

<http://dx.doi.org/10.7236/JIIBC.2014.14.3.185>

JIIBC 2014-3-26

## 트위터 기반 고객의 관심도 분석을 통한 마케팅 조언 시스템의 설계 및 구현

### Design and Implementation of Marketing Advisement System through the Concern Degree Analysis of Customers Based on Twitter

이기영\*, 김혜영\*\*, 김아름\*\*, 김성배\*\*\*

Ki-Young Lee\*, Hye-Young Kim\*\*, Aluem Kim\*\*, Sung-Bae Kim\*\*\*

**요약** 최근 급속한 스마트폰 이용자의 증가와 무선인터넷 서비스의 확장과 함께 SNS 이용자도 급증하고 있다. 많은 SNS 중에서도 트위터(Twitter)는 대한민국 내 SNS 시장을 주도하고 있다. 트위터 이용자들은 트위터를 통해 자신의 생각과 감정을 표현한다. 본 논문에서는 유통업체 인근의 트윗을 오피니언 마이닝을 통해 분석하고 긍정, 중립, 부정의 정도와 트윗의 개수를 이용한 관심도의 분석을 통해 마케팅자에게 마케팅 메시지를 전달한다. 그 결과, 우리는 해당 유통업체의 마케팅과 운영에 인근 고객의 요구를 반영하는 시스템을 제안한다.

**Abstract** With the fast increment of smart phone users and extension of wireless internet the number of SNS user is also increasing. Twitter among lots of SNS takes the lead in SNS market. Twiter users express their thinking and feelings. In this paper, by analyzing twitts near the distribution enterprise using opinion mining. And by analyzing concern degree using the number of twitts and positive, neutral, negative degree we deliver marketing message to marketer. As the result, we propose that marketing and management of this distribution enterprise can reflect the demand of customer who is near there.

**Key Words** : Mobile Phone, Twitter, Marketing, Distribution Enterprise

## 1. 서 론

최근 들어 스마트폰 이용자의 증가와 무선인터넷 서비스의 확장과 더불어 SNS의 이용자 또한 급증하고 있다. 대한민국내 SNS 시장을 주도하고 있는 페이스북(Facebook)과 트위터(Twitter) 이용자 수는 이미 2011년에 1천만 명을 돌파했으며,<sup>[1]</sup> 그 지속적인 증가 추세는 당

분간 멈추지 않을 것으로 예상된다. SNS는 광범위하고 동시에 특정 성향의 집단으로 분류될 수 있는 서비스 이 사용자들을 데이터베이스에 의해 파악하고 관리할 수 있다는 점에서 마케팅 활용가치가 날로 부상하고 있다.<sup>[2]</sup>

트위터(Twitter)는 140자 이내의 글만 적으면 되는 간편성과 엄청난 전파력을 지닌 확장성의 특징이 있다. 사용자들은 트윗을 통해 자신의 생각을 표현 하곤 한다.<sup>[3]</sup>

\*종신회원, 을지대학교 의료IT마케팅학과

\*\*정회원, 을지대학교 의료IT마케팅학과

\*\*\*정회원, 을지대학교 의료IT마케팅학과

접수일자 2014년 4월 25일, 수정완료 2014년 5월 20일

게재확정일자 2014년 6월 13일

Received: 25 April, 2014 / Revised: 20 May, 2014

Accepted: 13 June, 2014

\*\*\*Corresponding Author: ksb@eulji.ac.kr

Dept. of Medical IT and Marketing, Eulji University, Korea

이처럼 SNS에서 영화나 맛집 리뷰 및 제품 후기 등이 끊임없이 쏟아지고 있기 때문에 사용자들이 방대한 데이터에서 원하는 정보를 찾는 작업은 점차 어려워지고 있으며 그 효율성 측면에서도 많은 문제를 발생시키고 있다. 따라서 SNS의 대량의 정보로부터 사용자가 원하는 정보를 빠르게 분석해 주고, 유의미한 정보를 지능적으로 유추해내는 오피니언마이닝(opinion mining) 기술의 중요성은 그 어느 때보다도 커지고 있는 실정이다.<sup>[4]</sup>

본 논문에서는 트위터의 이러한 특성을 기반으로 유통업체의 인근에서 발생한 트윗을 오피니언 마이닝을 이용해 분석하여 입점 고객의 요구를 마케팅과 전반적인 유통업체의 운영에 반영하는 시스템을 제안한다.

## II. 관련 연구

### 1. SNS

가까운 지인 또는 유명인들과 소소한 일상을 공유하고 정보를 획득할 수 있는 SNS(Social Network Service)는 이미 주요한 인터넷 서비스로 자리잡고 있다. 최근 몇 년간 다양한 SNS 서비스가 등장하고 있으며, 세계적으로 SNS의 사용자도 급증하고 있다. 이러한 성장은 SNS가 콘텐츠를 쉽게 만들 수 있고, 신뢰가 높은 관계 혹은 친분을 중심으로 사용자의 일상, 생각, 콘텐츠를 공유할 수 있다는 장점이 있기 때문이다.<sup>[5]</sup>

### 2. LBSNS로서의 트위터

트위터는 구조적으로 사용자들 사이의 관계가 명확하고, 사용자가 작성한 트윗에는 개인의 일상, 사회 이슈, 사건 및 사고와 같은 내용과 함께 GPS 좌표 정보도 선택적으로 포함할 수 있기 때문에 이를 이용하여 다양한 위치 기반 응용 시스템을 개발할 수 있다.<sup>[6]</sup>

본 논문에서는 이러한 특성을 지닌 트위터를 이용하여 유통업체 인근에서 발생한 트윗 분석을 통해 마케팅에 입점 가능 고객의 요구가 반영되도록 하는 시스템을 제안한다.

### 3. 트위터의 오피니언 마이닝

SNS를 이용하여 여론 분석을 위한 오피니언 검색의 핵심 이슈는 주어진 주제에 대해서 긍정과 부정 내용을 분리하는데 있다. 이러한 감성 분류 기술은 기존의 오피

니언 마이닝에서 이미 많은 연구가 진행되었다. 기존의 오피니언 마이닝 연구에서는 주어진 문서가 모두 오피니언 문서라는 가정이 있었다. 그러나 트위터는 오피니언 문서의 비율이 그렇지 않은 문서(객관적인 사실만을 언급한 문서)에 비해 현저히 작은 경우가 흔하다. 따라서 트위터 환경에서 오피니언 문서를 검색하기 위해서는 우선적으로 오피니언 문서만을 분리하는 것이 핵심 기술이라고 볼 수 있다.<sup>[7]</sup>

## 4. 유통업체의 SNS 마케팅 활용

현재 유통업체에서는 SNS 활용이 기존의 홍보방법과 온오프라인 광고의 시각에서 벗어나지 못하고 있다. 이제 소비자들의 행태와 우리 사회 인식은 점차 소셜미디어 채널에서 소통과 대화를 나누는 추세로 나아가고 있다. 지금 소셜미디어에 동참하지 않는다면 SNS의 장기적 안목에서 볼 때 향후 마케팅 시장에서 고전하게 될 것으로 추측된다.<sup>[8,9]</sup>

본 논문의 궁극적인 목표는 유통업체의 새로운 SNS 활용 전략으로 기존의 기업→고객 위주의 마케팅에서 벗어나 고객→기업으로의 변화를 꾀하는 것이다.

## III. 시스템 설계

### 1. 시스템 흐름도

본 논문에서는 트위터에서 점포 인근 고객의 트윗을 분석하여 점포의 운영과 마케팅에 고객의 성향과 욕구를 반영하는 시스템을 제안한다.

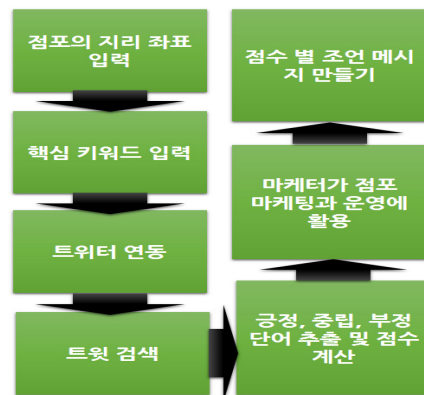


그림 1. 시스템 흐름도  
Fig. 1. System Flow Chart

그림 1은 본 논문의 시스템 흐름도이다. 우선 사용자는 점포 지리의 좌표와 고객의 성향을 반영하는 핵심 키워드를 입력한다. 그 후 트위터와 자동으로 연동되어 입력한 점포 인근 고객의 트윗 중 핵심 키워드가 들어가 있는 트윗이 검색된다. 이 트윗에서 긍정과 부정적인 단어를 추출하여 해당 트윗의 긍정, 중립, 부정 정도를 파악한다. 모든 트윗의 점수가 파악되면 순위를 매기고 순위에 따라 조언 메시지를 만들게 된다. 이 메시지들은 사용자(마케터)에게 전달되어 사용자가 마케팅과 점포 운영에 활용할 수 있도록 한다.

## 2. 시스템 설계도

그림 2는 본 논문의 시스템 설계도이다. Linking Twitter Module은 트위터와 본 시스템을 연결해주는 모듈로 지리 정보와 핵심 키워드를 입력하는 Search Geographic Information & Main Keyword 기능과 입력한 값에 해당하는 트윗을 검색하는 Search Twit으로 구성되어 있다.

Opinion Mining Module은 Linking Twitter Module에서 추출된 트윗을 이용하여 해당 트윗의 긍정, 부정 정도 점수를 계산하는 모듈로 긍정, 부정 단어가 저장되어 있는 Positive & Negative Words DB와 이 DB와 연동하여 트윗에서 긍정, 부정 단어를 검색하는 Searching Positive & Negative Words, 이를 토대로 긍정, 부정 정도 점수를 계산하는 Extracting Score of Positiveness & Negativeness로 구성되어 있다.

Advising to Marketer Module은 Opinion Mining Module에서 추출한 트윗에 대한 긍정, 부정 정도 점수에

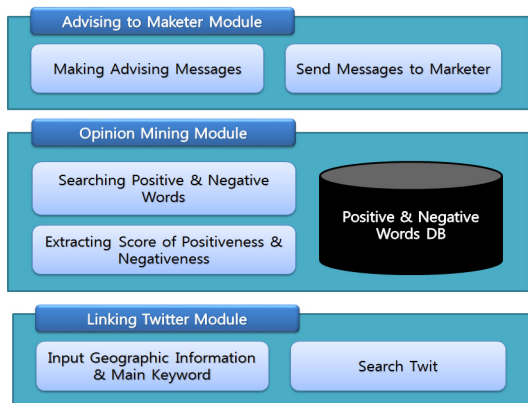


그림 2. 시스템 설계도  
 Fig. 2. System Architecture

따라 마케팅자에게 조언 메시지를 만들어 보여주는 모듈로 조언 메시지를 만드는 Making Advising Messages와 이 메시지를 마케팅자에게 전송하는 Send Message to Marketer 기능으로 이루어져 있다.

## IV. 시스템 구현

실험에 사용된 시스템은 표 1과 같으며, 하드웨어 사양은 Intel(R) i5 2.50GHz, 8GB RAM이며, 운영체제는 Windows 7 SP1를 사용하였다. 또한 실험을 위해 자바 기반 한글 형태소 분석기인 KOMORAN을 사용하였다. KOMORAN은 오픈 소스이다.

본 논문에서 구현된 시스템은 프로토타입으로 트위터의 트윗을 오픈피니언 마이닝을 이용하여 분석하고 이를 수치화하였다. 향후 사용자(마케터)에게 실질적으로 도움이 되는 마케팅 메시지를 전달할 것이다.

표 1. 실험 환경

Table 1. Experiment Environment

RAM	8GB
Programming Language	Java
Database	MySQL
Korean Morpheme Analyzer	KOMORAN
CPU	Intel i5 2.50GHz

다음 그림 3은 점포의 지리 좌표와 핵심 키워드를 이용해 트위터의 트윗을 검색하는 소스이다. 트위터의 OAuth 인증을 통과하기 위해 본인의 트위터 계정을 Twitter Developer에 등록하여 ConsumerKey와 ConsumerSecret, Access Token, AccessScret을 얻고 연동시켰다.

표 2은 트윗의 긍정, 중립, 부정 점수 계산법이다. 이 점수는 10에 가까울수록 핵심 키워드에 대해 긍정적이며, 0에 가까울수록 부정적이다.

표 2. 긍정, 중립, 부정 점수 계산표

Table 2. Table for the Score of Positiveness, Neutrality, Negativeness

긍정	+10 ~ +6
중립	+5
부정	0 ~ +4

```

try {
    ConfigurationBuilder builder = new ConfigurationBuilder();
    builder.setOAuthConsumerKey(ConsumerKey);
    builder.setOAuthConsumerSecret(ConsumerSecret);

    Configuration config1 = builder.build();
    OAuthAuthorization oauth = new OAuthAuthorization(config1);
    oauth.setOAuthConsumer(ConsumerKey, ConsumerSecret);

    AccessToken accessToken = new AccessToken(accessToken,
        AccessSecret);

    TwitterFactory factory = new TwitterFactory(config1);

    Twitter twitter = factory.getInstance(accessToken);

    for(int i=0; i< name.length; ++i) {
        String categoryName = name[i];
        Query query = new Query(categoryName);

        System.out.println("====="+categoryName+"=====");

        query.setGeoCode(new GeoLocation(37.579986, 126.898457), 5,
            Query.KILOMETERS);
        query.setSince("2013-12-27");
        query.setUntil("2014-01-02");
        QueryResult result = twitter.search(query);

        List<Status> tweets = result.getTweets();

        for(int j = 0; j < tweets.size(); ++j) {

            koranalyzer.koranalyze(categoryName, tweets.get(j).getText());

        }
    }
}
    
```

그림 3. 트위터 연동을 위한 소스코드  
 Fig. 3. Source Code for Connecting Twitter

그림 4는 점수를 계산하는 알고리즘을 나타낸 것이다. 긍정과 부정은 DB에 저장되어 있는 단어가 있을 경우에 점수를 부과하며, DB에 단어가 없을 경우 5점을 부과한다. 마지막으로, 총 평균을 내어 핵심 키워드에 대한 긍정, 중립, 부정 정도를 구한다.

```

while EOF //파일의 끝까지 반복한다.
if word not exist then score += 5
else score += word
count ++
totScore = score / count
print totScore
if totScore > 5 then print "positive"
else if totScore = 5 then print "neutral"
else totScore < 5 then print "negative"
    
```

그림 4. 점수 계산을 위한 알고리즘  
 Fig. 4. Algorithm for Calculating Scores.

이를 토대로 총합계를 통해 해당 핵심 키워드에 대한 관심도 즉, 인근 지역 사람들이 얼마나 이 키워드에 대해 관심을 가지고 있는가를 알 수 있다. 사용자는 이 수치들을 이용하여 어떻게 마케팅에 적용시키고, 유통업체의

운영을 어떻게 할 것인가에 대한 고려를 할 수 있다.

## V. 성능평가

본 논문에서는 유통업체 인근 지역 트윗을 특정 핵심 검색어로 검색하여 관심도와 긍정, 부정 정도를 알 수 있는 시스템을 제안하였다. 본 논문에서의 관심도는 모든 트윗에 나타난 서술어의 긍정, 중립, 부정 정도를 합한 값이다. 구현한 시스템을 성능평가하기 위한 척도로 수색역 근처에 위치한 유통업체인 E사의 판매량과 수색역 근처에서 발생한 트윗의 핵심 키워드에 대한 관심도를 비교해 보았다.

그림 5는 E사의 대용식, 유제품, 과자, 커피음료의 판매량과 이에 대응하는 핵심 키워드에 대한 트위터의 관심도를 비교한 결과이다. 대표적인 핵심키워드로 대용식은 시리얼을, 유제품은 우유, 과자는 과자, 커피음료는 커피로 선정하였다. 성능 평가 결과로 보았을 때 판매량과 인근 트윗의 관심도는 대체적으로 비슷한 양상을 띠는 것으로 나타났다.

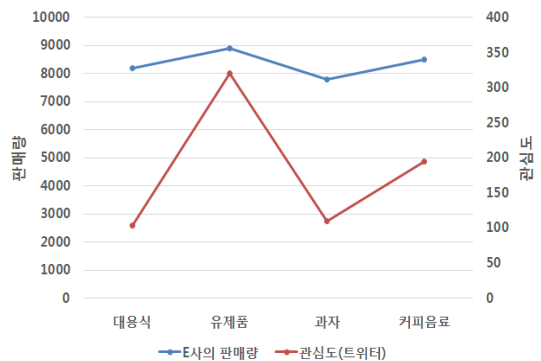


그림 5. 성능평가 결과  
 Fig. 5. The Result of Performance Evaluation

트윗에서 키워드에 대한 긍정, 부정, 중립의 정도도 파악해 보았는데, 대용식과 과자, 커피음료는 중립적으로 나타났으며, 유제품은 부정적으로 나타났다.

## VI. 결론

본 논문에서는 유통업체 인근 지역의 핵심 키워드에 대한 트윗을 오피니언 마이닝을 통해 긍정, 중립, 부정인

지 검색하여 분석하고, 관심도를 계산하여 사용자(마케터)가 해당 유통업체의 마케팅과 운영에 고객의 요구를 반영할 수 있도록 하는 시스템을 제안하였다.

본 논문의 시스템 구현은 프로토타입으로 트윗을 분석하여 긍정, 중립, 부정 정도와 관심도를 계산하였다. 향후, 실무에 있는 마케터와 협력하여 본 수치들을 이용해 마케팅적 메시지를 전달하는 것을 목표로 할 것이다. 또한, 트위터는 오피니언 문서의 비율이 적어 연구가 어려웠으므로, 트위터에서 오피니언 문서를 구별하거나 핵심 키워드에 대한 트윗을 요약하는 시스템을 개발할 예정이다.

## References

- [1] <http://www.mt.co.kr/view/mtview.php?type=1&no=2012011816300582344&outlink=1>
- [2] Wikipedia(Social Network Service)
- [3] Sang-Hee Huri, Kyu-Soo Choi, "A Study on Characteristics and Typed of Tweet in Twitter," Journal of Hanminjok Emunhak, Vol. 61, pp. 455-494, August 2012.
- [4] Kyoung-Mi Park, Ho-Kun Park, Hyuong-Gon Kim, Hyeon-Dong Go, "Research of Opinion Mining in SNS", The Journal of the Korean Institute of Information Scientists and Engineers, Vol. 23, No. 11, pp.54-60, November 2011.
- [5] Chang-Woo Choi, "The Process of Development and Main Cases of SNS", The Journal of ie Magazine, Vol. 18, No. 1, pp.20-23, March 2011.
- [6] Bumsuk Lee, Seokjung Kim, Eyung-Yeon Hwang, "Analyzing the Credibility of the Location Information Provided by Twitter Users", Journal of Korea Multimedia Society, Vol. 15, No. 7, pp. 910-919, July 2012.
- [7] Jae-Young Chang, "Automatic Retrieval SNS Opinion Document Using Machine Learning Technique", The Journal of the Institute of Internet, Broadcasting and Communication(IIBC) , Vol. 13, No. 5, October 2013.
- [8] Jaeyong Choi, Sanghyun Kim, Jusang Kim, YongCheol Hwang, "Study on the Patterns of Social Media Marketing in Inbound Retail

Componies", The Proceeding of Korea Distribution Association in Fall, pp.229-253, October 2012.

- [9] Yoo-Jung Kim, Jae-Young Kim, Jae-Min Han, "The structural relationships among user citizenship behavior, aberrant user behavior, social connectedness, privacy concern, and user satisfaction", Journal of the Korea Academia-Industrial cooperation Society(JKAIS), Vol. 13, No. 11, pp.4994-5004, 2012.

## 저자 소개

### 이 기 영(중신회원)



- 제10권 1호 참조
- 2009년~현재 : 한국인터넷방송통신학회 상임이사
- 1991년~현재 : 을지대학교 의료IT마케팅학과 교수
- 주관심분야 : u-Healthcare, 공간 데이터베이스, GIS, LBS, USN, 텔레메틱스, 빅데이터 등

• E-mail : kylee@eulji.ac.kr

### 김 아 름(정회원)



- 2011년~현재 : 을지대학교 의료IT마케팅학과 학생
- 주관심분야 : 오피니언 마이닝, 인공지능, 빅데이터 등
- E-mail : tpfldkfm@naver.com

### 김 성 배(정회원): 교신저자



- 홍익대학교 경영대학 경영학과
- 홍익대학교 대학원 경영학과(마케팅 전공, 경영학 석사)
- 홍익대학교 대학원 경영학과(마케팅/유통 전공, 경영학 박사)
- 2009년~현재 : 한국유통경영학회 이사
- 2000년~현재 : 을지대학교 의료IT마케팅학과 교수

• 주관심분야 : IT마케팅, 데이터베이스마케팅, u-Healthcare 관련 마케팅, 마케팅전략, 상품개발 및 관리, 촉진전략, 빅데이터 등

• E-mail : ksb@eulji.ac.kr

김 혜 영(정회원)



- 을지대학교 유통경영학사
- 현재 : 을지대학교 의료IT마케팅학과 석사과정중
- 1993 : 신세계백화점 공채입사
- 2008년~현재 : 신세계인터넷내셔널 선임수퍼바이저 근무
- 주관심분야 : IT마케팅, 빅데이터, 고객행동분석, 마케팅전략, 시장분석 등

• E-mail : rbuja2@nate.com