

## 텍스트마이닝을 통한 댓글의 공감도 및 비공감도에 영향을 미치는 댓글의 특성 연구\*

김정훈\*\* · 송영은\*\* · 진윤선\*\* · 권오병\*\*\*

### Applying Text Mining to Identify Factors Which Affect Likes and Dislikes of Online News Comments\*

Jeonghun Kim\*\* · Yeongeun Song\*\* · Yunseon Jin\*\* · Ohbyung kwon\*\*\*

#### ■ Abstract ■

As a public medium and one of the big data sources that is accumulated informally and real time, online news comments or replies are considered a significant resource to understand mentalities of article readers. The comments are also being regarded as an important medium of WOM (Word of Mouse) about products, services or the enterprises. If the diffusing effect of the comments is referred to as the degrees of agreement and disagreement from an angle of WOM, figuring out which characteristics of the comments would influence the agreements or the disagreements to the comments in very early stage would be very worthwhile to establish a comment based eWOM (electronic WOM) strategy. However, investigating the effects of the characteristics of the comments on eWOM effect has been rarely studied. According to this angle, this study aims to conduct an empirical analysis which understands the characteristics of comments that affect the numbers of agreement and disagreement, as eWOM performance, to particular news articles which address a specific product, service or enterprise per se. While extant literature has focused on the quantitative attributes of the comments which are collected by manually, this paper used text mining techniques to acquire the qualitative attributes of the comments in an automatic and cost effective manner.

Keyword : eWOM, Text Mining, Online News Comments, Online Marketing

Submitted : May 21, 2015

1<sup>st</sup> Revision : June 2, 2014

Accepted : June 6, 2015

\* This work was supported by the ICT R&D program of MSIP/IITP(R0126-15-1007, Curation commerce based global open market system development for personal happiness enhancement).

\*\* 경희대학교 일반대학원 경영학과

\*\*\* 경희대학교 경영대학, 교신저자

## 1. 서 론

인터넷 기사에 대한 댓글은 기사에 담겨 있는 내용에 대한 일반인들의 의견을 파악할 수 있는 가장 개방적이고 풍부한 방법이다. 기업은 댓글을 통해서 자신의 기업이나 제품, 서비스에 대한 의견을 파악할 수 있으며, 부정적인 의견이 접수되는 경우 신속하고 적극적인 대응을 통해 해결해 나가야 한다. 또한 댓글의 해당 기업에 대한 홍보 역할도 하게 되므로 댓글은 중요한 eWOM(Electronic Word of Mouth) 수단이라고 볼 수 있다. 특히 최근 모바일 단말기 보급의 증가로 댓글의 가시성이 더욱 증가함으로 인해 일반적인 WOM 매체보다 더욱 영향력이 크다(Bickart and Schindler, 2001).

그런데 인터넷 상에는 수많은 기사와 그에 대한 더욱 많은 댓글들이 실시간으로 스트리밍되기 때문에 자신의 기업에 관련한 댓글을 실시간으로 모니터링하고 대응하기란 쉬운 일이 아니다. 더욱이 댓글들 중에는 주변인의 관심을 받지 못하고 새롭게 생성되는 다른 댓글에 의하여 사라지거나 비공감을 유발하여 오피니언으로서의 효력을 상실하는 댓글도 상당수이기 때문에 발생하는 모든 댓글에 대해 일일이 관리하는 것은 매우 많은 비용이 소요된다. 따라서 댓글의 발생 초기에 댓글이 추후 얼마나 큰 영향력을 행사하는 오피니언이 될지를 효율적이면서도 정확하게 파악하는 것이 댓글을 통한 eWOM 관리에 매우 중요하다고 하겠다(Christodoulides et al., 2012).

기사에 대한 댓글 분석에 대한 기존의 연구는 어느 한 시점에서 댓글의 감성을 파악하여 긍부정성을 판단하고 오피니언 마이닝하는 형태가 주를 이루었다. 물론 댓글의 공감 수 등에 대한 시계열 자료를 통해 분석하는 연구도 진행된 바 있으나(O'Connor et al., 2010), 댓글의 일정 시점 후의 예상되는 공감수를 추론하는 수준에는 미치지 못하고 특정 이슈에 관련된 키워드를 수집하는 수준이었다. 하지만 키워드만을 대상으로 분석할 경우 특정 이슈에 연관되지 않은 데이터까지도 무분별

하게 수집하게 되므로 미래 시점에서의 공감수를 예측하는 정확도가 떨어지게 된다. 댓글이 중요한 마케팅 수단임을 고려할 때, 댓글에 대한 공감 혹은 비공감의 횟수도 댓글 기반 마케팅에 중요한 성과변수이므로 이를 예측하는 연구는 매우 중요하다. 하지만 이를 본격적으로 연구한 문헌은 거의 없다. 한편 댓글 내용의 긍부정성(Fan and Sun, 2010)이나 법적 규범적 평가(Maynard, 2006), 내용에 담긴 writer의 태도 혹은 주관적 웰빙(Thelwall, 2008), 내용으로 인해 야기하는 공감 정도(Saveski et al., 2014)는 모두 댓글이 미치는 영향에 대한 연구이나 모두 단편적인 요인을 다루고 있어 댓글의 여러 특성을 반영한 종합적인 영향 연구가 거의 없다.

따라서 본 연구의 목적은 텍스트마이닝 관점에서 뉴스 기사의 댓글 특성을 정립하고 그 특성들을 통하여 그 댓글이 앞으로 어떻게 평가 될지를 예측하는 방법을 제안하는 것이다. 첫째, 뉴스 기사의 특성을 정립하기 위하여 텍스트마이닝 기법을 통해 확인하고자 한다. 둘째, 이러한 특징들이 온라인 뉴스 댓글의 성과에 영향을 미치는 요인이 무엇인지를 다양한 통계기법을 통해 확인하는 것이다. 즉, 텍스트 마이닝 관점에서 뉴스 기사의 댓글 특성을 정립하고 그 특성들을 통하여 그 댓글이 앞으로 어떻게 평가 될지를 예측하는 방법을 제안하는 것이다.

이 논문은 다음과 같이 구성되어 있다. 제 2장에서는 eWOM과 온라인 뉴스 댓글에 대한 선행연구를 기술하였다. 제 3장과 제 4장은 각각 연구모형을 제안하고 그에 따르는 실증분석 결과를 제시하였다. 마지막으로 제 5장은 연구의 시사점, 한계 및 향후 연구 방향을 기술했다.

## 2. eWOM으로서의 댓글

WOM은 “어떤 소비자가 특정한 제품, 서비스 및 판매기업의 소유권, 활용도 또는 특성 등에 대해 다른 소비자에게 제공하는 정보 커뮤니케이션(Westbrook, 1987: 261)”이라고 정의된다. 소비자

는 일반적으로 WOM이 기업이 스스로 제공하는 마케팅이나 광고보다도 더 신뢰할 만한 것으로 받아들이고 있으므로(De Matos and Rossi, 2008; Katz and Lazarsfeld, 1955; Lau and Ng, 2001; Newman and Staelin, 1972), 매우 중요한 마케팅 방법이라고 볼 수 있다.

특히 온라인을 통해 이루어지는 WOM을 의미하는 eWOM은 공간과 시간의 제약 없이 전세계적으로 더 빨리 확산되는 능력을 가지고 있기 때문에 오프라인 상의 WOM보다도 더 영향력이 있다고 알려져 있다(Hennig-Thurau et al., 2004; Phelps et al., 2004). eWOM의 수단으로는 블로그나, 리뷰 사이트, 추천 사이트 등이 있을 것이다. 그러나 특정 기업의 제품, 서비스, 기업 자체에 관련한 온라인 기사에 대한 eWOM로서의 댓글도 매우 중요하다.

소비자들은 댓글이라는 플랫폼을 통해서 특정 대상에 대해서 다양한 정보를 나눌뿐더러 평가도 할 수 있다(Yang, 2008). 댓글을 통해서 소비자들은 서로 편리하고 적극적인 커뮤니케이션이 가능해졌으며, 이를 통해 여론이 생성할 수 있다. 특히 기사는 어떤 의도를 가지고 작성될 수 있는데 그 기사를 보는 독자이자 소비자들은 댓글 행위 및 그 댓글에 대한 공감(likes)으로서 반응할 수 있으며(Moon et al., 2013), 실제로 다른 eWOM 매체에 비하여 그 영향은 매우 크다(Kim, 2006). 또한 댓글이 해당 기업에 긍정적이라면 해당 기업에 대해 긍정적인 인기, 태도 및 행위변화가 이루어 질 수 있다(Kim, 2006; Bae, 2009).

연구결과를 통해 알 수 있듯이 댓글을 통해 여론의 동향과 소비자의 반응, 관심, 의견 등을 손쉽게 감지할 수 있기 때문에 기업에서도 댓글의 중요성을 인식하여 소비자와 의견을 공유할 수 있는 하나의 도구로 활용 할 수 있어야 한다(Moon et al., 2013). 활용의 범위와 방법에 따라 기업의 이미지를 손쉽게 높일 수도 있고 손상시킬 수도 있기 때문이다.

댓글을 통해 발생하는 기업에 관한 여론을 긍정적인 방향으로 조성하고 이미지 훼손 방지를 위해

서 우호적인 댓글을 게시하거나 관리하고, 악성 댓글과 부정적인 여론에 대한 적절한 대책마련을 해 두어야 한다. 또한 일반 댓글보다 상위에 공감 또는 비공감을 많이 얻은 댓글이 사용자들의 주의를 더 많이 끌기 때문에(Moon et al., 2013), 큰 영향력을 가질 확률이 높음을 이해하여야 한다.

하지만 기존 논문에서는 일반 댓글과 베스트 댓글을 구분하는데 대한 연구가 부족한 실정이다. 즉, 댓글에 대한 공감 혹은 비공감을 중요한 성과 변수라 판단하고 온라인 뉴스 댓글의 성과에 영향을 미치는 요인을 밝힐 필요가 있다.

### 3. 온라인 뉴스 댓글의 성과에 영향을 미치는 요인

#### 3.1 이론 모형

온라인 뉴스 기사에 대해 업로드하게 되는 각 댓글은 그 공감수나 비공감수, 해당 댓글에 대한 댓글의 수와 그 찬반의 정도 등에 의하여 얼마나 영향력 있는, 또는 동의를 얻는 댓글인지를 판단할 수 있다. 물론 일부 사이트의 경우에는 댓글을 올리는 기능은 없고 대신 기사에 대해서 동의, 비동의를 하거나 자신의 감정(기뻐요, 화나요, 슬퍼요, 혼썬해요 등)을 표현하게 하여 댓글의 기능을 대신하게 한다. 물론 이 경우에는 자신의 동의여부나 감정을 표현한 이에 대한 동의 비동의를 표현할 수는 없다. 그러나 <Figure 1>과 같이 대부분의 뉴스 사이트에서는 댓글은 대체로 댓글 내용과 댓글 저자에 대한 간략한 정보, 댓글 시점에 대한 정보 외에 그 댓글을 보는 다른 독자들이 그 글에 얼마나 동의(recommend, likes 등)하는지에 대한 정보를 제공한다. 그리고 일부는 동의하지 않는 정도(dislikes)를 제공하기도 한다. 따라서 댓글에서의 WOM은 댓글에 대한 동의의 수 및 비동의의 수로 표현된다. 그리고 기사를 활용하여 사회적 영향을 미치려는 개인이나 기업이라면 이러한 댓글의 성과는 중요한 요인이 된다.



<Figure 1> Online News Comments

eWOM 행동으로서의 댓글은 댓글을 활용하여 사회적 영향을 미치는 성과를 내려는 기업의 입장에서 중요한 요인이다. 일반적으로 eWOM의 성과는 주로 eWOM 수용의도로 측정한다(Gupta and Harris, 2010). 그런데 본 연구의 초점은 간접적인 질문 방법이 아닌 객관적이고 공개된 자료를 바탕으로 텍스트마이닝 한 결과물로 분석하는 것에 맞추어져 있기 때문에, eWOM에 대한 수용 정도는 사용자를 대상으로 하는 설문 방식에 의존하지 않고 공개된 댓글 관련 정보로부터 측정하려고 했다.

이를 온라인 뉴스 기사에 적용하면 <Table 1>과 같이 한 댓글에 여러 특성들이 획득 가능하다. 이중 온라인 뉴스 기사에 대해 업로드하게 되는 eWOM으로서의 각 댓글에 대한 수용 정도는 그 공감수나 비공감수, 해당 댓글에 대한 댓글의 수와 그 찬반의 정도 등에 의하여 얼마나 영향력 있는, 또는 동의를 얻는 댓글인지와 관련 있다.

이렇게 자가 설문식이 아닌 개방된 자료로부터의 댓글의 성과를 고려해 볼 때, 이에 영향을 미치는 요인에 대해 본 연구에서는 구문론적 속성(syntactic

features)과 의미론적 속성(semantic features), 그리고 초기 반응(initial response)으로 나누었다. 구문론적 속성은 문장의 구조와 관련된 것으로 작성된 댓글의 단어 수(message length)를 대표적 속성으로 보았으며, 의미론적 속성으로는 메시지에 대한 호감도와 사용하는 언어의 강도, 그리고 initial response로서는 시작하지 초기 시점에서의 댓글에 대한 공감수 및 비공감수이다. 결국 <Figure 2>와 같은 연구모형을 제시한다.

### 3.2 댓글에 대한 반응(Reader's WOM)

본 연구에서 고려 중인 종속변수, 즉 댓글에 대한 독자의 반응은 <Table 2>와 같이 댓글에 대한 공감수와 비공감수, 변곡점 도달 시점과 변곡점 도달시의 공감수/비공감수 등 여러 관점에서 볼 수 있다. 여기서 변곡점 도달 시점이란 공감/비공감수의 변화가 정체되기 시작하는 시점까지 소요 시간으로서, 변곡점 도달에 소요되는 시간이 길면 길수록 댓글을 보는 사람들이 더 오랫동안 활발하

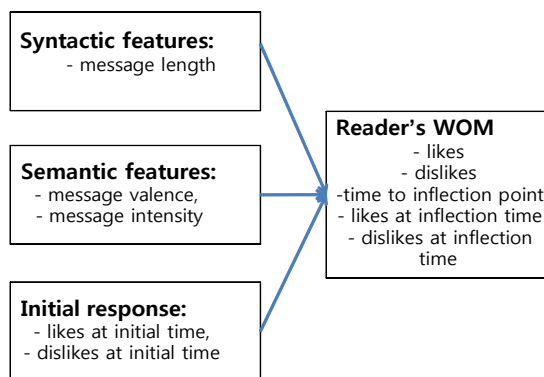
<Table 1> Source of Obtaining Comments(e.g Naver News)

Characteristics of comments	Sample Code
Comment ID	<li class = “_snsUser(0MWq3kG25iA =, naver) _commentReplyNo(1088954)”> <div class = “comment_area _comment_area comment_area_over”>
User ID	<div class = “author_wrap”> <div class = “author”> <span class = “maskname”> 3765**** </span> </div> </div>
Content	<p class = “txt”> Again! Seoul Gyeonggi Chungcheong Cheju. We are not different from each other. Please do not say bad things about each other. Please, stop it! </p>
Time	<div class = “func”> <span class = “date”> 2015.01.08 8 : 22pm </span>
Likes	<span class = “btn_recomm”> <a href = “#” class = “_check_layerlogin _sympathy(1088954) _nclicks(rpl.sym)”> Likes </a> <em> 0 </em> </span>
Dislikes	<span class = “btn_recomm”> <a href = “#” class = “_check_layerlogin _antipathy(1088954) _nclicks(rpl.opp)”> Dislikes </a> <em> 3 </em> </span>

<Table 2> Dependent Variables and Operationalization

Dependent Variables	Operationalized definition	Type
Likes	The number of likes at the finish time of data collection	Integer
Dislikes	The number of dislikes at the finish time of data collection	Integer
Time to inflection point	Elapsed time from start point to inflection point	Time (minutes)
Likes/dislikes at inflection point	The number of likes/dislikes at inflection point	Integer

게 공감/비공감했다는 것을 의미한다. 즉, 해당 댓글의 영향력 또는 파급효과가 긍정적이든 부정적이든 크다는 것을 나타낸다. 변곡점 도달시점과 같은 측면을 분석하는 것으로 변곡점 도달시의 공감수/비공감수도 있다.



<Figure 2> Research Model

### 3.3 공감수 및 비공감수의 증가 유형

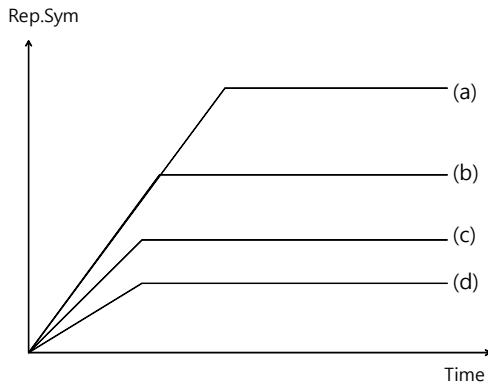
본 논문에서 댓글의 성공 여부는 댓글에 대한 공감수의 증가 속도(velocity), 포만(saturation)되지 않고 증가하는 기간(duration), 포만 후 재도약하는 여부(revitalization)로 판단할 수 있다. 그러나 일반적으로 한번 사라진 댓글이 다시 읽혀져서 재도약하는 경우는 그 댓글을 복사하여 다시 게시하기 전에는 거의 불가능하다고 볼 때, 증가속도와 증가기간으로 판단할 수 있을 것이다. 그렇다면 <Figure 3>과 같이 증가기간이 동일한 경우에는 증가속도가 더 빠른 것이 좋으며( $c > d$ ), 증가속도가 동일하다면 증가기간이 더 긴 것이 성공한 댓글이라고 볼 수 있을 것이다( $a > b$ ).

그렇다면 시간의 흐름에 따른 댓글에 대한 공감수 증가 패턴은 상당시간 시간이 경과한 후에는 모든 댓글의 공감수는 정체한다고 볼 때, 처음에 급

상승 후 후에 정체되는 유형과 처음부터 지속적으로 증가가 거의 이루어지지 않고 정체되는 두 가지 유형으로 나눌 수 있다.

H1-1 : 댓글의 공감수 증가유형이 증가형일 경우 정체형인 경우보다 최종 공감수가 더 많다.

H1-2 : 댓글의 비공감수 증가유형이 증가형일 경우 정체형인 경우보다 최종 비공감수가 더 많다.



〈Figure 3〉 Type of Likes/Dislikes Curve

### 3.4 댓글의 긍부정성(Message Valence)

댓글의 긍부정성(message valence)은 평가적 리뷰 내용의 방향성으로서 칭찬과 같이 긍정적이거나 불평과 같이 부정적일 수 있다(Cheung et al., 2009). 댓글의 긍부정성은 작성자의 감성 정도에 관련되어 있다(Richins, 1983).

댓글의 긍부정성에 대해서 독자들의 반응은 비대칭적일 수 있다. 특히 부정적인 댓글에 대해 더 강렬하게 반응하는 경향이 있는데, 이의 근거로 부정적 편이 이론(Theory of Negativity Bias)을 고려할 수 있다. 부정적 편이 이론이란 부정적 자극이나 경험을 받을 때가 긍정적인 경우보다 더 많이 기억되고 더 큰 파급효과가 있다는 것이다(Fiske, 1980; Rozin and Royzman, 2001; Baumeister et al., 2001; Christodoulides et al., 2012). 일

부 연구에서는 이러한 부정적 편이가 eWOM에서도 관찰됨을 보이고 있다(Park and Lee, 2009). 따라서 부정적 댓글은 긍정적 댓글보다 더 많은 반응을 보일 것이다.

H2-1 : 댓글의 부정성이 높을수록 최종 공감수가 증가할 것이다.

H2-2 : 댓글의 부정성이 높을수록 최종 비공감수가 증가할 것이다.

### 3.5 댓글 사용 언어의 강도(Message Intensity)

댓글에 대한 감성 정도와 함께 댓글의 긍부정성을 결정하는 요인인 댓글 사용 언어의 강도는 자신의 긍부정성을 얼마나 강하게 표현하느냐를 의미한다. 언어적으로 보면 형용사의 수식하는 부사로 사용 언어의 강도가 결정된다.

H3-1 : 댓글 사용 언어의 강도가 높을수록 최종 공감수가 증가할 것이다.

H3-2 : 댓글 사용 언어의 강도가 높을수록 최종 비공감수가 증가할 것이다.

### 3.6 댓글의 양(Message Length)

기사에 대한 댓글은 제품평가 사이트 등의 댓글에 비하여 장황한 분석보다는 간결하게 감정을 전달하는 체계이다. 따라서 기사의 댓글은 문장이 짧고 주관적이다. 그러므로 기사의 댓글은 감정을 극적으로 전달하는 은어나 약어가 많이 쓰이기도 하는 것이다.

댓글의 양은 정보과잉이론(Theory of Information Overload)과 연계하여 고려할 수 있다. 정보과잉 이론에 의하면 정보 과잉이란 너무나 많은 정보를 제공받는다라는 주관적 생각과 연관된 스트레스에 관한 것이다(Chen, 2012). Blom이 지적한 것처럼 정보 과잉은 정신적 피로나 관심 저하로 이어져 역효과가 일어난다(Blom, 2011). 따라서 댓글

글에서 사용하는 단어의 수가 증가할수록 그 댓글에 대한 공감을 이끌어내기 어려울 것이며 공감수의 증가는 쉽게 정체될 것이다.

H4-1 : 댓글의 양이 많을수록 최종 공감수가 감소할 것이다.

H4-2 : 댓글의 양이 많을수록 공감수의 변곡점에서의 공감수는 감소할 것이다.

### 3.7 초기 시점의 공감수/비공감수

(Likes/Dislikes at initial time)

본 연구는 댓글이 게시된 시점에서 초기에 그 댓글이 일정 시간이 경과한 후 얼마나 많은 공감수나 비공감수를 획득할 것인지에 초점이 맞추어져 있다. 따라서 댓글 초기 시점에 획득 가능한 공감수와 비공감수가 과연 이후의 공감수나 비공감수에 영향을 주는지를 파악하기 위하여 초기 시점의 공감수 및 초기 시점의 비공감수를 독립변수로 고려하였다.

H5-1 : 초기 시점의 공감수가 많을수록 최종 공감수가 증가할 것이다.

H5-2 : 초기 시점의 비공감수가 많을수록 최종 비공감수가 증가할 것이다.

이상의 독립변수들에 대한 조작적 정의 및 변수의 유형을 <Table 3>에 정리하였다.

## 4. 텍스트마이닝

자료수집은 IT 관련 뉴스 기사에 달린 댓글로 국한하였다. 특정 기사에 대해서 댓글은 10개씩 무작위로 수집하였으며 총 387개를 선정했다. 수집 직후에 댓글에 대한 내용 특성을 인식하였다. 매 5분 간격으로 24시간 동안 댓글에 대한 공감 비공감수의 변화를 수집하여, 특정 댓글 하나 당 288회씩 자료를 수집하였다. 수집 시간을 24시간으로 한정하는 이유는 Tatar(2011)의 주장과 같이 대부분의 기사는 및 댓글에 대한 평가는 하루 만에 종료되기 때문이다.

한편 기사 및 댓글에 대한 메타 정보(기사 게재 시간, 댓글의 양, 댓글 기록시간, 공감수, 비공감수) 및 내용(text)은 소스에서 읽어올 수 있다. 여기에서 News의 ID번호를 확보한 후(예 : news003,0006294540), 그것을 gno라는 argument에 넣어주면 된다.

또한 비정형화된 문서 원천으로부터 정형화 자료로 정보저장소에 저장하는 과정은 다음과 같다.

(단계 1) 비정형 데이터 수집

(단계 2) 불필요 태그 정보(stopping words) 삭제 후 단위 댓글 별로 획득 정보 인식(<Table 4> 참조)

(단계 3) 정보저장소에 저장

이때 댓글의 특성 중 댓글의 길이, 언어강도, 댓글

<Table 3> Independent Variables

Independent Variables	Operational Definition	Type
Message valence	The positiveness/negativeness of a comment published by a reader	Real
Message intensity	Degree of strength of the words used in a comment	Real
Message length	The number of words used in a comment	Integer
Type of likes/dislikes curve	Types of the pattern of increase of likes/dislikes about a comment from publishing time(1 : increasing, 2 : stagnant)	Integer(binary)
Likes/dislikes at initial time	The number of likes/dislikes of a comment at 10 minutes since published by a reader	Integer

글의 긍부정성 측정을 위해서 텍스트마이닝 도구인 Semantria@를 사용하였다. 또한 Word Cloud 분석은 R을 사용하였다.

한편 공감수 및 비공감수의 증가유형은 변곡점의 인식과 관련되며 Java application을 통해 자동 계산하였다. 우선 변곡점은 현시점 기준 과거 m개 자료, 즉 최근 5×m 분간의 공감수 혹은 비공감수의 평균치를 활용하여 그 평균치와 이전 평균치 사이의 기울기가 일정 각도 이하로 n-1회 지속되면 n회 전에 변곡점이 도래했다고 판정한다. 이에 따라 댓글이 처음 게시된 원점에서 변곡점 이전 시점까지의 기울기가 일정 각도 이상이면 증가형, 그 이하이면 정체형으로 판정하였다.

## 5. 결 과

### 5.1 Word Cloud 분석

먼저 R을 사용하여 획득된 댓글에 등장하는 단어들의 빈도 분석을 실시하고 그 결과를 <Figure 4>와 같이 Word cloud로 나타내 보았다. 먼저 빈

도분석을 한 결과 다음과 같은 단어들이 상위에 올랐다. 단어들을 보는 것처럼 특정 기업의 명칭, 일반 명사들 외에 이모티콘이나 은어들도 같이 사용 되는 것을 볼 수 있다. 따라서 댓글의 특징 분석을 할 때 이러한 부분들을 고려하여 분석해야 한다.

### 5.2 증가유형별 통계

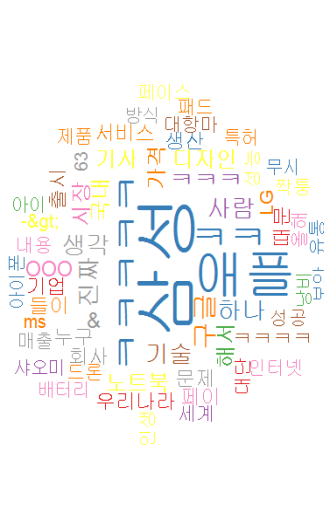
다음은 한 댓글에 대해 공감수 증가유형과 비공감수 증가유형이 어떻게 짝지어지는지를 분석한 것이다. 그 결과 총 396개의 댓글에 대해서 <Table 4>와 같은 결과가 도출되었다. 예를 들어 공감수 및

<Table 4> Frequency of Each Combination of Increasing Likes/Dislikes Types

		Dislikes		Total
		Increasing	Stagnant	
Likes	Increasing	Debates 89	Liked 130	219
	Stagnant	Disliked 10	Apathy 167	177
Total		99	297	396



(a) Min.freq = 2



(b) Min.freq = 10

<Figure 4> Screenshot of Word Cloud



<Table 5> Simple Statistics of Likes/Dislikes for Each Increasing Type

	Debates	Liked	Disliked	Apathy	Total
Likes	225.595 (283.959)	102.746 (220.006)	22.100 (26.231)	13.222 (13.620)	90.566 (202.225)
Dislikes	51.033 (113.788)	8.323 (9.633)	11.700 (14.164)	2.970 (3.637)	15.750 (57.605)
Time to inflection point	76.764 (37.816)	71.669 (37.259)	66.500 (42.498)	53.683 (46.239)	65.098 (42.710)
Likes at inflection point	198.258 (265.009)	98.677 (216.356)	21.500 (25.180)	12.874 (13.440)	82.924 (190.994)
Dislikes at inflection point	61.7980 (30.329)	48.854 (43.991)	23.200 (24.032)	19.952 (32.725)	38.927 (40.203)
The number of dislikes at inflection point	44.899 (94.016)	7.962 (9.354)	11.300 (13.864)	2.904 (3.645)	14.215 (48.003)
Likes at initial time	71.910 (143.042)	53.785 (128.862)	15.100 (23.015)	10.581 (13.017)	38.662 (103.946)
Dislikes at initial time	14.281 (25.423)	5.523 (8.103)	7.100 (10.986)	2.311 (3.218)	6.177 (13.980)
Message valence	-6.299 (1.145)	-0.346 (1.226)	-0.400 (1.114)	-0.425 (1.029)	-0.444 (1.130)
Message intensity	3.329 (2.504)	3.020 (2.460)	2.151 (1.276)	3.054 (3.146)	3.082 (2.766)
Message length	10.955 (8.752)	10.077 (9.049)	6.700 (4.360)	10.497 (11.018)	10.366 (9.809)

비공감수의 증가유형이 모두 증가형이라는 의미는 공감과 비공감이 모두 활발하게 증가하였음을 나타내는 것으로 특정 댓글에 대해 찬반이 격론 중이라는 뜻이다. 한편 증가형-정체형이라는 의미는 해당하는 130개의 댓글은 공감수가 증가하는 동안 비공감 수는 정체 상태라는 의미이니 대체로 게시한 댓글에 수긍하는 여론이 형성된 것으로 볼 수 있다.

또한 공감수 및 비공감수의 증가유형 조합별 기초통계량은 <Table 5>와 같다.

다음으로 공감수 및 비공감수의 증가유형에 따라서 최종공감수와 최종 비공감수의 차이가 있는지를 판단하기 위해 t-test를 실시하였으며 그 결과는 <Table 6>과 같다.

결국 연구 모형에서 도출된 10개의 가설들에 대해서 <Table 7>과 같은 결과가 도출되었다.

### 5.3 다중회귀분석

앞의 연구모형 및 가설에서 제시된 종속변인과 독립변인에 대해서 인과관계를 검증하기 위해 회귀분석을 실시하였다. 공감수에 영향을 미치는 요인에 대해서 최종공감수, 공감수의 변곡점 도달에 소요되는 시간, 공감수의 변곡점에서의 공감수의 세가지를 종속변인으로 보았으며, 비공감수에 영향을 미치는 요인에 대해서도 최종 비공감수, 비공감수의 변곡점 도달에 소요되는 시간, 비공감수의 변곡점에서의 비공감수의 세가지를 종속변인으로 보아 이들 각각에 대해서 별도로 회귀분석을 실시하였다.

$$L_j = \alpha_j + \beta_{ij}X_i$$

단,  $L_j$ 에서  $j = 1$ (최종 공감수),  $2$ (변곡점 도달에

〈Table 6〉 Difference of Likes/Dislikes for Each Increasing Type

Type of likes/dislikes increasing	Likes (increasing-stagant)	Dislikes (increasing-stagant)
The number of likes	138.948***	-52.928***
The number of dislikes	22.217***	41.747***
Time to inflection point	53.604***	25.296***
Likes at inflection point	70.361***	36.387***
Likes at initial point	50.314***	36.680***
Dislikes at initial point	6.500***	9.838***
Message valence	-.042	-.082
Message length	.151	.212
Message intensity	.143	.171

〈Table 7〉 Summary of Hypotheses Testing Results

ID	Hypothesis	Results
H1-1	Increasing type has more number of likes than stagnant type.	Supported
H1-2	Increasing type has more number of dislikes than stagnant type.	Supported
H2-1	Message negativeness will be positively associated with the number of likes.	Supported
H2-2	Message negativeness will be positively associated with the number of dislikes.	Supported
H3-1	Message intensity will be positively associated with the number of likes.	Supported
H3-2	Message intensity will be positively associated with the number of dislikes.	Not supported
H4-1	Message length will be negatively associated with the number of likes.	Supported
H4-2	Message length will be negatively associated with the number of likes at the inflection point.	Supported
H5-1	The number of likes at initial time will be positively assoaited with the number of likes.	Supported
H5-2	The number of dislikes at initial time will be positively assoaited with the number of dislikes.	Supported

소요되는 시간, 3(변곡점에서의 공감수)를 의미하며  $X_i$ 에서  $i$ 는 각각 1(초기 공감수), 2(초기 비공감수), 3(언어강도), 4(댓글길이), 5(댓글의 긍부정성)를 가리킨다.

〈Table 8〉은 위의 여섯 가지 종속변인 각각에 대해 다중회귀 분석을 통해 추정된 결과이다. 먼저 첫 번째 열에서 최종 공감수에는 초기 공감수와 초기 비공감수가 긍정적인 영향을, 댓글의 긍부정성은 부정적인 영향을 주는 것으로 나타났다. 한편 공감수의 변곡점 도달에 소요되는 시간에는 초기 공감수와 초기 비공감수외에 언어의 강도가 긍

정적인 영향을 주며, 단어의 수는 부정적인 영향을 주고 있다. 그리고 공감수의 변곡점에서의 공감수에는 초기 공감수, 초기 비공감수가 긍정적인 영향을, 한편 댓글길이와 댓글의 긍부정성은 부정적인 영향을 주는 것으로 나타났다.

비공감수에 대해서는 먼저 최종 비공감수는 오직 초기 비공감수만이 긍정적인 영향을 주며 초기 공감수의 경우에는 증가할수록 최종 비공감수는 감소하는 것으로 나타났다. 비공감수의 변곡점 도달에 소요되는 시간의 경우 초기 비공감수는 긍정적인 영향을 주지만, 댓글의 긍부정성은 부정적인 영향을 준

〈Table 8〉 Effects of the Characteristics of Comment on WOM Response

Dependent Variable	Likes			Dislikes		
	The number of likes	Time to inflection point(likes)	The number of likes at inflection point	The number of dislikes	Time to inflection point(dislikes)	The number of dislikes at inflection point
Const.	10.206 (6.693)	62.531*** (3.403)	6.235 (5.427)	-3.497 (2.582)	34.007*** (3.190)	-2.067 (2.088)
Likes at initial time	1.598*** (.043)	.039* (.022)	1.539*** (.036)	-0.096*** (.016)	.031 (.021)	-.071*** (.014)
Dislikes at initial time	2.692*** (.324)	.435*** (.167)	2.662*** (.266)	3.708*** (.125)	.449*** (.156)	3.093*** (.102)
Message valence	-10.725*** (4.178)	.206 (1.898)	-6.339** (3.027)	-1.956 (1.612)	-5.406*** (1.779)	-1.345 (1.164)
Message intensity	3.528 (4.700)	4.136* (2.143)	5.479 (3.349)	0.542 (1.813)	.204 (2.262)	.774 (1.480)
Message length	-1.339 (1.318)	-1.377** (.677)	-1.830* (1.080)	-0.241 (.508)	-.202 (.635)	-.294 (.415)
R-Square	.916	.054	.880	.840	.062	.718
F-Value	409.844***	4.490***	570.419***	187.904***	5.172***	198.855***

Note) \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ , parentheses indicate standard deviations.

다. 그리고 비공감수의 변곡점에서의 비공감수는 초기비공감수와는 정의 비례관계가 있으나 초기 공감수와는 반비례하는 관계인 것으로 인식되었다.

변곡점 도달에 소요되는 시간이 길면 길수록 댓글을 보는 사람들이 더 활발하게 공감 또는 비공감했다는 의미이다. 분석 결과 초기 공감수, 초기 비공감수, 언어의 강도는 모두 변곡점 도달에 소요되는 시간과 정의 상관관계가 있는 것으로 나타났다. 또한 댓글의 댓글길이는 통계적으로 유의하게 도달시간에 반비례한다. 이는 가설에서 밝힌바 대로 정보 과잉 이론과 일치하는 결과이다. 그러나 댓글의 공부정성은 공감수의 변곡점 도달시간과 관련이 없다.

공감수의 변곡점에 도달하는 시간에 영향을 주는 요인을 확인 하기 위한 회귀분석에서는 초기 비공감수와 댓글길이가 유의하게 나타났으며( $p < 0.05$ ), 댓글길이는 부의 영향을 미치는 것으로 나타났는데( $\beta = -1.377$ ), 댓글을 읽는 독자가 짧은 시간 내에 글쓴이의 의도를 파악하여 공감할 수 있는 글이 인기가 많다는 것을 확인 할 수 있었다. 또한 초기비공감수가 유의미하게 나타났다( $p < 0.05$ ). 비공감

이라는 행위가 댓글이 가진 의견에 대해 부정적인 의견만을 표현하는 것에 그치는 것이 아니라 독자의 또 다른 관심의 표현으로 볼 수 있다. 즉, 댓글이 비공감이라는 관심까지 끌어낼 수 있을 만큼 쟁점이 되는 댓글이 됨과 동시에 그댓글은 다른 독자들까지 흥미를 유발 시킬 수 있는 댓글이 되기 때문이다. 그렇기 때문에 초기 비공감수가 유의미하게 나타난 것은 초기에 댓글이 갖는 관심에 따라서 변곡점에 도달하는 시점에 영향을 준다고 할 수 있다. 댓글의 공부정성이 공감이 변곡점에 도달하는데 많은 영향을 줄 것으로 기대했지만 회귀분석결과에서는 유의미하게 나타나지 않았다. 그 이유는 댓글이 갖고 있는 특징이 영향을 주는 것으로 생각된다. 댓글은 빠른 속도로 방대한양이 증가 된다. 댓글을 읽는 독자 입장에서는 많은 양의 댓글을 짧은 시간 내에 빠르게 읽기 때문에 댓글이 갖고 있는 감정보다는 직관적으로 알 수 있는 정보에 대해서 빠르게 반응 하는 것으로 추측 된다. 또한 언어강도 역시 댓글의 감성과 연관된 것으로 유사한 결과를 나타낸 것으로 보인다.

한편, 변곡점에서의 비공감수에 영향을 미치는 요인들을 살펴본 결과 초기 공감수, 초기 비공감수가 영향을 주는 것으로 나타났다. 하지만 재미있는 점은 초기 공감수가 부의 영향을 미치고 있다는 것이다. 즉, 초기 공감수가 증가하면 변곡점에서의 비공감수는 감소한다. 또한 초기에 발생된 공감의 수가 적을 수록 변곡점의 비공감수에 영향을 더 준다는 것이다. 하지만, 초기 비공감수는 정의 영향을 미치는 것으로 보았을 때 초기 공감수가 초기비공감수에 비해 상대적으로 적은 상태일 때 영향을 더 많이 주는 것으로 보여진다. 또한 결과가 위와 같이 나온 것은 댓글을 읽는 독자들의 심리적인 요소가 영향을 주고 있는 것으로 판단된다. 초기에 발생한 공감의 수보다 비공감의 수가 많을 때 독자들은 심리적으로 부정적인 의사를 갖게 되어 공감보다는 비공감을 누르게 되는 것으로 보여진다.

## 6. 토론 및 결론

### 6.1 시사점

본 논문의 의의는 텍스트마이닝 기법을 통해 효율적으로 댓글의 특성을 파악하여 댓글의 eWOM으로서의 성과인 공감수 및 비공감수에 미치는 영향 요인을 실증 규명한 것이다. 과거 연구는 사람을 통하여 수작업으로 내용분석의 과정을 거쳐 연구한 것으로 주관적인 요인이 개입되며, 또한 방대하고 지속적으로 게시, 변동되는 댓글의 공감수 및 비공감수 추이에 대한 시계열적 연구를 위해서는 많은 비용이 소요되었다.

본 논문에서 제안한 연구모형에 대한 검증을 통해 다음과 같은 이론적, 실무적 시사점을 얻을 수 있었다. 먼저 댓글의 길이의 관점에서 보면 댓글의 길이는 공감부분에서 변곡점 도달시간과 변곡점에서의 공감수와 반비례의 관계를 가지는 것으로 나타났다으며, 그 외에는 통계적으로 유의한 영향을 주지 않는 것으로 나타났다. 즉, 댓글을 읽는 독자가 짧은 시간 내에 글쓴이의 의도를 파악하여 공감할

수 있는 글이 인기가 많다는 것을 확인 할 수 있었다. 한편 댓글은 빠른 속도로 방대한양이 증가하기 때문에, 댓글을 읽는 독자 입장에서는 많은 양의 댓글을 짧은 시간 내에 빠르게 읽기 때문에 댓글이 갖고 있는 감성보다는 직관적으로 알 수 있는 정보에 대해서 빠르게 반응 하는 것으로 추측 된다. 이는 Amazon.com 에서의 리뷰의 길이가 해당 도서에 대한 판매 성과에 긍정적인 영향을 주는 것과는 다른 결과이나, bn.com에서는 리뷰의 길이와 성과가 부정적이면서 통계적으로 유의하지 않은 것과는 비교적 유사한 결과이다(Chevalier and Mayzlin, 2006). Chevalier and Mayzlin(2006)가 밝힌대로 리뷰의 길이는 그 리뷰를 올린 사람의 열정을 반영한 것으로 볼 수 있으나 상대적으로 긴 리뷰에는 평가 대상(책 또는 기사)에 대한 긍정적 글과 부정적 글을 모두 실게 되기 때문에 혼합적인 결과가 나올 수 있다. 그런데 기사의 경우 댓글은 대부분 일관되게 찬성 혹은 반대의 의사를 밝히는 경우가 대부분이며, 유용한 정보를 전달하는 행위보다는 자신의 공부정적 의견을 피력하는 경우가 많기 때문에 긴 댓글은 반복적으로 의사를 표현하는 것이므로 공감을 얻기가 어려운 것으로 보인다. 실제로 공감수와 비공감수 모두에 대해서 통계적 유의미성을 막론하고 모두 댓글의 길이는 공감수나 비공감수 모두에게 부정적인 영향을 주는 것으로 나타났다.

또한 Chevalier and Mayzlin(2006)의 연구에서는 효과적으로 통제하지 못한 게시자의 열정은 본 논문에서는 언어강도와 댓글의 공부정성에서 반영된다. 먼저 댓글의 공부정성은 대체로 공감수나 비공감수에 부정적인 영향을 주는 것이 흥미롭다. 특히 최종 공감수, 공감수의 변곡점에서의 공감수, 그리고 비공감수의 변곡점 도달에 소요되는 시간에는 통계적으로 유의미하게 부정적인 영향을 준다. 구체적으로 보면, 공감수 관점에서 보면 긍정적인 댓글보다는 부정적인 댓글이 더 많은 공감을 더 빨리 이끌어내고 있다. 이로 보건대 부정적인 댓글에 대한 공감은 중간 시점이나 최종 시점을 가리지 않고 지속적으로

이루어지는 것으로 보인다. 이는 부정적인 메시지에 대해 더 강렬하게 기억하고 반응하는 부정적 편향이 댓글에서도 마찬가지로 나타나는 것을 보이고 있는 것이다(Park and Lee, 2009). 즉, 기사에 대한 댓글들은 대부분 부정적인 의견이나, 조롱 등 부정적인 성격을 가지며, 기사와 댓글을 읽는 독자들에게는 자극적이고 선정적인 댓글이 비교적 흥미를 유발한다고 볼 수 있다. 한편 변곡점 도달 시간은 댓글의 긍부정성이 영향을 주지 못한다.

반면 비공감수의 경우에는 댓글이 긍정적이든 부정적이든 비공감수에는 별 영향을 주지 못하고 있으며 도리어 비공감수 변곡점 도달 시간에는 부정적 댓글이 더 길다. 즉, 댓글이 긍정적일수록 비공감수는 더욱 빨리 정체된다. 결국 댓글에 표출된 감성의 긍부정성은 공감수에 영향을 주나, 비공감수에는 영향을 주지 못하며, 이는 특정 댓글의 댓글의 긍부정성에 대해서는 공감과 비공감이 비대칭적 반응을 보여주고 있는, 즉 조건부 부정적 편향(conditional negative bias)의 존재를 말하고 있다. 즉, 긍정적 댓글과는 달리, 부정적 댓글에 대해서 공감하는 수가 증가하는 양상을 보면 비공감의 사표현을 잘 하려고 하지 않는다 라고 이야기할 수 있다. 이는 첫째, negative bias를 가지고 부정적 의사를 가진 집단에 대해서 갈등을 조장하는 반대 의견을 강하게 피력하지 않으려는 즉, 대인관계 갈등에서의 회피적 갈등대처 방식이 작동되었다고 볼 수 있다(Folkman, 1980). 특히 온라인 의사소통의 경우 갈등 해소가 개인이 통제할 수 없는 경우에는 회피적인 대처를 하게 된다는 주장도 있다(Park et al., 2004). 왜냐하면 본인의 의사에 반하는 댓글과 그에 대한 타인의 공감에 적극적으로 비공감을 피력할 정도로 자신의 이해관계에 해를 입히지는 않기 때문이다. 둘째로는 온라인 기사에 대한 냉소적 반대자(cynical opponent)에 대한 동조 행동이라고 해석할 수도 있다. Cho and Hong(2009)에 의하면 기업의 CSR 활동에 대해 냉소적 반대자들은 대체로 부정적인 댓글을 달며 이는 그렇지 않은 댓글에 비해 더 많은 공감을 이끌어낸다는 연구와

일맥상통한다.

댓글 작성자의 열의를 보여주는 또 다른 요인인 언어강도에 대해서는 공감수나 비공감수 모두 정비례하는 것으로 나타났으나 대부분 통계적으로 유의하지는 않았다. 이는 상당수 댓글들이 본인의 주장을 강조하려고 은어나 비언어적 자극(이모티콘, 느낌표, 대문자로 표현 등)을 포함하게 되는데 이것이 열정에 대한 표출보다는 그러한 댓글을 작성하는 사람에 대한 신뢰나 동질의식을 떨어뜨려 강한 공감을 자아내지 못하기 때문인 것으로 보인다. 이는 댓글의 품질과도 연결되는데, 반말이나 욕설, 인터넷 용어가 포함된 글이 높은 관심을 받지 못하고 있음을 보이고 있다.

초기공감수는 최종 공감수와 매우 밀접한 정의 관계를 가지고 있다. 본 연구에서 초기공감수는 댓글 게시 후 10분이 경과한 시점의 공감수인 바 댓글 게시 시점으로부터 10분 정도가 경과하면 그 댓글이 최종적으로 얼마나 많은 공감수를 획득하게 될 지, 얼마나 오랫동안 지속적으로 공감을 얻을지, 그리고 변곡점에 다다르기 전에 얼마나 많은 공감을 확보할지를 모두 잘 설명하는 요인이다. 즉, 댓글에 대한 성과에 가장 일관되고 중요한 변수인 것이다. 더욱이 초기 공감수는 비공감수를 저하시키는 데에도 효과가 있는 것으로 나타났다. 이를 통해 특정 기사에 대해서 10분 정도의 공감 활동을 통해서도 기업 관련 기사에 대한 일반인들의 eWOM을 상당부분 통제할 수 있는 것을 실증하였다.

또한 초기 비공감수의 경우 최종 비공감수에 영향을 주는 것으로 나타났다. 이를 통해 기업 등이 자신의 기업과 관련된 기사에 대해서 일반인들의 댓글의 비공감수 추이를 오랫동안 기다리지 않고 10분 경과 후 살펴보는 것만으로도 파악이 가능하며, 초기에 대응(위기관리 등)이 가능함을 실증하였다. 흥미로운 사실은 초기 비공감수가 공감수와도 비례적인 관계를 보인다는 점이다. 이는 어느 정도 상식에 배치된 것처럼 보이나 많은 댓글이 공감수와 비공감수가 같이 증가하는 격론형임을 미루어볼 때 (<Table 4> 참조), 이는 격론형 댓글의 영향 때문

인 것으로 파악된다. 또한 회의적 반대자(Cynical opponent)들은 긍정형 댓글에 대해서도 부정형 반응을 펼쳐서 비공감수를 늘리는 때문인 것으로도 추측할 수 있다. 실제로 일부 온라인 뉴스 사이트는 댓글에 다시 댓글을 달 수 있는 기능을 제공하고 있는데, 이때 보면 적지 않은 수가 공감형 댓글에 대한 반대의 댓글인 것을 볼 수 있다.

본 논문에서의 실증 분석 결과는 기업에게 유용할 것이다. 첫째, 새로운 eWOM 매체로서 댓글을 활용할 수 있을 것이다. 본 논문의 결과대로 공감 및 비공감을 유발하는 요인들이 있음을 염두에 두어 기업이 일반인의 댓글이 기업에 미칠 영향을 초기에 파악하여 기업의 대외 관계 관리에 보다 더 기민하게 반응할 수 있을 것이다. 두 번째로, 아직도 많은 기업들이 기사 및 댓글로 인한 기업 이미지의 위기에 대해서 별다른 대책을 마련하고 있지 않거나 매우 느리게 반응한다. 본 연구결과는 댓글 게시 초기 시점에 댓글이 미칠 긍정적 혹은 부정적 영향을 예측할 수 있다는 점에서 유용할 것이다. 셋째, 댓글에 대한 공감 활동은 사회적 규범이나 행동양식을 대다수의 의견에 따라 순응해 나가려는 순응 행위로 설명되는데, 이에 의하면 부정적 댓글에 공감하는 독자들은 해당 기업에 대한 신뢰도와 해당 제품 및 브랜드에 대한 구매의도도 떨어지는 경향을 보인다는 연구(Kim, 2012)와 일관된 결과를 보이고 있으므로, 기업 입장에서는 부정적 댓글에 대한 섬세한 관리가 필요할 것이다.

이상의 연구 시사점을 통해서 리뷰(review)가 eWOM에서 일반적으로 대상으로 바라보는 것과 비교하여, 본 논문은 댓글(comment)도 리뷰처럼 eWOM의 요소가 될 수 있는지를 실증한 것이다. 이를 위해 분석하는 댓글의 기사도 IT기업 또는 IT제품이나 서비스에 대한 것으로 한정되었다.

## 6.2 결론

기존의 연구에서는 뉴스나 뉴스의 댓글과 같은 비정형화된 자료에 대한 데이터 분석을 위해 내용 분석 등 사람에 의한 질적 분석을 수행했다. 이러

한 이유로 비록 내용분석에 대한 가이드라인 등이 존재하나(Krippendorff, 1989; Landis and Koch, 1997) 이것은 내용분석 수행자(coder)들의 주관성을 완전히 배제하기 어렵고 분석에 소요되는 시간 및 비용이 큰 방법이었다.

이에 비해 본 연구에서는 댓글 공감수 및 비공감수의 유형과 댓글에 대한 초기 공감수 및 비공감수, 댓글의 길이와 언어 강도 등 여러 댓글의 특징들이 실제로 댓글에 대한 최종 공감수 및 비공감수에 유의한 영향을 미친다는 사실을 실증하고, 실증 결과를 활용하여 보다 더 객관적이고 비용효율적이며 기민한 댓글이라는 형태의 비정형데이터 분석 시스템을 개발할 수 있음을 보였다.

본 연구는 몇 가지 한계점을 가지고 있다. 먼저 특정 뉴스사이트만을 대상으로 댓글을 수집했다. 물론 한국에서는 가장 대표적인 뉴스 사이트이며 각 사이트별로 댓글의 기능이 차이가 있기 때문에 연구의 일관성을 유지하기 위한 방법이었다. 하지만 여기서의 결과를 일반화하기에는 주의가 필요하다. 또한 댓글의 길이와 최종 공감수 사이의 상관관계에서 댓글이 너무 짧은 경우에서의 상관관계의 변경 가능성이 있는데 본 연구에서는 다루지 않았으며 후추 보완해야 할 이유이다. 향후에는 각 사이트별로 댓글의 기능의 차이도 고려하여 각 기능이 어떠한 역할을 하는지를 규명하고, 그것이 각 뉴스 사이트의 활용도와 가지는 인과관계를 살펴본다면 뉴스 사이트에 대한 평가 방법도 개발할 수 있을 것이다. 또한 보다 더 다양한 댓글을 수집하고 댓글의 공감수 및 비공감수를 예측하는 전산화된 분석 시스템을 구현해야 할 것이다.

## References

- Bae, J.Y., "The Influence of Online News Replies on Netizen' CSR Perception, Attitude, Purchase Intention, and Word of Mouth Intention", *The Korean Journal of Advertising*, Vol.20, No.5, 2009, 7-37.

- (배지양, “넷글의 품질 및 방향성이 영리기업의 사회공헌활동에 대한 공중의 책임성인식, 태도, 구매의도, 구전활동의도에 미치는 영향”, *광고학연구*, 제20권, 제5호, 2009, 7-37.)
- Baumeister, R.F., E. Bratslavsky, C. Finkenauer, and K.D. Vohs, “Bad Is Stronger Than Good”, *Review of General Psychology*, Vol.5 No.4, 2001, 323-70.
- Bickart, B. and R.M. Schindler, “Internet Forums as Influential Sources of Consumer Information”, *Journal of Interactive Marketing*, Vol.15 No.3, 2001, 31-40.
- Blom, F., “Information Overload and the Growing Infosphere : a Comparison of the Opinions and Experiences of Information Specialists and General Academics on the Topic of Information Overload”, 2011. Available at <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A438746&dsid=-867>(Accessed May 18. 2015).
- Chen, H.L., “Identifying Factors of Online News Comments”, *Proceedings of the American Society for Information Science and Technology*, Vol.49, No.1, 2012, 1-4.
- Cheung, M.Y., C. Luo, C.L. Sia, and H. Chen, “Credibility of Electronic Word-of-Mouth : Informational and Normative Determinants of On-Line Consumer Recommendations”, *International Journal of Electronic Commerce*, Vol.13, No.4, 2009, 9-38.
- Chevalier, J.A. and D. Mayzlin, “The Effect of Word of Mouth on Sales : Online Book Reviews”, *Journal of Marketing Research*, Vol.43, No.3, 2006, 345-354.
- Cho, S.Y. and Y.S. Hong, “Netizen’s Evaluations of Corporate Social Responsibility : Content analysis of CSR News Stories and Online Reader’s Comments”, *Public Relations Review*, Vol.35, No.2, 2009, 147-149.
- Christodoulides, G., N. Michaelidou, and E. Argyriou, “Cross-National Differences in eWOM Influence”, *European Journal of Marketing*, Vol.46, No.11, 2012, 1689-1707.
- De Matos, C.A. and C.A.V. Rossi, “Word-of-Mouth Communications in Marketing : A Meta-Analytic Review of the Antecedents and Moderators”, *Journal of the Academy of Marketing Science*, Vol.36, No.4, 2008, 578-96.
- Fan, W. and S. Sun, “Sentiment Classification for Online Comments on Chinese News”, *In Computer Application and System Modeling (ICCSM), 2010 International Conference*, Vol.4, 2010, 4-740.
- Fiske, S.T., “Attention and Weight in Person Perception : The Impact of Negative and Extreme Information”, *Journal of Personality and Social Psychology*, Vol.38, No.6, 1980, 889-906.
- Folkman, S. and R.S. Lazarus, “An Analysis of Coping in a Middle-Aged Community Sample”, *Journal of Health and Social Behavior*, Vol.21, No.3, 1980, 219-239.
- Gupta, P. and J. Harris, “How eWOM Recommendations Influence Product Consideration and Quality of Choice : A Motivation to Process Information Perspective”, *Journal of Business Research*, Vol.63, No.9, 2010, 1041-1049.
- Hennig-Thurau, T., K. Gwinner, G. Walsh, and D. Gremler, “Electronic Word-of-Mouth Via Consumer Opinion Platforms : What Motivates Consumers to Articulate Themselves on the Internet?”, *Journal of Interactive Marketing*, Vol.18, No.1, 2004, 38-52.
- Katz, E. and P.F. Lazarsfeld, *The Part Played by*

- People in the Flow of Mass Communications*, Personal Influence, Transaction Publishers, 1955.
- Kim, E.M. and Y.H. Sun, "The Effect of Replies in Internet News on the Audience", *Korean Society For Journalism And Communication Studies*, Vol.50, No.4, 2006, 33-64.
- (김은미, 선유화, "댓글에 대한 노출이 뉴스 수용에 미치는 효과", *한국언론학보*, 제50권, 제4호, 2006, 33-64.)
- Kim, S.D., S.Y. Kim, and H.G. Cho, "A Model for Popularity Dynamics to Predict Hot Articles in Discussion Blog", *In Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication*, Vol.10, 2012, 1-10.
- Krippendorff, K., "Content Analysis", *International Encyclopedia of Communication*, Vol.1, 1989, 403-407.
- Landis, J.R. and G.G. Koch, "The Measurement of Observer Agreement for Categorical Data", *Biometrics*, Vol.33, No.1, 1977, 159-174.
- Lau, G. and S. Ng, "Individual and Situational Factors Influencing Word-of-Mouth Behaviour", *Canadian Journal of Administrative Sciences*, Vol.18 No.3, 2001, 163-178.
- Maynard, D.W., "Comment : Bad News and Good News : Losing vs. Finding the Phenomenon in Legal Settings", *Law and Social Inquiry*, Vol.31, No.2, 2006, 477-497.
- Moon, K.S., S. Kim, and S.J. Oah, "An Effect of the Valence of Best Reply on the Conformity of General Reply", *The Journal of the Korea Contents Association*, Vol.13, No.12, 2013, 201-211.
- (문광수, 김 슬, 오세진, "베스트 댓글의 방향성이 일반댓글의 동조효과에 미치는 영향", *한국콘텐츠학회논문지*, 제13권, 제12호, 2013, 201-211.)
- Newman, J. and R. Staelin, "Prepurchase Information Seeking for New Cars and Major Household Appliances", *Journal of Marketing Research*, Vol.9 No.3, 1972, 249-57.
- O'Connor, B., R. Balasubramanyan, B.R. Roulledge, and N.A. Smith, "From Tweets to Polls : Linking Text Sentiment to Public Opinion Time Series", *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, 2010, 122-129.
- Park, C.L., S. Armeli, and H. Tennen, "Appraisal-Coping Goodness of Fit : A Daily Internet Study", *Personality and Social Psychology Bulletin*, Vol.30, No.5, 2004, 558-569.
- Park, C. and T.M. Lee., "Information Direction, Website Reputation and eWOM Effect : A Moderating Role of Product Type", *Journal of Business Research*, Vol.62, No.1, 2009, 61-67.
- Phelps, J., R. Lewis, L. Mobilio, D. Perry, and N. Raman, "Viral Marketing or Electronic Word-of-Mouth Advertising : Examining Consumer Responses and Motivations to Pass Along E-Mail", *Journal of Advertising Research*, Vol.44, No.4, 2004, 333-348.
- Richins, M., "Negative Word of Mouth by Dissatisfied Consumers : A Pilot Study", *Journal of Marketing*, Vol.47, No.1, 1983, 63-79.
- Rozin, P. and E.B. Royzman, "Negativity Bias, Negativity Dominance, and Contagion", *Personality and Social Psychology Review*, Vol.5, No.4, 2001, 296-320.
- Saveski, M., D. Quercia, and A. Mantrach, "The Geography of Online News Engagement", *In Social Informatics*, Springer International Publishing, 2014, 279-289.
- Tatar, A., J. Leguay, P. Antoniadis, A. Limbourg, M.D. de Amorim, and S. Fdida, "Predicting



- the Popularity of Online Articles Based on User Comments”, *In Proceedings of the ACM International Conference on Web Intelligence, Mining and Semantics*, 2011, 67.
- Thelwall, M., “No Place for News in Social Network Web Sites?”, *Online Information Review*, Vol.32, No.6, 2008, 726-744.
- Westbrook, R.A., “Product/Consumption Based Affective Responses and Postpurchase Processes”, *Journal of Marketing Research*, Vol.24, No.3, 1987, 258-70.
- Yang, H.S., “The Effects of the Opinion and Quality of User Posting on Internet News Readers’ Attitude Toward the News Issue”, *Korean Society For Journalism And Communication Studies*, Vol.52, No.2, 2008, 254-281.
- (양혜승, “인터넷 뉴스 댓글의 견해와 품질이 독자들의 이슈에 대한 태도에 미치는 영향”, *한국언론학보*, 제52권, 제2호, 2008, 254-281.)

## ◆ About the Authors ◆



**Jeonghun Kim (adsky0719@khu.ac.kr)**

Jeonghun Kim is graduated from Mokwon University, Department of Information consulting. He is currently in Big data management master's degree program. Currently, his research focuses on text mining and big data, and CRM.



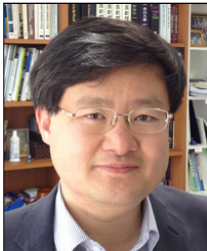
**Yeongeun Song (dudeun13@khu.ac.kr)**

Yeongeun Song is taking a master course in the Department of Management at Kyung Hee University. She received a degree in information consulting from Mokwon University in 2014. Currently, She is working at CaiTech laboratory as a researcher and taking part in "happiness index project." Her interests include text mining, big data analytics, Sentimental Analysis.



**Yunseon Jin (dudnrha@khu.ac.kr)**

Yoonsun Jin has completed a master degree major in e-business from Sookmyung Women's University. She is currently taking a doctor degree major in big data management at Kyung Hee University. She also has been employed in web sustainability institute as senior researcher and interested in are such big data, e-business and public data management.



**Ohbyung Kwon (obkwon@khu.ac.kr)**

Ohbyung Kwon is a professor in the School of Management at Kyung Hee University. He received his master's and Ph.D. degrees in MIS from Korea Advanced Institute of Science and Technology (KAIST), and a B.A. from Seoul National University. He published several articles in Journal of MIS, International Journal of Information Management, Decision Support Systems, etc. His research interests include ubiquitous computing, big data analytics, and decision support systems.