

코로나19 팬데믹 이후 컨테이너선 운임 상승 요인분석: 텍스트 분석을 중심으로⁺

(Analysis of Factors Affecting Surge in Container Shipping Rates in the Era of Covid19 Using Text Analysis)

나 진 성¹⁾
(Jin Sung Rha)

요약 코로나19 팬데믹 상황에서 컨테이너선 운임은 유례없는 큰 폭의 상승세를 보이고 있다. 컨테이너선 운임 상승 요인에 대해서 다양한 분석이 이루어지고 있으나, 비정형 데이터인 텍스트 자료를 활용한 분석은 전무한 상황이다. 따라서 본 연구에서는 관련 기사들을 대상으로 최근의 컨테이너선 운임 상승의 요인들을 텍스트 마이닝 기법중 하나인 네트워크 텍스트 분석과 LDA 토픽 모델링을 통해 파악하였다. 2020년 1월부터 2021년 7월까지 로이즈리스트에 게재된 기사들을 대상으로 텍스트 분석을 하였다. 분석 결과, 중국과 미국의 무역마찰, 글로벌 생산감소를 예측한 글로벌 선사들의 급격한 기항 횡수의 감소와 임시결항의 증가, 터미널 혼잡, 수에즈 운하 봉쇄와 같은 예기치 못한 사고들이 주요 원인으로 분석되었다.

핵심주제어: 컨테이너선 운임, 코로나19, 네트워크 텍스트 분석, 토픽 모델링

Abstract In the era of the Covid19, container shipping rates are surging up. Many studies have attempted to investigate the factors affecting a surge in container shipping rates. However, there is limited literature using text mining techniques for analyzing the underlying causes of the surge. This study aims to identify the factors behind the unprecedented surge in shipping rates using network text analysis and LDA topic modeling. For the analysis, we collected the data and keywords from articles in Lloyd's List during past two years(2020-2021). The results of the text analysis showed that the current surge is mainly due to "US-China trade war", "rising blanking sailings", "port congestion", "container shortage", and "unexpected events such as the Suez canal blockage".

Keywords: container shipping rates, covid19, network text analysis, topic modeling

* Corresponding Author: jinsungrha@dankook.ac.kr

+ 본 연구는 한국해양수산개발원이 2021년에 지원한 "수에즈 운하 등 주요 선박 통항로의 선박 좌초 및 폐쇄에 따른 해운 운임 변동 예측" 연구과제의 분석결과를 토대로 작성되었음.

Manuscript received January 17, 2022 / revised January 27, 2022 / accepted January 28, 2022

1) 단국대학교 경영학부, 교신저자, 제1저자

1. 서론

2020년 초부터 본격화한 코로나19 판데믹으로 인해서 컨테이너선 운임은 2021년 9월을 기준으로 사상 최고치를 경신하고 있다. 컨테이너선 운임은 지난 10여년동안 약간의 변동성은 있었으나, 큰 폭의 상승과 하락은 없었다. 그러나 판데믹 상황에서 컨테이너선 운임은 지속적인 상승세를 보이고 있다. 컨테이너선 운임의 대표지표인 상하이 컨테이너선 운임은 2020년 1월에 985.93포인트였으나, 2021년 9월에는 4,590.24포인트까지 상승하였다(Shanghai Shipping Exchange, 2021). 글로벌 공급망의 불확실성으로 인해 촉발된 운임상승은 공컨테이너 부족, 항만혼잡도 악화, 재난사고, 코로나로 인한 항만 폐쇄 등의 파급으로 인하여 상승세가 지속되고 있다(Ahn, 2021).

유례가 없는 컨테이너선 운임 상승 요인에 대해서 다양한 분석이 이루어지고 있으나, 비정형 데이터인 텍스트 자료를 활용한 분석은 전무한 상황이다. 비정형 텍스트 데이터 분석의 경우 기존의 정형 데이터에서 분석하기 어려운 다양한 이벤트들과 동향을 파악할 수 있으며 향후 머신러닝, 딥러닝 등과 같은 다양한 데이터 마이닝 분석에 활용될 수 있다. 따라서 본 연구에서는 컨테이너선과 관련한 기사들을 대상으로 최근의 컨테이너선 운임 상승의 요인들을 텍스트 마이닝 기법 중 하나인 네트워크 텍스트 분석과 토픽 모델링을 통해 파악하고자 한다.

다음 장에서는 코로나 판데믹 이후 에버기븐(Ever Given) 호의 좌초로 인한 수에즈 운하 봉쇄, 근로자의 코로나 감염으로 인한 항만폐쇄 등 컨테이너선과 관련한 주요 이슈들을 종합 정리한다. 3장에서는 본 연구에서 활용한 네트워크 텍스트 분석과 토픽 모델링 방법에 대해서 설명한다. 4장에서는 비정형 데이터인 기사 텍스트 자료를 활용하여 최근의 컨테이너선 운임 상승세와 관련한 주요 요인들과 배경을 파악한다. 네트워크 텍스트 분석 결과를 제시하고 토픽 모델링을 활용하여 토픽그룹별 주요 기사 내용을 분석한다. 컨테이너선 산업 분야의 동향을 파악하기 위해서 본 연구에서는 컨테이너선 운

임 상승의 변곡점들을 기준으로 2020년도 1월부터 2021년도 7월까지의 기간을 총 6개의 시기로 구분하여 분석을 시행하였다. 마지막으로 5장에서는 연구결과를 요약하고 연구의 시사점들과 향후 연구의 방향을 제시한다.

2. 코로나19와 컨테이너 운임 상승

코로나19 판데믹이 본격화한 2020년 상반기부터 컨테이너선 산업은 전방위적으로 코로나19의 영향을 받았다. 이로 인해 2020년도 상반기부터 공급사슬의 단절현상이 발생하였으며(Ahn, 2021), 이는 중국과 미국의 무역마찰과 코로나로 인한 중국 공장들의 생산감소를 예측한 글로벌 선사들의 급격한 기항 횟수의 감소와 임시결항의 증가로 이어지면서 글로벌 공급사슬의 불확실성은 더욱 심화되었다(Dirzka and Acciaro, 2021). 싱가포르와 같이 항만운영의 지능화를 통해 공급사슬의 유연성과 탄력성을 어느정도 확보한 경우에는 비교적 안정적으로 터미널을 운영할 수 있었으나, 대부분의 글로벌 터미널의 경우 혼잡상황을 피하지 못하고 컨테이너 처리율이 급감하게 되었다(Lam et al., 2021). 또한 미국을 중심으로 2020년도 하반기부터 이른바 보복 소비가 시작되면서 물동량에 대한 수요는 급증하게 되고 이는 컨테이너의 부족과 맞물리면서 해운운임의 가파른 상승을 이끌게 된다(Cullinane and Haralambides, 2021).

2021년 3월말 발생한 에버기븐 선박 좌초사고는 전세계 물동량의 30% 이상이 움직이는 수에즈 운하를 봉쇄하면서 배송지연으로 하루에 수십억 달러의 손실을 발생시키고 500여척의 달하는 선박들의 적체를 야기하였다(Lee and Wong, 2021). 이로 인한 컨테이너 박스의 부족과 터미널의 혼잡, 항만 근로자들의 근무여건 악화 등으로 해운운임은 사상 최고치를 경신하게 된다(Ramos et al., 2021). 수에즈 운하 봉쇄 외에도 예기치 않은 사건과 사고들로 인해 컨테이너선 운임은 가파른 상승세를 지속하였다. 중국의 옌텐항의 항만과 미국 서부 LA항만의 근로자 집단감염 사례는 터미널의 부분 폐쇄나 처리율의

급감으로 이어지면서 컨테이너의 적체현상을 심화시키게 된다(Song, 2021). 터미널의 용량부족 상황은 항만 근로자의 근무 여건도 악화시켜서 항만 운영관리에도 상당한 부담을 주고 있으며, 내륙 유통망의 지연도 물류대란에 영향을 주고 있었다(Cullinane and Haralambides, 2021). Fig. 1과 같이 이러한 이벤트들로 인해 상하이 운임지수는 2020년부터 2021년 상반기까지 가파른 상승세를 지속하였다.

컨테이너선 운임이 지속적인 상승세를 보이는 가운데 글로벌 선사들의 실적은 역대 최고치를 기록하고 있으며, 주가 역시 가파르게 상승하고 있었다. 화주들은 선사들과 계약을 맺을 시 컨테이너 부족으로 장기계약을 선호하는 추세이고 선사들은 장기고정계약과 스팟(Spot) 물량 조절에 대해서 전략적으로 접근하고 있다. 글로벌 선사들은 빠른 시일 내 컨테이너선 운임이 안정세로 접어들기는 어려울 것으로 전망하였다(Lam et al, 2021).



Fig. 1 Shanghai Containerized Freight Index (Shanghai Shipping Exchange, 2021)

3. 연구방법

3.1 네트워크 텍스트 분석

네트워크 텍스트 분석은 텍스트 마이닝과 네트워크 분석을 결합한 방법론으로 분석하고자 하는 텍스트 전문을 구성 단어들로 해체하고, 이들을 노드로 하여 다양한 네트워크를 구축하여 분석하는 방법이다. 네트워크 텍스트 분석을 통해서 네트워크 상의 단어 노드들의 위치와 연

결패턴을 통해서 주요 주제들을 유추하고, 텍스트의 의미를 이해할 수 있다(Rha, 2020). 텍스트 내의 단어 간 의미론적 연관성을 분석하는데 주력한다. 기존의 주제 동향 분석 연구들이 사전에 분류한 카테고리를 기반으로 하여 텍스트들을 구분하여 관련한 주제 동향을 분석하는데 중점을 두었다면, 네트워크 텍스트 분석은 여러 개의 텍스트의 분석을 통해 주제들 간 상호 관련성, 구조적 관계 및 중심성(Centrality), 클러스터 등을 네트워크 맵(Network Map)을 통해서 확인한다. 키워드에 기반한 네트워크 텍스트 분석을 통해서 텍스트들의 핵심 주제들과 하위 주제들을 파악할 수 있다(Pyun and Rha, 2021). 네트워크 맵을 도식화 하기 위해서 단어간 공출현(Co-Occurrence) 빈도수에 기반한 인접행렬을 구축하도록 한다.

본 연구에서 네트워크 텍스트 분석은 다음의 절차를 따른다. 첫째, 분석 대상 문헌자료를 선정한다. 문헌분석 단위는 제목, 초록, 본문 모두가 가능하다. 본 연구에서는 컨테이너선의 운임변동과 관련한 기사 자료들을 분석하기 위해 2020년도 1월부터 2021년도 7월까지 로이즈리스트(Lloyd's List)에 게재된 기사들을 검색하였으며 기사본문을 분석대상으로 하였다.

둘째, 수집된 기사들로부터 기사-단어의 이원 모드(Two Mode) 네트워크를 구축하기 위해 단어들 추출하였다. 추출 단어의 품사는 명사로 분석의 명확성을 위해서 한정하였다. 텍스트의 모든 단어들 추출할 수 없기 때문에 분석에 유의미한 단어들 추출하기 위한 사전을 구축하였다. 사전에는 분석에서 제외하기 위한 '제외어', 동일한 의미를 지닌 단어를 하나의 단어로 통일하는 '유의어', 복합명사인 단어를 구분하기 위한 '복합어'를 포함하였다. 제외어는 'Month', 'Time', 'Year', 'Lloyd's List Intelligence' 등과 같이 일반적으로 기사에서 많이 출현하는 단어로써 분석에서 특이점을 찾기 어려운 단어들 지정하였다. 제외어를 선정하기 위해서 Term Frequency-Inverse Document Frequency (TF-IDF) 분석을 시행하였다. TF-IDF 분석을 통해서 특정 문서에서 단어들의 중요성을 계산할 수 있다. $TF(d, t)$ 는 특정 문서 d 에서 특정

단어 t가 얼마나 자주 등장하는지를 계산한 값이고, IDF(t)는 특정 단어 t가 등장한 문서수를 의미하는 DF(t)에 반비례하는 수로 계산한 값이며, 이는 단순 출현 빈도가 높은 단어들의 영향력을 분석에서 제거하기 위함이다(Salton and Buckley, 1988). TF-IDF는 TF와 IDF의 곱으로 정의할 수 있으며, 값이 작아질수록 불용성의 단어일 가능성이 높다(Aizawa, 2003). 본 연구에서는 다수의 문서에서 TF-IDF 값이 0.5 이하인 단어들을 불용성의 단어로 판단하여 제외어 사전에 포함하였다. 유의어는 전문가들의 자문을 통해 구성하였다. 예를 들어 ‘Equipment Shortage’, ‘Box Shortage’, ‘Container Shortage’ 등은 유사한 의미를 지니고 있기 때문에 모두 ‘Container Shortage’로 통일하였다. 복합어를 구성하기 위해서는 단어 간 근접하여 출현하였는지 여부를 확인하였다. 예를 들어 ‘Supply’와 ‘Chain’은 전체 기사에서 200번 이상 바로 근접하여 출현하였다. 이와 같은 경우에 단어 간 근접성이 상당히 높다고 판단하여 단어 추출 시에 ‘Supply Chain’의 복합어로 추출되도록 하였다.

셋째, 추출한 단어들을 활용하여 기사-단어 이원모드 네트워크를 구축하였다. 네트워크는 이원모드이기 때문에 기사노드와 단어노드로 구성되어있다. 기사-단어 네트워크에서는 직접적으로 개별 단어의 영향력을 분석할 수 없다. 본 연구에서는 분석의 정확성을 위해서 개별 단어가 2편 이상의 기사에 출현한 경우에만 이원모드 네트워크에서 링크가 형성되도록 하였다. 기사-단어 네트워크에서는 개별 단어들이 기사에 출현한 빈도수를 확인할 수 있다.

넷째, 기사-단어 이원모드 네트워크를 단어-단어 일원모드(One Mode) 네트워크로 변환하였다. 일원모드 네트워크로의 변환은 단어 간 인접 관계를 기준으로 하였다. 텍스트에서 단어가 서로 근거리에서 있는 경우 공출현의 관계에 있고 상호 연관성이 높다고 할 수 있다. 본 연구에서는 단어 간 인접거리(Word Distance) 기준을 ‘4’로 하여 공출현 여부를 판단하고 공출현 횟수가 2이상인 경우에 네트워크에서 단어노드 간 링크를 형성하였다.

다섯째, 단어-단어 일원모드 네트워크에서 각

단어의 영향력을 분석하기 위해서 중심성을 파악하였다. 중심성의 종류는 연결중심성, 매개중심성, 위세중심성, 근접중심성 등이 있다. 본 연구에서는 연결중심성(Degree Centrality)을 분석하였다. 연결중심성은 하나의 개별 노드가 얼마나 많은 노드들과 링크로 연결되어있는지 파악하는 것이다. 특정 노드의 연결중심성은 노드가 다른 노드들과 연결된 링크 수를 전체 노드들과 연결 가능한 링크 수로 나누어주는 것이다. 식은 다음과 같다.

$$Deree Centrality = \frac{\sum Weight of Links}{Number of Nodes - 1}$$

단어 네트워크에서 연결중심성이 높은 단어 노드들은 텍스트에서 가장 활발히 언급되는 주제와 관련이 깊은 단어로 볼 수 있으며, 다른 하위 주제들과도 밀접한 관련이 있다. 시기별 텍스트 단어들의 연결중심성 변화를 파악하면 해당 분야의 동향에 대해서 파악할 수 있다.

여섯째, 단어-단어 일원모드 네트워크의 구조적 특성을 기반으로 노드들 간 클러스터가 형성되는지 분석한다. 이는 네트워크의 응집성(Cohesion)을 분석하는 것이다. 본 연구에서는 일원모드 네트워크의 주요 컴포넌트(Giant Component)를 먼저 추출하고 이후에 커뮤니티 분석을 시행하여 네트워크에서 노드들 간 클러스터 구조를 파악한다.

3.2 토픽모델링

주요 토픽그룹들을 파악하기 위해서 토픽 모델링을 시행하였다. 본 연구에서 토픽 모델링을 위해 Latent Dirichlet Allocation(LDA) 기법을 활용하였다. LDA에서는 기사-단어 이원모드를 구축하기 위해 구성된 인접행렬을 활용한다. 기사-단어 행렬을 기사-토픽 그리고 토픽-단어의 행렬의 곱으로 가정하고 각 토픽에 단어와 기사가 속할 확률을 계산하게 된다. 순서는 기사-토픽 그리고 토픽-단어의 행렬의 곱을 먼저 가정하고 원본인 기사-단어 행렬을 얼마나 잘 구현할 수 있는지를 평가하게 된다. 따라서 LDA 토픽

픽모델링을 통해서 각 토픽그룹별로 주요 단어와 기사들을 파악할 수 있다.

LDA에서는 크게 3가지 변수를 지정해야한다. Fig. 2와 같이 α , β , k 값을 지정해주면 문서-토픽 비중 확률인 θ 와 토픽-단어 비중확률인 ϕ 를 계산하여 원본을 잘 재현하는지를 구하게 된다 (Blei et al., 2003). k 는 토픽의 수이며, α 와 β 는 모델링을 위한 하이퍼 파라미터이다. 선행연구에 따라 α 는 $50/k$ 로 가정하고 β 는 0.1로 지정하였다(Griffiths and Steyvers, 2004).

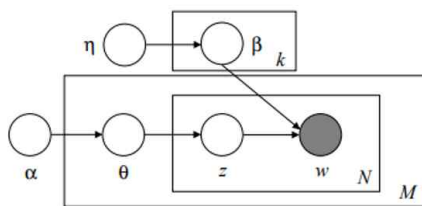


Fig. 2 LDA Topic Modeling (Blei et al., 2003)

4. 분석결과

본 연구에서는 최근의 컨테이너선 운임의 급격한 상승세를 분석하기 위해서 네트워크 텍스트 분석을 시행하였다. 네트워크 텍스트 분석을 통해서 컨테이너선 운임이 급상승하는 동안 있었던 주요 사건들과 요인들을 기사자료들로부터 파악하고자 한다. 분석 시기는 컨테이너선 운임 상승의 주요 변곡점을 기준으로 구분하였다.

본 연구에서 비정형 텍스트 자료로서 1734년에 창간한 세계적인 해운 전문지인 로이즈리스트를 활용하였다. 컨테이너선 관련 기사자료 중심으로 분석하기 위해서 로이즈리스트에서 2020년도 1월부터 2021년 7월까지 컨테이너선 카테고리 게재된 기사들을 수집하여 분석하였다. 분석을 위해서 NetMiner 4.0을 활용하였다.

4.1 2020년도 1월부터 6월까지의 기사분석

2020년도 1월부터 6월까지의 로이즈리스트에 게재된 컨테이너선 관련 기사 수는 총 187편이

다. 본 연구에서는 기사의 내용을 분석하기 위해서 기사로부터 단어들을 추출하여 기사-단어 이원모드 네트워크를 구성한다. 추출하는 단어 품사는 ‘명사’로 한정하였다. 분석의 정확성을 위해서 불용성의 단어들을 분석에서 제외할 필요가 있다. 다수의 기사에 빈출현하는 단어들은 의미가 관용적이거나 특별한 이슈를 설명하기에는 적합하지 않다. 불용성의 단어들을 파악하기 위해 IF-IDF 분석을 하여 0.5이하인 단어들은 분석에서 제외하였다. 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 7,463개이다.

4.1.1 네트워크 텍스트 분석

네트워크 텍스트 분석을 시행하기 위해서 기사-단어 이원모드 네트워크를 구성하였다. 기사와 단어 모드 사이에는 총 23,757개의 링크가 형성되었다. 이 기간 중에 가장 많이 출현한 단어는 ‘Coronavirus Pandemic’과 ‘China’이다. 2019년도 연말부터 코로나19 바이러스의 전파에 대한 우려가 커진 가운데, 2020년도 상반기에 코로나19 판데믹의 영향력이 지속되었음을 알 수 있다. 또한 이와 관련하여 중국의 항만, 공장 가동 상황이 컨테이너 선의 운임과 산업에 상당한 영향력을 주고 있음을 유추할 수 있다.

기사-단어의 이원모드 네트워크만으로는 직접적인 네트워크 텍스트 분석이 불가능하기 때문에 이원모드 네트워크를 단어-단어 일원모드 네트워크로 변환하였다. 변환 방법은 단어 간의 인접성을 기준으로 하였으며, 한 문장 내에서 단어 간의 거리가 ‘4’ 이하인 경우 단어 간의 공출현으로 인정하고, 2개 이상의 기사에서 공출현하는 경우에 해당 단어 노드 간의 링크가 형성되도록 하였다. 이와 같은 방법으로 단어-단어 일원모드 네트워크를 구성하였으며, 총 2425개의 링크가 형성되었다. 각 개별 단어 노드의 영향력을 파악하기 위해서 본 연구에서는 연결중심성을 계산하였다. 연결중심성은 개별 노드들이 다른 노드들과 얼마나 많이 링크로 연결되었는지를 상대적으로 계산한 지표값이다. 연결중심성이 높을수록 네트워크에서 활발히 다른 노드들에 영향을 준다. 단어-단어 네트워크에서 연결중심성이 높은 단어 노드들은 해당 시기에

중심적인 기사의 주제들과 밀접한 관련이 있는 단어로 파악할 수 있다. 연결중심성 분석결과 이 기간 중에 가장 영향력 있는 단어는 ‘Coronavirus Pandemic’, ‘China’, ‘Trade’, ‘Supply Chain’, ‘Us’, ‘Demand’, ‘Blank Sailings’, ‘Capacity’ 등 이었다.

기사 토픽들의 그룹을 파악하기 위해 네트워크 구조의 응집성을 파악하였다. 본 연구에서는 네트워크의 응집구조를 통해 적합한 토픽그룹의 수를 응집성 분석을 통해 도출하고, 세부적인 토픽 그룹과 그룹에 해당하는 기사에 대한 분석은 LDA 토픽 모델링을 통해 시행한다. 응집성 분석을 위해서 우선 단어 네트워크에서 중심이 되는 주요 컴퍼넌트를 파악하였다. 주요 컴퍼넌트에 해당하는 노드들만을 대상으로 커뮤니티 분석을 통해 클러스터 수를 파악하였다 (Girvan and Newman, 2002). 적합한 클러스터의 수는 NetMiner4에서 계산되는 Modularity Index 값을 참고하였으며, Modularity Index 값이 1.5 이상인 경우들을 고려하여 적합한 클러스터의 수를 정하였다. 이 기간의 단어 네트워크의 응집성을 분석한 결과 3개의 클러스터가 형성됨을 확인하여 토픽 모델링 분석에 활용하였다.

4.1.2 토픽 모델링 분석

토픽 모델링 분석을 한 결과, Table 1과 같이 토픽별로 단어들이 분포하고 있는 것으로 나타났다.

Topic1의 주요 단어는 ‘China’, ‘Coronavirus Pandemic’, ‘US’, ‘Trade’, ‘Effect’, ‘Blank Sailings’ 등 이었다. Topic1에 속하는 주요 기사들의 내용을 살펴보면, 코로나19가 글로벌 판데믹으로 확장하고 중국과 미국의 무역 갈등이 심화하면서 글로벌 선사들이 물동량 감소를 예측하여 기항 횟수를 줄이고 있으며, 중국을 비롯한 전 세계 공장의 가동이 중단되는 상황과 경기가 침체되는 상황에 대비하여 각국의 항만 당국은 통제를 시작하였음을 알 수 있다.

Table 1 Topics and Keyword (from 2020.1 to 2020.6)

Group	Keywords
Topic 1	China, Coronavirus Pandemic, US, Trade, Effect, Blank Sailings
Topic 2	Supply Chain, Coronavirus Pandemic, Capacity, Demand, Europe, Risk, Crisis
Topic 3	Terminal, Government, CMA CGM, DP World, Operations, Singapore

Topic2의 주요 단어는 ‘Supply Chain’, ‘Capacity’, ‘Demand’, ‘Risk’, ‘Crisis’ 등이다. Topic2에 속하는 주요 기사들의 내용은 전세계적으로 지역봉쇄가 계속되고, 중국 정부의 경제 정상화 주문으로 공장이 재가동 되면서 글로벌 공급사슬의 물동량의 증가와 감소에 대한 정확한 예측이 어렵다는 것이다. 공급사슬의 불안정성이 증폭됨에 따라서 공급사슬 단절현상이 이 시기에 많이 발생했다. 공급사슬 단절이란 갑작스런 재해 또는 사고에 의해서 공급사슬을 통해 자재들이 공급되지 못하고 일시에 중단되는 상황을 의미한다 (Tomlin, 2006). 이에 따라서 선사들이 선제적으로 기항 횟수를 줄이게 된다.

Topic3의 주요 단어들은 ‘Terminal’, ‘Government’, ‘CMA CGM’, ‘Singapore’, ‘DP World’, ‘Operations’ 등 이다. Topic3에 속하는 주요 기사들의 내용은 물동량이 계속줄어드는 가운데 CMA CGM, Maersk 등의 글로벌 선사들은 해운운임의 상승세 속에 실적을 개선하면서 터미널을 인수하는 등 공격적으로 투자를 진행하고 있으며, 터미널의 혼잡(Termial Congestion)이 증가하고 공컨테이너(Empty Container)의 확보가 어려워져 임시결항과 일괄 운임인상(General Rate Incarese) 등의 방법으로 수익성을 강화하기 시작했다는 것이다. 기항 횟수의 감소로 인한 여파가 이 시기부터 시작되었음을 기사를 통해 파악할 수 있다.

4.2 2020년도 7월부터 9월까지의 기사분석

2020년도 7월부터 9월까지의 로이즈리스트에 게재된 컨테이너선 관련 기사 수는 총 93편이

다. TF-IDF 값이 0.5이하인 불용성의 단어를 제거하고 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 4,855개이다.

4.2.1 네트워크 텍스트 분석

이 시기의 기사들에서 단어를 추출하여 기사-단어 이원모드 네트워크를 구축하였다. 이원모드의 링크의 수는 총 12,199개이다. 기사에서 가장 많이 출현한 단어들은 ‘Coronavirus Pandemic’, ‘Supply Chain’, ‘Maersk’, ‘Terminal’, ‘Beirut’ 등 이었다. 이전 시기에 비해서 상대적으로 ‘Blank Sailings’, ‘Congestion’, ‘Delays’ 등의 단어들의 빈도수가 증가하였음을 알 수 있다. 터미널의 혼잡상황이 악화되었음을 유추할 수 있다. 기사-단어 이원모드 네트워크를 단어-단어 일원모드 네트워크로 변환하였으며, 링크 수는 총 920개였다.

단어 노드의 영향력을 파악하기 위해서 연결 중심성을 분석하였다. 연결중심성 분석결과 베이르트항 폭발 관련 키워드들이 중심성이 높았다. 또한 지난 시기에 비해서 터미널의 혼잡, 임시결항과 관련한 키워드들의 중심성이 높아졌다. 싱가포르의 경우 유연성과 탄력성을 기반으로 화물 컨테이너 물동량 변동에 안정적으로 대처하고 있다는 기사들이 나오면서 관련한 단어들의 중심성도 높아졌다. 네트워크의 응집성 분석을 위해서 주요 컴퍼넌트를 찾고 커뮤니티 분석을 시행하였다. Modularity Index를 고려하여 3개의 클러스터가 형성되는 것으로 파악하였다.

4.2.2 토픽 모델링 분석

네트워크 응집구조를 참고하여 토픽의 수를 3으로 하여 토픽 모델링 분석을 시행하였다. 토픽별 주요 단어는 Table 2와 같다.

Table 2 Topics and Keyword (from 2020.7 to 2020.9)

Group	Keywords
Topic 1	Terminal, Supply Chain, Trade, US, Singapore, UK, Beirut
Topic 2	China, Hong Kong, Capacity, Cost
Topic 3	Coronavirus Pandemic, Maersk, Demand

Topic1의 주요단어는 ‘Terminal’, ‘Supply Chain’, ‘Trade’, ‘US’, ‘Singapore’, ‘UK’, ‘Beirut’ 등이다. Topic1에 속하는 주요 기사들의 내용을 살펴보면, 베이르트항의 폭팔과 관련한 이슈들과 영국의 Felixstowe 터미널은 혼잡을 이유로 공컨테이너의 보관을 거부하고 있으며 이유는 늘어나는 수요에 비해 근로자의 수가 매우 부족했던 상황으로 인한 컨테이너 야드의 혼잡을 이유로 공컨테이너의 보관을 거부하고 있다. 반면에 싱가포르항은 유연성과 탄력성을 기반으로 화물 컨테이너 물동량을 안정적으로 확보하고 있다는 내용의 기사도 중심을 이루고 있다.

Topic2의 주요단어는 ‘China’, ‘Hong Kong’, ‘Capacity’, ‘Cost’ 등이다. Topic2에 속하는 주요 기사들의 내용을 살펴보면, 중국의 수출감소와 이로인해 홍콩 터미널의 환적이 감소하고 있다는 것이다. Hutchison Port Holdings Trust의 물동량이 8% 감소하고 아시아 내의 미국의 물동량이 전반적으로 감소하고 있다는 내용의 기사들도 있다. 이로 인해 임시결항의 비율도 여전히 높음을 유추할 수 있다.

Topic3의 주요단어는 ‘Coronavirus Pandemic’, ‘Maersk’, ‘Demand’ 등이다. Topic3에 속하는 주요 기사들의 내용을 살펴보면, 글로벌 선사들이 코로나 판데믹의 여파로 물동량의 감소를 예상하고 수익성 개선을 위해 선복량을 감소시키는데 주력하고 있다는 것과 Maersk의 경우 2분기에 깜짝실적을 달성한 가운데 이러한 고공행진이 3분기에도 이어질 것으로 전망하고 있다는 것이다. 이는 컨테이너선 운임의 가파른 상승세가 글로벌 선사의 운영전략과 실적에 상당한 영향을 주고 있음을 알 수 있는 대목이다.

4.3 2020년도 10월부터 12월까지의 기사분석

2020년도 10월부터 12월까지의 로이즈리스트에 게재된 컨테이너선 관련 기사 수는 총 67편이다. 불용성 단어를 제거하고 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 3,657개이다.

4.3.1 네트워크 텍스트 분석

총 8,918개의 링크로 구성된 기사-단어 이원모드를 구축하였다. 이 시기의 빈출단어들은 ‘Congestion’, ‘Coronavirus Pandemic’, ‘Supply Chain’, ‘Demand’, ‘Equipment’ 등 이었다. 이원모드 네트워크를 총 632개의 링크로 이루어진 단어-단어 일원모드 네트워크로 변환하여 중심성 분석을 시행하였다.

연결중심성 분석 결과, 공급사슬의 단절현상과 관련한 키워드들이 중심성이 높아졌음을 확인할 수 있다. 또한 터미널의 혼잡이 전세계적으로 본격화되면서 관련한 단어들이 중심성이 높아지고, 컨테이너의 부족 현상에 대한 키워드들이 네트워크에서 주도적인 역할을 하고 있었다. 20년도 하반기로 접어들면서 컨테이너 물동량의 감소와 이로 인한 글로벌 선사들의 임시결항과 같은 선제적인 조치들이 시간차를 두고 공급사슬 단절과 컨테이너 부족과 같은 현상을 심화시키고 컨테이너선 운임이 상승되었음을 알 수 있다.

네트워크의 응집성 구조를 파악하기 위해서 커뮤니티 분석을 시행하였으며 3개의 클러스터가 형성되었다.

4.3.2 토픽 모델링 분석

네트워크 응집구조에 따라 토픽의 수를 3으로 하여 토픽 모델링 분석을 시행하였다. 토픽별 주요 단어는 Table 3과 같다.

Topic1의 주요단어는 ‘Asia’, ‘UK’, ‘US’, ‘Capacity’, ‘China’ 등이다. Topic1에 속하는 주요 기사들의 내용을 살펴보면, 아시아에서 미국이나 유럽으로 향하는 물동량은 계속 증가하고 있으나 돌아오는 백홀(Back Haul)물동량이 부족하여 컨테이너를 회수하는데 상당한 어려움이 있음을 알 수 있다. 영국의 Felixtow 터미널 혼잡은 어느정도 해소되었으나 만족할만한 유연성이나 탄력성을 보여주고 있지 못하고 있다는 기사들도 다수 있었다.

Table 3 Topics and Keyword (from 2020.10 to 2020.12)

Group	Keywords
Topic 1	Asia, UK, US, Capacity, China
Topic 2	Cost, Trade, Demand, High, Equipment
Topic 3	Supply Chain, Terminal, Congestion

Topic2의 주요단어는 ‘Cost’, ‘Trade’, ‘Demand’, ‘High’, ‘Equipment’ 등이다. Topic2에 속하는 주요 기사들의 내용을 살펴보면, 터미널 혼잡의 심화가 해운운임의 상승으로 계속 이어지고 있으며 특히 영국의 Felixtow 터미널 혼잡은 London Gateway와 Port of Southampton에 연쇄적으로 영향을 주는 동시에 근로자들의 근무환경과 불만과 임금인상 요구로 이어지고 있음을 알 수 있다.

Topic3의 주요단어는 ‘Supply Chain’, ‘Terminal’, ‘Congestion’ 등이다. Topic3에 속하는 주요 기사들의 내용을 살펴보면, 공컨테이너의 부족으로 해운 공급사슬의 단절과 불확실성이 역대 최고 수준이며 이에 따른 공급사슬의 탄력적 운영과 관리가 선사와 항만을 중심으로 잘 이루어지고 있지 않음을 지적하고 있다. 탄력적 공급사슬은 수요와 공급의 급격한 변화와 단절 속에서도 빠른 시간 내 위험관리와 협력을 통해 민첩하게 본래대로 정상 운영이 가능한 공급사슬을 의미한다.

4.4 2021년도 1월부터 2월까지의 기사분석

2021년도 1월부터 2월까지의 로이즈리스트에 게재된 컨테이너선 관련 기사 수는 총 151편이다. TF-IDF 값이 0.5 이하인 불용성의 단어를 제거하고 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 7,289개이다.

4.4.1 네트워크 텍스트 분석

추출한 단어들을 대상으로 기사-단어 이원모드를 구축하였으며 링크의 수는 총 19436개이다. 이 기간 동안 가장 많이 기사에서 언급한

단어는 ‘Coronavirus Pandemic’, ‘Maersk’, ‘Supply Chain’, ‘Demand’, ‘China’, ‘Crew’ 등이었다. 빈출단어로 ‘Crew’가 새로 추가되었으며, 이는 선원들의 코로나 감염과 자가격리 등으로 인한 물류지연 현상과 무관해보이지 않는다. 또한 중국의 춘절이 다가옴에 따라 급증한 운임이 중국공장들의 생산 감소로 운임의 상승세가 약화될 것이라는 전망도 있었다. 기사-단어의 이원모드 네트워크를 총 1,422개의 링크로 이루어진 단어-단어 일원모드 네트워크로 변환하였다. 이 시기에 미국의 LA항과 롱비치항에서 코로나 집단감염이 일어나면서 컨테이너 화물 처리가 적시에 이뤄지지 못하고 이에 따라 컨테이너와 컨테이너선이 부족으로 미국 서부 해안의 항만이 적체율이 최고 수준으로 다다르고 있는 상황과 관련한 단어들이 중심성이 높아졌다. 또한 국제해사기구의 환경규제에 따른 LNG 연료 사용, 탈황 등과 같은 키워드들이 중심성이 높아졌으며 이는 모두 운임상승으로 이어질 수 있는 요소들이다.

단어 네트워크의 응집구조를 파악하기 위해서 커뮤니티 분석을 시행하였으며 3개의 클러스터가 형성되는 것을 확인하였다.

4.4.2 토픽 모델링 분석

네트워크 응집구조 분석에 따라 주요 토픽의 수를 3으로 하여 토픽 모델링 분석을 시행하였다. 토픽별 주요 단어는 Table 4와 같다.

Topic1의 주요단어는 ‘US’, ‘Los Angeles’, ‘Trade’, ‘West Coast’ 등이다. Topic1에 속하는 주요 기사들의 내용을 살펴보면, 미국의 서부 해안의 로스엔젤레스와 롱비치항에서 항만 작업자 수백명이 코로나에 감염되면서 체선(Demurrage)현상이 심화되고 있으며 이는 운임상승으로 이어질 수 있다는 것이다. 또한 환경규제에 따른 탈황연료, LNG연료로의 전환이 광범위하게 논의되고 있음을 기사들을 통해 알 수 있다. 이는 모두 컨테이너선 운임상승의 기폭이 될 수 있는 요인들이다.

Table 4 Topics and Keyword (from 2021.1 to 2021.2)

Group	Keywords
Topic 1	US, Los Angeles, Trade, West Coast
Topic 2	Coronavirus Pandemic, China, Terminal, Production
Topic 3	Maersk, Crew, Fuel, Demand, Singapore

Topic2의 주요단어는 ‘Coronavirus Pandemic’, ‘China’, ‘Terminal’, ‘Production’ 등이다. Topic2에 속하는 주요 기사들의 내용을 살펴보면, 터미널과 관련된 이슈로서 컨테이너의 회수가 일정 부분 진행되고 있고 이에 따라 선박사용률(Vessel Utilisation Levels)이 증가하고 있다는 것이다. 이는 중국의 춘절로 인해 물동량이 줄어들 것이라는 전망과 맥락을 같이한다. Topic1이 해운운임의 상승을 예견한 내용이라면 Topic2는 이와는 달리 운임의 상승세가 안정세로 접어들 수 있음을 전망한 내용이다.

Topic3의 주요단어는 ‘Maersk’, ‘Crew’, ‘Fuel’, ‘Demand’, ‘Singapore’ 등이다. Topic3에 속하는 주요 기사들의 내용을 살펴보면, 글로벌 선사인 Maersk의 2020년 4분기 성과가 최대치이며 이는 지속적으로 컨테이너 운임을 인상한 결과로 분석한 것이다. 이는 컨테이너 부족으로 인한 유지비용 증가의 결과로 분석하고 있다. Maersk는 엔텐항을 비롯한 주요 터미널의 혼잡도의 지속적인 증가를 전망하여 컨테이너선 운임이 지속적으로 상승할 것으로 전망했다. 또한 싱가포르, 미국 등을 비롯한 주요 항만의 근로자와 선원들의 코로나 감염 또는 자가격리 등으로 인한 물류지연을 언급하는 기사들이 있었다.

4.5 2021년도 3월부터 4월까지의 기사분석

2021년도 3월부터 4월까지의 컨테이너선 관련 기사 수는 총 59편이다. 불용성의 단어를 제거하고 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 3,875개이다.

4.5.1 네트워크 텍스트 분석

총 7794개의 링크로 구성된 기사-단어 이원모

드를 구축하였다. 이 시기의 빈출단어들은 ‘Coronavirus Pandemic’, ‘Supply Chain’, ‘Capacity’, ‘Suez Canal’, ‘Disruption’ 등 이었다. 이 시기에 대만 국적 선사인 에버그린(Evergreen) 소속의 에버기븐(Ever Given) 호가 좌초하여 수에즈 운하를 봉쇄하면서 이와 관련한 단어들이 다수 출현하였다.

이원모드 네트워크를 활용하여 단어-단어 일원모드 네트워크를 구축하였으며, 링크의 수는 총 352개가 형성되었다. 연결중심성이 높은 단어들은 주로 수에즈 운하 봉쇄와 관련한 것들이다. 초대형 선박 에버기븐(Ever Given) 호가 수에즈 운하를 가로 막으면서 선박들이 대기하면서 하루에 수십억 달러씩의 손실이 발생하였고, 수에즈 운하의 사용이 재개되면서 선박들이 한번에 항만에 몰리면서 적체현상이 심화되고 있다는 내용의 키워드들이 주를 이루고 있다. 수에즈 운하 봉쇄로 인해서 급격한 운임상승세에서 벗어나 조정세에 접어들 것으로 전망되었던 컨테이너선의 운임이 수주 이상의 상승세를 지속하게 된다. 수에즈 운하 통행이 재개되면서 그동안 적체되어있던 선박들이 항만으로 향하고, 미국에서는 소비증가로 인한 물동량 증가에 따른 공컨테이너와 컨테이너선의 부족은 이 시기에 더욱 심각해졌다. 선박과 컨테이너의 부족으로 선박과 컨테이너에 대한 발주량은 역대 최고치에 이를 것으로 전망되었다. 또한 영국에서는 브렉시트(Brexit)로 인해 재고비축 차원에서 공컨테이너들을 계속 쌓아두면서 컨테이너 가용지수가 0.5를 상회하는 0.90으로 상승하는 등 해운운임 증가에 지속적으로 영향을 주고 있었다.

네트워크의 응집구조를 커뮤니티 분석을 통해 파악하였다. 전체 네트워크가 하나의 클러스터를 이루고 있음을 알 수 있었다. 이는 수에즈 운하 봉쇄라는 과급력이 큰 재난사고가 있었기 때문으로 분석된다. 따라서 이 시기의 기사에 대한 토픽 모델링 분석은 별도로 시행하지 않았다.

4.6 2021년도 5월부터 7월까지의 기사분석

2021년도 5월부터 7월까지의 로이즈리스트에

게재된 컨테이너선 관련 기사 수는 총 263편이다. TF-IDF 값이 0.5 이하인 불용성의 단어를 제거하고 복합어, 제외어, 유의어를 지정하여 추출한 단어 수는 총 10,039개이다.

4.6.1 네트워크 텍스트 분석

추출한 단어들을 대상으로 기사-단어 이원모드를 구축하였으며 링크의 수는 총 34176개이다. 이 기간 동안 가장 많이 기사에서 언급한 단어는 ‘Coronavirus Pandemic’, ‘Supply Chain’, ‘Demand’, ‘China’, ‘Suez Canal’, ‘Yantian’ 등 이었다. 다른 시기와는 달리 빈출단어로 ‘Yantian’이 나타났다. 이는 중국 선전시 옌텐(Yantian)항의 근로자 4명이 코로나 감염 양성 반응을 보이면서 터미널이 부분적으로 폐쇄된 것과 관련이 있어 보인다.

기사-단어의 이원모드 네트워크를 총 3,629개의 링크로 이루어진 단어-단어 일원모드 네트워크로 변환하였다. 연결중심성 분석결과에 따르면 중국 옌텐항의 부분 폐쇄와 관련한 단어들이 중심성이 높은 것을 확인할 수 있다. 옌텐항의 부분 폐쇄가 2주 정도 지속되는 경우 그 여파는 거의 없으나, 장기화되는 경우 처리량의 부족을 심화시켜 유럽과 북미 물동량에 심각하게 영향을 줄 것으로 전망하고 있었다. 또한 선원들의 코로나 감염사례가 증가하면서 선박이 목적지까지 이르지 못하고 중간지점에서 다른 항구에 정박하여 선원을 교체하는 상황까지 발생하면서 근무자의 코로나 감염에 대한 키워드들이 중심성이 높아지고 있음을 확인할 수 있다. 단어-단어 일원모드 네트워크의 응집구조를 파악하기 위해서 커뮤니티 분석을 시행하였으며 3개의 클러스터가 형성되는 것을 확인하였다.

4.6.2 토픽 모델링 분석

네트워크 응집구조에 따라 토픽의 수를 3으로 하여 토픽 모델링 분석을 시행하였다. 토픽별 주요 단어는 Table 5와 같다.

Topic1의 주요단어는 ‘Coronavirus Pandemic’, ‘Supply Chain’, ‘Demand’, ‘Disruption’ 등이다. Topic1에 속하는 주요 기사들의 내용을 살펴보면, 전문가들이 공급사슬의 단절 현상이 안정세

에 접어들 것이라는 전망을 했지만 21년도 들어서 공급사슬의 단절현상이 더욱 악화되고 있다는 것이다. 주요 터미널의 혼잡으로 인해서 화물의 절반 이상이 예정된 일정을 따르지 못하고 다음 선박으로 이월되고 있다고 기사들이 전하고 있다. 이에 따른 공급사슬에 대한 구조적 조정과 가시성 확보를 위한 노력이 병행되어야 함을 전문가들이 지적하고 있다.

Table 5 Topics and Keyword (from 2021.5 to 2021.7)

Group	Keywords
Topic 1	Coronavirus Pandemic, Supply Chain, Demand, Disruption
Topic 2	Order, China, US, Yantian, Capacity
Topic 3	Ever Given, Suez Canal, Crew, Seafarer

Topic2의 주요단어는 ‘Order’, ‘China’, ‘US’, ‘Yantian’, ‘Capacity’ 등이다. Topic2에 속하는 주요 기사들의 내용을 살펴보면, 중국 선전시 엔텐항에서 근로자들의 집단감염으로 부분 폐쇄가 이뤄졌고 다시 정상적으로 운영이 재개되었지만 누적된 컨테이너들이 2만개가 넘어서 물류망을 회복하기까지 최소 2달정도 걸릴 것으로 전망하고 있다는 것이다. 또한 향후 해운운임 계절적인 성수기에 접어들면서 상승세가 지속될 것이라는 전망을 전하는 기사들도 있었다.

Topic3의 주요단어는 ‘Ever Given’, ‘Suez Canal’, ‘Crew’, ‘Seafarer’ 등이다. Topic3에 속하는 주요 기사들의 내용을 살펴보면, 수에즈 운하의 여파에 대한 분석들이 주를 이루고 있다. 일반적으로 중국의 춘절을 전후하여서는 운임의 하락이 예상되지만 2021년도에는 공급사슬의 단절, 수에즈 운하 폐쇄 등으로 사상 최고치를 기록했다. 수에즈 운하 운행 재개이후 상승세가 잠시 멈추었으나 아시아 항만의 공컨테이너 부족이 심화되면서 다시 상승 국면에 접어들었음을 기사들이 전하고 있다. 또한 터미널 적체 현상이 심화되면서 근로자들의 근무여건 개선과 임금상승에 대한 요구도 이어지고 있어 물류지연과 운임상승은 당분간 지속될 것으로 전망되었다.

5. 결론

본 연구에서는 2020년도부터 시작된 컨테이너선 운임지수의 급격한 상승세와 관련한 요인들을 파악하고 향후 컨테이너선 운임 예측모형 개발을 위한 기초 분석 자료들을 제공하고자 하였다. 이를 위해서 컨테이너선과 관련한 문헌을 대상으로 텍스트 마이닝 분석을 시행하였다.

분석대상 문헌은 대표적인 해운 전문지인 로이저리스트에 2020년 1월부터 2021년 7월까지 게재된 컨테이너선 기사들이었다. 컨테이너선 관련 기사들을 취합하여 2020년 1월부터 2021년 7월까지 컨테이너선 운임지수의 상승의 변곡점이 되는 구간을 고려하여 총 6개의 기간으로 구분하여 텍스트 분석을 시행하였다.

Table 6 Main Clusters of Keywords

Main Clusters of Keywords	Related Articles
Coronavirus Pandemic, China, Disruption, Trade, US	US-China trade war due to Coronavirus Pandemic
Capacity, Crew, Demand, Seafarer, Terminal, Congestion	Terminal congestion worsens as global shipping container shortage continues
Equipment, Suez Canal, Yantian	Unexpected events such as the Suez Canal blockage and the shutdown of Yantian port lead to an unprecedented surge in shipping rates

6개 기간의 텍스트 분석결과를 토대로 시간 흐름에 따른 운임상승의 주요 키워드들과 토픽 그룹을 요약하면 Table 6과 같다. 첫째, 코로나 19 이후 미국과 중국의 무역분쟁으로 인한 물동량의 감소, 중국의 춘절로 인한 휴업 및 생산감소로 인한 컨테이너 가용감소, 이로 인한 글로벌 공급사슬의 불확실성 증대되었다. 이에 따라서 글로벌 선사들이 글로벌 공급사슬의 불확실성에 대응하기 위한 임시결항이 증가하면서 회복되는 물류량을 감당하지 못하였다. 둘째, 코로나 19로 인해 항만근로자의 코로나 확진, 업무과중으로 인한 파업 등으로 터미널 운영이 원활하지 못하여 혼잡도가 증가하였고, 물류지연으로

인해서 공컨테이너 부족(Equipment Shortage) 현상이 장기화 되었다. 셋째, 수에즈 운하 봉쇄, 엔텐항 폐쇄 등과 같은 예기치 않은 이벤트로 인해서 물류지연이 가속화 되어 운임상승의 폭이 커졌다.

본 연구의 시사점은 다음과 같다. 첫째, 컨테이너선의 운임과 관련한 문헌들을 대상으로 하는 네트워크 텍스트 분석 방법론을 체계화하여 분석한 것이다. 기존의 해운관련 텍스트 마이닝 연구들이 단어들의 추출과 출현빈도를 중심으로 분석을 해왔으나, 본 연구에서는 기사-단어 네트워크, 단어-단어 네트워크 등을 구성하여 텍스트에 포함된 숨겨져 있는 의미들을 체계적으로 분석하였다. 본 연구에서는 로이즈리스트에 게재된 기사들만을 대상으로 네트워크 텍스트 분석을 시행하였으나, 다양한 국내외의 문헌들을 동시에 분석할수록 텍스트 마이닝 분석의 의의가 있다고 할 수 있다. 제시한 방법론이 향후 다양한 해운운임 관련 문헌분석에 활용될 수 있을 것으로 기대한다.

둘째, 2020년부터 지속되고 있는 컨테이너선의 운임상승과 관련한 요인들을 시기별로 구분하여 분석하였기 때문에 컨테이너선 운임상승과 관련한 향후 연구들에 활용될 수 있다. 본 연구에서는 컨테이너선 운임지수의 변곡점에 따라 6개의 시기로 구분하여 기사-단어, 단어-단어 등의 네트워크를 구축하여 분석을 시행하였고, 시기별로 빈출단어와 단어 노드의 연결중심성이 변화하는 것을 보여주었다. 또한 LDA 토픽 모델링 기법을 활용하여 토픽그룹에 속하는 기사들의 내용변화를 시간의 흐름에 따라 제시하여 향후 텍스트 마이닝 분석결과를 토대로 운임지수의 변동과 관련지어 시계열 분석 및 다양한 데이터 마이닝 분석이 가능하다. 본 연구에서 제시한 토픽 그룹들과 주요 키워드들의 중심성 분석 결과는 기존의 정형 데이터 분석 결과들과 큰 틀에서 차이는 없지만, 정형 데이터 분석과는 달리 시간의 흐름에 따른 주요 토픽의 변화를 탐색적으로 제시한 점에서 의의가 있다.

셋째, 텍스트라는 비정형 데이터를 활용하여 2020년 이후의 가파른 컨테이너선 운임의 상승세를 분석하였다. 기존의 연구들이 정형적이고

정량적인 데이터를 활용하여 수치 관계 분석을 주로 시행하였다면, 본 연구에서는 비정형 데이터인 텍스트를 그대로 활용하여 다양한 분석을 시도하였다. 향후 컨테이너선 운임과 관련한 다양한 비정형 데이터들이 활용될 수 있기를 기대한다.

본 연구는 다음의 한계점을 갖고 있다. 분석 대상 문헌을 로이즈리스트의 컨테이너 카테고리에 2020년부터 2021년도 7월까지 게재된 기사들만을 대상으로 하였기 때문에 컨테이너 관련 이슈 이외의 다양한 사회문화적 이슈들과 이벤트를 반영하지 못했을 수 있다. 또한 코로나19 팬데믹 이전의 문헌자료를 대상으로 분석을 시행하지 않았기 때문에 코로나19 이전과 이후의 컨테이너선 운임 관련한 이슈들을 체계적으로 종합하여 비교하지 못했다. 향후 연구에서는 이러한 한계점들을 보완할 필요가 있다.

References

- Ahn, Y. G. (2021). Strategies of the Korean Shipping Industry in the Post-Corona Era, *Korean Logistics Review*, 31(4), 29-40.
- Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent dirichlet allocation, *The Journal of Machine Learning Research*, 3, 993-1022.
- Cullinane, K., Haralambides, H. (2021). Global Trends in Maritime and Port Economics: the COVID-19 Pandemic and Beyond, *Maritime Economics and Logistics*, 23(3), 369-380.
<https://doi.org/10.1057/s41278-021-00196-5>.
- Dirzka, C., Acciaro, M. (2021). Global Shipping Network Dynamics During the COVID-19 Pandemic's Initial Phases, *Journal of Transport Geography*, <https://doi.org/10.1016/j.jtrangeo.2021.103265>.
- Girvan, M., Newman, M. E. (2002). Community structure in social and biological networks, *Proceedings of the*

- National Academy of Sciences*, 99(12), 7821-7826.
- Griffiths, T. L., Steyvers, M. (2004). Finding Scientific Topics, *Proceedings of the National Academy of Sciences*, 101(1), 5228-5235.
- Lam, J. S. L., Li, Q., Pu, S. (2021). Volatility and Uncertainty in Container Shipping Market. In *New Maritime Business, WMU Studies in Maritime Affairs*, https://doi.org/10.1007/978-3-030-78957-2_2.
- Lee, J. M. Y., Wong, E. Y. C. (2021). Suez Canal blockage: an analysis of legal impact, risks and liabilities to the global supply chain, *In MATEC Web of Conferences (Vol. 339)*, EDP Sciences.
- Pyun, J., Rha, J. S. (2021). Review of Research on Digital Supply Chain Management Using Network Text Analysis, *Sustainability*, 13(17), 9929. <https://doi.org/10.3390/su13179929>.
- Ramos, K. G., Rocha, I. C. N., Cedeño, T. D. D., dos Santos Costa, A. C., Ahmad, S., Essar, M. Y., Tsagkaris, C. (2021). Suez Canal Blockage and its Global Impact on Healthcare Amidst the Covid19 Pandemic, *International Maritime Health*, 72(2), 145-146. <https://doi.org/10.5603/IMH.2021.0026>.
- Rha, J. S. (2020). Trends of Research on Supply Chain Resilience: A Systematic Review Using Network Analysis, *Sustainability*, 12(11), 4343. <https://doi.org/10.3390/su12114343>.
- Salton, G., Buckley, C. (1988). Term-weighting Approaches in Automatic Text Retrieval, *Information Processing and Management*, 24(5), 513-523.
- Shanghai Shipping Exchange. (2021). *Shanghai Containerized Freight Index*, <https://en.sse.net.cn/indices/scfinew.jsp> (Accessed on Oct. 15th, 2021)
- Song, D. (2021). A Literature Review, Container Shipping Supply Chain: Planning Problems and Research Opportunities, *Logistics*, 5(2), 41. <https://doi.org/10.3390/logistics5020041>.
- Tomlin, B. (2006). On the Value of Mitigation and Contingency Strategies for Managing Supply Chain Disruption Risks, *Management Science*, 52(5), 639-657.



나 진 성 (Jin Sung Rha)

- 종신회원
- 연세대학교 교육학학사/경영학사
- 미국 네브라스카 주립대학교 경영학 석사/경영학 박사
- 단국대학교 경영경제대학 경영학부 부교수
- 관심분야 : 공급사슬리스크관리, 친환경 공급사슬, 네트워크 분석