

Original Article

<https://doi.org/10.12985/ksaa.2023.31.4.019>
ISSN 1225-9705(print) ISSN 2466-1791(online)

Text Mining 기법을 활용한 항공안전관리 이슈 분석

권문진*, 이장룡**

Analysis of Aviation Safety Management Issues using Text Mining

Moonjin Kwon*, Jang Ryong Lee**

ABSTRACT

In this study, a total of 2,584 domestic research papers with the keywords "Aviation Safety" and "Aviation Accidents" were subjected to Text Mining analysis. Various text mining techniques, including keyword frequency analysis, word correlation analysis, network analysis, and topic modeling, were applied to examine the research trends in the field of aviation safety. The results revealed a significant increase in research using the keyword "Aviation Safety" since 2015, with over 300 papers published annually. Through keyword frequency analysis, it was observed that "Aircraft" was the most frequently mentioned term, followed by "Drones" and "Unmanned Aircraft." Phi coefficients were calculated for words closely related to "Aircraft," "Aviation," "Drones," and "Safety." Furthermore, topic modeling was employed to identify 12 distinct topics in the field of aviation safety and aviation accidents, allowing for an in-depth exploration of research trends.

Key Words : Air Accident(항공사고), Aviation Safety(항공안전), Safety Management(안전관리), Text Mining(텍스트마이닝), Topic Modeling(토픽모델링)

1. 서 론

항공안전관리란 비행 또는 이와 관계가 있는 운용 및 작업 등에서 발생할 수 있는 사고를 예방하기 위해 승무원, 항공기 및 기타 자원의 안전위험(Risk)을 선제적으로 조치하는 행위이다(박원화, 2009). 국제민간항공기구(ICAO)는 세계적 항공 활동의 복잡성이 지속 증가함에 따라 전통적 수단으로는 항공분야에서의 안전 위험을 수용 가능한 수준 이내로 관리하기에 그 효과

성이 감소하고 비효율적이 되었기 때문에 이러한 안전 위험을 이해하고 관리할 수 있는 새로운 수단을 Annex 19을 통해 전세계 회원국에게 제시하였다(ICAO, 2016).

ICAO Annex 19 Safety Management가 제시한 안전관리는 안전데이터 및 정보에 기반하여 수행하는 활동으로써 전통적 안전관리와의 가장 큰 차이점이라고 할 수 있다.

항공안전관리의 성공적 운영을 위해서는 안전데이터 및 정보의 수집과 분석이 필수적이다(소재선 외, 2014). 우리나라도 항공안전법에 항공안전데이터 수집 및 처리에 관한 조항을 명문화하였고, 항공안전데이터 처리 및 활용에 관한 규정을 제정하여 운영하고 있다(국토교통부, 2021).

한편, 최근 항공산업은 전통적인 항공기 제조, 항공 정비(MRO, Maintenance · Repare · Operation), 항

Received: 16. Oct. 2023, Revised: 24. Oct. 2023,
Accepted: 26. Oct. 2023

* 한국항공대학교 항공운항관리학과 박사과정

** 한국항공대학교 항공운항학과 교수

연락처 E-mail : jrherky@kau.ac.kr

연락처 주소 : 경기도 고양시 덕양구 항공대학로 76, 본관
524호

공운송사업뿐만 아니라, 무인항공기 및 도심항공교통(UAM, Urban Air Mobility) 등이 새롭게 부상하면서 성장세가 두드러지고 있다. 세계 항공산업의 규모는 COVID-19의 여파로 잠시 주춤하기는 하였지만 2030년경 약 9,281달러로 본격적인 성장궤도 진입이 전망된다(이성일, 2022). 항공산업의 성장과 비례하여 항공 안전의 중요성도 더욱 커지고 있다. 광범위한 항공산업 영역에서 어떠한 안전 이슈(Issue)에 항공관계인들이 관심을 가져야 하는지를 파악할 수 있다면 효율적 항공안전관리에 큰 도움이 될 것이다.

항공안전관리에 주목해야 할 안전 이슈(Issue)를 파악하기 위해서는 ICAO가 강조한 데이터 및 정보의 수집 및 처리에 관한 지침을 적용해 볼 수 있다. 더욱이 빅데이터 기술의 발달로 방대한 양의 자료 분석이 가능해졌고 이를 통해 항공안전 이슈를 분석할 수 있는 기술적 환경도 마련되어 있다. 이에 본 연구는 국내 항공안전 분야 연구 자료들에 대해 텍스트 마이닝 기법을 활용한 토픽 모델링을 실시하여 항공안전관리에 관심을 가져야 할 이슈들을 분석하였다.

II. 이론적 고찰 및 선행연구

2.1 안전데이터 및 정보

ICAO(2013)는 항공안전관리에 안전데이터, 정보의 분석 및 활용은 무엇보다 중요하다고 강조하였다. 안전데이터는 관찰 또는 측정의 결과로 보고되거나 기록된 것을 의미하고, 주어진 상황에 맞춰 처리, 조직, 통합 또는 분석될 때 안전정보로 변환되어 안전관리에 유용하게 사용되기 때문에 효과적인 안전관리는 안전데이터의 수집, 분석, 관리 기능에 크게 영향을 받는다(ICAO, 2018)

항공안전관리에 활용하기 위해 안전데이터는 정제(Clean)되어야 하고, 활용 목적에 적합해야 하는데, 이러한 것들이 데이터의 품질에 큰 영향을 미치기 때문이다. 높은 데이터 품질을 보장하기 위해서는 다음과 같은 4가지 요소에 대한 고려가 필요하다(ICAO, 2018).

- 정제성(Cleanliness): 손상, 부정확, 관련 없는 데이터는 대체, 수정, 삭제되어야 한다.
- 관련성(Relevance): 조직의 요구와 문제를 나타낼 수 있어야 한다.
- 적시성(Timeliness): 현재의 이슈를 다룰 수 있어야 한다.

- 정확성(Accuracy & Correctness): 주어진 상황(given scenario)을 설명할 수 있어야 한다.

2.2 안전 분석

안전 분석(Safety Analysis)은 유용한 정보를 발견하고 결론을 제안하며, 의사결정을 지원하기 위해 안전 데이터를 확인, 검사, 기술, 변형, 압축, 평가, 통계 또는 기타 분석 기법을 적용하는 과정을 의미한다(ICAO, 2018).

ICAO(2018)는 항공안전관리를 위해 활용하는 주요 안전 분석 방법으로 4가지 유형을 제시하였다(Fig. 1).

첫째는 기술적 분석(Descriptive analysis)이다. 이 방법은 수집한 데이터를 표, 그래프, 차트 등 의미있고 유용한 방식으로 서술하고 요약하는 것으로써 패턴이나 트렌드 등을 설명, 표시, 요약할 수 있어서 해석이 용이하고, 사례연구, 기회, 문제점을 명확하게 정의하는데 활용된다. 그러나, 이 방법은 데이터에 관한 정보는 제공하지만, 분석한 정보를 넘어서 결론까지 도출하기에는 제한을 가진다.

둘째는 추론적 분석(Inferential analysis)이다. 귀납적 분석이라고도 하는 이 방법은 데이터의 표본을 통해 모집단에 대한 이해를 목적으로 하며, 데이터들이 가지는 일반화 된 결론을 도출할 수 있도록 하는 기법으로 볼 수 있다. 모수 추정(Estimating parameter), 통계적 가설 검정(Testing statistical hypothesis) 등이 여기에 해당한다.

세 번째는 예측적 분석(Predictive analysis)이다. 이 분석은 과거 및 현재 데이터에서 추출한 정보로 향후 추세 및 패턴을 예측하는 것으로써 떠오르는 위험 및 기회 식별에 도움을 준다. 이미 발생한 사건 변수들 사이의 관계를 포착하고, 이를 활용하여 알려지지 않은 결과를 예측하는 데 활용되며, 때로는 서로 다른 리스크 결과에 대한 시나리오를 모델링을 할 수 있게 한다.

네 번째는 통합적 분석(Combined analysis)이다.

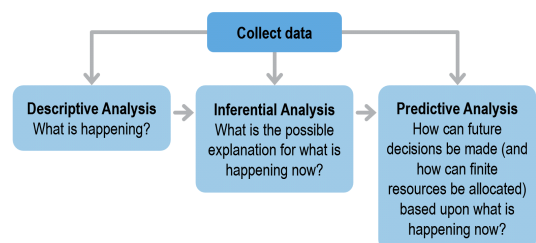


Fig. 1. Common statistical analysis types

이 방법은 위에 언급한 기술적 분석, 추론적 분석, 그리고 예측적 분석 방법 중 2개 이상을 함께 활용하여 분석하는 방법이다. 통계분석 방법들은 서로 연관되어 있어 종종 함께 활용된다. 예를 들어, 추론적 분석은 일련의 데이터와 관련된 결론 도출시 사용되는 주요 수단이지만, 여기에는 기술적 분석이 함께 사용되기도 하고, 추론적 분석 결과는 예측적 분석의 기초자료로 사용되기도 한다.

2.3 텍스트 마이닝(Text Mining)

텍스트 마이닝이란 자연어로 구성된 비정형 텍스트 데이터(Unstructured Text Data)에 숨겨진 패턴 또는 관계를 추출하여 의미 있고 활용 가치가 높은 정보 또는 지식을 찾아내는 분석 기법으로서 자연어 처리(Natural Language Processing) 기술을 기반으로 한다(김현정 외, 2015). 텍스트 마이닝은 비정형 데이터 수집에서 데이터 전처리, 정보 추출, 정보 분석 과정으로 진행된다. 텍스트 마이닝의 주요 기법 중 하나인 토픽 모델링은 문서를 구성하는 키워드를 기반으로 추출한 주제(Topic)를 기준으로 문서를 분류(Clustering)하고 분석한다. 문서에 내재된 잠재적 주제를 도출하여 새로운 문서의 주제에 대한 예측 및 분석이 가능하기 때문에 다양한 분야에서 비정형 텍스트를 분석하기 위해 사용하고 있다(김진국 외, 2022).

해외 선행연구에 의하면 항공기 운항 및 유지보수 등 항공안전관리 분야에 텍스트 마이닝 기법이 활용되고 있다. Koteeswaran(2019)은 1919년부터 2014년까지 미연방항공청(FAA, Federal Aviation Regulation)의 항공사고/준사고 데이터베이스 내 사고조사보고서 분석에 데이터 마이닝 기법을 적용하였다. 인공신경망, SVM(Support Vector Machine), K-NN(Nearest Neighbor) 등 지도학습 머신러닝 알고리즘을 이용하여 항공기 사고를 발생시키는 잠재적인 요인을 분류하고, 주요 인자의 개수를 확인함으로써 항공사고를 줄이고, 항공안전관리시스템을 지원할 수 있는 방안을 제시하였다.

Anderson(2017)은 일반항공(General Aviation)의 조종능력상실(LOC, Loss of control) 사고 보고서에 대한 NTSB(National Transportation Safety Board)의 사고 원인을 항공기 감항인증 기준에 따라 분석하였다. FAA Part 23, CAR(Civil Air Regulation) 3, LSA(Light Sport Aircraft), E-AB(Experimental-Amateur Built)로 구분된 항공기의 사고조사 보고서

에 서술된 원인을 텍스트마이닝 기법을 이용하여 분석하였다. 그 결과, CAR 3에 해당하는 항공기에 계기 이상 조건에 대한 시계비행과 관련된 단어 클러스터를 확인하였고, 1960대 이전에 설계 및 제조된 항공기는 LOC 사고예방에 도움이 되는 기술이 적용되지 않았음을 확인하였다.

2.4 토픽 모델링(Topic Modeling)

토픽 모델링의 대표 알고리즘인 LDA(Latent Dirichlet Allocation, 잠재 디리클레 할당)는 문헌 내의 잠재된 토픽을 추정하는 것으로써 어떤 단어들을 어떠한 토픽에서 선택하여 추출할 것인지에 대해 각각의 파라미터(parameter: α , β)를 이용하여 모델링하는 기법이다. LDA는 디리클레 확률 분포(Dirichlet Probability Distribution)를 통해 문서의 잠재적인 확률을 검토한 후 깁스 샘플링(Gibbs sampling) 알고리즘을 통해 해당 문서의 토픽과 단어에 대한 확률값을 추정한다. 또한, LDA는 어떤 토픽과 단어에 대한 사전 정보가 없는 비지도 생성 모델(Unsupervised Generative Model)이기 때문에 각 토픽과 관련한 모든 단어를 찾을 수 있다는 장점이 있다(김진국 외, 2022).

김현정 외(2015)는 항공산업의 전반적인 분야에 빅데이터 분석 방법인 텍스트 마이닝 기법을 적용하여 항공 분야의 연구 동향과 미래 유망분야를 전망하였다. 그 결과, 항공안전정책과 항공운임(저가항공), 친환경 고연비 연료가 미래 유망분야로 도출하고, 텍스트 마이닝 기반의 트렌드 분석의 활용 확대 가능성을 제시하였다.

최동현 외(2022)는 COVID-19가 항공산업에 미친 영향과 동향을 살펴보기 위해 다양한 국내 뉴스 기사들의 항공사에 관한 키워드를 중심으로 트렌드 분석을 실시하여 COVID-19 발생 전과 후의 항공산업 주요 토픽이 특정 국가의 문제로 인한 영향에서 항공여객과 항공화물 산업에 대한 문제, 그리고 여행 및 관광 관련 토픽으로 변화하고 있음을 확인하였다.

이처럼 방대한 비정형 자료를 대상으로 텍스트 마이닝 기법을 이용한 분석 결과가 의미있게 활용될 수 있음을 선행연구 결과들을 통해 확인할 수 있었다.

III. 연구 방법 및 내용

3.1 연구 모형

본 연구에서는 ‘항공안전’ 및 ‘항공사고’와 관련된

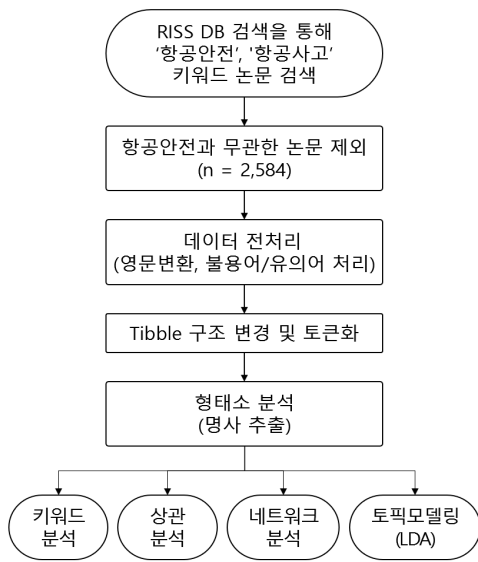


Fig. 2. Study flow chart

논문을 대상으로 항공안전관리 분야에서 이슈가 되고 있는 토픽을 살펴보기 위해 Fig. 2와 같은 순서로 연구를 진행하였다.

3.2 데이터 수집과 전처리

본 연구에서는 최근 항공안전관리 이슈를 분석하기 위해 항공안전과 관련된 논문들을 수집하였다. 한국교육학술정보원에서 운영 중인 학술연구정보서비스(RISS, Research Information Sharing Service)의 데이터베이스에서 '항공안전'과 '항공사고'를 키워드로 사용하는 국내 논문을 대상으로 하였다. 웹 크롤링을 통해 논문 제목, 저자, 초록, 키워드, 게재일, 학회지명 정보를 수집하였으며, 그 중 '항공안전'과 전혀 무관하거나 영문으로 작성된 논문을 제외하고, 총 2,584편의 논문을 분석 대상으로 선정하였다.

총 2,584편의 논문 제목, 초록, 키워드를 하나의 파일로 문서화 한 후 텍스트 마이닝 분석 기법을 적용하기 위해 데이터 전처리 과정을 가장 먼저 진행하였다. 일부 영문으로 작성된 단어는 한글로 변경하여 분석 과정에서 누락 되는 단어가 없도록 수정하였으며, 기본적인 연구 논문 주제로 사용되는 '연구', '분석', '고찰'과 같은 단어와 '관한', '위한', '대한' 등과 같은 조사는 불용어 처리하였다. '항공 안전관리'와 '항공안전관리'와 같이 띄어쓰기로 인해 구분되는 명사는 유의어로 판단하고 동일한 명사로 통일하였다(Table 1).

Table 1. Data preprocessing

전처리항목	내 용
영문 변환	<ul style="list-style-type: none"> • SMS → 항공안전관리시스템 • UAM → 도심항공모빌리티 • UAV → 무인항공기 • Checklist → 체크리스트
불용어	<ul style="list-style-type: none"> • 명사 : 논문, 연구, 분석, 고찰 등 • 조사 : 관한, 위한, 대한, 미치는, 이용한, 통한, 활용한, 위해, 관해, 대해, 하지, 하기, 이를 등 • 학회명, 학회지명
유의어	<ul style="list-style-type: none"> • 항공 안전관리, 항공안전관리 • 헬리콥터, 헬기 • 운항 승무원, 운항승무원

특히, 항공 용어 특성상 영문으로 작성된 단어와 부품, 장비에 관한 영문 명칭이 많았으며 이에 대한 의미가 왜곡되지 않도록 단어를 직접 검색 및 해석하여 수정하는 과정을 진행하였다. 또한, 영문을 번역하여 설명한 단어가 중복되어 표기되지 않도록 전처리하였으며, 한문으로 표기된 단어도 변환하는 과정을 거쳤다.

3.3 분석 방법

전처리 과정이 완료된 데이터는 Tibble 구조로 변경하고 토큰(Token)으로 구분하였다. 형태소(Morphem) 분석을 통해 명사로 판별된 단어를 추출하였으며, 한국어 데이터 전처리 과정에 기본적으로 활용되는 KoNLP(Korean Natural Language Processing) 패키지과 세종 한국어 사전을 사용하였다.

명사를 기준으로 토큰화된 결과에 기반하여 연도별 항공안전 분야 논문 실적을 확인하고, 키워드 분석과 단어 간 상관분석, 네트워크 분석을 실시하였다. 또한, 토픽 모델링의 분석기법 중 하나인 LDA를 적용하여 항공안전 분야의 주요 토픽(이슈)을 추출하였다. LDA 모델을 구축하는 과정에서 적정 토픽 수를 결정하기 위해 하이퍼파라미터 튜닝을 진행하였다. 토픽 수를 2에서 20까지 바꾸어 가며 모델을 만든 다음 성능 지표를 보고 최종 모델을 선정하였다. Fig. 3에서 확인되는 바와 같이 토픽 수가 12개일 때까지는 성능 지표가 점진적으로 증가하다가 그 이후는 변동이 발생한다. 이에 최종 토픽 수를 12개로 결정하고 모델을 선정하였다.

위와 같은 연구 방법은 통계프로그램 R 4.3.0 버전을 사용하여 분석을 수행하였다.

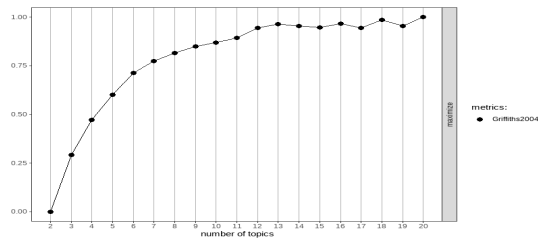


Fig. 3. Hyper-parameter tuning (topic number : 12)

IV. 연구 결과

4.1 연도별 항공안전 분야 연구 추이

‘항공안전’을 키워드로 하는 국내 논문은 1958년에 처음 발표되었고, 2006년까지 두 자리 수를 유지하다가 2007년부터는 Fig. 4와 같이 연도별 100건이 넘어 서면서 급격히 증가하는 추세를 보이고 있다. 특히, 2015년부터는 연간 약 300여 건 이상의 논문이 발표되는 등 해마다 ‘항공안전’을 키워드로 발표되는 연구 논문이 꾸준히 증가하고 있음을 확인할 수 있고, 2021년에는 총 363건의 논문이 발표되며, 역대 가장 많은 항공안전에 관한 연구 논문이 발표되었다.

4.2 키워드 빈도 분석

분석 대상 논문에서 가장 많이 사용된 단어를 추출한 결과는 Fig. 5에서 볼 수 있듯이 ‘항공기’가 1,568 회로 압도적으로 많이 언급된 것을 알 수 있다. 뒤를 이어 ‘항공’이 593회, ‘드론’이 582회, ‘안전’이 544회 언급되었다. 그 중 ‘드론’과 ‘무인항공기’의 경우 도합 1,050회 언급되며 항공안전분야에서 무인기가 새로운 이슈로 주목받고 있음을 확인할 수 있다.

4.3 단어 간 상관분석 및 네트워크 분석

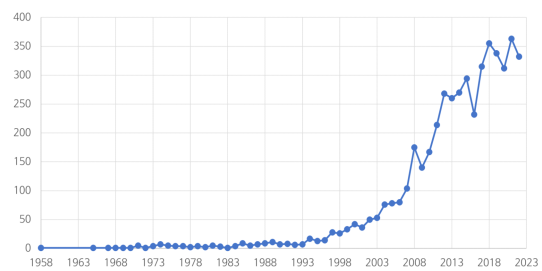


Fig. 4. Paper numbers in aviation safety by year

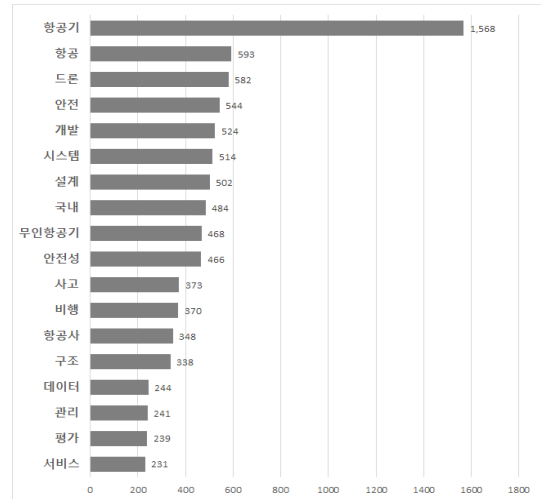


Fig. 5. Keyword frequency (top 20)

키워드 분석 결과로 나타난 상위 4개 단어를 대상으로 단어 간 상관분석(Phi Coefficient)과 네트워크 분석을 실시하였다. 파이 계수는 -1에서 +1 사이의 값을 가지며, +1에 가까울수록 두 단어의 상관성이 높은 것으로 해석할 수 있으며, 상대적으로 관련성이 큰 단어를 확인할 수 있다. ‘항공기’, ‘항공’, ‘드론’, ‘안전’ 등 4개 키워드와 관련성이 큰 단어를 분석한 결과는 Fig. 6과 같다.

‘항공기’의 경우, ‘조종사’(0.17)와 가장 높은 상관성을 보이고 있으며, ‘항공’은 ‘철도’(0.30), ‘드론’은 ‘사생활’(0.37) 및 ‘항공안전법’(0.35) 및 ‘안전’은 ‘승객’(0.18)과 상대적으로 높은 상관관계를 나타내었다.

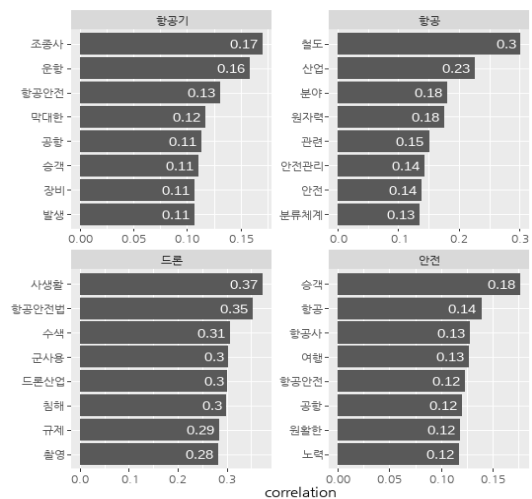
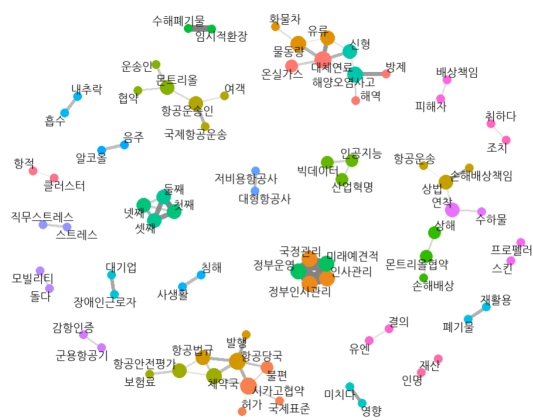


Fig. 6. Phi coefficient by keyword

Fig. 7. Word network (Phi Coefficient ≥ 0.5)

파이 계수를 이용하여 다른 단어에 비해 상대적으로 자주 함께 사용된 단어쌍을 네트워크 그래프로 도식한 결과, Fig. 7과 같이 나타났다. 파이 계수를 이용한 그래프는 관련성이 큰 단어끼리만 연결되어 있으므로 단어 군집이 명확하게 시현된다. ‘저비용항공사’는 ‘대형항공사’와 밀접하게 관련되어 사용되고 있으며, ‘감항인증’과 밀접한 관련이 있는 단어는 ‘군용항공기’로 나타났다. ‘항공법규’의 경우 ‘항공당국’과 ‘발행’과 동일한 커뮤니티를 이루었고, ‘계약국’ 커뮤니티와 ‘시카고협약’ 커뮤니티와의 관련성을 나타내었다.

4.4 LDA 분석

‘항공안전’과 ‘항공사고’를 키워드로 사용한 논문들은 어떤 핵심 단어를 내포하고 있으며, 어떤 주제를 논의한 것인지 확인하기 위하여 토픽 모델링을 실시하였는데, 가장 널리 사용되는 토픽 모델링 알고리즘인 LDA 모델을 사용하였다. 하이퍼파라미터 튜닝 결과에 의해 12개의 토픽이 선정되었고, LDA 모델 결과는 Table 2, Fig. 8과 같이 도출되었다.

토픽 1에서 베타값이 가장 높은 단어는 ‘수행’, ‘해석’으로써 항공기 구조와 설계, 피로수명 해석 등이 항공안전 및 항공사고와 관련된 이슈임을 확인할 수 있다.

토픽 2에서 베타값이 가장 높은 단어는 ‘사회’, ‘국가’, ‘대응’ 등이었으며, 해당 논문 제목을 확인한 결과, 항공테러와 테러방지, 보안검색이 안전관리의 중요 이슈로 여겨지고 있었다.

토픽 3에서 가장 큰 베타값을 가지는 단어는 ‘대상’, ‘영향’, ‘요인’, ‘서비스’ 등으로 나타났으며, 해당 논문 제목 및 초록을 통해 저비용 항공사, 서비스 품질, 객실승무원 등의 토픽이 항공안전 및 항공사고와 관련된 이슈임을 알 수 있었다.

토픽 4에서 가장 큰 베타값을 가지는 단어는 ‘활용’, ‘다양’, ‘분야’ 등으로 드론 및 무인기, 도심항공모빌리티와 관련된 주제가 항공안전 및 항공사고 예방에 중

Table 2. Result of LDA topic modeling

Topic	1 st		2 nd		3 rd		4 th		5 th		6 th	
	Keyword	Beta	Keyword	Beta	Keyword	Beta	Keyword	Beta	Keyword	Beta	Keyword	Beta
1	해석	0.024	수행	0.024	구조	0.018	조건	0.017	적용	0.016	설계	0.016
2	사회	0.014	국가	0.013	활동	0.011	이후	0.011	대응	0.009	세계	0.008
3	대상	0.019	영향	0.015	요인	0.015	서비스	0.014	파악	0.012	실시	0.011
4	활용	0.031	다양	0.029	분야	0.027	필요	0.026	기술	0.024	문제	0.020
5	발생	0.055	상황	0.024	위험	0.015	사고	0.013	때문	0.013	이상	0.013
6	정보	0.022	기반	0.021	이용	0.021	제안	0.019	가능	0.017	사용	0.015
7	규정	0.013	우리나라	0.012	문제	0.011	보호	0.009	국제	0.009	책임	0.009
8	산업	0.014	경제	0.011	변화	0.010	세계	0.010	방향	0.010	추진	0.009
9	비행	0.029	시스템	0.027	검증	0.019	시험	0.018	지상	0.018	성능	0.017
10	개발	0.053	요구	0.041	기술	0.024	적용	0.021	비행	0.019	시스템	0.017
11	중요	0.037	고려	0.026	방법	0.022	가지	0.016	일반	0.015	요소	0.015
12	국내	0.031	체계	0.021	방안	0.020	관리	0.020	사례	0.017	운영	0.016

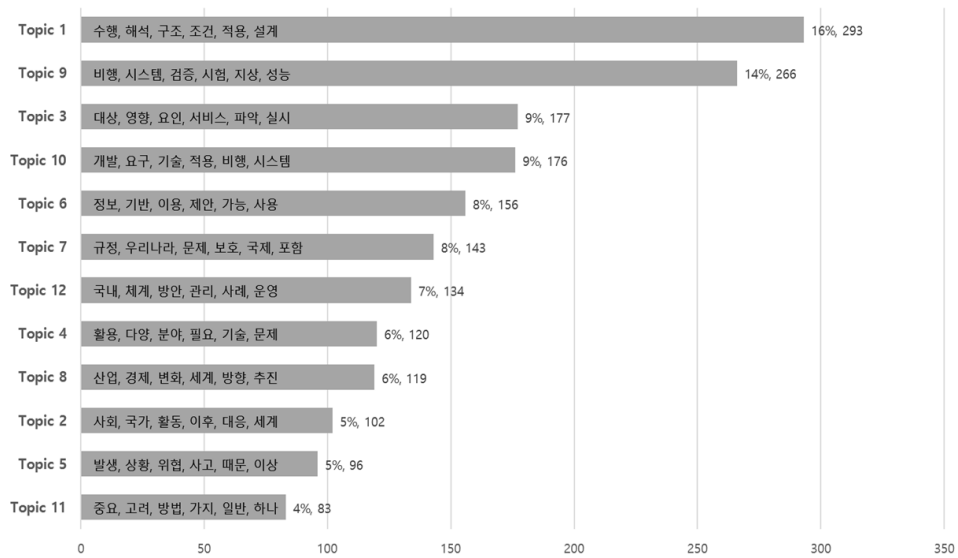


Fig. 8. Frequency of keywords and documents by topic

요하게 고려되고 있는 것으로 파악되었다.

토픽 5에서는 ‘발생’, ‘상황’, ‘위험’, ‘사고’ 등이 베타값이 높은 단어로 나타났으며 조종사의 인지능력과 인적오류, 재난 대응 등에 관한 주제를 다루어 항공안전 및 항공사고 예방에 중요하게 고려하고 있는 이슈임을 확인하였다.

토픽 6에서는 ‘정보’, ‘기반’, ‘이용’ 등이 베타값이 높았으며 항공사진 측량과 원격탐사, 지도제작을 위해 항공사진을 사용함에 있어 항공안전 및 항공사고가 고려되고 있음을 확인할 수 있었다.

토픽 7에서는 ‘규정’, ‘우리나라’, ‘문제’, ‘보호’, ‘국제’ 등의 단어가 베타값이 높게 나타났다. 해당 논문 주제를 살펴본 결과, 항공법규에 대한 고찰과 불법행위에 따른 법적 판례, 국제협약 등에서 항공안전 및 항공사고와 관련된 이슈들을 다루고 있었다.

토픽 8에서는 ‘산업’, ‘경제’, ‘변화’ 등이 베타값이 높았으며 항공정책 활성화 및 4차산업혁명에 따른 항공안전 및 항공사고의 이슈가 논의되고 있음을 알 수 있었다.

토픽 9의 경우 ‘비행’, ‘시스템’, ‘검증’, ‘시험’, ‘성능’ 등이 나타나 우주발사체의 안전, 비행안전분석, 비행시험의 이슈가 항공안전과 관련되어 진행되고 있다는 것이 파악되었다.

토픽 10의 경우 ‘개발’, ‘요구’, ‘기술’ 등의 단어가 베타값이 높았고 감항인증과 품질인증, 형식증명, 안정성 평가 관련 이슈가 항공안전 및 항공사고와 관련되

어 있음을 확인하였다.

토픽 11에서 베타값이 높은 단어는 ‘중요’, ‘고려’, ‘방법’ 등으로 항공안전을 평가하는 기준과 기술이 중요한 이슈로 언급되고 있었다.

마지막으로 토픽 12에서 높은 베타값을 가진 단어는 ‘국내’, ‘체계’, ‘방안’, ‘관리’ 등으로 국제표준과 국내도입방안, 법적고찰, 체계개선 등이 항공안전에 중요한 이슈로 논의되고 있음을 확인할 수 있었다.

V. 논의 및 결론

본 연구는 성장과 변화가 가속화되고 있는 항공산업 분야의 안전관리 이슈들을 파악하기 위해서 ‘항공안전’ 및 ‘항공사고’를 키워드로 사용하는 국내 논문 2,584편을 대상으로 키워드 빈도 분석, 단어 간 상관 분석 및 네트워크 분석, 토픽 모델링 등 텍스트 마이닝 기법을 활용한 다양한 분석을 실시하였다.

항공안전을 키워드로 사용하는 연구는 2015년 이후 급격히 증가하여 연 300건 이상의 논문이 발표되고 있고, 이 논문들의 키워드를 빈도 분석한 결과, ‘항공기’ 빈도수가 가장 높았으며(1,568회), 다음으로는 ‘드론’ 및 ‘무인항공기’ 빈도수가 높았다(10,50회). 키워드 분석 빈도가 높았던 4개 단어(항공기, 항공, 드론, 안전)를 대상으로 단어 간 상관분석을 실시한 결과, ‘항공기’는 ‘조종사’와, ‘항공’은 ‘철도’, ‘드론’은 ‘사생활’, ‘안전’은 ‘승객’과 높은 상관성을 가지는 단어로 확인되었다.

본 연구를 통해 수행한 토픽 모델링 결과, 식별한 항공안전관리에 관한 이슈는 총 12가지로 아래와 같이 요약할 수 있다.

- 이슈 1 : 항공기 구조와 설계에 관한 사항
- 이슈 2 : 항공테러 및 보안에 관한 사항
- 이슈 3 : 저비용항공사 및 항공서비스 관련 사항
- 이슈 4 : 드론 및 무인항공기에 관한 사항
- 이슈 5 : 조종사의 인지능력과 인지오류 문제
- 이슈 6 : 항공사진 측량 및 공간정보에 관한 사항
- 이슈 7 : 항공법규 및 불법행위에 관한 사항
- 이슈 8 : 항공정책 및 4차 산업혁명에 관한 사항
- 이슈 9 : 비행안전분석 및 시험비행에 관한 사항
- 이슈 10 : 감항인증 및 형식증명에 관한 사항
- 이슈 11 : 항공안전 평가기준 및 기술 관련 사항
- 이슈 12 : 국제표준 도입방안에 관한 사항

드론, 도심항공모빌리티 등 새로운 항공 운용 분야의 부상과 COVID-19 이후 항공산업의 뚜렷한 회복세에 따라 항공안전관리의 범위는 더욱 커질 것으로 전망된다.

본 연구를 통해 확인한 항공안전에 관한 이슈들은 향후 항공분야에서 안전관리에 대한 선택과 집중이 필요한 분야로 판단되며, 특히, 드론 및 무인항공기 분야 이슈들은 기존 항공안전체계에 새로운 패러다임으로 부상하고 있다는 점을 주목할 필요가 있다고 생각한다.

본 연구는 텍스트 마이닝 기법을 항공안전 이슈 분석에 활용하였다. 하지만, 데이터 전처리 과정에 일부 한계점이 존재하였다. 항공분야 특성상 영어로 된 단어 및 용어가 다수 존재하지만 이를 한글로 대체하기 위해 참조할 수 있는 항공용어 정의 자료가 부족했다는 점이다. 텍스트 마이닝의 전처리 과정은 결과의 신뢰도에 영향을 미치기 때문에 향후 전문가 집단과의 토의 등과 같은 수단을 통해 충분한 국문 항공용어의 정의 자료를 확보하고 연구를 수행한다면 좀 더욱 구체적인 항공안전 이슈들을 확인할 수 있을 것으로 예상된다.

후 기

본 논문은 2023년도 한국항공운항학회 춘계학술대회 발표논문을 발전·적용하였습니다.

References

1. Park, W. H., "Aviation safety : International

and domestic legal action", Jeju Peace Institute, 2009(16), 2009, pp.2.

2. ICAO Annex 19, "Safety Management second Edition", International Civil Aviation Organization, Montreal, Canada, 2016.
3. So, J. S., and Lee, C. K., "A legal study on safety management system", Korea Society of Air & Space Law and Policy, 29(1), 2014, pp.3-32.
4. MOLIT, Aviation Safety Law (No. 18789), Sejong: Ministry of Land, Infrastructure, and Transportation, 2021.
5. Lee, S. I., "Aviation Industry Status and Development Direction", Korea Development Bank, KDB Monthly, No. 802, 2022, pp.32-63.
6. ICAO Doc 9859, "Safety Management Manual (SMM) Fourth Edition", International Civil Aviation Organization, Montreal, Canada, 2018.
7. Kim, J. G., Yang, C. H., and Park, S. B., "A study on the research trends for tunnel traffic safety using text mining", Journal of Digital Contents Society, 23(10), 2022, pp. 2075-2083.
8. Koteeswaran, S., Malarvizhi, N., Kannan, E., Sasikala, S., and Geetha, S., "Data mining application on aviation accident data for predicting topmost causes for accidents", Cluster Computing, 22(5), 2019, pp.11379-11399.
9. Anderson, C., and Smith, M. O., "Qualitative analysis of loss of control aircraft accidents using text mining techniques", International Journal of Aviation, Aeronautics, and Aerospace, 4(4), 2017, pp.1-26.
10. Kim, H. J., Jo, N. O., and Shin, K. S., "Text mining-based emerging trend analysis for the aviation industry", Korea Intelligent Information Systems Society, 21(1), 2015, pp.65-82.
11. Choi, D. H., Song, B. M., Park, D. H., and Lee, S. W., "Keyword trends analysis related to the aviation industry during the co-

- vid-19 period using text mining”, Journal of the Korea Industrial Information Systems Research, 27(2), 2022, pp.115-128.
12. Kim, J. H., Lim, J. J., and Lee, J. R., “A study on the analysis of aviation safety data structure and standard classification”, Journal of the Korean Society for Aviation and Aeronautics, 28(4), 2020, pp.89-101.
 13. Kim, J. Y., Na, H. S., and Park, K. H., “Topic modeling of profit adjustment research trend in Korean accounting”, Journal of Digital Convergence, 19(1), 2021, pp.125-139.