

소셜네트워크 분석을 통한 마케팅 전략 : 트위터의 검색네트워크

Marketing Strategies using Social Network Analysis : Twitter's Search Network

유병국, 김순홍
인천대학교 무역학부

Byong-Kook Yoo(bkyoo@incheon.ac.kr), Soon-Hong Kim(snow8817@hanmail.net)

요약

입소문효과 극대화를 위한 유력자의 역할은 트위터 네트워크에서도 매우 중요하다고 볼 수 있다. 본 논문에서는 기업 마케팅의 관점에서 트위터 유력자를 파악하고자 하였다. 기업의 마케팅 메시지가 가능한 많은 사람들에게 노출되기 위해서는 특정인의 팔로어 수뿐만 아니라 지속적인 리트윗을 통해 입소문이 많이 확산되어야 할 것이다. 즉, 팔로어 수는 많은데 리트윗되지 않거나 리트윗은 많이 되는데 소수자에게 주로 리트윗된다면 전체적인 노출의 정도는 미약하게 될 것이다. 트위터의 특정 검색네트워크 데이터를 이용한 그래프 비교를 통해 다음과 같은 사실을 확인할 수 있었다. 첫째, 리트윗이 없는 팔로어 유력자에 비하여 팔로어 유력자의 리트윗을 받는 상대적인 소수자들의 노출도가 크다는 것을 알 수 있었다. 둘째, 리트윗 유력자중에서도 팔로어 유력자에 의하여 리트윗을 받지 못하는 사용자의 노출도는 매우 미미하다는 것을 알 수 있었다. 위와 같은 사실은 노출도를 높이기 위해서는 실제 유력자도 중요하지만 그런 유력자의 리트윗을 유도하는 사용자의 파악이 더 중요하다는 점을 시사해준다고 볼 수 있다. 이와 함께 노출도의 경우 대부분의 중앙성 척도와 고른 상관관계를 유지하고 있어 노출도가 높은 이용자일수록 네트워크구조상에서도 중심적인 위상을 차지한다는 점도 살펴볼 수 있었다.

■ 중심어 : | 트위터 유력자 | 팔로어 수 | 리트윗 수 | 노출도 | 네트워크 중앙성 |

Abstract

The role of influentials to maximize word-of-mouth effect can be seen to be very important. In this paper, we have the perspective of corporate marketing to understand Twitter influentials. We start from the point of view of who can induce eventually most exposure of tweets when he tweets the company's specific marketing messages. From this perspective, we observe both the follower influentials who have many followers and the retweet influentials who induce many retweets by visualizing graphs from network data collected via Twitter Search API. Although some users have small followers they may bring much more exposure than follower influentials if they can induce retweets by follower influentials. On the contrary, some retweet influentials who don't induce retweets by follower influentials may bring very little exposure. This suggests the fact that some small users who can induce retweets by influentials might have more important role than influentials themselves in order to increase the exposure of tweets. These users also are seen to have high centrality measures in the network structure.

■ keyword : | Twitter Influentials | Followers | Retweets | Exposure | Network Centrality |

* 이 논문은 인천대학교 2011년도 자체연구비 지원에 의하여 연구되었음.

접수번호 : #130219-009

접수일자 : 2013년 02월 19일

심사완료일 : 2013년 03월 22일

교신저자 : 김순홍, e-mail : snow8817@hanmail.net

1. 서론

입소문(Word-of-mouth)은 정보가 다수의 사람들에게 전달되어 여론을 형성하거나 새로운 기술이나 상품이 채택되게 하는 중요한 수단으로 간주되어 왔다. 그러나 이러한 입소문에 대한 경험적 연구는 매우 바람직스럽기는 하지만 두 가지 점에서 어려움을 가지고 있다 [1]. 첫째, 입소문이 전파되는 네트워크는 일반적으로 관찰되지 않기 때문에 영향력의 정도를 측정하는 것이 어렵다. 둘째, 그나마 입소문으로 관찰되는 데이터는 성공적인 경우에 국한되는 경우가 많아 편향적(biased)일 수 있다. 이러한 점에서 최근 사용층이 확대되고 있는 소셜 네트워크 서비스(Social Network Service; SNS)로서 140자 단문 서비스인 트위터(Twitter)는 입소문연구를 위한 대안적인 대상이 될 수 있다[2][3]. 트위터는 사용자가 다른 사용자가 보내는 정보의 수신에 가입 혹은 팔로우(follow)함으로써 정보가 유포되는 본질적인 특성을 가지고 있다. 즉, 팔로어(follower)관계를 통해 정보유포자와 정보수신자간의 네트워크가 구축된다. 이에 따라 입소문의 성공여부에 관계없이 유포되는 데이터의 관찰에 따른 표본추출이 가능해진다.

최근에 이러한 입소문의 효과를 최대화할 수 있는 특정한 개인으로서 유력자(influentials)에 관한 연구들이 많은 관심을 받고 있다. 소셜 네트워크(social network)에서 유력자의 정의를 한마디로 정의하기 어렵지만 일반적으로 유력자는 “의견지도자(opinion leader)나 트렌드 형성자(trend setter)로서 중요한 역할을 하고 새로운 정보기술의 채택 속도에도 직접적인 영향을 미치는 특별한 개인들”로 평가받고 있다[4].

입소문 마케팅에서 가장 중요한 것 중 하나가 바로 누가 유력자인가를 찾아내는 것이다. 영향력 있는 대상에게 마케팅을 집중하는 것이 효과도 좋고 비용도 절감할 수 있기 때문이다. 그러나 유력자의 의미는 관찰하는 네트워크의 종류와 보고자 하는 영향의 종류에 따라 달라질 수 있기 때문에 누가 유력자인지를 알아내고자 하는 연구는 아직 진행중인 논의라고 볼 수 있다. 2단계 유통이론(two-step flow theory)에 따르면 정보는 매스 미디어로부터 의견선도자들에게로 전해지며, 이들에

의해 대다수의 국민 층으로 전달된다고 한다[5]. 여기서 는 사람들의 태도를 변화시키는 데에 매스미디어의 영향보다 접촉한 의견지도자의 영향이 더 큰 것으로 관찰되고 있다. 혁신 혹은 신제품의 시장 확산에 있어서 초기 채택자인 혁신자(innovators) 혹은 초기수용자(early adopters)의 역할이 대단히 중요하다[6]. 이들은 다른 집단에 비해서 합리적이며 변화에 능동적이며 불확실성에 대한 대응능력이 뛰어나고 대중매체와의 접촉이 잦아 초기 다수 수용자(early majority), 후기 다수 사용자(late majority), 지각 수용자(laggards) 등의 선택에 지대한 영향을 미치게 된다.

그러면 트위터 네트워크에서 유력자는 어떻게 규정 지을 수 있을까? 흔히 트위터 네트워크에서 유력자는 팔로어 수로 측정되는 경향이 있다. 우리나라에서도 팔로어 수가 많은 유명 인사들은 여러 분야에서 상당한 영향력을 가지고 있다고 볼 수 있다. 기업의 입장에서도 해당 트위터의 팔로어는 기업의 메시지를 직접적으로 전달할 수 있는 수요층이므로 팔로어 수의 증가에 많은 노력을 기울이는 것이 사실이다. 그러나 트위터 사용자들은 개개인이 불균형한 영향력을 가지고 있기 때문에 단순히 팔로어 수가 많다고 해서 반드시 유력자라고 할 수는 없다[7]. 또 팔로어 수보다는 리트윗(retweet)등과 같은 이용자들 간의 소통적 상호작용 수준으로 측정될 필요도 있다[8]. 리트윗을 통해 많은 사람들에게 자신의 메시지가 전달되는 트위터의 특성을 고려할 때, 자주 리트윗되는 글을 작성하는 사람 역시 많은 사람들에게 영향력을 미칠 수 있을 것이다.

그렇다면 기업의 입장에서 트위터 유력자는 어떻게 규정 지을 수 있을 것인가? 기업의 마케팅 메시지가 가능한 한 많은 사람들에게 노출되기 위해서는 특정인의 팔로어 수뿐만 아니라 지속적인 리트윗을 통해 입소문이 많이 확산되어야 할 것이다. 즉, 팔로어 수는 많은데 리트윗되지 않거나 리트윗은 많이 되는데 소수자에게 주로 리트윗된다면 전체적인 노출의 정도는 미약하게 될 것이다. 지금까지 트위터의 영향력과 노출의 관계를 직접적으로 살펴본 연구는 거의 없었다고 볼 수 있다. 본 논문에서는 팔로어 수나 리트윗이 아닌 트윗의 노출 관점에서 유력자의 개념을 새롭게 조명해 보고자 한다.

즉, 유력자가 가지는 입소문의 영향력을 노출의 관점에서 평가해보고자 한다. 이를 위해 우선 트위터 검색네트워크를 통해 특정 주제어와 관련된 네트워크데이터를 수집한다. 그래프분석을 통해 팔로어 수, 리트윗에 따라 유력자를 구분하고 이를 노출의 관점에서 비교해 본다. 마지막으로 이러한 관점에서 유력자들이 네트워크에서 차지하는 구조적 위상을 검토해 본다.

II. 자료 수집

트위터에서는 특정인의 트위터 네트워크뿐만 아니라 특정주제의 네트워크데이터를 추출해 분석할 수 있다. 이 경우 팔로어 수 등 이용자개인의 정보뿐만 아니라 특정주제에 대한 리트윗 수의 수집이 가능해 입소문에 관여하는 유력자의 역할을 조명할 수 있다. 기존의 검색엔진마케팅(search engine marketing)이 키워드를 입력한 소비자가 공급자를 찾아가는 방식이라면 트위터의 검색네트워크는 공급자가 소비자를 찾아간다는 점에서 차이가 있다.

일반적으로 트위터 자료를 추출하기 위해서는 트위터에서 공식적으로 제공하는 공개 API(Application Programming Interface)를 사용해야 한다. API는 트위터 네트워크의 외부에서 내부의 데이터에 쉽게 접속할 수 있도록 하는 프로그램 함수를 말한다. 트위터 API를 호출했을 경우에 나오는 결과에는 다음 두 가지 성격의 중요한 정보가 포함된다. 첫째는 다양한 사용자 정보이다. 여기에는 사용자의 이름, 아이디(user ID)뿐만 아니라 자신이 위치한 도시정보, 소개글, 트윗 수, 홈페이지 주소 등 수십 가지의 정보가 포함된다. 두 번째는 팔로어관계 등 사용자들 간의 관계 정보이다. 사용자들은 서로의 트윗을 구독(following)하는데, 사용자들의 구독 여부 관계를 '소셜 링크(social link)'라고 부르며 이 정보 또한 API를 통해 접근이 가능하다.

본고에서는 API를 통한 트위터 검색네트워크의 접근을 위하여 NodeXL을 이용하고자 한다. NodeXL은 마

이크로소프트사의 엑셀 템플릿 (Excel template)중의 하나로 네트워크를 분석할 수 있는 툴이다². NodeXL은 엑셀의 인터페이스를 가지며 트위터, 유튜브 등 SNS 데이터를 수집할 수 있고 다양한 네트워크 척도(metrics)들을 쉽게 계산할 수 있다는 장점을 가지고 있다[9]. NodeXL이 제공하는 검색메뉴를 이용하면 뉴스, 광고, 제품 등 트윗에 나타날 수 있는 모든 구절에 대한 검색이 가능하며 그러한 구절들이 어떻게 전파되고 있는지도 관찰할 수 있다.

이러한 트위터의 검색네트워크를 조사할 때는 다음과 같은 점에 유의할 필요가 있다. 첫째, 트위터는 빠르게 증가하는 메시지들의 리스트를 핫이슈(trending topics)로 분류하여 보관하고 있다. 따라서 검색에 사용되는 주제어는 최근의 사건과 관련되어야 하며 다른 사람들에게 전파될 수 있도록 충분한 전염성을 가져야 한다. 둘째, 트위터 자체의 제한정책으로 인해 특정주제어가 전파되는 망의 전체적인 구조를 파악하는 데에 어려움이 있다. 즉, 트위터에서는 외부 정보요구에 따른 망 부하를 줄이기 위해서 자체적으로 시간당 150~200노드(사용자)로 제한(rate limit)하고 있다. 따라서 특정시점에 해당되는 트윗의 양이 방대할 경우 부분적인 수집으로 제한될 가능성이 있다.

검색 주제어는 사용목적에 따라 다양하고 구체적으로 정해질 수 있다. 본고에서는 트위터를 통한 입소문 확산의 일반적인 사례를 조사하기 위하여 '모바일앱(mobile app)'을 주제로 정하였다. 모바일 앱 혹은 모바일 애플리케이션(application)은 모바일기기 운영체제(mobile operating system)에 적합하게 디자인되어 모바일기기 자체의 기능을 확장 및 향상시키는 소프트웨어를 지칭한다[10]. Apple에 의해 2007년 처음 소개되었고, 이후 스마트폰과 태블릿PC의 등장에 따라 급속히 확산되어 데스크톱 컴퓨팅 시대에서 모바일 컴퓨팅 시대로 전환하는 계기를 제공하고 있다. 2011년 전 세계 모바일 앱 스토어 매출이 151억 달러 이상을 기록하였는데 이는 2010년에 비해 190% 이상 증가된 것으로 오는 2015년에는 1,440억 달러에 이르는 가파른 성

1 대부분의 소비자들은 홍보형 트윗을 무시하는 경향이 있는데 홍보형 메시지를 리트윗하는 소비자가 있다면 이러한 네트워크 분석을 통해 적절한 보상을 해주는 것도 가능해진다.

2 NodeXL은 다음 사이트(<http://nodexl.codeplex.com/>)에서 무료로 다운 받을 수 있다.

장세가 전망되고 있다[11]. 제품군도 게임, 음악, SNS, 교육, 도서, 엔터테인먼트 등으로 다양하게 확대되고 있다. 따라서 기존 모바일앱의 제공자나 신규참여자의 입장에서 전방적인 모바일앱시장에 대해 조망해 볼 필요가 있을 것이다.

[그림 1]은 NodeXL의 'Import from Twitter Search Network'의 메뉴를 이용하여 '모바일앱'이 들어간 트윗들을 수집하기 위해 작성된 메뉴의 예를 보여준다.

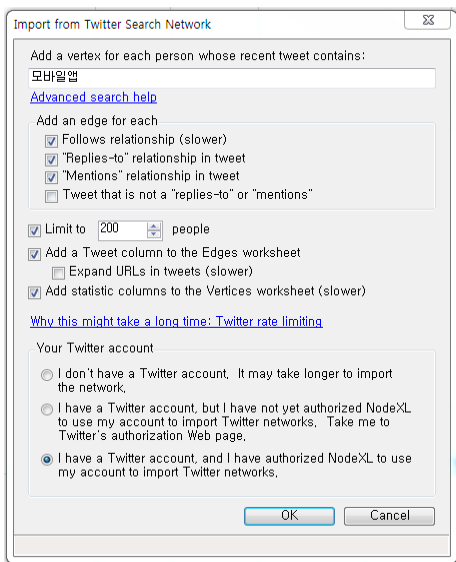


그림 1. 검색메뉴의 작성예

즉, 메뉴작성에는 다음과 같은 내용들이 포함된다. 첫째, 검색어로 '모바일앱'을 명기하면 현재 핫이슈로 분류되어 보관중인 트윗들 중에서 메시지내용에 '모바일앱'이 포함된 트윗의 작성자가 노드(node)로서 추출된다. 둘째, 이러한 트윗 작성자들간 관계가 포함된다. 즉, 특정주제(모바일앱)를 언급한 작성자들을 대상으로 팔로어관계, 멘션(mention)관계, 리플라이(reply)관계를 조사하여 관계가 있으면 링크(link 혹은 edge)로 연결하게 된다. 여기서 멘션은 사용자A가 사용자B의 이름을 자기의 트윗에 언급한 경우를 말하며 리플라이는 사용자A가 자신의 트윗 첫머리에 사용자B의 이름을 언급하므로 사용자B의 트윗에 대답하는 경우를 말한다. 셋째, 트위터의 rate limit 정책에 따라 추출되는 트

윗 수를 제한해야 한다.

2012년 11월 15일 '모바일앱'으로 검색한 결과 다음의 [표 1]과 같이 115노드와 153개의 링크로 이루어진 네트워크 데이터가 수집되었다. 여기서 노드는 특정 트위터사용자를 나타내는데 [표 1]에서 보는 바와 같이 115개의 노드중 다른 노드와 연결이 없는 단일 노드(single-node)는 33개이며 나머지 82개의 노드는 서로 연결되어 하나의 커다란 구성체(component)를 이루고 있었다. 또 이러한 노드를 이어주는 링크는 각 사용자간 팔로어, 멘션의 관계를 나타내는데 이 네트워크에서의 모든 멘션은 리트윗을 나타내는 멘션이었다.

표 1. 네트워크 데이터 요약

그래프 개요	값
그래프 유형	방향성있음 (Directed)
노드의 수	115
링크의 수	153
전체 구성체의 수	34
단일노드 구성체의 수	33
구성체의 최대 노드의 수	82
구성체의 최대 링크의 수	205
가장 큰 최단거리 (Geodesic Distance)	8
평균 최단거리	3.84652
그래프 밀집도 (Graph Density)	0.013653699

III. 트위터 검색네트워크 분석

위에서 수집된 네트워크데이터를 그래프로 표시하면 다음의 [그림 2]와 같다. 단일노드(H1wony 등)들은 왼쪽 하단에 따로 모아서 표시하였다. 여기서 링크의 화살표방향은 예를 들어 $A \rightarrow B$ 이면 사용자A가 사용자B를 팔로어(혹은 리트윗)한다는 것을 의미한다. 이 경우 정보의 흐름은 화살표 반대방향으로 흐른다고 할 수 있다. 즉, 사용자B가 트윗하면 사용자A에게 전달되며 사용자A는 경우에 따라서 사용자B의 트윗을 리트윗할 수 있다. 연결된 구성체에서의 노드의 크기는 해당 사용자의 팔로어 수에 비례하여 다르게 표시하였다. 즉, 검색된 네트워크에서는 SKTworld와 SocialCaster의

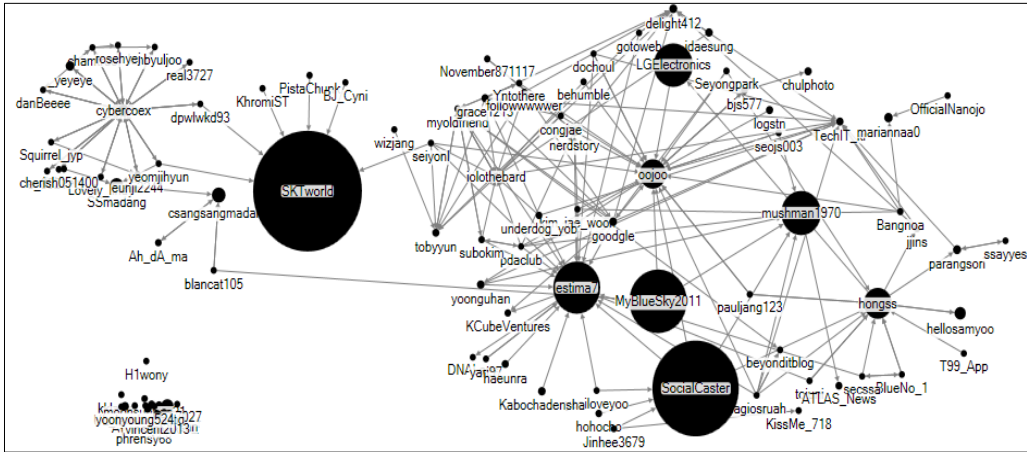


그림 2. 모바일앱 검색네트워크의 팔로어 유력자

팔로어 수가 각각 154,525명, 120,756명으로 가장 많은 것으로 나타났다. 그 다음에 MyBlueSky2011, estima7 등이 있음을 알 수 있다. 위의 특정 검색 네트워크에서 SKTworld와 SocialCaster는 팔로어 수의 관점에서 보면 유력자라고 할 수 있다. 우리는 이들을 팔로어 유력자라고 하자. 즉, 이들이 작성하거나 리트윗한 트윗은 각각의 많은 팔로어들에게 전달되게 된다.

그러나 앞서 거론한 데로 유력자는 팔로어 수보다는 리트윗 등과 같은 이용자들 간의 소통적 상호작용 수준으로 측정될 필요가 있다. 그렇다면 이 네트워크에서 가장 많은 리트윗을 받고 있는 사용자는 누구일까? [그림 3]의 그래프는 노드간 연결되는 링크 중에서 리트윗에 해당하는 링크를 굵게 표시하고 리트윗 수에 따라

노드의 크기를 달리하여 표시한 것이다. 아래의 네트워크에서 cybercoex, iolothebard, oojojo 등은 각각 13개, 10개, 9개의 리트윗을 받고 있어 다른 사용자에 비해 리트윗의 관점에서는 유력자라고 할 수 있다. 이들을 리트윗 유력자라고 하자. 위의 네트워크 관찰을 통해 다음과 같은 사실을 확인할 수 있다. 첫째, 유력자측면에서 리트윗 수는 팔로어 수와 정의 관계를 보이고 있지 않았다. 즉, 팔로어 수가 많은 SKTworld와 SocialCaster의 경우 리트윗을 거의 받지 못한 반면에 상대적으로 팔로어 수가 적은 cybercoex, iolothebard, oojojo 등의 사용자들은 리트윗을 많이 받고 있었다. 둘째, SocialCaster를 제외하면 대부분의 사용자가 1개를 초과하여 리트윗하지 않았다. 반면에 팔로어 유력자인

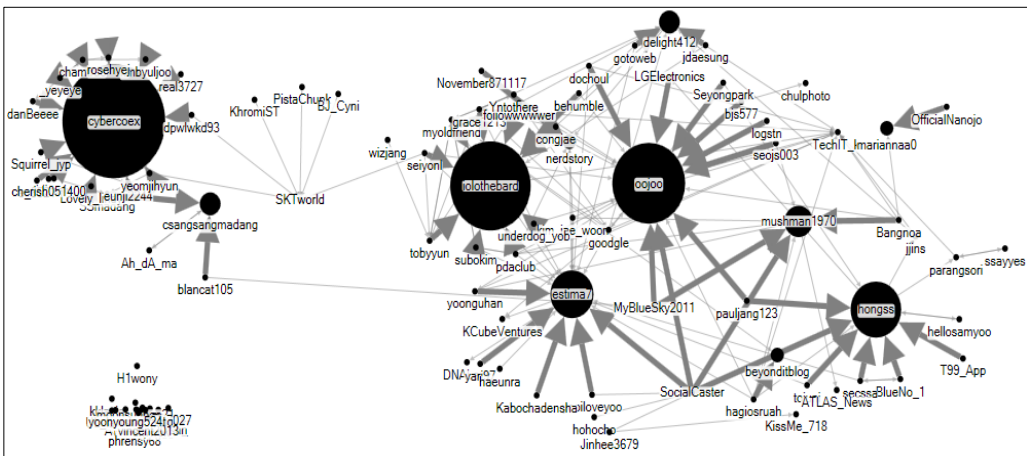


그림 3. 모바일앱 검색네트워크의 리트윗 유력자

SocialCaster의 경우 다른 사용자들의 트윗을 활발히 리트윗하는 특징을 보이고 있었다.

[그림 2]와 [그림 3]을 비교해보면 몇 가지 흥미있는 점을 관찰할 수 있다. 첫째, 팔로어 유력자중 SKTworld는 트윗의 내용을 스스로 생산하여 주로 리트윗을 받는데 반해 SocialCaster는 자신의 트윗보다는 타인의 트윗을 리트윗하는데 주력하고 있다. 둘째, 리트윗 유력자에서는 누구에게 리트윗되느냐로 구분할 수 있다. cybercoex, ioiothebard 등의 트윗은 팔로어 수가 비교적 적은 사용자가 리트윗하는데 반하여 oojoo, hongss, estima7 등의 트윗은 비교적 팔로어 수가 많은 사용자에게 의해서 리트윗됨을 알 수 있다. 실제로 이들의 트윗은 팔로어 유력자중 하나인 SocialCaster에게 모두 리트윗되고 있음을 알 수 있다.

이러한 두 가지 사항들은 트윗의 노출에 많은 영향을 미치게 된다. 그러면 어떤 사용자의 트윗이 가장 노출이 많이 이루어진 것인가? 일반적으로 트윗의 노출도는 두 가지 측면에서 살펴볼 수 있다. 즉, 노출도는 트윗 작성자의 팔로어 수와 그 트윗이 리트윗되는 경우 리트윗하는 사용자들의 팔로어 수의 합으로 이루어지게 된다.

$$\text{노출도} = \text{작성자의 팔로어 수} + \text{리트윗 유저들의 팔로어 수}$$

이러한 트윗의 노출도는 위에서 언급한 팔로어 유력

자의 개념과 리트윗 유력자의 개념을 모두 고려한 것이라 볼 수 있다. 즉, 비록 팔로어 유력자라 하더라도 리트윗을 적게 받거나 리트윗 유력자일지라도 팔로어 수가 많은 사용자의 리트윗을 받지 못하면 노출도는 미미할 수 있다. 반면에 팔로어 수가 적은 작성자라도 팔로어 유력자 등 많은 팔로어를 가진 응답자들에 의해 리트윗 받을 경우 노출도는 비약적으로 향상될 수 있는 것이다.

표 2. 유형별 사용자 순위

순위	팔로어 수	리트윗 수	노출도
1	SKTworld (154,525)	cybercoex (13)	oojoo (282,143)
2	SocialCaster (120,756)	ioiothebard (10)	mushman1970 (249,170)
3	MyBlueSky2011 (77,219)	oojoo (9)	estima7 (186,959)
4	estima7 (61,507)	hongss (6)	SKTworld (154,525)
5	mushman1970 (50,646)	estima7 (5)	hongss (154,165)
6	LGElectronics (49,807)	mushman1970 (3)	SocialCaster (120,756)
7	hongss (32,692)	csangsangmadang (2)	MyBlueSky2011 (77,219)
8	oojoo (30,579)	delight412 (2)	LGElectronics (49,807)
9	csangsangmadang (11,303)	nerdstory (1)	csangsangmadang (21,748)
10	SSmadang (10,269)	mariannaa0 (1)	cybercoex (11,303)

(*) 괄호안의 수는 각각의 값을 나타낸다.

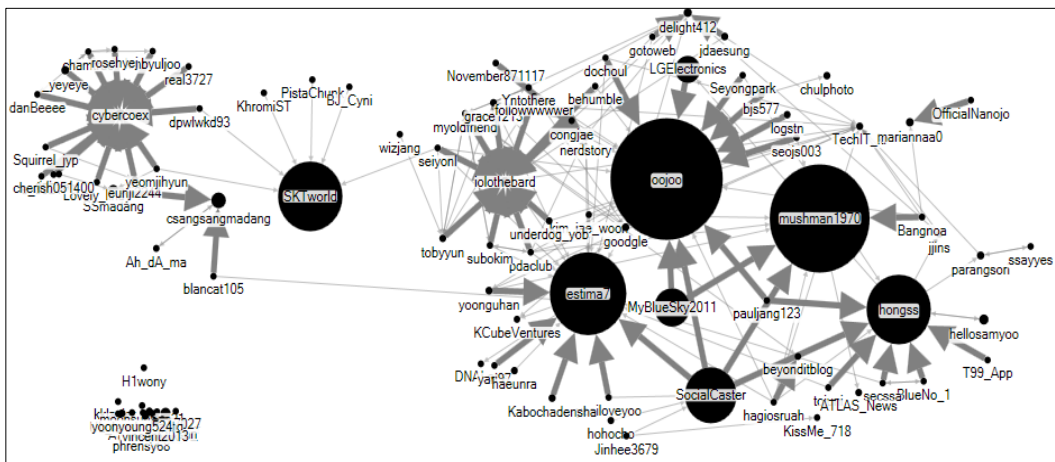


그림 4. 모바일앱 검색네트워크의 노출도 크기

[그림 4]의 그래프는 노출수에 따라 노드의 크기를 변화시킨 것이다. 리트윗관계는 굵은 실선으로 나타내었다. 그래프에 따르면 팔로어 수 혹은 리트윗 수가 곧 높은 노출도를 보장하지 않는다는 사실을 알 수 있다. [표 2]는 조사된 네트워크내의 전체 사용자중에서 팔로어 수, 리트윗 수, 노출도 측면에서 상위 10위 이내의 사용자를 나타낸 것이다.

이상에서 살펴본 바와 같이 노출도의 증가를 위해서는 자신의 팔로어 수도 중요하지만 유력자의 리트윗을 많이 받는 것이 매우 중요하다는 것을 알 수 있다. 따라서 기업의 입장에서는 인위적인 팔로어 수 증가보다는 유력자의 리트윗을 유도할 수 있는 열성적인 트위터 사용자들을 활용하는 것이 마케팅 전략적인 측면에서 바람직하다고 볼 수 있다. 위의 네트워크 예에서는 트윗당 1회 이내의 리트윗이 관찰되었다. 그러나 트위터 네트워크의 속성상 동일한 트윗이 네트워크를 따라 여러 번 리트윗될 수 있는 가능성은 충분하다고 볼 수 있다. 이 경우 자신과 직접적인 팔로어 관계에 있지 않은 유력자의 리트윗도 가능하게 되어 노출도는 비약적으로 상승할 수 있다. 이러한 가능성을 다음의 네트워크의 사례들에서 찾아볼 수 있다.

1. 애니팡의 검색사례

[그림 5]의 그래프는 2012년 10월26일에 인기 모바일 앱 중하나인 ‘애니팡’을 검색어로 하여 나타난 트위터 검색네트워크이다. 검색된 노드와 링크는 각각 200개와 5개이며 그룹 밀도가 0.001로 매우 낮아 다른 어떠한 노드와도 연결되지 않은 노드는 173개였다. 즉, 이때 애니팡의 검색네트워크에서는 대부분의 노드들이 다른 노드와 팔로어, 멘션 등의 관계에 있지 않았다. 이 그래프는 이 중에서 가장 많은 노드가 연결된 구성체를 나타낸 것이다. 하단의 아이디 AllInV4를 가진 사용자가 트윗한 내용(“으익ㅋㅋㅋㅋ 애니팡 스티커 나옴ㅋㅋㅋㅋ http://t.co/xe2ujh6G”)이 리트윗되는 경로를 굵은 실선으로 표시하였다. 여기서 노드의 크기는 해당 사용자의 팔로어 수를 나타내는 것으로 아이디 AllInV4의 팔로어는 전체 네트워크에서 비교적 적은 187명이다. 그러나 그의 트윗은 팔로어 관계에 있는 RAONJENA,

ppasic, 2021950, DS_toqui에 의해 리트윗되고 있다. 이들의 팔로어 수는 모두 AllInV4의 팔로어 수보다 많을 뿐 만 아니라 AllInV4와 직접적으로 연결관계에 없는 다른 노드들로 AllInV4의 트윗을 리트윗하고 있다. 특히 RAONJENA의 경우, 위 네트워크에서 가장 많은 팔로어 수를 가지고 있는 Oooong_Gom(3,586명)에 AllInV4의 트윗을 연결해주고 있다. 그 결과 AllInV4의 트윗이 노출되는 노출도는 자신의 팔로어 수(187명)에 리트윗되는 팔로어 수를 모두 합하여 6,876로 대폭 늘어나게 된다. 이 같은 수치는 가장 많은 Oooong_Gom의 팔로어 수보다도 2배에 가까운 크기이다. 위 네트워크에서 RAONJENA의 역할은 앞서 논의했던 ‘모바일 앱’ 검색네트워크에서의 oojo0 등의 역할과 유사하다고 볼 수 있다. 즉, 팔로어 수에 비하여 리트윗받는 비중이 크며 특히 팔로어 유력자에 의해 리트윗되어 노출도를 크게 향상시키는 특징을 가지고 있다.

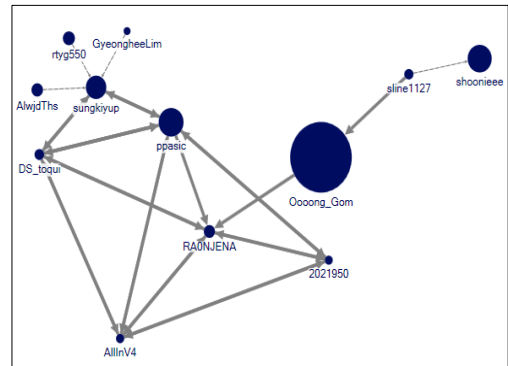


그림 5. 애니팡 검색네트워크

2. 싸이의 강남스타일 사례

가수 싸이의 ‘강남스타일’은 최근 유튜브 사상 첫 10억뷰를 돌파하는 기록을 세우며 저스틴 비버의 ‘베이비’ 뮤직비디오를 제치고 유튜브 사상 가장 많이 본 동영상 1위에 올랐다. 이러한 기록적인 조회수를 달성한 이면에는 트위터의 영향력이 지대했으며 특히 유튜브 관련 사이트를 링크하는 트윗이 전 세계적으로 광범위하게 전파되었다는 데에 연유한다. 즉, 강남스타일의 유튜브 뮤직비디오 조회수는 관련 트윗 수와 거의 같은 곡선을 그리며 증가하였다[12]. 강남스타일이 트위터를 통해

처음 언급된 시점은 2012년 7월 11일이며 강남스타일이 발표된 2012년 7월15일 0시에 소속사인 YG에서는 공식 트위터를 통해 인기가수인 G-Dragon 등 다수의 유명가수들이 강남스타일이 수록된 6집앨범을 피쳐링했다고 발표하였다. 이러한 홍보가 시작되어 입소문을 타면서 하루 3,000건 정도이던 강남스타일에 대한 언급건수는 7월 15일 들어 6,922건까지 상승하였다[13]. 이러한 상승세를 주도한 사람은 같은 소속사인 가수 산다라박의 팬인 WeLoveDaRa(팔로어 수 약 3만 명)라는 해외 트위터 사용자였다. 이후 추츨하였으나 8월 1일 저스틴 비버(Justin Bieber)를 발굴했다는 스쿠터 브라운(Scooter Braun)(팔로어 수 2,009,443)이 강남스타일에 대해 “어떻게 내가 이 친구와 계약을 안했던 것이지?”라고 언급하며 유튜브사이트를 트윗에 링크하자 또다시 상승하는 국면을 맞기 시작했다. 이후 티페인, 케이티 페리 등 서구권 유명 인사들의 트윗으로 강남스타일은 트윗 수가 처음으로 8만 건을 돌파하였으며 미국 주류 언론과 라이언 시크리스트(‘아메리칸 아이돌’ 심사위원) 등 유력인사들도 ‘강남스타일’을 일상용어로 사용하기 시작했다. 이상에서 보는 바와 같이 강남스타일의 돌풍에는 스쿠터 브라운과 같은 해외 유명 인사들의 트윗이 결정적으로 작용했다는 것을 알 수 있다. 하지만 이러한 해외 유력자의 트윗은 처음부터 가능했던 것은 아니고 산다라박의 팬처럼 효과적인 홍보에 따른 주변인들의 활발한 리트윗(혹은 트윗)으로부터 촉발되었다고 추론해 볼 수 있다. 특히 이 과정에서 아시아권 K팝 팬들이 역할이 지대하였다고 지적되고 있다. 예를 들어 ‘SGTeacherSays’라는 싱가포르의 한 트위터 사용자(팔로어 수6,000명)가 언급한 강남스타일에 대한 트윗이 74만 명의 팔로어가 있는 유력 한류 매체인 ‘올케이팝(allkpop)’에 소개되면서 강남스타일의 전 세계적 확산에 커다란 도화선이 되기도 하였다. 이러한 강남스타일의 노출도 상승에 있어 SGTeacherSays 혹은 WeLoveDaRa의 역할은 ‘애니팡’에서의 RA0NJENA의 역할 혹은 ‘모바일앱’ 검색네트워크에서의 oojoo 등의 역할과 매우 유사하다고 볼 수 있다.

위의 두 가지 네트워크 사례에서 다음과 같은 점을 확인할 수 있다. 첫째, 사용자가 보내는 트윗의 노출도

는 유력자의 트윗(혹은 리트윗)에 크게 의존한다는 점이다. 실제로 한국인 트위터 유력자의 경우 단순히 팔로어 수가 많은 사람이라기보다는, 리트윗과 같이 정보의 유통 및 확산이라는 미디어적 기능과 역할을 적극적으로 수행하는 사람이라는 특징을 가지고 있다고 한다[14]. 따라서 노출도를 높이기 위해서는 실제 유력자도 중요하지만 그런 유력자의 리트윗을 유도하거나 많이 받는 사용자의 파악이 더 중요하다고 볼 수 있다. 둘째, 유력자의 리트윗을 유도하기 위해서 트윗의 작성자가 유력자와 반드시 이웃관계(팔로어 등)에 있을 필요는 없다는 점이다. 팔로어가 많은 유력자는 여기에 상응해서 팔로잉(following)도 늘어나게 된다. 따라서 이웃 등 주변에 리트윗을 활발히 하는 이용자가 많을수록 트위터 네트워크의 확산경로에 따라 존재하는 유력자의 리트윗을 받을 확률이 증가하게 될 것이다. 예를 들어 포드자동차는 자동차 분야 소셜 미디어 활동이 활발한 100명을 모집해 이들을 통해 매달 1개씩 총 6개월 간 미션을 수행하도록 하여 트위터 등 SNS와의 연계 프로모션 활동을 전개한 바 있다[15]. 이들의 활발한 활동으로 유튜브 조회수가 700만 건을 넘어서고 트위터에 400만회 노출되는 실적을 올렸다고 한다.

IV. 노출도의 네트워크 중앙성

앞에서 우리는 그래프를 이용하여 입소문의 영향력을 노출도의 관점에서 살펴보았다. 일반적으로 네트워크에 위치하는 특정 노드의 위상은 정보의 전파와 확산에 중요한 원인을 제공한다고 한다[16]. 그러면 노출도는 이러한 노드의 위상과 어떠한 연관관계를 가지고 있을까? 노드의 위상과 관련한 네트워크구조연구에 있어 많은 부분이 중앙성(centrality)의 개념과 관련이 있다[17]. 이런 연구들은 “네트워크에서 가장 중요한 혹은 가장 중심적인 노드는 무엇인가”라는 물음에서 출발한다. 중앙성을 측정하는 척도에는 다음과 같은 것들이 있다.

첫째, 가장 단순한 중앙성 측정치로 노드의 연결정도 중앙성(degree centrality)을 들 수 있다. 연결정도(degree)는 노드에 연결되어 있는 링크의 수를 말하며

흔히 그 노드가 가지는 인기도(popularity)를 반영한다. 가장 간단한 측정치이지만 중앙성에 관한 여러 가지 시사점을 줄 수 있다. 개인의 친구수가 많은 개인은 그렇지 못한 사람에 비해서 더 영향력이 있거나 더 많은 위신, 더 많은 정보 접근력을 가질 수 있다. 트위터 네트워크에서 팔로어 수나 논문의 피인용지수 등이 여기에 해당된다고 볼 수 있다.

둘째, 연결정도중앙성을 확장한 개념이 고유벡터중앙성(eigenvector centrality)이다. 고유벡터중앙성은 얼마나 중요한 사람들과 연결되어 있는냐는 개념이다. 노드가 가지고 있는 연결정도 뿐만 아니라 연결된 노드들의 연결정도도 고려한다. 즉, 어느 특정노드에 연결되어 있는 노드들은 모두 동등하지 않다는 전제를 가진다. 이에 따라 노드의 중요성을 연결되어 있는 다른 노드들의 중요성에 따라서 달리 측정하는 방식이다. 예를 들어 트위터의 경우 두 사람의 팔로어 수가 같을지라도 그중 비중있는 팔로어를 많이 가지고 있는 사람을 더 중요하게 간주하는 방식이다. 각 노드 i 의 중앙성을 x_i 라 하면 노드 i 의 이웃들의 중앙성의 합은 식(1)과 같이 표시할 수 있다.

$$x_i = \alpha \sum_j A_{ij} x_j \quad (1)$$

여기서 α 는 양의 상수이며 A_{ij} 는 인접행렬(adjacency matrix)의 요소로 노드 i 와 노드 j 사이에 링크가 있으면 1이고 없으면 0의 값을 가진다.

셋째, 구글의 웹검색 순위기술인 페이지랭크(PageRank)는 고유벡터중앙성을 보완한 것으로 식(2)와 같이 표시할 수 있다.

$$x_i = \alpha \sum_j A_{ij} \frac{x_j}{k_j^{out}} + \beta \quad (2)$$

즉, 모든 노드에 고정된 중요성(β)를 부여하며, 특정 노드의 중앙성에 인접 노드의 중앙성이 모두 반영되는 것이 아니라 그 인접 노드로부터 나가는 연결정도(k_j^{out})를 고려하여 조정하는 것이다. 예를 들어 야후

(Yahoo)가 어떤 특정 사이트를 링크했다면 야후의 높은 중앙성으로 인해 그 사이트의 중앙성은 매우 크게 나타날 것이다. 그러나 실제 야후가 링크하고 있는 사이트는 수백만 개에 달하므로 개별 링크 페이지 각각에 대한 기여도는 매우 작다고 할 수 있다. 이러한 점을 반영하기 위하여 인접 노드의 중앙성을 외부 링크수로 나누어 조정하는 것이다.

넷째, 사이중앙성(betweenness centrality)을 들 수 있다. 이것은 하나의 노드가 다른 노드의 사이에 위치하는 정도(brokerage)를 측정한다[18]. 즉, 하나의 노드가 다른 노드들 간의 최단거리에 위치하는 빈도가 높을수록 사이중앙성은 높아진다. 만일 노드 i 가 노드 s 에서 노드 t 까지의 최단거리상에 위치한다면 n_{st}^i 의 값이 1이고 그렇지 않으면 0이라고 할 때 노드 i 의 사이중앙성 x_i 는 다음의 식(3)과 같이 표시할 수 있다.

$$x_i = \sum_{st} n_{st}^i \quad (3)$$

높은 사이중앙성을 가진 노드들은 다른 그룹 혹은 다른 노드들 사이에 흐르는 정보에 대한 통제를 가짐으로 상당한 영향력을 가질 수 있다. 즉, 특정 주제에 관한 가장 많은 정보나 메시지가 흐르는 노드이며 제외될 경우 전체적인 네트워크의 연결에 지장을 초래할 가능성이 높은 노드이다.

다섯째, 인접중앙성(closeness centrality)은 네트워크 상에서 한 노드가 다른 노드들과 얼마나 가까이 있는가를 나타낸다. 만일 정보가 네트워크 전체에 퍼져야 한다면 인접중앙성이 높은 노드들은 상대적으로 빨리 그 정보를 전체에 전달할 수 있다. 인접중앙성은 다른 노드들에 도달하는 평균거리로 측정한다. d_{ij} 를 노드 i 에서 노드 j 까지 가는 최단경로의 길이라고 할 때 노드 i 의 인접중앙성 C_i 는 식(4)와 같이 구할 수 있다. 즉, C_i 는 평균거리의 역수로 측정된다.

$$C_i = \frac{n}{\sum_j d_{ij}} \quad (4)$$

표 3. 중앙성 순위비교

순위	연결정도중앙성	고유벡터중앙성	페이지랭크	사이중앙성	인접중앙성
1	cybercoex	oojoo	estima7	seiyonl	oojoo
2	estima7	estima7	oojoo	SKTworld	seiyonl
3	oojoo	TechIT_kr	cybercoex	oojoo	estima7
4	iolothebard	goodgle	hongss	estima7	goodgle
5	hongss	iolothebard	iolothebard	cybercoex	TechIT_kr
6	TechIT_kr	pdaclub	TechIT_kr	yeomjihyun	iolothebard
7	myoldfriend	delight412	SKTworld	dpwlwkd93	hagiosruah
8	goodgle	grace1213	mushman1970	SocialCaster	SocialCaster
9	delight412	hagiosruah	goodgle	iolothebard	pdaclub
10	mushman1970	myoldfriend	SocialCaster	hongss	nerdstory

사용자별 중앙성 정도를 조사하기 위하여 ‘모바일웹’ 검색데이터를 이용하였다. [표 3]은 총 115개의 노드에 대하여 위에서 언급된 5가지 척도를 NodeXL에서 각각 계산하여 얻은 결과, 상위 10개의 노드(사용자)를 나타내고 있다. 여기서 다음과 같은 점을 확인할 수 있다.

첫째, 5가지 중앙성 척도에서 고루 높은 순위를 보이고 있는 사용자는 oojoo, estima7 등이다. 이들은 앞서 거론한 노출도가 매우 높게 나온 사용자들로 노출도가 네트워크의 중앙성 등 구조적 위상과 일정부분 연관성이 있음을 시사해 준다고 볼 수 있다. 둘째, 일반적인 트위터 네트워크에서의 연결정도중앙성은 해당 노드의 팔로어 수와 관련있다고 볼 수 있다. 그러나 검색네트워크의 경우 특정 주제를 언급한 팔로어 관계만이 링크로 나타나게 된다. 이러한 점을 반영하여 팔로어 수가 많은 사용자보다는 해당 주제에 관한 리트윗 등이 많은 사용자가 연결정도중앙성이 높은 것으로 나타나고 있다. 셋째, 사이중앙성과 인접중앙성이 높게 나온 seiyonl은 [그림 4]에서 보는 바와 같이 매우 미약한 노출도를 보이고 있다. 즉, seiyonl는 그래프 중간에서 SKTworld와 iolothebard사이에 위치하여 양쪽 하위네트워크를 연결하는 중요한 구조적 위상을 가지고 있으나 모든 링크가 주로 다른 노드를 팔로어하고 있어 자신의 트윗을 노출시키지 못하고 있다. 이 같은 측면은 seiyonl가 고유벡터, 페이지랭크 등의 척도에서 낮은 값을 보이는 사실에서도 확인할 수 있다.

앞에서 유력자의 유형을 팔로어 수, 리트윗 수, 노출도로 나누어 비교해 보았다. 이제 이러한 유력자 유형

들이 네트워크의 구조적 척도와 어떤 연관성을 갖는지 좀 더 살펴보기 위하여 전체 115개 노드를 대상으로 유력자 유형과 5가지 척도간의 상관관계를 조사하였다. [표 4]를 통해 다음과 같은 사실을 확인할 수 있다. 첫째, 검색네트워크에서 팔로어 수는 대부분의 척도에서 리트윗 수나 노출도보다 낮은 상관관계를 보이고 있다. 즉, 검색네트워크에서 사용자의 팔로어 수는 중앙성 척도와 큰 관련이 없는 것으로 나타나고 있다. 둘째, 리트윗은 연결정도 중앙성과 페이지랭크, 사이중앙성과 높은 상관관계를 보이고 있다. 셋째, 노출도는 인접중앙성을 제외한 나머지 4가지 척도와 비교적 높은 상관관계를 보이고 있다. 이 같은 사실로부터 본 논문의 트위터 검색네트워크에서는 팔로어 수보다는 리트윗 수나 노출도가 구조적 척도와 높은 상관관계를 가지고 있음을 확인할 수 있었다. 특히 노출도의 경우 대부분의 척도와 고른 상관관계를 유지하고 있었는데 단순한 팔로어 수가 아니라 유력자와의 상호작용에 비례하기 때문에 고유벡터중앙성 및 페이지랭크와의 상관관계가 비교적 높게 나타나고 있다는 점이 특징이라고 할 수 있다.

표 4. 중앙성과의 상관관계 비교

	팔로어 수	리트윗 수	노출도
연결정도 중앙성	0.308**	0.793**	0.568**
고유벡터 중앙성	0.215*	0.390**	0.454**
페이지랭크	0.414**	0.754**	0.652**
사이 중앙성	0.544**	0.550**	0.615**
인접 중앙성	0.242**	0.210*	0.275**

(주) *, **는 각각 5%, 1%수준에서 통계적으로 유의함을 나타냄

V. 결론

입소문효과 극대화를 위한 유력자의 역할은 트위터 네트워크에서도 매우 중요하다고 볼 수 있다. 그러나 누구를 유력자라고 할 수 있는가하는 논의는 아직 진행중이라고 볼 수 있다. 본 논문에서는 기업 마케팅의 관점에서 트위터 유력자를 파악하고자 하였다. 즉, 기업의 특정 마케팅 메시지를 누가 가장 많이 노출시킬 수 있는가라는 관점에서 시작하였다. 이러한 관점에서 현재 유력자의 주요한 기준으로 거론되고 있는 팔로어 수와 리트윗 수를 트위터의 특정 검색네트워크 데이터를 이용하여 관찰해 보고자 하였다. 이를 위하여 트윗의 노출도를 두 가지 측면에서 살펴보았다. 즉, 노출도를 트윗 작성자의 팔로어 수와 그 트윗이 리트윗되는 경우 리트윗하는 사용자들의 팔로어 수의 합으로 보았다. 검색결과와 그래프비교를 통해 다음과 같은 사실을 확인할 수 있었다. 첫째, 리트윗이 없는 팔로어 유력자에 비하여 팔로어 유력자의 리트윗을 받는 상대적인 소수자들의 노출도가 크다는 것을 알 수 있었다. 둘째, 리트윗 유력자중에서도 팔로어 유력자에 의하여 리트윗을 받지 못하는 사용자의 노출도는 매우 미미하다는 것을 알 수 있었다.

이와 같은 사실은 노출도를 높이기 위해서는 실제 유력자도 중요하지만 그런 유력자의 리트윗을 유도하는 사용자의 파악이 더 중요하다는 점을 시사해준다고 볼 수 있다. 최근 리트윗을 활발히 하는 미디어적인 역할이 유력자의 중요한 특징으로서 부각되고 있기도 하다. 따라서 이를 이용하여 이웃 등 주변에 리트윗을 활발히 하는 이용자를 적극 활용하는 것이 중요하다고 볼 수 있다. 이러한 점들의 중요성은 앞에서 언급한 포드자동차의 사례나 싸이의 뮤직비디오사례에서도 발견되고 있다.

결론적으로 트위터를 이용한 입소문 마케팅을 실행하는 데에 있어서 단순한 팔로어 수 증가보다는 유력자의 리트윗을 적극적으로 유도할 수 있는 열성적인 트위터 사용자들을 전략적으로 발굴하고 활용하는 것이 마케팅 효과 측면에서 매우 중요하다고 볼 수 있다. 이와 함께 노출도의 경우 대부분의 중앙성 척도와 고른 상관

관계를 유지하고 있어 노출도가 높은 이용자일수록 네트워크구조상에서도 중심적인 위상을 차지한다는 점도 살펴볼 수 있었다. 특히 고유벡터중앙성 및 페이지랭크와 비교적 높은 상관관계를 보이고 있어 노출도가 단순한 팔로어 수가 아니라 유력자와의 긴밀한 관계를 보여준다는 점이 부각되고 있다. 이러한 결론들은 비록 특정주제어에 한정된 네트워크 데이터를 기반으로 하고 있지만 향후 보다 다양한 주제들을 통해서 비교 검증할 필요가 있다고 보여진다.

참고 문헌

- [1] 안지혜, 민병현, “영화 마케팅 채널로서 소셜미디어의 가능성”, 한국콘텐츠학회논문지, 제11권, 제6호, pp.228-241, 2011.
- [2] 송용태, “프리어나온싱 된 신제품 정보의 생동감이 인터넷커뮤니티 유저의 e-WOM와 인터넷 포스팅의도에 미치는 영향”, 한국콘텐츠학회논문지, 제11권, 제2호, pp.312-321, 2011.
- [3] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, “Everyone is an Influencer: Quantifying Influence on Twitter,” WSDM'11, February 9 - 12, Hong Kong, China, 2011.
- [4] D. J. Watts and P. S. Dodds, “Influentials, networks, and public opinion formation,” J. of Consumer Research, Vol.34, No.4, pp.441-458, 2007.
- [5] E. Katz and P. Lazarsfeld, *Personal influence : The part played by people in the flow of mass communications*, Free Press, 1955.
- [6] E. Rogers, *Diffusion of innovations*, Free Press, 1962.
- [7] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, “Measuring user influence in Twitter: The million follower fallacy,” Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, Washington D.C., USA, 2010.

[8] A. Leavitt, E. Burchard, D. Fisher, and S. Gilbert, *The influentials : New approaches for analyzing influence on Twitter*, Web Ecology Project, 2009.

[9] D. Hansen, B. Shneiderman, and M. A. Smith, *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*, Morgan Kaufmann, 2010.

[10] Pew Internet & American Life, "Half of adult cell phone owners have apps on their phones," Retrieved December 3, from <http://pewinternet.org/Reports/2011/Apps-update.aspx> 2011.

[11] 한국콘텐츠진흥원, "모바일 애플리케이션 비즈니스 현황과 전망", KOCCA포커스, 통권48호, 2012.

[12] 중앙일보, "강남스타일, 최다 검색 국가 어딘지 보니..." 2012(8.27).

[13] 이성규, "싸이 <강남스타일>이 성공한 진짜 이유?", 프레시안, 2012.8.16.

[14] 이원태, 차미영, 양혜륜, "소셜미디어 유력자의 네트워크 특성 한국의 트위터 공동체를 중심으로", 언론정보연구, 제48권, 제2호, pp.44-79, 2011.

[15] 안중배, "소셜미디어의 특징과 마케팅활용사례", 자동차경제, 11월호, pp.54-57, 2011.

[16] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.

[17] M.E.J. Newman, *Networks : An Introduction*, Oxford, 2010.

[18] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, Vol.40, pp.35-41, 1977.

저 자 소 개

유 병 국(Byong-Kook Yoo)

정회원



- 1985년 2월 : 연세대학교 경제학과(경제학사)
- 1988년 2월 : 한국과학기술원 경영과학과(공학석사)
- 1992년 8월 : 한국과학기술원 경영정책과(공학박사)

▪ 1994년 ~ 현재 : 인천대학교 무역학부 교수

<관심분야> : e-비즈니스, 네트워크경제, 환경경제

김 순 홍(Soon-Hong Kim)

정회원



- 1985년 2월 : 중앙대학교 산업경제학과(경제학사)
- 1987년 2월 : 중앙대학교 산업경제학과(경제학석사)
- 1987년 2월 : 중앙대학교 산업경제학과(경제학박사)

▪ 1991년 ~ 현재 : 인천대학교 무역학부 교수

<관심분야> : 유통, 물류, CRM