

Analysis of Technology Trends in Smart Port Technologies Using Topic Modeling and Generative Topographic Mapping

Hee-Kyeong Nam*, Jeong-Yoon Hong**, Min-Su Woo*, Go-Eun Han*,
Jae-Min Song*, Yul-Kyeong Yun*, In-Yeong Lee*, Sung Won Cho***

*Student, Dept. of Management Engineering, Dankook University, Cheonan, Korea

**Student, Dept. of Industrial Engineering, Dankook University, Cheonan, Korea

***Professor, Dept. of Management Engineering, Dankook University, Cheonan, Korea

[Abstract]

As the digitalization of the port industry accelerates, this study aims to identify core and vacant technologies related to smart ports using machine learning methods and propose development strategies for each technology. We applied TF-IDF (Term Frequency-Inverse Document Frequency) to patent data on smart port technologies to extract key terms. Using LDA (Latent Dirichlet Allocation) topic modeling, we identified core technologies such as container scanning systems, port simulation, smart containers, shipping scheduling algorithms, object detection for port security, remote control-based autonomous ships, and autonomous vehicles in ports. Furthermore, we generated a technology map using GTM (Generative Topographic Mapping), identifying autonomous vehicles in ports as a vacant technology.

▶ **Key words:** Smart port, Technological forecasting, Patent analysis, Machine learning, Topic modeling, Generative Topographic Mapping (GTM)

[요 약]

본 연구는 스마트 항만 기술이 급격히 발전하는 것에 대응하여 스마트 항만의 핵심기술 및 공백기술을 파악하고 각 기술별 개발 전략을 제안한다. 스마트 항만 기술에 대한 특허 데이터에 TF-IDF (Term Frequency-Inverse Document Frequency)를 적용하여 주요 용어를 추출하였다. 이후 LDA (Latent Dirichlet Allocation) 기반의 토픽 모델링을 사용하여 컨테이너 내부 스캐닝 시스템, 항만 시뮬레이션 기술, 스마트 컨테이너 기술, 선적 계획 수립 알고리즘, 객체 탐지 기술, 원격관제 기반의 자율운항선박 기술, 항만 내 차량의 자율주행 기술과 같은 핵심기술을 도출하였다. 또한, GTM (Generative Topographic Mapping) 기반의 스마트 항만 기술 맵을 생성하여 항만 내 차량의 자율주행 기술이 공백기술임을 확인하였다.

▶ **주제어:** 스마트 항만, 기술 예측, 특허 분석, 기계학습, 토픽 모델링, Generative Topographic Mapping (GTM)

- First Author: Hee-Kyeong Nam, Corresponding Author: Sung Won Cho
- *Hee-Kyeong Nam (heekyunnam132@naver.com), Dept. of Management Engineering, Dankook University
- **Jeong-Yoon Hong (jyoon333@dankook.ac.kr), Dept. of Industrial Engineering, Dankook University
- *Min-Su Woo (mwoosu@naver.com), Dept. of Management Engineering, Dankook University
- *Go-Eun Han (hangoeun0418@naver.com), Dept. of Management Engineering, Dankook University
- *Jae-Min Song (song71jpp@naver.com), Dept. of Management Engineering, Dankook University
- *Yul-Kyeong Yun (dreamkhan20@naver.com), Dept. of Management Engineering, Dankook University
- *In-Yeong Lee (dldlsdud0870@naver.com), Dept. of Management Engineering, Dankook University
- ***Sung Won Cho (sungwon.cho@dankook.ac.kr), Dept. of Management Engineering, Dankook University
- Received: 2024. 11. 01, Revised: 2024. 12. 24, Accepted: 2024. 12. 26.

I. Introduction

세계 무역의 급성장에 따라 해상 운송은 더욱 중요해지고 있다. 세계 해상 무역의 규모는 2022년 약 120억 톤에 이르렀으며, 2024~2028년에는 2% 이상 성장할 것으로 예상된다[1]. 해상 운송 수요가 증가함에 따라 효율적인 항만 운영의 중요성이 대두되고 있으며, 항만의 스마트화가 가속화되고 있다[2].

국제해사기구(International Maritime Organization; IMO)에서도 항만을 포함한 해상 산업에서 디지털화의 중요성을 인식하고 있다[3]. IMO는 외국 항만에 입항하는 선박에 대한 정보를 전자적으로 교환하는 Maritime Single Windows 사용을 의무화하는 등 해상 운송 행정의 단순화 및 효율화를 목표로 항만의 디지털화를 추진하고 있다. 또한, 데이터 교환에 대한 Reference Data Model을 정의하여 다양한 시스템과 국가 간 데이터 교환을 표준화하고 일관성을 유지하기 위한 노력을 기울이고 있다[4,5]. 국내의 경우 기존 항만 운영 시스템에서 4차 산업혁명 기술들이 단기간에 적용될 수 있는 영역에 초점을 맞추어 따라 스마트 항만의 추진이 제대로 이루어지지 않고 있다. 스마트 항만은 자체적으로 의사결정 능력을 갖춘 지능화 항만으로서, 정보화와 자동화가 선제되어야 하며 항만 전체가 아니라도 항만의 일부는 스마트화가 진행되어야 한다[6].

이에 국내에서도 스마트 항만의 구축을 위한 움직임이 보이고 있다. 해양수산부는 '2030 항만정책 방향 및 추진 전략'을 제시하여 2030년도 부산항에 한국형 스마트 항만을 본격적으로 도입하는 것을 목표로, 항만의 스마트화·지능화를 적극적으로 추진하고 지속적인 인프라 확충을 통해 항만의 글로벌 경쟁력을 강화하고자 한다[7]. 그러나 이러한 노력에도 불구하고, 여전히 국내의 스마트 항만 수준은 글로벌 선진 항만 대비 미흡한 수준이다. 국내 스마트 항만은 글로벌 선진 항만들에 비해 자동화 수준이 낮고, 디지털화된 물류 데이터의 수집과 연계가 어렵기 때문에 항만 내 공급 사슬 전체의 운영 효율화를 위한 분석과 개선에 한계가 존재한다. 이에 따라, 항만 내외부의 정보 연계가 부족하여 다양한 이해 관계자들이 각각의 역할을 제대로 수행하지 못한다는 문제점이 지속되고 있다[8,9]. 더불어, 하역, 이송, 무인 자동화 도입뿐만 아니라 항만 내의 모든 구성 요소를 초연결하는 항만의 디지털화를 꾸준히 추진하고 있는 네덜란드, 영국, 중국 등의 세계 주요 항만들과 달리, 국내 스마트 항만은 현재 부산과 인천, 광양 등 일부 지역에서만 항만 자동화를 추진하는 실정이다. 전 세계적으로 항만이 디지털화의 중요성을 인식하고 있음에

따라, 국내 항만에서도 모든 요소를 균형 있게 발전시키는 스마트 항만의 구축이 필요하다[10].

따라서, 본 연구에서는 국내 스마트 항만 기술에 대한 트렌드를 분석하여 기술 개발 전략을 제안하고자 한다.

기술의 트렌드를 분석하는 방법에는 전문가 패널 바탕의 델파이 기법, 리서치 보고서, 학술 논문 등의 문헌 조사, 기술 관련 키워드의 데이터 분석 등이 존재한다. 본 연구에서는 특히 데이터에 대한 분석을 진행한다. 특히 데이터는 발명 명칭, 내용, 배경 기술, 기술 분야, 기술 분류 코드 등 기술에 대한 구체적인 정보를 포함한다. 따라서, 특히 데이터를 활용하여 기술의 진보 수준과 기술 생태계를 이해하는 등 기술을 다각적으로 분석할 수 있다.

먼저, 스마트 항만과 관련된 특허 데이터를 대상으로 TF-IDF(Term Frequency-Inverse Document Frequency)를 적용하여 키워드를 추출한 후, 토픽 모델링 기법인 LDA(Latent Dirichlet Allocation)를 적용하여 핵심 기술을 제시한다. 이후 토픽 모델링 결과를 바탕으로 GTM(Generative Topographic Mapping) 기반의 기술 맵을 생성하여 공백 기술을 도출한다. 이러한 결과를 바탕으로 항만 디지털 전환과 기술의 경쟁력 향상을 위한 기술 개발의 방향성을 제시하고자 한다.

본 논문의 구성은 다음과 같다. 제2장에서는 스마트 항만 기술 개발 현황과 관련 연구에 대해 조사한다. 제3장에서는 본 연구의 프레임워크와 사용한 방법론에 대해 설명한다. 제4장에서는 연구 결과에 대해 설명하며, 제5장에서는 연구의 결론과 시사점, 그리고 한계점을 제시한다.

II. Related works

1. Trends in smart port technologies

스마트 항만 기술은 1960년대 전자 선적 문서와 전자 데이터 교환 체계의 도입으로 시작되었다. 이후 1990년대에는 레이저, RFID, GPS 등의 기술이 적용되며 항만 자동화 시스템이 현실화되었다. 2010년 이후로는 클라우드 플랫폼, 인터넷, 인공지능 등의 기술이 활용되면서 빅데이터의 축적과 비즈니스 의사결정 지원이 가능해졌다[11].

네덜란드의 로테르담 항은 'Smart Port' 조직을 설립하여 물류, 인프라, 지속 가능한 운영 전략 등 다양한 스마트 항만 프로젝트를 추진하고 있으며, Pronto 플랫폼을 통해 운송회사, 에이전트, 터미널 등과의 실시간 데이터 공유를 가능하게 하였다[12]. 이는 항만 운영의 효율성 증대와 비용 절감에 기여한다. 영국은 2019년 'Maritime 2050'을

발표하여 로봇과 AI 기술의 적용을 통한 스마트 항만 시스템 도입의 필요성을 강조하였으며, 이를 통해 운영 비용의 절감과 환경오염에 영향을 미치는 요인의 개선을 기대하고 있다[13].

국내에서는 부산항이 스마트 항만 구축의 선도적인 역할을 하고 있다. 2023년 완공된 부산 신항의 완전 자동화 부두에서는 무인 크레인이 선박의 컨테이너를 자율운송 차량에 실어 야적장의 크레인으로 옮기는 자동화 시스템이 운영되고 있다[14]. 또한 디지털 트윈 기술을 도입하여 항만 운영의 효율성을 17% 향상시키고, 선박 배출 탄소를 33% 저감하는 효과를 확인하였다. 현재는 이 기술의 적용 범위를 부산항 전체로 확대하기 위한 통합 시스템을 구축하고 있다[15]. 이는 국내의 항만 디지털화와 자동화 기술의 발전이 특정 지역에 국한되지 않고, 국내에서 통합적으로 운영됨으로써 글로벌 경쟁력을 확보할 수 있다는 가능성을 보인다.

이에 본 연구는 항만 내의 시스템 일부뿐만 아니라 항만 전체가 스마트화될 수 있도록 스마트 항만 기술 발전의 트렌드를 분석하고자 한다. 이를 위해 특허 데이터를 활용하여 스마트 항만 기술의 주요 기술 영역과 흐름을 체계적으로 분석하고, 항만 기술 발전을 위한 구체적인 방향성을 제안하고자 한다.

2. Technological forecasting using patent analysis

특허란 공인된 정부 기관 또는 지정된 권한에 의해 발행되는 문서로서, 특허권을 가진 발명자 또는 양수인은 해당 장치 또는 기술에 대해 타인의 생산 및 사용권의 통제 권한을 갖는다[16]. 특허의 내용은 원칙적으로 공개되어야 하므로 분석을 위한 지표로도 사용될 수 있다. 더욱이 특허는 많은 양의 기술 정보를 포함하며, IPC(International Patent Classification) 코드와 같은 표준화된 시스템에 따라 분류되어 있어 기술의 분석이 용이하다.

과거에는 특허 정보를 기술 혁신에 대한 연구나 경쟁 기업 모니터링, 기술 포트폴리오 측정 등에 활용하였다. 오늘날 글로벌 경쟁이 심화되고 기술이 빠르게 진화함에 따라 이러한 변화를 모니터링하는 것은 기업에게 있어 필수적인 요소가 되었다[17]. 특허 정보를 활용하여 기업이나 국가의 경쟁력을 넘어 산업과 기술의 동향을 파악하고 이해하는 것을 목적으로, 다양한 연구가 활발히 진행되고 있다[18]. 미국 웰니스케어 산업에서는 기술의 잠재력을 예측하기 위해 미국 특허청(USPTO)의 특허 데이터를 수집하였다. 특허 지표를 분석하고 기술 매핑을 수행하여 향후 기술 발전

방향에 대한 연구를 수행하였다[19]. 또한, 특허 데이터를 인공지능으로 학습한 챗봇을 출시하여 연구 토픽의 분포를 파악하고, 기술 발전의 동향을 분석하였다[20].

III. Methodology

1. Research framework

본 연구는 Fig. 1과 같이 진행된다. 먼저, 스마트 항만 기술과 관련된 특허 데이터를 수집하고 데이터 전처리를 수행한다. 전처리된 데이터를 바탕으로 LDA 기반의 토픽 모델링을 적용하여 스마트 항만과 관련된 핵심기술을 추출한다. 마지막으로, 핵심기술로 정의된 토픽에 대해 GTM을 적용하여 스마트 항만 기술에 대한 기술 맵을 생성한 뒤, 공백기술을 추출하고 해석한다.

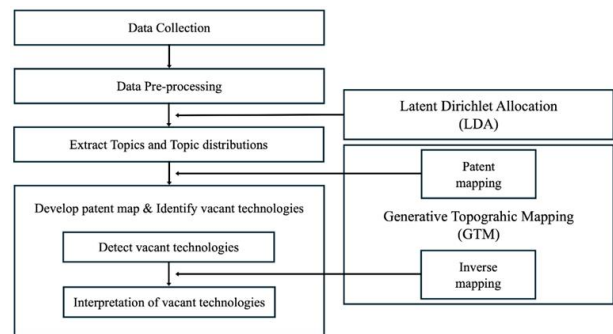


Fig. 1. Research Framework

2. Data collection

KIPRIS를 통해 수집한 2004년 3월 5일부터 2024년 7월 2일까지 한국에 등록된 특허를 분석 대상으로 한다. 선행 연구 논문 및 R&D 보고서를 바탕으로 스마트 항만 기술과 관련된 키워드를 추출하여 정의한 검색식으로 특허 데이터를 추출하였다. 특허 데이터는 키워드를 기반으로 수집되기 때문에 특정 키워드에 대해 과도한 데이터 집중이 발생하여 불안정한 데이터 수집이 이루어질 수 있다. 이를 보완하기 위해 추출된 raw data에 대해서 발명자, 발명의 명칭, 요약을 기준으로 중복된 데이터를 제거하였다. 또한 데이터 분석의 신뢰성을 높이기 위해 특허 데이터의 정보 중, 비어있는 열이 존재하는 데이터를 제거한 뒤, 청구항에 대해서 전수조사를 실시하였다. 이 과정에서 스마트 항만 기술과 무관한 데이터를 식별하여 제거함으로써 데이터의 정확성을 높였으며, 최종적으로 총 401건의 유효 특허 데이터를 분석에 활용하였다.

3. Data preprocessing

본 연구에서는 특허의 핵심 내용을 기술한 청구항을 사용한다. 한글 텍스트로 이루어진 비정형 데이터를 분석하기 위해 파이썬 KoNLPy 라이브러리의 Komoran 형태소 분석기를 이용하였다. Komoran의 형태소 분석의 성능 향상을 위해 ‘기계학습’, ‘인공지능’, ‘자율운항선박’과 같이 청구항 데이터에 자주 등장하는 복합명사들을 사전에 추가하였다. 이는 형태소 분석 과정에서 띄어쓰기가 반영되어 발생할 수 있는 복합명사 분리 문제를 해결하기 위함이다. 예를 들어, ‘기계학습’이라는 복합명사에 대해 komoran 형태소 분석기를 적용하면 ‘기계’와 ‘학습’으로 분리되기 때문에, 단일 의미 단위로써의 ‘기계학습’에 대한 정확한 분석이 어렵다. 이러한 문제는 본 연구에서 기대하는 분석의 성능을 저하시키는 결과를 초래할 수 있다. 이를 해결하고자 복합명사에 대해 띄어쓰기를 제거한 형태로 사전에 추가하였고, 이를 통해 형태소 분석의 전반적인 성능을 개선시키고자 하였다.

청구항에 등장하는 기술명은 한글과 영어가 병기되는 경우가 있다. 그러나 KoNLPy 라이브러리는 한국어만 처리할 수 있기 때문에 영어로 작성된 기술명을 분석하기 어렵다. 본 연구에서는 한글과 영어로 작성된 기술명 모두를 효과적으로 반영하기 위해서 다음과 같이 전처리 과정을 진행하였다. 먼저, 영어 단어를 별도로 추출하고 영어 불용어를 제거하여 보존하였다. 이후, 한글과 공백을 제외한 모든 문자와 중복된 공백을 제거하였다. 이를 바탕으로 형태소 분석과 명사 추출을 진행하였으며, 추출된 결과에서 한글 불용어를 제거하였다. 마지막으로 이전 단계에서 추출한 영어 단어와 한국어 명사를 결합하여 최종적으로 분석에 사용할 입력 데이터를 생성하였다. 이와 같은 과정은 영어와 한국어가 혼재된 청구항 데이터를 보다 정확하게 분석할 수 있도록 하며, 기술명에 대한 분석의 타당성을 확보하는데 기여한다.

4. Topic modeling

4.1 TF-IDF

TF-IDF는 문서 내 단어의 빈도와 역문서 빈도를 통해 단어의 가중치를 계산하여 해당 단어의 중요도를 파악하는 기법이다[21]. 이는 특정 문서에서 단어의 출현 빈도가 높을수록 가중치는 증가하지만, 전체 문서 집합에서 단어 출현 빈도가 높을수록 그 값이 상쇄된다. 이를 통해 전체 문서에 자주 나타나는 단어의 중요도를 낮추고, 특정 문서에만 자주 나타나는 단어의 중요도를 상대적으로 높여서 특허 분석 과정에서 특정 키워드의 잠재적 편향성을 방지할 수 있다. TF-IDF를 통해 문서로부터 단어의 가중치를

계산하는 수식은 식 (1)과 같다.

$$w_{i,j} = \frac{tf_{i,j}}{C_d} \times \log\left(\frac{n}{df_i}\right) \tag{1}$$

Table 1. Notation of TF-IDF

Symbol	Definition
$w_{i,j}$	Weights of words t_j in document d_i
$tf_{i,j}$	Number of occurrences of word t_j in d_i
df_i	Number of documents include word t_j
n	Number of documents

4.2 LDA

LDA 알고리즘은 문서-단어 행렬을 생성하여 문서 내 단어의 중요도를 계산한다. 이후 Gibbs 샘플링 방법을 사용하여 문서 내 단어의 Dirichlet 분포를 통해 토픽을 선정하는 방법이다[22].

LDA 알고리즘은 Fig. 2의 과정으로 진행된다. Fig. 2의 외부 사각형은 문서를 나타내며, 내부 사각형은 문서 내에서의 토픽과 단어의 반복적 선택을 나타낸다. LDA 알고리즘의 목표는 관찰된 단어인 $W_{d,n}$ 를 통해 잠재 변수 θ_d , $Z_{d,n}$, ϕ_k 를 찾는 것이다. 해당 알고리즘의 파라미터 α 는 θ_d 에 관한 것이며, Dirichlet 분포를 따른다. 문서 d 의 n 번째 단어에 대한 토픽 할당 $Z_{d,n}$ 은 θ_d 와 $W_{d,n}$ 에서 파생되며, 전체 토픽은 θ_d 로부터 파생된다.

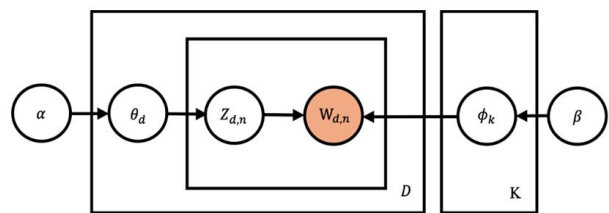


Fig. 2. LDA process

Table 2. Notation of LDA

Symbol	Definition
D	Number of documents
d	Number of documents of each topic
k	Number of topics
ϕ_k	Word frequency of n th topic
θ_d	Topic ratio of document d
$Z_{d,n}$	Topic assignment for n th word of document d
$W_{d,n}$	n th word appearing in document d
α	Parameters for topics per document
β	Parameters for words per topic

언어 모델의 perplexity 알고리즘에 따라, 토픽 모델링을 통해 생성할 토픽의 개수 k 값을 결정한다. Perplexity는 모델이 실제 확률 분포에 얼마나 적합한지를 측정하는 정량적 지표로, 모델의 일반화 능력을 나타내기도 한다 [23]. Perplexity 값이 작을수록 모델의 안정성이 높아지며, 이는 더 나은 주제 탐지 결과와 함께 시간 및 공간적 비용의 감소를 가져온다. 본 연구에서는 식 (2)를 통해 perplexity 값을 계산하였다.

$$\begin{aligned} & \text{Perplexity}(W) \\ &= \exp\left(-\frac{\sum_{d=1}^M \sum_{w_i \in d} \log\left(\sum_{p \in K} P(z_p|d_j)P(w_i|z_p)\right)}{\sum_{d=1}^M N_d}\right) \end{aligned} \quad (2)$$

Table 3. Notation of Perplexity

Symbol	Definition
N_d	Number of words
M	Size of document d
N	Size of word set
K	Size of topic set
$P(z_p d_j)$	Probability of topic Z_p appearing in document d_j
$P(w_i z_p)$	Probability of word w_i appearing in topic Z_p

Perplexity는 모델의 성능을 평가하는 정량적 지표이지만, 해석 가능성을 보장하지 않는다는 한계가 있다. 이를 보완하기 위해 coherence score가 사용된다. Coherence Score는 토픽 내 상위 단어들이 얼마나 의미적으로 밀접하게 관련되어 있는지를 평가하는 질적 지표이며, 이는 분포 가설에 기반한다[24].

Coherence Score는 식 (3)과 같이 Normalized Pointwise Mutual Information (NPMI)을 기반으로 계산되며, 상위 N_{top} 단어 쌍 간의 상호 정보를 정규화한 값을 사용한다.

$$\begin{aligned} & \text{NPMI}(w_i, w_j) \\ &= \frac{\log(P(w_i, w_j)) - \log(P(w_i) \cdot P(w_j))}{-\log(P(w_i, w_j))} \end{aligned} \quad (3)$$

$$\begin{aligned} & \text{Coherence}(C_v) \\ &= \frac{1}{|N_{\text{top}}|} \times \sum_{w_i, w_j \in W, i < j} \text{NPMI}(w_i, w_j) \end{aligned} \quad (4)$$

Table 4. Notation of Perplexity

Symbol	Definition
w_i, w_j	Words i and j
$P(w_i, w_j)$	Probability of words w_i and w_i, w_j co-occurring
$P(w_i)$	Probability of word w_j occurring
$P(w_j)$	Probability of word w_i, w_j occurring
$\text{NPMI}(w_i, w_j)$	Normalized Pointwise Mutual Information between words w_i , and w_j
N_{top}	Number of top words in a topic
W	Set of words in a topic
$i < j$	Condition where i precedes j in the word pair w_i, w_j

식 (4)에 따라 coherence score C_v 는 상위 N_{top} 단어 쌍의 평균 NPMI값을 통해 산출된다. Coherence Score는 값이 높을수록 토픽 내 단어들이 의미적으로 더 밀접하게 연결되어 있음을 의미하며, 모델의 해석 가능성을 높이는 중요한 지표로 작용한다.

따라서 perplexity와 coherence score를 함께 고려하여 최적의 토픽 수 k 를 결정하면, 모델의 설명력과 인간의 해석 가능성을 동시에 충족할 수 있다. 본 연구에서는 perplexity와 coherence score를 비교 분석하여 최적의 k 값을 설정하였으며, 이를 바탕으로 생성된 토픽의 질적 유효성을 확보하였다.

5. GTM

GTM은 데이터의 시각적 탐색과 고차원 데이터의 차원 축소를 위한 확률 모델 기반의 비선형 매핑 기법이다. 기존의 SOM(Self-Organizing Map), PCA(Principal Component Analysis) 기반의 공백기술 추출 기법은 전문가의 주관적인 판단에 의존하여 공백 영역을 식별하게 된다. 또한, 기술 맵의 공백 영역에 대한 기술 해석을 위해 많은 시간과 노력이 소요되고, 연구자에 따라 상이한 결과가 도출될 수 있다[25]. 따라서, 본 연구에서는 공백기술을 식별하기 위해 GTM을 사용한다.

GTM은 식 (5)과 같이 잠재 변수 x , 가중치 행렬 W , 노이즈 β 를 통해 추정된 확률 분포를 데이터 공간으로 매핑한다. 여기서 x 는 잠재 공간(R^L)에서 차원이 축소된 데이터 벡터를 나타내고, t 는 데이터 공간(R^D)에서 관측된 데이터 벡터를 나타낸다. 변화된 잠재 데이터 $y(x, W)$ 와 관측 데이터 t 사이의 거리 측면에서의 근접성을 식별하기 위해 식 (3)과 같이 가우시안 혼합 분포를 사용한다. 데이터 공간에서 데이터 t 가 $y(x, W)$ 에 가까우면 확률

$p(t|x)$ 가 증가한다.

$$\begin{aligned}
 p(t|x, W, \beta) &= N(y(x, W), \beta) \\
 &= \left(\frac{-\beta}{2\pi}\right)^{-D/2} \exp\left(\frac{-\beta}{2} \sum_d (t_d - y_d(x, W))^2\right)
 \end{aligned}
 \tag{5}$$

Table 5. Notation of GTM

Symbol	Definition
$y(x; W)$	Function that convert x into data space
x	Latent variables
t	Data variables
D	Dimension of t
W	$D \times M$ weights matrix
β	Noise
K	Number of grid points
x_K	Grid points in latent space
$\phi(x)$	Fixed basis function of M latent variables

그리고 식 (6)와 같이 t -공간에서의 데이터 분포가 x -공간의 데이터 분포를 기반으로 한 적분으로 표현된다.

$$p(t|W, \beta) = \int P(t|x, W, \beta)p(x)dx
 \tag{6}$$

그러나 $p(t)$ 는 연속적인 분포이기 때문에 추론하기 어렵다. 이를 해결하기 위해 $p(x)$ 를 여러 개의 이산형 가우시안 분포의 혼합으로 근사한다. GTM은 t -공간과 x -공간 사이의 매핑을 위해 함수 $y(x; W)$ 를 적용한다.

Fig. 3은 잠재 공간과 데이터 공간에서 3×3 격자를 나타낸 것이다. 녹색의 구는 잠재 공간의 각 격자점이 데이터 공간으로 매핑한 결과를 나타낸다. 해당 과정에서는 데이터의 복잡한 고차원 구조를 저차원 공간에 투영하며, 가우시안 분포의 확률적 모델을 통해 잠재 공간에 데이터를 매핑한다.

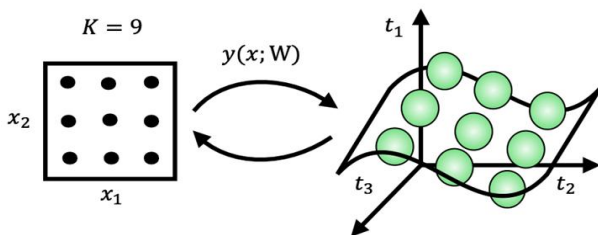


Fig. 3. Example of data mapping from latent space grid to data space

잠재 공간의 격자점에 대해 가우시안 혼합 모델을 적용한 데이터 공간에서의 확률 분포는 아래의 식 (7)와 같이 표현된다.

$$p(t|W, \beta) = \frac{1}{K} \sum P(t|x_K, W, \beta)
 \tag{7}$$

이때, 가중치 행렬과 노이즈를 포함한 모수들은 EM(Expectation-Maximization) 알고리즘을 사용해 추정된다. GTM을 데이터 집합에 적용하면, 식 (8)과 같이 베이지 정리를 사용하여 잠재 데이터에 의해 생성된 데이터 포인트의 확률을 추정한다. 이 과정을 통해 관측된 데이터 포인트를 잠재 변수에 할당한다.

$$p(x_K|t_n, W, \beta) = \frac{p(t_n|x_K, W, \beta)p(x_K)}{\sum_{K'} p(t_n|x_{K'}, W, \beta)p(x_{K'})}
 \tag{8}$$

마지막으로 관찰된 데이터는 잠재 공간으로 매핑하여 공백기술을 확인할 수 있다. 공백기술의 의미를 파악하기 위해 식 (9)을 통해 역매핑을 수행한다.

$$y(x, W) = \phi(x) W
 \tag{9}$$

공백 영역의 좌표를 초기에 매핑할 때 사용한 기저함수의 입력 값을 받아 변환 공식을 적용하면 해당 영역을 고차원으로 표현하는 벡터를 얻을 수 있다.

앞서 도출한 토픽 모델링의 결과를 바탕으로 원본 데이터의 차원으로 투영하여 각 공백기술을 정의하고 의미를 해석한다.

IV. Results

1. LDA

LDA 기반의 토픽 모델링을 적용하기에 앞서 언어 모델의 perplexity와 coherence score를 통해 토픽 모델링의 적절한 토픽 수 k 를 결정하였다. Perplexity는 모델이 주어진 데이터 세트를 얼마나 잘 설명하는지에 대한 척도로, Fig. 4는 k 값에 따른 perplexity 결과를 나타낸다. Perplexity 값은 토픽 수가 증가함에 따라 감소하는 경향을 보이지만, 7, 8, 9 지점에서 감소 폭이 완만해지는 것을 확인할 수 있었다. Perplexity는 모델의 정량적 설명력에

초점을 맞추는 반면, coherence score는 각 토픽에서 추출된 키워드 간의 일관성을 평가하는 질적 지표이다. Fig. 5는 k 값에 따른 coherence score 결과를 나타낸다. Coherence Score는 토픽 수 7에서 가장 높은 값을 보였으며, 이는 7개의 토픽이 키워드 간의 의미적 일관성이 가장 높음을 의미한다. Perplexity와 Coherence Score를 종합적으로 고려한 결과, 토픽 수 7이 모델의 설명력과 질적 유효성을 동시에 만족하는 최적의 값으로 확인되었다. 최적의 토픽 수 선정 이후, 5,000회의 반복 샘플링을 통해 최종적으로 7개의 주요 토픽을 도출하였다.

각 토픽별 빈도 수를 기준으로 추출한 상위 단어에 대해 토픽명을 정의한 결과는 Table 6과 같다. 그 결과 Topic 0은 컨테이너 내부 스캐닝 시스템, Topic 1은 항만 시뮬레이션 기술, Topic 2는 IoT 기반의 스마트 컨테이너 기술, Topic 3은 선적 계획 수립 알고리즘, Topic 4는 객체 탐지 기술, Topic 5는 원격제어 기반의 자율운항선박 기술 그리고 Topic 6은 항만 내 차량의 자율주행 기술로 정의된다. 각 토픽명은 각 키워드와 관련된 상위 특어를 바탕으로 정의하였다.

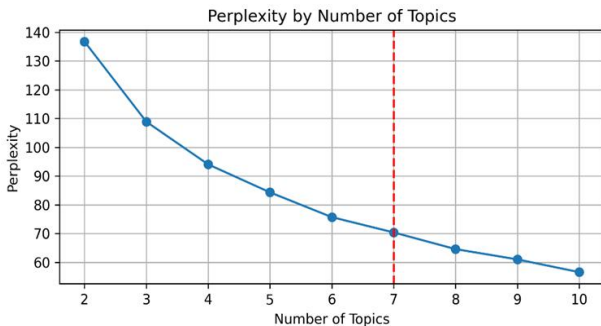


Fig. 4. Graph of Perplexity for optimal k

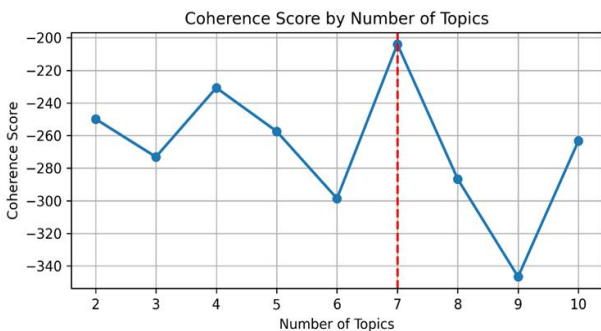


Fig. 5. Graph of Coherence score for optimal k

Table 6. Result of LDA topic modeling

No	Topic
Topic 0	Container internal scanning system
Topic 1	Port simulation technology
Topic 2	Smart container technology
Topic 3	Shipping planning algorithm
Topic 4	Object detection algorithm
Topic 5	Remote-controlled autonomous ship technology
Topic 6	Autonomous driving technology

1.1 Core technologies

Topic 0은 스캐닝 기술을 적용하여 컨테이너 내부를 검사하는 시스템이다. 과거에는 작업자가 컨테이너를 직접 열어서 모든 수화물을 육안으로 검사하였기 때문에 수화물의 훼손이 자주 발생하였으며, 컨테이너 반입 및 처리 시간이 지체되어 항만 운영의 병목현상을 초래하였다. 이러한 문제점을 개선하기 위해 x-ray, 방사선 등의 데이터를 융합하는 지능형 기술을 적용하여 컨테이너 내부 수화물을 검사하는 기술이 개발되었다. 이를 통해 작업자의 시각에 의존했던 검사 과정을 자동화함으로써 인건비 절감과 동시에 검사 능력 수준을 향상시켜서 효율적인 보안 및 불법성 관리를 가능하게 하였다[26].

Topic 1은 항만 시설과 자동화 장비, 이들의 운영 및 통제를 위한 정보시스템을 연계한 시뮬레이션 기술이다. 이 기술은 장비 구성, 운영, 신규 장비 도입 등에 따른 다양한 시나리오를 가상으로 구현함으로써 관리자의 의사결정을 지원한다[27]. 특히, 항만 시설 및 장비에 다양한 센서가 도입되면서 디지털 트윈 기반의 시뮬레이션 기술로 고도화되고 있다.

Topic 2는 IoT 및 센서 등을 통해 컨테이너의 위치, 내부 온도 및 습도, 외부 충격 등의 다양한 정보를 실시간으로 모니터링하고 데이터를 전송하는 스마트 컨테이너 기술이다. 이에 따라 선사는 실시간으로 화물 상태 추적 및 관리가 가능하여 공급망 관리의 효율성을 높이고 서비스 품질을 향상시킬 수 있다[28]. 스마트 컨테이너는 선박과 항만 간 그리고 항만과 항만 간에 단절된 정보를 유기적으로 연계하며, 항만물류의 전반적인 효율성 향상을 위한 핵심기술로 자리 잡고 있다.

Topic 3은 선적 계획 수립과 관련된 지능형 알고리즘 기술이다. 이 기술은 화물의 특징을 반영하여 선박의 공간을 효율적으로 사용하고, 화물의 선적지와 하적지를 고려하여 효율적인 선·하적 동선을 확보함으로써 물류비용을 최소화한다. 과거엔 휴리스틱 알고리즘을 통해 학습 및 예

측하고 수학적 모델링을 적용하여 최적의 해를 도출함으로써 효율적인 선적 계획을 수립하였다[29-31]. 점차 인공지능 기술이 발전함에 따라 인공지능망과 결합한 선적 계획 알고리즘이 개발되고 있다[32].

Topic 4는 항만의 보안 구역 감시 및 위험 감지를 위한 객체 탐지 기술이다. 카메라, 센서 등을 통해 관리자가 외부인 및 외부 차량의 위치를 식별하고 위험 구역으로의 출입 여부를 파악함으로써 항만의 안정성을 높인다. 특히, 컴퓨터 비전 분야의 폭발적인 발전으로 딥러닝 기반의 객체 탐지 알고리즘이 다양한 산업에 적용되고 있다[33]. 항만의 위험 구역에 설치된 CCTV에 객체 탐지 기술을 적용하여 근로자와 외부인을 구분하고 안전 장비 착용의 여부를 감지하는 등으로 개발이 이루어지고 있다.

Topic 5는 자율운항선박의 상태를 모니터링하여 충돌 등의 위험 상황이 발생하면 사람이 개입하여 원격으로 제어하는 기술이다. 자율주행과 관련된 다양한 기술이 개발됨에 따라 자율운항 기술도 발전하고 있다. 선박의 자율운항 단계는 4단계로 이루어지며, 4단계에 해당하는 완전 무인 자율운항선박을 개발하고 상용화하기 위해서는 본 기술이 필수적이다. 특히, AIS(Automatic Identification System), VDES(VHF Data Exchange) 등의 대량으로 생성된 데이터를 클라우드로 관리하고, 원격제어를 위한 인공지능 기반의 추론 알고리즘이 도입됨에 따라 의사결정 속도가 점차 빨라지고 있다[34].

Topic 6은 항만 내 차량에 자율주행 기술을 적용하여 항만 운영 및 관리를 지원하는 기술이다. 이 기술은 야드 트랙터와 로드트랙터 등 항만 내에서 주행하는 차량에 무선 통신 기반의 자율주행 기술을 접목함으로써 주변 상황을 인식하여 스스로 목적지까지의 최적 경로를 찾아 주행할 수 있게 한다. 이동 시 차량 및 물체의 위치 인식과 장애물 인식 및 회피를 수행함으로써 작업의 안전성과 효율성을 높이는 데 기여한다.

2. GTM

GTM은 다차원 데이터 공간을 저차원 잠재 공간에 매핑하고 역매핑이 가능한 확률적 접근 방법이다. 본 연구에서 GTM의 하이퍼파라미터로 k 는 9로 설정하였으며, k 는 GTM이 사용하는 가우시안 모델의 성분 수를 의미한다.

본 연구에서는 토픽 모델링 결과에 대하여 GTM 기반의 기술 맵을 생성하고 공백기술을 도출하였다. 공백기술이란 잠재적 개발 가치가 있는 기술로 새로운 기술 개발을 위한 중요한 기회를 제공하며 연구개발과 투자를 통해 기술적 혁신의 가능성을 가진 기술을 의미한다. GTM 기반 스마

트 항만의 기술 맵은 LDA 기반의 토픽 모델링으로 도출된 토픽에 포함되는 확률을 바탕으로 나타낼 수 있다. GTM으로 도출된 기술 맵은 공백기술을 식별하기 위해 유용하며, 2차원 공간에서 분포를 시각화하는 것이 필수적이다. GTM 기반의 기술 맵은 Fig. 6, Fig. 7과 같이 생성된다.

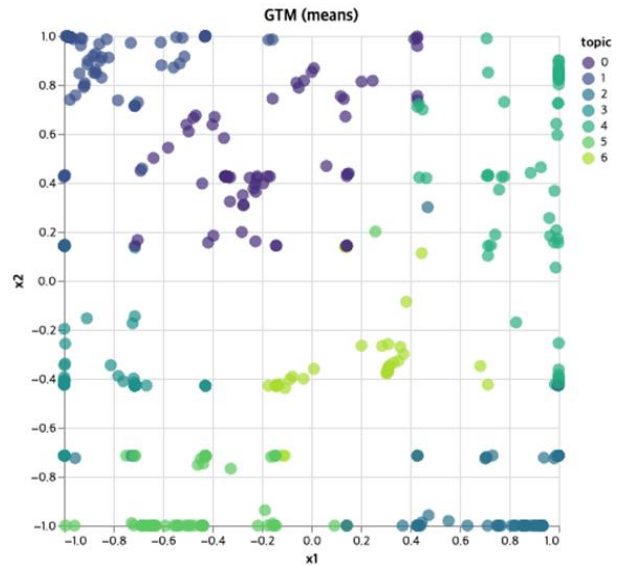


Fig. 6. Technology map(means)

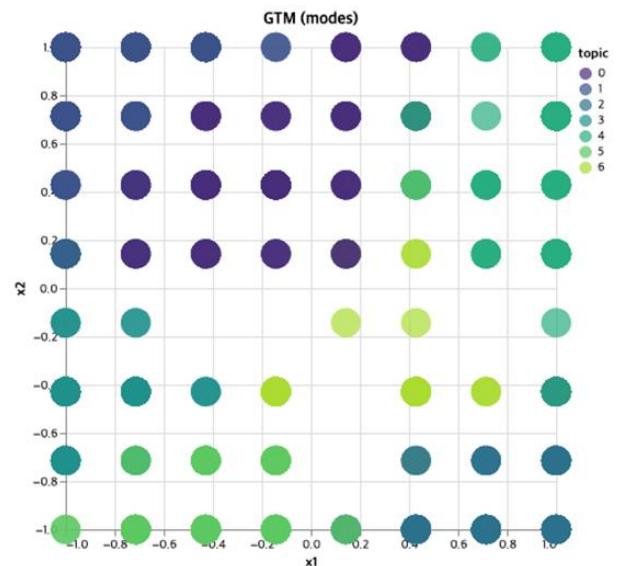


Fig. 7. Technology map(modes)

Fig. 6은 사후 평균 투영을 보여준다. 이는 특히 데이터를 잠재 공간상의 위치로 매핑했을 때, 특히의 원래 위치와 잠재 공간 내 위치 사이의 상대적인 거리를 나타낸다. 그러나, 이는 공백 영역을 명확하게 식별하지 못한다는 어려움이 존재한다. 이와 달리, Fig. 7은 동일한 데이터를 격자 기반으로 재구성하여 나타낸 것으로, Fig. 6에서의 연

Table 7. Weights for each topic per vacuum area

	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
Vacuum space 1	0.33604	0.05437	0.05624	0.2858	0.08706	-0.1298	0.31037
Vacuum space 2	0.27760	0.02376	0.03639	0.10331	0.03054	-0.0807	0.60899
Vacuum space 3	0.03521	0.04282	0.2230	0.02315	-0.0228	0.21568	0.48292
Vacuum space 4	0.00097	-0.0159	0.02202	0.04214	-0.0682	0.02736	0.99157
Vacuum space 5	0.00304	0.17278	-0.1014	0.03845	0.41465	0.05138	0.42108

Table 8. Result of topic assignment to vacuum area

	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
Vacuum space 1	0	0	0	0	0	0	0
Vacuum space 2	0	0	0	0	0	0	1
Vacuum space 3	0	0	0	0	0	0	0
Vacuum space 4	0	0	0	0	0	0	1
Vacuum space 5	0	0	0	0	0	0	0

속적인 분포를 격자점에 매핑함으로써 구조화된 형태를 제공한다. 데이터 포인트들이 잠재 격자점에 매핑된 사후 모드 투영을 보여주며, 각 데이터의 최빈값 위치를 계산하여 매핑하였다. 해당 기술 맵은 각 특허 데이터가 잠재 공간의 격자점에 적절히 매핑되어 공백 영역을 보다 명확하게 표현한다. 따라서, Fig. 8과 같이 본 연구에서는 최빈값을 이용한 기술 맵을 통해 총 64개의 격자 포인트 중 5개의 공백 영역을 추출하였다.

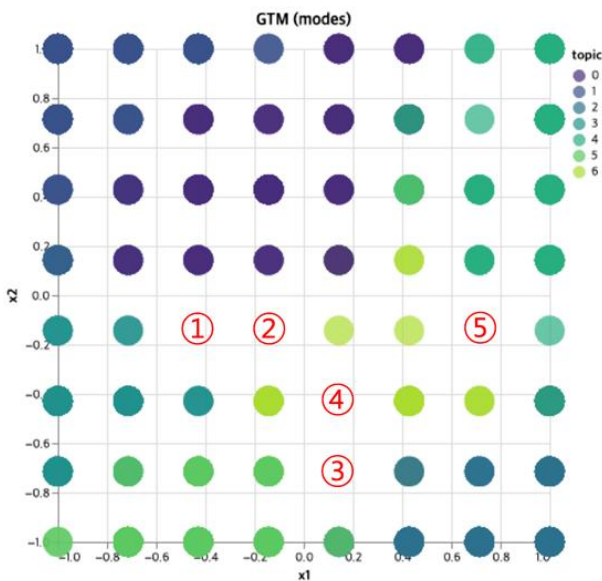


Fig. 8. Vacuum areas in technology map

2.1 Vacant technologies

공백 영역의 역매핑을 통해 의미를 해석할 수 있다. 식별한 각 공백 영역에 대하여 영역별로 각 토픽에 대한 가

중치(W)를 Table 7에 나타내었다. 이러한 가중치는 LDA에서 도출한 토픽들을 GTM의 각 노드에 매핑한 가중치를 표현한다. 공백 영역별로 할당된 토픽의 가중치 절댓값이 클수록 해당 토픽과 관련성이 높음을 의미한다.

공백 영역을 해석하기 위해 각 노드별 토픽 벡터의 분산을 분석하였다. 노드별 토픽 벡터는 각 노드가 다양한 토픽에 할당된 가중치 값들로 구성되며, 분산이 높을수록 특정 토픽에 대한 가중치 값이 두드러지게 나타나 특정 토픽과의 연관성이 명확해진다. 반대로 분산이 낮을수록 가중치 값이 비슷하게 분포되어 특정 토픽에 대한 할당이 불분명해지는 경향을 보인다.

이를 기반으로 상대적으로 높은 가중치 값의 집중도를 반영하기 위해 임계값을 0.5로 설정하였다. 이는 특정 토픽에 대한 가중치가 뚜렷하게 구분되는 기준으로, 공백 영역에 대한 명확한 해석을 가능하게 한다. Table 8에 제시된 결과와 같이 설정된 임계값을 기준으로 가중치를 이진화하여 표현하였으며, 이를 통해 공백 영역 2와 4가 Topic 6과 연관되어 있음을 알 수 있다.

Topic 6의 항만 내 차량 자율주행 기술과 관련하여 이 기술이 적용될 경우 운송작업의 효율성 및 안정성을 보장할 수 있을 뿐만 아니라 작업자의 작업 효율성, 안전성, 그리고 에너지 절감에도 기여할 것으로 기대된다[31]. 정부는 지난 2019년도 '스마트 해운항만물류 시스템' 구축 전략에 따라 5G 기반의 스마트 해상물류사업을 본격화하였다[36]. 이후 2024년 해양수산부가 자동화 야드크레인을 시작으로 광양항에 항만 자동화 테스트베드 구축 사업에 본격 착수하였으며, 정부와 여수광양항만공사는 2029년까지 완전 자동화 항만 하역 장비의 도입을 목표로 삼고 있

대[37]. 인천항에서는 '자율협력주행기반 화물운송시스템 개발' 프로젝트를 통해 level 4 수준의 자율협력주행 시스템을 개발함으로써 항만 경쟁력을 확보하려는 노력을 하고 있다[38]. 또한, 부산항은 국립한국해양대 RIS 스마트 항만물류사업단과 함께 AI 기반 무인자동화터미널 연구개발을 통해 부산항 첨단 자동화터미널 전환과 스마트 항만물류 혁신 인재 양성에 기여하고 있다[39]. 2024년 12월 항만 자동화와 항만 인프라 구축을 선도하는 글로벌엔씨 기업이 항만 내 차량사물통신(V2X) 기반 자율협력주행차의 단체표준을 획득하는 등 V2X 기반의 한국형 항만 자율주행시스템의 개발 또한 순조로이 이루어지고 있다[40].

그러나 자율주행을 위한 통신 기술은 다양한 장애물이 있는 환경과 수백 대의 차량이 밀집한 조건에서 성능이 저하된다는 문제점이 존재한다[41]. 따라서 다양한 구조물이 존재하고, 많은 차량이 운행되는 항만에 적용하기 위해서는 기술 개발이 요구된다. 또한, 항만 내 자율운행 차량의 에너지원으로 고려되고 있는 배터리는 에너지 밀도가 낮으며, 사용 가능한 전압의 범위가 제한적이다. 이러한 한계점을 해결하기 위해 국내외에서 관련 연구가 진행되고 있으나, 아직 초기 단계로서 각 요소 기술의 최적화가 정립되지 않은 상황으로 기술적 과제 해결이 요구된다[42]. 따라서 공백 영역으로 도출된 Topic 6은 여전히 활발한 기술 개발이 필요한 단계로 보이며, 이는 새로운 기술 개발을 위한 기회로 판단할 수 있다.

V. Conclusions

세계 무역의 성장과 제4차 산업혁명의 기술 혁신으로 스마트 항만으로의 변화가 중요해지고 있다. 본 연구에서는 이러한 기술 변화에 대한 트렌드 분석을 통해 핵심기술과 공백기술을 도출하고자 한다.

기술의 동향 분석 및 파악에 용이한 특허의 청구항 데이터에 TF-IDF를 적용하여 문서 내 빈도가 높은 용어를 파악하였고, 이를 바탕으로 LDA 기반의 토픽 모델링을 적용하여 총 7개의 유망기술을 도출하였다. Topic 0은 컨테이너 내부 스캐닝 시스템, Topic 1은 항만 시뮬레이션 기술, Topic 2는 IoT 기반의 스마트 컨테이너 기술, Topic 3은 선적 계획 수립 알고리즘, Topic 4는 객체 탐지 기술, Topic 5는 원격관제 기반의 자율운항선박 기술 그리고 Topic 6은 항만 내 차량의 자율주행 기술로 정의된다. 이후 GTM 기반의 기술 맵을 생성하여 5개의 공백 영역을 식별하였고, 최종적으로 Topic 6의 항만 내 차량의 자율

주행 기술이 공백기술임을 확인하였다. 항만 내 차량 자율주행 기술은 일련의 운송작업과 작업자의 작업 효율성 및 안정성, 에너지 절감에 기여할 수 있다. 이에 해당 기술의 상용화를 위한 다양한 노력이 이루어지고는 있으나, 통신 성능과 배터리 성능 저하 등 기술 개발의 어려움이 존재한다. 결론적으로 항만 내 차량 자율주행 기술은 기술 개발의 초기 단계에 머무르는 유망기술로서, 향후 기술의 고도화된 발전이 필요하다는 방향성을 제시하였다.

본 연구의 결과는 스마트 항만 기술과 관련된 다양한 기업의 기술 개발 전략 및 투자 계획 수립에 활용할 수 있을 것으로 기대한다. 또한, 핵심기술에 대한 투자를 지원함으로써 기업의 기술 경쟁력을 향상시킬 수 있고, 공백기술을 혁신의 기회로 포착하여 기술을 선제적으로 개발함으로써 산업 내 경쟁우위를 확보하고 시장을 선점할 수 있다.

그러나 본 연구는 다음과 같은 한계점을 가진다. 먼저 데이터 수집 범위를 국내 특허로 한정하였다. 글로벌의 측면에서 더욱 범용적인 유망기술을 도출하기 위하여 미국, 유럽 등의 국가 내 특허를 포괄한다면 더욱 신뢰도 높은 결과가 추출될 것으로 사료된다. 또한, 특허 데이터 외 연구 논문, 정책 보고서 등의 문헌을 포함하여 분석 시, 더욱 복합적인 인사이트를 도출할 수 있을 것이다. 더불어 본 연구에서는 수집한 모든 시간 내 특허를 동시에 활용하였으나, 향후 연구에서 연도별 시간을 나누고 각 기술에 대한 시간적 트렌드를 고려한 연구 방법론을 활용한다면 연구의 결과를 보다 고도화할 수 있을 것으로 기대한다.

REFERENCES

- [1] Review of Maritime Transport 2023, <https://unctad.org/publication/review-maritime-transport-2023>
- [2] S. W. Cho, H. J. Park, A. Kim, and J. H. Park, "GMM-based online optimization for container stacking in port container terminals," *Computers & Industrial Engineering*, Vol. 173, 2022. DOI: 10.1016/j.cie.2022.108671
- [3] M. Fruth, F. Teuteberg, and S. Liu, "Digitization in maritime logistics-What is there and what is missing?," *Cogent Business & Management*, Vol. 4, No. 1, 2017. DOI: 10.1080/23311975.2017.1411066
- [4] N. D. Cauwer, M. Fontanet, J. A. Garcia, H. Greven, J. S. Juhl, S. Probert, and Ø. J. Rødseth, "Maritime Informatics: Additional Perspectives and Applications," Springer International Publishing, pp. 61-77, 2021.
- [5] H. J. Park, S. W. Cho, A. Nanda, and J. H. Park, "Data-driven dynamic stacking strategy for export containers in container

- terminals,” *Flexible Services and Manufacturing Journal*, Vol. 35, No. 1, pp. 170-195, 2023.
- [6] T. Lee, "Smart Port Policy Trend of Europe and Singapore and Its Political Implications," *The Korea Port Economic Association*, Vol. 36, No. 1, pp. 77-90, 2020. DOI: 10.38121/kpea.2020.03.36.1.77
- [7] Build smart ports by 2030 to strengthen global competitiveness, <https://www.mof.go.kr/doc/ko/selectDoc.do?docSeq=36220&menuSeq=971&bbsSeq=10>
- [8] S. Won, and E. Lee, "A Study on Smart Technologies for Port-Centric Logistics," *Electronic Trade Institute*, Vol. 19, No. 2, pp. 1-25, 2021.
- [9] E. Lee, and S. Lee, "A Comparative Analysis on the Level of the Korean Smart Port in the Era of 4th Industrial Revolution," *The Korean Association of Shipping and Logistics*, Vol. 35, No. 2, pp. 323-348, 2019.
- [10] W. Ryu, and H. Nam, "A Study on Determining the Priority of Introducing Smart Ports in Korea," *The Korea Port Economic Association*, Vol. 40, No.1, pp. 31-59, 2024. DOI: 10.38121/kpea.2024.3.40.1.31
- [11] M. Anwar, L. Henesey, and E. Casalicchio, "Digitalization in container terminal logistics: A literature review," 27th annual conference of international association of maritime economists, Athens, pp. 1-25, 2019.
- [12] K. L. A. Yau, S. Peng, J. Qadir, Y. C. Low, and M. H. Ling, "Towards smart port infrastructures: Enhancing port activities using information and communications technology," *Ieee Access*, pp. 83387-83404, 2020. DOI: 10.1109/ACCESS.2020.2990961
- [13] Virtual Model (Digital Twin) Port Construction Begins in Full Swing, <https://www.mof.go.kr/doc/ko/selectDoc.do?docSeq=56229&menuSeq=971&pwise=mMain&bbsSeq=10&pWiseMain=N2>
- [14] M. Heikkilä, J. Saarni, and A. Saurama, "Innovation in smart ports: Future directions of digitalization in container ports," *Journal of Marine Science and Engineering*, Vol. 10, No. 12, 2022. DOI: 10.3390/jmse10121925
- [15] T. Inkinen, R. Helminen, and J. Saarikoski, "Technological trajectories and scenarios in seaport digitalization," *Research in Transportation Business & Management*, Vol. 41, pp. 100633, 2021. DOI: 10.1016/j.rtbm.2021.100633
- [16] P. Hanel, "Intellectual property rights business management practices: A survey of the literature," *Technovation*, Vol. 26, pp. 895-931, 2006. DOI: 10.1016/j.technovation.2005.12.001
- [17] P. I. Lee, and H. N. Su, "Framing patent indicators for innovation study," 2016 Portland International Conference on Management of Engineering and Technology (PICMET), pp. 1021-1030, 2016.
- [18] H. Fujii, and S. Managi, "Trends and priority shifts in artificial intelligence technology invention: A global patent analysis," *Economic Analysis and Policy*, Vol. 58, pp. 60-69, 2018. DOI: 10.1016/j.eap.2017.12.006
- [19] G. Kim, and J. Bae, "A novel approach to forecast promising technology through patent analysis," *Technological Forecasting and Social Change*, Vol. 117, pp. 228-237, 2017. DOI: 10.1016/j.techfore.2016.11.023
- [20] E. Pantano, and G. Pizzi, "Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis," *Journal of Retailing and Consumer Services*, Vol. 55, 2020. DOI: 10.1016/j.jretconser.2020.102096
- [21] S. Qaiser, and R. Ali, "Text mining: use of TF-IDF to examine the relevance of words to documents," *International Journal of Computer Applications*, Vol. 181, No. 1, pp. 25-29, 2018.
- [22] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, pp. 993-1022, 2003.
- [23] L. Huang, J. Ma, and C. Chen, "Topic detection from microblogs using T-LDA and perplexity," 2017 24th Asia-Pacific software engineering conference workshops (APSECW), pp. 71-77, Nanjing, China, Dec. 2017. DOI: 10.1109/APSECW.2017.11
- [24] S. Syed, and M. Spruit, "Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation," *International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 165-174, 2017. DOI: 10.1109/DSAA.2017.61
- [25] C. Son, Y. Suh, J. Jeon, and Y. Park, "Development of a GTM-based patent map for identifying patent vacuums," *Expert Systems with Applications*, Vol. 39, No. 3, pp. 2489-2500, 2012.
- [26] J. H. Hwang, Y. S. Kim, and J. N. Kim, "Freight Inspection System using X-ray Image for Automatic Port Logistics," *Korean Multimedia Society Academic*, pp. 627-630, 2006.
- [27] J. S. Lee, "Simulation Methodology for Automation of Port Systems : Example of Container Terminal," *Journal of Advanced Marine Engineering and Technology (JAMET)*, Vol. 34, No. 1, pp. 154-162, 2010. DOI: 10.5916/jkosme.2010.34.1.154
- [28] Study on ways to activate digital transformation of domestic container shipping companies, https://library.kmi.re.kr/%24/10140/contents/6088097?utm_source=chatgpt.com
- [29] H. J. Park, S. W. Cho, and C. Lee, "Particle swarm optimization algorithm with time buffer insertion for robust berth scheduling," *Computers & Industrial Engineering*, Vol. 160, 2021. DOI: 10.1016/j.cie.2021.107585
- [30] A. Kim, H. J. Park, J. H. Park, and S. W. Cho, "Rescheduling strategy for berth planning in container terminals: An empirical study from Korea," *Journal of Marine Science and Engineering*, Vol. 9, No. 5, 2021. DOI: 10.3390/jmse9050527
- [31] S. W. Cho, H. J. Park, and C. Lee, "An integrated method for berth allocation and quay crane assignment to allow for reassignment of vessels to other terminals," *Maritime Economics & Logistics*, Vol. 23, pp. 123-153, 2021.

- [32] S. H. Woo, H. J. Park, S. W. Cho, and K. H. Kim, "Proactive berth scheduling with data-driven buffer time in container terminals," *International Transactions in Operational Research*, Vol. 31, No. 5, pp. 2875-2902, 2024. DOI: 10.1111/itor.13412
- [33] A. Nanda, S. W. Cho, H. Lee, and J. H. Park, "KOLOMVERSE: Korea Open Large-Scale Image Dataset for Object Detection in the Maritime Universe," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 25, No. 12, pp. 20832-20840, 2024. DOI: 10.1109/TITS.2024.3449122
- [34] W. Lee, and S. W. Cho, "AIS trajectories simplification algorithm considering topographic information," *Sensors*, Vol. 22, No. 18, 2022. DOI: 10.3390/s22187036
- [35] B. Kim, E. Kim, W. Park, J. Seo, and S. Lee, "Autonomous Transport Robot Using Optimized Container Pickup Algorithm," *Korea Information Processing Society*, pp. 1006-1007, 2024.
- [36] M. Kang, "In the 5G era, the shipping and port industry map will be changed: Smart ports, autonomous ships, and high-speed maritime communication 'Commercialization Stretch'," *Maritime Korea*, Vol. 2019, No. 6, pp. 36-42, 2019.
- [37] Establishment of 'Automated Port' at Gwangyang Port... 746.4 billion investments by 2029, <https://biz.newdaily.co.kr/site/data/html/2024/08/12/2024081200013.html>
- [38] Development of autonomous cooperative driving-based cargo transportation system, <https://www.icpa.or.kr/mobile/article/view.do?articleKey=27029&boardKey=393&menuKey=3101¤tPageNo=1>
- [39] National Maritime University RIS Smart Port Logistics Project Group, Busan Port AI-based unmanned automated terminal research and development support agreement, <http://ihaesa.com/news/article.html?no=84178>
- [40] Korean Port Autonomous Driving System Development, <http://www.maritimepress.co.kr/news/articleView.html?idxno=323641>
- [41] H. S. Oh, D. W. Kang, S. Y. Lee, J. W. Lee, and Y. S. Song, "Current Status of V2X Communication Technology for Cooperative Driving," *The Journal of The Korean Institute of Communication Sciences*, Vol. 34, No. 6, pp. 20-26, 2017.
- [42] Supercapacitors: Next-generation, high-efficiency, high-capacity energy storage devices for renewable energy solutions, <https://ssl.pstatic.net/imgstock/upload/research/industry/1630547314317.pdf>

Authors



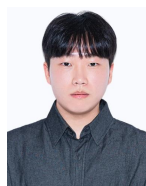
Hee-Kyeong Nam is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. She joined the Computational Optimization for Intelligent Systems Lab at

Dankook University, Cheonan, Korea, in 2023. She is interested in natural language processing(NLP), supervised learning, and data analysis.



Jeong-Yoon Hong received the B.S. degree in Management Engineering from Dankook University, Republic of Korea, in 2024. She joined the Computational Optimization for Intelligent Systems Lab at Dankook

University, Cheonan, Korea, in 2023. She is currently on a master's course in the Department of Industrial Engineering, Dankook University. She is interested in data analysis, supervised learning.



Min-Su Woo is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. He joined the Computational Optimization for Intelligent Systems Lab at Dankook

University, Cheonan, Korea, in 2024. He is interested in optimization in inventory management.



Go-Eun Han is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. She joined the Computational Optimization for Intelligent Systems Lab at Dankook

University, Cheonan, Korea, in 2023. She is interested in data-driven decision making models and object detection of computer vision.



Jae-Min Song is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. He is interested in production management and quality control.



Yul-Kyeong Yun is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. He is interested in IoT and data mining.



In-Yeong Lee is on a bachelor's course in the Department of Management Engineering, Dankook University, Republic of Korea. He is interested in artificial intelligence, IoT, and digital twin technology.



Sung Won Cho received the B.S. degree in Logistics from the Korea Aerospace University, Gyeonggi-do, Republic of Korea, in 2016, and the Ph.D. degree in Industrial and Management Engineering from Korea

University, Seoul, Republic of Korea, in 2021. From June 2019 to August 2023, he was a Senior Researcher at the Maritime Digital Transformation Research Center, Korea Research Institute of Ships and Ocean Engineering, Republic of Korea. Since 2023, he has been an Assistant Professor with the Department of Management Engineering, Dankook University. His research interests include maritime logistics, intelligent transportation systems and optimization methods.