

# 소셜 네트워크 빅데이터 기반 택배업체 고객만족도에 관한 연구

## A Study on Customer Satisfaction for Courier Companies based on SNS Big data

이동준(DongJun Lee)\*, 원종운(JongUn Won)\*\*,  
권용장(YongJang Kwon)\*\*\*, 김미례(MiRye Kim)\*\*\*\*

### 초 록

최근 국내외 택배업체들은 가격경쟁에 의한 수익성 악화로 서비스의 차별화를 통한 고객 및 수익성 확보에 많은 노력을 기울이고 있다. 따라서, 택배 서비스의 품질 개선을 통한 고객 만족도를 높이는 노력이 어느 때 보다 중요한 실정이다. 하지만 기존의 택배서비스 품질 측정은 오프라인 설문조사로 이루어져 많은 시간과 비용을 들어간다는 한계를 지니고 있다. 이러한 한계는 온라인상의 소셜 빅데이터 분석을 활용한다면 보다 적은 비용과 노력으로 극복 가능할 것이며, 택배업체 경쟁력 강화에 크게 도움이 될 수 있을 것이다. 따라서, 본 연구에서는 SNS상의 빅데이터를 활용하여 국내외 택배업체들에 대한 의견을 수집하고, R을 통해 각 택배업체들의 고객만족도를 분석하고 이를 미국고객만족도(ACSI), 한국국가고객만족도(NCSI)와 비교하여 검증을 실시하였다. 그 결과 SNS 분석 결과와 고객만족도가 뚜렷한 양적 선형관계를 형성하였다. 이는 향후 실시간 소셜 네트워크의 정보를 이용하여 간편하게 고객만족도 결과를 예측할 수 있다는 것을 의미한다.

### ABSTRACT

Global courier companies have been devoting to get more customers and profits with different service because of the worse profits from price competition. So, the effort of improving satisfaction of customers through improving courier service qualities is more important than any other time. However, the previous way to measure courier service has limitation that costs lots of time and money from off-line survey. This limitation could be overcome with less effort and costs if utilizing on-line social big data analysis and it is so helpful to improve competitiveness of courier companies. Therefore, I have collected comments from domestic and international courier companies from big data on social network service, analyzed the

---

본 연구는 한국철도기술연구원 2016년도 주요사업(PK1604A)의 연구비 지원으로 수행되었음.

\* First Author, Transportation and Logistics System, University of Science and Technology (dongjun@krii.re.kr)

\*\* Corresponding Author, Logistics system Research Team, Korea Railroad Research Institute (juwon@krii.re.kr)

\*\*\* Co-Author, Logistics system Research Team, Korea Railroad Research Institute(yjkwon@krii.re.kr)

\*\*\*\* Co-Author, Transportation and Logistics System, University of Science and Technology (heyday2015@krii.re.kr)

Received: 2016-10-25, Review completed: 2016-11-09, Accepted: 2016-11-18

satisfaction of customers by R and verified the result by comparing with American Customer Satisfaction Index (ACSI) and Korea National Customer Index (NCSI) in this research. I found out the result depicts clear correlation between SNS analysis and customer satisfaction. This study can be the foundation to predict customer satisfaction easily by utilizing real time SNS information.

**키워드** : 택배, 고객만족도, 빅데이터, 소셜 네트워크

Courier Company, Big Data, Customer satisfaction, Social Network

## 1. 서 론

택배업계의 변화 동향을 살펴보면, 기업택배고객의 급속한 증가로 인하여 1992년 한진택배를 시작으로 2000년 이후 현대, 동부, 동원, 그리고 유진그룹 등과 같은 대기업의 중견 택배기업 인수가 지속적으로 이루어져 왔을 뿐만 아니라 2011년에는 CJ GLS가 금호그룹의 대한통운을 인수하여 택배물동량 1위를 차지하게 되었다. 이와 같은 택배업계 M&A의 근본적인 원인은 원가절감에 기반을 둔 지나친 가격경쟁에 따른 수익성 악화에서 찾아볼 수 있을 것이다. 즉, 무분별한 택배단가의 하락은 중견 및 소규모 택배업체의 사업포기로 이어지게 되었으며, 고객서비스에도 매우 큰 영향을 미치게 되었다[21].

최근 국내외 택배업체들은 서비스 차별화를 위한 노력을 다방면에서 전개하고 있다. 국내 소셜 커머스 업체인 쿠팡은 2014년 3월부터 24시간 이내 배송인 '로켓배송'을 운영하고 있고, 미국 우정청(USPS)은 2014년 10월부터 뉴욕과 샌프란시스코 등 일부 대도시에서 새벽 3시부터 우유와 달걀을 비롯한 식료품을 배달해주는 서비스를 시작하였다. 일례로 드론을 이용한 무인 배송 서비스, 고객이 미리 주문한 물품을 차 안에서 바로 수령할 수 있는 드라이브 스루 서비

스 등을 제공하고자 노력하고 있다[10].

최근 빅데이터의 관심 및 활용도가 더욱 증가하고 있다. 유통업계에서는 인터넷상의 빅데이터 자료를 활용하여 마케팅에 적극 활용하고 있다. 고객의 니즈를 실시간으로 파악하여 사업 모델에 적극반영하고 있는 것이다. 택배관련 사업에서도 주변에 많은 데이터가 존재하고 있다. 실제로 택배업체에서는 이러한 많은 데이터를 활용하여 배송추적관리시스템 등의 고객 편의를 위한 서비스를 제공하고 있다. 하지만 택배 서비스에 대한 만족도 및 고객의 니즈를 파악하고 대응하는데 빅데이터의 활용은 미약하다.

그동안 택배서비스의 고객만족도와 관련하여 많은 연구가 있었다. 하지만 대부분의 연구들이 택배서비스 품질을 측정하기 위해서 고객만족도를 활용하는 방식으로 학문적 연구가 진행되었다. 이때 고객만족도는 주로 설문자료를 통해 수집하였다. 설문의 경우는 비용도 많이 들어가고, 많은 표본을 확보하는데 제한이 있다. 물론 국가고객만족도를 활용할 수도 있지만, 국가고객만족도는 연 1회 브랜드에 대한 만족도를 나타내므로 택배서비스 개선요인을 세부적으로 분석하기에는 그 활용도에 제한이 있다. 고객의 만족도를 실시간 혹은 월 및 분기단위로 정확한 분석이 가능하다면 택배업체의 서비스 차별화 전략의 방향을 제시해 줄 있을 것이다.

본 연구에서는 이러한 고객만족도 설문조사를 소셜 네트워크(SNS, Social Network Service) 빅데이터를 활용한 분석으로 대체할 수 있는 부분이 있다는 생각에서 출발하였다. SNS상의 빅데이터를 활용하면 택배 이용 고객들의 만족도 및 요구사항을 예측할 수 있을 것이라는 가정하고, 미국 및 국내 택배업체들에 대한 SNS상의 트윗 의견을 수집하여 R을 통해 각 택배업체들의 소비자 이미지를 분석하였다. 이를 미국고객만족도(ACSI, American Customer Satisfaction Index), 한국국가고객만족도(NCSI, National Customer Satisfaction Index)와 비교하여 검증 실시하여 SNS 빅데이터를 활용한 고객만족도 분석의 가능성을 진단하였다. 분석 대상 택배업체는 검증을 위하여 ACSI, NCSI 택배분야에서 고객만족도지수를 제공하는 업체로 한정하였다.

## 2. 이론적 배경

### 2.1 빅데이터

빅 데이터(Big Data)란 기존 데이터베이스 관리도구로 데이터를 수집, 저장, 관리, 분석할 수 있는 역량을 넘어서는 대량의 정형 또는 비정형 데이터 집합 및 이러한 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 의미한다. 다양한 종류의 대규모 데이터에 대한 생성, 수집, 분석, 표현을 그 특징으로 하는 빅 데이터 기술의 발전은 다변화된 현대 사회를 더욱 정확하게 예측하여 효율적으로 작동케 하고 개인화된 현대 사회 구성원 마다 맞춤형 정보를 제공, 관리, 분석 가능케 하며 과거에는

불가능했던 기술을 실현시키기도 한다[6].

빅데이터 기술은 크게 분석 기법과 인프라 기술 측면으로 나누어 볼 수 있다. 빅데이터 분석 기법으로 데이터 마이닝, 텍스트 마이닝, 평판 분석, 소셜 분석, 클러스터 분석, 현실 마이닝 등의 크게 6가지로 나누어 볼 수 있다[20]. 본 연구에서는 소셜 빅데이터를 분석하기 위해서 오피니언 마이닝(Opinion Mining) 기법을 활용하였다.

오피니언 마이닝은 웹 사이트와 소셜 미디어에 나타난 여론과 의견을 분석하여 실질적으로 유용한 정보로 재가공하는 기술로, 텍스트를 분석하여 네티즌들의 감성과 의견을 통계/수치화하여 객관적인 정보로 바꿀 수 있는 기술이다. 오피니언 마이닝은 리뷰 데이터와 같은 대량의 정보 속에서 유용한 정보를 찾아낼 수 있고, 묻고 답하는 형식을 넘어 이용자들의 생각과 표현의 파편을 모아 일정한 법칙성을 찾아내어 새로운 의견 형성을 발굴하고 탐사하는 방식이다[24].

빅데이터 인프라 기술 중 널리 활용되는 기술은 대부분 오픈소스 형태로 데이터양의 증가에 발맞추어 지속적으로 발전하고 있다. 대표적인 인프라 기술은 분산 처리 시스템인 하둡, 데이터 마이닝 적합한 R, 비정형 데이터 처리에 강한 NoSQL 등이다. 본 연구에서는 오픈소스 프로젝트 R을 활용하였다.

R은 통계 계산 및 시각화를 위한 언어 및 개발 환경을 제공하며, R 언어와 개발환경을 통해 기본적인 통계 기법부터 모델링, 최신 데이터 마이닝 기법까지 구현/개선이 가능하다. 이렇게 구현한 결과는 그래프 등으로 시각화할 수 있으며, Java나 C, Python 등의 다른 프로그래밍 언어와 연결도 용이하다. 또 R의 경우에는 GNU 프로젝

트로 만들어져서 무료로 사용될 수 있고, 다양하게 활용할 수 있는 패키지들이 빠르게 추가되고 있어 최근 들어 더욱 주목을 받고 있다. 위의 장점들로 인하여 R은 통계 분석 분야에서 인지도를 높여왔으며, 하둡 환경 상에서 분산처리를 지원하는라이브러리 덕분에 구글, 페이스북, 아마존 등의 빅데이터 분석이 필요한 기업에서 대용량 데이터 통계분석 및 데이터 마이닝을 위해 널리 사용되고 있다[8].

## 2.2 택배서비스와 고객만족도

택배란 ‘사람이나 업체가 포장된 상품이나 물품 등을 요구하는 장소까지 직접 배달해 주는 것’을 의미한다[6]. 일반적으로 택배서비스는 기업과 기업 간 서비스(B2B), 기업과 개인 간 서비스(B2C), 개인과 개인 간 서비스(C2C) 형태를 보이고 있으며, 다양한 품종의 소량을 운송한다.

고객만족도란 고객이 어느 상품이나 서비스를 구매했을 때 구입한 상품과 서비스에 만족하는 정도이다. 상품의 품질과 성능만이 아니라 점원의 서비스, 애프터서비스 등도 포함되며 고객이 실제로 느낀 만족이 어느 정도 높은가 하는 것이다. 고객만족도가 높으면 고객에 의한 동일 상품의 구매율도 높아진다.

택배 등 서비스업과 고객만족도에 관한 선행연구는 서비스 품질을 측정하기 위해서 고객만족도를 활용하는 방식으로 학문적 연구가 시작되었다. PZB(1988)의 SERVQUAL모형[18]은 지금까지도 광범위하게 이용되어온 서비스 품질 측정도구로 많은 연구자들이 이를 토대로 분석 모형을 만들어 연구를 진행하였다. 선행연구(27건)들은 설문을 통해 고객만족도

및 만족요인을 분석하고 이를 토대로 서비스 품질 결정요인으로 경제성, 안전성, 정확성, 편리성, 신속성, 친절성 등을 제시하였다.

이러한 연구는 택배서비스 품질 개선을 위해서 중요한 요인이 어느 분야인지 밝히게 됨으로 택배산업의 변화와 혁신을 해야 하는 분야를 제시한다. 다만 그 분석 자료가 설문에 의한 수집된 자료이며 그 표본의 수가 약 100~300건으로 제한적이며, 설문조사를 위해서 많은 시간과 비용이 들어간다는 점이 단점으로 제시할 수 있을 것이다. 또한 시간과 비용의 제약성으로 실시간은 불가능하고, 주기적인 조사와 자료 수집 또한 어려운 실정이다.

## 3. 연구방법

빅데이터 분석 기술 중 R패키지를 활용하여 오피니언 마이닝을 실시하였다. 미국과 국내 택배업체 관련 소셜미디어 데이터를 추출하여 감성분석을 통해 고객들의 평가를 도출하고, 이를 ACSI와 NCSI의 실제 분석된 고객만족도와 비교하였다.

많은 소셜 네트워크 서비스 중 본 연구에 적합한 데이터 소스를 선정하고, 데이터를 수집한 후 URL 및 불용어를 제거하는 전처리과정을 수행한다. 그 다음 긍정어와 부정어로 구성된 감성사전을 구축하고, 감성사전을 기초로 스코어링 알고리즘을 만든 후 정확도 검사를 수행한다. 그 후 전처리 한 데이터를 극성분석을 수행하고 이를 바탕으로 극성정도 분석을 실시하여 ACSI, NCSI와 결과를 비교 분석한다. 위 수행과정은 Jeffrey Breen의 Mining twitter for Airline consumer sentiment[8] 과정을 참

고하였으며, 본 연구에서는 한글 트윗에 대한 분석과 스코어합수 결과값의 정확도 검사를 추가적으로 실시하였다.

<Table 1> Research Process Flow

1	Selection of Data source
2	Collection and preprocessing of Data
3	Setting and Accuracy testing of Sentiment Dictionary
4	Polarity Analysis for Data
5	Analysis: Degree of Polarity
6	Comparison with ACSI and NCSI

### 3.1 데이터 소스 선정

택배업체 고객만족도를 분석할 비정형데이터 획득을 위해 수집할 SNS 서비스 종류를 선정하였다.

소셜 네트워크는 전 세계적으로 트위터와 페이스북으로 대표된다. 국내에서는 카카오톡이 대중적인 인기를 얻고 있으나 미국 내 택배업체 고객만족도 분석에 활용하기에는 부적당하다. 유튜브, 인스타그램의 경우 동영상 및 사진 데이터를 위주로 하는 SNS로 본 연구에 데이터로 활용하기에는 제약이 많다. 트위터와 페이스북 모두 본 연구의 분석틀인 R패키지를 활용하여 쉽게 테스트 데이터를 가져올 수 있는 장점도 있다.

트위터는 개방형으로 누구나 트윗(tweet)에 접근할 수 있고, 글자 수 제한으로 인해 짧은 문장으로 의견을 표현하는 특징이 있다. 이러한 특징으로 인해 본 연구에서 개인별 긍정 및 부정의 감성을 분석하기 용이하다. 페이스북의 경우는 폐쇄형으로 공개된 글에 대해서만 접근이 가능하고, 이미지 형태의 글들이 많으며

글자 수 제한이 최대 420자로 특정 개인의 글에서 하나의 감정을 분석하는데 트위터에 비해서 상대적으로 어려움이 있다. 따라서 본 연구에서는 최근 트윗(tweet)을 Jeff Gentry에 의해 만들어진 Twitter R패키지[5]를 활용하였으며 수집하였다.

<Table 2> Characteristics of SNS

Classification	Twitter	Facebook	KakaoStory
Type	Open type	Closed type	Closed type
Characteristic	Sharing information	Human networking, Timeline	Human networking,
Remarks	Limited to 140 characters	Limited to 420 characters	Domestic SNS

### 3.2 데이터 수집/전처리

데이터 수집의 범위는 많은 미국과 국내 택배업체 중에서 차후 ACSI, NCSI와의 비교를 위하여 ACSI와 NCSI에서 고객만족도를 제공하는 택배분야 기업으로 한정하였다. 이 기업들은 해당국가의 주요 택배업체들로 고객들에게 많은 인지도를 가진 기업들이다.

데이터 수집은 트위터 앱 매니저 가입 후 TwitterR 패키지를 활용하여 키워드, 기간, 추출 MAX값을 설정하여 수집하였다. <Table 3>과 같이 데이터 추출 결과에서 볼 수 있듯이 미국에 비해 국내 택배업체 관련 트윗이 상대적으로 적고, 업체별로 차이가 많이 발생하여 사용 데이터는 미국 택배업체의 경우는 1,000건씩, 국내 택배업체의 경우는 100건씩 무작위로 선정하였다.

〈Table 3〉 Collection of Tweets Data

Classification	US			Korea			
	UPS	FedEx	USPS	Hanjin express	CJ korea express	Korea Post	Hyundai logistics
Keyword	“UPS”	“FedEx”	“USPS”	“한진택배”	“대한통운”	“우체국택배”	“현대택배”
Search period	2015. 1. 1~2016. 6. 20						
MAX Sampling	2,000						
Result	2,000	2,000	2,000	126	903	354	115
Using data	1,000	1,000	1,000	100	100	100	100

트위터 데이터는 특성상 분석 대상이 되는 텍스트 내용 이외의 아이디, URL, 문장부호, 숫자, 철자 오타, 불용어 등이 혼합되어 있어 컴퓨터를 통해 분석하기 어렵다. 따라서 텍스트 데이터를 컴퓨터가 처리하기 쉽도록 변환하는 작업인 전처리 과정이 필요하다. 데이터 전처리 과정은 다양한 비정형 데이터를 분석할 수 있도록 해주는 처리 단계라고 할 수 있다.

본 연구에서 데이터 전처리 과정은 아이디 제거, URL 제거, 문장부호 제거, 불용어 제거 등을 실시하였다. 전처리 과정의 예시는 <Table 4>

〈Table 4〉 Example of Data Preprocessing

Process	Contents
Text	“im_locked: 한진택배 또 경비실에 놓고감.. \n아니 저기... 저 집에 있어요 s://t.co/ppdlszgP4T”
Removing ID	“한진택배 또 경비실에 놓고감.. \n아니 저기... 저 집에 있어요 s://t.co/ppdlszgP4T”
Removing URL	“한진택배 또 경비실에 놓고감.. \n아니 저기... 저 집에 있어요”
Removing Punctuation mark	한진택배 또 경비실에 놓고감 \n아니 저기 저 집에 있어요
Removing Stopword	한진택배 또 경비실에 놓고감 아니 저기 저 집에 있어요

와 같다. 사용할 데이터 전체에 대해서 R패키지를 활용하여 데이터 전처리 실시하였다.

### 3.3 감성사전 구축/정확도 검사

트윗 문장을 단어 단위로 구성한 후 필요한 단어를 추출하기 위해서 긍정 및 부정어 사전을 구축하였다. 이는 트윗 문장이 주관적인 성향을 나타낸다는 가정 하에 연구를 진행한 것이다. 긍정 및 부정어 사전 내 단어와 트윗 문장의 단어와의 매치 여부를 확인 하여 문장 전체의 긍정 및 부정 성향을 파악한다. 분석을 위해 사용한 감성 사전은 미국 일리노이 시카고 대학[2]의 긍정 및 부정 사전을 활용하였다. 사전의 단어 수는 긍정어 4,783개, 부정어 2,006개이다. 한글 감성 사전은 오픈 소스를 활용하기 제한이 되어, 영문 감성 사전을 한글로 변환한 기초 자료를 바탕으로 중복문 삭제 및 택배 관련 의견에 주로 등장하는 긍정 및 부정어를 추가하여 자체적으로 사전을 구축하였다. 그 결과 한글 감성 사전의 단어 수는 긍정어 880개, 부정어 1,670개이다.

위에서 구축한 감성사전을 기반으로 하여 트윗 문장의 긍정 및 부정 스코어링이 정확하지 않음을 실시하였다. 수집된 데이터에서 미국업체 2곳과 국내 업체 2곳의 각각 100건씩의

데이터를 영문 및 한글 감성사전을 바탕으로 스코어 함수를 생성하여 분석한 결과 값과 실제 문장의 긍정 및 부정 결과를 비교하였다. 미국업체는 FedEx와 UPS의 트윗 100건씩을 비교 분석한 결과 영문 감성사전은 <Table 5>와 같이 80%, 84%로 평균 82%의 정확도를 보였다.

<Table 5> Accuracy test for English Sentiment Dictionary

Classification	FedEx	UPS
Tweets tested	100	100
Matching tweets	80	84
Rate	80%	84%

한국 업체는 우체국택배와 한진택배 트윗 100건씩을 비교 분석한 결과 한글 감성사전은 <Table 6>과 같이 50%, 47%로 평균 49%의 정확도를 보였다. 이는 인터넷상 사용되는 한글 특성상 띄어쓰기, 조사, 반어법, 비속어, 분야별 은어 등으로 인하여 감성사전을 바탕으로 한 긍정 및 부정 스코어링의 어려움을 나타내었다.

<Table 6> Accuracy test for Hangul Sentiment Dictionary

Classification	Korea Post	Hyundai logistics
Tweets tested	100	100
Matching tweets	50	47
Rate	50%	47%

미국 업체 트윗 분석 시에는 영문 감성사전을 바탕으로 한 스코어함수의 결과 값을 바탕으로 극성분석을 진행하였다. 한국 업체 트윗 분석 시에는 정확도가 상대적으로 너무 낮아 극성분석 결과의 신뢰성 향상을 위해서 본 연

구에서는 4개 업체 트윗 400건에 대해서 문장의 긍정 및 부정 성향을 확인하여 일부 값을 보정해 주었다.

### 3.4 데이터 분석

극성분석은 수집된 트윗 데이터가 주관적인 의견을 포함한다는 가정 하에 문장의 긍정 혹은 부정을 분류하는 과정이다. 본 연구에서는 미국 및 국내 택배업체 7곳에 대해서 극성분석을 실시하였다. 각 트윗에서 긍정 단어와 부정 단어의 빈도수가 동일하면 0점으로 중립으로 표현한다.

극성의 정도 분석은 각 택배업체에 대한 긍정, 중립, 부정 이미지의 비율을 분석하여 중립 이미지를 제외하고 긍정과 부정 이미지의 비율을 비교하여 점수화 하여 전체적인 업체별 고객의 만족도를 분석한다. 업체별 만족도는  $(\text{긍정 비율} / (\text{긍정 비율} + \text{부정 비율})) / 100$ 으로 환산한다.

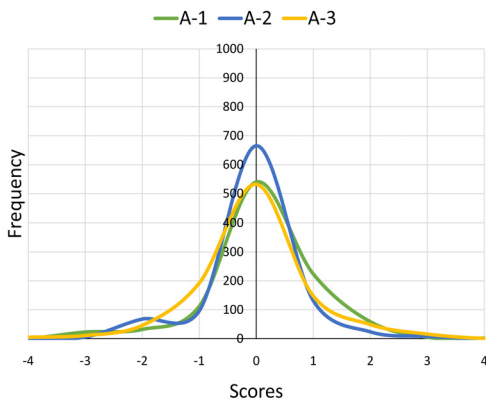
## 4. 데이터 분석 결과 및 비교

본 연구에서는 트위터 데이터 중 미국 주요 택배업체 3곳(UPS, FedEx, USPS)과 관련된 트윗 각 1,000건과 국내 주요 택배업체 4곳(한진택배, CJ대한통운, 우체국택배, 현대택배)과 관련된 트윗 각 100건을 기초 데이터로 하여 각 택배업체별 고객만족도를 분석하였고, 이를 ACSI 및 NCSI와 비교하였다. 분석결과에서 업체명 비공개를 위해 미국업체는 A-1, A-2, A-3으로, 국내업체는 K-1, K-2, K-3, K-4로 임의 표기하였다.

### 4.1 극성분석

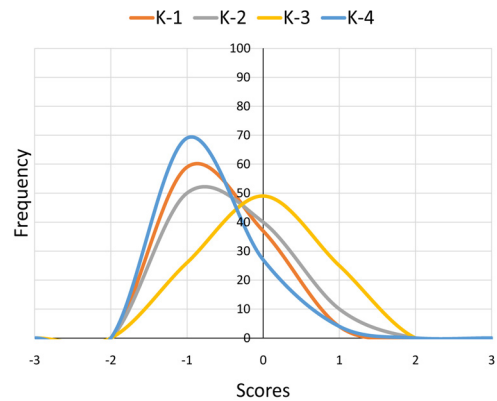
미국 택배업체의 스코어 점수를 정리하여 가로축은 각 업체별 스코어를 세로축은 빈도수를 나타내는 <Figure 1>과 같은 극성분석 그래프를 얻을 수 있었다.

극성분석 그래프가 우편향으로 나타나면 해당 택배업체에 대해서 긍정적인 이미지가 많다는 것을 의미하며, 좌편향으로 나타나면 부정적인 이미지가 많다는 것을 알 수 있다. 따라서 가장 긍정적인 이미지 택배업체는 A-1업체이며, 상대적으로 가장 부정적인 이미지 업체는 A-3업체임을 알 수 있다. 또한 분석한 택배업체 대부분이 0을 기준으로 균등하게 분포되어 있는 것으로 유추해 보아 미국 택배업체에 대한 소비자들의 전반적인 이미지가 중립적임을 알 수 있다.



<Figure 1> Polarity Analysis Result for American Courier Companies

국내 택배업체의 스코어 점수를 정리하여 가로축은 각 업체별 스코어를 세로축은 빈도수를 나타내는 <Figure 2>와 같은 극성분석 그래프를 얻을 수 있었다.



<Figure 2> Polarity Analysis Result for Korea Courier Companies

국내 택배업체 중 가장 긍정적인 이미지는 K-3 업체이며, 가장 부정적인 이미지는 K-4 업체임을 알 수 있다. 또한 분석 대상 중 K-3업체를 제외한 대부분의 택배업체들이 0을 기준으로 좌측에 분포되어 있는 것으로 유추해 보아 고객 대부분이 택배업체에 대해서 부정적인 이미지를 가지고 있다는 것을 알 수 있다.

### 4.2 극성의 정도분석

극성의 정도 분석은 트윗 문장에 대해 긍정적인 정도와 부정적인 정도를 측정하는 분석이다. 극성분석의 결과를 바탕으로 각 트윗 문장의 스코어 값을 확인하여 1 이상은 긍정, 0은 중립, -1 이하는 부정으로 구분하였다.

<Table 7>은 트윗 분석을 통하여 나타난 고객들의 택배업체에 대한 긍정, 중립, 부정의 이미지 비율을 나타낸다. 정확한 비교를 위해서 중립을 제거하고 긍정 이미지와 부정 이미지의 빈도를 바탕으로 각 기업별 긍정 및 부정 이미지 점수를 산출하였다.



<Table 7> Polarity Analysis Result

Classification		Negative (%)	Neutral (%)	Positive (%)
US	A-1	16.8	54.0	29.2
	A-2	17.0	66.5	16.5
	A-3	25.4	53.2	21.4
Korea	K-1	59.0	37.0	4.0
	K-2	50.0	40.0	10.0
	K-3	27.0	49.0	24.0
	K-4	69.0	27.0	4.0

<Table 8> Score of Customer Preference

Classification		Negative	Positive	Total	Score
US	A-1	16.8	29.2	46.0	63.5
	A-2	17.0	16.5	33.5	49.3
	A-3	25.4	21.4	46.8	45.7
Korea	K-1	59.0	4.0	63	6.3
	K-2	50.0	10.0	60	16.7
	K-3	27.0	24.0	51	47.1
	K-4	69.0	4.0	73	5.5

\* Score: (Positive/(Positive+Negative))/100.

미국 택배업체 중 가장 높은 점수를 받은 업체는 A-1업체로 63.5점이며, 국내 택배업체 중 가장 높은 점수를 받은 업체는 47.1점의 K-3업체이다.

### 4.3 국가고객만족도와 비교

매년 ACSI와 NCSI는 소비자 인터뷰를 통해 얻은 자료를 분석하여 각 산업 분야별 기업별 상품 및 서비스에 대해 고객이 평가한 만족의 정도를 측정해 계량화한 지표를 공시하고 있다. ACSI는 실질적으로 미국시장에서 점유율을 가지고 있는 외국 기업과 국내 기업에 대

해서 제품과 서비스의 질에 대한 고객의 평가를 기초로 고객만족도를 측정하여 서비스 만족도를 비교할 수 있도록 자료를 제공하고 있다. 우리나라에서도 미국의 ACSI 모델을 기초로 하여 국내 실정에 맞도록 NCSI 모델을 만들어서 동일한 측정방법론을 적용하여 국내 산업 및 기업들을 대상으로 고객만족도를 측정하여 발표하고 있다.

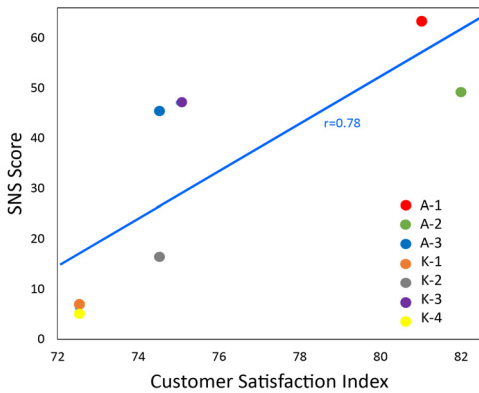
<Table 9> ACSI / NCSI

Classification		Year 2015	Year 2016	Average
US	A-1	82	80	81
	A-2	82	82	82
	A-3	75	74	74.5
Korea	K-1	72	73	72.5
	K-2	74	75	74.5
	K-3	74	76	75
	K-4	72	73	72.5

Source: <http://www.theacsi.org>,  
<http://www.ncsi.or.kr>.

<Table 9>는 택배업체들의 국가고객만족도를 나타낸다. A-2업체가 미국 내에서는 평균 82점으로 가장 높은 만족도를 보이며, 상대적으로 A-3업체는 평균 74.5점으로 낮은 고객만족도를 보인다. 국내 택배업체 중 K-3업체가 75점으로 가장 높은 만족도를 보이며, K-1, K-4업체가 상대적으로 낮은 만족도를 보인다.

트윗 문장 극성정도 분석 결과와 국가고객만족도 두 데이터 사이의 상관관계를 분석하였다. <Figure 3>의 X축은 국가고객만족도를 나타내고, Y축은 극성분석을 통해 도출된 점수를 나타낸다. 극성정도분석 결과와 고객만족도 간의 상관계수는 0.78로 트윗을 분석한 긍정 및 부정 이미지 점수와 실제 고객만족도가 뚜렷한 양적 선형관계를 형성한다.



〈Figure 3〉 Correlation of Customer Satisfaction Index and SNS analysis score

이러한 결과는 실시간으로 나타나는 소셜 네트워크의 정보를 이용하여 택배업체들의 고객만족도를 예측하는데 활용될 수 있다는 것을 의미한다.

## 5. 결 론

빅데이터의 관심 및 활용도는 점점 증가하고 있고, 실제 우리의 생활 가까이 와 있다. 온라인 포털 및 쇼핑사이트에 들어가지만 하여도 과거에 검색했던 혹은 클릭했던 상품 정보 및 광고들의 팝업을 쉽게 볼 수 있다. 택배업계에서도 이러한 내부 데이터를 활용하여 추가적인 서비스를 제공하고 있다. 예를 들어 배송 추적 서비스, 안내 메시지 등이다. 택배 서비스에 대한 고객들의 만족도 확인 및 서비스 품질 측정을 위한 소셜 네트워크 빅데이터의 활용은 미약하다.

기존의 많은 선행연구에서 택배서비스 품질 측정 및 고객만족도에 관하여 설문조사를 기초자료로 하여 분석을 실시하였다. 이는 비용

과 시간이 많이 들어가며 많은 표본의 수를 확보하기에 어려움이 있다. 설문조사를 대신하여 SNS 빅데이터를 활용한다면 고객만족도 분석 시스템을 구축할 수 있을 것이고, 이를 통해서 분기, 월별, 혹은 주 단위, 일 단위, 더 나아가 실시간 고객들의 만족도를 확인 및 측정할 수 있을 것이다.

이러한 가정을 확인하기 위해서 본 연구는 SNS 빅데이터를 활용하여 택배업체 고객만족도를 분석에 활용할 수 있다는 가능성을 제시하였다. 본 연구에서는 빅데이터 분석 기법과 택배서비스, 고객만족도에 대해서 이론적 고찰을 선행하였다. 그 후 연구를 설계하고, SNS 빅데이터를 수집하고, 실험을 진행하여 산출한 결과 값과 실제 고객만족도와 비교하였다. 그 결과 각 택배업체들의 SNS 트윗 극성 정도분석 결과와 실제 공인기관에서 발표되는 고객만족도는 강한 상관관계를 보였다. 실시간의 생성되는 소셜 네트워크의 데이터상의 소비자 감정 정보는 향후 고객만족도 결과를 예측하는데 활용할 새로운 정보로 이용 가능하다는 점을 검증하였다.

하지만 트윗 문장 자체가 주관적이라는 가정 하에 연구를 진행하여 결과의 신뢰성에 한계가 있다. 영문 감성사전의 경우는 비교적 높은 정확도를 보여 많은 데이터 실시간 처리도 가능하다는 것을 확인하였으나 한글 감성사전 구축 및 분석 간에 여러 제한 사항이 있었다. 우선 국내 택배업체 관련 트윗 데이터양이 적었고, 한글의 특성과 인터넷 상의 은어, 약어 등으로 인하여 공부정 점수화 과정에서 낮은 정확도를 보여 일부 값을 보정작업 후 비교 분석을 실시하였다. 최근에는 SNS 빅데이터를 활용하기 위해서 자연어 처리 기법 및 각 분야

별 특징을 반영한 감성사전 구축이 많이 진행되고 있으나 택배 및 물류분야의 감성사전 구축은 이루어지지 않고 있어 실시간 소셜 네트워크의 방대한 소비자 데이터를 분석하기 위해서는 추가적인 연구가 필요하다.

또한 본 연구에서는 세부적인 고객만족도 및 만족요인별로 구체적인 분석을 하지는 못하였다. 그리고 1회성이 아닌 주기적인 분석을 위해서는 고객만족도 분석 시스템을 구축해야 하는데 연구의 시간 및 비용 등 여러 제약조건으로 제한되었다. 차후 온라인상의 고객들의 다양한 의견인 빅데이터 기반으로 택배서비스의 고객만족도 분석을 위해서는 체계적인 분석 시스템 구축이 필요하다.

최근 사회·경제 전 분야에서 빅데이터의 활용이 두드러지고 있으며, 특히 고객 마케팅에 많은 활용을 하고 있다. SNS 및 웹사이트의 빅데이터를 활용하여 택배서비스 품질을 측정 및 예측할 수 있다면 이를 통해 택배산업의 전반적인 서비스 품질 향상을 도모할 수 있을 것이다. 본 연구에서는 이러한 빅데이터의 활용 가능성을 제시하였으며, 앞으로 택배산업의 발전을 위해서 빅데이터 기반 택배서비스 품질 측정 및 고객만족도 분석에 관한 많은 연구가 진행되기를 희망한다.

---

## References

---

- [1] Bae, K. B., "The Analysis of Economic Effects of Door to door transport Industry," Korea Logistics Review, Vol. 22, No. 1, pp. 163-181, 2012.
- [2] <http://www.cs.uic.edu>.
- [3] <http://www.ncsi.or.kr>.
- [4] <http://www.theacsi.org>.
- [5] <https://cran.r-project.org/web/packages>.
- [6] <https://ko.wikipedia.org>.
- [7] <https://mran.revolutionanalytics.com/posts/twitter.html>.
- [8] Kim, J. S., "Utilize big data and analytics review," Korea University MA Dissertation, 2012.
- [9] Kim, J. W., Lee, S. G., and Choi, J. H., "Measuring Parcel Service Quality," Korean management review, Vol. 31, No. 2, pp. 283-294, 2002.
- [10] Kim, M. K. and Jung, H. S., "An Analysis of Customer Perception on the Differentiated Parcel Service using Kano Model and Timko's Customer Satisfaction Coefficient: Focusing on Online Shopping Customers," Korean Journal of Logistics, Vol. 24, No. 2, pp. 33-47, 2016.
- [11] Koo, H. Y., "A Method of Economic Analysis for RFID Applications Focused on Courier Sector," Journal of Society for e-Business Studies, Vol. 15, No. 1, pp. 119-137, 2010.
- [12] Lee, C. K. and Lee, M. K., "Antecedents of Satisfaction with Internet News Sites," Journal of Society for e-Business Studies, Vol. 9, No. 3, pp. 17-37, 2004.
- [13] Li, B., Riley, M. W., and Hsieh, C.-T., "Assessing customer satisfaction in parcel delivery industry: An empirical study among university customers," Interna-

- tional Journal of Services and Standards, Vol. 1, No. 2, pp. 172-192, 2004.
- [14] NIA, "3 elements for successful utilizing Big Data," Seoul, 2012.
- [15] NIA, "Technology classification and status of Big Data," Seoul, 2013.
- [16] Oh, P. H. and Hwang, B. Y., "Real-time Spatial Recommendation System based on Sentiment Analysis of Twitter," The Journal of Society for e-Business Studies, Vol. 21, No. 3, pp. 15-28, 2016.
- [17] Oh, S. G. and Lee, C. S., "The effect of the quality determinant of Korea-China Courier Service on Customer value and Satisfaction," Korea research academy of distribution and management review, Vol. 19, No. 4, pp. 5-19, 2016.
- [18] Parasuraman, A., Zeithaml, V. A., and Berry, L. L., "SERVQUAL-A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," Journal of Retailing, Vol. 64, No. 1, pp. 12-40, 1988.
- [19] Park, J. H. and Lee, S. H., "The Effects of Logistics Service Quality, Relationship Quality among Customer Loyalty in Courier Company," Korea Marketing Review, Vol. 23, No. 4, pp. 23-50, 2008.
- [20] PWC, "Big Data: The next frontier for innovation," Price waterhouse Coopers, 2012.
- [21] Seo, M. K., Kwon, J. H., and Choi, Y. J., "Impact of Logistics Service Quality on Customer Satisfaction and Loyalty in Parcel Service," Korea Logistics Review, Vol. 23, No. 5, pp. 239-262, 2013.
- [22] Seo, S. M. and Lee, J. H., "An Introduction on the Customer Reaction Encouragement System for E-Commerce Customer Satisfaction: Extension of Customer Response System," Journal of Society for e-Business Studies, Vol. 9, No. 1, pp. 237-254, 2004.
- [23] Song, J. G. and Kim, T. R., "The Performance Formation Model of Service Quality Factors for Courier Service," Journal of Distribution Science, Vol. 10, No. 4, pp. 37-45, 2012.
- [24] Yoon, H. K., "Research on the Application Methods of Big Data within the Cultural Industry," Academic association of global cultural contents, No. 10, pp. 157-180, 2013.

## 저 자 소개



이동준 (E-mail: dongjun@krii.re.kr)  
2005년 육군사관학교 국제관계/군사학 (학사)  
2016년~현재 UST 교통물류 및 ITS 공학과 석사과정  
관심분야 물류운영, 데이터마이닝, 군물류 등



원중운 (E-mail: juwon@krii.re.kr)  
1998년 한국해양대학교 제어계측공학 (석사)  
2004년 경북대학교 전자공학 (박사)  
2005년~현재 한국철도기술연구원 책임연구원  
2015년~현재 UST 교통물류 및 ITS 공학 겸임교수  
관심분야 IoT, 영상처리, 인공지능, 빅데이터 등



권용장 (E-mail: yjkwon@krii.re.kr)  
1997년 한양대학교 경영학 (석사)  
2007년 한양대학교 경영학 (박사)  
1997년~현재 한국철도기술연구원 수석연구원  
2004년~2007년 건설교통부 물류정책자문관  
2011년~2012년 기술표준원 국가물류표준코디네이터  
2012년~현재 UST 교통물류 및 ITS 공학 겸임교수  
관심분야 국가물류표준화, 물류기술개발, 철도물류 등



김미례 (E-mail: heyday2015@krii.re.kr)  
2014년 UST 교통물류시스템 (석사)  
2015년~2016년 한국철도기술연구원 석사후연수연구원  
2016년~현재 UST 교통물류 및 ITS 공학과 석사과정  
관심분야 철도물류, 교통공학, 철도토목 등