에 비해 트위터만의 가장 큰 특징이다. 이외에도 트위터는 기존의 SNS와는 다른 몇 가지 특징을 가지고 있다.

트위터는 타인과의 관계를 맺음에 있어서 호혜성(reciprocity)을 원칙으로 하지 않는다. 트위터 이용자가 다른 사람과 관계를 맺고자 하면 “팔로우”(follow)를 클릭하면 된다. 자신을 팔로우하는 이용자는 팔로워(follower)가 되며, 자신이 다른 이용자를 팔로우한다면 자신은 다른 이용자의 팔로잉(following)이 된다. 팔로잉과 팔로워가 서로 일치할 필요가 없기 때문에 호혜적이지 않다. 또한 트위터는 리트윗(retweet)이라는 독특한 커뮤니케이션 수단을 갖고 있다. 이용자가 작성한 트윗은 팔로워(follower)의 게시판인 타임라인(timeline)에 노출이 되고, 그것을 본 이용자가 리트윗을 하면 다시 자신의 모든 팔로워에게 노출되는 방식이다. 이러한 일방향 관계 맺기와 리트윗의 메커니즘(mechanism)은 트위터의 신속하고 광범위한 정보 전파를 더욱 촉진시켰다. 트위터의 광범위한 정보 전파력으로 인해 나타나는 다양한 현상들은 사회정치적으로 논의를 불러왔고, 학술적으로도 주요한 연구대상으로 부상했다.

트위터의 정보 전파력에 대한 연구는 정보 전파 유발 주체의 관점에서 유력자(influential)를 규명하는 연구(Java, et al., 2007; Krishnamurth, et al., 2008; Backshy, et al., 2010; Cha, et al., 2010; Kwak, et al., 2010; Romero, et al., 2010; 이원태 외, 2010; 2011)와 트위터 정보 전파(information diffusion)의 특징을 분석한 연구(Asur, & Huberman, 2010; Lerman, & Ghosh, 2010; Suh, et al., 2010; Yang, & Counts, 2010; 장덕진·김기훈, 2011)로 크게 분류할 수 있다.

두 가지 범주의 선행 연구들은 초기 트위터 내 정보 전파와 이를 유발하는 주체에 대한 사람들의 관심을 학문적으로 접근했다는 점에서 그 의의를 찾을 수 있다. 하지만 트위터의 정보 전파력과 관련한 선행 연구들을 분석한 결과, 대부분 연구들은 몇 가지 한계를 지니고 있었다. 이 논문은 트위터의 정보 전파에 관한 선행 연구의 한계를 극복하고, 이론적, 방법론적으로 트위터 연구에 기여하는데 그 목적을 갖고 있다. 선행 연구의 한계와 이 논문의 목적을 보다 구체적으로 기술하면 다음과 같다.

첫째, 방법론적 측면에서, 기존 트위터 연구들은 단순한 지수만을 활용하여 트위터의 이용자, 정보 전파 특성의 현상을 기술하는데 그쳤다는 한계가 있다. 기존 연구에서는 정보 전파의 특성을 분석하기 위해 주로 리트윗 받은 수를 그 지수로 활용하고 있다. 그러나 리트윗 받은 수로는 전체적인 정보 전파의 대략적 특성을 볼 수는 있지만, 정보 전파의 시간적 차

원, 정보 전파의 넓이, 깊이 등 정보 전파의 다차원적인 특성을 알 수 없다. 따라서 이 논문에서는 선행 연구에서 정보 전파를 분석하는데 사용된 다양한 지수를 종합, 복합적으로 활용하여 선행 연구의 방법론적 한계를 극복하고자 한다.

둘째, 이론적 측면에서, 기존 트위터 연구는 정보 전파에 영향을 미치는 속성으로 유력자에 집중하여, 정보 전파에 영향을 미치는 또 다른 속성을 간과하고 있다. 이 논문은 선행 연구에서 주목하지 않은 메시지의 속성, 하이퍼링크(hyperlink)를 중점적으로 분석했다. 트윗을 비하이퍼링크 트윗(non-hyperlink tweet)과 하이퍼링크가 포함된 하이퍼링크 트윗(hyperlink tweet)으로 구분하고, 하이퍼링크가 트윗의 정보 전파에 어떠한 영향을 미치는지 분석했다. 정보 전파에 영향을 미치는 독립 변인으로서 이용자 속성과 더불어 메시지 속성을 고려하여 선행 연구의 이론적 한계를 극복하고자 한다.

셋째, 연구 대상의 측면에서, 기존 연구는 한국인 트위터 네트워크 내 정보 전파에 대한 연구가 적다는 한계가 있다. 한국인 트위터 네트워크를 연구한 기존 연구(장덕진·김기훈, 2011)에서 밝혔듯이, 한국인 트위터 네트워크의 특성은 평균 경로간 거리가 짧고 호혜성이 높은 등 전 세계 트위터 네트워크와 다르다. 따라서 한국인 트위터 네트워크 내 정보 전파의 특징은 전 세계의 네트워크의 특징과 다를 것으로 예측할 수 있고, 이에 대한 분석이 필요하다고 할 수 있다.

위와 같이, 이 논문은 초기에 시도되었던 트위터 정보 전파 연구의 한계를 극복하고, 트위터 정보 전파 연구에 방법론적, 이론적으로 기여하는데 그 목적이 있다. 이를 위해 다양한 지수를 복합적으로 활용하고, 이를 통해 정보 전파의 다차원적 특징을 살펴보고자 한다. 또한 정보 전파에 영향을 미치는 메시지 속성을 분석함으로써 이용자 속성만 주목한 선행 연구의 이론적 한계를 극복하고자 한다. 마지막으로, 한국인 트위터 네트워크의 특징이 전 세계 트위터 네트워크와 다르다고 보고, 한국인 트위터 네트워크 내 정보 전파의 특징을 연구하고자 한다.

2. 이론적 논의

1) API를 활용한 트위터 연구

아직 초창기 서비스인 현 시점에서 트위터에 대한 이해, 그리고 그런 이해를 포함하는 연구 경향에 대한 정리 및 평가를 시도하는 것은 중요한 작업이다. 트위터는 API(Application Programming Interface)가 공개돼 있기 때문에, 많은 연구가 API를 활용한 데이터를 수집하여 연구를 수행해 오고 있다. 이재현(2012a)에 따르면, API를 활용한 기존 연구들은 트위터의 이용자 행태와 트위터 네트워크를 보는 관점에 따라 두 가지 차원으로 분류할 수 있다.

〈표 1〉 관점에 따른 트위터 선행 연구 분류

		이용자 행태	
		팔로잉 행태	트윗 행태
네트워크 관점	개인 네트워크	A 개별 네트워크	B 개인 트윗
	전체 네트워크	C 전체 네트워크	D 개인 트윗

첫 번째 차원은 이용자의 행태다. 이는 다시 다른 트위터 이용자와 관계를 맺는 팔로잉 행태, 그리고 트윗 작성, 리트윗(retweet), 멘션(mention) 등의 트윗 행태로 구분된다. 두 번째 차원은 연구 대상으로서 트위터를 보는 관점이다. 전통적인 사회관계망 연구에서처럼 이것은 개인 네트워크(ego-centered) 관점과 전체 네트워크(whole network) 관점으로 구분할 수 있다(Garton, et al., 1999/2000, pp. 322~326). 이러한 기준에 따라 트위터의 선행 연구를 네 가지로 분류한 것을 표로 나타내면 〈표 1〉과 같다.

A유형은 개별 트위터 이용자가 팔로잉(following)을 통해 형성한 대인적 관계망을 분석한 연구들이다. A유형에 해당하는 대표적인 연구는 개별 이용자의 팔로잉 수, 팔로워 수 등으로 개인의 영향력(Java, et al., 2007; Krishnamurth, & Arlitt, 2008; Cha, et al., 2010; Kwak, et al., 2010; Romero, et al., 2010; 이원태 외, 2010), 능동성(Huberman, et al., 2008)을 측정하고, 이용자를 유형화(Krishnamurth, & Arlitt, 2008; Chu, et al., 2010)하는 연구들이다. 대부분 A유형의 연구는 이용자의 프로필(profile) 정보를 수집하여 팔로잉 수, 팔로워 수를 지수로 사용한 연구다. A유형의 연구는 개별 사용자들의 특성을 파악했다는 점에서 의의를 갖고 있다.

B유형은 개별 트위터 이용자가 작성한 트윗, 리트윗, 멘션 등으로 구성되는 트윗 줄기(tweets threads), 그리고 트윗의 전파과정을 다룬 연구들이다. B유형에 해당하는 대표적인 연구는 개별 이용자의 트윗 수, 리트윗 받은 수, 리트윗 한 수 등을 지수로 활용하여 유력자를 검증한 연구들(Java, et al., 2007; Huberman, et al., 2008; Krishnamurth, & Arlitt, 2008; Cha, et al., 2010; Kwak, et al., 2010; Lerman, & Ghosh, 2010; Romero, et al., 2010; Yang, & Counts, 2010; 이원태 외, 2010)이다. 그리고 개별 이용자가 작성한 트윗의 전파를 분석한 연구들(Cha, et al., 2010, Suh, et al., 2010)도 B유형에 해당된다. B유형의 연구는 개별 트윗의 정보 전파과정을 분석하였다는 점에서 의의가 있다. 하지만 대부분의 A, B 유형 연구는 이론적 논의가 부족하고 단순히 현상을 기술했다는 한계가 있다.

C유형은 전체 트위터 네트워크 혹은 특정 속성을 공유하는 집단에서 팔로잉을 통해 형성된 관계망을 분석한 연구다. C유형에 해당하는 대표적인 연구는 전 세계 트위터 네트워크를 분석한 연구(Kwak, et al., 2010), 한국인 트위터 네트워크를 분석한 연구(장덕진·김기훈,

2011)가 있다. 이들 연구들은 개별 트위터 이용자의 정보를 토대로 이용자 분포, 밀도, 호혜성(reciprocity), 유사성(homophily), 이용자 사이의 거리(degree of separation) 등을 분석했다. C유형 연구들은 대규모 데이터를 수집하여 네트워크 분석을 통해 트위터 네트워크의 지형을 파악하였다는 점에서 의의가 있다.

D유형은 집합체 수준에서의 트윗 내용과 트윗의 전파과정을 분석한 연구들이다. D유형의 대표적인 연구는 집합체 수준에서 이용자들 트윗의 전파 특성(Asur, & Huberman, 2010; Lerman, & Ghosh, 2010; Yang, & Counts, 2010), 트윗 내용의 특성(Kwak, et al., 2010), 트윗의 감성(sentiment)에 대한 분석(Asur, & Huberman, 2010; Tumasjan, et al., 2010; Thewell, et al., 2011) 등이 있다. 대규모 데이터를 활용한 자연어 처리, 감성 분석 등의 내용분석(content analysis)이 D유형 연구에서 주로 활용된 방법론이다.

이 논문은 개별 이용자 트윗의 전파과정을 분석하기 위한 지수로 구성되어 있다. 따라서 B유형에 속한다고 할 수 있다. 하지만 이 논문은 정보 전파에 미치는 영향으로서 메시지 속성, 그 중에서도 하이퍼링크에 초점을 맞췄으며, 기존 연구에서 주목한 사용자 속성과의 상호작용 효과까지 분석했다는 점에서 기존 B 유형의 연구와 차별성을 갖고 있다. 그리고 다양한 지수를 활용하여 정보 전파과정을 다각도로 분석했다는 것도 기존 B 유형의 연구와의 차별점이다.

2) 트위터의 정보 전파

리트윗(retweet)은 트위터 정보 전파의 핵심적인 메커니즘(mechanism)이다(Nam, et al., 2011). 트위터에서 사용자가 작성한 트윗은 팔로워(follower)의 게시판인 타임라인(timeline)에 노출이 되고, 그것을 본 이용자가 리트윗을 하면 자신의 모든 팔로워에게 다시 노출된다. 이러한 리트윗 방식으로 트윗은 빠르고 광범위하게 전파된다. 게다가 트위터는 실명으로 가입할 필요가 없다. 트위터 이용자는 실명을 노출할 필요 없이 트윗을 작성할 수 있고, 다른 사람의 트윗 내용을 재전송할 수 있다. 이러한 트위터의 익명성과 개방성은 정보의 전파를 더욱 빠르고 광범위하게 만든다. 트위터의 정보 전파력은 양면의 효과를 낳았다. 날씨, 교통 등의 긴급 정보를 전파, 공유하는 것은 트위터 정보 전파의 긍정적인 예시이다. 하지만 많은 양의 정보가 진위 여부가 제대로 검증되지 않은 채 큰 책임 없이 전달되어 괴담과 같은 근거 없는 정보가 빠르고 광범위하게 유포돼 사회적으로 혼란을 가져오는 부정적 효과를 낳기도 한다.

트위터를 통한 정보 전파의 영향력이 확인되면서 이와 관련한 많은 연구가 활발하게 이루어지고 있다. 대부분의 연구들(Cha, et al., 2010; Kwak, et al., 2010; Romero, et al., 2010; 이원태 외, 2010)은 트위터 정보 전파에 영향을 미치는 사용자 속성에 중점을 두고 있다. 이들 연구에서 공통적으로 주장하고 있는 바는, 트윗이 많이 리트윗된 이용자가 유력자

라는 것이다. 이외에도 정보 전파에 영향을 미치는 메시지 속성에 대한 연구도 있다. 서 외 (Suh, et al., 2010)는 메시지 속성 측면에서 URL, 해시태그 등이 담긴 트윗이 리트윗될 확률과 이용자 속성 측면에서 팔로워 수, 팔로워 수, 트윗 수 등을 고려해 작성한 트윗이 리트윗될 확률을 “리트윗 가능성”(retweetability)으로 개념화하고, 메시지 속성과 이용자 속성이 정보 전파에 미치는 영향을 검증했다. 이 외에도 다양한 관점에서 트위터의 정보 전파와 관련한 연구가 이루어지고 있다. 마케팅 관점에서 특정 상품의 인기도를 측정한 연구(Asur, & Huberman, 2010), 트위터 내 스토리의 인기도를 측정한 연구(Lerman, & Ghosh, 2010), 특정 이슈의 전파과정과 특징을 분석한 연구(장덕진·김기훈, 2011) 등이 대표적이다. 또, 양과 카운츠(Yang, & Counts, 2010)는 다양한 지수를 활용하여 정보 전파의 특징을 보다 다각도로 분석하였다. 이들은 트윗 작성자와 팔로워 관계를 맺고 있으면서 리트윗을 한 이용자를 1단계, 트윗 작성자와 팔로워 관계는 없지만 1단계 이용자들과 팔로워 관계가 있고 리트윗을 한 이용자를 2단계 등으로 정보 전파의 단계를 순차적으로 나누어서 트위터 정보 전파의 특징을 분석했다. 이렇듯 트위터 정보 전파와 관련한 기존 연구들에서 사용된 지수를 정리한 내용은 다음 <표 2>와 같다.

트위터의 정보 전파력과 관련한 선행 연구들을 분석한 결과, 기존 연구들은 몇 가지 한계를 지니고 있었다. 기존 연구의 한계와 이를 토대로 설정한 본 논문의 목적과 방향을 제시하면 다음과 같다.

첫째, 기존 트위터 연구들은 리트윗 받은 수만을 정보 전파를 분석하는 지수로 활용하였다는 방법론적 한계를 지니고 있다. 따라서 많은 연구들은 정보가 얼마나 멀리 전파되었는가, 그리고 얼마나 빠르고 오래 유지되는가와 같은 정보 전파의 다양한 차원을 분석하지 못하였다. 그리고 리트윗 받은 수만을 지수로 활용한 것은 이용자 속성이 정보 전파에 미치는 영향을 분석하는데 한계로 작용하였다. 다시 말해서, 리트윗 받은 수가 많은 이용자를 유력자로 정의하여, 리트윗 받은 수가 많은 이용자가 작성한 트윗은 리트윗 받은 수가 많다는 동어반복(同語反覆)의 오류를 일으키고 있다. 따라서 이 논문에서는 선행 연구에서 정보 전파를 분석하는데 사용된 다양한 지수를 종합, 복합적으로 활용하였다. 이를 통해, 정보가 얼마나 빨리, 멀리, 넓게 전파되는지와 얼마나 오래 지속되는지와 같은 정보 전파의 다양한 차원을 분석했다.

둘째, 기존 트위터 연구는 정보 전파에 영향을 미치는 속성으로 유력자에만 집중하여, 정보 전파에 영향을 미치는 메시지 속성을 간과하고 있다는 이론적 한계를 지니고 있었다. 이 논문은 선행 연구에서 주목하지 않은 메시지의 속성 중에서 하이퍼링크(hyperlink)를 중점적으로 분석했다. 트윗을 비하이퍼링크 트윗(non-hyperlink tweet)과 하이퍼링크가 포함된 하이퍼링크 트윗(hyperlink tweet)으로 구분하고, 하이퍼링크가 트윗의 정보 전파에 어떠한 영향을 미치는지 분석했다. 추가적으로 이 논문에서는 정보 전파에 영향을 미치는 이용자 속성과 메시지 속성의 상호작용을 분석했다.

〈표 2〉 선행 연구의 정보 전파와 관련된 주요 개념과 지수

구성적 개념	지수		수식	연구
	개념	조작적 정의		
일련의 스토리 사이즈 분포	팔로워(팬)의 리트윗 수 (1단계)	팔로워 리트윗 수와 이에 해당하는 스토리 수		Lerman, & Ghosh (2010)
	팬 팔로워의 리트윗 수 (2단계)	팬의 리트윗 수와 이에 해당하는 스토리 수		
스토리 인기도	스토리가 리트윗된 전체 수	특정 수가 리트윗된 스토리 수		
리트윗 가능성	특정 콘텐츠 속성을 포함하는 리트윗 비율(도메인, 해시태그)	특정 콘텐츠 속성을 포함하는 트윗이 리트윗될 확률	특정 콘텐츠 속성을 포함하는 리트윗 수/ 특정 A를 포함하는 트윗 수*전체 트윗 수/ 전체 리트윗 수 특정 이용자 속성에 해당하는 리트윗 수/ 특정 이용자 속성 해당하는 트윗 수* 전체 트윗 수/전체 리트윗 수	Suh, et al. (2010)
	특정 이용자 속성에 해당하는 리트윗 비율 (팔로워 수, 팔로워 수, 트윗 수, 이용 기간, favorite 수)	특정 이용자 속성에 해당하는 트윗이 리트윗 될 확률		
정보 전파	스피드(speed)	첫 번째 전파가 언제 생기는지, 생성 여부		Yang, & Counts (2010)
	범위(scale)	첫 번째 단계에서 영향 받은 노드의 수		
	거리(range)	전파 체인이 계속될 수 있는 길이의 수		
토픽 관심도	개봉 전 관심도	특정 영화에 대해 언급한 URL, 리트윗 비율	특정 영화에 대해 언급한 URL, 리트윗/ 전체 트윗×100 트윗 수/시간	Asur, & Huberman (2010)
	개봉 주 관심도	트윗비율, 시간 당 특정 영화를 언급한 트윗 수		
유저 참여도	토픽에 참여하는 유저 수	토픽에 대해 언급한 이용자 수		

셋째, 한국인 트위터 네트워크 내 정보 전파에 대한 연구가 적다는 한계가 있었다. 한국인 트위터 네트워크를 연구한 기존 연구(장덕진·김기훈, 2011)에서 밝혔듯이, 한국인 트위터 네트워크의 특성은 평균 경로간 거리가 짧고 호혜성이 높은 등 전 세계 트위터 네트워크와 다르다. 즉, 한국인 트위터 네트워크는 전 세계 네트워크와 비교해서 더 동질적인 팔로워를 공유하고 있는 특징을 보이고 있다는 것이다. 이는 한국어라는 공통된 언어를 사용하는 한국인 트위터 네트워크는 그 수가 적고, 지리적으로도 가깝기 때문이다. 따라서 한국인 트위터 네트워크 내 정보 전파의 특징은 전 세계의 네트워크의 특징과 다를 것으로 예측할 수 있고, 이에 대한 분석이 필요하다.

3) 트위터의 유력자

정보 전파에 영향을 미치는 이용자 속성으로 유력자(influential)에 관한 논의의 출발은 전통적 유력자로서 의견지도자(opinion leader)이다. 라자스펠트 외(Lazarsfeld, et al., 1948)는 커뮤니케이션 2단계 흐름(two step flow)을 파악하는 과정에서 의견지도자의 특성을 규명했다. 이들은 미디어에서 생성된 정보가 의견지도자를 거쳐 수용자에게 전달된다고 주장했다. 이후 연구에서 카츠와 라자스펠트(Katz, & Lazarsfeld, 1955)는 의견지도자들이 대개 높은 사회경제적 위치에 있으며, 일반인들보다 미디어의 이용 정도가 더 높다고 밝힌바 있다. 전통적으로 유력자는 그들이 생성 혹은 중계하는 메시지가 광범위하게 전파되어 영향력을 행사하는 사람들로 이해된다. 이 같은 유력자 개념은 트위터에도 적용된다. 트위터에서는 리트윗, 전통적 의미로 달리 말하면 소위 구전(word of mouth)을 통해 정보가 전파되는데(Nam, et al., 2011), 유력자가 생성한 정보가 구두 전파를 통해 많은 사람에게 영향을 미친다는 관점은 기존 의견지도자의 역할과 그 맥을 같이 한다.

그러나 트위터에서 유력자는 기존의 유력자와는 다른 특징을 보이기도 한다. 이외수, 김제동 등과 같이 오프라인(off-line) 유명인들이 트위터의 유력자도 되는 경우도 많지만, 트위터 유력자의 상당수는 기존 전통 매체 의제 설정자와는 달리 일반인이다. 누구나 유력자가 될 수 있는 개방성은 트위터 유력자의 행위에도 영향을 미친다. 누구나 트위터의 유력자가 될 수 있는 상황에서 트위터의 유력자는 자신과 자신이 생성한 정보에 대한 팔로워들의 평판에 민감하며(이원태 외, 2010), 자신의 주장을 효과적으로 부각시키기 위해 노력한다(김은미, 이준용, 2004). 또한 트위터 유력자는 스스로 정보를 개조, 창작하는 “창작자”(creator)이면서(Li, & Bernoff, 2008; 이원태 외, 2010, p. 41에서 재인용), 다른 이용자의 트윗을 적극적으로 전파하는 매개자의 모습(intermediary influential)을 보이기도 한다(이원태 외, 2011).

트위터 정보 전파력과 이를 유발하는 주체, 즉 유력자에 대한 관심을 반영하듯 많은 연구자들이 유력자를 규명하기 위해 연구를 진행하였다. 기존 트위터 연구에서 유력자를 규명하

는 시도는 두 가지로 나눌 수 있다. 첫째는 이용자의 팔로잉 행태 차원이고, 둘째는 이용자의 트윗 행태 차원이다. 먼저, 팔로잉 행태 차원에서는 팔로워 수를 지수로 활용하여 유력자를 규명한 연구(Krishnamurth, et al., 2008; Cha, et al., 2010; Kwak, et al., 2010; Romero, et al., 2010)가 있다. 자바 외(Java, et al., 2007)는 팔로잉 행태에 기반한 HIT 알고리즘(algorithm)을 개발, 이용자의 허브(hub) 값과 권위(authority) 값을 측정하여 유력자를 검증하였다. 둘째, 트윗 행태 차원에서 트윗 수를 지수로 활용하여 유력자를 규명한 연구(Krishnamurth, et al., 2008)가 있다. 그러나 대부분의 연구(Backshy, et al., 2010; Cha, et al., 2010; Kwak, et al., 2010; Romero, et al., 2010; 이원태 외, 2010; 2011)들은 리트윗 받은 수만을 지수로 활용하여 유력자를 규명했다는 한계를 갖고 있다. 리트윗한 메시지의 속성까지는 고려하지 못하고 있다.

4) 트위터의 하이퍼링크

하이퍼링크(hyperlink)는 하이퍼텍스트를 정의함에 있어서 가장 핵심적인 개념(이재현, 2004)으로 인터넷이 등장하면서부터 의미론적 기능에서 그 중요성이 부각돼 왔다. 버블스(Burbules, 2001)는 하이퍼링크가 두 노드 사이의 의미적 관계를 독자에게 강하게 암시하여, 결국 특정 정보에 접근하는 독자를 통제(control)하는 역할을 한다고 주장한다. 하이퍼링크의 기능을 제시한 모건(Morgan, 2002)은 하이퍼링크가 독자의 읽는 순서를 결정하며, 두 노드 사이의 간극(gap)을 이어주고, 독자로 하여금 끊임없이 읽기의 위치를 전환하는 위치 재설정(repositioning)의 감각을 부여하고, 의미를 생성하려는 독자의 욕구를 활성화한다고 주장했다.

하이퍼링크가 이용자의 인지에 어떤 영향을 미치는가에 대한 연구는 컴퓨터 매개 커뮤니케이션(computer-mediated communication, CMC) 관점에서 주로 수행되었다. CMC 분야에서 하이퍼링크는 정보의 신뢰도(credibility)에 영향을 미치는 요소(Eysenbach, & Kohler, 2002; Freeman, & Spyridakis, 2004)로 연구되었다. 일련의 연구(Flanagin, & Metzger, 2000; Fogg et al., 2001; Flanagin, & Metzger, 2007) 결과, 하이퍼링크는 정보에 대한 이용자의 인지된 신뢰도에 정적 영향을 미치는 것으로 나타났다. 하이퍼링크로 인하여 정보의 진위 여부를 바로 입증(verification)할 수 있기 때문이다(Flanagin, & Metzger, 2007). 즉, 수많은 정보가 있는 인터넷에서 신뢰도 있는 정보를 소비하려는 이용자에게 하이퍼링크는 긍정적인 영향을 미친다.

트위터에서 하이퍼링크는 트윗 메시지 속성 중 하나이다. 특히, 트위터에서 하이퍼링크가 중요한 이유는 트위터가 갖고 있는 기술적 한계 때문이다. 트위터에서 한 번에 올릴 수 있는 내용은 140자 이내로 한정돼 있기 때문에, 음악, 비디오, 사진과 같은 다른 미디어를 포함시키거나 정보를 더 많이 담기 위해서는 하이퍼링크의 사용이 필수적이다. 이런 점에서

하이퍼링크 트윗은 확장된 트윗(extended tweet)이다(김찬균, 2012). 이재현(2012b)도 하이퍼링크가 SNS에서 텍스트를 확장하는 중요한 요소라고 주장하였다. 그리고 하이퍼링크 자체가 의미론적 기능을 갖고 있기 때문에 하이퍼링크 트윗의 사용은 의미론적으로도 논의될 수 있다. 하이퍼링크의 사용은 트윗을 통해 자신의 의견을 주장하고, 하이퍼링크를 통해 그 근거를 제시하는 바람직한 논의 구조를 낳게 한다(장덕진·김기훈, 2011). 이를 수사적(rhetorical) 차원으로 보면, 하이퍼링크 트윗은 준거의 트윗(referent tweet)이다. 다시 말해, 하이퍼링크는 트윗 내에서 주장을 뒷받침하는 근거를 연결하는 논리적 준거(logical reference)와 관련된 추가 사실적 정보를 연결하는 정보적 준거(informational reference)로 기능한다(김찬균, 2012).

결국, 하이퍼링크는 트윗 정보의 내용과 양에 모두 영향을 미치는 중요한 요소이다. 하지만 이러한 중요성에도 불구하고 하이퍼링크가 정보 전파에 미치는 영향에 대해 주목한 선행 연구는 많지 않으며, 아직 기술적(descriptive)인 수준에 머물러 있다. 장덕진·김기훈(2011), 백사이(Bakshy, et al., 2011) 등은 트윗에 포함된 하이퍼링크의 출처와 유형을 분류했다. 이외에 하이퍼링크 출처와 유형의 비율을 근거로 트위터 이용자를 유형화한 연구(Java, et al., 2007; Chu, et al., 2010), 하이퍼링크 트윗의 리트윗 정도로 이용자의 영향력을 측정하는 연구(Romero, et al., 2010; Bakshy, et al., 2011), 하이퍼링크가 정보 전파에 미치는 정도를 측정하는 연구(Suh, et al., 2010; Yang, & Counts, 2010) 등이 있다. 이들 모두 단순히 트윗에 포함된 하이퍼링크의 기능과 유형 등을 정리했을 뿐 정보 전파에 미치는 영향을 직접적으로 다루지는 않고 있다.

5) 연구 문제 및 가설

트위터에서 유력자는 많은 수의 팔로워 혹은 추종자(supporter)를 보유하고, 사회정치적 의제설정 및 여론 형성 과정에서 큰 영향력을 발휘하는 사람이다(이원태 외, 2011). 트위터와 관련한 기존 연구들 중 유력자 여부가 정보 전파에 미치는 영향을 분석한 경우는 드물었다. 이는 단순히 리트윗 받은 수만으로 정보 전파를 측정하면 리트윗 받은 수를 기준으로 유력자를 규명하는 시도가 동어반복(同語反覆)의 오류가 되기 때문이다. 따라서 이 논문에서는 정보 전파를 측정함에 있어서 리트윗 받은 수뿐만 아니라 양과 카운츠(Yang, & Counts, 2010)가 사용한 거리, 속도, 범위를 지수로 활용했다. 이에 더해, 허니컷과 헤링(Honeycutt, & Herring, 2009)이 사용한 길이(length) 개념을 차용했다. 이들은 트위터 대화를 분석함에 있어서 대화가 시작된 시점과 끝난 마지막 시점의 차이를 길이로 개념화했다. 이 논문에서는 길이를 활동 기간으로 재개념화하고, 정보 전파의 전체 시간을 분석하기 위해 활용했다.

선행 연구 분석을 통해 살펴봤듯이 정보 전파에 결정적 역할을 하는 유력자는 트위터에서

도 마찬가지로일 것으로 추정된다. 따라서 이 논문에서는 유력자가 작성한 트윗은 비유력자가 생산한 트윗보다 정보 전파의 모든 측면에서 우세할 것이라고 가정했다. 이러한 논의를 바탕으로 이 논문은 다음과 같은 연구 문제 및 검증 가설을 설정했다.

- 연구문제 1: 유력자의 여부는 트윗의 정보 전파에 어떠한 영향을 미치는가?
- 연구가설 1-1: 유력자의 트윗은 비유력자가 생산한 트윗에 비해 리트윗의 범위가 더 넓을 것이다.
- 연구가설 1-2: 유력자의 트윗은 비유력자가 생산한 트윗에 비해 리트윗의 거리가 더 멀 것이다.
- 연구가설 1-3: 유력자의 트윗은 비유력자가 생산한 트윗에 비해 리트윗의 속도가 더 빠를 것이다.
- 연구가설 1-4: 유력자의 트윗은 비유력자가 생산한 트윗에 비해 리트윗의 활동 기간이 더 길 것이다.

하이퍼링크 트윗은 트윗에 하이퍼링크가 포함된 트윗을 말한다. 트위터는 140자 이내의 텍스트로 한정되어 있기 때문에, 비디오, 사진, 음악 등과 같은 다른 미디어를 포함시키거나 140자 이상 장문의 정보를 작성하기 위해서는 하이퍼링크가 필수적이다. 이러한 점에서 하이퍼링크 트윗은 확장된 트윗 (extended tweet) 이다(김찬균, 2012; 이재현, 2012b). 또한 하이퍼링크는 트윗을 통해 자신의 의견을 주장하고 하이퍼링크를 통해 그 근거를 제시하는 바람직한 논의 구조를 낳게 한다는(장덕진·김기훈, 2011) 점에서, 하이퍼링크 트윗은 준거의 트윗 (referent tweet) 이다(김찬균, 2012). 이와 함께, 컴퓨터 매개 커뮤니케이션(CMC) 연구(Flanagin, & Metzger, 2000; Fogg et al., 2001; Flanagin, & Metzger, 2007)에서는 하이퍼링크가 포함된 정보가 이용자들의 인지된 신뢰도에 정적 영향을 미치는 것으로 나타나고 있다.

이러한 논의를 통해, 수많은 정보가 타임라인에 노출되는 트위터에서 하이퍼링크 트윗은 더 많은 정보를 담고 있고, 논리적이며 신뢰도가 높다고 할 수 있다. 이를 바탕으로 하이퍼링크 트윗이 비하이퍼링크 트윗에 비하여 정보 전파에 있어서 우세할 것으로 가정했다.

- 연구문제 2: 하이퍼링크 여부는 트윗의 정보 전파에 어떠한 영향을 미치는가?
- 연구문제 2-1: 하이퍼링크 트윗은 비하이퍼링크 트윗에 비해 리트윗의 범위가 더 넓을 것이다.
- 연구문제 2-2: 하이퍼링크 트윗은 비하이퍼링크 트윗에 비해 리트윗의 거리가 더 멀 것이다.
- 연구문제 2-3: 하이퍼링크 트윗은 비하이퍼링크 트윗에 비해 리트윗의 속도가 더 빠를 것이다.

- 연구문제 2-4: 하이퍼링크 트윗은 비하이퍼링크 트윗에 비해 리트윗의 활동 기간이 더 길 것이다.

마지막으로, 이 논문에서는 정보 전파에 영향을 미치는 이용자 속성과 메시지 속성의 상호작용 효과를 분석하고자 한다. 발화자의 속성과 메시지의 속성이 정보 전파에 미치는 영향은 커뮤니케이션학의 주요한 고전적 논쟁 중 하나이다. 이 논문에서 설정한 트위터 이용자의 속성은 팔로워 수, 트윗 한 개 당 평균 리트윗 받은 수로 정의된 유력자의 여부이며, 메시지의 속성은 트윗에 포함된 하이퍼링크의 유무이다.

- 연구문제 3: 정보 전파에 영향을 미치는 이용자 속성과 메시지 속성은 상호작용 효과가 있는가?

3. 연구 방법

1) 연구 대상

이 논문은 국내 트위터 이용자 1,002명을 대상으로 2011년 7월 11일부터 7월 17일까지 작성한 트윗 내용을 수집하였다. 조사 기간 동안 발생한 이슈들을 중심으로 95개의 메인 키워드와 서브 키워드를 선정했다.¹⁾ 수집한 트윗 내용 중 95개의 키워드를 포함하고 있는 내용을 추출한 결과 최종적으로 1,538개의 이니셜 트윗을 분석대상으로 선정했다. 이니셜 트윗의 정보 전파 트리 구조 작성을 위하여 이니셜 트윗에 대한 4,310개의 반응 트윗도 수집했다. 반응 트윗은 API 크롤링의 기술적 한계 상, 공식 리트윗과 공식 멘션의 경우로 한정했다. 공식 리트윗, 멘션은 트위터 인터페이스에서 제공하는 각각의 버튼을 클릭하여 생성된 것을 의미한다. 공식 리트윗과 멘션은 트위터 서버에 저장되어 API로 공개된다. 반면, 비공식 리트윗, 멘션은 각각에 해당되는 기호인 “RT”, “@”를 직접 입력하는 방법으로, 트위터 서버에 저장되지 않아 API 크롤링 방법을 통해 수집할 수 없었다.

조사 대상자의 전체 트윗 내용이 아니라 특정 키워드를 포함한 내용만을 수집한 이유는 한국인 트위터 이용자의 특성과 한국인 트위터 네트워크 내 정보 전파의 특성을 보려는 목적이다. 한국인 트위터 이용자가 생성하는 트윗의 양은 엄청나며, 이를 모두 분석할 수 없기 때문에 키워드와 이니셜 트윗의 생성시기를 통제했다. 또한, 키워드를 통제하고 이니셜 트윗의 생성 시간을 한정함으로써 정보 전파의 특징을 명확히 분석해보고자 했다. 시사적 내용이 담겨 있을 경우 좀 더 정보에 민감하고 파급력이 있기 때문이다.

1) 95개의 키워드는 첨부된 <부록 1>을 참고할 것.

2) 데이터 수집 및 분석 방법

정보 수집을 위한 API 호출은 IP 주소 당 한 시간에 350회로 제한된다. 따라서 연구를 위한 데이터 수집에는 오랜 시간, 많은 수의 컴퓨터가 필요했다. 이에 따라 데이터 수집을 위해 사용된 총 컴퓨터는 366대로 총 4개월에 걸쳐 데이터를 수집했다.²⁾

API 호출을 통한 데이터의 수집은 파이썬(Python) 언어 기반의 데이터 추출 모듈(module)을 만들어 이용했다. 정보 전파과정에서 드러나는 특징을 좀 더 명확하게 분석하기 위해 최초 트윗 발생을 시드(seed)로 간주하고 그를 기반으로 확산하는 전략을 채용해 정보 전파의 트리 구조(tree structure)를 조직했다. 트리 구조는 이니셜 트윗 작성자와 반응 트윗 작성자 간 맺고 있는 인맥 관계(following)를 기반으로 구성했다. 이니셜 트윗에 대해 반응한 작성자의 인맥 관계를 모두 추적하여, 선행 트윗 작성자를 팔로잉하고 있는지 여부를 확인했다. 이 과정을 통해, 분석 대상인 이니셜 트윗은 0단계, 이니셜 트윗 작성자를 팔로잉하고 있는 이용자의 반응 트윗은 1단계, 1단계 트윗을 작성한 이용자를 팔로잉하고 있는 이용자의 이니셜 트윗에 대한 반응 트윗은 2단계 등으로 트리 구조를 구성했다.

데이터는 총 3차례로 나누어 수집했다. 1차 수집 기간에는 트윗 수집 모듈을 준비하고, 시험 수집을 하였다. 2차 수집 기간에는 대규모 컴퓨터를 활용하여 분산 수집 모듈을 준비하고 반응 트윗의 확산 종료 시점까지 수집을 진행했다. 3차 수집 과정에서는 수집된 트윗을 작성 시간 순으로 재배열하고, 이를 이용자 사이의 팔로잉 관계에 기반 한 전파 단계의 트리 구조로 전환하는 작업을 진행했다.

API의 기술적 한계로 인하여, 리트윗 받은 수, 생성 리트윗 수, 하이퍼링크 트윗 수 등의 정보는 이용자가 계정 생성 이후 데이터 수집 당시 시점까지 생성한 트윗에서 일부를 표집해 수집했다. API 크롤링을 통해서만 이용자의 누적 트윗 수, 즉 이용자가 계정을 생성한 후에 수집 당시 시점까지 작성한 트윗 수만 수집할 수밖에 없기 때문이다. 이용자의 생성 트윗 수는 한번 시도 시 현재 시점으로부터 3,200개까지 수집이 가능했으며, 각 이용자의 트윗이 리트윗된 수는 100개 이내에서만 수집이 가능했다. 하이퍼링크 트윗 수는 “http” 문자열 검색을 통해 수집하였다.

구성된 트리 구조 데이터베이스는 SPSS 구문(syntax)을 통해 SPSS에서 통계 처리할 수 있는 형태로 변환했다. 마지막으로, 연구 문제 및 가설을 검증하기 위하여 다변량 분산분석(MANOVA)과 선형 회귀분석(linear regression)을 실시했다.

2) 이렇게 수집된 본 데이터는 황유선·이재현(2011)의 연구에도 활용되었음.

3) 주요술어의 정의와 지수의 구성

이 논문의 <연구 문제 1>, <연구 문제 2>에 해당하는 검증가설의 독립변인과 <연구 문제 3>의 주요 변인은 각각 유력자 여부, 하이퍼링크 유무이다. 유력자는 팔로워 수와 리트윗 영향력 모두 전체 이용자의 상위 25%에 해당하는 이용자로 정의하였다. 반대로 비유력자는 팔로워 수와 리트윗 영향력 모두 전체 이용자의 하위 25%에 해당하는 이용자로 정의하였다. 이처럼 집단을 사분위수로 나눈 이유는 집단 간의 차이를 보다 극명하게 보기 위함이다. 팔로워 수는 이용자의 프로필을 근거로 하여 특정 이용자의 트윗을 구독하는 이용자의 수로 정의했다. 리트윗 영향력은 이용자가 생성한 트윗 한 개당 평균 리트윗 받은 수로 정의했다. 하이퍼링크 트윗은 문자열 검색을 통해 “http”가 트윗에 포함된 트윗으로 정의했다. 반대로 비하이퍼링크 트윗은 “http”가 포함되어 있지 않은 트윗으로 정의하였다.

이 논문의 <연구 문제 1>, <연구 문제 2>에 해당하는 검증가설의 종속변인은 정보 전파의 범위, 거리, 속도, 활동 기간이다. 이 지수들은 정보 전파의 시발(始發)이 되는 ‘이니셜 트윗(initial tweet)’을 생산한 이용자의 팔로워 관계에 기반 한 트리 구조(tree structure)를 기초로 정의됐다. 범위는 트윗을 작성한 이용자와 팔로워 관계가 있으면서 트윗을 리트윗한, 즉 1 단계에서 생성된 리트윗 수로 정의했다. 거리는 트리 구조에서 특정 트윗이 전파된 최대 단계의 수로 정의했다. 속도는 이니셜 트윗의 생성 시간과 1 단계에서 가장 빨리 생성된 리트윗의 생성 시간의 차로 정의하였다. 활동 기간은 이니셜 트윗의 생성 시간과 정보 전파 트리 구조에서 가장 늦게 생성된 리트윗의 생성 시간의 차로 정의했다. 각각의 종속변인을 지수화하면 다음과 같다.

- 범위 = 트윗 작성자 팔로워의 리트윗 수
- 거리 = 트윗 작성자(a)와 트윗 작성자(a)의 팔로워가 아닌 이용자 중 트윗 작성자(a)의 트윗을 리트윗한 이용자(b) 간의 단계 수
- 속도 = (트윗 생성 시간) - (최초 리트윗의 생성 시간)
- 활동 기간 = (트윗 생성 시간) - (마지막 리트윗 생성 시간)

4. 연구 결과

1) 주요 변인의 기술통계

이 논문의 연구 대상은 1,002명의 이용자와 이들이 작성한 1,538개의 이니셜 트윗이다. 연구 대상에 대한 기술적 통계 결과를 간략히 제시하면 <표 3>, <표 4>와 같다.

〈표 3〉 이용자의 속성

(단위 = 명, 건, %, N = 1,002)

행태	지수	M	SD	비고
팔로잉 행태	팔로워 수	12,129.8	23714.7	이용자의 계정 생성 이후부터 누적된 총 수치임.
트윗 행태	리트윗 영향력	1.4	4.0	이용자의 계정 생성 이후부터 누적된 총 트윗 에서 일부를 표집하여 수집함.(N: 1,971,842)
	하이퍼링크 트윗 포함 비율	43.6	29.3	

〈표 4〉 이니셜 트윗의 정보 전파 특성

(단위 = 건, 단계, 시: 분: 초, N = 1,538)

	Min	Max	M	SD
리트윗 받은 수	1.0	81.0	3.1	6.1
범위	1.0	74.0	2.8	5.1
거리	1.0	7.0	1.1	0.5
속도	0:00:07	248:39:58	2:04:14	13:38:29
활동 기간	0:00:07	248:39:56	6:20:17	18:29:30

〈표 3〉에서 보는 바와 같이, 1,002명의 평균 팔로워 수는 약 12,130명, 트윗 한 개당 받은 평균 리트윗 수는 1.4개다. 이처럼 연구 대상의 평균 팔로워 수가 상대적으로 큰 것은 이 논문이 적어도 한 번 이상 리트윗된 트윗을 작성한 이용자를 중심으로 연구대상을 수집했기 때문이다. 즉, 리트윗 여부를 기준으로 삼았기 때문에 어느 정도 영향력이 있는 이용자가 수집되어 평균 팔로워 수가 일반적 수치보다 높게 나타난 것이다. 이들이 생산한 트윗 중 43.6%는 하이퍼링크 트윗이었다. 1,538개 이니셜 트윗의 정보 전파 특성을 살펴본 결과, 평균 리트윗 받은 수는 3.1개이며, 범위는 평균 2.8개였다. 이를 통해 한국인 트위터 네트워크에서 정보 전파는 거의 대부분 1단계에서 끝남을 알 수 있다. 이니셜 트윗의 평균 거리는 1.1이었고, 속도는 2시간 4분 14초, 활동 기간은 6시간 20분 17초로 나타났다. 이러한 수치는 팔로워 수를 제외하면, 장덕진·김기훈(2011)이 분석한 수치와 흡사하다. 이들 연구에서 보고한 한국인 평균 팔로워 수는 72명, 평균 리트윗 받은 수는 2.5개, 활동 기간은 7시간이었다.

다음으로, 이 논문에서는 수집된 하이퍼링크 트윗의 출처를 분석하였다. 수집된 1,538개 이니셜 트윗 중 하이퍼링크를 포함하고 있는 트윗은 모두 1,176개였다. 이 하이퍼링크들의 출처는 ‘뉴스 미디어’, ‘블로그/커뮤니티/SNS’, ‘트윗 서비스’, ‘기타’, ‘알 수 없음’ 등으로 분

〈표 5〉 하이퍼링크의 출처

(단위 = 건, %, N = 1,176)

	뉴스 미디어	트윗 서비스	블로그/커뮤니티/ SNS	기타	알 수 없음
개수	817.0	179.0	107.0	31.0	42.0
비율	69.5	15.2	8.9	2.6	3.6

〈표 6〉 출처와 정보 전파

(단위 = 건, 단계, 시: 분: 초, N = 1,003)

	뉴스 미디어	블로그/커뮤니티/ SNS/트윗 서비스	T 값	전체
분석 트윗 수	817.0	286.0	-	1,176.0
범위	2.4	1.7	2.608**	2.4
거리	1.0	1.0	-0.105	1.0
속도	1:05:10	2:55:58	-1.868	10:01:04
활동 기간	4:18:16	6:47:56	-1.624	17:29:21

주: “기타”, “알 수 없음”으로 분류된 73개 트윗은 분석에서 제외함.

* $p < .05$, ** $p < .01$, *** $p < .001$

류했다. ‘뉴스 미디어’는 전체적으로 웹 페이지가 뉴스 장르별로 구분되어 있으며, 기사를 기자가 작성한 웹 페이지를 말한다. ‘블로그/커뮤니티/SNS’는 개인의 블로그, 카페 등 커뮤니티, 페이스북(facebook), 유튜브(Youtube)와 같은 SNS 등이다. ‘트윗 서비스’는 트위터 API를 이용한 서비스로 대표적인 예로 와이프로그(Yfrog) 등이 있다. ‘기타’는 정당, 회사 등의 기관 홈페이지, 혹은 외국 웹으로 연결된 것 등이다. 하이퍼링크와의 연결이 끊어져 출처를 알 수 없는 것은 ‘알 수 없음’으로 분류하였다. 하이퍼링크의 출처를 분류한 결과는 〈표 5〉와 같다. 분석 결과, 하이퍼링크 트윗 중 69.5%의 출처가 뉴스 미디어였다. 또, 수집된 하이퍼링크 각각의 정보 전파 정도를 측정했다. “기타”와 “알 수 없음”으로 분류된 73개의 트윗은 분석에서 제외한, 1,003개의 트윗을 분석한 결과는 〈표 6〉과 같다.

2) 분석 결과 및 논의

〈연구 문제 1〉에서 제시한 가설은 트위터 네트워크에서 유력자 여부가 정보 전파에 미치는 영향을 검증하기 위한 것이다. 〈연구 문제 2〉와 이에 대한 가설은 하이퍼링크의 유무가 정보 전파에 미치는 영향을 검증하기 위하여 구성되어 있다.

〈표 7〉 주요 변인의 집단 별 기술 통계치 I

(단위 = 건, 단계, 시: 분: 초)

	유력자(N = 131)		비유력자(N = 139)		하이퍼링크 유(N = 173)		하이퍼링크 무(N = 97)	
	M	SD	M	SD	M	SD	M	SD
범위	4.2	8.6	0.7	0.7	1.6	2.9	3.9	9.5
거리	1.0	0.4	0.7	0.8	0.9	0.5	0.9	0.9
속도	1:40:38	5:05:24	1:56:54	7:08:08	2:08:17	7:17:46	1:14:39	3:31:35
활동 기간	8:46:33	19:08:04	2:45:01	8:14:08	3:49:23	10:20:16	8:58:28	20:15:07

- 주: a) 유력자: 팔로워 수 12,190 이상, 리트윗 영향력 1.01 이상인 집단.
 b) 비유력자: 팔로워 수 1,766 이하, 리트윗 영향력 0.07 이하인 집단.
 c) 하이퍼링크 유, 무: 유력자와 비유력자가 생산한 트윗 중에서 하이퍼링크가 있는 트윗과 없는 트윗으로 분류함.

〈연구 문제 1〉에 대한 분석 결과, 〈표 7〉에서 볼 수 있듯이 유력자 트윗의 범위는 4.2, 거리는 1.0, 속도는 1시간 40분 38초, 활동 기간은 8시간 26분 33초로 나타났다. 비유력자 트윗의 범위는 0.7, 거리는 0.7, 속도는 1시간 56분 54초, 활동 기간은 2시간 45분 01초로 나타났다. 각각의 지수에 대한 통계적 검증을 실시한 결과, 유력자의 트윗은 비유력자의 트윗보다 범위가 컸고($F = 35.2, p < .001$), 거리가 멀었으며($F = 14.5, p < .001$), 활동 기간이 더 길었다($F = 18.4, p < .001$). 그러나 속도는 유력자와 비유력자 간의 차이가 통계적으로 유의미하지 않는 것으로 나타났다($F = 0.0, p > .05$). 따라서 〈연구 문제 1〉에 해당하는 〈연구 가설 1-1〉, 〈연구 가설 1-2〉, 〈연구 가설 1-4〉는 지지되었으나, 〈연구 가설 1-3〉는 검증되지 않았다. 전체적으로 유력자가 생산한 트윗은 비유력자의 트윗보다 더 멀리 전파됐으며, 트윗 내용이 더 오래 전파되고 있음을 알 수 있다.

위 분석 결과에서 주목할 점은 정보 전파의 거리가 유력자와 비 유력자 간에 차이가 크지 않다는 것이다. 이는 한국인 트위터 네트워크의 동질성(同質性) 때문으로 풀이된다. 〈표 4〉에서 보듯이, 한국인 트위터 이용자 트윗의 평균 전파 거리는 1이며, 그 최댓값도 7이다. 즉, 유력자의 트윗을 리트윗하는 팔로워와 비유력자의 팔로워가 비교적 동질하다는 것이다. 트위터에서 정보는 이질적인 이용자 집단에게 전파될 때 전파 거리가 커진다. 하지만 위의 결과는 대부분의 이용자 집단이 비교적 동질적인 인맥을 갖고 있다는 것을 보여준다. 이는 한국인 트위터 네트워크의 이용자 수가 적고, 한국어라는 동일한 언어를 사용하는 집단으로 한정되어 있기 때문으로 해석할 수 있다. 또 하나 주목할 점은 유력자와 비유력자 간의 전파 속도가 차이가 통계적으로 의미가 없다는 것이다. 이는 전파의 속도가 유력자의 여부를 떠나 정보의 내용에 큰 영향을 받는다고 해석할 수 있다. 전파 속도는 이용자

가 즉각적으로 정보를 가치 있다고 판단하느냐에 따라 크게 달라진다. 이용자들은 정보가 가치있고 새롭거나 자극적일 경우 이를 빠르게 리트윗한다. 이때 누가 트윗을 했느냐는 이때 중요하지 않다. 트위터에서 피답이 빠르게 퍼지는 이유도 이를 통해 설명이 가능하다. 피답은 자극적이기 때문에 누가 작성했는지, 타당성이 있는지는 크게 고려대상이 되지 않는 것이다.

〈연구 문제 2〉에 대한 분석 결과, 〈표 7〉에서 볼 수 있듯이 하이퍼링크 트윗의 범위는 1.6, 거리는 0.9, 속도는 2시간 08분 17초, 활동 기간은 10시간 20분 16초였다. 비하이퍼링크 트윗의 범위는 3.9, 거리는 0.9, 속도는 1시간 14분 39초, 활동 기간은 8시간 58분 28초였다. 통계적 차이를 검증한 결과, 비하이퍼링크 트윗은 하이퍼링크 트윗에 비해 범위가 컸으며 ($F = 12.5, p < .001$), 활동 기간도 길었다 ($F = 9.7, p < .01$). 반면, 거리와 속도는 하이퍼링크의 유무에 따라 유의미한 통계적 차이를 보이지 않았다. 결국, 〈연구 문제 2〉에 해당하는 모든 검증 가설은 검증되지 않았으나, 〈연구 가설 2-1〉, 〈연구 가설 2-4〉에 해당하는 결과는 모두 유의미하게 반대로 나왔다. 다시 말하면, 하이퍼링크는 정보 전파의 거리와 속도에는 영향을 미치지 못했으나, 정보 전파의 범위와 활동 기간에는 부적 영향을 미친 것이다.

이러한 결과는 하이퍼링크가 정보 전파의 과정에 있어서 “거리의 마찰”(friction of distance)을 강화하는 기제로 작용한다고 해석할 수 있다. 이는 하이퍼링크의 기술적 속성과 내용적 속성, 두 가지 차원에서 해석할 수 있다. 기술적 속성 차원에서 보면, 모바일 환경의 경우 많은 이용자들은 하이퍼링크를 클릭하고 이를 확인하는 과정을 번거롭게 느낀다. 이 때문에 정보 전파에 부적 영향을 미치는 것이다. 내용적 속성 차원에서 보면, 이용자는 기존 미디어에서 생산한 정보에 대해 낮은 신뢰도를 보이며, 자신의 인적 네트워크 내에 있는 유력자의 정보를 더 가치 있는 것으로 판단한다고 해석할 수 있다.

마지막으로 〈연구 문제 3〉은 정보 전파에 영향을 미치는 속성인 이용자 속성과 메시지 속성의 상호작용 효과를 검증하기 위한 것이다. 〈표 8〉, 〈표 9〉은 〈연구 문제 3〉을 검증하기 위한 다변량 분산분석(MANOVA)의 결과를 제시한 것이다.

다변량 분산분석(MONOVA) 결과, 유력자 여부, 하이퍼링크 유무인 두 독립 변인이 범위, 거리, 속도, 활동 기간 등 종속변인들에 대한 상호작용 효과를 유발하고 있었다(Wilks' $\lambda = .94, p < .01$). 이를 각 지수 별로 구체적으로 살펴보면, 속도 ($F = 1.8, p > .05$), 거리 ($F = 0.6, p > .05$)는 상호작용 효과가 존재하지 않는 것으로 나타났다. 하지만 범위에서는 두 변수의 상호작용 효과가 나타났다. 즉, 유력자의 트윗이 비유력자의 트윗보다 하이퍼링크 유무에 따라 정보 전파의 범위에 큰 영향을 받고 있다 ($F = 12.3, p < .01$). 또한 활동 기간도 두 변수의 상호작용 효과가 있었는데, 유력자의 트윗이 비유력자의 트윗보다 하이퍼링크 유무에 따라 정보 전파의 활동 기간에 큰 영향을 받는 것으로 나타났다 ($F = 18.4, p < .001$). 이를 그림으로 표현하면 〈그림 1〉과 〈그림 2〉와 같다.

〈표 8〉 주요 변인의 집단 별 기술 통계치 II

(단위 = 건, 단계, 시: 분: 초)

	유력자 (N = 131)				비유력자 (N = 139)			
	하이퍼링크 유 (N = 86)		하이퍼링크 무 (N = 45)		하이퍼링크 유 (N = 87)		하이퍼링크 무 (N = 52)	
	M	SD	M	SD	M	SD	M	SD
범위	2.5	3.9	7.6	13.0	0.7	0.5	0.7	1.0
거리	1.0	0.4	1.1	0.4	0.7	0.5	0.8	1.0
속도	1:36:20	5:12:33	1:48:53	4:54:31	2:39:52	8:53:34	0:45:01	1:26:37
활동 기간	4:59:13	11:33:24	16:00:59	27:14:24	2:40:21	8:53:26	2:52:49	7:05:10

주: a) 유력자: 팔로워 수 12,190 이상, 리트윗 영향력 1.01 이상인 집단.
 b) 비유력자: 팔로워 수 1,766 이하, 리트윗 영향력 0.07 이하인 집단.

〈표 9〉 사용자 속성, 하이퍼링크 속성이 정보 전파에 미치는 영향

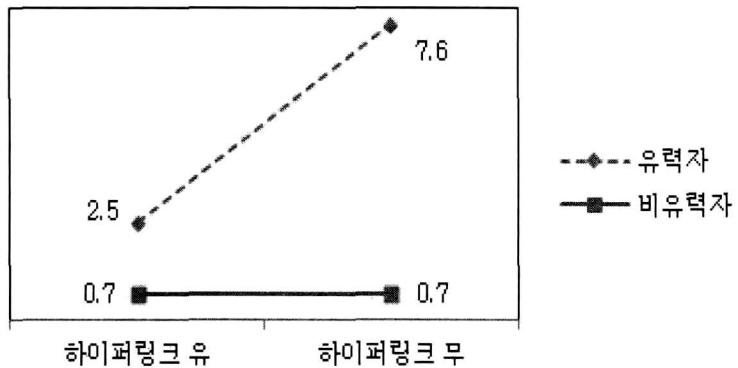
	SS	df	MS	F
중속 변인: 범위				
유력자 여부(A)	1163.3	1	1163.3	35.2***
하이퍼링크 유무(B)	414.7	1	414.7	12.5***
A × B	405.3	1	405.3	12.3***
중속 변인: 거리				
유력자 여부(A)	5.6	1	5.6	14.5***
하이퍼링크 유무(B)	0.3	1	0.3	0.9***
A × B	0.2	1	0.2	0.6***
중속 변인: 속도				
유력자 여부(A)	5812.2	1	5812.2	0.0***
하이퍼링크 유무(B)	5.8	1	5.8	1.2***
A × B	9.1	1	9.1	1.8***
중속 변인: 활동 기간				
유력자 여부(A)	4.8	1	4.8	18.4***
하이퍼링크 유무(B)	2.5	1	2.5	9.7***
A × B	2.4	1	2.4	9.0***

주: a) 유력자: 팔로워 수 12,190 이상, 리트윗 영향력 1.01 이상인 집단.
 b) 비유력자: 팔로워 수 1,766 이하, 리트윗 영향력 0.07 이하인 집단.

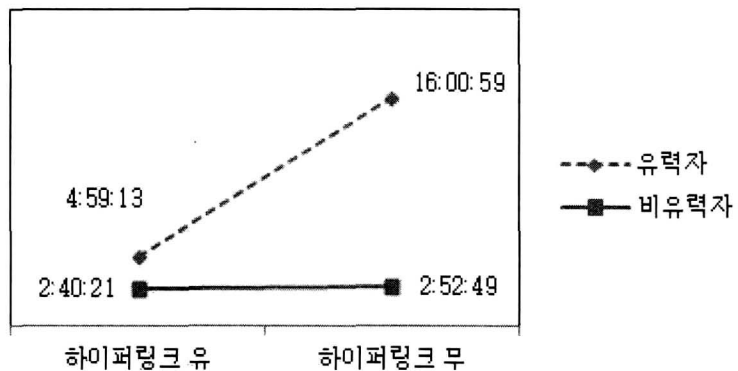
*p<.05, **p<.01, ***p<.001

〈그림 1〉과 〈그림 2〉에서 보듯이, 비유력자의 트윗은 하이퍼링크 유무에 따라 큰 차이를 보이지 않았으나, 유력자의 트윗은 하이퍼링크 유무에 따라 정보 전파의 범위와 활동 기간에 영향을 받는 것으로 나타났다. 비 유력자의 트윗은 하이퍼링크에 크게 영향을 받지 않았지만, 유력자의 트윗은 하이퍼링크에 따라 정보 전파에 큰 차이를 나타낸다. 이는 이용자들이 유력자가 생산한 비 하이퍼링크 트윗을 더 많이 전파한다는 것을 보여준다. 요약하면, 전체적으로 하이퍼링크는 정보 전파에 부적 영향을 미치지만, 비유력자 트윗의 정보 전파에는 크게 영향을 미치지 않는고 유력자 트윗의 정보 전파에 큰 영향을 미치는 것으로 나타났다.

추가적으로 이 논문에서는 이용자 속성과 메시지 속성 중 어떤 속성이 정보 전파에 더 큰 영향을 주는 것인지 검증하기 위하여 선형 회귀분석(linear regression)을 실시했다. 분석 결과는 〈표 10〉과 같다.



〈그림 1〉 정보 전파의 범위에 미치는 이용자 속성과 메시지 속성의 상호작용 효과



〈그림 2〉 정보 전파의 활동 기간에 미치는 이용자 속성과 메시지 속성의 상호작용 효과

〈표 10〉 정보 전파에 영향을 미치는 변인들의 회귀분석

변수	범위	거리	속도	활동기간
유력자 여부	.290***	.223***	-.024***	.209***
하이퍼링크 유무	-.190***	-.054***	.070***	-.173***
R ²	.117***	.052***	.005***	.071***
Adj. R ²	.110***	.045***	-.002***	.064***

*p<.05, **p<.01, ***p<.001

분석 결과, 범위, 활동 기간에서 두 독립변인 모두에 대한 유의미한 통계 값이 도출되었다. 〈표 10〉에서 보듯이, 정보 전파의 범위(R² = .117)는 유력자의 여부(β = .290, t = 5.044, p < .001)가 하이퍼링크 유무(β = -.190, t = -3.307, p < .01)보다 더 큰 영향을 미쳤다. 활동기간(R² = .071)도 유력자의 여부(β = .209, t = 3.534, p < .001)가 하이퍼링크 유무(β = -.173, t = -2.936, p < .01)보다 더 큰 영향을 미쳤다. 거리(R² = .052)는 유력자 여부(β = .223, t = 3.749, p < .001)와의 관계는 존재했지만, 하이퍼링크의 유무(β = -.054, t = -.905, p > .05)와의 관계는 존재하지 않았다. 속도(R² = .005)의 경우 유력자 여부(β = -.024, t = -.394, p > .05), 하이퍼링크 유무(β = .070, t = 1.144, p > .05)와의 관계는 존재하지 않았다.

지금까지의 결과를 토대로 논의해 보면, 세 가지 해석이 가능하다. 첫째, 정보 전파에 영향을 미치는 것은 유력자 여부였다. 즉, 트위터에서 정보 전파는 누가 정보를 생산했느냐는 정보 생산의 주체에 더 큰 영향을 받는다. 이와 같은 결과는 트위터 이용의 익명성, 개방성에 따른 정보의 양 차원에서 해석할 수 있다. 트위터는 익명성을 기반으로 누구나 정보를 생산할 수 있기 때문에 많은 양의 정보가 생성된다. 이때 이용자는 수많은 정보 중에서 전파의 판단 기준을 유력자의 여부로 선택한다고 해석할 수 있다.

둘째, 정보 전파에 영향을 미치는 것은 단순히 하이퍼링크 유무가 아닌, 하이퍼링크의 수사적 기능과 내용이다. 하이퍼링크는 입증 가능한 정보를 첨부해주는 기능을 수행한다(Flanagin, & Metzger, 2007). 즉, 하이퍼링크는 사실 혹은 주장에 대한 추가 정보나 근거를 첨부하는 수사적 기능을 한다(김찬균, 2012). 수사적 기능을 수행하는 하이퍼링크 트윗은 입증 가능한 정보가 첨부되어 논의의 여지가 비하이퍼링크 트윗보다 적다. 반면, 주장 혹은 사실에 대한 근거가 없는 비하이퍼링크 트윗은 논의의 여지가 많아 정보 전파의 범위와 활동 기간이 긴 것으로 해석된다. 특히, 유력자의 비하이퍼링크 트윗은 유력자라는 정보 생산 주체의 영향력과 그들의 의견, 주장과 같은 내용상의 영향력이 결합되어 더 많이 전파되고 더 오래 지속되는 것으로 풀이할 수 있다.

두 번째 해석은 트윗과 하이퍼링크의 내용 분석을 수행한 연구의 결과를 통해서도 뒷받침된다. 남 외(Nam, et al., 2011)는 사실 정보가 첨부된 하이퍼링크 트윗이 가장 전파 속도가

빠르며, 의견이 담긴 트윗, 의견과 하이퍼링크가 동시에 있는 트윗이 활동 기간이 더 길다고 주장했다. 김찬균(2012)의 연구 결과도 유사하다.³⁾ 그의 분석 결과, 정보 전파의 속도는 사실 정보가 첨부된 정보적 하이퍼링크가 가장 빠르지만 활동 기간은 정보적 하이퍼링크가 가장 느렸다. 결국, 사실 혹은 사실 정보에 첨부된 하이퍼링크는 논의의 여지가 없어 빠르게 전파되고, 의견 혹은 의견 정보에 첨부된 하이퍼링크는 논의의 여지가 많아 멀리 전파되고 오래 지속되는 것이다.

셋째, 트위터는 대안적 뉴스 미디어(alternative news media)의 성격을 일정 부분 지니고 있다. 분석 결과에서 봤듯이, 하이퍼링크와 연결된 출처는 대부분이 뉴스 미디어이고, 뉴스 미디어를 출처로 하는 하이퍼링크 트윗의 전파 범위와 활동 기간은 비하이퍼링크 트윗보다 전파 범위가 작고 활동 기간이 짧다. 이러한 차이는 유력자가 생산한 하이퍼링크, 비하이퍼링크 트윗의 정보 전파 결과와 비교했을 때 더 크게 나타났다. 이러한 결과는 트위터 이용자가 뉴스 미디어에서 생산한 정보가 아닌 자신의 인적 네트워크 내에 있는 이용자가 생산한 정보를 더 선호한다는 것을 나타낸다. 이는 트위터의 대안적 뉴스 미디어의 성격을 간접적으로 보여준다.

5. 맺음말

이 논문은 유력자의 여부, 하이퍼링크의 유무가 정보 전파에 미치는 영향, 그리고 유력자 여부와 하이퍼링크 유무가 정보 전파에 미치는 상호작용 효과를 검증했다. <연구 문제 1>에 대한 분석 결과, 유력자의 트윗은 비 유력자의 트윗보다 정보 전파의 범위가 더 넓고, 거리가 더 멀었으며, 활동 기간이 더 길었다. 이에 <연구 가설 1-1>, <연구 가설 1-2>, <연구 가설 1-4>는 지지되었으나, <연구 가설 1-3>은 검증되지 않았다. <연구 문제 2>에 대한 분석 결과, 하이퍼링크는 정보 전파의 범위와 활동 기간에 부적 영향을 미치는 것으로 나타났다. 이에 <연구 가설 2-1>, <연구 가설 2-4>는 기각되었으며, <연구 가설 2-2>, <연구 가설 2-3>은 검증되지 않았다. 마지막으로 <연구 문제 3>의 분석 결과, 유력자 여부와 하이퍼링크 유무 간의 상호작용은 존재했다. 그리고 선형 회귀 분석 결과를 통해 이용자 속성이 메시지 속성보다 정보 전파에 더 큰 영향을 미친다는 점을 알 수 있었다.

이러한 가설 검증 결과를 바탕으로 이 논문에서는 4가지의 결론을 도출하였다. 첫째, 한국인 트위터 이용자의 네트워크는 동질적이며, 이는 정보 전파에 영향을 미쳤다. 한국인 트위터 이용자는 한국어라는 동일한 언어를 공유하고 이용자 수가 적기 때문에 한국인 트위터 네트워크는 호혜성이 높고, 평균 경로 간 거리가 짧다(장덕진 · 김기훈, 2011). 이는 이용자

3) 위에 대한 보다 자세한 결과는 <부록 2> 혹은 김찬균(2012)의 서울대학교 석사 학위논문을 참고할 것.

들이 비교적 유사한 팔로워를 보유하고 있다는 점을 뜻한다. 이와 같이 동질적인 네트워크는 유력자 여부, 하이퍼링크 유무에 관계없이 1에 가까운 짧은 정보 전파의 거리에 영향을 미치고 있었다.

둘째, 유력자 여부는 하이퍼링크 유무보다 정보 전파에 더 큰 영향을 미쳤다. 트위터의 익명성과 개방성 때문에 트위터에서는 수많은 정보가 생산된다. 이용자들은 수많은 정보 중에서 전파 가치가 있는 트윗을 선택해서 리트윗한다. 이때 이용자들은 자신의 인적 네트워크 내에 있는 유력자를 판단의 유력한 기준으로 삼는다. 유력자가 생산한 트윗의 선호와 링크를 클릭하고 확인하는 번거로운 과정은 정보 전파에 있어서 하이퍼링크가 “거리의 마찰”을 강화하는 기제로 작용한다.

셋째, 하이퍼링크의 유무보다는 하이퍼링크 수사적 기능과 트윗의 내용이 정보 전파에 더 큰 영향을 미쳤다. 하이퍼링크는 의견, 사실 등을 담은 트윗에 근거 혹은 추가적 정보를 연결시킨다. 따라서 사실과 관계없는 의견, 주관적 사실 등의 트윗은 논란의 여지가 많기 때문에 전파의 활동 기간이 길다. 반면, 사실을 담고 있는 트윗은 논란의 여지가 없기 때문에 전파의 활동 기간이 짧다. 여기서 하이퍼링크는 일종의 촉매 작용을 한다. 의견, 주관적 사실을 담은 트윗에 하이퍼링크가 있는 경우에는 더욱 활동 기간을 길게 하며, 사실을 담고 있는 트윗에 하이퍼링크가 있는 경우에는 전파 속도를 더욱 빠르게 한다.

넷째, 트위터는 대안적 뉴스 미디어로 평가할 수 있다. 하이퍼링크로 연결된 출처는 대부분이 뉴스 미디어이고, 뉴스 미디어를 출처로 하는 하이퍼링크 트윗의 전파 범위와 활동 기간은 비하이퍼링크 트윗보다 전파 범위가 작고, 활동 기간이 짧다. 이러한 차이는 유력자가 생산한 하이퍼링크, 비하이퍼링크 트윗의 정보 전파 결과와 비교했을 때 더 크게 나타났다. 이러한 결과는 트위터 이용자가 뉴스 미디어에서 생산한 정보가 아닌 자신의 인적 네트워크 내에 있는 이용자가 생산한 정보를 더 선호한다는 것을 나타낸다. 이는 트위터의 대안적 뉴스 미디어의 성격을 간접적으로 보여준다.

마지막으로, 이 논문의 한계를 언급하면서 후속 연구자들을 위한 몇 가지 제언을 제시하면 다음과 같다. 첫째, 이 논문은 메시지 속성에 집중하여 메시지의 질을 간과하였다. 남 외 (Nam, et al., 2011), 김찬균 (2012)의 연구에서 볼 수 있듯이, 트윗의 내용, 하이퍼링크와 트윗 텍스트 간의 수사적 맥락도 트위터의 정보 전파에 큰 영향을 미친다. 이 같은 점을 고려할 때, 향후 연구는 자연어 처리, 감성 분석 등과 같은 내용 분석을 통해 보다 심도 있는 사회과학적 함의를 도출해야 할 것이다. 특히, 하이퍼링크와 웹사이트 간의 상호연결성이 인터넷 내의 사회적 커뮤니케이션 (sociocommunication)의 단면을 밝힐 수 있다 (Park, & Thelwall, 2008). 는 점에서 하이퍼링크와 트윗 텍스트 간의 수사적 기능을 고려한 내용 분석 연구를 향후 연구자들을 위해 제안한다.

둘째, 데이터 수집 방법이다. 이 논문은 한국인 트위터 이용자와 그들이 작성한 트윗을 수집하기 위하여 키워드를 중심으로 데이터를 수집했다. 이러한 방법은 연구의 목적과 기술

적 한계로 인하여 선택한 방법이기도 하지만 다소 왜곡된 연구 대상이 수집될 수밖에 없었다. 수집된 이용자의 팔로워 수의 평균값이 일반적인 트위터 이용자의 평균값보다 높았고, 시사적 내용의 트윗만을 수집했기 때문이다. 결국, 연구 대상 선정의 문제로 인하여, 이 논문의 결과에 대한 일반화에 주의 깊은 해석이 요구되었다. API를 활용한 데이터 수집은 비용과 이용의 편리함 차원에서는 효과적이지만, 데이터의 신뢰도 차원에서는 주의를 기울여야 한다(Park, 2011). 따라서 후속 연구자들은 대표성 있는 이용자와 트윗을 수집하는데 주의를 기울여야 할 것이다.

■ 참고문헌

- 김은미·이준용 (2004). 새로운 공론장으로서의 인터넷 토론 공간에 관한 소고. 『2004년 한국언론학회 전환기의 한국 언론 세미나』, 117~147.
- 김찬균 (2012). 『한국인 트위터 이용자의 하이퍼링크 트윗 분석』. 서울대학교 대학원 석사학위논문.
- 박한우 (2010). e-사이언스 시대의 인문사회학 연구하기: 인터넷 연구 방법을 중심으로. 『사회과학연구』, 30권 2호, 195~211.
- 윤성이·이원태 (2009). 디지털시대 시민의식 제고를 위한 정책 방안. 『한국지역정보학회지』, 12권 1호, 59~84.
- 이원태·차미영·박현유 (2010). 『모바일 소셜미디어에서 유력자(influential)의 역할』. (디지털 컨버전스 기반 미래 연구(II) 시리즈 10~26), 서울: 정보통신정책연구원.
- 이원태·차미영·양해륜 (2011). 소셜미디어 유력자의 네트워크 특성: 한국의 트위터 공동체를 중심으로. 『언론정보연구』, 48권 2호, 44~79.
- 이재현 (2004). 『멀티미디어와 디지털 세계』. 서울: 커뮤니케이션북스.
- _____ (2012a). 프롤로그: 트위터란 무엇인가. 이재현(편) (2012), 『트위터란 무엇인가: 다학제적 접근』(1~23쪽), 서울: 커뮤니케이션북스.
- _____ (2012b). 글쓰기 공간으로서의 SNS: 재매개, 환유, 에크프라시스. 『커뮤니케이션이론』, 8권 1호, 323~351.
- 장덕진·김기훈 (2011). 한국인 트위터 네트워크의 구조와 동학. 『언론정보연구』, 48권 1호, 59~86.
- 최서영 (2010). 『글쓰기로서의 블로그 링크 분석: 하이퍼링크의 수사적 이용을 중심으로』. 서울대학교 대학원 석사학위논문.

- 황유선 · 이재현 (2011). 『트위터에서의 뉴스 생산과 재생산: 8개 언론사와 일반인의 트윗 및 전파행태에 관한 연구』. 서울: 한국언론진흥재단.
- Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. Retrieved from <http://arxiv.org/abs/1003.5699v1>
- Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). *Everyone's an influencer: Quantifying influence on Twitter*. Paper presented at the WSDM'11.
- Burbules, N. C. (2001). The web as a rhetorical place. In Snyder, I.(ed.). *Silicon lieracies*(pp. 75~84). London: Routledge.
- Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010). Measuring user influence in Twitter: The million follower fallacy. Paper presented at the 4th International AAAI Conference on Weblogs and Social Media.
- Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2010). Who is tweeting on Twitter: Human, bot, or cyborg? Paper presented at the ACSAC' 10.
- Eysenbach, G., & Kohler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, 324, 573~577.
- Flanagin, A., & Metzger, M. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515~540.
- _____ (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(12), 319~342.
- Freeman, K., & Spyridakis, J.(2003). An examination of factors that affect the credibility of online health information. *Technical Communication*, 51(2), 239~263.
- Fogg, et al. (2001). What makes web sites credible? A report on a large quantitative study. Paper presented at SIGCHI' 10.
- Garton, L., Haythornthwaite, C., & Wellman, B. (1999). Studying on-line social networks. In S. Jones (Ed.). *Doing internet research: Critical issues and methods for examining the net*(pp. 75~105). Thousand Oaks: Sage. 이재현 역 (2000). 『인터넷 연구 방법: 쟁점과 사례』 (312~362쪽). 서울: 커뮤니케이션북스.
- Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. Paper presented at the 42nd Hawaii International Conference on System Sciences.
- Huberman, B. A., Romero, D. M., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. Retrieved from <http://arxiv.org/pdf/0812.1045>.
- Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why We Twitter: Understanding microblogging usage and communities. Paper presented at the 9th WEBKDD and 1st SNA-KDD Workshop.
- Katz, E. & Lazarsfeld, P. (1955). *Personal influence: The part played by people in the flow of mass communication*. New York: The Fress Press.
- Krishnamurthy, B., Gill, P., & Arlitt, M. (2008). A few chirps about Twitter. Paper presented at the WOSN' 08.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media?. Paper presented at the WWW 2010.
- Lazarsfeld, P., Berelson, B, & Gaudet, H. (1948). *The people's choice*. New York: Columbia University Press.

- Lerman, K., & Ghosh, R. (2010). Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. Paper presented at the 4th International AAAI Conference on Weblogs and Social Media.
- Li, C., & Bernoff, J. (2008). *Groundswell: Winning in a world transformed by social technologies*. Harvard Business Press.
- Morgan, W. (2002). Heterotropes: Learning the rhetoric of hyperlinks. *Education, communication and information*, 2(2/3), 215~233.
- Nam, Y., Son, I., & Lee, D. (2011). The Impact of message characteristics on online viral diffusion in online social media service. 『지능정보연구』, 17(4), 75~94.
- Park, H., & Thelwall, M. (2008). Developing network indicators for ideological landscapes from the political blogosphere in South Korea. *Journal of Computer-Mediated Communication*, 13(4), 856~879.
- Park, H. (2011). How do social scientists use link data from search engines to understand internet-based political and electoral communication. *Quality & Quantity*, 46(2), 679~693.
- Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2010). Influence and passivity in social media. Retrieved from <http://ssrn.com/abstract=1653135>
- Suh, B., Hong, L., Pirolli, P., & Chi, E. D. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. Paper presented at the Second IEEE International Conference on Social Computing.
- Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2), 406~418.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. Paper presented at the 4th International AAAI Conference on Weblogs and Social Media.
- Yang, J., & Counts, S. (2010). Predicting the speed, scale, and range of information diffusion in Twitter. Paper presented at the 4th International AAAI Conference on Weblogs and Social Media.

최초 투고일 2012년 2월 10일

게재 확정일 2012년 5월 18일

논문 수정일 2012년 5월 28일

<부록 1> 분석 대상 추출을 위한 키워드와 해당 트윗 수

메인 키워드	서브 키워드	트윗 수	메인 키워드	서브 키워드	트윗 수
무상급식	주민투표	254	물폭탄	소강상태	2
총기사고	해병대	16	폭우	서울	21
해병대	기자	79	남부 폭우	-	3
한나라당 사무총장	김정권	6	뽀로로	디즈니	224
복지포퓰리즘	오세훈	6	백곰 박테리아	-	3
홍준표	막말	119	서울택시	시외요금	6
한선교	귀빈실	9	분당	16분	24
평창올림픽	공동개최	29	갤럭시S2	-	85
법무장관	권재진	168	패스트패션	명품	1
김상병	해병대	12	PT	평창	23
문재인	특전사	47	고속철	베이징	7
주민투표	박근혜	5	귀신 버섯	야광	1
해병대사령관	유낙준	9	최홍만	노조미	4
류경호텔	쓰레기	1	롤러코스터	최고 경사	6
저축은행	로비	39	세습	쇠고기	19
최병렬	노무현	6	아틀란티스	북대서양	3
J-10	중국	4	브루니	임신	1
한상대	검찰총장	146	화산	칠레	1
평창수	코카콜라	7	시간 여행자	1911년	1
이건희	품질	2	달라이 라마	오바마	5
휘발유값	3000원	6	카카오톡	상반기	1
역세권	용산	1	4G폰	LTE	1
펀드	월지급	1	스마트폰	5인치	4
SC제일은행	파업	89	이동통신사업	삼성	2
도시광산	50조	1	iOS 5	베타3	8
경인운하	2조	7	윈도8	사양	1
면세	한도	1	규혁롬	악플	8
목사	입양말	25	위치추적	집단소송	34
닌자	경복궁	25	네시	괴물	1
테크노마트	원인	20	애플	대만	15
오징어	대왕	9	네이버 야구	통신사	13
태풍	도카게	25	갤탭 10.1	아이패드2	2

〈부록 1〉 계속

메인 키워드	서브 키워드	트윗 수	메인 키워드	서브 키워드	트윗 수
터널	택시화재	16	트레이드	LG	22
양지마을	밀양	2	윤석민	11	11
학파라치	주부	2	한기주	756일	3
막말	맞선	2	손흥민	프리킥 골	6
나서영	여대생	2	코리	퇴출	3
호우주의보	서울	27	덕 매티스	삼성	4
유민	가문의 수난	1	박주영	세비야	5
손석희	손목시계	18	조용형	면제	1
용준형	저작권료	2	차두리	22인	3
최지우	지고는 못살아	3	신기생면	귀신	7
영거	임성민	1	박유천	리플리	38
정만호	이훈	4	김주현	100평	3
정재형	폭언녀	2	박진영	섬데이	3
권리세	키이스트	1	강혜정	리플리	3
이혜영	하와이	2			

〈부록 2〉 트위터 정보 전파의 특성

하이퍼링크	하이퍼링크 무	하이퍼링크 유	
		정보적	논리적
수사적 기능	-		
리트윗 받은 수			
범위	+++	++	+
거리			
속도	++	+++	+
활동 기간	+++	+	++

+++ > ++ > +
출처: 김찬균 (2012).

Analysis of the Information Diffusion Process on Twitter: Effects of Influentials and Hyperlinks

Jae Hyun Lee

Professor, Dept. of Communication, Seoul National University

Chan Kyun Kim

M.A, Dept. of Communication, Seoul National University

Reflecting its great social and cultural influence, a number of researches on the information diffusion on Twitter have been conducted. Previous studies, however, used the number of retweets only as an index of the level of diffusion, causing insufficient results in the aspect of the multidimensional process of information diffusion. The present paper aims to fill up a lacuna in this point, applying a variety of indices to the research and analyzing the very process of propagation. The properties of tweets, which could influence on the information process, have also been overlooked, so this paper tries to differentiate the process according to the property of Twitter user-whether or not (s)he is an influential, and according to the property of tweet messages-the usage of hyperlink. The study shows that the tweets posted by influentials are spread more widely and deeply, lasting longer. Also, the tweets with hyperlink are propagated more widely and longer compared to those without hyperlink. The tweets posted by influentials and containing hyperlink turn out to be the most widely propagated ones. Finally, the linear regression analysis tells us that the user factor has a stronger influence than the message factor when it comes to the spread of tweets.

Key words: Twitter, retweet, RT, information diffusion, influential, hyperlink