creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공기업정책학 석사 학위논문

빅데이터 기반 공항서비스 만족도 분석

- 고객의 소리(VOC) 분석을 중심으로 -

2021년 8월

서울대학교 행정대학원

공기업 정책학과

구 현 규

빅데이터 기반 공항서비스 만족도 분석

- 고객의 소리(VOC) 분석을 중심으로 -

지도교수 엄 석 진

이 논문을 공기업정책학 석사 학위논문으로
제출함
2021년 3월

서울대학교 행정대학원
공기업정책학과
구 현 규

구현규의 석사 학위논문을 인준함
2021년 6월

위 원 장 _____ 김 윤 지

부위원장 _____ 최 종 원

위 원 _____ 엄 석 진

국문초록

최근 공항산업 분야에서도 주변 경쟁공항들이 앞다투어 시설 확장에 과감한 투자를 하고 있고, IOT, AI, 빅데이터를 기반으로한 첨단 스마트 시스템을 선제적으로 도입하고 있어 한반도라는 지정학적 리스크를 가지고 있는 인천공항이 시설 규모나 여객수로는 세계 유수 공항들과의 경쟁력에서 비교우위를 점하기는 점차 쉽지 않을 것으로 보인다.
이에 본 연구는 인천공항만이 가지고 있는 서비스 경쟁력을 확보하고자 고객의 서비스 불만족 요인을 파악하고 그에 따른 고객만족을 위한 서비스 개선 및 회복 전략을 수립하고자 고객만족도 영향요인에 대한 혼합방법적인 분석을 위해 고객민원 데이터인 VOC(고객의소리)에 대한 빅데이터 분석을 시행하였다. 먼저 고객의 불만족 요인을 파악하기 위해 텍스트 빈도분석과 토픽분석을 시행하였고, 빅데이터 분석결과와 실제 고객만족도 조사결과에 대한 상호 비교 검증을 실시함으로써, 공통적으로 일치하는 서비스 품질 영향요인을 재확인하고, 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출하는 것을 목적으로 하며, 이 연구의 연구문제는 다음과 같다.

첫째, 최근 5년간 VOC 내용의 주요 키워드들과 토픽들은 무엇인가 ?
둘째, 최근 5년간 주요 토픽들은 시간의 흐름에 따라 어떻게 변화하며, 관심이 증가하는 토픽과 감소하는 토픽은 무엇인가 ?
셋째, 전체 문서에 대한 감성분석 결과를 통해 최근 5년간 VOC 내용의 긍정적 키워드와 부정적 키워드는 각각 무엇인가 ?
넷째, VOC 빅데이터 분석결과와 실제 고객만족도 조사 간에는 어떠한 공통점과 차이점이 있는가 ? 품질개선이 필요한 사항은 무엇인가 ?

먼저 인천공항 고객의 불편·불만족과 관련한 데이터를 수집·분석하기 위해 최근 5년('15년 하반기～'20년 상반기) 동안 인천공항 VOC(고객의 소리) 시스템에 접수된 약 33,000건의 고객 민원 자료를 활용하였다. 또한 기존의 고객만족도 분석결과와의 상호 비교검증을 위해 최근 5년('15년～'20년) 동안 인천공항 이용객을 대상으로 설문조사를 실시한 고객만족도 조사결과를 활용하였으며, 연구의 주요결과는 다음과 같다.

첫째, VOC 문서에 대한 키워드 빈도 분석결과 직원, 사람, ‘버스’, ‘주차’, ‘주차장’, ‘차량’, ‘택시’, ‘대행’이라는 키워드의 출현빈도가 높아 인적서비스와 관련된 문서가 많은 것을 확인할 수 있으며, 서비스 유형으로는 접근교통 및 주차 서비스와 관련된 문서가 많았음을 확인할 수 있었다.
둘째, VOC 문서에 대한 토픽모델링(LDA)분석을 시행하였고, 문서에 잠재된 토픽을 추출하기 위해 정합도(coherence score)를 이용하여 최적토픽수 8개를 결정하였으며, 토픽을 구성하는 상위 30개의 단어를 바탕으로 8개의 토픽(① 주차서비스, ② 인적서비스(단순불만), ③ 인적서비스(강한불만), ④ 상담/문의사항, ⑤ 상업시설서비스, ⑥ 공항시설서비스, ⑦ 공항공사 개선노력, ⑧ 접근교통서비스)을 추출하였다.
먼저 직원응대와 관련한 인적서비스 관련 토픽(②번 토픽, ③번 토픽)의 비중이 가장 높았고, 다음으로 인프라 서비스인 공항시설과 관련한 ③번 토픽의 비중이 높았다.
8개 주요토픽의 시간의 흐름에 따른 시계열적 변화를 분석해본 결과, 대부분의 토픽이 ‘17년도에 비해 ‘18년도에 급격하게 증가하였고, 특히 공항시설과 관련한 6번토픽의 비중이 57%이상 증가한 것을 확인할 수 있는데 이는 인천공항 제2여객터미널이 2018년 1월 그랜드 오픈함에 따라, 공항운영 불안정성으로 전체적인 토픽수 및 공항시설과 관련된 6번토픽의 비중이 급격하게 증가했음을 보여준다.
셋째, 문서의 극성정도(긍정/부정/중립)를 파악하기 위해 별도의 감성분석을 시행하였고, 고객 불편·불만 등의 부정적 감성을 나타내는 문서가 전체 문서의 60.2%, 상담·문의, 의견제안 등 중립적 감성을 나타내는 문서가 37.5%, 칭찬·격려 등의 긍정적 감성을 나타내는 문서는 2.3%로 가장 낮은 비율로 나타났다. 부정적 감성의 문서에서는 ‘직원’, ‘말’, ‘태도’, ‘생각’, ‘교통’, ‘주차’, ‘버스’, ‘검색’ ‘불편’ 등의 키워드가 상위를 차지함에 따라 접근교통 및 주차서비스와 보안검색 서비스 분야에서 공항직원의 인적서비스에 대해 공항이용객의 불편과 불만족이 많았음을 확인할 수 있었고, 긍정적 문서에서는 ‘직원’, ‘감사’, ‘칭찬’, ‘마음’, ‘도움’, ‘친절’, ‘핸드폰’, ‘가방’, ‘안내서비스’, ‘안내데스크’ 등의 키워드가 상위를 차지함에 따라 공항안내 직원의 친절한 안내 서비스에 고객들이 감사해하는 내용이 많았음을 확인할 수 있었다.

넷째, 고객의 소리(VOC) 분석결과와 실제 고객만족도 조사결과와 상호 비교분석을 위해 고객의 소리(VOC) 문서에 대한 서비스 유형별(7개분야) 서비스 긍정지수를 산출하여 실제 고객만족도 점수와 비교하였다. 분석 결과, 5개 서비스 유형(① 교통/주차, ④ 체크인서비스, ⑤ 상업시설, ⑥ 안내서비스, ⑦ 편의시설)은 고객만족도 점수와 서비스 긍정지수 간에 공통적으로 일치하는 경향을 보였으며, 2개 서비스 유형인 ② 보안검색과 ③ 출입국심사 분야는 고객만족도 점수와 서비스 긍정지수 간에 차이점을 나타내는 것을 확인할 수 있었다. 출입국심사 분야는 고객만족도 점수는 가장 높은 점수를 획득하였으나, 서비스 긍정지수는 세 번째로 낮은 점수를 획득하였는데 이는 자동출입국심사 도입에 따라 출입국 심사 이용 편리성 및 여권/비자 심사의 신속성과 정확성 등은 크게 개선되었으나, 출입국 심사관의 불친절한 응대 및 대기시간 지연에 따른 고객의 불편·불만사항이 많았던 것으로 분석되었다. 또한 공항 서비스 종합만족도에 미치는 영향요인을 분석하기 위해 고객만족도 조사에 대한 다중회귀분석결과를 시행하였고, 통계적으로 유의미한 영향을 미치는 변수는 편의시설, 안내시설, 상업시설(쇼핑매장), 보안검색 서비스로 나타났다. 이는 VOC데이터의 경우 공항직원의 응대, 친절도 등 인적서비스와 관련된 텍스트 및 토픽의 비중이 높은 반면에 고객만족도 조사의 경우 여객터미널내 시설 이용의 편리·편의성 등 시설서비스가 공항서비스 종합만족도에 큰 영향을 미치는 것으로 해석할 수 있을 것이다.

본 연구는 고객의 소리(VOC) 자료에 대한 키워드 분석 및 토픽분석을 통해 고객의 불편·불만족사항에 대한 주요 이슈를 도출해냈으며, 토픽의 시계열적 분석을 통해 고객의 주요 관심도의 추이를 확인하였다. 또한 VOC 문서에 대한 극성정도(긍정/부정/중립)를 파악하기 위해 감성분석을 시행하였고, 고객만족도 조사결과와 상호 비교 분석을 통해 공통적으로 일치하는 서비스 품질 영향요인과 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출해냈다는데에 연구의 의의가 있다.

주요어 : 빅데이터, 텍스트마이닝, VOC(고객의 소리), 키워드분석, 토픽모델링, 감성분석, 서비스 긍정지수, 고객만족도

학 번 : 2020-23375

목 차

제 1 장 서론 ····· 1

제 1 절 연구의 배경 및 목적 ····· 1

제 2 절 연구의 대상과 범위 ····· 4

제 3 절 연구의 방법 ····· 5

제 2 장 이론적 배경 및 선행연구 검토 ····· 7

제 1 절 이론적 배경 ····· 7

1. 공항서비스의 정의 및 중요성 ····· 7
2. 공항서비스 품질 결정요인 및 측정 ····· 9
3. 공항서비스 고객 만족도 평가 현황 ····· 11
4. 빅데이터와 분석기법 ····· 14

제 2 절 선행연구 검토 ····· 20

1. 공항서비스 품질과 고객 만족도에 관한 연구 ····· 20
2. 빅데이터 분석기반 고객만족도 연구 ····· 24
3. 본 연구의 차별성 ····· 26

제 3 장 연구모형 및 설계 ····· 28

제 1 절 연구의 분석틀 ····· 28

1. 연구 모형 ····· 28
2. 연구 문제 ····· 29

제 2 절 자료의 수집 및 분석방법 ····· 29

1. 표집 방법 ····· 29
2. 표집 현황 ····· 30
3. 분석 방법 ····· 36

제 4 장 연구 결과 ····· 40
제 1 절 데이터 분석결과 ····· 40
1. 데이터 전처리 과정 ····· 40
2. 텍스트 빈도 분석 ····· 41
3. 토픽모델링(LDA) 분석 ····· 45
4. 감성분석 ····· 59
제 2 절 고객만족도 조사 분석 ····· 62
1. 고객만족도 조사 분석결과 ····· 62
2. 고객만족도 조사 다중회귀분석 ····· 64
제 3 절 상호비교 분석 ····· 66
1. 서비스 유형별 만족도 비교 ····· 66
2. 종합비교 ····· 69

제 5 장 결 론 ····· 70
제 1 절 연구의 요약 ····· 70
제 2 절 연구의 의의 및 시사점 ····· 73
제 3 절 연구의 한계 및 향후 연구방향 ····· 75

참고문헌 ····· 76
Abstract ····· 80

표 목 차

[표 2-1] 공항서비스 분류 ····· 8
[표 2-2] ASQ 평가 영역 및 세부 평가 항목 ····· 12
[표 2-3] ASQ, SKYTRAX 서비스 평가 비교 ····· 13
[표 2-4] 빅데이터 분류 ····· 15
[표 2-5] 고객만족도 관련 선행연구 검토 ····· 23
[표 2-6] 텍스트마이닝 분석 선행연구 검토 ····· 25
[표 3-1] 인천공항 VOC (고객의 소리) 현황 ····· 30
[표 3-2] 인천공항 고객만족도 조사 현황 ····· 30
[표 3-3] 고객의 소리(VOC) 접수채널 ····· 31
[표 3-4] 고객의 소리(VOC) 요구 유형별 현황 ····· 32
[표 3-5] 고객의 소리(VOC) 서비스 유형별 현황 ····· 33
[표 3-6] 공항서비스 만족도의 평가지표 ····· 35
[표 3-7] 텍스트마이닝(단어-문서 행렬 예시) ····· 37
[표 4-1] 상위 50개 빈출 키워드와 빈도('15년~'20년) ····· 41
[표 4-2] 상위 50개 빈출 키워드와 빈도('20년) ····· 43
[표 4-3] 토픽분석 종합(토픽명, 포함단어) ····· 56
[표 4-4] 연도별 토픽의 추이분석 ····· 57
[표 4-5] 인천공항 고객만족도 조사 세부 평가항목 ····· 62
[표 4-6] 인천공항 고객만족도 회귀분석 결과 ····· 65
[표 4-7] 출입국 심사분야 고객만족도 세부 평가결과 ····· 68
[표 4-8] 출입국 심사 불만족 사례 ····· 68

그 림 목 차

[그림 2-1] 고객만족도의 선행 및 결과 변인 구조 ····· 9
[그림 2-2] 텍스트 마이닝의 주요과정 ····· 17
[그림 3-1] 공항서비스 고객만족도 분석 연구모형 ····· 28
[그림 3-2] 인천공항 고객만족도 조사 개요 ····· 34
[그림 3-3] 최근 3년간 인천공항 고객만족도 조사결과 ····· 35
[그림 3-4] 고객의소리(VOC) 분석 연구방법과 절차 ····· 36
[그림 3-5] 텍스트마이닝(LDA 토픽 모델링 예시) ····· 37
[그림 3-6] 텍스트마이닝(키워드 네트워크 구성 예시) ····· 38
[그림 3-7] 텍스트마이닝(감성분석 예시) ····· 39
[그림 4-1] 상위 50개 키워드 워드클라우드('15년~'20년) 42
[그림 4-2] 상위 50개 키워드 워드클라우드('20년) ····· 44
[그림 4-3] 토픽수 산출 관련 정합도 변화 그래프 ····· 45
[그림 4-4] 토픽모델링(LDA) 시각화 결과(전체) ····· 47
[그림 4-5] 토픽모델링(LDA) 1번토픽(주차서비스) ····· 48
[그림 4-6] 토픽모델링(LDA) 2번토픽(인적서비스) ····· 49
[그림 4-7] 토픽모델링(LDA) 3번토픽(인적서비스) ····· 50
[그림 4-8] 토픽모델링(LDA) 4번토픽(상담/문의서비스) ··· 51
[그림 4-9] 토픽모델링(LDA) 5번토픽(상업시설서비스) ···· 52
[그림 4-10] 토픽모델링(LDA) 6번토픽(공항시설서비스) ·· 53
[그림 4-11] 토픽모델링(LDA) 7번토픽(공항공사 개선노력) 54
[그림 4-12] 토픽모델링(LDA) 8번토픽(접근교통서비스) ·· 55
[그림 4-13] 토픽모델링(LDA) 시계열 분석 ····· 59
[그림 4-14] 고객의소리(VOC) 감성분석 결과 ····· 60
[그림 4-15] 고객의소리(VOC) 감성분석 매트릭스 ····· 61
[그림 4-16] 인천공항 고객만족도 조사결과(최근5년간) ···· 63
[그림 4-17] 인천공항 서비스 유형별 만족도 비교분석 ····· 66

제 1 장 서 론

제 1 절 연구의 배경 및 목적

전세계적으로 4차 산업혁명 시대가 열리면서 시장과 고객 모두가 변화하는 격변기에 돌입함에 따라 기업간에 지속가능한 경쟁우위를 확보하기 위한 치열한 경쟁이 발생하고 있으며, 차별화된 경쟁력을 확보하는 것이 기업 생존의 필수요소로 작용하고 있다. 공항산업 분야에서도 동아시아 경제 협력체의 출범과 함께 동아시아권 항공 체계가 재정비되고 있을 뿐 아니라 주변국가의 공격적인 공항투자로 인해 공항간 경쟁이 치열해 지고 있다.

최근 중국이 기존 북경의 서우두 공항 이외에 베이징에 연간 1억명의 여객을 수용할 수 있는 다싱공항을 개항하였고, 싱가포르 창이공항도 연간 1,600만명을 수용할 수 있는 제4여객터미널을 오픈하는 등 글로벌 허브공항으로서의 입지를 강화하기 위해 동아시아 주요 공항들간의 경쟁구도가 다시금 심화되고 있다.

외부 환경요인을 종합적으로 분석해보면 주변 경쟁공항들이 앞다투어 시설 확장에 과감한 투자를 하고 있고, IOT, AI, 빅데이터를 기반으로 한 첨단 스마트 시스템을 선제적으로 도입하고 있어 한반도라는 지정학적 리스크를 가지고 있는 인천공항이 시설 규모나 여객수로는 세계 유수 공항들과의 경쟁력에서 비교우위를 점하기는 점차 쉽지 않을 것으로 보인다. 따라서 이제는 주변 공항과는 분명히 차별화된 인천공항만이 가지고 있는 서비스 경쟁력을 적극적으로 활용하여 고객 특성에 맞는 새로운 서비스 개발에 대한 노력이 필요한 때이다.

인천공항은 국제항공협의회(Airports Council International, ACI)가 주관하는 세계 공항서비스 평가(ASQ)에서 12년 연속 1위(2005년~2016년)를 차지하며 명예의 전당에 올랐으며, Skytrax가 매년 발표하는 세계 공항

서비스 순위에서도 꾸준히 상위권을 유지하며 주변국들의 벤치마킹 대상이자 경쟁상대로 꾸준히 거론되어 왔다.

최병권(2007)은 탁월한 기업의 서비스 품질이 서비스를 제공하는 기업의 경쟁우위와 경제적인 성공에서 핵심적인 요인이며, 서비스 품질이 기업의 성공적인 수익성을 보장해주는 가장 유력한 요인이라고 밝혔다. 또한 임성빈(2013)은 인천국제공항의 종합적인 경쟁력을 다이아몬드 모델을 통해 분석한 결과 인천국제공항이 서비스 경쟁력 측면에서 경쟁공항 대비 비교 우위에 있다고 밝히고 있는데 이는 공항서비스의 고객만족도가 공항의 국제경쟁력을 확보하는 데 있어서 가장 중요한 요소임을 보여주고 있는 결과이다.

Reicheld and Sasser(1990)은 앞선 연구를 통해 서비스 기업의 고객 이탈 비율이 5% 감축된 경우 기업의 이익이 25%내지 85%까지 증가하기 때문에 서비스 기업에게 있어 기존 고객의 유지는 매우 중요하다고 주장하였다. 따라서 서비스 제공자는 고객의 만족도를 유지 혹은 강화시켜 고객의 이탈을 최대한 방지해야 할 것이며 서비스의 불만족은 고객이 이탈할 수 있는 위기 상황인 동시에 서비스 개선을 통해 고객만족을 회복한다면 장기적인 관계를 구축할 수 있는 기회가 될 수도 있다

최근에도 인천공항은 세계 공항서비스 평가 1위의 자리를 지키기 위해 입국장 면세점 개장, 출입국 서비스 자동화 개편, 스마트 공항 구현 노력 등 고객 맞춤형 서비스 제공을 통해 고객만족도 향상을 위해 노력하고 있다. 또한 고객만족도 향상을 위한 서비스 품질개선 수단으로서 매년 분기별로 고객만족도 조사를 시행하여,「인천국제공항 주요 서비스」에 대한 품질수준을 진단하고 불만족 요인을 발굴해 나가고 있다. 그러나 이러한 노력에도 불구하고 인천공항에서 운영하고 있는 고객의 소리(VOC)[1] 시스템에 접수되는 고객 불만족, 불편 등과 관련된 민원건수는 '14년 5,145건, '15년 5,763건, '16년 5,426건, '17년 5,974건, '18년 7,380건 등으로 매년 점증적으로 증가하는 추세를 보이고 있다.

---

1) VOC 시스템이란 고객의 각종 상담문의, 불편불만, 칭찬격려, 의견제안 등을 수집, 배분, 처리, 분석, 활용하는 업무를 지원하는 전산시스템을 말한다.

인천공항은 고객의 불만해결과 고객만족이 매우 중요하다고 인식함에 따라 對 고객 서비스 개선사항을 발굴하고, 고객의 다양한 의견 및 제안을 반영하기 위해 VOC 시스템을 구축하여 운영중에 있다.

고객의 소리(Voice of Customer, VOC)는 고객으로부터 제안된 필요와 욕구 내지는 고객의 선호도와 회피하고 싶은것들에 대한 기술(description)이며(Lee et al.,2014b), 공항이용객이 어떤 구체적인 문제에 대해 공항측에 해결해 줄 것을 자발적으로 요구하는 통로로서, 고객이 남긴 글의 제목과 본문에는 고객의 불만사항이 주로 어떤 측면에서 제기되고 있는지에 대해 유용한 정보를 담고 있다. 따라서 고객들이 직접적으로 제기한 서비스 불만족 사항을 해소하고, 고객만족을 실현하기 위해서는 서비스 불만족의 원인을 명확히 파악하고 그에 따른 고객만족을 위한 서비스 개선 및 회복 전략을 수립하는 것이 필요하다.

그러나 인천공항 개항 이후 20년 넘게 고객 불만족과 관련한 다양한 의견이 여러 채널을 통해 축적되어 왔음에도 불구하고, 아직까지 고객만족도와 관련성이 높으며 고객 서비스 개선에 필요한 유용한 정보를 도출하여 서비스 개선에 적극적으로 반영하지 못하고 있는 실정이다. 이는 실제 VOC시스템이나 각종 SNS를 통해 생성되는 데이터의 대부분이 비정형 데이터로 존재하여 실제 자료 활용 및 분석이 어렵고, 데이터 분석에 필요한 기술력 및 관련 인프라가 많이 부족했기 때문이다.

그러나 최근에는 빅데이터 분석 기술의 발전과 함께 대규모 자료를 저장·관리 할 수 있는 클라우드 컴퓨팅 등 효율적인 장비활용이 가능해짐에 따라 다양한 분야에서 비정형의 방대한 데이터에 대한 주요 키워드 빈도수 분석, 토픽모델링 분석, 감성분석 등의 텍스트 마이닝(text mining) 분석을 활용한 인사이트 도출이 가능해 지고 있다.

빅데이터는 기존의 일반적인 데이터베이스 소프트웨어가 수행할 수 있는 데이터 수집 방식이나 저장, 관리 및 분석의 정도를 뛰어넘는 대량의 데이터를 의미(Manyika et al., 2011)하며, 기존의 데이터, 이론, 방법과 해석을 발전시키고 더욱 견고하게 만드는 잠재력을 갖고 있다고 인식되고 있다.(White & Breckenridge,2014)

이러한 빅데이터를 활용한 분석은 잘 드러나지 않았던 현장의 요구나 역동성, 의사결정을 위한 가치 있는 패턴이나 정보 등의 발견을 제공한다(Adrian et al., 2106)는 점에서 큰 의의가 있겠고, 국가나 기업의 미래 경쟁력 확보를 위해 중요한 자원으로 활용될 수 있다는 점에서 최근 크게 주목받고 있다. 따라서 공공기관인 인천국제공항의 입장에서도 빅데이터의 등장과 분석기술의 발전은 공공서비스인 공항서비스를 이용하는 고객이 직접적으로 느끼는 서비스 불만 요인과 고객니즈에 대한 정확한 분석이 가능해짐에 따라 사회적 갈등비용을 감소시키고, 공항서비스의 품질 수준을 향상시키는 기회로 작용할 수 있다.

이에 본 연구에서는 인천공항 고객만족도 영향요인에 대한 혼합방법적인 분석을 위해 고객민원 데이터인 VOC(고객의소리)에 대한 빅데이터 분석을 먼저 시행하고, 기존의 고객만족도 설문조사결과와 상호 비교 검증을 실시함으로써, 공통적으로 일치하는 서비스 품질 영향요인을 재확인하고, 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출하고자 한다.

또한, VOC(고객의 소리) 빅데이터 분석을 통해 서비스 유형별 고객만족도 점수의 예측가능성을 진단해보고자 하며, 이를 통해 기존 설문조사 방식의 고객만족도 평가의 여러 제약조건들을 상호 보완함으로써 공항서비스 고객만족도 향상에 크게 기여할 수 있을 것으로 생각한다.

제 2 절 연구의 대상과 범위

본 연구는 공항의 핵심 서비스 분야인 접근교통·주차, 탑승수속, 보안검색, 출입국심사, 상업시설 등에 있어서 고객 불편·불만족의 원인을 진단해 보고자 VOC(고객의 소리)에 대한 빅데이터 분석을 시행하고, 분석결과를 바탕으로 서비스 분야별 실제 고객만족도 점수와의 상관관계에 대해 분석해보고자 한다.

먼저 인천공항 고객의 불편·불만족과 관련한 데이터를 수집·분석하기 위해 최근 5년('15년 하반기~'20년 상반기) 동안 인천공항 VOC(고객의

소리) 시스템에 접수된 약 30,000건의 고객 민원 자료를 활용하고자 한다. VOC자료는 인천국제공항공사의 서비스를 제공받은 고객이 공사에 제기하는 각종 불편불만, 칭찬격려, 상담문의, 의견제안 등을 말하며, 홈페이지, 이메일, 엽서 등 온라인 및 오프라인을 통하여 접수되고 있다.

또한 공항서비스 분야별 고객만족도 설문 조사자료는 최근 5년('16년~'20년) 동안 인천공항 이용객을 대상으로 시행한 조사결과를 활용하고자 한다. 고객만족도 조사는 인천공항 주요서비스의 품질수준 진단 및 불만족 요인을 발굴하기 위해 매년 분기별로 4차에 걸쳐 1:1면접방식의 설문조사로 진행되며 주요프로세스별 유형 5가지(출국/입국/환승/항공사/문화예술)와 서비스 분야별 유형 7가지(교통/주차, 탑승수속, 보안검색, 상업시설, 공항시설, 안내서비스, 편의시설)로 구분하여 고객 체감 전반에 대한 만족도 조사를 실시한다.

제 3 절 연구의 방법

본 연구는 인천공항 고객 만족도에 대한 영향요인을 분석하기 위해 인천국제공항공사에서 운영하고 있는 고객의 소리(VOC) 시스템에 접수된 민원자료에 대한 빅데이터 분석(텍스트 빈도분석, 토픽모델링 분석, 감성분석 등)을 시행하고자 한다. 해당 자료는 고객의 개인정보가 일부 포함되어 있어 고객정보를 제외한 민원내용만을 수집·분석하고자 하며, 데이터분석하고자 하는 구체적인 내용은 다음과 같다.

- 인천공항 고객의 소리(VOC) 빅데이터 분석
· 텍스트 빈도분석(단어빈도, N-gram, 연결중심성 등)
· 토픽분석(LDA 토픽 모델링)
· 감성분석(감성어 사전 기반 감성추출)

분석프로그램은 텍스톰(Textom)의 빅데이터 분석 솔루션인 텍스트 마이닝 기술을 이용하여 VOC(고객의 소리) 문서에 대한 데이터 분석을 시행하고자 하며, 가장먼저 기존 보유 데이터인 VOC(고객의 소리)에

대한 데이터 정제및 형태소 분석 등의 데이터 전처리 과정을 거쳐 데이터를 가공한다. 다음으로 데이터 분석과정으로서 텍스트 빈도분석을 통해 VOC 문서에 등장한 주요 키워드를 도출할 것이며, 토픽분석을 통해서는 VOC 문서내 동시출현 빈도가 높은 단어를 배열함으로써 숨겨진 주제를 발견하게 될 것이다. 또한 감성분석을 통해서는 전체 문서에 대한 극성정도(긍정/부정/중립)의 구분이 가능해짐으로써 긍정문서와 부정문서 각각의 주요 키워드 도출이 가능할 것이다.

마지막으로 VOC(고객의 소리) 문서에 대한 데이터 분석결과와 실제 고객만족도 점수와는 어떤 관계가 있는지 상호 비교분석을 위해 공항 서비스 핵심 유형별로 비교해보고자 한다.

상호 비교검증을 위해서는 VOC 문서에 대한 데이터 분석결과를 점수화하는 작업이 필요하며, 이를 위해 전체 문서에 대한 긍정, 부정, 중립의 정도를 측정하는 감성분석 결과를 활용하고자 한다.

전체문서를 공항 서비스 유형별로 구분하고, 각각의 서비스 유형별 서비스 긍정지수[2]를 산출하여 고객만족도 조사점수와 상호 비교검증을 실시함으로써 일치하는 서비스 품질 영향요인을 재확인하고, 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출하고자 한다.

2) 서비스 긍정지수 : (긍정+중립건수/전체건수)×100%

제 2 장 이론적 배경 및 선행연구 검토

제 1 절 이론적 배경

1. 공항서비스의 정의 및 중요성

공항서비스란 항공사업법 제2조 및 항공교통서비스 평가업무지침 제2조에 의하면 공항운영자(공항운영의 권한을 부여받은 자 또는 그 권한을 부여받은 자로부터 공항운영의 권한을 위탁·이전받은자) 및 공항상주 기관·업체가 공항에서 제공하는 서비스 등을 말한다. 장대성(2004)에 의하면 공항서비스란 공항에서 수행되는, 그리고 공항과 관련된 모든 서비스의 총체라 정의하고 있다. 또한 국제공항은 단순히 여객과 화물을 실어 나르는 운송터미널의 개념이 아닌 21세기의 국가 경쟁력을 대표하는 복합 서비스 산업이라 정의하고 있다.

우리나라의 다양한 서비스산업 중 공항 서비스 분야는 다양한 조직체들이 서로 유기적으로 연계되어 항공기 운항 계획 수립, 항공기 안전운항을 위한 관제, 보안검색 절차를 통한 수하물 처리, 출입국 관리 및 세관 수속 업무 등 공항의 제반 업무를 수행하는 서비스 컴플렉스 유형의 산업으로 분류된다고 하였고. Ndoh & Ashford(1994)는 공항서비스를 아래 <표 2-1>와 같이 여객 및 화물처리, 대기 장소, 공항 내 이동시설, 부대시설, 접근교통시설 등 5개 부분으로 나누어 분류하였다.

<표 2-1> 공항서비스 분류(Ndoh & Ashford, 1994)

공항서비스 시스템	내용
여객과 화물수속	티켓팅, 체크인, 보안절차, 수하물 처리, 통관 및 세관검사
대기장소	출발장소, 도착장소, 게이트 및 환승라운지
여객터미널내 이동	curb지역, 탑승, 복도, 환승
각종 부대시설	안내, 화장실, 음식점, 은행, 우체시설
접근 교통성	접근속도, 편리함, 비용, 터미널과 연계성

이처럼 공항서비스는 접근 교통 및 주차시설, 탑승수속을 위한 체크인, 보안검색, 출입국 심사, 상업시설 및 편의시설 서비스 등 공항 이용객인 고객이 공항에 도착해서 항공기를 탑승하여 공항을 떠나기 까지 다양한 조직으로부터 제공받는 일련의 서비스를 모두 포함한다고 볼수 있다

앞서 서론에서 논의한 바와 같이, 동아시아의 주변 경쟁 공항들이 과감한 투자를 통한 시설확장과 첨단 스마트 시스템을 선제적 도입하는 등 서비스 경쟁력을 지속적으로 확보하려는 이유는 글로벌 공항서비스 시장에서 도태되지 않고, 기존 고객의 이탈을 최대한 방지함으로써 서비스경쟁 우위를 확보하기 위함이다.

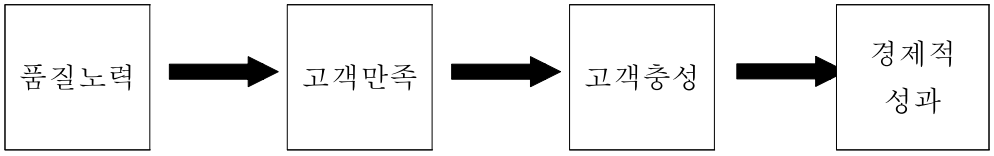
곽상규(2017)는 인접 국가 경쟁공항 대비 공항서비스 수준이 떨어지면, 공항의 인지도 및 구전에 의한 홍보 효과 등에 부정적 영향을 주게 되고 이는 고객의 재방문(재구매) 의도 등을 약화시킬 것이며, 그 결과 경쟁력 저하 및 유휴 시설 발생에 따른 공항 운영 효율 저하로 이어져 막대한 자본이 투입된 국가기반시설의 지속가능한 발전은 어려워질 수 있다고 하였다. 또한 정윤희·박용화·백소라(2013)는 서비스 기업의 고객 이탈 비율이 감소된 경우 이익이 크게 증가할 수 있기 때문에 기존 고객의 유지는 매우 중요하다는 Reicheld and Sasser(1990)의 연구 결과를 인용해 서비스 제공자는 고객의 만족도를 유지 혹은 증대시켜 고객의 이탈을 방지해야 한다고 하였다.

최근의 연구들을 살펴보더라도 서비스 품질이 고객만족의 선행변수로 보는 것이 타당하다는 의견이 지배적이며, 서비스 품질이 고객 만족에 영

향을 미치고 이러한 만족이 재구매 의사나 주변사람에 대한 호의적인 추천, 구전효과에 영향을 미친다고 보는 것이다.(Cronin and Taylor, 1992) 결국 해당 서비스 품질에 대한 고객만족은 고객의 요구와 기대에 부응한 결과로서 재구매나 재방문으로 이어질 것이고 이것은 기업에게 수익성과 함께 지속가능한 경영이 가능하게 할 것이다.

김영찬과 차재성(2003)은 고객만족이란 그 스스로가 다차원적 개념이며 그렇기 때문에 고객만족 개념은 기업의 제품이나 서비스품질과 직접 관련되어 있고 고객을 유지함으로써 기업의 수익성에 기여한다고 하면서 그 측정은 고객만족의 수준과 더불어 그에 의한 경제적 보상과의 관계까지 측정하는 것이 보다 중요하다고 결론 내렸다. 이와 같이 선행변수와 결과변수 그리고 고객만족도간의 관련성에 관한 이론적 근거를 도식화하면 <그림 2-1>의 고객만족도의 선행 및 결과 변인 구조(김영찬,차재성 2005)로 나타낼 수 있다.

<그림 2-1> 고객만족도의 선행 및 결과 변인 구조(김영찬,차재성 2005)



2. 공항서비스 품질 결정요인 및 측정

Gronroos(1984)에 의하면 서비스 품질은 기대된 서비스와 지각된 서비스에 의해 이루지며, Parasuraman·Zeithmal and Berry는 Gronroos의 연구를 기초로보다 세부적으로 서비스 품질 구성요인을 분석하였는데, 신뢰성, 반응성, 능력, 접근성, 예절, 의사소통, 신용도, 안전성, 고객의 이해, 유형성의 10가지 주요 요인을 도출하였고, 이러한 구성요인에 대한 기대된 서비스 수준과 지각된 서비스 수준의 차이가 곧 서비스 품질의 결정요인이 된다고 하였다.(장대성, 2004)

그리고 이후 그들은 품질 구성요인에 대한 실증적인 타당성 검증을 통해 유형성, 신뢰성, 응답성, 반응성, 공감성의 5개 구성요인을 서비스 품질의 척도로서 제시하였으며 이를 측정하는 방법으로서 SERVQUAL을 개발하였는데 이 방법은 현재 서비스 품질 측정 중 가장 일반화된 모형으로 서비스 마케팅 영역에서 널리 사용되고 있다.

그러나 이 SERVQUAL을 사용하기 위해서는 기대수준이 성과를 경험하기 이전에 측정되어야 하며 종단조사 분석을 실시해야 한다고 주장하는 학자도 있다.(이학식, 1997)

장대성(2004)은 기대된 서비스와 지각된 서비스를 측정하려면 동일인을 대상으로 종단분석을 해야 하는데 공항이용자들을 대상으로 동일인을 추적하여 기대수준과 지각된 서비스 수준을 종단 분석한다는 것은 기술적으로 사실상 불가능하다고 보고 있다.

즉 기존의 공항서비스 조사가 설문지에 기대수준과 성과수준을 구분하지 않고 동시에 설문조사를 실시하고 있으며, 서비스를 모두 경험한 이후에 예상 기대수준을 측정하는 것이 되어 이는 서비스 수준 측정에 있어 오류로 작동할 수 있다고 보는 것이다.

이청림·라선아(2015)도 공항서비스 품질을 제대로 측정하고 관리하기 위해서는 공항의 개별 서비스 항목이 아닌 실제로 고객이 공항 서비스를 이용하면서 접하는 다양한 경험들을 종합적으로 고려하여 공항서비스의 품질을 정의해야 한다고 하였다.

현재 국제항공협의회(ACI)의 ASQ 평가, 스카이트랙스(SKY-TRAX)의 공항서비스 평가 그리고 인천국제공항공사에서 자체적으로 시행하는 고객만족도 조사 모두 공항 터미널 내에서 출국, 입국, 환승서비스 이용고객을 대상으로 어권별/시간대별/장소별로 표본을 임의할당하여 일괄적으로 이루어지고 있다. 설문조사를 위한 전문 인력 및 조사 방법을 갖춘 전문기관에서 시행하지만 고객들이 경험하는 다양한 서비스에 대한 만족도를 깊이있게 파악하기에는 한계가 있으며, 또한 서비스 수준에 대한 종단 분석을 하기에는 현실적 어려움이 있음을 알수 있다.

3. 공항서비스 고객 만족도 평가 현황

가. 국제항공협의회(ACI)의 공항서비스 평가(ASQ)

공항과 관련된 가장 대표적인 고객 만족도 평가는 국제항공협회 ACI(Airport Council International)에서 주관하는 ASQ(Airport Service Quality) 평가이다. ACI는 지난 1993년부터 매년 전 세계 356개 주요 공항의 서비스와 시설, 운영 분야 등 34개 부문에 대한 승객들의 즉각적인 인지를 포착하는 공항서비스평가를 실시하고 있다.

ASQ 평가는 전 세계 약 550,000명의 공항 이용객을 대상으로 서비스 전반에 대한 만족도를 조사하는 신뢰 있는 모니터로서 많은 공항에서 공항서비스의 주요 성과지표중 하나로 사용하고 있으며, 광범위하고 정기적인 피드백으로 공항들을 그들의 승객들이 어떻게 평가하고 있는지에 대한 통찰력을 즉각적으로 가질 수 있도록 지원해주고 있다.

또한 각국의 대표 공항들이 자국 내 최신의 시설과 최상의 서비스를 지향하고 있음을 감안할 때 ASQ는 각국의 기술력과 서비스 수준을 가늠할 수 있는 척도이기도 하다. 따라서 각 국의 공항들은 ASQ를 통해 세계 공항의 서비스 실태를 서열화하여 파악하는 것뿐만 아니라 조사에 참여하는 세계 공항들과 정확하게 비교 및 평가하고, 결과에 대한 원인 파악 및 해결방안에 대해 모색할 수 있는 유용한 정보를 얻을 수 있다.

국제항공협의회(ACI)에서는 공항서비스 측정에 있어 공항 이용 여객의 이동 및 편의성, 항공기 운영 및 조정에 주안점을 둠으로써 공항의 전반적이고 효율적인 프로세스를 강조하는 특징이 있다.(임재필, 2011)

인천공항은 2001년부터 ASQ 평가에 참여한 이후 지속적인 서비스 품질에 대한 개선 노력으로 2005년에 1위를 달성한 이후 2017년까지 12년 연속으로 세계 공항서비스 평가 1위를 달성했다. 세계공항서비스평가(ASQ) 12년 연속 세계 1위와 세계 최고 공항상 11회 수상이라는 영예는 세계 공항 역사에 길이 남을 대기록이다.

현재 인천공항은 ASQ 12년 연속 1위라는 대기록을 달성한 이후, 2017년부터 ASQ 공식 평가에는 참여하지 않기로 결정하였다. 그간 ASQ 평가를 통해 세계 최고 공항이라는 이미지를 확고히 했고, 서비스 개선효과도 충분히 누렸다고 판단한 이유가 가장 크지만, ASQ 평가가 아주 미세한 점수 차로 공항별 순위가 결정됨에 따라 현재의 평가 시스템으로는 더 이상 공항서비스 개선에 크게 도움이 되지 않는다고 판단한 측면도 없지 않다.

<표 2-2> ASQ 평가 영역 및 세부 평가 항목

서비스 영역	평가 기준 항목
전반적(종합) 만족도	1. 공항에 대한 전반적 만족도
공항진입	2.공항 목적지/출발지 간 육상교통, 3 주차시설 4 요금대비 주차시설 만족도, 5 수하물 카트 이용의 편리성
탑승수속	6 탑승수속 대기 시간, 7 탑승수속 직원의 능률 8 탑승수속 직원의 친절과 도움
여권 등 비자심사	9 여권/비자 심사 대기시간, 10 여권/비자 심사 직원의 친절성
보안 검색	11 보안검색요원의 친절성, 12 보안검색의 정확성 13 보안검색의 신속성, 14 공항 안전 및 보안의신뢰성
길 찾기 서비스	15 안내표지판을이용한공항내길찾기용이성 16 항공편 안내 스크린 17 터미널 내 이동 편의성 18 환승객 길 찾기 용이성
공항 시설	19 공항직원의 친절/도움, 20 식당시설 21 음식점 만족도, 22 은행/ATM/환전시설 이용 용이성 23 쇼핑시설, 24 가격대비 쇼핑 만족도 25 인터넷 액세스/Wi-Fi, 26 항공사 라운지 27 화장실 이용편의성, 28 화장실 청결성 29 탑승 대기시설의 편안함
공항 환경	30 터미널 청결도, 31 공항의 분위기
도착 서비스	32 여권 및 비자심사, 33 수하물 찾는 속도, 34 세관 심사

나. 스카이트랙스(SKY TRAX) 공항 서비스 평가

1989년에 설립된 스카이트랙스(SkyTrax)는 영국의 항공서비스 전문 리서치 기관으로서 영국 런던에 본사를 두고 있으며, 항공 서비스 리서치 분야에서 세계 최대의 조사 규모와 평가의 객관성으로 세계적인 권위를 인정받고 있다. 홈페이지에서 매년 전 세계 Top 100 공항을 순위로 발표하고 있는데, 공항의 조사 설문에 2000년도에는 약 102만 명이 참가하였는데 반해서 2019~2020년도 조사(조사기간 2019.5월~2020.2월)에서는 100개 이상의 국적에 1,000만 명이상이 참가하였으며 약 550개 공항을 대상으로 서비스를 평가하였다.
스카이트랙스(SkyTrax) 평가는 공항의 출발, 도착, 체크인, 환승, 쇼핑, 출입국 심사, 보안, 등에 포함되어있는 하위 39개 핵심성과지표에 대해서 온라인 조사를 통해 이루어지고 있다. 모든 조사를 온라인으로 진행하다 보니 해당 조사 공항의 서비스를 직접 경험해서 그 지각된 서비스 품질에 대해 평가를 바로 할 수가 없다는 단점이 있으나, 아래 표2-3에서 보는 것처럼 평가 방법이 용이함에 따라 더 많은 설문을 위한 참가자를 확보될 수 있다는 것은 가장 큰 장점이라고 생각된다

<표 2-3>. ASQ, SKYTRAX 서비스 평가 비교

구분	ASQ	SKYTRAX
주관	국제항공협의회(ACI)	영국 SKYTRAX 社 (1989년 설립 항공운송컨설팅사)
평가대상	공항(356개)	공항(550개), 항공사, 공항호텔
평가항목	34개	39개
평가방법	면접원에 의한 1:1 인터뷰	온라인 설문, 인터뷰, 전화 등
설문	출국여객(35만명)	여객, 전문가 집단, 여행사 등(1000만명 이상)
평가기간	매년 1월 ~12월(12개월)	전년 5월~ 당해년 2월(10개월)
분야 및 시상	8개 카테고리 34개 부문시상	6개 카테고리 51개 부문시상

4. 빅데이터와 분석기법

가. 빅데이터의 정의와 빅데이터 분석의 의의

빅데이터란 일반적으로 데이터의 양이나 단위의 관점에서 기존의 관리 및 분석 체계로는 감당할 수 없을 정도의 거대한 데이터의 집합을 의미(Manyika et al., 2011)한다. 빅데이터의 특징은 3V로 요약하는 것이 일반적이며, 데이터의 양을 의미하는 규모(volume), 데이터의 다양한 형태를 나타내는 다양성(variety), 데이터가 생산, 처리 및 분석되는 속도(velocity)에 주로 초점을 맞추어 3V를 갖춘 경우를 빅데이터로 정의하였고,(Laney, 2001) 이후에 날씨나 경제 등과 같은 데이터 자체의 본질적인 불확실성 개념인 정확성(veracity) 개념을 추가하여 4V로 정의하기도 하고(정지선, 2011), 최근에는 어떤 정보를 발견하는데 사용할수 있는가라는 목적으로서의 가치(value) 측면을 추가하여 빅데이터를 5V로 표현하기도 한다(White, 2012).

또한 사회 과학에서 빅데이터는 연구자가 전체 데이터의 대표적인 표본(arepresentative sample of the entire dataset)을 직접 수동 코딩(manual coding)하는 것이 어려울 정도의 거대한 양을 의미한다(Riffe, Lacy & Fico, 2014, p.92)고 알려져 있다.

빅데이터는 생성주체, 데이터 유형과 데이터 수집방식에 따라 아래의 표 2-4 와 같이 분류될 수 있다. 최근에는 모바일, 클라우드, 소셜 네트워크 서비스 등의 등장과 확대로 인해 숫자 형태의 정형데이터뿐만 아니라 인간관계 데이터, 위치, 행태, 인식 등 사람과 사회 현상을 이해할 수 있는 텍스트나 이미지, 동영상 등의 비정형데이터가 다양하고 활발하게 생성되고 있다(배동민, 박현수, 오기환, 2013) 또한 기술의 발전과 함께 이러한 다양한 형태의 데이터를 효과적으로 축적하고 공유할 수 있게 되면서 이를 분석하고 예측, 전망하여 새로운 가치를 창출하고자 하는 요구가 증가하고 있는데, 이에 따라 데이터의 규모나 단위 측면에서보다 원하는 가치를 얻을 수 있는 데이터를 빅데이터로 해석하려는 경향이 증대되고 있다(송상영, 김사혁, 2012)

빅 데이터 분석기술은 기존 데이터베이스 관리도구로 데이터를 수집, 저장, 관리, 분석할수 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 작동하게 하고 개인화된 현대 사회 구성원 마다 맞춤형 정보를 제공, 관리, 분석 가능하게 하며 과거에는 불가능했던 기술을 실현시키기도 한다

이처럼 다양하고 방대한 규모의 데이터는 미래 경쟁력의 우위를 좌우하는 중요한 자원으로 활용될 수 있다는 점에서 최근 주목받고 있으며 현재의 빅데이터 환경은 과거와 비교해서 데이터의 양은 물론 질과 다양성 측면에서 패러다임의 전환을 의미하고 있다.

<표 2-4> 빅데이터 분류

분류기준	만족도 평가지표	종류
생성주체	기계데이터 사람데이터 관계데이터	어플리케이션 서버 로그데이터, 센서 데이터, 위치데이터 등 트위터, 블로그, 이메일, 게시판 등 트위터 페이스북 링크드인 등
데이터 유형	정형데이터 비정형데이터 반정형데이터	고정된 필드에 저장된 데이터: 데이터베이스, 스프레드 시트 등 고정된 필드에 저장되어 있지 않은 데이터: 텍스트 문서, 이미지/동영상 등 고정된 필드에 저장되어 있지는 않지만 메타 데이터나 스키마를 포함하는 데이터: XML, HTML등
데이터 수집 방식	내부데이터 외부데이터	자체적으로 보유한 내부 파일시스템이나 데이터베이스 관리 시스템 등에 접근하여 데이터를 수집 웹크롤링 엔진을 통해 인터넷 링크의 모든 페이지복사본을 생성함으로써 데이터를 수집

출처: 이지영(2015). 빅데이터의 국가통계 활용을 위한 기초연구.

나. 텍스트 마이닝 기법의 개념 및 주요절차

본 연구에서는 빅데이터 분석기법 중 텍스트 마이닝(Text Mining) 기법을 활용하고자 한다. 텍스트 마이닝(Text Mining)은 텍스트(Text) 데이터에서 유용한 정보를 캐낸다(Mining)는 의미로, 비정형 텍스트로 구성된 각 문서에 포함된 특징(feature)을 추출하고 이들의 패턴이나 연관성을 파악하여 활용하는 비정형 데이터 분석이다.

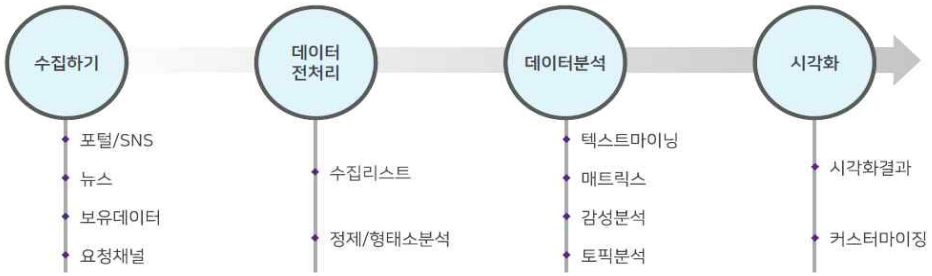
텍스트 마이닝은 책, 문서, 신문기사, 블로그, SNS 등의 비정형데이터를 분석 대상으로 하며, 자연어처리 기술을 기반으로 텍스트에 나타나는 단어를 이용하여 분석을 수행하는데 텍스트 요약, 문서검색, 정보검색과 같은 텍스트 마이닝 작업을 통해 정보 추출이 가능해지고, 텍스트 분류, 문서 군집화, 언어인식, 핵심문구 식별 등 문서의 유사성을 측정하는 방식을 확대되기도 한다.(Khan & Kanth, 2016).

텍스트 마이닝은 대량의 텍스트를 활용할 수 있어, 기존 연구들이 연구자의 역량에 의존하여 일부의 텍스트에 대해 제한적으로 접근하므로 이전에는 볼 수 없었던 숨은 지식을 찾아내는 것이 가능한 연구방법으로서 핵심적인 가치가 있다(Mahgoub, Röner, Ismail & Torkey, 2008). 특히 정확하게 파악하기 어려운 사회 현상 및 이슈와 관련된 다양한 논의 및 그 변화를 파악하는데 유용하게 사용된다(Hale, 2005).

또한 텍스트 마이닝(Text Mining) 기법은 다양한 웹 환경에서 원하는 기간의 데이터를 수집하고, 데이터 전처리 과정(정제/형태소 분석) 및 데이터 분석과정(텍스트 마이닝)을 통해 매트릭스, 감성분석, 토픽모델링, 시각화 등 원하는 결과물을 산출할 수 있다는 장점이 있다.

아래 <그림 2-2>은 일반적인 텍스트 마이닝 기법의 처리과정을 나타낸다.

<그림 2-2> 텍스트 마이닝의 주요과정



자료 : 텍스톰(텍스트 마이닝 기법) 소개서 발췌

1) 키워드 분석

텍스트 마이닝은 기본적인 텍스트 분석 기법인 키워드(keyword) 분석으로부터 시작되며, 이를 기초로하여 보다 발전된 형태의 분석을 수행하여 기존에 발견하지 못한 숨겨진 지식이나 정보를 찾아내는 특징이 있다. 대표적인 키워드 분석 방법은 다음과 같이 분류할 수 있다.

첫째, 단어 빈도 분석(term frequency analysis)은 가장 기초적인 텍스트 분석 기법으로 미리 정의된 분류 항목이나 단어 정보에 기초하여 텍스트를 분석하는 것이다(Li, 2010). 텍스트로부터 자연어 처리(NLP) 과정을 거쳐 전처리를 통해 추출된 단어들을 활용해 주어진 문서들에서 가장 많이 혹은 가장 적게 나타난 단어를 찾아내는 방법으로, 추출된 단어들과 빈도는 분류 항목 간 비교나 요약 정보 등을 나타내는 지표가 될 수 있다(Krippendorff, 2004, p.306). 해당 단어를 포함하고 있는 문서의 수를 구할 수도 있고, 하나의 문서에 해당 단어가 포함된 빈도를 구할 수도 있다.

둘째, 단어 연관성 분석(term association analysis)도 기초적 키워드분석 기법 중 하나로 분류된다. 어떠한 단어를 포함한 문서 내에서 해당 단어와 함께 나타나는 단어들을 분석, 추출하여 해당 단어들과 어떤 단어들이 관련되어 연결성을 갖고 내용을 구성하는지를 볼 수 있는 방법이다(이미숙, 이창훈, 김지연, 2014, p.41).

워드 클라우드 이미지 등으로 시각화하여 직관적으로 단어의 상대적 출현 정도(Benoit & Herzog, 2015) 등을 제시할 수 있다.

셋째, 정보 검색(IR) 및 기계 학습에서 많이 사용되는 TF-IDF도 문서 집합 안에서 사용된 단어의 중요성을 나타내는 측정치로서 키워드 분석에 주로 사용된다.

TF-IDF는 단어별 가중치로서, 여러 문서 집합으로부터 특정 단어가 특정 문서에서 얼마나 중요한가를 판단할 수 있는 값으로 간단하지만 효과적인 값으로 알려져 있다(Salton & Buckley, 1988).

구체적으로 살펴보면 TF(Term Frequency)는 특정 단어의 문서 내에 등장 빈도수로, 해당 값이 높을수록 문서 내 중요도가 높다고 판단할 수 있고, DF(Document Frequency)는 특정 단어의 여러 문서 내 등장 빈도수로, 해당 값이 높을수록 해당 단어가 보편적으로 사용된다고 판단할 수 있다. IDF(InverseDocument Frequency)는 DF의 역수이다.

정리하자면 TF-IDF는 TF의 값과 IDF의 값의 곱으로 특정 단어의 문서 내 중요한 정도를 의미한다.

본 연구에서는 TF-IDF 방법으로 키워드의 중요도를 계산하고, 각 키워드를 순위화 하여 분석 결과를 도출하고자 한다.

2) 토픽 분석(토픽 모델링)

텍스트 마이닝의 보다 발전된 형태인 토픽 모델링(topic modeling)은 통계에 기반을 두고 주어진 문서나 텍스트에서 잠재된 토픽을 찾아내는 방법론이다. 각종 비정형 문서로부터 정형화된 정보 및 학습 규칙, 개체명 등을 추출하기 위해 텍스트로부터 모형(model)을 만들어 기존의 키워드 분석으로 찾아낼 수 없었던 의미를 찾고자 하는 시도로부터 시작된 기법이다. 이렇게 추출된 정보는 주어진 텍스트에 나타나는 단어만을 이용하기 때문에 그 자료들을 해석하고 이용하기에 적합도가 높고 효율적일 수 있지만, 다른 텍스트나 유사한 상황의 다른 자료들에서도 항상 최적의 결과가 나오지는 않는다(Provost & Fawcett, 2013). 따라서 문서에 존재하는 복잡성을 고려할 필요가 있고 단순히 정확한 구체

적 단어와의 불일치로 인해 분석한 텍스트 정보를 효율적으로 이용하기 어려운 경우에는, 문서와 단어만을 이용하여 모형을 만드는 것이 아니라 문서와 문서에 나타나는 단어 사이에 토픽 계층이 존재한다고 가정하고 이 잠재 정보(latent information)를 추출하여 모형을 만드는 토픽 모델링(topic modeling)을 이용해 토픽(topic)을 분석한다. (Blei, Ng & Jordan, 2003)

토픽 모델링은 최근에 정립, 발전하고 있는 텍스트 마이닝 분석 방법이며 가장 주목받는 방법으로 알려져 있다(Blei & Lafferty, 2009)

본 연구에서는 토픽 모델링 방법 중 현재 가장 널리 사용되고 있는 Blei et al.(2003)이 개발한 Latent Dirichlet Allocation(LDA) 모델을 이용한 토픽 클러스터링 분석을 활용하여 문서에 포함된 토픽을 도출하였다. LDA 모델은 문서에 수동으로 레이블링 작업을 할 필요가 없는 비지도 학습 알고리즘으로, 문서-토픽-단어 3계층으로 베이지안 모형이며, 기존 토픽 모델링 알고리즘보다 성능이 우수한 것으로 알려져 있다. (Chiru et al., 2014, pp.255-262).

LDA는 토픽의 개수(K)를 입력하여 문서에 포함된 단어의 분포, 즉 단어-문서 행렬을 가장 잘 설명할 수 있는 문서-토픽 행렬과 토픽-단어 행렬을 계산하는 방법론이다.

이 행렬을 바탕으로 하여 각 문서가 토픽에 포함될 확률 및 각 단어가 토픽에 포함될 확률을 구할 수 있으며 최종 결과로 각 토픽에 포함된 단어들이 나타나며, 이 단어들을 보고 연구자가 주관적으로 판단하여 토픽을 결정하는 방식의 알고리즘이다. 또한 토픽 모델링은 대용량의 텍스트에서 토픽에 포함되어 있는 단어들을 바탕으로 잠재되어 있는 토픽의 추출이 가능할 뿐만 아니라, 그 토픽들이 시간의 흐름에 따라 어떻게 변화하는지를 시계열적으로 분석이 가능한 통계기법이다.

토픽모델링 분석에서는 최적 토픽수의 결정은 분석 결과의 해석에 영향을 주기 때문에 매우 중요하다.(Steyvers & Griffiths, 2007) 따라서 분석하고자 하는 문서에 대한 최적의 토픽수를 결정하기 위해서는 토픽의 개수를 조정해 가면서 복잡도(Perplexity)와 정합도(Coherence)를 기준으

로 하여 적정 토픽수를 연구자가 판단하여 결정해야 한다. 일반적으로 토픽의 개수를 늘려갈수록 복잡도(Perplexity)의 값이 작아지는 경향을 보이며, 토픽모델링 분석으로부터 양질의 학습결과를 얻어낼 수 있다.
최근에는 복잡도(Perplexity)만 고려해서는 토픽모델링 분석으로부터 얻어진 모형의 결과가 좋다고 단정하기 어려워, 주제의 일관성을 반영하는 정합도(Coherence)도 함께 고려해야 한다는 연구결과가 있다. 정합도는 토픽모델링 분석결과로부터 나타난 상위 빈출 키워드간의 유사도를 기준으로 결정되며, 정합도가 높을수록 토픽과 키워드들간의 일관성이 높고 해당 토픽을 잘 반영하고 있음을 보여준다.
토픽모델링(LDA) 분석기법을 활용한 연구들은 최근 정책학, 교육학, 언론정보학, 사회과학 등 다양한 학문 분야에서 이루어지고 있다. 호텔 레스토랑 같은 개인서비스 뿐만 아니라 공공서비스 분야에 있어서도 고객만족도를 분석하기 위해 SNS 이용후기나 민원데이터를 활용하고 있으며, 정책학 관련 분야에서도 정책의 수요와 관련 이해주체의 동향을 파악하기 위해 각종 언론자료 및 기사, 소셜 네트워크 데이터에 대한 토픽 모델링 분석을 시행하고 있다.

제 2 절 선행연구 검토

1. 공항서비스 품질과 고객 만족도에 관한 연구

공항서비스와 고객만족에 대한 선행연구를 살펴보면, 대부분 서비스품질이 고객 만족의 선행요인으로 밝히고 있으며, 어떠한 품질요인이 고객만족에 가장 큰 영향을 미치고 있는지에 대한 연구가 이루어져 왔다. 즉 서비스 품질에 대한 고객의 기대 서비스 수준의 내용과 고객 만족에 대한 내용으로 구분하여 살펴볼 수 있고, 주로 공항의 서비스 품질에 대한 고객의 기대 서비스 만족도에 영향력이 큰 서비스 요인이 무엇인지를 밝히고자 하였다.

장대성(2002)은 Cronin & Taylor(1992)의 SERVPERF 모형에서 말한 대로, 서비스품질은 태도에 의해 개념화되어야 하고 측정하는 것이 고객의 만족의 관점에서 측정하는 것보다 더 일관성이 있다는 주장을 근거로 하여 인천국제공항과 김포국제공항을 이용하는 내국인을 대상으로 설문조사를 통해 서비스품질이 고객만족에 미치는 영향에 대해 연구하였다. 서비스 품질은 공항시설, 운영절차, 신뢰성, 반응성, 보증성, 공감성, 지리적 위치, 안전성 등 8개 요인으로 나누었는데, 회귀분석 결과 공항이용객의 만족에 중요한 영향을 끼치는 서비스 품질은 공항 근무자의 반응성과 공항의 안전성이었다.

최현정(2005)은 연구를 SERVQUAL의 서비스 결정 요인들에 대해서 SERVPERF에 적용 가능한 연구모형에 맞게 측정항목을 도출하여 공항서비스의 실정에 맞도록 보완, 수정해서 항목을 구성하였으며, 인구통계학적 특성은 공항서비스 만족도 차이는 직업에 의해 차이가 있는 것으로 나타났다. 시설 품질의 효율, 인적 서비스 품질, 운송시스템 품질의 효율이 더 높을수록 공항서비스 만족도 또한 높아지는 것으로 나타났다.

한편, 최병권(2007)은 그간의 공항관련 연구가 여객의 서비스 품질 구성요인에 치중되어 있었으므로, 이와 다르게 공항서비스 이용자 중 국제특송 업체, 물류 업체 등 항공운송업체를 대상으로 하여 서비스품질과 만족도에 관한 연구를 실시하였다. 항공운송 업체를 대상으로 서비스 품질요인 및 만족도에 관한 설문을 실시하였고 서비스품질 요인으로는 유형성, 신뢰성, 응답성, 확신성, 공감성 등 다섯가지 품질요인으로 나누었다. 회귀분석 결과 서비스품질 중, 공감성, 확신성이 전반적인 만족도에 영향을 가장 많이 끼치는 것으로 나타났다.

홍석진, 이재환(2007)은 인천국제공항의 서비스 품질을 여객터미널, 상업시설, 여객터미널 접근의 편리성, 항공사, 인천국제공항공사의 지역사회 공헌 등으로 구분하고 공항서비스 종합 만족도에 영향을 주는 요인은 무엇인지 분석한 결과 '여객터미널 서비스 품질'이 전체 만족도에 가장 큰 영향을 미친다고 보았다.

또한, 장순자·김기웅(2010)은 인천국제공항과 김포국제공항을 이용하는 한·중·일3개국 여객을 대상으로 설문조사를 통해 공항의 서비스품질 차

이와 국적별로 공항 서비스품질에 대한 평가를 비교하였고, 전반적인 만족도와 국적별 공항별 차이 분석 및 만족도에 영향을 주는 주요 품질요인은 무엇인지 연구하였다. 전반적인 만족도에 영향을 주는 품질요인 세 가지는 정보의 정시성, 시설운영서비스, 제공정보의 내용 순이었다.

연구 수행측면에서 살펴보면 기존의 선행연구들은 주로 서비스 품질의 측정도구 개발 및 서비스 품질요인이 서비스 만족도에 미치는 영향에 대한 분석과 연구가 다수를 차지하였고, 서비스 불만족 및 서비스 실패에 대한 분석 및 회복노력에 대한 연구는 매우 드물었다.

서비스 만족도에 대한 영향요인을 분석하기 위해 보편적으로 활용되고 있는 서비스 품질 항목들에 연구자가 몇가지 항목들을 추가하여 서비스 품질요인을 분석하였고 그 연구결과는 각기 다양하게 나타났으며 객관적이기 보다는 주관적인 측면이 많이 작용하였음을 확인할 수 있고, 서비스 품질요인의 우선순위(또는 중요도)에 있어서도 연구논문마다 각기 다른 결과를 보여주고 있다.

또한 대부분의 연구가 설문조사 방식으로 진행됨에 따라, 많은 표본수를 확보하기가 어렵고, 단기간에 조사가 이루어져 조사 데이터가 모집단의 특성을 정확하고 깊이있게 반영하기 힘들며 시계열적으로 분석하는데 있어 많은 한계점을 가지고 있었다.

그리고 고객만족도 조사의 대상 또한 서비스 이용에 큰 불편함이 없는 일반고객층을 대상으로 하다 보니 고령자, 장애인, 영·유아동반자, 임산부 등 교통약자 입장에서 서비스 불만족의 원인을 파악하기 힘들었으며, 그동안 그들을 위한 서비스 개선은 거의 이루어지지 않았다.

최근에 우리나라도 초고령화 사회에 진입함에 따라 통행에 불편함을 겪는 교통약자의 이동권 증진에 관심이 고조되고 있으며, 이들의 이동권 확보와 원활한 사회활동을 지원하기 위한 연구와 정책수립의 필요성이 점차 중요시되고 있다.

<표 2-5> 고객만족도 관련 선행연구 검토

논문 제목	연도 및 저자	내용
국제공항 서비스 품질 측정도구 개발에 관한 연구	장대성, 2004	독립변수: 공항시설,운영절차,신뢰성,반응성, 보증성, 공감성, 지리적위치,안전성 종속변수: 전반적 고객만족도
복수공항 이용객의 서비스 만족에 영향을 미치는 품질요인의 비교연구	장순자, 김기웅, 2010	독립변수: 운영서비스, 인적서비스, 정보서비스 종속변수: 고객만족도
공항서비스 품질이 이용객 만족도에 미치는 영향에 관한 연구	백운일, 2009	독립변수: 시설편의/외적심미성, 반응/보증성, 공감성, 신속정확성, 교통편의성 종속변수: 이용객만족도
공항 서비스품질 및 고객만족 모형의 개발	이청림, 라선아 2015	독립변수: 전반적 서비스 품질(도착,체크인,보안검색,출입국심사,부대시설,탑승수속,세관심사,정보확인) 종속변수: 고객만족도
공항서비스 만족도 영향요인에 관한 연구	곽상규, 2017	독립변수:출입국심사유형(자동OR일반), 서비스 품질요인(시설/인적) 종속변수: 서비스 전반적 만족도
공항의 기대 서비스 수준 및 고객만족에 관한 연구	천정이, 2019	독립변수: 기업요인(서비스직원,기업이미지), 개인요인(니즈,과거경험),사회적요인(구전, 사회적상황) 종속변수: 기대서비스 수준
스마트서비스가 여객만족도에 미치는 영향	김지은, 2020	독립변수: 스마트서비스(셀프체크인,셀프백드롭,로봇안내) 종속변수: 고객만족도, 사용의도

2. 빅데이터 분석기반 고객만족도 연구

빅데이터를 활용한 국내 연구는 서비스 산업 분야 뿐만 아니라, 정책학, 교육학, 언론정보학, 사회과학 등 다양한 학문 분야에서 고객만족도 제고 및 정책 수요나 관련 주제의 동향을 살펴보기 위해 주로 소셜 데이터(social data)와 뉴스기사, 개인 SNS 등을 활용하거나 혹은 기존 기업에서 보유하고 있는 데이터를 기반으로 연구가 진행중에 있다.

공효순(2014)은 서비스 산업의 대표 기업인 호텔기업의 고객피드백을 파악하기 위해서 국내의 대표적인 특급호텔 3개를 선정하여 개인블러그나 SNS상의 호텔 관련 빅데이터로부터 호텔 이용후기 등 고객의 의견들을 수집·분석하여 호텔 고객관계관리(CRM)에 필요한 개선점을 도출하였다.

이동준(2016)은 국내외 택배업체들의 서비스 품질개선을 위해 SNS상의 빅데이터를 활용하여 국내외 택배업체들에 대한 고객의견을 수집하고, 이를 국가간 고객만족도 점수와 비교·분석함으로써 SNS 빅데이터를 활용하여 택배업체 고객만족도 분석에 활용할 수 있다는 가능성을 제시하였다

한주형(2019)은 가격이 중저가 호텔과 게스트하우스의 고객만족도 평점에 미치는 영향을 분석하기 위해 온라인 리뷰 빅데이터를 활용하였고, 다른 형태의 숙박시설을 비즈니스 단위에서 비교 연구함으로써 마케팅 전략수립의 기초자료로 활용하고자 하였다.

엄준석(2020)은 지방 상수도 고객 민원자료에 대한 빅데이터 분석을 통해 고객민원과 고객만족도를 상호 연계함으로써 고객만족도 예측 모델과 고객 불만을 이용한 불만 조기 경보 시스템을 개발하고자 하였다.

또한 한국환경정책·평가연구원(2019)에서는 생활밀착형 환경 이슈를 파악하기 위해 국민들의 소통창구인 민원 데이터를 수집 및 분석하여 환경 이슈에 대한 국민들의 정책수요를 파악하였고, 궁극적으로 공공정책 수립시 정책 의제 설정 및 정책평가에 반영 가능함을 시사하였다.

이와 같이 기존 빅데이터를 활용한 연구들은 고객의 관심사항과 불만사항 등의 주요 이슈를 파악하는데 있어 정제되지 않은 빅데이터로 부터 고객 서비스 개선에 유용한 정보를 추출해냄으로써 고객만족도 제고를 위한 다양한 이슈를 도출하고자 하였다.

<표 2-6> 텍스트마이닝 분석 선행연구 검토

논문 제목	연도 및 저자	내용
빅데이터 기반 호텔고객 평판분석에 관한 연구	공효순, 2014	국내의 대표적인 특급호텔 3개를 선정하여 SNS상의 고객 이용후기 데이터 수집, 분석하여 호텔 고객관계관리(CRM)에 필요한 개선점 도출에 필요한 개선점 도출
호텔의 여가소비 확장에 관한 연구	엄혜란, 2018	호텔의 여가소비공간적 역할과 도시적 의미에 대한 이해를 도목하기 위해 호텔공간에 대한 도시민의 이용 행태 분석
가격이 중저가 호텔과 게스트하우스의 고객만족도 평점에 미치는 영향	한주형, 2019	온라인 리뷰를 활용하여 서로다른 형태의 숙박시설을 비즈니스 단위에서 비교 연구함으로써 마케팅 전략의 기초자료로 활용
소셜 네트워크 빅데이터 기반 택배업체 고객만족도에 관한 연구	이동준, 2016	국내외 택배업체들의 서비스 품질개선을 위해 SNS 빅데이터 활용하여 택배업체의 고객만족도 분석에 활용할 수 있다는 가능성 제시
텍스트 마이닝을 활용한 도시철도 이용자 서비스 니즈 분석	조시환, 2020	도시철도와 관련된 SNS, 뉴스기사,논문 등을 대상으로 텍스트 마이닝 분석을 통해 도시철도 이용자의 잠재적 니즈 파악
지방 상수도 고객민원 빅데이터 분석	엄준석, 2020	상수도 고객민원자료에 대한 빅데이터 분석을 통해 고객만족도 예측모델 제시
생활밀착형 환경이슈에 대한 수요반영개선연구(민원빅데이터 분석중심)	진대용, 2019	환경 이슈에 대한 민원데이터를 수집 분석하여 공공정책 수립시 정책의제 설정 및 정책평가 반영 가능성 시사
청와대 국민청원의 현황과 과제(빅데이터 분석결과를 중심으로)	엄석진, 2019	청와대 국민청원 및 제안 사이트에 접수된 빅데이터를 수집하고, 청원성격/동의추세/핵심어 등을 분석함으로써 국민청원의 현황 및 향후 과제 제시
북한 김정은 집권 전후의 외교정책변화 추이분석	박성렬, 2017	북한의 대내외 선전매체인 노동신문에 대한 내용분석을 통해 북한 원문자료에 대한 디지털 아카이빙과 일반공개의 필요성 제시
환경복지 정책모형 설정에 관한 탐색적 연구(텍스트 마이닝 중심)	채미현, 2018	환경복지를 주제로한 관련문헌을 질적분석하여 공통의견을 추출하고 모형으로 재구성하는 탐색적 기법수행
텍스트 마이닝을 활용한 고용정책 프레임 연구	김경동, 이시영, 고길곤, 2020	언론기사와 국회회의록 자료를 바탕으로 토픽모델링과 네트워크 방법을 통해 고용정책의제와 프레임에 대한 시계열적 분석시행

3. 본 연구의 차별성

그 동안 인천공항의 공항서비스 만족도와 관련해서는 다양한 연구가 있었다. 하지만 대부분의 연구가 공항서비스 만족도에 관한 품질요인을 측정하기 위해 기존의 고객만족도 조사결과를 활용하거나 새롭게 설문조사를 시행하는 방식으로 연구가 진행되었으며, 서비스 불만족 및 서비스 실패에 대한 분석 및 회복노력에 대한 연구는 매우 드물었다.

고객만족도 조사에 있어서 설문조사방식의 연구는 많은 표본수를 확보하는것이 실질적으로 어렵고, 단기간에 이루어짐에 따라 시계열적 분석이 어려우며, 조사 데이터가 대상 주제에 대해 세부적인 설명력 부족으로 모집단인 고객의 실제 불만사항을 정확하고 깊이있게 반영하지 못할 가능성이 높아 연구의 타당성 확보에 어느정도 한계를 가지고 있다.

본 연구는 선행연구와 차별성을 두어 최근 5년간 인천공항 VOC(고객의 소리)시스템에 접수된 고객 민원내용을 서비스 유형별로 혼합방법적으로 분석함으로써 공항서비스에 대한 고객의 실질적인 불만사항과 그 원인을 분석해보고자 한다.

특히 그동안 설문조사가 실질적으로 어려웠던 교통약자(고령자, 장애인, 영유아동반자, 임산부 등)의 민원내용에 대한 데이터 분석이 가능하므로 교통약자의 입장에서 그들이 필요로 하는 이동 편의 서비스 유형과 불편·불만족 사항이 무엇인지를 직접적으로 파악함으로써 교통약자의 이동편의성 증진에 크게 기여할 수 있을 것이다.

또한 기존의 빅데이터 분석 관련 연구의 대부분은 주로 특정 키워드를 중심으로 연관성이 있는 단어들의 출현 빈도가 높은 키워드를 중심으로 해당시기의 토픽을 추론함으로써 그 자료에서만 정확도가 높고 유사한 상황의 다른 자료에서는 효율적으로 기능하지 못할 수 있었다.(Provost & Fawcett, 2013)

본 연구는 선행연구와 차별성을 두어 고객민원자료에 대한 빅데이터 분석시 통계적으로 대규모 자료를 종합하는데 가장 효율적인 토픽 모델링 기법을 사용함으로써 단순 키워드 분석의 한계를 극복하고 잠재된 토픽

계층을 설정하고자 하며, 토픽계층의 시계열적 분석을 통해서는 고객 불만족 유형의 변화 추이를 파악할 수 있을 것이다.
또한 빅데이터 분석을 통해 고객 불편·불만족사항에 대한 주요 이슈를 도출하는데 그치지 않고, 주요 서비스 유형별로 감성분석을 통해 고객들의 서비스 평가지수를 도출해보고 이를 실제 고객만족도 점수와 비교해 보고자 한다. 이를 통해 VOC(고객의 소리) 데이터를 이용하여 고객만족도 조사 결과를 미리 예측할 수 있는 연구기반을 마련하고자 하며, 이는 향후 기존 설문조사 방식의 고객만족도 평가의 여러 제약조건들을 보완할 수 있을 것으로 생각한다.

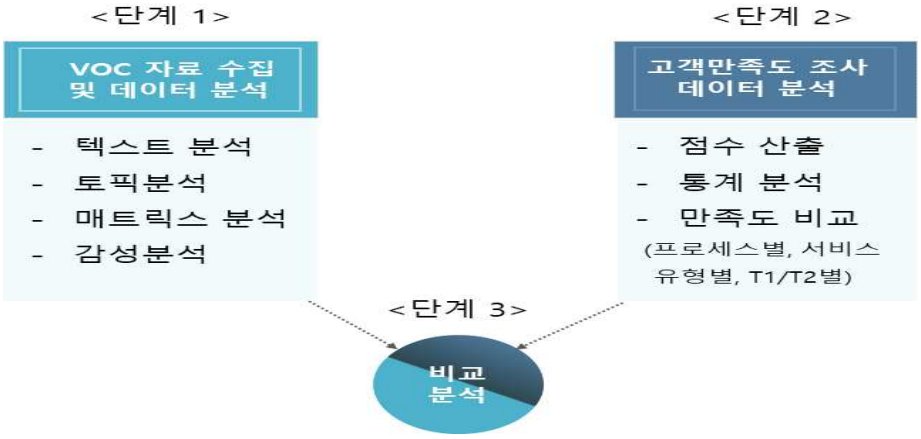
제 3 장 연구모형 및 설계

제 1 절 연구의 분석틀

1. 연구 모형

본 연구는 인천국제공항에서 제공하는 공항서비스의 핵심요소인 체크인, 보안검색, 출입국심사, 교통/주차 등의 서비스를 중심으로 해당 서비스의 고객만족도에 영향을 미치는 서비스 품질요인을 분석하기 위해 먼저 인천공항 VOC(고객의 소리) 시스템에 접수된 고객 불만족과 관련한 데이터를 텍스트마이닝 기법을 활용하여 분석하고자 한다.
다음으로 VOC 데이터 분석결과와 설문조사 방식의 고객만족도 조사결과 간에는 어떠한 공통점과 차이점이 있는지를 확인하기 위해 상호 비교검증을 실시하고자 한다. 이를 위해 공항 핵심 서비스 유형별로 해당 문서의 서비스 긍정지수를 산출하고 고객 만족도 점수와의 상관관계를 분석하고자 하며 구체적인 연구모형은 아래 <그림 3-1>과 같다.

<그림 3-1> 공항서비스 고객만족도 분석 연구모형



2. 연구 문제

본 연구를 통해 알아보고자 하는 연구 문제는 구체적으로 다음과 같다.

① 최근 5년간 VOC(고객의 소리) 시스템에 접수된 문서내용의 주요 키워드들과 토픽들은 무엇인가 ?
② 주요 토픽들은 시간의 흐름에 따라 어떻게 변화하며, 관심이 증가하는 토픽들과 감소하는 토픽들은 무엇인가 ?
③ 최근 5년간 VOC(고객의 소리) 시스템에 접수된 문서내용의 부정적 키워드, 긍정적 키워드는 각각 무엇인가 ?
④ VOC 빅데이터 분석결과와 실제 고객만족도 조사 간에는 어떠한 공통점과 차이점이 있는가 ? 품질개선이 필요한 사항은 무엇인가 ?

제 2 절 자료의 수집 및 분석방법

1. 표집방법

본 연구는 먼저 빅데이터 분석을 위해 아래의 <표 3-1>과 같이 최근 5년간 인천공항공사 홈페이지에 등록된 고객의 소리(VOC) 약 30,214건의 데이터를 활용하고자 한다. VOC자료는 인천국제공항공사의 서비스를 제공받은 고객이 공사에 제기하는 각종 불편불만, 칭찬격려, 상담문의, 의견제안 등의 내용으로 홈페이지, 이메일, 엽서 등 온라인 및 오프라인을 통해 접수되고 있다.
고객만족도 조사는 아래의 <표 3-2>와 같이 최근 5년간 인천국제공항공사에서 매년 분기별로 4차에 걸쳐 공항 이용객을 대상으로 실시한 결과를 분석대상으로 활용하고자 한다. 설문조사는 구조화된 설문지에 의한 1:1 개별면접조사로 이루어졌으며, 5점척도로 서비스유형별 만족도를 측정한 후 100점 만점으로 환산하여 만족도 점수를 산출하였다. 연간 표본수는 내·외국인을 포함하여 출국여객, 입국여객, 환승여객, 항공사 서비스 이용여객 등 총 20,357명을 대상으로 설문조사를 진행하였다.

<표 3-1> 인천공항 VOC (고객의 소리) 현황

구분	'15년 하반기	'16년	'17년	'18년	'19년	'20년 상반기	합계
건수	2,228	5,426	5,974	7,380	6,589	2,617	30,214

<표 3-2> 인천공항 고객만족도 조사 현황

구 분	고객만족도 조사
조사대상	출국, 입국, 환승, 항공사, 문화예술 서비스 이용고객
조사기간	매년 1~4분기 (7일/분기당)
표 본 수	전체 20,357명
조사방법	1:1 면접조사
조사내용	고객체감 전반 만족도

2. 표집현황

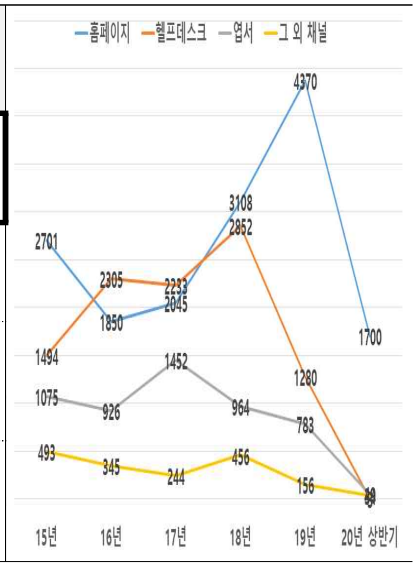
가. 인천공항 VOC(고객의 소리) 현황

인천공항 고객의 소리(VOC) 시스템으로 2015년 부터 2020년 상반기까지 5년동안 약 3만건 이상의 민원이 접수되었으며, 아래 <표 3-1>을 통해 민원 건수가 매년 증가하고 있음을 알 수 있다. 2018년에는 인천공항 제2여객터미널 오픈('18. 1월)에 따라 여객터미널 분리운영에 따른 혼선으로 민원이 급증하였고, '20년 상반기에는 코로나 사태로 인하여 여객수가 급감함에 따라 민원 발생건수가 일부 감소하였음을 알 수 있다. 또한 VOC 접수 채널은 인천공항 홈페이지, 헬프데스크, 엽서 등을 통해 접수하고 있으며, 고객이 민원 접수방법에 대해 문의

시 홈페이지를 통한 VOC 등록을 홍보함에 따라 홈페이지를 통한 접수 비중은 점차 늘어가는 반면에, 헬프데스크를 통한 민원 접수건수는 점차 감소추세에 있음을 알 수 있다.

<표 3-3> 고객의 소리(VOC) 접수채널 [단위 : 건]

구분	'15년	'16년	'17년	'18년	'19년	'20년 상반기
홈페이지	2,701 (47%)	1,850 (34%)	2,045 (34%)	3,108 (42%)	4,370 (66%)	1,700 (95%)
헬프 데스크	1,494 (26%)	2,305 (42%)	2,233 (37%)	2,852 (39%)	1,280 (19%)	3 (0%)
엽서	1,075 (19%)	926 (17%)	1,452 (24%)	964 (13%)	783 (12%)	49 (3%)
기 타	493 (8%)	345 (6%)	244 (4%)	456 (6%)	156 (3%)	34 (2%)



또한 VOC에 접수되는 민원의 유형은 크게 상담문의, 불편불만, 칭찬격려, 의견제안으로 구분할 수 있다. 최근 5년간 VOC 요구 유형별 접수현황을 살펴보면, 불편불만에 관한 민원의 비중이 과반수(50%~60%) 이상을 차지하는 가운데, 공항시설과 서비스 이용과 관련한 상담문의 건수도 계속해서 증가하는 추세를 보이고 있다.

특히 '20년 상반기에는 코로나 19 비상상황에 따라 검역, 세관, 출입국 심사와 관련한 상담문의가 급증하였음을 확인할 수 있다.

<표 3-4> 고객의 소리(VOC) 요구 유형별 현황 [단위 : 건]

상담문의	불편불만	칭찬격려	의견제안	구분	요구유형별 접수추이
1,107 (62%)	470 (26%)	119 (7%)	90 (5%)	'20년 상반기	
2,178 (33%)	3,242 (50%)	748 (11%)	421 (6%)	'19년	
1,128 (15%)	4,838 (66%)	676 (9%)	738 (10%)	'18년	
736 (12%)	3,737 (63%)	1,010 (17%)	491 (8%)	'17년	
597 (11%)	3,448 (64%)	755 (14%)	626 (12%)	'16년	
970 (16%)	2,950 (51%)	1,024 (18%)	819 (14%)	'15년	

다음으로는 VOC 접수된 불편불만과 관련한 내용을 서비스 유형별로 구분해 보면, 매년 교통/주차분야의 불편불만에 대한 접수건수가 가장 큰 비중을 차지하는 것을 알 수 있다. 특히 18년에는 제2여객터미널 오픈에 따른 여객터미널 분리운영으로 목적터미널을 사전에 확인하지 않아 대중교통(리무진,공항철도) 이용시 많은 불편사항이 발생한 것으로 판단된다. 다음으로는 상업시설, 탑승수속, 보안검색 순으로 불편불만 접수건수가 많으며, 보안검색 서비스 분야의 민원제기건수가 지속적으로 증가하고 있음을 알수 있다.

<표 3-5> 고객의 소리(VOC) 서비스 유형별 현황

[단위 : 건]

구분	'15년		'16년		'17년		'18년		'19년		'20년 상반기	
교통주차	665	(23%)	820	(24%)	940	(25%)	1367	(28%)	807	(25%)	93	(20%)
상업시설	553	(19%)	471	(14%)	590	(16%)	665	(14%)	503	(16%)	85	(18%)
탑승수속	326	(11%)	403	(12%)	533	14%	558	(11%)	282	(8%)	28	(6%)
보안검색	269	(9%)	286	(8%)	401	(11%)	471	(10%)	407	(13%)	83	(18%)
공항시설	220	(7%)	439	(13%)	253	(7%)	482	(10%)	247	(8%)	41	(9%)
편의시설	192	(6%)	233	(7%)	104	(3%)	232	(5%)	183	(6%)	27	(6%)
안내 서비스	163	(5%)	162	(5%)	202	(5%)	352	(7%)	218	(7%)	34	(7%)
CIQ	243	(8%)	314	(9%)	303	(8%)	396	(8%)	333	(10%)	53	(11%)
기타	319	(11%)	320	(9%)	411	(11%)	315	(6%)	262	(8%)	26	(5%)
합계	2,950	(100%)	3,448	(100%)	3,737	100%	4,838	(100%)	3,242	(100%)	470	(100%)

나. 인천공항 고객만족도 조사 현황

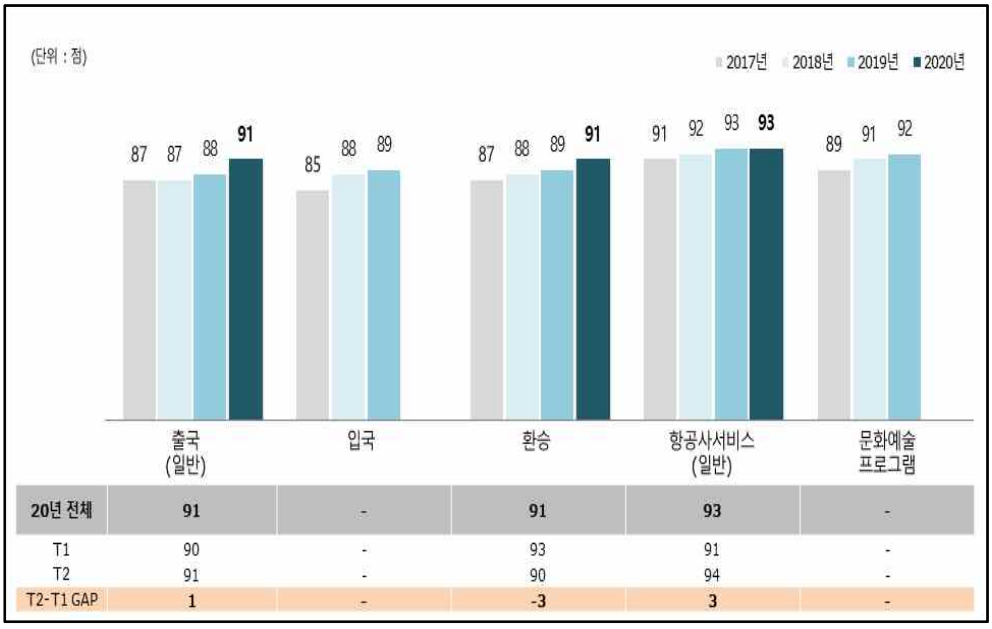
인천공항공사에서는 자체적으로 매년 분기별로 4차에 걸쳐 공항이용객을 대상으로 고객만족도 조사를 시행해 오고 있다. 인천공항에서 제공하는 주요서비스(출국, 입국, 환승, 항공사서비스, 문화예술프로그램)의 품질 수준을 진단하고, 불만족 요인을 분석하여 지속적인 서비스 개선 추진을 위한 기초 자료를 마련하고자 하는 것이 조사의 주된 목적이다. 고객만족도 조사는 전체 표본수 20,357명을 대상으로 1;1 면접조사를 통해 공항서비스 유형별 고객체감에 대한 전반적인 만족도를 조사한다. 최근 3년간 고객만족도 조사결과 전체만족도는 항공사 서비스 분야가 가장 높은 평가를 받았으며, 전반적으로 매년 점수가 소폭 상승하는 경향을 보이고 있다.

<그림3-2> 인천공항 고객만족도 조사 개요

	출국	입국	환승	항공사	문화예술 프로그램
조사 대상	출국서비스 이용고객	입국서비스 이용고객	환승서비스 이용고객	항공사 체크인 서비스 이용고객	문화 및 예술 프로그램 이용고객
조사 기간	1분기 : 3월 28일(목) ~ 4월 03일(수) 2분기 : 6월 24일(월) ~ 6월 30일(일) 3분기 : 08월 16일(금) ~ 08월 22일(목) 4분기 : 10월 14일(월) ~ 10월 20일(일)				
조사 방법	구조화된 설문지에 의한 1:1 개별면접조사(※ Tablet PC에 설문 문항을 탑재하여 진행)				
표본추출 방법	어권별/시간대별/장소별 임의할당			항공사별 임의할당	프로그램별/장소별 임의할당
조사 장소	출국게이트	입국장	탑승터미널, 환승라운지	항공사 체크인 카운터 /셀프체크인 KIOSK	여객터미널 1층, 3층 탑승게이트, 탑승동4층 한국문화전시관/박물관
진행 표본 (단위·명)	4,966 (FT: 220) - 내국인: 2,934 외국인: 2,032 - T1: 3,476, T2: 1,490	4,914 - 내국인: 2,923 외국인: 1,991 - T1: 3,425, T2: 1,489	2,653(외국인) - T1: 1,051, T2: 1,602	6,164 - 내국인: 3,506 외국인: 2,658 - T1: 5,014, T2: 1,150	1,660 - 내국인: 855 외국인: 803 - T1: 1,005, T2: 655
분기별 표본 (단위·명)	▪ 1분기 : 1,264 ▪ 2분기 : 1,247 ▪ 3분기 : 1,233 ▪ 4분기 : 1,222	▪ 1분기 : 1,232 ▪ 2분기 : 1,230 ▪ 3분기 : 1,221 ▪ 4분기 : 1,231	▪ 1분기 : 665 ▪ 2분기 : 662 ▪ 3분기 : 665 ▪ 4분기 : 661	▪ 1분기 : 1,545 ▪ 2분기 : 1,541 ▪ 3분기 : 1,559 ▪ 4분기 : 1,519	▪ 1분기 : 421 ▪ 2분기 : 403 ▪ 3분기 : 415 ▪ 4분기 : 421

※ 인천국제공항공사 내부자료 (고객만족도 조사결과보고서, 2019)

<그림3-3> 최근 3년간 인천공항 고객만족도 조사결과(전체만족도)



<표 3-6> 공항서비스 만족도의 평가지표

구분	만족도 평가지표
접근 교통 및 주차서비스 만족도	공항 접근 교통수단의 편리함, 장애인 주정차 편의, 주차대행서비스 만족도, 휠체어 리프트 서비스, 전동차 PICK UP 서비스 등
쇼핑, 식당 등 상업시설 만족도	매장 청결, 직원 청결, 음식 맛, 다양한 품목 등
보안검색 서비스 만족도	보안검색 대기시간, 보안검색 요원의 친절성, 보안검색의 정확성/신속성 등
편의시설 서비스	위치 편리성, 청결, 직원 친절도, 내부 쾌적성 등
체크인 서비스 만족도	체크인 대기시간, 체크인 업무의 신속성, 체크인 직원의 친절성 등
안내 서비스	직원친절도, 원하는 정보 제공 등
출국심사 서비스 만족도	교통약자 출국 우대 서비스, 출국심사 대기시간, 출국심사 정확성, 출국심사 신속성, 출국심사 직원 친절성 등

3. 분석방법

가. VOC 빅데이터 분석

본 연구에서는 고객 민원자료에 대한 빅데이터 분석을 위해 텍스톰(Textom)의 빅데이터 분석 솔루션인 텍스트 마이닝 기술을 이용하여 비정형 텍스트들에서 각 데이터에 포함된 특징(featue)을 추출하고 이들의 패턴이나 연관성을 조사하고자 한다. 텍스트 마이닝 기술의 구체적인 방법과 순서는 다음의 <그림 3-5>와 같다.

<그림 3-4> 고객의소리(VOC) 분석 연구방법과 절차

자료 : 저자 작성

(1) 텍스트 빈도 분석

텍스트 마이닝은 키워드(keyword) 빈도분석으로부터 시작되며, 키워드 빈도분석은 각 문서에 등장한 키워드의 출현 회수를 세어서 주요 키워드를 도출하는 방법이다. 이는 문서에 등장한 키워드의 빈도수(TF: term frequency)를 중요도로 활용하여 행렬의 형태로 나타낸 것으로서 문서에 빈번하게 나타나는 단어가 해당문서에서 중요하다는 것을 의미하며 아래 <표 3-5>은 단어-문서 행렬의 예를 나타낸다.

<표 3-7> 단어-문서 행렬 예시

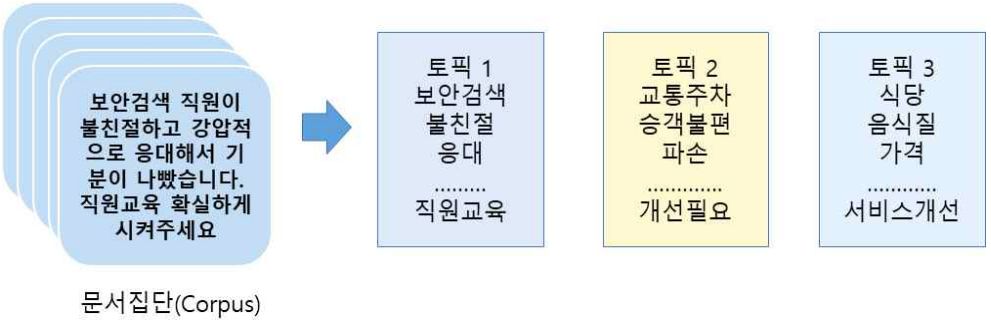
구분	교통	버스	직원	말	불친절	태도
문서 1	5	3	2	1	4	2
문서 2	3	4	4	5	2	3
문서 3	2	1	5	4	5	4

자료 : 저자작성

(2) 토픽분석(LDA 토픽 모델링)

토픽 모델링 방법 중 현재 가장 널리 사용되고 있는 LDA 모델을 이용한 토픽 클러스터링 분석을 활용하여 문서에 포함된 주요 토픽을 도출하고자 한다. VOC(고객의 소리) 데이터에 대한 토픽분석을 통해서는 아래 <그림 3-4>와 같이 공항서비스 고객 불만족 관련 주요 토픽을 도출할 수 있을 것이다.

<그림 3-5> LDA 토픽 모델링 예시



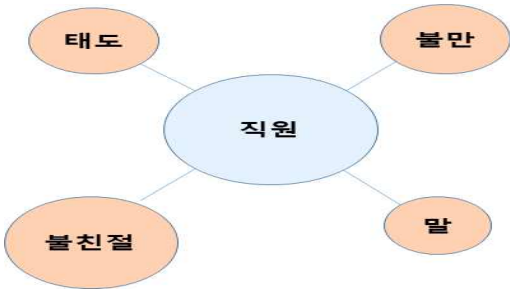
자료 : 저자 작성

(3) 키워드 네트워크

키워드 네트워크를 위해서는 노드와 엣지를 정의해야 하며, 각 LDA로부터 나오는 키워드들을 노드로서 활용해야 하며, 이에 본 연구는 분석 키워드와 연관성이 거의 없다고 판단한 키워드를 직접 제거한 후 최종 네트워크를 구성하고자 한다.

아래 <그림 2-5>는 '공항직원'과 연관된 키워드 네트워크 구성방법의 예시를 나타낸 것이다. '불만', '불친절', '태도' 등은 LDA 분석 결과로 도출된 키워드이고, 각 노드의 크기는 노드로 들어오는 엣지 갯수의 합이다.

<그림 3-6> 키워드 네트워크 구성 예시



자료 : 저자 작성

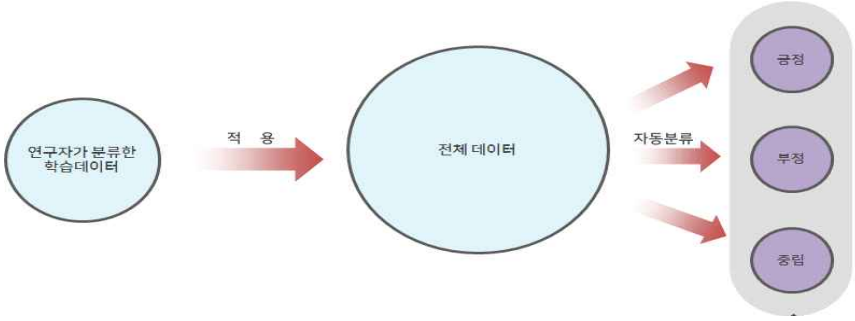
(4) 감성분석

감성분석(sentiment analysis)은 텍스트에 포함된 주관적인 감성을 추출해 내는 분석방법론으로서 연구자가 직접 학습데이터를 구성하여 적용함으로써 분석 주제의 제한없이 모든 분야의 데이터에 대한 감성분석이 가능하다. 따라서 본 연구에서는 주요 토픽인 7가지 서비스 유형별로 긍정의 정도를 수치화하기 위해 긍정의 감성지수를 산출하고자 한다. 7가지 서비스유형과 서비스 긍정지수의 산출방법은 아래와 같다.

* 7가지 서비스 유형 : ① 교통/주차, ② 탑승수속, ③ 보안검색, ④ 공항시설, ⑤ 상업시설 , ⑥ 안내서비스, ⑦ 편의시설
* 서비스 긍정지수 산출방법

$$\text{긍정지수} = \frac{\text{긍정건수}}{\text{긍정건수} + \text{부정건수}} \times 100(\text{단위}\%)$$

<그림 3-7> 감성분석 예시



나. 고객만족도 조사결과 분석

최근 5년간 공항 서비스 고객만족도 조사결과를 활용하여 주요 서비 유형별 7가지(교통주차, 보안검색, 출입국심사, 항공사 체크인, 공항 편의시설,상업시설,안내서비스)에 대한 연도별 고객 만족도 점수를 산출하여 분석하고자 한다. 설문조사결과에 대한 통계적 분석은 SPSS 통계 프로그램을 사용하고자 하며, 7가지 서비스 유형별 만족도 점수는 세부 평가 항목별/연도별 점수의 평균점수를 적용하고자 한다. 또한 고객만족도 조사결과를 활용하여 공항 서비스 종합만족도에 미치는 영향요인을 실증분석하기 위해 다중회귀분석을 시행하고자 한다.

다. 상호 비교분석 시행 (VOC 분석결과, 고객만족도 평가결과)

고객의 소리(VOC) 민원내용에 대한 극성정도 분석결과(서비스 긍정지수)와 공항서비스 고객만족도 조사 결과를 상호 비교 검증하기 위해 공항 서비스 분야를 다음과 같이 7개 분야 ① 교통/주차. ② 탑승수속, ③보안검색, ④ 공항시설, ⑤ 상업시설, ⑥ 안내서비스, ⑦ 편의시설로 구분하여 두 데이터 사이의 공통점과 차이점을 분석하고자 한다. 이를 통해 공통적으로 일치하는 서비스 품질 영향요인을 재확인하고, 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출해보고자 한다.

제 4 장 연구 결과

제 1 절 데이터 분석결과

1. 데이터 전처리 과정

비정형 텍스트에 대한 데이터 분석에서 가장 중요하며, 먼저 진행되는 것이 데이터 전처리 과정이다. 데이터 수집은 2015년 7월부터 2020년 9월까지 인천국제공항 고객의소리(VOC) 시스템에 접수된 문서 30,214건을 수집하였고, 데이터 전처리 과정중 정제과정으로서 고객의 소리 데이터 중 제목, 등록일, 유형분류 등의 정보는 제외하고, 고객의 소리 포스팅 본문내용만을 분석대상으로 하였다. 그리고 등록일 정보는 토픽분석시 시간의 흐름에 따른 토픽의 변화를 살펴볼 때 추가로 사용하였다.
다음은 형태소분석 설정 과정으로서 고객의 소리 포스팅 내용 중 영문, 중문, 일본어 등 한글 이외의 문자 및 숫자, 특수문자를 분석대상에서 제외하였다. 그리고 Open Korean Text(Okt) 한글 형태소 분석기를 사용하여 명사를 추출하여 명사만을 텍스트 분석에 사용하였다.
추출된 명사 중 출현빈도는 높으나 텍스트의 의미를 파악하는 데에는 기여도가 현저히 낮은 단어들은 불용어로 지정하여 분석대상에서 제외하였으며, 고객의 소리(VOC) 각 포스팅이 하나의 문서(document)이며, 각 문서는 불용어 목록에 포함되지 않은 한글 명사 단어들로 이루어져있다.

* 주요 불용어 목록 : 가지, 개, 건, 경우, 공항, 관련, 국제공항, 다시, 달라, 당시, 몇, 바람, 번, 부분, 분, 분경, 사항, 상황, 시, 시분, 안녕하십니까, 얘기, 여기, 오늘, 원, 월일, 이건, 인천, 인천공항, 자기, 저, 저희, 정도, 첨부, 첨부파일, 측, 편, 회

2. 텍스트 빈도 분석

가. 고객의 소리 전체문서('15년 하반기~'20년 상반기)

인천공항 VOC(고객의 소리) 전체 30,214건의 문서에 대한 텍스트 빈도 분석 결과는 다음의 <표 4-1>와 같다.

<표 4-1> 상위 50개 빈출 키워드와 빈도('15년 하반기~'20년 상반기)

키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도
직원	40815	버스	9418	주차장	6057	출발	4917	연락	3849
이용	16801	승객	9182	탑승	6046	문의	4762	업체	3765
답변	15098	주차	8699	문자	5916	문제	4727	수하물	3756
시간	13078	고객	8355	터미널	5565	요청	4492	이해	3553
생각	11981	도착	7470	보안	5398	기분	4489	항공	3538
사람	10530	전화	6840	개선	5396	기사	4337	외국인	3518
민원	10129	차량	6621	비행기	5232	한국	4326	처리	3513
안내	9960	항공사	6456	서비스	5156	태도	4245	근무	3368
출국	9833	입국	6346	접수	4939	여행	4244	대행	3265
확인	9814	검색	6110	택시	4936	카운터	4211	사용	3151

키워드 빈도에 대한 분석결과, '직원', '사람'이라는 키워드의 출현빈도가 높은 것으로 보아 인적서비스와 관련된 문서가 많은 것을 확인할 수 있으며, '버스', '주차', '주차장', '차량', '택시', '대행'의 키워드의 출현빈도가 높은 것으로 보아 접근교통 및 주차서비스, 그리고 '검색', '보안', '수하물' 키워드의 출현빈도가 높아 보안검색 서비스와 관련된 문서가 많았음을 확인할 수 있었다.

텍스트 마이닝 중 단어 빈도분석을 통해 문서에 가장 많이 출현한 상위 50개 키워드를 워드 클라우드로 시각화한 결과는 다음의 <그림 4-1>과 같다.

<그림 4-1> 상위 50개 빈출 키워드의 워드클라우드

나. 고객의 소리 '20년 상반기 문서

'20년 단년도로 한정해서 살펴보더라도 상위 50개 빈출 키워드중 '직원', '사람'이라는 키워드의 빈도가 높은 것을 확인할 수 있으며, '버스', '주차장', '보안', 검색' 키워드의 출현 빈도가 높은 것으로 보아 전체문서에 대한 빈도분석결과와 동일하게 교통 및 주차 서비스, 그리고 보안검색 서비스의 관심도가 높은 것을 확인할 수 있다. 그리고 전년도 문서에는 출현빈도가 낮거나 나타나지 않았던 '코로나', '마스크', '격리', '검사' '문의', '확인', '이동'과 같은 키워드의 출현빈도가 높은 것으로 보아 코로나 바이러스(COVID-19)의 비상상황속에서 검역(격리지침 및 규정), 환승(가능여부 및 절차), 출입국 심사와 관련한 문의사항이 증가한 것을 확인할 수 있었다.
'20년 문서에 대한 상위 50개 키워드를 워드 클라우드로 시각화한 결과는 다음의 <그림 4-2>와 같다.

<표 4-2> 상위 50개 빈출 키워드와 빈도(2020년 문서)

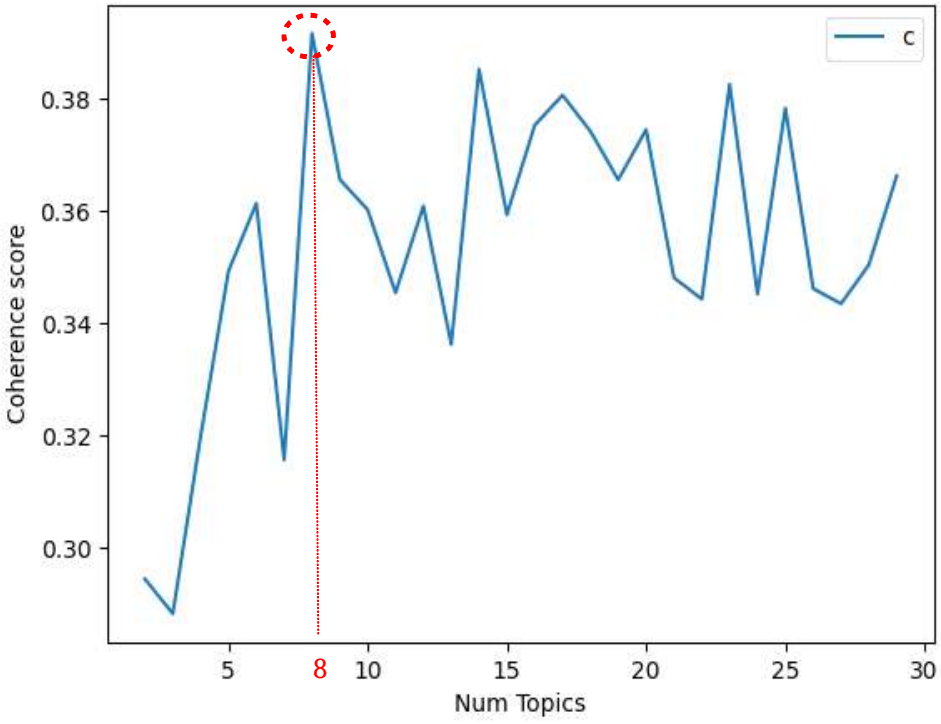
키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도	키워드	빈도
직원	1785	마스크	501	검색	344	보안	279	항공	231
이용	943	한국	492	주차장	336	사용	278	차량	224
입국	758	비행기	410	격리	333	예정	276	조치	212
시간	684	버스	404	승객	314	서비스	264	민원	208
사람	651	안내	401	근무	311	시설	254	여행	206
확인	598	터미널	400	주차	311	검사	244	전화	205
코로나	593	답변	380	현재	306	연락	242	공사	200
출국	576	탑승	369	고객	306	출발	242	이동	200
생각	550	도착	361	문제	292	기분	237	처리	195
문의	514	항공사	352	상주	281	개선	233	업무	193

<그림 4-2> 상위 50개 빈출 키워드의 워드클라우드

3. 토픽모델링(LDA) 분석

가. 최적 토픽수 결정

인천공항 VOC(고객의 소리) 전체문서에 대한 잠재된 토픽을 추출하기 위해 정합도(coherence score)를 이용하여 적정 토픽수를 결정하였다. 토픽이 2개인 경우부터 1개씩 수를 늘려 가장 높은 정합도를 보이는 토픽 개수를 적정 토픽수로 판단하였다.
정합도 분석결과 아래 <그림 4-2>와 같이 토픽의 개수가 8개일때 가장 높은 값을 보여, 토픽수를 8개로 지정해 LDA 기반으로 토픽모델링 분석을 수행하였다.



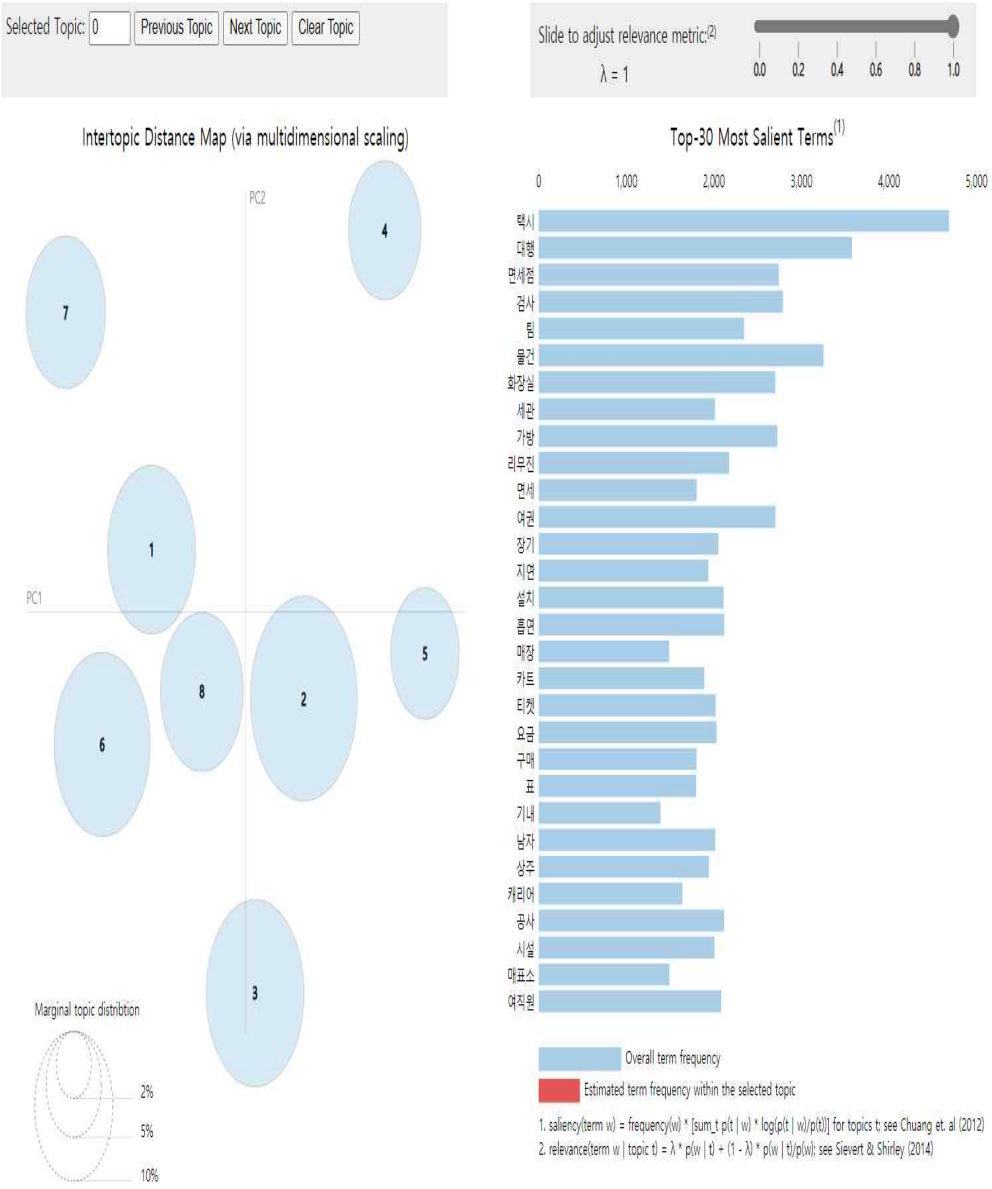
<그림 4-3> 토픽수 산출 관련 정합도 변화 그래프

나. 토픽별 키워드 분석

토픽 분석결과에 따라 최적 토픽 8개에 대한 동시출현빈도가 높은 단어들의 목록을 도출하였다. 토픽분포도에서 토픽을 직접 클릭하거나, ‘Selected Topic’에 토픽을 직접 입력하면 각각의 토픽을 구성하는 상위 30개의 단어를 직접 확인할 수 있다. 토픽분포도에서 토픽 원의 크기가 클수록 높은 빈도수의 단어들로 구성되어 있음을 알수 있으며, 토픽원간의 거리가 멀수록 판별 타당도가 높고 주제가 또렷하게 구분됨을 알 수 있다. 만약 토픽간의 거리가 가깝거나 겹쳐져 있다면 판별 타당도가 낮아 비슷한 주제를 나타내게 된다.

아래 <그림 4-4>와 같이 본 연구내용에 대한 토픽모델링 결과 8개의 토픽간에 서로 중첩되는 부분이 없이 적정한 거리를 두고 있어 판별 타당도가 높고 주제가 또렷하게 구분되어있음을 확인할 수 있다.

전체토픽에 대한 분석결과 ‘택시’, ‘대행’, ‘면세점’, ‘검사’, ‘물건’, ‘화장실’, ‘세관’, ‘리무진’, ‘지연’, ‘흡연’, ‘카트’, ‘캐리어’, ‘매표소’ 등의 키워드가 공통적으로 많이 등장하였음을 확인할 수 있다. ‘택시’, ‘대행’, ‘리무진’, ‘매표소’ 등의 키워드가 자주 등장한 것으로 보아 공항서비스 분야중 “교통 및 주차서비스의 토픽과 관련된 단어가 자주 등장하였고, 또한 ‘화장실’‘흡연’, ‘카트’, ‘캐리어’ 등의 키워드가 자주 등장한 것으로 보아 공항편의시설 서비스의 토픽과 관련된 단어가 자주 등장하였음을 알 수 있다. 1번 토픽부터 8번 토픽까지의 토픽별 키워드를 바탕으로 각각의 토픽을 유추해 보았으며, 토픽별 키워드 분석결과는 다음과 같다.

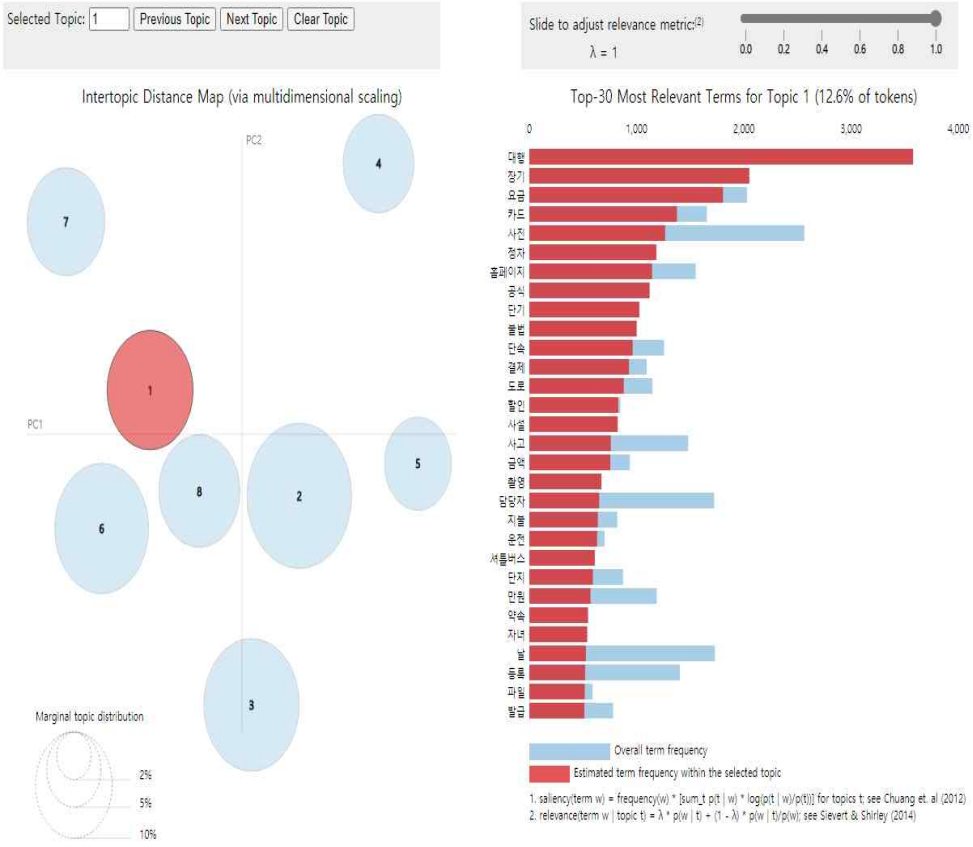


자료 : 저자 작성.

<그림 4-4> 토픽모델링(LDA) 시각화 결과(전체)

(1) 1번 토픽 : 주차 편의 서비스

1번 토픽은 아래 <그림4-5>에서와 같이 전체 토픽 중 약 12.6%의 비중을 차지하고 있다. 1번 토픽에서는 ‘대행’, ‘장기’, ‘요금’, ‘카드’, ‘정차’, ‘공식’, ‘불법’, ‘단속’, ‘할인’, ‘사설’ ‘만원’, ‘자녀’, ‘등록’ 등의 키워드가 나타났고, 주로 불법주차대행, 차량파손 보상, 주차요금 불만, 다자녀 주차요금 감면절차 등과 관련한 문서가 대부분으로 1번 토픽의 주제를 ‘주차편의 서비스’로 유추하였다.



자료 : 저자 작성.

<그림 4-5> 토픽모델링(LDA) 1번 토픽(주차서비스)

(2) 2번 토픽 : 인적서비스(단순불만)

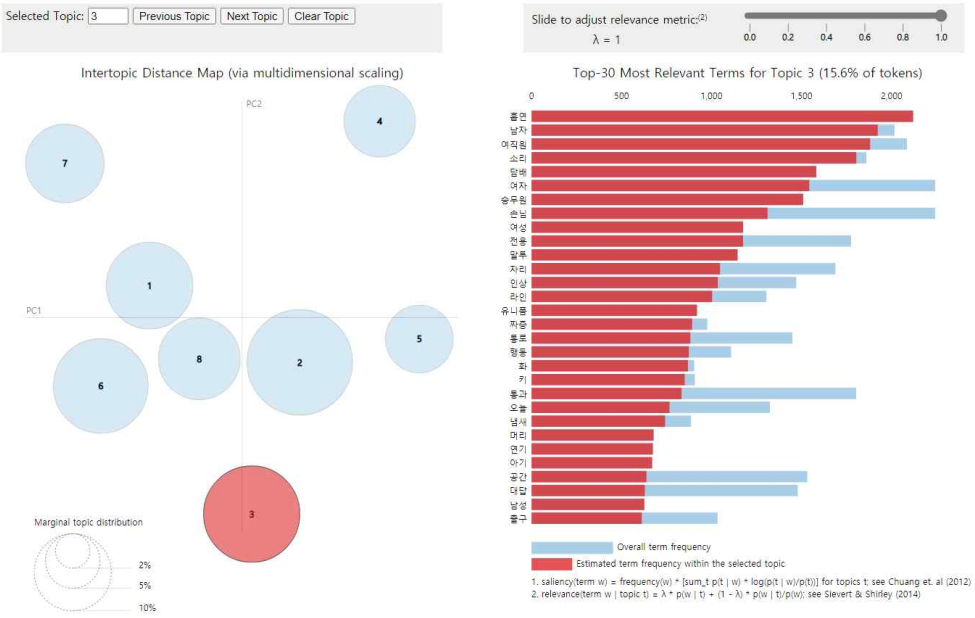
2번 토픽은 아래 <그림4-6>에서와 같이 전체 토픽 중 약 18.7%로 가장 큰비중을 차지하고 있다. 2번 토픽에서는 '여권', '핸드폰', '분실', '유실물', '칭찬', '심사', '대한항공', '휴대폰' 등의 키워드가 나타났고, 주로 고객분실물 처리, 입출국과정에서 공항직원의 응대서비스 등과 관련한 인적 서비스와 관련한 문서가 대부분으로 2번 토픽의 주제를 '인적서비스(단순불만)'으로로 유추하였다.

자료 : 저자 작성.

<그림 4-6> 토픽모델링(LDA) 2번 토픽(인적서비스)

(3) 3번 토픽 : 인적서비스(강한불만)

3번 토픽은 아래 <그림4-7>에서와 같이 전체 토픽 중 약 15.6%로 두 번째로 큰 비중을 차지하고 있다. 3번 토픽에서는 ‘흡연’, ‘소리’, ‘담배’, ‘말투’, ‘인상’, ‘유니폼’, ‘짜증’, ‘행동’, ‘화’, ‘키’, ‘냄새’, ‘머리’, ‘연기’ 등의 키워드가 나타났고, 주로 흡연구역의 담배냄새로 인한 불편사항, 직원의 불친절 응대 등 공항시설과 인적 서비스에 대한 강한 불만사항을 내포하고 있다. 2번 토픽과의 차이점은 2번 토픽은 인적서비스에 대한 개선 또는 교육 시행 등을 요구하는 내용이 많았으나, 3번 토픽은 해당직원의 인적사항(유니폼, 키, 인상)을 기입해 직접 사과를 요구하거나, 공항시설의 즉각적인 시정조치 등 강한 불만사항을 내포하고 있어, 3번 토픽의 주제를 ‘인적서비스(강한불만)’으로로 유추하였다.

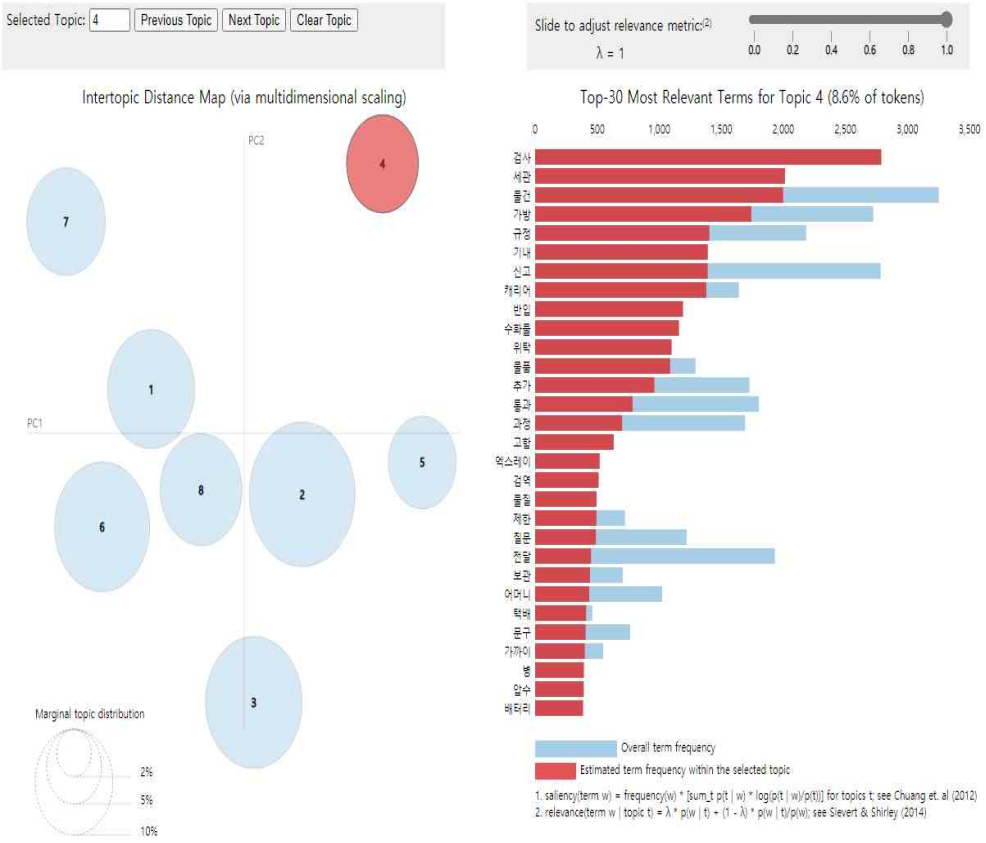


자료 : 저자 작성.

<그림 4-7> 토픽모델링(LDA) 3번 토픽(인적서비스)

(4) 4번토픽 : 상담/문의 응대서비스

4번 토픽은 아래 <그림4-8>에서와 같이 전체 토픽 중 약 8.6%의 비중을 차지하고 있다. 4번 토픽에서는 ‘검사’, ‘세관’, ‘가방’, ‘규정’, ‘기내’, ‘신고’, ‘캐리어’, ‘반입’, ‘통과’, ‘엑스레이’, ‘검역’, ‘제한’, ‘병’, ‘압수’, ‘베터리’ 등의 키워드가 나타났고, 주로 세관통관 규정, 기내위험물 및 액체류 반입 등 공항 이용을 위한 상담·문의와 관련한 문서가 대부분으로 3번 토픽의 주제를 ‘상담/문의 서비스’로 유추하였다.

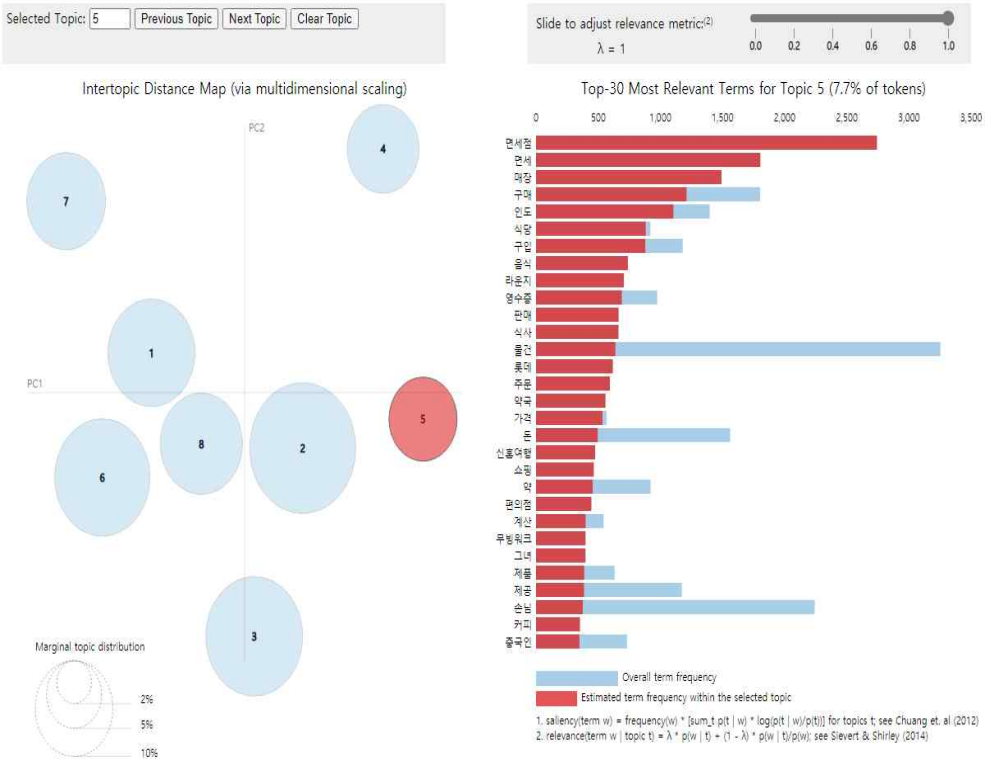


자료 : 저자 작성.

<그림 4-8> 토픽모델링(LDA) 4번 토픽(상담/문의 응대서비스)

(5) 5번토픽 : 상업시설 서비스

5번 토픽은 아래 <그림4-9>에서와 같이 전체 토픽 중 약 7.7%의 비중을 차지하고 있다. 5번 토픽에서는 '면세점', '매장', '식당', '음식', '영수증', '식사', '약국', '가격', '쇼핑', '약', '편의점', '제한' 등의 키워드가 나타났고, 주로 식당 위생상태, 음식 품질, 상업시설(면세점,식당,약국 등) 가격 불만 등 상업시설(쇼핑/식당/은행) 이용시 불편·불만사항과 관련한 문서가 대부분으로 5번 토픽의 주제를 '상업시설 서비스'로 유추하였다.

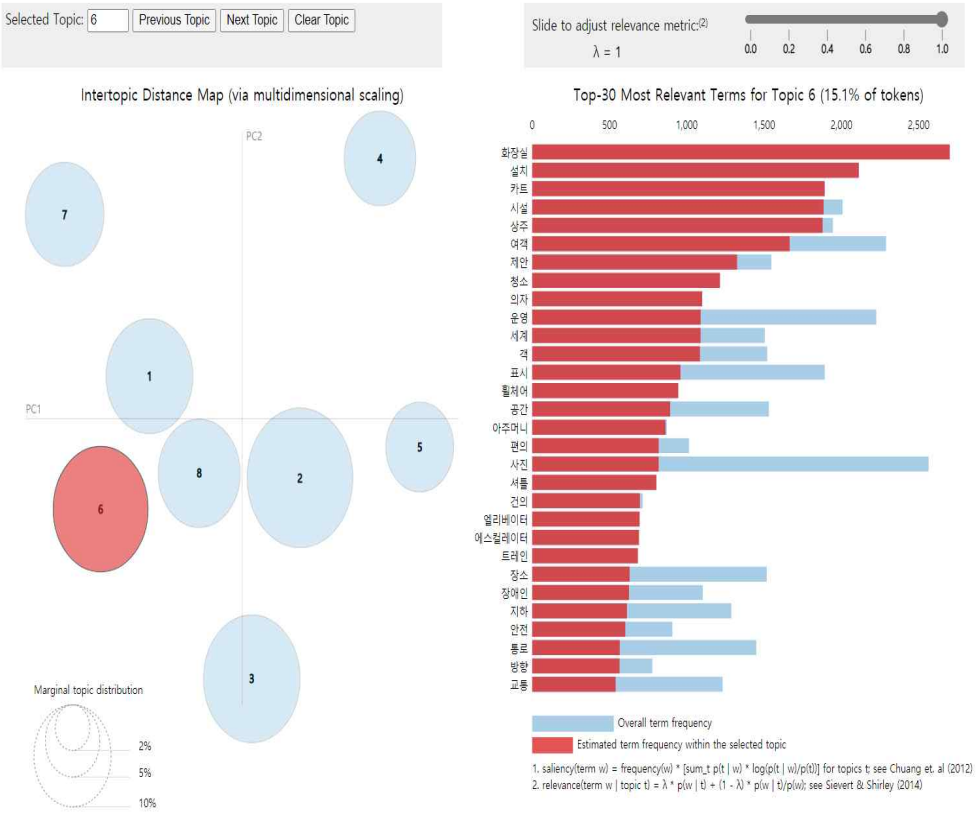


자료 : 저자 작성.

<그림 4-9> 토픽모델링(LDA) 5번 토픽(상업시설 서비스)

(6) 6번토픽 : 공항시설(편의시설/환경)

6번 토픽은 아래 <그림4-10>에서와 같이 전체 토픽 중 약 15.1%의 비중을 차지하고 있다. 6번 토픽에서는 '화장실', '카트', '청소', '의자', '휠체어', '편의', '엘리베이터', '장애인', '안전' 등의 키워드가 나타났고, 주로 화장실 청소상태 불량, 카트이용 불편, 교통약자 이동불편 등 공항내 편의시설 이용시 불편·불만사항과 관련한 문서가 대부분으로 6번 토픽의 주제를 '공항시설(편의시설/환경)'로 유추하였다.

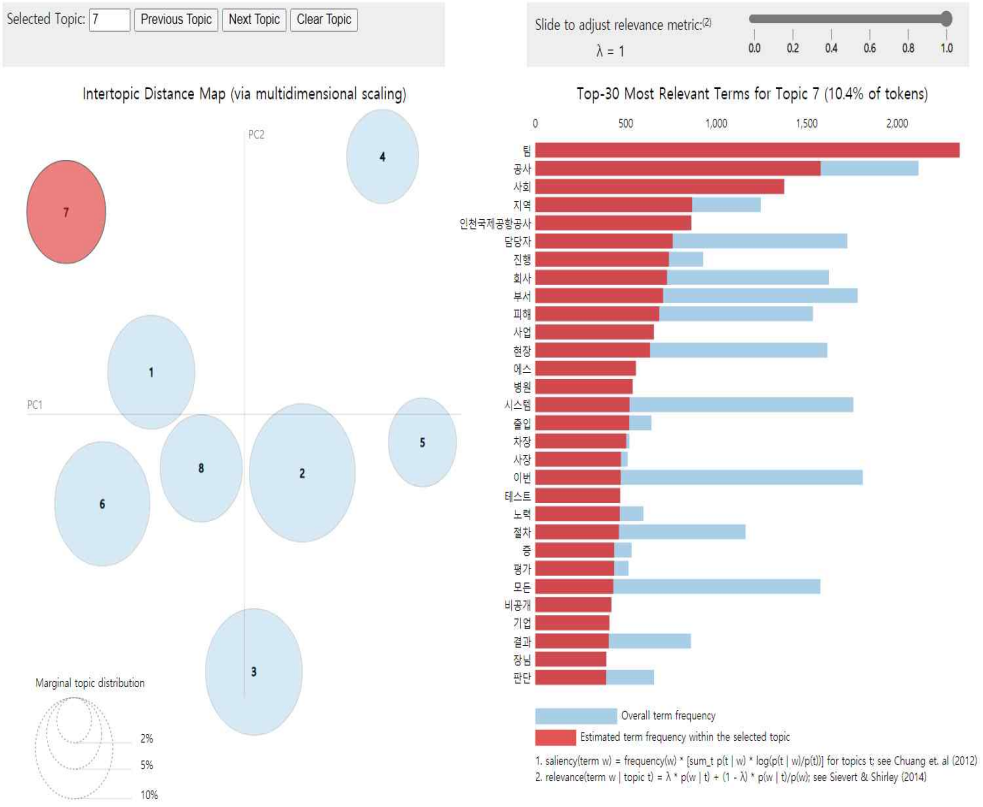


자료 : 저자 작성.

<그림 4-10> 토픽모델링(LDA) 6번 토픽(공항시설 서비스)

(7) 7번토픽 : 공항공사 개선노력

7번 토픽은 아래 <그림4-11>에서와 같이 전체 토픽 중 약 10.4%의 비중을 차지하고 있다. 7번 토픽에서는 ‘인천국제공항공사’, ‘담당자’, ‘부서’, ‘시스템’, ‘사장’, ‘노력’, ‘절차’, ‘약국’ 등의 키워드가 나타났고, 주로 공항시설이용과 관련한 제도 및 절차 개선요구 사항 뿐만 아니라, 항공기소음, 건설현장 피해 등 공항공사에 직접적으로 개선을 요구하는 문서가 대부분으로 5번 토픽의 주제를 ‘공항공사 개선노력’으로 유추하였다.

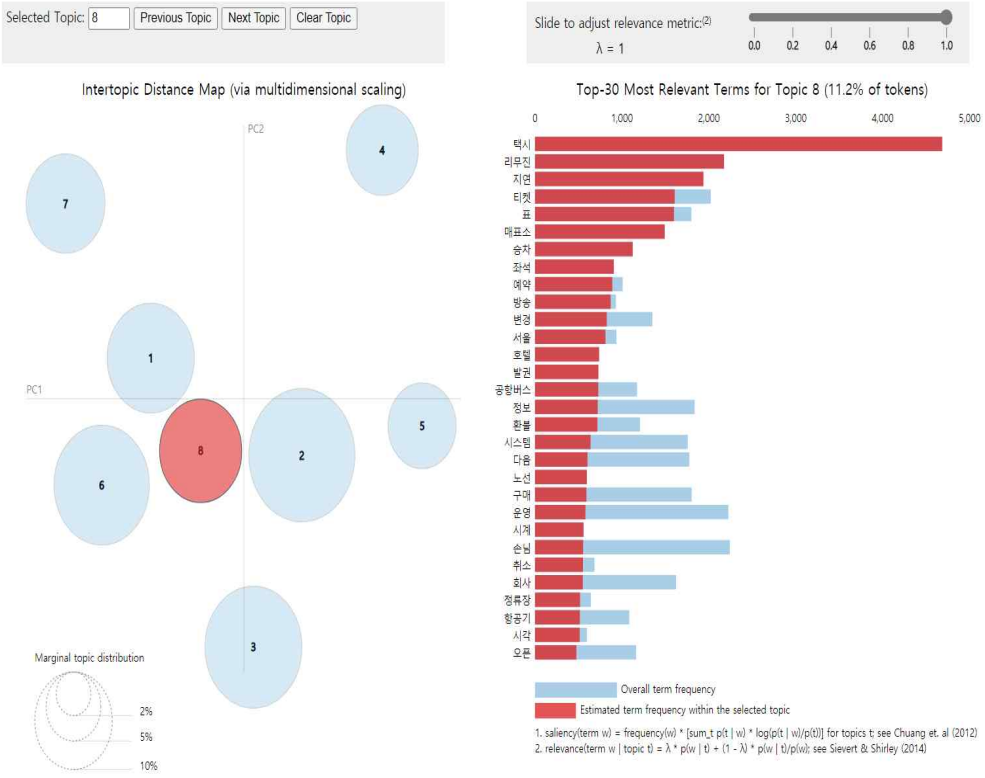


자료 : 저자 작성.

<그림 4-11> 토픽모델링(LDA) 7번 토픽(공항공사 개선노력)

(8) 8번토픽 : 접근교통 서비스

8번 토픽은 아래 <그림4-12>에서와 같이 전체 토픽 중 약 11.2%의 비중을 차지하고 있다. 8번 토픽에서는 ‘택시’, ‘리무진’, ‘티켓’, ‘매표소’, ‘승차’, ‘좌석’, ‘호텔’, ‘공항버스’, ‘환불’, ‘노선’, ‘정류장 등의 키워드가 나타났고, 주로 공항버스 지정좌석제 운영에 따른 승차거부, 교통안내 정보 미흡, 버스매표소 안내 불편 등 대중교통이용과 관련한 불편사항에 대한 문서가 대부분으로 8번 토픽의 주제를 ‘접근교통 서비스’로 유추하였다.



자료 : 저자 작성.

<그림 4-12> 토픽모델링(LDA) 8번 토픽(접근교통 서비스)

(9) 토픽분석 종합

앞에서 살펴본 인천공항 VOC(고객의 소리) 문서에서 추출한 8개 토픽과 토픽별 주요 포함단어 및 비중을 아래 <표4-3>과 같이 정리하였다.

토픽명	포함 단어	토픽비중
① 주차서비스	대행/장기/요금/카드/사진/정차/홈페이지/공식단기/불법/단속/결제/도로/할인/사설/사고/금액/촬영/담당자/지불/운전 /셔틀버스/단지/만원/약속/자녀/날/등록/파일/발급	12.6%
② 인적서비스 (단순불만)	여권/도움/핸드폰/아이/수속/분실/유실물/칭찬/우리/센터/자동/가족/심사/중국/친구/인사/가방/덕분/사무실/대한항공/휴대폰/말씀/등록/환승/대답/정신/기억/미국/출입국/이번	18.7%
③ 인적서비스 (강한불만)	흡연/남자/여직원/소리/담배/여자/승무원/손님/여성/전용/말투/자리/인상/라인/유니폼/짜증/통로/행동/화/키/통과/오늘/냄새/머리/연기/아기/공간/대답/남성/출구	15.6%
④ 상담/문의	검사/세관/물건/가방/규정/기내/신고/캐리어/반입/수화물/위탁/물품/추가/통과/과정/고함/엑스레이/검역/물질/제한/질문전달/보관/어머니/택배/문구/가까이/병/압수/배터리	8.6%
⑤ 상업시설 서비스	면세점/면세/매장/구매/인도/식당/구입/음식/라운지/영수증판매/식사/물건/롯데/주문/약국/가격/돈/신혼여행/쇼핑/약/편의점/계산/무빙워크/그녀/제품/제공/손님/커피/중국인	7.7%
⑥ 공항시설 서비스 (편의시설)	화장실/설치/카트/시설/상주/여객/제안/청소/의자/운영/세계/객/표시/휠체어/공간/아주머니/편의/사진/셔틀/건의/엘리베이터/에스컬레이터/트레인/장소/장애인/지하/안전/통로/방향/교통	15.1%
⑦ 공항공사 개선노력	팀/공사/사회/지역/인천국제공항공사/담당자/진행/회사/부서/피해/사업/현장/에스/병원/시스템/출입/차장/사장/이번/테스트/노력/절차/증/평가/모든/비공개/기업/결과/장님/판단	10.4%
⑧ 접근교통 서비스	택시/리무진/지연/티켓/표/매표소/승차/좌석/예약/방송/변경/서울/호텔/발권/공항버스/정보/환불/시스템/다음/노선/구매/운영/시계/손님/취소/회사/정류장/항공기/시각/오픈	11.2%

<표 4-3> 토픽분석 종합(토픽명, 포함단어)

다. 토픽의 시계열적 변화 경향 분석

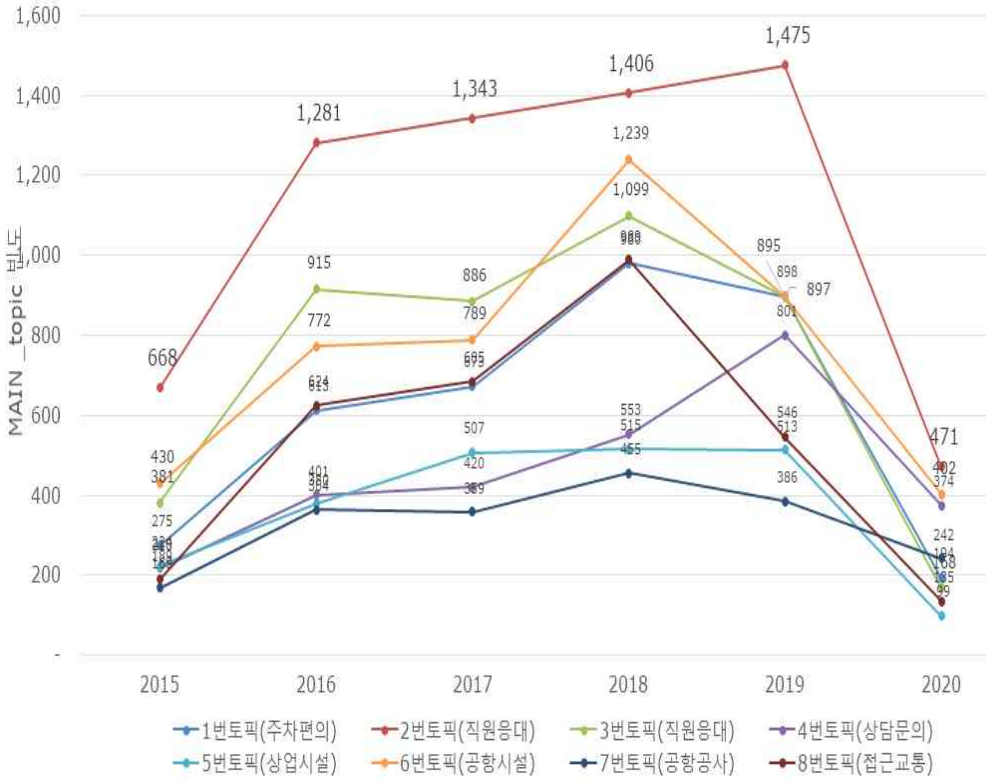
다음으로 인천공항 VOC(고객의 소리) 문서의 8개 토픽이 최근 5년간 시간의 흐름에 따라 어떻게 변화하였는지의 추이를 분석하기 위해 토픽별 비중변화를 아래 <표 4-4>와 같이 산출하였다. 고객의 소리 접수문서가 인천공항 VOC 시스템상에 최근 5년간의 자료만 보존되고 기간이 경과된 문서는 자동 삭제됨에 따라, 15년도(하반기)와 20년도(상반기) 문서는 표본수가 상대적으로 적음에 따라, '16년도에서 '19년도 까지의 문서에 대한 토픽별 비중 추이를 분석해보고자 한다.

연도	2015		2016		2017		2018		2019		2020	
토픽	빈도	비율	빈도	비율	빈도	비율	빈도	비율	빈도	비율	빈도	비율
1	275	11%	613	11%	673	12%	980	14%	898	14%	194	9%
2	668	26%	1,281	24%	1,343	24%	1,406	19%	1,475	23%	471	23%
3	381	15%	915	17%	886	16%	1,099	15%	895	14%	168	8%
4	218	9%	401	8%	420	7%	553	8%	801	12%	374	18%
5	224	9%	380	7%	507	9%	515	7%	513	8%	99	5%
6	430	17%	772	14%	789	14%	1,239	17%	897	14%	402	19%
7	169	7%	364	7%	359	6%	455	6%	386	6%	242	12%
8	189	7%	624	12%	685	12%	989	14%	546	9%	135	7%
계	2,554	100%	5,350	100%	5,662	100%	7,236	100%	6,411	100%	2,085	100%

<표 4-4> 연도별 토픽의 추이분석

아래의 <그림 4-13>에서 확인할 수 있듯이 대부분의 토픽별 비중이 '18년에 급격하게 증가한 것을 확인할 수 있다. 이는 2018년 인천공항 제2여객터미널이 그랜드 오픈함에 따라, 공항 이용객들의 불편·불만족 또는 문의사항과 관련한 문서가 급증하였음을 알 수 있다. 특히 노란색 선인 6번토픽(공항시설)이 57%이상 큰 폭으로 증가한 것을 확인할 수 있는데, 이는 제 2여객터미널 운영 초기 시스템 및 시설 안정화가 미흡하다 보니 운영상에 많은 문제점이 발생하여 관련 토픽의 빈도수가 급증하였음을 알 수 있다.

그러나 '19년도에는 공항운영이 점차 안정화되다 보니 대부분 토픽의 빈도수가 감소하였으나, 보라색 선인 4번토픽(상담/문의건)의 빈도수가 증가하였음을 알 수 있다. 이는 전년 대비하여 식품/액체류 반입, 세관신고 및 출입국 심사, 환승절차, 유실물 신고, 주차요금 문의 등이 증가하였음을 확인 할 수 있다. 또한 특이한 점은 파란색선인 1번토픽(주차편의)과 갈색선인 8번토픽(접근교통)이 연도별로 시간의 흐름에 따라 유사하게 변화하는 경향을 보이고 있음을 확인할 수 있는데, 특히 '18년도에 제2여객터미널 오픈에 따른 터미널의 분리운영으로 공항 이용객의 목적지 터미널 혼선에 따른 교통/주차분야의 문서 접수건수가 급증하였음을 알 수 있다. 그러나 '19년도에 접근교통 관련 토픽의 비중이 급감하였는데 이는 공항버스 노선 증설 및 배차시간 조정 등 대중교통 불편을 대폭 개선하였고, 택시의 불법 호객행위 및 부당요금에 대한 단속 등으로 관련 민원건수가 감소하였음을 보여준다. 반면에 주차편의 서비스와 관련된 토픽의 비중은 소폭으로 감소하였음을 확인할 수 있는데, 이는 불법 주차대행 서비스로 인한 차량파손, 과속운전, 다자녀가구 및 저공해차량 등 주차요금 감면절차 복잡 등으로 인한 민원건수는 크게 감소하지 않아 주차편의 서비스에 있어서 서비스 개선노력이 필요할 것으로 생각된다.



자료 : 저자 작성.

<그림 4-13> 토픽모델링(LDA) 시계열 분석

4. 감성분석

가. 감성분류

앞서 시행한 텍스트 빈도분석 및 토픽분석은 VOC(고객의 소리)의 문서 내용이 긍정적 감성을 나타내는지 부정적 감성을 나타내는지에 대해서는 구별해주지 않아, 문서의 극성정도를 파악하기 위해 별도의 감성분석을 시행하였다. 감성분석은 고객의 소리(VOC) 원문 데이터에 대해 직접 극성을 부여한 학습데이터(1000여개)와 텍스톰이 보유

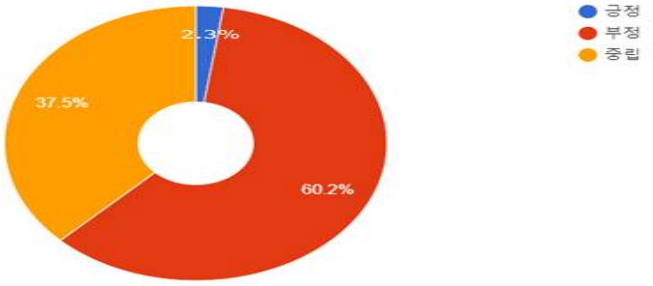
하고 있는 감성어 사전을 기반으로 하여 전체데이터에 대한 극성정도(긍정/부정/중립)를 자동으로 분류하였으며 그 결과는 아래 <그림 4-14>와 같다.

감성분석결과 고객 불편·불만 등의 부정적 감성을 나타내는 문서가 전체 문서의 60.2%, 상담·문의, 의견제안 등 중립적 감성을 나타내는 문서가 37.5%, 칭찬·격려 등의 긍정적 감성을 나타내는 문서는 2.3%로 가장 낮은 비율로 나타났다.

감성분석 결과 (문서기반) - 전체

구분	빈도(건)	비율(%)
전체	29195	100.0
긍정	678	2.32
중립	10950	37.51
부정	17567	60.17

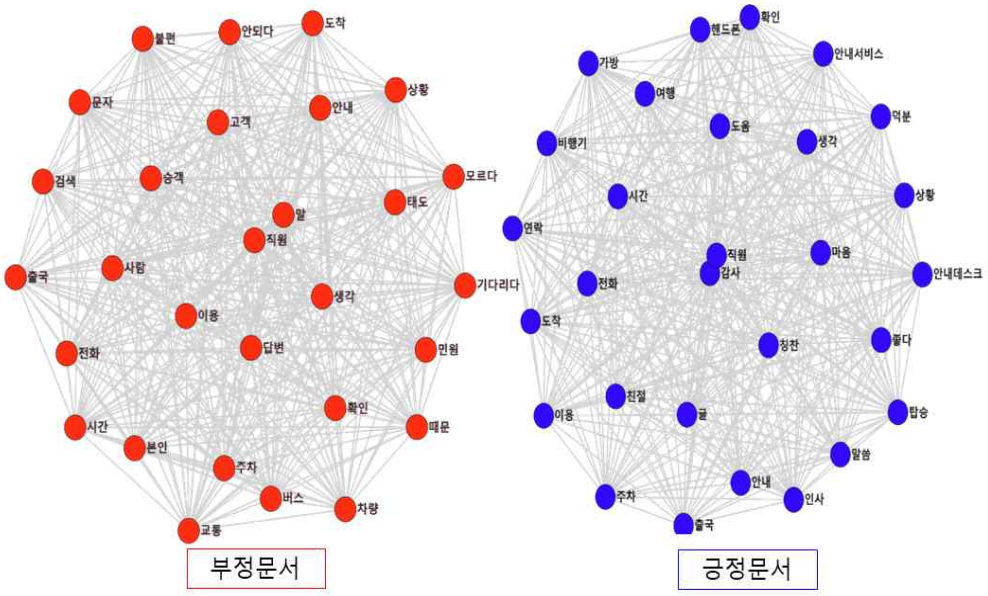
성의없는 태도 쇼핑/식당 > 면세점 > 출국장면세점 신라면세점 바비브라운 직원분들 정말 불친절을 넘어 불쾌했습니다. 오전 11시30분 오후...	부정
Bicycle Packing Service 쇼핑/식당 > 택배/수화물보관 This year on June 5, my wife and I wi...	중립
출국심사시 핵 반입 되나요?? 보안/검색 > 검색대 검색 친에어 이용 세부 갈 예정입니다저녁 시간이 걸려서. 아이 먹을 떡을 집에서 가져가서...	중립
장기주차 교통/주차 > 주차대행 > 공식 주차대행 장기주차 이용시 주차대행을 이용하곤 합니다. 사설 주차대행에 대한 워낙 안좋은 뉴스들이 많...	부정
출입국심사에 대한 안내 부적절 출입국심사 > 입국심사 > 자동 심사 안녕하세요, 지난 밤1월 1일 11시 경 대한항공 비행기 편으로 대만에서...	부정
직원이 승객을 몰래 촬영 후 개인 SNS에 상습적으로 게시 보안/검색 > 출입/촬영허가 물류 단지 쪽에서 일하는 것으로 추정 되는 92년생 ...	부정
뱀부항공 기내에 아기 식품에 대해서 문의합니다. 보안/검색 > 기내반입 불가 물품관리 검색해서 찾아보고 뱀부항공에 전화도 하고이메일도 보냈는...	중립
세부 막탄공항 반입음식 검역 > 동식물 검역 세부 막탄 공항 입국시 라면.고기볶음고추장.짜장.카레.물.우유.김치등 반입이 되는지 알고 싶습니...	중립
탑승게이트 옆 동전수집함 설치 공항공사 > 사회공헌 1. 제안배경여행을 마치고 돌아왔는데 내지갑에 그 나라의 동전이 몇개 남아있었다. 이제...	긍정



<그림 4-14> 고객의소리(VOC) 감성분석 결과

나. 감성 분석 매트릭스(상위 30개)

감성분석 결과를 바탕으로 부정적 감성을 나타내는 문서와 긍정적 감성을 나타내는 문서를 구분하여 추가로 텍스트 분석을 시행하였으며, 그 결과는 아래 <그림 4-15>와 같다. 긍정과 부정의 문서 모두에서 '직원' 이라는 키워드가 가장 많이 출현하였음을 알 수 있고, 먼저 부정적 감성의 문서에서는 '말', '태도', '생각', '교통', '주차', '버스', '검색' '불편' 등의 키워드가 상위를 차지함에 따라 접근교통 및 주차서비스와 보안검색 서비스 분야에서 공항직원의 인적서비스에 대해 공항이용객의 불편과 불만족이 많았음을 확인할 수 있다. 반면에 긍정적 문서에서는 '감사', '칭찬', '마음', '도움', '친절', '핸드폰', '가방', '안내서비스', '안내데스크' 등의 키워드가 상위를 차지함에 따라 공항안내 직원의 친절한 안내 서비스에 고객들이 감사해하며 해당직원을 칭찬하는 내용이 많았음을 확인할 수 있다.



<그림 4-15> 고객의소리(VOC) 감성분석 매트릭스(상위 30개)

제 2 절 고객만족도 조사 분석

1. 고객만족도 조사 분석결과

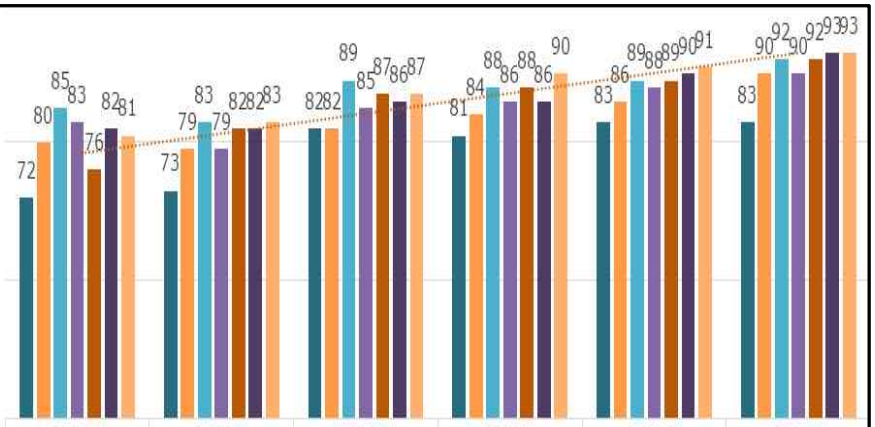
최근 5년간('15년～'20년)의 인천공항 고객만족도 조사결과를 활용하여 7개 공항 핵심 서비스 유형별 만족도 점수를 산출하였고, 연도별 변화 추이를 분석해보았다. 7개 서비스 유형별 세부 조사항목은 아래 <표 4-5>와 같다

서비스 유형		세부 평가항목
① 접근교통/ 주차		1. 공항접근 교통, 2. 주차시설 찾기 용이, 3. 직원 친절도, 4. 셔틀버스 이용 편리성, 5. 주차시설 이용편리성, 6. 이용요금 적절성,7. 주차대행 직원 친절성, 8. 주차시설 만족도, 9. 전반적 만족도
② 보안검색		1. 대기시간 적정, 2. 보안검색 정확성, 3. 보안검색 신속성, 4. 직원친절도, 5. 상황에 맞는 서비스 제공, 6. 사전안내, 7 전반적 만족도
③ 출입국 심사		1. 대기시간 적정, 2. 정확성, 3. 신속성, 4. 직원친절도, 5. 자동출입국심사 이용 편리성, 6. 공간 및 설비 충분, 7. 전반적 만족도
④ 체크인 서비스		1. 카운터 위치 안내, 2. 대기시간 적정, 3. 정확성, 4. 신속성, 5. 직원 친절도, 6. 원활한 수화물 처리, 7. 공간 및 설비 충분, 8. 전반적 만족도
⑤ 상업시설	식음료	1. 매장 청결, 2. 직원친절도, 3. 음식 맛, 4. 전반적 만족도
	쇼핑	1. 매장 청결, 2. 직원친절도, 3. 다양한 품목, 4. 전반적 만족도
	은행/환전	1. 이용 편리성, 2. 위치 및 안내 적절성, 3. 직원친절도
⑥ 안내서비스		1. 직원 친절도, 2. 외국어 이용 편리성, 3. 원하는 정보 제공, 4. 공항내 길찾기 용이, 5. 안내스크린 비치, 6. 이동수단 편리, 7. 전반적 만족도
⑦ 편의시설	항공사 라운지	1. 위치안내, 2. 매장충분/청결, 3. 직원친절도, 4. 이용편리성, 5. 내부 쾌적성
	화장실	1. 이용 편리성, 2. 청결
	무선인터넷	1. 접근 환경 우수, 2. 이용속도 적절성, 3. 충전시설 충분
	유아휴게실	1. 위치안내, 2. 이용편리성

<표 4-5> 인천공항 고객만족도 조사 세부 평가항목

서비스 유형별 평가점수는 세부 평가항목별 점수의 평균으로 산출하였고 그 결과는 아래 <그림 4-16>과 같다.

공항핵심 서비스유형별 고객만족도 점수를 살펴보면, 교통 및 주차 서비스 분야의 만족도 점수가 가장 낮고, 상업시설, 보안검색 분야 순으로 점수가 낮음을 확인할 수 있다. 반면에 출입국 심사 분야의 만족도 점수가 높고, 편의시설, 안내서비스 분야 순으로 점수가 높음을 확인할 수 있다. 또한 대체적으로 7개분야의 만족도 점수가 전년도 대비 매년 상승하는 경향을 보이고 있으며, '20년도에는 교통 및 주차서비스 분야를 제외한 나머지 분야의 고객만족도 점수가 모두 분야별 편차없이 90점이상의 점수를 기록하고 있음을 확인할 수 있다.



	2015년	2016년	2017년	2018년	2019년	2020년
교통/주차	72	73	82	81	83	83
보안검색	80	79	82	84	86	90
출입국심사	85	83	89	88	89	92
체크인서비스	83	79	85	86	88	90
상업시설	76	82	87	88	89	92
안내서비스	82	82	86	86	90	93
편의시설	81	83	87	90	91	93

<그림 4-16> 인천공항 고객만족도 조사결과(최근 5년간, 2015년~ 2020년)

2. 고객만족도 조사 다중회귀분석

인천공항 고객만족도 조사 데이터를 활용하여 고객만족도에 미치는 영향 요인이 무엇인지 실증분석하기 위해 아래의 <표 4-6>과 같이 다중회귀 분석을 시행하였다. 분석에 활용한 데이터는 2017년 1사분기부터 2020년 4분기까지의 고객만족도 조사 데이터이며, 4년치 조사데이터 약 1만개의 관측치중 결측치가 있는 관측치를 제외하여 최종 분석에 활용된 관측치는 1,110개이다. (2020년도는 코로나19 영향으로 3,4분기만 고객만족도 조사 시행) 종속변수는 인천공항 서비스의 전반적 만족도, 인천공항에 대한 좋은 이미지, 인천공항이 다른 공항에 비해 우수하다는 인식의 세가지 변수를 평균한 종합만족도이다. 독립변수는 공항의 핵심 서비스 유형인 다음의 7가지(① 접근교통 및 주차, ② 보안검색, ③ 출입국심사, ④ 체크인(탑승수속), ⑤ 상업시설, ⑥ 안내시설, ⑦ 편의시설)로서 공항의 서비스 유형별 만족도가 종합만족도에 미치는 영향을 분석하고자 하였고, 통제변수로는 성별, 연령, 내국인 여부, 거주지 변수를 포함하였다.

다중회귀 분석결과, 통제변수중에서 종속변수인 종합 만족도에 통계적으로 유의미한 영향을 미치는 변수는 없는 것으로 나타났으며, 7가지 독립변수중 종속변수인 공항서비스 종합만족도에 통계적으로 유의미한 영향을 미치는 변수는 다음의 4가지 서비스 유형인 보안검색만족도, 상업시설(쇼핑매장) 만족도, 안내시설 만족도, 편의시설 만족도로 나타났다. 공항서비스 종합 만족도에 가장 큰 영향을 미치는 변수는 편의시설 서비스이며, 다음으로 보안검색 만족도, 안내시설 만족도, 상업시설(쇼핑매장) 만족도 순으로 유의수준 5%하에서 통계적으로 유의미한 영향을 미치는 것으로 분석되었는데, 주로 직원응대와 같은 인적서비스 측면보다는 여객터미널내 시설 이용의 편리·편의성과 같은 시설서비스가 공항서비스 종합만족도에 영향을 미치는 것으로 해석할 수 있을 것이다.

또한 서비스 유형별 고객만족도 점수와 연계해 분석해보더라도, 편의시설과 안내서비스 분야는 고객만족도 점수와 종합 만족도에 미치는 영향 모두 큰 것으로 분석되었으며, 접근교통 및 주차서비스 분야의 경우에는 고

객만족도 점수는 가장 낮았지만 종합만족도 점수에는 유의미한 영향을 미치지 않는 것으로 분석되었는데 이는 접근교통 및 주차서비스에 불만이 있는 고객도 공항도착 이후 여객터미널내 각종 여객 편의시설과 안내서비스에서 좋은 인상과 만족를 느끼면 전반적인 공항서비스 만족도가 높아지는 것으로 해석할 수 있을 것이다.

<표 4-6> 인천공항 고객만족도 조사 회귀분석 결과

변수		회귀계수	p-value
상수항		0.928	0.277
① 접근교통/주차 만족도		0.038	0.290
② 탑승수속 만족도		0.002	0.957
③ 보안검색 만족도		0.139 ***	0.002
④ 출입국 심사 만족도		0.061	0.189
⑤ 안내시설 만족도		0.112 ***	0.010
⑥ 상업시설	식음료매장 만족도	0.035	0.315
	쇼핑매장 만족도	0.087 **	0.022
	은행/환전 만족도	0.020	0.452
⑦ 편의시설 만족도		0.275 ***	<.0001
남성		-0.026	0.615
10대		0.285	0.734
20대		0.266	0.747
30대		0.263	0.750
40대		0.235	0.777
50대		0.291	0.727
60대		0.468	0.582
내국인		0.038	0.729
서울		0.001	0.996
부산		-0.091	0.711
대구		0.210	0.409
인천		0.069	0.641
광주		0.155	0.495
대전		0.013	0.945
울산		-0.414	0.160
경기		0.036	0.770
강원		-0.068	0.742
충남		0.102	0.564
충북		0.173	0.392
전남		0.293	0.277
전북		0.293	0.160
경남		-0.364	0.153
경북		0.050	0.826
제주		0.030	0.917
세종		0.154	0.336

*: p<0.1, **: p<0.05, ***: p<0.01

제 3 절 상호비교 분석

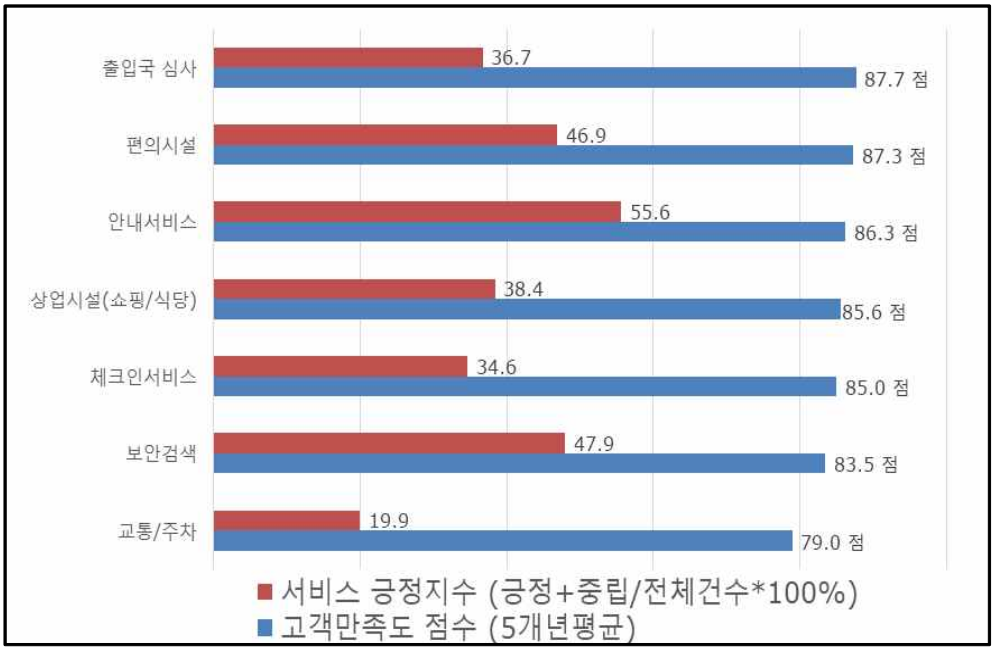
1. 서비스 유형별 만족도 비교(고객만족도 점수VS서비스 긍정지수)

공항 서비스 유형별 만족도를 분석하기 위해 인천공항 고객만족도 조사결과와 고객의 소리(VOC) 문서간의 상호비교 분석을 시행하였고 그 결과는 아래 <그림 4-17>과 같다. 상호비교 분석을 위해 고객만족도 조사는 최근 5년간 조사결과의 평균값을 적용하였고, 고객의 소리(VOC) 문서는 비정형 데이터로 존재함에 따라 서비스 유형별로 긍정의 정도를 수치화하기 위해 서비스 긍정지수를 산출하였으며, 산출방법과 7가지 서비스 유형은 아래와 같다.

* 서비스 긍정지수 산출방법

$$\text{긍정지수} = \frac{\text{긍정} + \text{중립건수}}{\text{전체 문서건수}} \times 100(\text{단위}\%)$$

* 7가지 서비스 유형 : ① 교통/주차, ② 보안검색, ③ 출입국심사, ④ 체크인서비스, ⑤ 상업시설 , ⑥ 안내서비스, ⑦ 편의시설



<그림 4-17> 인천공항 서비스 유형별 만족도 비교분석

상호비교 분석결과, 다음의 5가지 서비스 유형(① 교통/주차, ④ 체크인 서비스, ⑤ 상업시설, ⑥ 안내서비스, ⑦ 편의시설)은 고객만족도 점수와 서비스 긍정지수 간에 공통적으로 일치하는 경향을 확인할 수 있다.
① 교통/주차 분야는 고객만족도 점수와 서비스 긍정지수 모두 가장 낮은 점수를 보이고 있으며, ⑥ 안내서비스와 ⑦ 편의시설 분야는 고객만족도 점수와 서비스 긍정지수 모두 상위점수를 기록하고 있으며, 또한 ④ 체크인서비스과 ⑤ 상업시설 분야는 모두 중간정도의 점수를 기록하고 있는 것을 확인할 수 있다.
반면에 다음의 2가지 서비스 유형(② 보안검색, ③ 출입국심사)은 고객만족도 점수와 서비스 긍정지수 간에 차이점을 나타내는 것을 확인할 수 있다.
② 보안검색 분야는 고객만족도 점수는 ① 교통/주차 분야 다음으로 두 번째로 낮았으나, 서비스 긍정지수는 ⑥ 안내서비스 다음으로 두 번째로 높은 점수를 획득하였다. 이는 보안검색과 관련된 고객의 소리(VOC) 문서중 항공기내 위험 수화물, 식품, 액체류 반입과 관련한 문의사항이 많았고, 해당문서들이 중립건수로 분류됨에 따라 서비스 긍정지수는 상대적으로 높은 점수를 획득한 것으로 분석된다.
다음으로 ③ 출입국심사 분야는 고객만족도 점수는 가장 높은 점수를 획득하였으나, 서비스 긍정지수는 ① 교통/주차, ④ 체크인 서비스 다음으로 세 번째로 낮은 점수를 획득하였음을 확인할 수 있다. 이는 자동출입국 심사 도입에 따라 출입국 심사 이용 편리성 및 여권/비자 심사의 신속성과 정확성 등은 크게 개선되었으나, 출입국 심사관의 불친절한 응대 및 대기시간 지연에 따른 불편·불만사항이 많았던 것으로 분석된다. 실제 <표 4-7>의 출입국 심사분야 고객만족도 조사결과(최근5년간)와 <표 4-8>의 고객 불만족 사례를 살펴보더라도 자동시스템 이용 편리성과 설비 공간의 충분성과 관련해서는 높은 점수를 획득한 반면 직원의 친절도에 대한 평가점수는 매년 가장 낮은 점수를 획득하였음을 알 수 있고, 자동출입국 시설사용 편리성과 신속성 등 시설서비스는 크게 개선된 반면에 직원의 부적절한 응대, 강압적 태도 등 인적서비스는 개선이 시급함을 알 수 있다.

구분	심사항목	평가 점수					
		'16	'17	'18	'19	'20	평균
출입국 심사	자동시스템 이용편리	85점	94점	95점	97점	96점	93점
	공간 및 설비 충분('18년 신규)	-	-	89점	91점	93점	91점
	신속성	84점	90점	88점	89점	92점	89점
	정확성	83점	90점	88점	89점	92점	88점
	대기시간 적정	82점	89점	86점	88점	91점	87점
	직원 친절도	82점	86점	84점	86점	91점	85점

<표 4-7> 출입국 심사 분야 고객만족도 세부 평가결과('16년~'20년)

먼저 인천공항의 장점을 나열하자면 최고의 시설과, 신속한 출입국심사, 다양한 면세점, 세계에서 유일하게 공짜로 물을 주는 음수대를 운영, 화장실도 세계에서 가장 많은 공항,공짜로 샤워도 가능한거같고, 공짜로 편한 쇼파에서 쉴수 있게해주는등 세계어딜내놓아도 최고의 공항에 틀림없다고 생각합니다. 그런데 출입국심사 공무원들의 근무태도는 그렇지 못하다는것이 조금 아쉽습니다 물론 미국이나 다른 어디나라를 가더라도 출입국심사 공무원은 무뚝뚝하고 인사도 잘안하고 하지만 그건 외국인으로서 그나라에 갈 때 충분히 감수해야 입국과 출국을 할수있으니 참을수있다고 생각합니다. 하지만 대한민국 국민으로 출입국심사를 할때는 좀 달라야한다고 생각합니다. 단 한번도 먼저 안녕하세요 감사합니다. 인사하는것을 본적이 없습니다.마치 심사대에서 나는 판사이니 당신은 가만있으시요하는 느낌입니다. 근엄한 근무자세가 매뉴얼인지는 모르겠으나 요즘 공무원 및 기업에서 친절을 강조하는 시대에 유일하게 불친절하게 아직 근무하는곳인듯 합니다.말을 거의 한마디도 하지 않습니다. 귀찮아 하는 것 같기도 하고 항상 피곤에 찌들어 있어보입니다. 주민,센터와 파출소에 들어가도 인사를 하는 시대인데 아직 출입국심사공무원들은 90년대에 살아가고있는것 같습니다 외국인에게 친절하는것을 바라는것이 아니라 자국민에게 따스한 말한마디를 하면 피곤에 쩔어 귀국할때 더욱 힘나고 감사하겠습니다.

<표 4-8> 출입국 심사 고객 불만족 사례

2. 종합 비교

고객의 소리(VOC) 데이터에 대한 텍스트분석 및 토픽분석 결과에서는 주로 공항직원의 응대, 태도 ,답변, 안내와 같은 인적서비스와 관련된 텍스트 및 토픽의 비중이 높았다. 그러나 고객만족도 조사의 경우에는 다중회귀 분석결과 여객터미널내 편의시설, 안내시설, 상업시설(쇼핑)과 같은 시설서비스가 공항서비스 종합만족도에 큰 영향을 미치는 것으로 나타났다. 이는 고객이 공항이용시 직접적으로 느끼는 서비스 불만요인은 직원의 부적절한 응대, 불친절과 같은 인적서비스에 의해 표출되고 제기되는 반면에, 공항이용시 고객이 느끼는 서비스 만족도의 경우에는 공항시설의 편리/편의성, 신속성, 정확성과 같은 시설 사용성에 의해 큰 영향을 받는다는 것을 확인 할 수 있었다.

인천공항은 이미 세계 최고 수준의 공항시설과 최첨단 운영시스템을 바탕으로 세계에서 가장 편리하고 안전한 공항으로 자리매김해왔다. 그러나 이러한 공항 인프라 시설의 우수성과는 다르게, 직원의 응대 및 안내태도의 적절성 등 인적서비스에 있어서는 공항종사자에 대한 지속적인 서비스 교육과 고객만족도 조사를 통한 모니터링 등 서비스개선 노력에도 불구하고 고객들의 불편·불만사항은 매년 증가해왔다.

인천공항을 구성하는 이해관계자는 정부기관, 상주업체, 시설운영 협력사 등 다양한 기관이 존재하며, 이들 이해관계자가 고객이 공항을 이용할 때 경험하는 행정, 상업, 편의 서비스 등 다양한 서비스를 제공하고 있다. 따라서 고객의 긍정적인 공항이용 경험 강화를 위해서는 기관간의 상호협력을 통해 공통의 CS 전략체계 확립이 필요하며, 이를 위해 공항운영자인 공항공사는 고객서비스 품질 향상을 위해 고객 친화적 서비스 확대, 고객중심의 CS관리체계 강화, 전사차원의 CS 교육 프로그램 운영 등의 많은 노력을 기울여야 할 때이다.

제 5 장 결론

제 1 절 연구의 요약

본 연구는 '공항 서비스에 대한 고객 불만족 요인'을 파악하기 위해 인천공항을 이용하는 고객들의 소통창구인 VOC(고객의 소리) 시스템에 접수된 민원 데이터를 수집 및 분석하였다. 분석프로그램은 텍스톰(Textom)의 빅데이터 분석 솔루션인 텍스트 마이닝 기술을 이용하였고, 주요 키워드 도출을 위한 텍스트 빈도분석, 문서내 주요 토픽을 도출하기 위한 토픽모델링 분석, 문서에 포함한 주관적인 감성을 추출해 내기 위한 감성분석을 순서대로 진행하였다. 마지막으로 감성분석 결과를 바탕으로 공항서비스 유형별 서비스 긍정지수를 산출하였고 이를 실제 고객만족도 조사결과와 비교함으로써 공통적으로 일치하는 서비스 품질 영향요인을 재확인하고, 기존의 조사결과와 일치하지 않는 서비스 품질에 대한 요인을 도출하였다.

먼저 키워드 빈도에 대한 분석결과 직원, 사람이라는 키워드의 출현빈도가 높은 것으로 보아 인적서비스와 관련된 문서가 많은 것을 확인할 수 있으며, 서비스 유형별로는 접근교통 및 주차, 보안검색 서비스와 관련된 문서가 많았음을 확인할 수 있었다.

다음으로 토픽모델링(LDA)분석을 시행하였고, 문서에 잠재된 토픽을 추출하기 위해 정합도(coherence score)를 이용하여 최적토픽수 8개를 결정하였으며, 토픽을 구성하는 상위 30개의 단어를 바탕으로 8개의 토픽(① 주차서비스, ② 인적서비스(단순불만), ③ 인적서비스(강한불만), ④ 상담/문의사항, ⑤ 상업시설서비스, ⑥ 공항시설서비스, ⑦ 공항공사 개선노력, ⑧ 접근교통서비스)을 구성하였고, 토픽별 비중에 따른 분석결과는 다음과 같다.

먼저 직원응대와 관련한 인적서비스인 ②번 토픽(18.7%)과 ③번 토픽(15.6%)의 비중이 가장 높았으며, 합계 비중도 전체 토픽의 34%이상을 차지하였다. 다음으로 시설서비스인 공항시설과 관련한 ③번 토픽(15.1%)의 비중이 높았다. 또한 유사 서비스 분야인 주차서비스와 접근교통서비스에 대한 토픽비중이 각각 12.6%, 11.2%를 차지하였다.

8개 주요토픽의 시간의 흐름에 따른 시계열적 변화를 분석해본 결과, 대부분의 토픽이 '17년도에 비해 '18년도에 급격하게 증가하였고, 특히 공항시설과 관련한 6번토픽의 비중이 57%이상 증가한 것을 확인할 수 있었다. 이는 인천공항 제2여객터미널이 2018년 1월 그랜드 오픈함에 따라, 전체적으로 토픽별 비중이 전년도에 비해 증가했으며 공항시설과 관련된 6번토픽의 비중이 급격하게 증가했음을 보여준다.

'19년도에는 공항운영 안정화에 따라 전체적인 토픽의 빈도수가 감소하였으나, 직원응대와 관련한 2번토픽의 비중은 여전히 증가하고 있음을 보여준다. 또한 1번토픽(주차편의)과 8번토픽(접근교통)이 서로 밀접한 연관이 있으며 시간의 흐름에 따라 유사하게 변화하는 경향을 보이고 있음을 확인할 수 있다. '18년도에 제2여객터미널 오픈 및 터미널 분리운영으로 여객의 목적 터미널 혼선에 따른 교통 및 주차분야의 민원건수가 급증하였으나, 대중교통 안내 시스템 개선, 셔틀버스 노선 증설 및 배차시간 조정, 불법 호객행위 택시 및 부당요금에 대한 단속 등으로 접근교통서비스 관련한 토픽 빈도수는 급감한데 비해, 불법주차 대행, 주차차량 파손, 주차요금 감면절차 복잡 등 주차 서비스 불편·불만은 지속됨에 따라 주차편의 서비스는 크게 개선되지 않았음을 보여주었다.

다음으로 문서의 극성정도(긍정/부정/중립)를 파악하기 위해 별도의 감성분석을 시행하였다. 감성분석결과 고객 불편·불만 등의 부정적 감성을 나타내는 문서가 전체 문서의 60.2%, 상담·문의, 의견제안 등 중립적 감성을 나타내는 문서가 37.5%, 칭찬·격려 등의 긍정적 감성을 나타내는 문서는 2.3%로 가장 낮은 비율로 나타났다.

감성분석 결과를 바탕으로 부정적 감성을 나타내는 문서와 긍정적 감성을 나타내는 문서를 구분하여 추가로 텍스트 분석을 시행하였다.

부정적 감성의 문서에서는 ‘직원’, ‘말’, ‘태도’, ‘생각’, ‘교통’, ‘주차’, ‘버스’, ‘검색’ ‘불편’ 등의 키워드가 상위를 차지함에 따라 접근교통 및 주차서비스와 보안검색 서비스 분야에서 공항직원의 인적서비스에 대해 공항이용객의 불편·불만족이 많았음을 확인할 수 있었다. 반면에 긍정적 문서에서는 ‘직원’, ‘감사’, ‘칭찬’, ‘마음’, ‘도움’, ‘친절’, ‘핸드폰’, ‘가방’, ‘안내서비스’, ‘안내데스크’ 등의 키워드가 상위를 차지함에 따라 공항안내 직원의 친절한 안내 서비스에 고객들이 감사해하는 내용이 많았음을 확인할 수 있었다.
다음으로 최근 5년간(‘15년～‘20년)의 인천공항 고객만족도 조사결과를 활용하여 7개 공항 핵심 서비스 유형별 만족도 점수를 산출하였고, 연도별 변화 추이를 분석해보았다. 교통 및 주차서비스 분야의 점수가 매년 가장 낮은 만족도 점수를 기록하였고, 대체적으로 7개분야의 만족도 점수가 전년도 대비 매년 상승하는 경향을 보이고 있다. 또한 고객만족도 조사에 대한 다중회귀 분석결과, 독립변수인 7개 공항서비스 유형중 종속변수인 공항서비스 종합만족도에 통계적으로 유의미한 영향을 미치는 변수는 보안검색만족도, 상업시설(쇼핑매장) 만족도, 안내서비스 만족도, 편의시설 만족도로 나타났다. 공항서비스 종합 만족도에 가장 큰 영향을 미치는 변수는 편의시설 서비스이며, 다음으로 보안검색 만족도, 안내서비스 만족도, 상업시설(쇼핑매장) 만족도 순으로 유의수준 5%하에서 통계적으로 유의미한 영향을 미치는 것으로 분석되었다. 고객만족도 조사의 경우 직원응대와 같은 인적서비스 측면보다는 여객터미널내 시설사용의 편리·편의성, 신속·정확성 등의 시설서비스가 공항서비스 종합만족도에 영향을 미치는 것으로 해석할 수 있을 것이다.
마지막으로 고객의 소리(VOC) 분석결과와 실제 고객만족도 조사결과의 서비스 유형별(7가지) 상호비교분석을 시행하였다. 고객만족도 조사는 최근 5년간 조사결과의 평균값을 적용하였고, 고객의 소리(VOC) 문서는 서비스 유형별로 서비스 긍정지수를 산출하여 상호 비교하였다. 비교 분석결과, 5개 서비스 유형(① 교통/주차, ④ 체크인서비스, ⑤ 상업시설, ⑥ 안내서비스, ⑦ 편의시설)은 고객만족도 점수와 서비스 긍정지수 간에 공통적으로 일치하는 경향을 보였다. 반면에 다음의 2개 서비스 유형(②

보안검색 , ③ 출입국심사 분야)은 고객만족도 점수와 서비스 긍정지수 간에 차이점을 나타내는 것을 확인할 수 있었다. 특히 출입국심사 분야는 고객만족도 점수는 가장 높은 점수를 획득하였으나, 서비스 긍정지수는 세 번째로 낮은 점수를 획득하였는데 이는 자동출입국심사 도입에 따라 출입국 심사 이용 편리성 및 여권/비자 심사의 신속성과 정확성 등은 크게 개선되었으나, 출입국 심사관의 불친절한 응대 및 대기시간 지연에 따른 고객의 불편·불만사항이 많았던 것으로 분석되었다.

연구결과를 종합해보면, 고객의 소리(VOC) 데이터에서는 주로 공항직원의 응대, 태도 ,답변, 안내, 말투 등 인적서비스 관련 텍스트 및 토픽의 비중이 높았으나, 고객만족도 조사의 경우는 다중회귀 분석결과 여객터미널내 편의시설, 안내시설, 상업시설(쇼핑)과 같은 시설서비스가 공항서비스 종합만족도에 큰 영향을 미치는 것으로 나타났다. 이는 고객이 공항이용시 직접적으로 느끼는 서비스 불만요인은 직원의 부적절한 응대, 불친절과 같은 인적서비스에 의해 적극적으로 표출되는 반면에, 고객이 느끼는 공항서비스에 대한 만족도의 경우에는 시설사용의 편리/편의성, 신속성, 정확성 등에 의해 큰 영향을 받는다는 것을 확인할 수 있었다.

제 2 절 연구의 의의 및 시사점

본 연구는 인천공항 서비스 만족도 분석을 위해 고객의소리(VOC)에 접수된 문서에 대한 키워드 분석, 토픽분석과 감성분석을 수행하여 기존의 고객만족도 조사결과와의 유사점과 차이점을 살펴보았다. 다음에서는 분석결과를 바탕으로 한 본 연구의 의의 및 인천공항 서비스 만족도 제고를 위한 정책적 제언을 제시하고자 한다.

첫째, 인천공항 고객의 소리(VOC) 문서에 나타난 주요 토픽을 예측해보고, 시간의 흐름에 따른 토픽의 변화를 분석함으로써 고객만족도 제고를 위해 서비스 개선이 필요한 분야를 파악할 수 있었다.

기존의 인천공항 고객만족도 관련 연구는 대부분 설문조사 방식으로 진행됨에 따라 조사 데이터가 모집단의 특성을 정확하고 깊이있게 반영하

기 힘들며 시계열적으로 분석하는데 있어 많은 한계점을 가지고 있었다. 이에 본 연구는 고객의 소리(VOC) 자료에 대한 토픽분석을 통해 고객의 불편·불만족사항에 대한 주요 이슈를 도출해내었다는 데에 그 의의가 있다. 둘째, 고객의소리(VOC) 텍스트에 포함된 감성(긍정/부정/중립)을 추출하기 위해 감성분석을 시행함으로써, 해당 문서가 긍정적인지 부정적인지 중립적인지를 구분하였고, 긍정문서의 주요 키워드와, 부정문서의 주요 키워드를 각각 도출하였다. 기존의 텍스트마이닝 분석의 연구는 대부분 텍스트분석, 토픽분석까지 진행되었으나, 본 연구는 연구자가 직접 학습 데이터를 구성하여 적용함으로써 전체문서에 대한 감성을 추출하였다는 데에 그 의의가 있다.

마지막으로 감성분석 결과를 바탕으로 공항 핵심 서비스 유형별로 서비스 긍정지수를 산출하였고, 이를 실제 고객만족도 조사결과와 상호 비교분석을 시행하였다. 이를 통해 공통적으로 일치하는 서비스 품질 영향요인을 재확인하였고, 기존의 연구결과에서 확인하지 못했던 서비스 품질에 관한 새로운 이슈를 도출하였다.

최근의 고객만족도 조사결과를 분석해보면 서비스 유형별 만족도 점수는 매년 상승하는 경향을 보이고 있으며, 평균점수도 90점 이상으로 유형별로 큰 편차가 없어 고객만족도 제고 및 개선에 큰 변별력이 없다는 한계점을 지니고 있다. 따라서 매년 분기별로 진행되는 현행 고객만족도 조사의 평가방법 및 세부 평가항목에 대한 재검토 및 개선이 필요할 것이다. 앞으로는 고객의 소리(VOC) 시스템에 빅데이터 분석기법을 도입함으로써 고객의 불편·불만족과 관련한 세부적인 분석이 가능할 것이고, 기존 고객만족도 조사와 상호 보완적으로 활용한다면 공항서비스 고객만족도 향상에 큰 기여를 할 수 있을 것이다.

제 3 절 연구의 한계 및 향후 연구방향

이 연구는 인천공항의 고객만족도 영향요인을 분석하기 위해 기존의 설문조사방식의 연구와 다르게 인천공항에 고객이 직접적으로 제기한 고객의 소리(VOC) 민원자료를 활용하여 고객의 서비스 불만족 유형을 확인하고 고객만족도 조사결과와 상호 비교 검증함으로써 기존의 연구에서 확인하지 못했던 새로운 서비스 품질요인을 도출하고자 하였으나, 다음과 같은 한계점을 가진다.

우선 본 연구는 인천공항 고객의 소리(VOC) 시스템에 접수되는 문서만을 분석 대상으로 함에 따라, 공항이용과 관련한 각종 언론이나 뉴스, 개인 SNS(Social network service) 이용후기 등 다양한 자료들을 수집하지 못해 다양한 이해관계자들의 불만족 사항을 분석 및 파악하는데는 한계점을 가지고 있다. 따라서 향후에는 인천공항 뿐만 아니라 전국 공항의 고객 불만족 관련 자료들을 수집하고 활용한다면 다양한 이해관계자들의 관심사항을 파악함으로써 공항 서비스 만족도 제고 및 개선에 큰 도움이 될 것이다.

다음으로 고객의 소리(VOC) 시스템에 접수된 문서는 고객이 인천공항을 이용하면서 느끼는 불편·불만과 관련된 부정적인 내용이 과반수 이상으로, 일반 고객을 대상으로 하는 고객만족도 조사결과와 상호 비교 검증한다는 것은 표본선택에 있어 편향이 발생할 수 있다는 문제점이 있다.

마지막으로 텍스트마이닝 분석기법 자체가 가지고 있는 한계점으로 문서내 동시출현빈도가 높은 단어배열과 단어-토픽, 문서-토픽의 분포를 고려하여 해당 토픽의 비중에 따른 우선순위에 대한 추론은 가능하나, 어떤 토픽이 상대적으로 중요한지를 판단하는 기준은 아직 정립되지 않았다는 점이다. 즉 고객의 소리(VOC)에 접수된 문서들에 대한 주요 이슈들과 관련한 토픽을 예측해보고 토픽의 시간의 흐름에 따른 변화를 파악하는 것은 가능하나, 어떤 토픽이 고객만족도에 미치는 영향이 큰지를 분석하는데는 한계가 있다는 점이다.

향후에는 빅데이터 분석기법인 머신러닝 기술의 발달과 함께 데이터의 패턴 분석에 따른 유용한 정보의 제공뿐만 아니라 구조적 분석을 통해 상관관계나 인과관계까지 규명할 수 있기를 희망해 본다.

참 고 문 헌

<단행본>

김병섭(2010), 「편견과 오류 줄이기- 조사연구의 논리와 기법」, 법문사
남궁근(2016), 「행정조사방법론」, 법문사
인천공항공사(2017), 「인천공항 CS 중장기 전략수립 용역」, 한국생산성본부
인천공항공사(2019), 「인천국제공항 고객만족도 조사 결과보고서」
인천공항공사(2015~2020), 「인천국제공항 고객의 소리(VOC) 운영결과」

<신문·잡지 기타>

인천국제공항공사 홈페이지 : http://www.airport.kr
국제항공협의회(ACI) 홈페이지 : http://www.aci.aero
스카이트랙스(SKYTRAX) 홈페이지 : http://www.worldairportawards.com
텍스톰 홈페이지 : http://www.textom.co.kr

<논문>

김영찬, 차재성(2003) "고객만족도 측정방법론과 전략적 활용"(마케팅 연구 제18권 제1호)
장대성(2004) "국제공항 서비스품질 측f정도구 개발에 관한 연구(인천 및 김포국제공항을 대상으로)"(품질경영학회지Vol.32(2))
최현정(2005) "대구국제공항 이용고객의 만족도 제고방안"(계명대학교 석사학위논문)
최병권(2007) "공항서비스에 대한 인천국제공항 이용자의 지각된 서비스 품질과 만족에 관한 연구"(무역학회지 제32권 제2호)
장순자, 김기웅(2010) "복수공항 이용객의 서비스 만족에 영향을 미치는 품질요인의 비교 연구"(한국항공운항학회지, Vol.18(3))
임재필(2011) "ASQ 경쟁력 강화를 위한 인천국제공항의 서비스 전략 방안에 관한 연구" (경영경제연구 제34권 제2호)

정지선(2011) “신가치창출 엔진, 빅데이터의 새로운 가능성과 대응 전략, 새로운 미래를 여는 빅데이터 시대”(한국정보화진흥원, 빅데이터 전략연구센터)

송상영, 김사혁(2012) “빅데이터 시대의 새로운 정책이슈와 이용자 중심의 활용방안 연구”(정보통신정책연구원 12-진흥-097)

정윤희, 박용화, 백소라(2013) “인천공항 여객터미널 서비스의 실패와 고객만족에 미치는 회복 공정성의 영향 분석”(한국항공경영학회지 Vol.11(5))

임성빈(2013) “허브공항으로서의 인천국제공항의 경쟁력 분석” (서울대 석사학위 논문)

배동민, 박현수, 오기환(2013) “빅데이터 동향 및 정책 시사점”(정보통신방송정책, 25(10))

공효순(2014) “빅데이터 기반 호텔고객 평판 분석에 관한 연구”(한국디지털콘텐츠학회, 15(2))

이청림, 라선아(2015) “공항 서비스 품질 및 고객만족 모형의 개발: 인천공항 출입국 고객의 MOT분석을 통한 접근”(한국경영학회지 Vol.13(5))

이지영(2015) “빅데이터의 국가통계 활용을 위한 기초연구”(2015년 상반기 연구보고서 제 Ⅱ권 3장)

이동준(2016) “소셜 네트워크 빅데이터 기반 택배업체 고객만족도에 관한 연구”(한국 전자거래학회,21(4))

김지은(2017) “빅데이터를 활용한 대학구조개혁 평가의 키워드 및 토픽분석” (서울대 교육학박사학위논문)

진대용(2019) “생활밀착형 환경이슈에 대한 수요반영 개선연구”(한국환경정책·평가연구원 정책보고서)

한주형(2019) “가격이 중저가호텔과 게스트하우스의 고객만족도 평점에 미치는 영향:온라인리뷰 빅데이터 분석을 중심으로”(관광경영학회 제23권 제2호)

엄준석(2020) “지방 상수도 고객 민원 빅데이터 분석”(한국통신학회 학술대회논문집)

Adrian, C., Sidi, F., Abduliah, R., Ishak, I., Affendey, L. S., & Jabar, M. A.(2016). Big data analytics implementation for value discovery: asystematic literature review. *Journal of Theoretical and AppliedInformation Technology*, 93(2), 385-393.

Blei, D., & Jordan, M. (2003). Modeling annotated data. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval* (pp. 127-134).ACM.

Blei, D., & Lafferty, J. D. (2009). Topic models. *Text mining : Classification, Clustering, and Applications*, 10(71), 34.

Benoit, K., & Herzog, A. (2015). *Text analysis: estimating policy preferences from written and spoken words*. Retrieved from http://www.kenbenoit.net/pdfs/HerzogBenoit_bookchapter.pdf.

Cronin & Taylor(1992) "Reexamination and Extension", (Journal of Marketing, p.58~59)

Gronroos(1984) "A Service Quality Model and Its Marketing Implications" (European Journal of Marketing, p.18).

Hale, R. (2005). Text mining: Getting more value from literature resources. *Drug Discovery Today*, 10(6)

Khan, R. A., & Kanth, S. (2016). Text Mining: Knowledge Discovery from Unstructured Data. *Artificial Intelligent Systems and Machine Learning*, 8(2), 71-77.

Krippendorff, K. (2004). *Content analysis: an introduction to tis methodology.* 2nd Ed., Thousand Oaks, CA: Sage Publication.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6 February.

Li, F. (2010). Textual analysis of corporate disclosures: a survey of the literature. *Journal of Accounting Literature*, 29, 143-165.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburghh,

C & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute Report*, May 2011. Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

Mahgoub, H., Rösner, D., Ismail, N., & Torkey, F. (2008). A text mining technique using association rules extraction. *International journal of computational intelligence*, 4(1)

Ndoh, Ashford(1994) "Evaluation of Transportation Level of Service Using Fuzzy Sets", Transportation Research Record, 1461, p 31-37

Parasuraman, Berry & Zeithaml (1988) "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perception of Service Quality," (Journal of Retailing, 64(1). p.12～40)

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data_analytic thinking.* (강권학, 옮김). 서울: 한빛미디어(주). (원서출판 2013)

Riffe, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research*. New York, NY: Routledge.

Salton, G. & Buckley, C. (1988). Term-weighing approache sin automatic text retrieval. *Information, Processing & Management*, 24(5): 513-523.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis: A Road to Meaning,* pp. 427-448. Mahwah, NJ: Lawrence Erlbaum Associates.

White, M. (2012). Digital workplaces: Vision and reality. *Business Information Review*, 29(4), 205-214.

White, P., & Breckenridge, R. S. (2014). Trade-offs, limitations, and promises of big data in social science research. *Review of Policy Research*, 31(4), 331-338

Abstract

A Study on Big data-based Airport Service Satisfaction Analysis
: Focusing on the VOC(Voice of Customer)

Hyunkyu, Koo
Department of Public Enterprise Policy
The Graduate School of Public Administration
Seoul National University

Recently, in the airport industry, the surrounding competitive airports are making bold investments in expanding facilities, and preemptively introducing advanced smart systems based on IOT, AI, and Big Data. Therefore, it is expected to be very difficult that Incheon International Airport, which has geopolitical risks of the Korean peninsula, will secure a comparative advantage in competitiveness with the world's leading airports in terms of facility size and number of passengers.
In this research, the service dissatisfaction factors were identified to secure the service competitiveness of Incheon International Airport.
And the big data analysis of VOC(Voice of Customer) was conducted for the mixed method study on the factors affecting customer satisfaction to establish service improvement and recovery strategy for customer satisfaction according to dissatisfaction factors
First, text frequency analysis and topic analysis were conducted to identify customer dissatisfaction factors, and then the comparative analysis on big data analysis and the results of actual customer satisfaction survey were performed.
The purpose of this study is to reaffirm the factors affecting service quality that are common, and to derive new issues related to service quality that have not been confirmed in the existing research.

The research questions of this study are as follows.

1. What are the main keywords and topics of VOC contents in the last five years?

2. How have the major topics changed over time over the last five years, and what are the topics that increase interest and decrease interest?

3. What are positive and negative keywords of VOC contents in the last 5 years through the emotional analysis results of the entire document?
4. What are the similarities and differences between the VOC big data analysis and actual customer satisfaction scores?

First, In order to collect and analyze the data related to inconvenience and dissatisfaction of Incheon International Airport passengers, about 33,000 customer complaints data received through the VOC system of Incheon International Airport in the last 5 years(the second half of 2015 to the first half of 2020) were used.
In addition, the results of customer satisfaction survey of Incheon International Airport passengers for the last 5 years (2015 to 2020) were used to compare and verify the existing customer satisfaction survey results. The main results of this study are as follows.
1. As a result of frequency analysis of keywords for VOC documents, it was confirmed that there were many documents related to human services due to the high frequency of keywords such as 'employee', 'people', 'bus', 'parking', 'parking lot', 'vehicles', 'taxi', and 'valet services', and as for the type of service, many documents were related to access transportation and parking service.

2. The topic modeling (LDA) analysis of VOC documents was performed, 8 optimal number of topics were determined using the coherence score to extract the latent topic in the document, and 8 topics(①Parking service, ②Human service(simple complaint), ③ Human service(strong complaint), ④Counseling/Inquiry, ⑤Commercial facility service, ⑥Airport facility service, ⑦Airport improvement effort, ⑧Access transportation service) were extracted based on the top 30 words that consist topics.

First, The topic with the highest proportion was the topic related to human service(topic number ②, and topic number ③) such as employee response, followed by the topic number ③ related to airport facilities, which is an infrastructure service.

As a result of analyzing the time series change according to the time of the 8 major topics, most of the topics increased sharply in 2018 compared to 2017, and especially the proportion of the topic number ⑥related to airport facilities increased by more than 57%.

This shows that as the Incheon International Airport Terminal 2 opened in January 2018, the total number of topics increased, and as the airport facilities and systems were still unstable, the proportion of the topic number ⑥ increased sharply.

In addition, the overall frequency of topic decreases in 2019 due to the stabilization of airport operation, but the proportion of topic number ② related to employee service, which is human service, is still increasing.

3. In order to identify the degree of polarity(positive/negative/neutral) of the document, separate emotional analysis was conducted.

60.2% of the total document showed negative emotions such as customer inconvenience and dissatisfaction, 37.5% of the document showed neutral emotions such as counseling, inquiry, and opinion proposal, and 2.3% of the document showed positive emotions such as praise and encouragement.

In the negative documents, as the keywords such as 'employee', 'words', 'attitude', 'thought', 'traffic', 'parking', 'bus', 'search' and 'inconvenience' were ranked at the top, it was confirmed that airport passengers were inconvenient and dissatisfied with the human service of airport employees in the fields of access traffic, parking service and security search service.

In the positive documents, as the keywords such as 'employee', 'thanks', 'praise', 'mind', 'help', 'kindness', 'mobile phone', 'bag', 'information service', 'help desk' were ranked at the top, it was confirmed that customers were grateful for the friendly guide service of airport service representative.

4. For the comparative analysis of VOC and actual customer satisfaction survey, the service positive index for VOC was calculated by each service type and compared with actual customer satisfaction score. As a result of the analysis, the five service types(①Traffic/Parking, ④Check-in service, ⑤Commercial facility, ⑥information service, ⑦Amenity) showed a common tendency that customer satisfaction score and service positive index tended to match.

And the two service types, ②security search and ③immigration, showed differences between customer satisfaction score and service positive index.

In the immigration field. the customer satisfaction score was the highest , but the service positive index was the third lowest.

This showed that the convenience of immigration and the speed and accuracy of passport/visa screening were greatly improved due to the introduction of smart entry service, but there were many inconveniences and complaints of customers due to unfriendly response of immigration judges and delay of waiting time.

This study derived major issues about customer's inconvenience and dissatisfaction through keyword analysis and topic analysis on VOC

data, and confirmed the trend of customer's main interest through time series analysis of topic.
In addition, emotional analysis was conducted to identify the degree of polarity(positive/negative/neutral) of VOC documents, and service positive index was calculated by the core service types of airports based on the results of sentiment analysis.
This study is meaningful in that it conducted a comparative analysis with the actual customer satisfaction survey results, and derived the factors affecting service quality that were commonly matched and new issues related to service quality that were not confirmed in the existing research results.

Keywords : Big data, Text mining, VOC(Voice of customer), Keyword analysis, Topic modeling, Sentiment analysis, Service positive index, Customer satisfaction

Student Number : 2020-23375