

# Sentiment Expression on SNS According to Gender and User Activity Level

Phil-Sik Jang

Sehan University, Department of Air Traffic and Logistics, Choongnam, 31746

## 성별, 사용자 활성 수준에 따른 SNS 감성표현 차이 연구

장 필 식

세한대학교 항공교통물류학과

### Corresponding Author

Phil-Sik Jang

Sehan University, Department of Air

Traffic and Logistics, Choongnam, 31746

Mobile: +82-10-9725-9707

Email : philsjang@gmail.com

Received : March 15, 2019

Revised : March 30, 2019

Accepted : March 31, 2019

**Objective:** The primary aim of this study was to examine the sentiment expression according to gender and user activity level with a particular focus on the social networking site of Twitter.

**Background:** With the vast and ever-growing popularity of social networking services, sentiment analysis on SNS has become widely used in data science for analysis of user's feedback and opinion mining.

**Method:** The messages posted from 13.4 million twitter users (6.9 million male users) for 11 months were analyzed and the message polarity, sentiment strength and the type (evaluative/factual) of messages were explored.

**Results:** The results showed that the average sentiment strength of female users was significantly ( $p < 0.01$ ) higher than that of male users and the average sentiment strength of highly active male users was lowest among groups. It also showed that the highly active male users tend to post more negative messages than other groups and female users express significantly more positive sentiment than male users for both highly and less active user levels.

**Conclusion:** It implies that the gender and activity level differences of users should be considered in sentiment analysis on social media data to gain valuable insights.

**Application:** The results of this study could be used as quantitative, empirical data and design guidelines for developing sentiment analysis and opinion mining systems on SNS.

**Keywords:** Gender difference, User activity level, SNS, Sentiment expression, Message polarity

Copyright©2019 by Ergonomics Society of Korea. All right reserved.

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

월드와이드웹 기반으로 활동과 관심을 공유하는 사람들 간의 관계망을 구축해 주는 서비스인 SNS (Social Network Service)는 새로운 구전(word of mouth) 커뮤니케이션 수단으로써 강력한 위력을 발휘하고 있다. 최근 다양한 분야에서 이러한 새로운 쌍방향 커뮤니

니케이션을 이용하고자 하는 시도들이 계속되고 있는데(Chu and Kim, 2011; Lee and Jang, 2013), 이러한 시도들 중 최근 주목받고 있는 분야가 오피니언 마이닝(opinion mining)으로도 불리는 감성분석(sentiment analysis)이다(Ravi and Ravi, 2015; Lee and Jang, 2017). 현재, 감성분석은 주로 자연어처리 및 의미분석과 연관되어 분석의 정확도를 향상시키기 위한 기술개발에 초점이 맞추어져 있다(Jang, 2018). 하지만 감성분석 결과가 진정한 통찰력(insight)을 가지려면, 데이터에 대한 기계적인 분석뿐만 아니라, 분석 데이터 즉, 실제 감성이 포함된 메시지들을 작성하고 주고받는 사용자들의 특성과 감성표현 양상에 대한 연구가 병행되어야 한다.

최근, 사용자의 언어 및 문화적 특성에 따른 감성표현 양상에 대한 연구(Kim and Jang, 2016)와 십대 소녀들의 과도한 SNS(Facebook, Twitter 등) 활동 및 그에 따른 부정적 영향 등에 대한 연구(Tiggemann and Slater, 2013)가 발표되었으나, SNS 사용자들의 성별, 활동 정도와 감성표현의 양상, 특성 등의 연관관계에 대한 체계적 연구는 찾기 힘들다. 본 연구에서는 대표적인 소셜미디어인 트위터 사용자들을 대상으로, 사용자 성별, 활동(활성: user activity) 수준에 따른 메시지 감성표현 양상과 특성을 정량적으로 분석하였다.

## 2. Background

### 2.1 Message polarity

온라인 환경에서는 구전이 오프라인과 다르게 텍스트 형태로 전달되기 때문에 메시지 특성이 중요한 역할을 한다(Park and Kim, 2008; Lee and Jang, 2013). 구전은 긍정적인 정보뿐만 아니라 부정적이고 상세한 정보까지 전달될 수 있는데 이러한 감성적 특성을 정보의 방향성(polarity, orientation)이라 부르며, 일반적으로 부정, 긍정, 중립으로 구분한다(Lima et al., 2015). 현재 대부분의 감성분석 알고리즘에서 이 세 가지 방향성 분류가 많이 사용되고 있다. 메시지의 방향성 효과에 대한 관련 연구들(Baumeister et al., 2001; Chiou and Cheng, 2003)에서는 긍정적 감정, 정보보다 부정적 감정과 정보가 더 높은 가중치를 가지며, 더 큰 영향을 미친다고 주장하는데, 이러한 영향을 부정 편향성(negativity bias) 또는 부정 효과(negativity effect)라 칭한다. 하지만, 온라인 구전에서 부정적 구전과 긍정적 구전 간 영향력의 유의한 차이가 없거나 반대로 긍정적 구전 효과가 더 크다는 연구결과들 또한 발표되고 있다(Doh and Hwang, 2009; Xue and Zhou, 2010). 이와 같이 온라인 구전의 방향성을 다룬 연구결과들이 일관성 없는 결과들을 보여주고 있지만, 메시지 내에 포함된 감성, 감정의 방향성이 구전에 유의한 영향을 미친다는 것이 상기 연구들에서 공통적으로 확인된다.

이와는 별개로, 구전 메시지 및 정보 유형에 대한 분류 방법들은 연구에 따라 차이를 보이는데, 주관적 속성/사실적 속성으로 분류하기도 하며(Darley and Smith, 1993), 정보 및 메시지가 객관적으로 입증가능한가(verifiability)에 따라 평가적(evaluative)/사실적(factual) 메시지로 나누기도 한다(Edell and Staelin, 1983; Lee and Jang, 2013). 본 연구에서는 감성분석에 많이 이용되며, 다양한 연구들(Edell and Staelin, 1983; Lee and Jang, 2013)에서 사용된 분류 방법을 차용하여, SNS 정보 유형을 평가적 메시지와 사실적 메시지로 구분하였다.

### 2.2 Gender and activity level of user

SNS 활용에 있어 성별 차이를 다룬 최근 연구들에 따르면, 여성 사용자들이 남성들에 비해 특정 SNS들을 더 편중하여 사용하며, 페이스북에 더 열성적으로 사진을 업로드하는 것으로 나타났다(Pew Research Center, 2015). 또한 여성들이 남성들에 비해 SNS 상에 가족, 친구, 연인관계에 대한 이야기들을 더 많이 포스팅하며(Jones et al., 2008), 남성들이 페이스북을 새로운 인간관계를 형성하는 의도로 주로 이용하는 반면, 여성들은 관계 유지를 위해 이용하는 것으로 파악되었다(Muscanell and Guadagno, 2012).

이와 함께 소셜미디어 데이터를 포함한 빅데이터 분석, 활용 연구(Lee and Jang, 2017)에 있어 고려해야 할 요인으로 새롭게 주목받고 있는 것이 사용자 활성(user activity) 수준이다(Jang, 2018; Tedjamulia et al., 2005). 트윗 패턴 등이 사용자 활성정도에 따라 구분된 집단 간에 많은 차이를 보인다는 연구결과(Tedjamulia et al., 2005)에 따라, 소셜 미디어 데이터 분석 시 이를 고려하는 연구들이 최근 발표되고 있다(Jang, 2018). SNS에서는 소수의 고활성 사용자들(highly active users)이 작성한 데이터가 전체 데이터의 상당 부분을 차지하는 멱법칙(power law)을 따른다(Anderson, 2006). 따라서 사용자 활성도에 따라 사용자를 분류할 때, 1:9 또는 1:90의 비율을 이용하게 된다. 본 연구에서는 사용자 활성 수준에 따라 사용자 집단을 1:9 비율로 분류하고 이들 사용자 집단과 성별, 메시지 방향성과 유형 간의 연관성을 분석하였다.

### 3. Method

본 연구에서는 대표적인 SNS 중 하나이며, 데이터 취합이 비교적 용이한 트위터를 대상으로 11개월(2016년 10월~2017년 8월) 동안 작성된 트윗을 실시간 수집하였으며, 이를 바탕으로 13,426,304명의 트위터 이용자들을 분석하였다. 데이터 수집에는 트위터에서 제공하는 Streaming API (Twitter.com, 2019)를 이용하였으며, R과 MonetDB (MonetDB, 2019)를 활용하여 데이터를 분석, 저장하였다. 감성분석에는 비교적 높은 정확도와 빠른 분석 속도로 최근 다양한 연구(Thelwall et al., 2012)에 이용되고 있는 SentiStrength (SentiStrength, 2019)를 R과 연동하여 이용하였다.

취합된 사용자들의 성별 판단에는 R의 gender package (Lincoln, 2018)를 활용하였으며, 트위터 메시지 게재 수 상위 10% 활성 사용자들을 고효성 사용자(highly active user)로, 하위 90% 사용자들을 저활성 사용자(less active user)로 구분하여 분석하였다. 메시지 방향성(polarity) 판정을 위해서는 Rsentiment package (Subhasree and Saptarsi, 2018)를 이용하여 각 메시지들을 긍정(positive), 중립(neutral), 부정(negative)으로 분류하였다.

### 4. Results

#### 4.1 Twitter activity

Table 1은 취합된 트윗 사용자를 성별, 사용자 활성 수준 별로 분류한 것이다. 사용자들 중 남성은 약 6백8십만 명, 여성은 6백5십만 명인 것으로 집계되었다.

**Table 1.** A summary table of users

	Highly active	Less active	Sum
Male	764,446	6,109,029	6,873,475
Female	578,224	5,974,605	6,552,829
Sum	1,342,670	12,083,634	13,426,304

사용자들의 성별, 활성정도에 따른 사용자 군을 독립변수로, 트윗과 리트윗을 포함한 포스팅 활동 횟수(statuses count)를 종속변수로 시행한 robust ANOVA 결과는 Table 2와 같다. 분석결과 유의수준 0.01에서 교호작용이 통계적으로 유의한 것으로 나타났는데, Figure 1은 이들 교호작용을 그래프로 표시한 것이다. 그래프에서 알파벳 a~d는 절사평균(trimmed mean)의 쌍별 비교(pairwise comparison)를 이용한 사후 검정결과를 나타내며, 동일한 알파벳은 유의수준 0.01에서 통계적으로 유의한 차이가 없음을 의미한다.

**Table 2.** Robust ANOVA summary table (Dependent variable: statuses count)

Source	Statistic	Sig.
Gender	19,791.9	0.001**
User activity	8,717,318.0	0.001**
Gender×User activity	27,494.23	0.001**

\*\* $p < .01$

Figure 1에서 볼 수 있는 것처럼, 하위 90% 저활성 사용자 군에서는 큰 차이가 관찰되지 않았지만 상위 10% 고효성 사용자 군에서 남

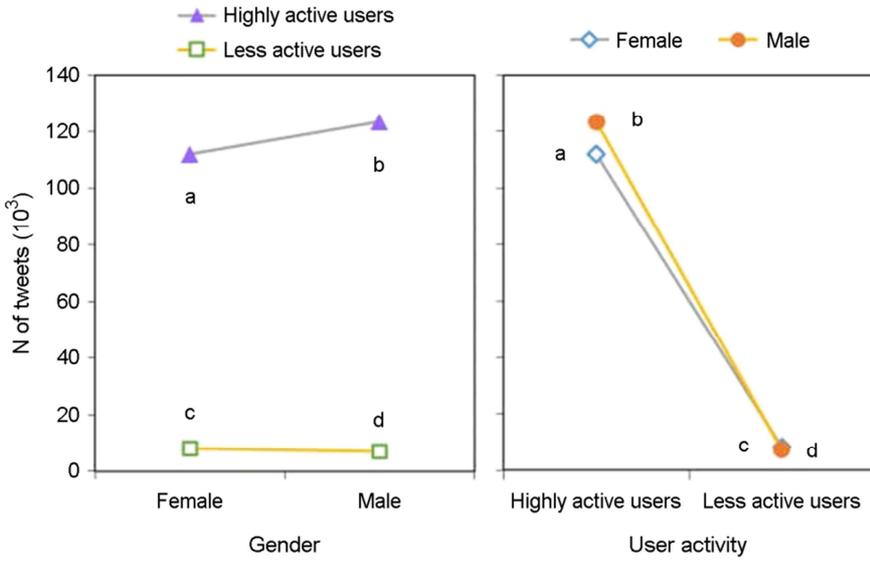


Figure 1. Interaction plot involving the trimmed means of the statuses count

Table 3. Robust ANOVA summary table (Dependent variable: Number of characters in tweet)

Source	Statistic	Sig.
Gender	1,424.4	0.001**
User activity	155,980.5	0.001**
Gender×User activity	973.2	0.001**

\*\* $p < .01$

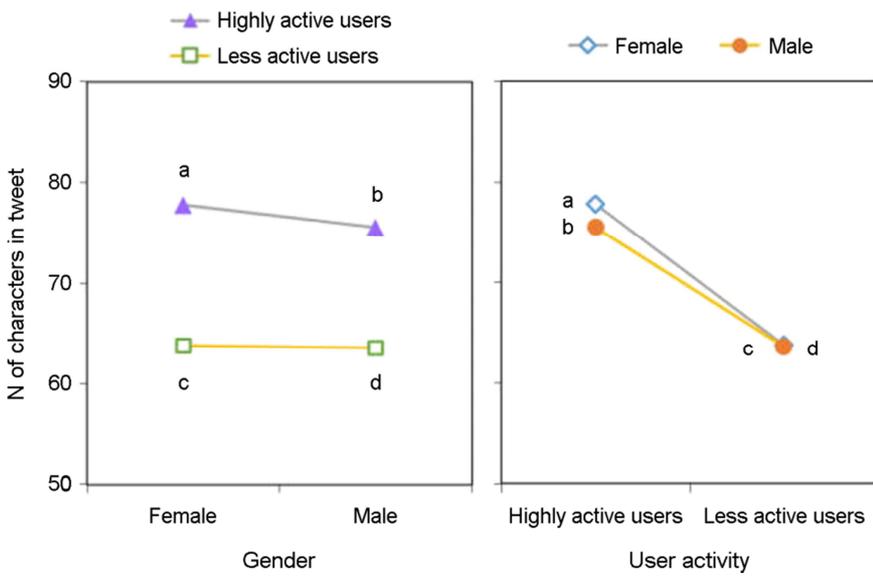


Figure 2. Interaction plot involving the trimmed means of the number of characters in tweet

성 사용자들의 포스팅 활동이 여성 사용자들보다 많음을 확인할 수 있다. Table 3은 사용자가 작성한 트윗 메시지의 문자 개수를 종속 변수로 하고, 성별, 사용자 활성 수준을 독립변수로 시행한 robust ANOVA 결과를 나타낸 것이다.

분석결과 유의수준 0.01에서 통계적으로 유의한 교호작용이 관찰되었으며, 이들 교호작용을 그래프로 표시하면 Figure 2와 같다. 상위 10% 고탈성 사용자들이 작성한 문자 수가 하위 사용자들 보다 많은 것으로 나타났으며, 이 상위 10% 고탈성 사용자 군에서는 여성 사용자들의 문자 수가 남성 사용자들의 문자 수보다 유의하게 많은 것으로 파악되었다.

## 4.2 Sentiment expression

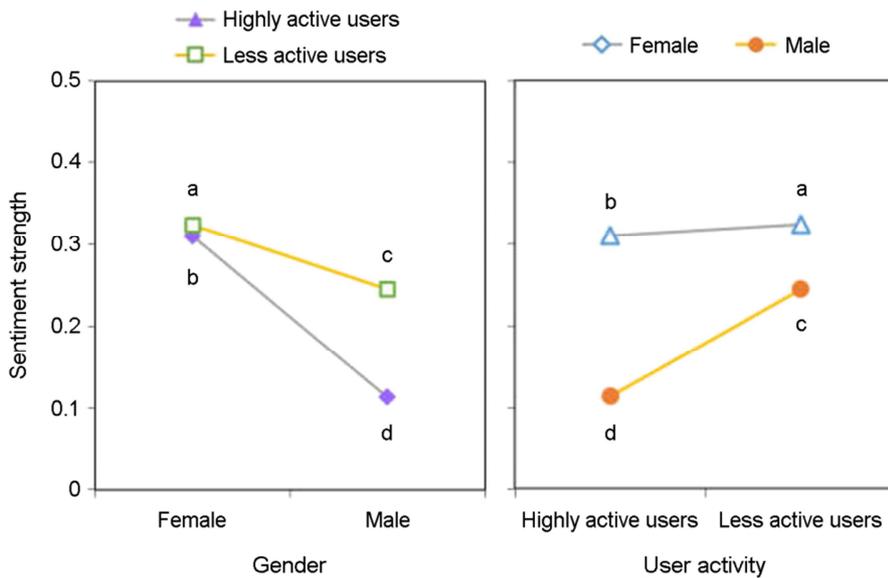
### 4.2.1 Sentiment strength

메시지의 감성강도를 종속변수로 하고, 성별과 사용자 활성 수준을 독립변수로 하여 robust ANOVA를 시행한 결과는 Table 4와 같다. 분석결과, 두 개 독립변수의 교호작용이 유의수준 0.01에서 통계적으로 유의한 것으로 확인되었으며, 이들 2인자 교호작용을 그래프로 나타내면 Figure 3과 같다.

**Table 4.** Robust ANOVA summary table (Dependent variable: sentiment strength)

Source	Statistic	Sig.
Gender	79,423.94	0.001**
User activity	22,150.49	0.001**
Gender×User activity	14,361.74	0.001**

\*\* $p < .01$



**Figure 3.** Interaction plot involving the trimmed means of the sentiment strength

절사평균의 쌍별 비교를 이용한 사후 검정결과, 두개 사용자 군 모두에서 여성 사용자 트윗의 평균 감성강도가 남성 사용자 트윗의

평균 감성강도보다 통계적( $p < 0.01$ )으로 높은 것으로 확인되었다. 여성 하위 90% 저활성 사용자의 메시지가 가장 높은 감성강도를 보이고 있으며, 남성 상위 10% 고탈성 사용자들의 메시지가 가장 낮은 감성강도 평균을 보이는 것으로 파악되었다. 여성의 경우 고탈성 사용자와 저활성 사용자 간의 차이가 크지 않지만, 남성의 경우 그 차이가 여성보다 큰 것으로 나타났다.

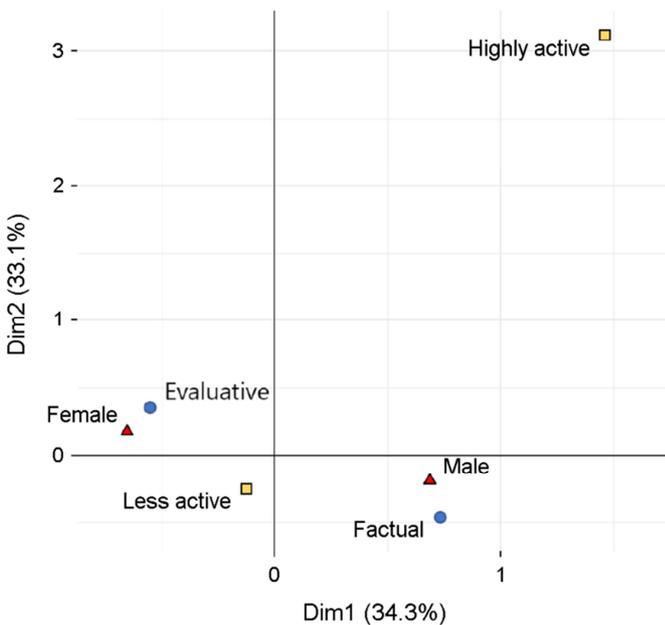
#### 4.2.2 Evaluative/Factual message ratio

사용자의 성별, 활성 수준과 사실적/평가적 메시지 비율의 연관관계를 파악하기 위하여 계층적 로그선형 분석을 실시하였으며,  $2 \times 2 \times 2$  포화(saturated) 모형으로부터 후진제거법을 이용하여 모형을 최적화한 결과는 Table 5와 같다. 분석결과, 세 개 요인의 교호작용이 유의수준 0.001에서 통계적으로 유의한 것으로 나타났다. Figure 4는 이들 3개 요인들의 상호관계를 파악하기 위해 시행한 대응분석 결과를 도시한 것인데, 여성 사용자의 경우, 평가적 메시지와 연관관계가 높으며, 남성 사용자의 경우 사실적 메시지와 연관관계가 높다는 것을 보여주고 있다.

**Table 5.** Test H0s for  $2 \times 2 \times 2$  stepwise backward hierarchical log linear analysis (Gender $\times$ User Activity $\times$ Evaluative/Factual)

If deleted simple effect is	D.F.	$\chi^2$	Sig.
Gender $\times$ User Activity $\times$ Evaluative/Factual	1	1654.94	0.000***

\*\*\* $p < .001$



**Figure 4.** Correspondence analysis of Gender $\times$ User Activity $\times$ Evaluative/Factual

Figure 5는 3개 요인들을 2개의  $2 \times 2$  분할표로 1/4원 씩 나누어 표시하는 fourfold plot으로 나타낸 것인데, 각 요인의 빈도가 1/4원의 면적으로 표현된다. 상위 10% 고탈성 사용자 군과 하위 90% 사용자 군 모두에서 사실적 메시지 보다 평가적 메시지가 상대적으로 더 많음을 볼 수 있다. 또한 상위 10% 고탈성 사용자 군의 여성들이 남성에 비해 평가적 메시지 작성 비율이 높은 것으로 나타났다.

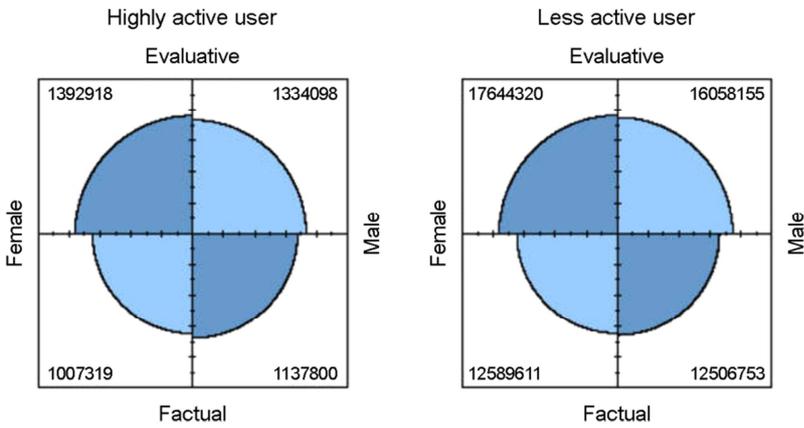


Figure 5. Fourfold plot of Gender×User Activity×Evaluative/Factual

### 4.2.3 Message

Table 6는 사용자의 성별, 활성 수준과 메시지 방향성(긍정/중립/부정)의 연관관계를 파악하기 위하여 계층적 로그선형 분석을 실시한 결과를 나타낸 것으로, 유의수준 0.001에서 3인자 교호작용이 통계적으로 유의함을 보여주고 있다. Figure 6는 이들 3개 요인들의 상

Table 6. Test H0s for 2×2×2 stepwise backward hierarchical log linear analysis (Gender×User Activity×Message Polarity)

If deleted simple effect is	D.F.	$\chi^2$	Sig.
Gender×User Activity×Message Polarity	1	528.41	0.000***

\*\*\* $p < .001$

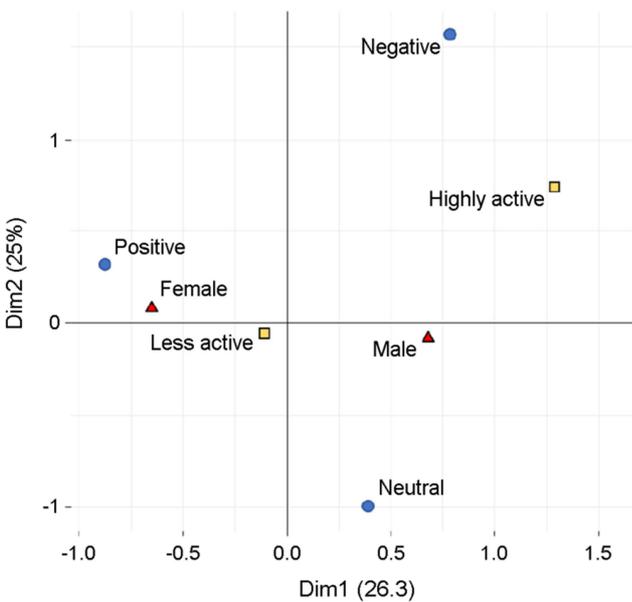


Figure 6. Correspondence analysis of Gender×Activity×Message polarity

호관계를 파악하기 위해 시행한 대응분석 결과이며, 여성, 하위 90% 저활성 사용자 군, 긍정 감성 요인들이 그룹을 이루고 있어, 하위 90% 저활성 여성 사용자가 긍정적인 감성을 표현하는 경향이 있음을 암시하고 있다.

이들 3개 요인들의 분할표를 fourfold plot 형태로 도시하면 Figure 7과 같은데, 상위 10% 활성 사용자 군의 경우, 남성 사용자가 여성 사용자들 보다 부정적인 메시지 작성 비율이 높음을 보여주고 있다. 또한 긍정적인 메시지들은 모든 사용자 군에서 남성들 보다 여성들이 상대적으로 더 높은 비율로 작성한 것으로 파악된다.

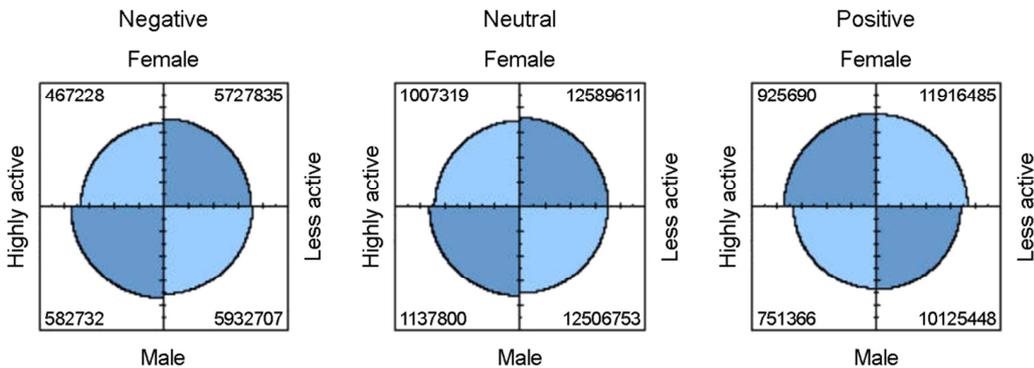


Figure 7. Fourfold plot of Gender×User Activity×Sentiment polarity

## 5. Conclusion

최근 SNS는 폭발적인 성장에 따라 학문적, 사회적으로 큰 관심을 모으고 있으며, 이를 다양한 분야에 접목하는 시도들이 이루어지고 있다. 이들 중 주목받는 분야가 감성분석 분야인데, 사용자에게 통찰을 줄 수 있는 감성분석 결과를 도출하기 위해서는 데이터에 대한 기계적인 분석뿐만 아니라 분석 데이터들을 양산하는 사용자들의 특성과 감성표현 양상에 대한 연구가 병행될 필요가 있다. 본 연구에서는 대표적인 SNS인 트위터 사용자들을 대상으로, 사용자들의 성별, 활성수준과 감성표현의 양상, 특성 등의 연관관계에 대한 분석을 시행하였다. 분석결과를 요약하면 다음과 같다.

첫째, 상위 10% 고활성 사용자 군에서 남성 사용자들의 포스팅 활동이 여성 사용자들보다 많음을 확인할 수 있다. 둘째, 메시지 작성 시, 상위 10% 고활성 사용자들이 작성 문자 수가 저활성 사용자들 보다 많은 것으로 나타났으며, 이 상위 10% 고활성 사용자 군에서는 여성 사용자들의 문자 수가 남성 사용자들의 문자 수보다 유의하게 많은 것으로 파악되었다. 셋째, 두개 활성수준 모두에서 여성 사용자가 작성한 메시지의 감성강도가 남성 사용자의 메시지 감성강도 보다 통계적( $p < 0.01$ )으로 높은 것으로 확인되었다. 하위 90% 저활성 여성 사용자가 작성한 메시지들이 가장 감성강도가 높은 것으로 나타났으며, 상위 10% 고활성 남성 사용자의 메시지들이 가장 감성강도가 낮은 것으로 파악되었다. 넷째, 여성 사용자의 경우, 평가적 메시지와 연관관계가 높으며, 남성 사용자의 경우 사실적 메시지와 연관관계가 높은 것으로 나타났다. 또한 두 개 활성수준 모두에서 사실적 메시지 보다 평가적 메시지가 상대적으로 더 많은 것으로 파악되었다. 다섯째, 하위 90% 저활성 여성 사용자가 긍정적인 감성메시지와와의 연관성이 높은 것으로 나타났으며, 사용자 군에 상관없이 여성들이 남성들 보다 상대적으로 긍정적인 메시지 작성 비율이 높은 것으로 파악되었다.

본 연구는 SNS 중 트위터를 대상으로 분석을 진행하여 텍스트 메시지를 이용한 감성표현에 국한되는 한계를 가진다. 하지만 이러한 결과는 개방형 소셜미디어 내에서의 성별, 활성화 수준 별 감성표현에 대한 실증적 자료를 제공함으로써, 감성분석 결과로부터 좀더 유용한 통찰을 얻을 수 있는 기초자료로서 의미를 가진다. 추후 사진, 이모티콘 등 다양한 멀티미디어를 통한 감성표현에 대한 분석이 추가된다면 한 층 활용도 높은 감성분석 시스템을 구축하는데 활용 가능할 것으로 기대된다.

## Acknowledgement

이 논문은 2014년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2014S1A5A2A03066219).

## References

- Anderson, C., *The long tail: Why the future of business is selling less of more*, New York, NY: Hyperion, 2006. ISBN: 1401302378
- Baumeister, R.F., Bratslavsky, E., Finkenauer, C. and Vohs, K.D., Bad is Stronger than Good, *Review of General Psychology*, 5, 323-370, 2001. doi: 10.1037/1089-2680.5.4.323
- Chiou, J. and Cheng, C., Should a Company Have Message Boards on Its Web Site?, *Journal of Interactive Marketing*, 17(3), 50-61, 2003. doi: 10.1002/dir.10059
- Chu, S.C. and Kim, Y., Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising*, 30(1), 47-75, 2011. doi: 10.2501/IJA-30-1-047-075
- Doh, S. and Hwang, J., How Consumers Evaluate eWOM Messages, *Cyber Psychology and Behavior*, 12(2), 193-197, 2009. doi: 10.1089/cpb.2008.0109
- Darley, W.K. and Smith, R.E., Advertising claim objectivity: Antecedents and effects. *Journal of Marketing*, 57(4), 100-113, 1993. doi: 10.2307/1252222
- Edell, J.A. and Staelin, R., The information processing of pictures in print advertisements. *Journal of Consumer Research*, 10(1), 45-61, 1983. doi: 10.2307/1131221
- Jang, P.S., Relationships Among User Group, Gender and Self-disclosure in Social Media, *Journal of The Korea Society of Computer and Information*, 23(4), 25-31, 2018. doi: 10.9708/jksci.2018.23.04.025
- Jones, S., Millermaier, S., Goya-Martinez, M. and Schuler, J., Whose space is MySpace? A content analysis of MySpace profiles. *First Monday*, 13(9), 2008. doi: 10.5210/fm.v13i9.2202.CS1
- Kim, H.H. and Jang, P.S., Differences in Sentiment on SNS: Comparison among Six Languages, *Journal of Digital Convergence*, 14(3), 165-170, 2016. doi: 10.14400/JDC.2016.14.3.165
- Lima, A.C.E., Castro, L.N. and Corchado, J.M., A polarity analysis framework for Twitter messages. *Applied Mathematics and Computation*, 270, 756-767, 2015. doi: 10.1016/j.amc.2015.08.059
- Lee, J.Y. and Jang, P.S., Effects of message polarity and type on word of mouth through SNS (Social Network Service), *Journal of Digital Convergence*, 11(6), 129-135, 2013. UCI: G704-002010.2013.11.6.019
- Lee, J.Y. and Jang, P.S., Airline Customer Satisfaction Analysis using Social Media Sentiment Evaluation: Full Service Carriers vs. Low Cost Carriers. *Journal of Digital Convergence*, 15(6), 189-196, 2017. doi: 10.14400/JDC.2017.15.6.189
- Lee, J.Y. and Jang, P.S., Study on Research Trends in Airline Industry using Keyword Network Analysis: Focused on the Journal Articles in Scopus. *Journal of the Korea Convergence Society*, 8(5), 169-178, 2017. doi: 10.15207/JKCS.2017.8.5.169

Lincoln, M., *gender: Predict Gender from Names Using Historical Data. R package version 0.5.2*, 2018, <https://github.com/ropensci/gender> (retrieved Jan 10, 2019).

MonetDB. *An Open-Source Database System*. <https://www.monetdb.org/> (retrieved Jan 10, 2019).

Muscanell, N.L. and Guadagno, R.E., Make new friends or keep the old: Gender and personality differences in social networking use. *Computers in Human Behavior*, 28(1), 107-112, 2012. doi: 10.1016/j.chb.2011.08.016

Park, D.H., and Kim, S., The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, 7(4), 399-410, 2008. doi: 10.1016/j.elerap.2007.12.001

Pew Research Center, *Men catch up with women on overall social media use*, 2015, <http://www.pewresearch.org/fact-tank/2015/08/28/men-catch-up-with-women-on-overall-social-media-use/> (retrieved Jan 10, 2019).

Ravi, K. and Ravi, V., A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*, 89, 14-46, 2015. doi: 10.1016/j.knosys.2015.06.015

*SentiStrength*. <http://sentistrength.wlv.ac.uk/> (retrieved Jan 10, 2019).

Subhasree, B. and Saptarsi, G., *RSentiment: Analyse Sentiment of English Sentences. R package version 2.2.2*, 2018, <https://CRAN.R-project.org/package=RSentiment> (retrieved Jan 10, 2019).

Tedjamulia, S.J., Dean, D.L., Olsen, D.R. and Albrecht, C.C., Motivating content contributions to online communities: Toward a more comprehensive theory. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 193b-193b, 2005. doi: 10.1109/HICSS.2005.444

Thelwall, M., Buckley, K. and Paltoglou, G., Sentiment strength detection for the social Web, *Journal of the American Society for Information Science and Technology*, 63(1), 163-173, 2012. doi: 10.1002/asi.21662

Tiggemann, M. and Slater, A., NetGirls: The Internet, Facebook, and body image concern in adolescent girls. *International Journal of Eating Disorders*, 46(6), 630-633, 2013. doi: 10.1002/eat.22141

*Twitter.com: Docs-Twitter Developers*. <https://dev.twitter.com/streaming/overview> (retrieved Jan 10, 2019).

Xue, F. and Zhou, P., The Effects of Product Involvement and Prior Experience on Chinese Consumers' Responses to Online Word of Mouth, *Journal of International Consumer Marketing*, 23(1), 45-58, 2010. doi: 10.1080/08961530.2011.524576

## Author listings

**Phil-Sik Jang:** philsjang@gmail.com

**Highest degree:** PhD, Department of Industrial Engineering, KAIST

**Position title:** Professor, Department of Air Traffic and Logistics, Sehan University

**Areas of interest:** Human-computer Interaction, Sentiment Analysis, Bigdata Analysis, Machine Learning