J. Soc. Korea Ind. Syst. Eng Vol. 41, No. 3 : 129-137, September 2018
https://doi.org/10.11627/jkise.2018.41.3.129
ISSN : 2005-0461(print)
ISSN : 2287-7975(online)

Analyzing Customer Feedback Differences between VOCs and External Channels

Sang Hyeon Ahn* · Dong Hyun Baek**†

*Korea Productivity Center
**Department of Business Administration, Hanyang University ERICA

VOC와 외부채널간의 고객 피드백 차이 분석

안상현* · 백동현**†

*한국생산성본부
**한양대학교 경상대학 경영학부

VOCs have been used as the most definitive resource to reflect customer feedback when developing products and services. However, due to the development of the Internet and the emergence of SNS, VOC is no longer the only channel that represents customer opinions. There are also a number of studies showing that many customers express complaints through channels other than VOCs. In this paper, we analyze the difference between the official VOC data and the data collected through the external channel, and suggest ways to reflect the various opinions of customers. To do this, this study uses keyword analysis that can identify differences according to frequency through social network, modular analysis to distinguish topics according to centrality and similarity, and emotional analysis to confirm word polarity (positive and negative). The results of this study show that the opinions of the customers were different depending on channels such as VOCs and external channels. Therefore, the collected data through VOC as well as external channels should be used in order to reflect the opinions of customers. In particular, this paper confirms that the results of one channel may vary depending on the channel characteristics even for the same channel. This confirms that collecting voc only on certain channels may differ from what real customers require. Therefore, data collected through VOCs as well as external channels must be used to reflect various customer feedback.

Keywords : VOC, Customer Opinion, Social Network Analysis, Keyword Analysis

1. 서 론

기업의 공식적인 채널을 통해 수집되는 VOC(Voice Of Customer, 고객의 소리 이하, VOC)는 수많은 기업들이 제품 및 서비스의 질을 높이기 위해 주요한 자료로써 활용되어 왔다. 기업들이 품질기능전개(Quality Function Deployment)와 전사적 품질경영(Total Quality Management)을 이용하여 고객의 요구를 잘 반영한 맞춤형 제품이나 서비스가 기업의 수익에 영향을 미치자 VOC의 비중은 지속적으로 증대되어 왔다.

그러나 인터넷과 같은 다양한 매체의 발달은 VOC가 더 이상 고객의 의견을 대표하는 채널로서의 기능을 떨어뜨렸다. 고객들은 제품이나 서비스에 대하여 말할 때, 공식적인 채널 보다는 다른 공간에서 말하고 있다. SNS의 등장은 이러한 현상을 더욱 가속화 시키고 있다.

기업들은 이러한 문제를 해결하기 위하여 인터넷 커뮤

Received 2 July 2018; Finally Revised 5 September 2018; Accepted 6 September 2018
† Corresponding Author : estarbaek@hanyang.ac.kr

니티나 공식 SNS 사이트를 개발하는 등의 노력을 기울이고 있지만 문제는 해결되지 않고 있다. 이러한 문제점은 지속적으로 지적되어 왔으며 여러 가지 방법으로 해결방안이 강구되기도 하였다. 최근 빅 데이터의 등장으로 기업들이 이를 이용하여 VOC를 개선하려는 노력이 진행되고 있다. 고객관리를 위한 전담 시스템 및 부서 개발, 사전경보시스템 개발들이 이에 해당한다.

그러나 이러한 노력들은 VOC와 외부채널과의 비교를 통한 차이점 분석 보다는 VOC의 발생원인, VOC 내부시스템 개선에 관련된 부분이 주를 이루고 있다. 학계에서도 이러한 경향이 주를 이루어 나타나고 있다. 국내 연구들을 보더라도 고객들의 의견을 정확하게 반영하고 있지 못하다는 점에서는 공통적으로 일치를 이루고 있으나 그 해결방안에 대해서는 주로 VOC 시스템 내부 개선이나 새로운 시스템 개발, 또는 새로운 평가지표를 반영하는 방법이 대부분을 차지하고 있다. 현재 VOC가 가지고 있는 가장 큰 문제들 중 하나는 VOC가 고객들의 의견을 반영하는 대표성이 떨어진다는 부분이다.

하지만 위에서 언급한 해결방안들은 VOC 시스템 내부를 개선하는 부분에서는 긍정적일 수 있으나, 대표성 문제를 해결한다고는 보기는 어렵다. 다른 채널들의 의견을 반영하여 VOC 시스템에 반영하는 방법이 아니기 때문이다. VOC에 대한 문제를 해결하기 위해서는 앞서 언급한 대표성 문제를 고려해야 한다. 본 연구의 목적은 VOC의 대표성 문제를 보완하여 고객의 다양한 의견을 반영할 수 있는 방법을 제시하고자 하였다.

본 연구에서는 기업이 인터넷 채널을 통해 공식적으로 고객의 의견을 수집하는 VOC 자료와 해당 주제에 대하여 다른 인터넷 채널 자료를 비교하여 차이점을 분석하기로 하였다. 분석을 위한 방법은 세 가지로 사회연결망 분석, 모듈성 분석, 감성 분석이다. 각각 빈도수, 중심성, 극성(긍정과 부정)을 통해 결과를 비교하고 이를 공통부분과 비 공통부분으로 나누어 차이점을 분석하였다.

본 연구는 고객의 다양한 의견을 반영하지 못하는 VOC의 문제점을 보완하기 위해 진행하였다. 본 연구의 결과는 VOC를 활용하여 제품 및 서비스 개발 및 개선을 해야 하는 기업에 참고할 자료로서 도움을 줄 수 있을 것이다.

2. 선행연구와 이론적 배경

2.1 VOC

VOC는 1972년 고베 미츠비시 선박회사가 공정의 효율성과 제품의 완성도를 위해 개발하여 1970년대 도요타에서 활용된 품질기능전개의 한 요소이다. 품질기능전개는 제조공정 전 단계에서 고객의 요구를 반영한다. 그렇기 때문에 VOC는 품질기능전개에서 고객의 요구를 파악하는데 있어 가장 중요한 역할을 차지한다. 품질기능전개의 프로세스가 약 60%의 디자인 제작비용과 40%의 디자인 제작 시간 감소를 가져오자 1986년 포드와 제녹스가 이 프로세스를 적극적으로 활용하기 시작했다. 이후 품질기능전개가 확대된 형태인 전사적 품질경영 다수의 미국기업들이 활용하기 시작했다. 이를 확인할 수 있는 방법에 하나인 Malcolm Baldridge 모델(말콤 볼드리지 모델, 이하 MB 모델)을 들 수 있다. 1990년 MB 모델의 최소 요구조건이 18만 건에서 1991년 19만 건으로 증가했지만 106개의 미국기업이 MB 모델 수상에 도전했다. 이는 다수의 미국 기업들이 전사적 품질경영에 초점을 맞추고 있다는 사실을 증명하는 결과이다.

거시적 관점의 마케팅 기법에서 품질기능전개는 시장조사의 여러 가지 관점 중 기능적 측면에서 도움을 준다. 이는 전사적 품질경영 역시 마찬가지이다. 그리고 이러한 기능들은 VOC에서 나온 데이터를 통해 각각의 이용정도와 수요를 측정할 수 있다.

그러나 VOC가 각광받기 시작하면서 이에 대한 문제점을 지적하는 연구도 발표되었다. Grrifin[10]이 휴대용 식기류를 사용하고 있는 고객들 중 30명을 무작위로 선발하여 해당 제품에 대한 인터뷰 자료를 분석하였다. 그 결과 VOC에서 고객들이 불만사항을 언급하는 빈도수와 실제 제품이 중요하게 고려해야 할 부분이 불일치한다는 사실을 발표했다.

Singh[42]은 고객이 사용하는 제품이나 서비스가 불만족스러운 경우, 이에 반응하는 고객의 유형을 연구하였다. 그 결과, 직접적으로 해당 기업의 VOC 채널에 항의를 하는 경우는 전체 37%에만 해당한다는 결과를 발표하였다.

국내에서는 이수열 외[33]가 소비자 트렌드 변화를 고려하지 않고 일반적인 분석 자료로써 활용되고 있는 VOC의 문제점을 지적하였다. 그리고 이를 해결하기 위해서 소비자 트렌드에 따라 데이터베이스를 재구축 한 후, 상품 및 월별 세부 분석을 통해 기업이 취할 수 있는 상품관리와 시설관리 등 제반 조치사항을 준비하도록 하는 사전경보 VOC 시스템을 제안하였다.

우지영 외[44]는 다양한 고객의 의견을 체계적으로 분석하지 못하는 VOC 시스템의 문제점을 지적하고 해결방안으로 고객의 소리 발생 원인을 분석하여 사후관리 뿐 아니라 사전에 예방할 수 있도록 발생 변화 요인에 따라 기존 요인에 VOC 발생에 영향을 많이 미치는 외부 환경적인 요인을 추가로 분석해야 한다고 주장하였다.

2.2 사회연결망 분석(Social Network Analysis)

사회연결망 분석은 개체들 간의 사회적 관계를 구조적으로 분석하여 내재된 관계를 파악하는 것을 말한다. 사회연결망 분석을 활용하면 다양한 사건 혹은 집단의 구성원들을 점으로 표현하고 이들 사의 관계를 선으로 나타내어 구성원 사이의 의사소통이나 구성원 간 영향을 미치는 정도를 정량화 하여 분석하는 것이 가능하다. 본 연구에서는 사회연결망 분석의 하나인 키워드 분석(Keyword Analysis)을 활용하기로 하였다. 키워드 분석은 개체들 사이의 중심성(노드와 노드 사이의 인접 정도와 영향성) 정도를 측정하여 빈도순으로 결과를 정리하는 방법이다.

2.2 모듈성 분석(Modularity Analysis)

모듈성 분석은 단어들 간의 유사도(similarity)를 계산해서 유사한 단어들을 몇 개의 집단으로 묶어주는 방법이다. 텍스트 데이터에 전통적인 군집 분석을 적용하는 것으로 문서 또는 텍스트 클러스터링(document, text clustering)이라고도 한다. 모듈성 분석은 유사중복문서 검출(near duplicates detection : 문서가 완전히 일치하지는 않으나 많은 범위 내에서 일치하는 문서), 검색 엔진 최적화(search engine optimization : SEO, 웹 사이트가 검색엔진에서 검색 시 상위에 링크되도록 하는 작업) 등에 응용되고 있다.

2.3 감성 분석(Sentiment Analysis)

감성 분석이란 문서를 작성한 사람들의 태도, 의견, 성향과 같은 주관적인 데이터를 가지고 특정 주제에 대해 긍정인지 또는 부정인지를 분류하는 기술로 오피니언 마이닝(Opinion mining)이라고도 한다. 감성 분석은 문서를 작성한 사람의 감정을 추출해 내는 분석 방법이기 때문에 문서의 주체가 무엇인지 추출하기보다는, 그 문서를 작성한 사람이 주제에 대하여 어떠한 감정을 가지고 있는가를 판단하는 분석이다. 현재까지 감성 분석이 주로 적용되는 분야로는 주로 온라인상에서 영화평, 도서평, 상품평 등에 대한 분석이고, SNS 등장 이후로는 선거, 주식시장 예측 등의 다양한 분야로 그 활용이 확대되고 있다. 감성 분석에는 다양한 기법이 있으나 본 연구에서는 Ridge(능형)와 Lasso (Least absolute shrinkage and selection operator) 회귀모형을 이용한 예측 분석(Elastic net)을 활용하기로 하였다.

3. 분석 방법

<Table 1>은 VOC 자료와 외부채널 간 자료 분석에 대한 절차를 6단계로 정리한 것이다. 그림에서 보는 바와 같이 6단계는 각각 분석방향 설정, 자료수집, 키워드 분석, 모듈성 분석, 감성 분석, 분석결과 비교로 이루어져 있다.

<Table 1> Analysis Process

Step 1.	Analysis Direction
Step 2.	Collecting Data
Step 3.	Keyword Analysis
Step 4.	Modularity Analysis
Step 5.	Sentimental Analysis
Step 6.	Compare Results

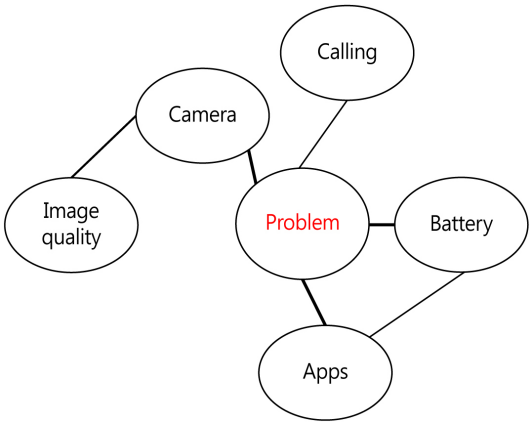
본 논문에서는 '제주도 관광'이라는 주제로 비교 대상 VOC 채널에서 빈도수, 군집정도, 긍정과 부정의 의미에 대하여 각각에 해당하는 기법들을 활용하여 차이가 있는지에 대해서 분석해보고자 하였다.

분석방향설정에서는 전체적인 분석의 틀을 정하는 단계이다. 본 연구에서는 기존 VOC 자료와 외부채널간의 비교를 통한 자료 간 차이점 분석을 실시하여 다양한 의견을 반영할 수 있는 VOC의 방향에 대해 연구하고자 하였다.

자료수집에서는 VOC와 외부채널 자료에 대한 수집방법을 정하는 단계이다. 본 연구에서 VOC는 기업에서 공식적으로 운영하는 인터넷 채널을 대상으로 하였다. 외부채널 역시 인터넷자료인 포털사이트와 SNS으로 한정하기로 하였다.

수집된 자료들은 총 3가지 분석을 거친다. 첫 번째로 키워드 분석에서는 각각의 자료를 주제와 관련된 단어들로 변환한 다음 빈도수를 기준으로 비교 분석을 실시한다.

<Figure 1>은 키워드 분석결과를 예시로 나타낸 것이다. 특정 단어의 빈도수가 타 개체에 비해 높고 상관 비율이 높은 해당 단어는 네트워크 가운데에 위치하여 연결선이 굵게 표시된다.

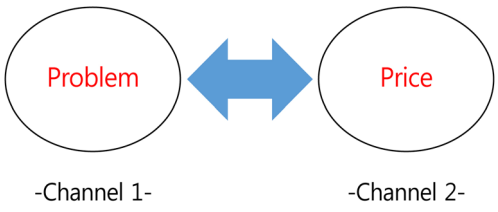


<Figure 1> Keyword Analysis

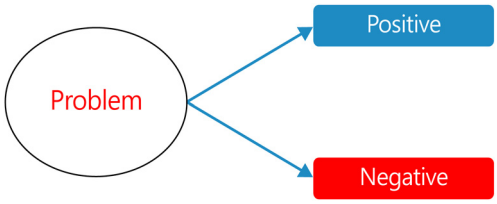
모듈성 분석에서는 단어 간의 유사도를 측정한 뒤, 중심성(centrality)을 기준으로 모듈화(군집화)시켜서 각각의 자료를 비교한다. <Figure 2>는 모듈성 분석결과를 예시로 나타낸 것이다. 모듈성 분석을 통해 공통주제에 대한 각 채널의 내용들을 확인한다.

감성 분석에서는 앞선 분석을 통해 추출된 단어들을 긍정단어와 부정단어로 나누고 이에 대한 회귀계수를 측정한다. 이를 통해 일반적으로 쓰이는 단어가 특정주제에서 긍정인지 부정인지를 확인 하고 앞서 실시한 두 분석결과와 비교하여 분석들의 신뢰성을 측정한다.

분석결과비교에서는 세 가지 분석결과를 나열하여 자료간의 특성을 확인하는 단계이다. 여기서는 각각의 자료에서 공통으로 도출된 부문과 한 채널에서만 도출된 비 공통부문으로 나누어 비교분석을 실시한다.



<Figure 2> Modularity Analysis



<Figure 3> Sentimental Analysis

4. 사례 분석

4.1 분석방향 설정

본 연구에서는 제주도 관광에 대한 문제점을 분석하기 위해 제주관광공사와 제주도에서 공식적으로 운영하는 VOC 채널인 '제주관광신문고'와 외부채널과의 비교분석을 통해 채널 간 특징을 파악하고 다양한 의견을 반영할 수 있는 VOC의 방향에 대해 알아보고자 하였다.

4.2 자료수집

VOC자료는 제주관광신문고에 있는 공시 자료를 대상으로 2014년 1월부터 2015년 12월까지의 자료 200건을 수집하였다. 외부채널 자료의 경우, 국내 포털사이트(네이버, 다음 등)와 SNS로 한정하고 VOC 자료와 동일 기간 동안 웹 스크래핑(Web-Scrapping)방식을 통해 자료를 수집하였다. 자료 수집 시, 검색어로서 긍정문서는 '제주도 관광 추천'이라는 키워드로, 부정문서는 '제주도 관광 문제'라는 키워드로 검색했다. 검색 결과, 각각의 자료 99건씩 198건을 수집하였다.

4.3 키워드 분석

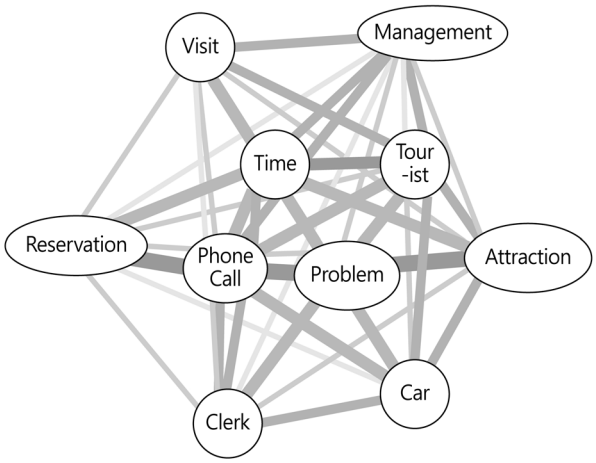
4.3.1 제주관광신문고

제주관광신문고에서 수집한 VOC 자료 200건에 키워드 분석을 실시한 결과 '제주도 관광 문제'에 대하여 언급한 단어의 총 개수는 162개였다. 해당 단어들의 빈도수를 모두 합한 건수는 2,820건으로 나타났다.

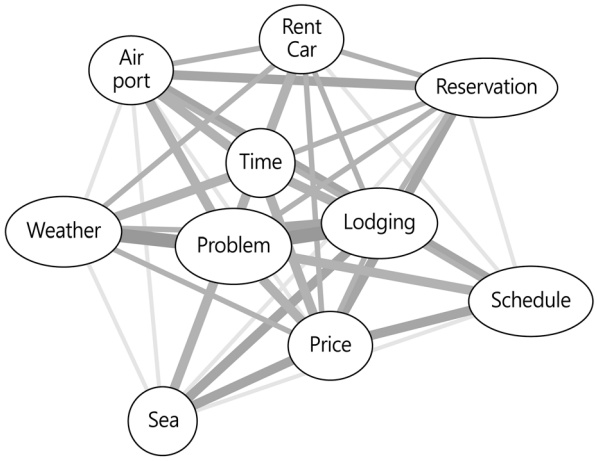
<Table 2> Jeju Tourism Complaints Keyword Analysis

No.	Jeju Tourism Complaints	Frequency
1	Time	51
2	Tourist	46
3	Phone Call	42
4	Problem	41
5	Visit	37
6	Attractions	32
7	Management	31
8	Reservation	29
9	Clerk	28
	Car	28

<Table 2>는 제주관광신문고의 키워드 분석결과를 상위 10개 키워드를 중심으로 정리한 것이다. 키워드 분석은 각각의 채널에서 나온 내용(문장)에서 명사를 기준으로 나온 단어들을 빈도수에 따라 내림차순으로 정리한 것이다. 표에서 보는 바와 같이 제주관광신문고에서 가장 많이 언급된 단어는 시간이었으며, 다음으로는 관광객, 전화, 문제, 방문 순이었다. 그리고 위 결과를 바탕으로 시각화 작업을 실시한 결과는 <Figure 4>와 같다. 그래프에서 보는 바와 같이 상위 10개 키워드들 중에서 가장 높은 빈도수를 보이는 시간은 전화, 문제, 관광객과 같은 단어들과의 연결선이 굵게 나타나고 있어 해당 키워드들과 연관성이 높다는 사실을 확인 할 수 있다. 또한 문제를 중심으로 연결망을 살펴보면, 차량, 관광지, 직원과 같은 단어들이 연관성이 높다는 사실을 확인 할 수 있다. 이를 통해 제주관광신문고에서는 시간이나 관광객 관련 내용이 자주 언급되고 있음을 확인 할 수 있다.



<Figure 4> Jeju Tourism Complaints Keyword Analysis Visualization



<Figure 5> Web-Scrapping Keyword Analysis Visualization

4.3.2 웹 스크래핑

제주관광신문고와 동일한 방법으로 외부채널을 웹 스크래핑을 통해 수집한 99건의 자료에 대해 키워드 분석을 실시하였다. 분석결과 단어의 총 개수는 43개이고 해당 단어들의 빈도수 총합은 788건으로 나타났다. <Table 3>은 웹 스크래핑의 키워드 분석결과를 상위 10개 키워드를 기준으로 정리한 것이다. 웹 스크래핑에서 가장 많은 빈도수를 기록한 단어는 문제였으며 그 뒤로는 시간, 숙소, 사격, 렌터카 순이었다.

<Table 3> Web-Scrapping Keyword Analysis

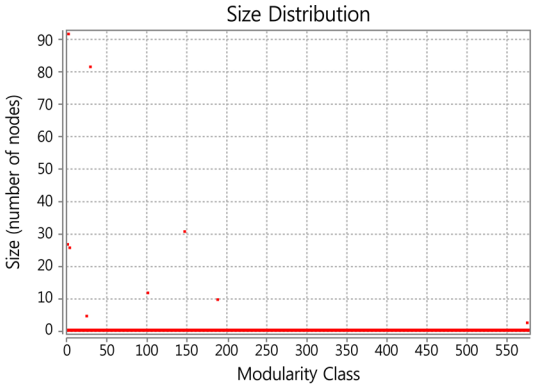
Division	Web Scrapping	Frequency
1	Problem	67
2	Time	57
3	Lodging	40
4	Price	33
5	Rent Car	32
6	Airport	28
	Weather	28
8	Reservation	26
9	Sea	25
10	Schedule	22

웹 스크래핑 역시 각각의 단어들의 연관성을 확인하기 위해 시각화 작업을 실시하였다. 제주관광신문고 자료와 비교하여 볼 때, 시간과 문제가 중심에 위치해 있다는 것은 동일하지만 그 외에 일치하는 단어가 예약밖에 없다는 것을 확인 할 수 있다. 또한 문제를 중심으로 연관성이 높은 단어가 숙소, 가격, 날씨로 제주관광신문고와 차이가 있다는 것을 확인 할 수 있다.

4.4 모듈성 분석

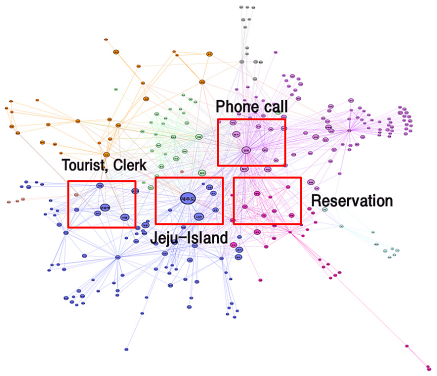
4.4.1 제주관광신문고

<Figure 6>은 제주관광신문고의 모듈성 분석결과를 그래프로 나타낸 것이다. 그래프의 X축은 각각의 모듈들의 항목 개수를 나타낸 것이고 Y축은 모듈들의 번호로 임의로 생성된다. 제주관광신문고 모듈의 총 개수는 399개이며 이중 높은 중심성을 보인 모듈은 총 9개였다. 모듈성 분석을 측정하는 Modularity는 0.466으로 나타났다. 이 수치는 1에 가까울수록 모듈화 정도가 높음을 나타낸다.

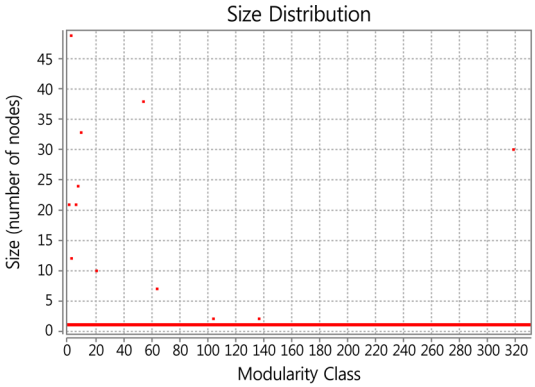


<Figure 6> Jeju Tourism Complaints Modularity Analysis

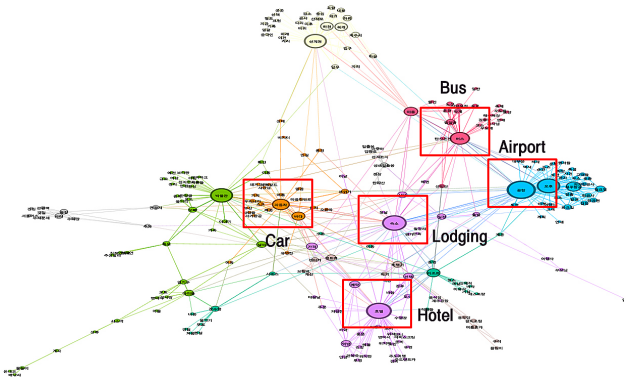
<Figure 7>은 제주관광신문고의 모듈성 분석결과를 구체적으로 확인하기 위해 시각화를 실시한 자료이다. 그림에서 동일한 색으로 연결된 모듈들은 같은 범주에 속해있다는 것을 의미한다. 이 중 중심성이 높은 9개의 모듈에서 제주도 관광 문제와 관련이 있는 단어를 분석한 결과, '관광객', '직원', '제주도', '전화', '예약'의 5가지인 것으로 나타났다. 키워드 분석결과와 비교했을 때 공통으로 도출된 단어는 '예약' 하나였다.



<Figure 7> Jeju Tourism Complaints Modularity Analysis Visualization



<Figure 8> Web-Scrapping Modularity Analysis



<Figure 9> Web-Scrapping Modularity Analysis Visualization

4.4.2 웹 스크래핑

<Figure 8>은 웹 스크래핑에 대한 모듈성 분석결과를 그래프로 나타낸 것이다. 모듈의 총 개수는 331개 이며 그 중 12개의 군집이 중심성이 높은 것으로 분석되었다. Modularity는 0.675였다.

다음으로 분석결과를 시각화 시킨 결과 중심성이 높은 12개의 모듈 중에 주제와 연관성이 있는 것은 5개로 나타났다. 5개의 모듈은 각각 '자동차', '숙소', '버스', '공항'이었으며 키워드 분석결과와 공통으로 도출된 단어는 '숙소' 하나로 나타났다. 이를 통해 웹 스크래핑 역시 제주관광신문고와 마찬가지로 키워드 분석결과와 차이를 보였다.

<Table 4> Comparison of Keyword & Sentiment Analysis

No.	Keyword
1	Time(51)
2	Tourist(46)
3	Phone Call(42)
4	Problem(41)
5	Visit(37)
6	Attraction(32)
7	Management(31)
8	Reservation(29)
9	Clerk(28)
	Car(28)

No.	Sentiment
1	Complaint(-15.2256)
2	Sincerity(-15.1985)
3	Humanly(-14.5295)
4	Words(-12.2775)
5	Reservation(-10.7483)
…	
13	Problem(-7.8554)
21	Management(-6.6804)
24	Tourist(-5.7711)
33	Phone Call(-3.6785)

<Table 5> Comparison of Modularity & Sentiment Analysis

No.	Modularity
1	Jeju Island
2	Tourist
3	Phone Call
4	Clerk
5	Reservation

No.	Sentiment
5	Reservation(-10.7483)
24	Tourist(-5.7711)
33	Phone Call(-3.6785)
43	Clerk(-2.5752)
…	

4.5 감성 분석

4.5.1 제주관광신문고

<Table 4>는 제주관광신문고의 키워드 분석과 감성 분석을 비교한 표이다. 제주관광신문고의 경우, 수집한 자료 200건 중 부정적 의미로 도출된 단어의 개수는 총 89개로 집계되었다. 분석결과 가장 높은 수치를 기록한 단어는 '항의'였다. 다음으로는 '진심', '인간적', '말씀'순이었다. 두 분석에서 동일하게 도출된 단어는 5개이며, '예약'이 두 분석 모두에서 높은 비중을 차지하고 있음을 확인 할 수 있다. 단, '문제', '관리', '관광객', '전화'는 회귀계수가 높진 않지만 부정의 의미로 쓰였다는 것을 확인 할 수 있다.

같은 방법으로 모듈성 분석과 감성 분석을 비교한 결과, 제주도를 제외한 나머지 4개 단어가 모두 감성 분석에 포함되어 있었다. 이를 통해 키워드 분석보다는 모듈성 분석이 감성분석결과와 유사하다는 사실을 확인 할 수 있다.

4.5.2 웹 스크래핑

웹 스크래핑을 통해 수집한 자료 중에서 공통어를 제외한 부정적 의미의 단어는 31개로 집계되었다. 분석결과에서 회귀계수가 가장 높게 나타난 단어는 '문제'였고

<Table 6> Comparison of Keyword & Sentiment Analysis

No.	Keyword
1	Problem(67)
2	Time(57)
3	Lodging(40)
4	Price(33)
5	Rent Car(32)
6	Airport(28)
7	Weather(28)
8	Reservation(26)
9	Sea(25)
10	Schedule(22)

No.	Sentiment
1	Problem(-96.1103)
2	Complaint(-25.9050)
3	Relief(-19.4615)
4	Reservation(-18.011)
5	Old days(-11.8714)
6	bad(-11.0879)
7	First day(-10.8449)
8	Rent Car(-10.5241)
…	
27	Lodging(-0.7965)

<Table 7> Comparison of Modularity & Sentiment Analysis

No.	Modularity
1	Jeju Island
2	Tourist
3	Phone Call
4	Clerk
5	Reservation

No.	Sentiment
5	Reservation(-10.7483)
24	Tourist(-5.7711)
33	Phone Call(-3.6785)
43	Clerk(-2.5752)
…	

다음으로는 ‘불만’, ‘안심’, ‘예약’ 순이었다. 키워드 분석결과와 동일하게 나온 단어의 수는 4개로 제주관광신문고와 비슷한 수준으로 일치했다. 단, ‘숙소’를 제외한 3개의 단어는 모두 감성 분석에서 높은 수준의 비중을 차지하고 있는 것으로 나타났다.

마찬가지로 웹 스크래핑의 모듈성 분석과 감성 분석의 내용을 정리한 결과 두 분석에서 공통으로 도출된 단어는 ‘숙소’와 ‘버스’였다. 이를 통해 웹 스크래핑의 자료가 제주관광신문고 보다는 감성분석결과의 일치 비율이 낮은 것을 알 수 있다.

4.6 분석결과 비교

<Table 8>은 앞서 실행한 3가지 분석결과를 채널을 기준으로 정리한 것이다. 각각의 채널에서 공통으로 2번 이상 나온 단어들은 ‘문제’와 ‘예약’이었다. ‘예약’의 경우 두 채널 모두 동일한 내용의 문제점을 지적하고 있었다. 관련 세부항목들을 보면 예약내용과 실제 서비스이용과의 차이가 다수를 차지했다. 이를 통해 제주도 관광 문제 중 예약과 관련한 사항이 가장 시급하게 개선되어야 할 사항이라는 것을 확인 할 수 있다. 단 ‘문제’의 경우, 세부 항목에서 자료 간 내용의 차이를 보이고 있다. 분석결과를 비교해 보면 제주관광신문고에서는 공공시설에 대한 내용이 다수를 이루고 있지만 웹 스크래핑에서는 사설 업체들에 대한 내용이 주를 이루고 있다. 이를 통해 각각의 채널에서 공통으로 언급된 단어라도 세부항목에서 내용상의 차이가 있을 수 있다는 사실을 알 수 있다.

<Table 8> Analysis Result

Jeju Tourism Complaint	Division	Web Scrapping
Time, Tourist, Phone Call, Problem, Visit, Attraction, Management, Reservation, Clerk, Car	Keyword	Problem, Time, Lodging, Price, Rent Car, Airport, Weather, Reservation, Sea, Schedule
Jeju Island, Tourist, Phone Call, Clerk, Reservation	Modularity	Lodging, Bus, Airport, Car, Hotel,
Complaint, Sincerity, Humanly, Words, Reservation, Correction, Internet, Kindness, Bad, No	Sentiment	Problem, Complaint, Relief, Reservation, Old days, Bad, First days, Rent Car, Arrangement, Lodging

비 공통부문에서는 이러한 차이가 좀 더 자세히 드러난다. 제주관광신문고에서 도출된 단어들 중 3가지 분석에서 2번 이상 나온 단어는 ‘관광객’, ‘전화’, ‘직원’이었다. 단어의 세부 항목들을 살펴보면 제주관광신문고에서는 웹 스크래핑에 비해 서비스를 제공하는 직원들의 친절에 관련된 사항이 많다는 것을 확인 할 수 있다.

웹 스크래핑 분석결과에서 2번 이상 도출된 단어는 ‘숙소’, ‘렌트카’, ‘예약’이었다. 공통부문인 ‘예약’을 제외하고 세부내용을 살펴보면 웹 스크래핑에서는 시설 및 서비스를 이용하는데 있어 발생하는 불만이 많다는 사실을 확인 할 수 있다. 이를 통해 공통주제라고 하더라도 채널에 따라 발생하는 내용에는 차이가 있다는 것을 확인 할 수 있다.

5. 결 론

본 연구에서는 VOC 자료와 외부채널 간의 데이터 비교를 통하여 VOC의 문제점을 확인하고 VOC가 다양한 의견을 반영할 수 있는 자로로써 활용될 수 있는 방향성에 대하여 연구해 보았다. 빈도수를 기준으로 자료 간 상호비교를 실시하는 키워드 분석결과, 공통으로 도출된 단어보다는 비 공통부분의 단어가 더 많이 도출되는 것을 확인 할 수 있었다. 또한 공통부분에서도 세부항목에서 차이를 보였다. 단어 간 유사도를 측정하여 중심성을 기준으로 각각의 자료의 차이를 확인하기 위해 활용한 모듈성 분석에서는 키워드 분석결과보다 비 공통부분의 비중이 높았다. 감성 분석에서는 회귀 분석을 통해 측정된 회귀계수와 앞선 두 분석결과와의 비교를 통해 분석 방법의 신뢰성을 검증하고자 하였다. 분석결과, 차이는 있으나, 대체적으로 키워드 및 모듈성 분석결과와 일치하는 것으로 나타났다. 본 논문은 각각의 분석의 기준(빈도수, 군집정도, 긍정과 부정)에 따라 분석을 실시하였다. 분석결과 빈도수와 군집정도(단어 간 공통적으로 도출 또는 유사성), 그리고 긍정과 부정정도에서 각 채널의 결과가 일치하지 않았다. 이를 통해

공통주제라 할지라도 채널에 따라 내용이 달라질 수 있다는 사실을 확인하였다. 본 연구의 결과는 VOC를 활용하고 있는 기업 및 공공기관에서 다양한 의견을 반영하여 VOC의 문제점을 보완하는 방법으로 활용될 수 있다. 단, 웹 스크래핑 자료 뿐 아니라 더 많은 외부채널들을 활용하고, 감성 분석 시에 사용하는 긍정 및 부정 단어들에 대한 기준 관련 연구가 필요하다.

Acknowledgement

This work was supported by the research fund of Hanyang University(HY-2017-G).

References

[1] Allen, J., James, A.D., and Gamlen, P., Formal versus informal knowledge network in R&D : a case study using social network analysis, *R&D Management*, 2007, Vol. 37, No. 3, pp. 179-196.

[2] Baek, B.S., Research of VOC operating system and the problems of the police service, *Korean Association of Police Science*, 2013, Vol. 15, No. 5, pp. 115-146.

[3] Blondel, V.D., Guillaume, J.L., Lambiotte, R., and Lefebvre, E., Fast unfolding of communities in large network, *Journal of Statistical Mechanics : Theory and Experiment*, 2008, pp. 1-12.

[4] Chang, J.Y., An Opinion Document Clustering Technique for Product Characterization, *Society for E-Business Studies*, 2014, Vol. 19, No. 2, pp. 95-108.

[5] Cho et al., Sentiment Analysis Using News Comments for Public Opinion Mining, *Korean Institute of Intelligent Systems Conference*, 2013, Vol. 23, No. 1, pp. 149-150.

[6] Cho, Y.H. and Bang, J.H., Social Network Analysis for New Product Recommendation, *Journal of Intelligent Information Systems*, 2009, Vol. 15, No. 4, pp. 183-199.

[7] Choi, C.H., A Study on the Informal Networks in Organizations : An Application of Social Network Analysis, *Seoul Association For Public Administration*, 2006, Vol. 17, No. 1, pp. 1-23.

[8] Choi, C.I. and Kang, B.Y., Social Network Analysis on the Climate Change Researches and its Implications, *The Korean Regional Development Association*, 2012, Vol. 24, No. 5, pp. 1-20.

[9] Freeman, L.C., Centrality in social networks conceptual clarification, *Social Networks*, 1979, Vol. 1, No. 3, pp. 215-239.

[10] Griffin, A. and Hauser, J.R., The Voice of Customer, *Marketing Science*, 1993, Vol. 12, No. 1, pp. 1-24.

[11] Han, S.H., A Study on Keyword Extraction From a Single Document Using Term Clustering, *Journal of the Korean Society for Library and Information Science*, 2010, Vol. 44, No. 3, pp. 155-173.

[12] Hui, Z. and Trevor, H., Regulariuzation and variable selection via elastic net, *Journal of the Royal Statistical Society : Series B*, 2005, Vol. 67, No. 2, pp. 301-320.

[13] Hyeon et al., Customers who take advantage of multi-dimensional clustering analysis of the topic, *Korea Intelligent Information Systems Society conference*, 2014, pp. 183-185.

[14] Jang, G.S., A Integrated VOC Management Schema in Large-Scale Manufacturing Companies : A Case Study on Implementation for Construction Equipment Division in 'H' Heavy Industry, *Journal of the Korea Society of Computer and Information*, 2009, Vol. 14, No. 8, pp. 127-136.

[15] Ju, J.M. and Hwang, S.G., Establish of VOC analysis system for efficient CRM, *The Korean Society for Quality Management*, 2003, Vol. 32, No. 1, pp. 75-91.

[16] Kang, B.I., Song, M., and Jho, W.S., A Study on Opinion of Newspaper Texts based on Topic Modeling, *Journal of the Korean Society for Library and Information Science*, 2013, Vol. 47, No. 4, pp. 315-334.

[17] Kim et al., A Social Network Analysis on Stakeholders of Watershed Management in Saemanguem Area, *The Korean Association For Environmental Sociology*, 2013, Vol. 17, No. 2, pp. 227-274.

[18] Kim et al., A Technique of the Approval Rating Analysis for Political Party Using Opinion Mining, *Journal of Korean Institute of Information Technology*, 2014, Vol. 12, No. 10, pp. 133-141.

[19] Kim, J.E., Kim, N.G., and Cho, Y.H., User-Perspective Issue Clustering Using Multi-Layered Two-Mode Network Analysis, *Journal of Intelligent Information Systems*, Vol. 20, No. 2, pp. 93-107.

[20] Kim, J.H. and Kwon, O.B., Business Intelligence from Voice of Customer, *Korean Academic Society of Business Administration Conference*, 2015, pp. 1-9.

[21] Kim, J.Y., An efficient acquisition method of VOC in new product development, *Korea Business Education Association*, 2012, Vol. 23, No. 1, pp. 54-76.

[22] Kim, S.H. and Jang, R.S., The Study on the Research Trend of Social Network Analysis and its Applicability to

Information Science, *Journal of the Korean Society for Information Management*, 2010, Vol. 27, No. 4, pp. 71-87.

[23] Kim, Y.G., Lee, W.H., and Kwon, H.C., Clustering of Web Document Exploiting with the Co-link in Hypertext, *Journal of Korean Library and Information Science Society*, 2003, Vol. 34, No. 2, pp. 233-253.

[24] Kim, Y.H., In, K.H., and Kim, E.M., The Blog Polarity Classification Technique using Opinion Mining, *Korean Intelligent Information Systems Society Conference*, 2012, pp. 225-228.

[25] Kim, Y.S. and Jeong, S.R., Intelligent VOC Analyzing System Using Opinion Mining, *Journal of Intelligent Information Systems*, 2013, Vol. 19, No. 3, pp. 113-125.

[26] Kwon, B.C. and Cho, N.W., A Study on the Relationship between Social Network of Code share and Performances in Airline Industries, *Journal of the Korean society for Quality Management*, Vol. 39, No, 2, pp. 271-280.

[27] Lee et al., Developing Corporate Valuation System with Opinion Mining Based on Big Data, *Proceedings of the 25th Annual Conference on Human and Cognitive Language Technology*, 2013, Vol. 25, pp. 126-128.

[28] Lee, D.R., A Social Network Analysis on Basic Old-age Pension Policymaking Process, *Korean Association For Policy Science*, 2012, Vol. 16, No. 3, pp. 109-135.

[29] Lee, H.J., Analysis of marketing research knowledge structure using social network analysis, *Korean Academic Society of Business Administration*, 2014, Vol. 43, No. 3, pp. 965-992.

[30] Lee, K.M., Text Characteristic Recognition Using by Clustering, *Korea Information Processing Society*, 2001, Vol. 8, No. 1, pp. 837-840.

[31] Lee, M.H., Kim, J.H., and Jang, J.Y., Intermediate support organizations analyzed using Social Network Analysis, *Seoul Association For Public Administration*, 2011, Vol. 22, No. 1, pp. 141-178.

[32] Lee, S.H., Hwang, H.K., and Kim, J.Y., A Study on Utilization of Customers(VOC) for Improvement in Information Services, *Journal of Korean Library and Information Science Society*, 2015, Vol. 46, No. 1, pp. 25-42.

[33] Lee, S.Y. and Kim, Y.E., A Study on the establishment of VOC system in compliance with the shift in customer trend, *Journal of Distribution Science*, 2009, Vol. 7, No. 2, pp. 89-119.

[34] Lee, S.Y., In, K.H., and Kim, U.M., Analyzing University Bulletin Board Data by using Opinion Mining, *Korea Information Science Society Conference*, 2012, Vol. 39, No. 2, pp. 104-106.

[35] Lee, W.K., Community centrality of power analysis through network analysis, The Korean Association For Public Administration Conference, 1997, pp. 377-394.

[36] Moody, J., The structure if a social science collaboration network : Disciplinary cohesion from 1963 to 1999, *American sociological review*, 2004, Vol. 69, No. 2, pp. 213-238.

[37] Moon, H.N. and Kim, J.W., Internet news research individual stock returns prediction models utilizing : the use of opinion mining techniques, *Korea Intelligent Information Systems Society Conference*, 2014, pp. 387-393.

[38] Park et al., Two-Level Clustering for Sub-Topic Labelling of Social Media Data, *Journal of KISS : Software and Applications*, 2014, Vol. 41, No. 3, pp. 225-232.

[39] Park, C.Y., Simple principal component analysis using Lasso, *Journal of the Korean Data and Information Science Society*, 2013, Vol. 24, No. 3, pp. 533-541.

[40] Park, W.J. and Yu, K.Y., Spatial Clustering Analysis based on Text Mining of Location-Based Social Media Data, *Journal of the Korean Society for Geo-Spatial Information System*, 2015, Vol. 23, No. 2, pp. 89-96.

[41] Seung, H.W. and Park, M.Y., A Clustering Technique Using Association Rules for The Library and Information Science Terminology, *Journal of the Korean Society for Library and Information Science*, 2003, Vol. 37, No. 2, pp. 89-105.

[42] Singh, J., A Typology Consumer Dissatisfaction Response Styles, *Journal of Retailing*, 1990, Vol. 66, No. 1, pp. 57-99.

[43] Song et al., An Integrated Data Mining Model for Customer Relationship Management, *Korea Information Science Society*, 2006, Vol. 33, No. 2, pp. 154-159.

[44] Woo, J.Y., Choi, M.S., and Park, S.C., Systematic research on the causes of changes occurring Voice of Customer, *Korean Institute of Industrial Engineers Fall Conference Proceeding*, 2004, pp. 33-37.

[45] Yun, S.J., Kim, S.H., and Shin, K.S., Development of the Accident Prediction Model for Enlisted Men through an Integrated Approach to Datamining and Textmining, *Journal of Intelligent Information Systems*, 2015, Vol. 21, No. 3, pp. 1-17.

ORCID

Sang-Hyeon Ahn | https://orcid.org/0000-0003-3899-3514

Dong Hyun Baek | https://orcid.org/0000-0002-3107-9511