

토픽모델링을 이용한 국내 산업경영공학 연구동향 분석

A Study on the Research Trends in Domestic Industrial and Management Engineering using Topic Modeling

김 상 겸* Kim, Sang Kyoum

장 성 용** Jang, Seong Yong

This study deals with analyzing the English abstracts from the Korea Management Engineers Society and other seven journals to identify the research trends in domestic industrial and management engineering. Topic modeling method based on LDA(Latent Dirichlet Allocation) was used to analyze abstract data from 2004 to 2015.

The results are as follows. First, 20 key topics including quality management, optimization, and system modeling are identified. Second, studies of healthcare and financial engineering analysis show upward trend, but those of process control and design of manufacturing systems show downward trend. Third, in analysis of network, studies on system modeling, supply chain management, and data mining are shown as the subjects with high connection centrality and they are thought to be essential themes in domestic industrial and management engineering. Also, the subject for evaluation of HW/SW reliability can be seen as an important one which plays a key role in the domestic industrial and management engineering since it shows higher betweenness centrality than the other subjects.

Keywords : research trends, text mining, topic modeling, LDA(Latent Dirichlet Allocation), domestic industrial and management engineering

이 연구는 서울과학기술대학교 교내연구비의 지원으로 수행되었습니다.

* 제1저자, 서울과학기술대학교 데이터사이언스학과, skkim85@seoultech.ac.kr

** 교신저자, 서울과학기술대학교 글로벌융합산업공학과 교수, syjang@seoultech.ac.kr

투고일 2016.06.16

수정일 2016.09.06

게재일 2016.09.30

1. 연구 배경 및 필요성

산업경영공학은(Industrial and Management Engineering)은 기존의 산업공학 영역에 경영학의 내용을 융합하여 공학적 기술과 경영에 대한 이해를 추구하는 학문이다. 국내의 대학에서는 산업정보시스템, 산업경영공학과 등의 명칭으로 학과가 개설되어 있으며 산업경영공학은 넓은 범위에서 산업공학이라는 학문에 포함되어 있다고 할 수 있기 때문에 연구의 범위를 확장하기 위해 산업공학이라는 큰 틀에서 연구를 수행한다.

산업공학은 우리 사회의 모든 산업에 관련되어 있으며 1차 산업인 농업, 2차 산업인 제조업, 3차 서비스 산업인 병원·은행·정부조직 등 광범위하게 응용되고 있다. 이러한 산업공학의 발전 배경에는 18세기에 나타난 산업혁명을 토대로 하고 있다. 산업혁명 이후 제품들의 상호교환과 노동의 전문화로 대량생산이 시작이 되면서 산업조직의 규모가 커지기 시작하였고 당시의 관리 방식은 소규모 조직에는 적합하였으나 기업 및 경영의 규모가 거대하고 복잡한 생산조직에는 적절하지 못하였다. 이로 인해 보다 나은 경영시스템이 필요하게 되면서 오늘날 산업공학이라고 명명되어진 학문이 탄생하게 되었다(한상찬, 2012).

한국교육개발원의 학과 전공 분류 자료집에 따르면 산업공학 학문은 시스템을 체계적으로 분석·설계하고 효율적인 방법으로 운영·개선하여 산업 및 공공조직의 경영합리화와 생산성 향상 및 경쟁력 강화에 기여하는 학문으로서 부존자원이 부족하고 기술수준이 정착되어 있지 않은 우리나라의 모든 산업계에 적용되어 생산성 제고와 기업 경영 합리화 및 기업 경쟁우위를 확보하기 위해 필요한 전략적 학문이다.

부존자원이 부족한 우리나라에서 기업의 경영성과를 향상시키기 위한 연구가 이루어져왔는데 차

영태(2015)는 경영혁신형 중소기업의 혁신활동이 기업의 경쟁우위와 경영성과에 미치는 영향에 대해 연구하였고 정대권과 윤원영(2015)은 중소 제조 기업의 현장 관리 및 개선 체계 구축에 대한 연구를 하였다. 또한 최민석 외(2011)는 기업의 교육서비스 품질이 구성원 만족과 경영성과에 미치는 영향에 관한 연구를 하였다. 이와 같이 기업의 경영활동과 관련하여 혁신을 통해 기업성과를 향상 시키는 연구가 수행되어 왔다.

국내 산업공학 학문의 역사는 1958년 한양대학교에 공업경영학과가 신설 되면서 시작되었다. 이후 건국대학교, 동국대학교, 동아대학교 등 국내의 많은 대학에 산업공학 관련학과가 신설되면서 국내 산업공학 학계는 수많은 학문적 발전을 이루었다. 대한산업공학회(대한산업경영공학 관련 전공 학과 현황에 대한 조사를 1994년과 2014년 두 차례 실시하였다. 2014년 기준으로 국내 산업공학 관련학과는 55개가 있으며 교수 수는 1994년에 321명에서 2014년에 476명으로 약 50% 정도 증가하였고, 학생 수는 입학 정원 기준으로 1994년에 3,050명에서 2014년 3,594명으로 약 17.8% 증가하였다. 학생 수에 비해 교수의 증가폭이 큰 것은 산업공학 관련 전공 교육 및 연구 분야가 다양해졌음을 알 수 있다(최진영, 2014).

이와 같은 국내 산업공학 학문의 발전에 건인차 역할을 한 것은 1974년 창립한 대한산업공학회를 비롯하여 한국SCM학회·한국경영공학회·한국정보지능시스템학회 등 산업공학 분야 주요 학회들의 활발한 연구 활동이라고 할 수 있다. 이들 학회는 학술지 발간 및 연구자들이 연구한 결과를 공유할 수 있는 자리를 마련할 뿐 아니라 다양한 연구 과제를 수행함으로써 국내 산업공학 학문 발전에 기여를 하고 있다.

산업공학의 역할은 시대의 발전에 따라 변화해 왔다. 1960 ~ 70년대는 제조업분야의 생산 혁신을

리드하였고 1980년대는 프로세스 혁신과 품질 경영을 리드하였다. 1990년대 들어서는 IT와 결합하여 새로운 제조 방법 및 서비스 산업을 구축하기 시작했으며 기존에 없었던 새로운 비즈니스 세계를 만들어내는데 기여하였다. 시대의 발전으로 산업공학의 적용 분야는 전통적인 제조 산업에서 벗어나 IT, 통신, 금융, 유통, 물류, 의료 등 다양한 서비스 산업 분야로 확대되었다(김형태, 2012).

오늘날 우리 사회의 산업은 제4의 물결을 맞이하고 있다. 인공지능과 머신러닝 같은 신기술이 발전함에 따라 이전에 없었던 새로운 산업 및 시대가 도래하고 있다. 2016년 1월에 세계경제포럼은 ‘제4차 산업혁명의 이해’라는 주제로 포럼을 개최하였으며 화두는 인공지능, 로봇공학, 사물인터넷(IoT) 등의 신기술이 우리사회의 산업에 어떠한 변화를 미칠지에 대한 것이었다. 제4의 물결은 과거의 산업혁명보다 더 큰 파급력으로 산업의 규모와 범위, 복잡성을 증가시킬 것으로 보고 있으며 우리 정부는 이 변화에 대해 ‘제조업혁신 3.0 전략’을 추진하고 있다. 이와 같은 제 4의 물결이 국내의 모든 산업에 영향을 미치게 됨에 따라 산업공학의 역할이 더욱더 확대가 될 것으로 보인다.

이러한 시점에서 산업공학 학계에서 연구되고 있는 동향을 분석함으로써 새로운 변화에 대해 대응할 필요가 있다. 그동안 국내 산업공학의 연구는 특정 분야의 개선 및 요인 분석, 예측 모델 개발 관련 연구가 많이 수행된 반면에 산업공학의 전체적인 흐름 및 변화를 살펴보는 연구는 미비한 실정이다. 대학 및 기관은 산학협력으로 다양한 분야의 연구를 수행하고 있고 그 결과를 산업경영공학 학계에 게재하기 때문에 산업경영공학 학계에서 연구되는 동향을 분석하면 새로운 변화에 대한 전략을 수립할 수 있을 것으로 기대된다.

본 연구는 국내의 주요 산업공학 학술지의 영문

초록 데이터를 이용하여 국내 산업경영공학의 동향 분석을 수행한다. 총 8개 학술지에서 2004 ~ 2015년의 영문 초록 데이터를 사용하여 LDA (Latent Dirichlet Allocation) 기반의 토픽모델링을 수행함으로써 국내 산업경영공학 분야에서 연구되고 있는 주요 주제 및 과거에 비해 새롭게 나타나고 있는 연구 주제를 도출한다. 또한 연구가 활발히 수행되고 있는 상향 주제(Hot Topic)와 점점 하락하고 있는 하향 주제(Cold Topic)를 분석하고 주제 간 네트워크 구조 분석을 통해 핵심 주제를 도출한다. 이를 통해 국내 산업경영공학 분야에서 연구자들이 관심을 갖는 주요 주제가 무엇인지 규명하고 산업경영공학의 핵심 주제가 무엇인지 제시하여 산업경영공학의 전반적인 연구 동향을 분석하고자 한다.

논문의 구성은 다음과 같다. 2장에서는 산업공학 관련 연구 대한 문헌 연구를 통해 기존 연구들을 살펴보고, 3장에서는 연구 방법을 설명한다. 4장에서는 국내 산업경영공학의 동향 분석 결과를 제시하고, 5장에서는 연구의 의의 및 한계점을 포함하여 결론을 맺는다.

2. 선행 연구

2.1 국내 산업경영공학의 동향 연구

김형태(2012)의 연구에 따르면 특정 분야를 전문적으로 연구하는 기계공학, 전기공학, 화학공학 등 타 공학 분야와는 다르게 산업공학은 여러 분야가 융합되어 있는 공학이라고 정의하였다. 타 공학에서는 각 분야의 전문적인 영역을 다루는 반면 산업경영공학은 기업의 보다 나은 경영 환경 개선을 위해 사회과학 분야의 지식과 함께 수학·물리학 및 공학적인 설계, 해석원리 등 자연과

학적 지식과 원리를 응용한다. 이를 통해 기업의 경영 환경에 대해 여러 요소를 과학적으로 분석하고 최적의 경영 시스템 형성에 기여하는 공학인 것이다.

김형태(2012)의 연구에서 산업공학이 다루는 분야를 크게 3가지의 범위로 구분하였다. 첫째, 기업의 장기경영예측 수립을 위한 각종 경영예측, MAPI 방식을 이용한 설비계획, 신제품의 연구·계획·평가, 정보시스템의 설계·관리, 사무의 자동화, 상품계획 분야이다. 둘째, 작업측정과 작업개선에 관한 것으로 현장작업 뿐만이 아니라 사무작업에도 적용되고 분야로 공정관리·작업관리·임금관리·재고관리·운반계획·품질관리 분야이다. 마지막으로 인공지능, 감성공학, 데이터마이닝, 생물정보학과 같은 분야이다. 이와 같이 산업경영공학은 여러 분야를 다루기 때문에 각 분야에서 필요로 하는 문제점에 대해 개선 및 요인 분석, 예측모델 수립을 위한 연구는 많았으나 산업공학의 전반적인 연구 동향을 분석한 연구는 소수이다.

여러 분야가 포함되어 있는 학문이기 때문에 체계를 구성할 필요가 있었고 이태억(2003)은 대학의 교과과정이 산업계의 요구에 충분히 대응하지 못하고 있다고 판단하여 정성적인 분석을 수행하였다. 대한산업공학회 표준교과과정위원회의 보유 자료, 국내외 우수 대학의 산업공학의 커리큘럼, 산업공학의 변화를 주도하고 있는 학계, 산업계 전문가의 의견을 종합하여 제조시스템, Business Process Modeling & Workflow System, Product/Engineering Information Management, ERP 및 기업 정보시스템, e-Business 및 전자상거래, 공급체인관리, 데이터베이스 및 모델링, 인터넷 응용 시스템, 운송 및 물류 시스템, 금융 및 공공 시스템 등 산업공학의 주요 교육방향을 제시하였다.

국내 산업공학은 지난 50여년 많은 발전을 이루어 다양한 산업계로부터 그 수요가 증가하였으

며 업무 영역 또한 넓어지게 되었다. 기업에서의 산업공학 전공자의 업무 영역이 제조업 중심의 생산, 품질의 영역 외에 재무, 마케팅, 기획 등으로 확장되면서 중요한 변화를 맞이하였다. 박양병(2006)의 연구에서는 양적 팽창과 급변하는 기업 환경의 변화에 대응하기 위해 창의적이고 미래지향적인 산업공학 교육체제와 교과내용을 개발하는 연구를 수행하였다. 국내외 대학 산업공학 교과과정 사례연구, 산업공학 전공 졸업생 대상 설문조사, 산학연 전문가의 자문을 통해 기초과학(수학, 과학, 전산), 시스템 분석, 생산/물류, 품질, 인간공학, 정보시스템, 경영공학의 산업공학의 7개 교육 분야를 구축하였다.

산업공학 동향 연구는 기술 발전에 따라 변하는 기업 환경에 대응하기 위해 교육 분야를 몇 개의 영역으로 정의하는 연구가 정성적으로 수행되었다. 교육기관에서는 이를 토대로 하여 각 분야에 맞는 지식을 제공함으로써 기업에서 필요로 하는 전문적인 인력을 양성해오고 있다. 이러한 교육 분야는 과거에 연구가 수행된 결과물이며 최근 급변하는 기업 환경에 맞추어 현재 시점에서 새롭게 나타나고 있는 교육 주제 및 산업공학에서의 핵심 주제를 연구할 필요가 있다. 본 연구는 산업공학 학계의 논문 텍스트 데이터를 이용하여 정량적으로 연구동향을 분석한다.

2.2. 텍스트 데이터를 이용한 산업경영공학 관련 연구

연구동향을 정량적으로 분석하여 추세를 파악하는 방법으로 문서에 나타난 텍스트 데이터를 이용하는 방법이 있다. 과거에는 매트릭스 구조에 숫자형태로 기록된 정형화된 데이터 구조를 이용한 데이터 분석이 주로 연구되었으나 최근에는 문서의 텍스트, 이미지, 동영상, 음성 등 비정형화

된 데이터를 이용하여 정보를 분석하는 방법에 관심이 증가하고 있다(Chakrabartij, 2002).

텍스트 데이터를 이용하는 분석은 문서의 분류 및 군집, 문서 필터 및 추천에 사용될 수 있으며 키워드 도출 및 추세를 예측하는 방법에도 사용될 수 있다(정근하, 2011). 또한 특허의 텍스트 데이터를 이용하여 기술동향을 분석(Yoon et al., 2012)하는 연구 및 학계의 문헌 텍스트 데이터를 이용하여 연구동향을 분석(Park et al., 2013)하는 연구에도 사용된다.

산업공학 분야에서도 텍스트 데이터를 활용한 연구가 수행되어 왔다. Cho et al.(2012)는 산업공학의 분야의 대표적인 국제 학술지인 IIE Transactions에 출판된 논문의 텍스트 데이터를 이용하여 논문 주제어간 연관관계를 분석하는 연구를 수행하였다. 1969년부터 2011년도의 논문에서 제목 및 초록데이터가 존재하는 2,527개의 논문을 분석 대상으로 하여 48개의 주제어를 선정하였다. 주제 간의 고차원의 행렬 데이터를 구성하여 상관관계 분석을 위해 시각화, 군집분석, 연관성 분석을 실시하여 과거에서 현재에 이르는 산업공학 내 연구의 현황과 추이를 관찰하였다.

Cho et al.(2012)가 해외 저널의 텍스트 데이터를 이용하여 산업공학의 연구 동향을 살펴보았다면 Cho et al.(2014)는 국내 대학의 산업공학 소속 전임교수들이 가장 많이 논문을 게재한 국내 학술지 상위 5개를 선택하여 이들 논문의 텍스트를 데이터를 이용하여 산업공학에서 주로 사용되는 연구기법을 분석하였다. 상위 5개 학술지는 대한산업공학회지, 산업공학(IE Interface), 한국산업경영시스템학회지, 한국경영과학회지, 한국인간공학회지로 1975년부터 2012년까지 게재된 3,857개의 논문에서 제목, 초록, 저자가 작성한 주제어를 수집하였고 단순빈도분석, K-평균군집화 기법, 연관성분석을 이용하여 산업공학 연구기법의 추세를

분석하였다. 또한 Jeong et al.(2013)은 대한산업공학회지와 IE Interfaces 논문의 공저(co-authorship) 정보를 활용하여 국내 산업공학 분야의 산학융합연구 네트워크 분석을 수행하였다.

국외에서도 산업공학의 연구동향을 분석하는 연구가 이루어졌는데 Dastkhan와 Owlia(2009)는 산업공학 분야와 관련된 25개의 해외저널을 선정하여 1980년부터 2007년까지 총 7114개 논문에서 제목, 저자, 년도, 저널명, 기관명, 나라, 키워드 등 데이터를 이용하여 분석하였다. 논문의 키워드, 저널의 영역, 전문가의 지식을 이용하여 상위 주제와 하위주제를 정하였고 이를 활용하여 연구동향을 분석하였다.

Uys et al.(2011)는 해외 저널 중 산업공학과 관련된 주요 저널인 Computers & Industrial Engineering(CaIE)에서 1977년도부터 2011년까지 4497개 논문의 텍스트 데이터를 이용하여 토픽모델링 방법을 통해 연구동향을 분석하였고 Dastkhan와 Owlia(2009) 연구 결과와 비교분석 하였다. Dastkhan와 Owlia(2009)가 전문가의 지식을 이용하여 도출한 주제에 대해 빈도분석을 사용하여 연구동향을 분석한 반면에 Uys et al.(2011)는 논문의 텍스트데이터를 이용하여 문서에 잠재된 주제를 도출하는 토픽모델링 방법을 활용하여 연구동향을 분석하였다.

텍스트 데이터를 활용한 국내 산업공학 동향 분석은 논문의 제목 및 초록, 키워드에서 특정 단어들을 기반으로 한 빈도분석 및 연관성분석, 군집분석을 사용하여 주제어에 대한 연관관계, 연구기법을 분석하였으나 산업공학 분야의 전체적인 주제 흐름에 관한 연구동향을 분석한 연구는 이루어지지 않았다. 반면에 국외에서는 논문의 텍스트 데이터를 이용하여 토픽모델링 기법을 활용하여 산업공학 분야에서 연구되고 있는 전체적인 주제 변화를 분석하였으며 이를 기존 선행연구와 비교

분석하였다.

따라서 본 연구는 산업공학의 연구 동향을 분석하기 위해 국내에서 연구가 이루어지지 않았던 산업공학 분야의 논문 텍스트 데이터를 활용하여 문서에 잠재된 주제를 도출할 수 있는 토픽모델링 기법을 통해 산업공학에서 연구되는 주요 주제를 도출하고 변화를 분석한다.

3. 연구방법

3.1 연구 프레임워크

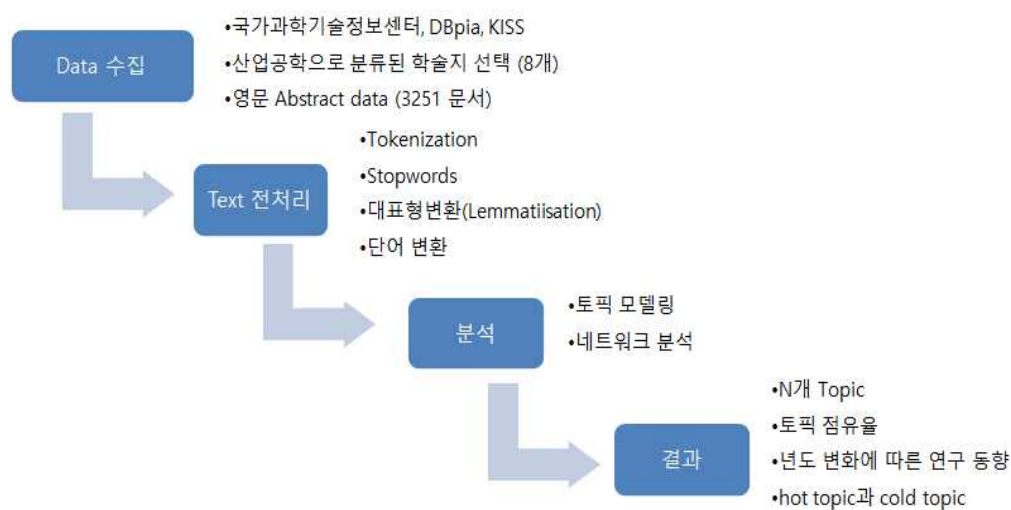
본 연구의 프레임워크는 <그림 3.1>과 같다. 한국학술지인용색인의 주제 분류 기준에 따라 산업공학으로 분류된 학술지 8개를 선택하였다. 그리고 국내학술 DB인 국가과학기술정보센터, DBpia, KISS에서 영문 초록이 게재되어 있는 3251개의 초록을 수집하였다.

수집된 초록데이터를 분석에 알맞은 형태로 변환하기 위해 R program을 이용하여 Tokenization, Stopwords, Lemmatization 등 데이터 전처리 작업을 진행하였다. 그리고 연구되고 있는 주제의 현황 분석을 위해 토픽모델링 기법을 사용하였고 주제간 연계 구조를 분석하기 위해 네트워크 분석을 진행하였다.

분석 결과로 산업공학 분야에서 연구되고 있는 20개의 주요 주제를 추출 및 정의를 하고 이 주제들에 대해 년도 변화에 따른 연구 동향을 살펴 보았다. 또한 연구가 점점 증가하고 있는 상향 주제(Hot Topic)와 하락하고 있는 하향 주제(Cold Topic)들을 살펴봄으로서 연구 트렌드를 분석하였다. 마지막으로 주제어간 코사인 유사도를 도출하여 주제 간의 관계 네트워크를 구축하여 산업공학의 핵심 주제를 도출하였다.

3.2 분석 데이터

논문의 초록 데이터를 이용하여 산업공학의 연



<그림 3.1> 연구 프레임워크

구동향을 분석하기 위해서는 산업공학과 관련된 적절한 학술지를 선택하는 것이 중요하다. 산업공학은 모든 산업에 적용될 수 있고 연구하는 분야가 다양하기 때문에 연구 논문이 여러 학술지에 게재가 된다. 국내의 많은 학술지에서 산업공학과 관련된 연구 논문이 게재된 학술지를 명확하게 선정하는 것은 어려움이 있다. 연구자의 주관으로 학술지를 선정하게 되면 데이터의 편차가 발생할 수 있기 때문에 학술지 선정에 있어서 객관적인 기준이 필요하다.

Cho et al.(2014)의 연구에서는 데이터의 객관성을 유지하기 위해 2011년 ~ 2013년 중앙일보의 학과평가에서 나타난 상위 25개 대학의 산업공학 관련학과를 선정하고 학과의 소속 전임교수들이 논문을 가장 많이 게재한 상위 5개의 국내저널인 대한산업공학회지, 산업공학(IE Interface), 한국시스템경영시스템학회지, 한국경영과학회지, 한국인간공학회지 등을 선정하였으나 한국인간공학회지는 분석 목적에 있어서 다른4개 학회지와 매우 상이하다고 판단하여 제외하였다.

이와 같이 적절하게 학회지를 선정하였다고 하더라도 분석의 목적에 따라 학회지 선정이 제외될 수 있다.

본 연구에서는 주관적인 기준을 배제하고 데이터의 편차를 줄이기 위해 한국학술지인용색인(Korea Citation Index)의 주제 분류 기준에 따라 산업공학과 관련된 학술지를 선정하였다. 한국학술지인용색인은 국내의 학술지 정보, 논문 정보 및 참고문헌을 DB화하여 논문 간 인용관계를 분석하는 시스템이며 한국연구재단에서 운영하기 때문에 공신력이 있다고 할 수 있다.

한국학술지인용색의 주제 분류 기준에 따라 선정된 학술지는 <표 3.1>과 같다.

6개의 발행기관에서 출간하고 있는 8개 학술지를 선정하여 <표 3.2>와 같이 학술 DB인 국가과학기술정보센터(NDL), DBpia, KISS에서 2004년부터 2015년까지 영문 초록이 게재된 논문 총 3251개를 수집하였다. 산업공학회지는 2013년2월부터 대한산업공학회지 통합되어 발간되기 때문

<표 3.1> 학회지 선정 대상

학술지	발행기관	분야
Industrial Engineering & Management Systems 대한산업공학회지 산업공학회지	대한산업공학회	최적화, 생산/물류관리, 확률/통계, 인간/인지공학, 정보시스템, 경영공학
한국SCM학회지	한국SCM학회	SCM, 물류/유통관리, 수송관리, e-business
한국경영공학회지	한국경영공학회	품질경영, 경영공학, 기술경영
지능정보연구	한국정보지능시스템학회	인공지능/데이터마ining, 지능형 인터넷, 지식경영
산업경영시스템학회지	한국산업경영시스템학회	생산/제조, 통계/품질/6시그마, 경제성/경영/서비스공학, 물류/SCM, 정보/IT, 최적화/알고리즘, 인간/감성/안전
품질경영학회지	한국품질경영학회	R&D 품질, 품질 경영, 서비스 품질, 경영 품질, 통계적 품질 경영, 품질 선진 기법, 국방 품질, 사이버 비즈니스

에 2004 ~2012년까지 수집되었고 한국경영공학회는 총 3개의 학술지를 포함하고
지는 2011년부터 2015년까지의 데이터가 수집되었다. 있기 때문에 데이터 개수가 가장 많았으며 2011

<표 3.2> 데이터 출처

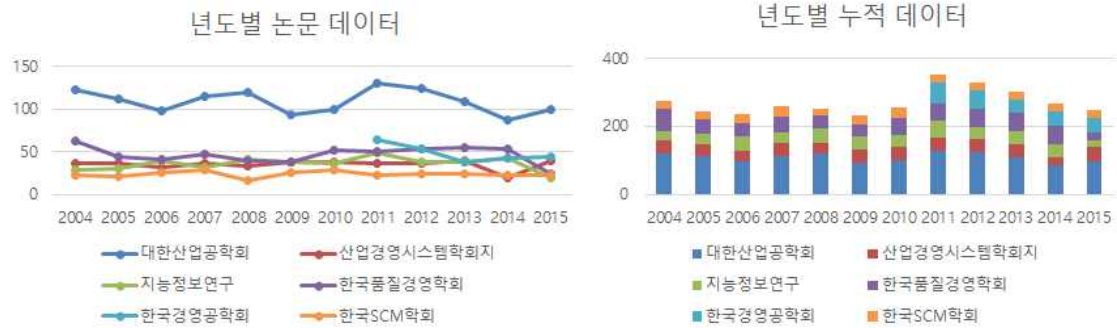
학술지	년도	문서개수	출처
Industrial Engineering & Management Systems	2004 ~ 2015	354	DBpia
대한산업공학회지	2004 ~ 2015	521	DBpia
산업공학회지	2004 ~ 2012	431	DBpia
한국SCM학회지	2004 ~ 2015	287	DBpia
한국경영공학학회지	2011 ~ 2015	242	KISS
지능정보연구	2004 ~ 2015	433	NDSL
산업경영시스템학회지	2004 ~ 2015	424	NDSL
품질경영학회지	2004 ~ 2015	559	NDSL
총		3251	

* 산업공학회지는 2012년 이후 대한산업공학회지와 통합됨

데이터를 수집하여 <표 3.3>과 같이 제목, 년도, 년에는 논문의 출간이 가장 많았다. 수집한 데이
초록의 데이터베이스를 구축하였다. 발행기관이 터의 현황은 <그림 3.2>과 같다.

<표 3.3> 산업공학 학술지 영문초록 데이터베이스

No	Title	Year	List
1	A Branch-and-price Approach to the ATM Switching Node Location Problem	2004	This paper presents a process plan selection model with multiple objectives. The process ...
2	Branch and Bound Approach for Single-Machine Sequencing with Early/Tardy Penalties and Sequence-Dependent Setup Cost	2004	A tree structure model has been proposed for representing the unequal-area facility layout. Each facility ...
3	Evolutionary Algorithm for Process Plan Selection with Multiple Objectives	2004	Studies on the optimal location of retail store have been made in case of no obstacle ...



<그림 3.2> 학회지별 분석 데이터 현황

3.3 분석 데이터 전처리

텍스트 데이터를 분석에 활용하기 위해 전처리 작업을 진행하였다. 데이터 전처리를 통해 정형화되어 있지 않은 텍스트 데이터를 분석이 가능하도록 하여 분석에 소요되는 시간을 줄이고 분석의 정확성을 높일 수 있다.

데이터 전처리에 대한 설명은 <표 3.4>와 같고 전처리 과정은 <그림 3.3>처럼 진행하였다.

3.4 분석 방법

3.4.1 토픽모델링(Topic Modeling)

토픽모델링은 수집된 문서에 대하여 각 문서에 어떤 주제들이 분포되어있는지를 찾아내는 확률 모형이다(Blei, 2012). 각 문서에는 토픽들의 집합이 확률적으로 존재하고 각 토픽을 단어의 분포로 나타냄으로서 문서의 구조를 분석하는 모델이

전처리 단계	내용
Tokenization	여러 문장에 대해 공백을 기준으로 단어를 분리 A/ tree /structure/ model/ has/ been proposed/ for
Stopwords	분석에 불필요한 단어 제거 tree/ structure/ model/ proposed/
Lemmatization	키워드 품사가 고려되어 의미적 정보를 가진 단어의 표제어 도출 tree/ structure/ model/ propose/
공통단어로 변환	같은 의미이지만 다른 형태로 표현된 단어들을 공통 단어로 변환 “supply chain , supply chain managemt” >> “supplychain”

<표 3.4> 텍스트 데이터 전처리 내용

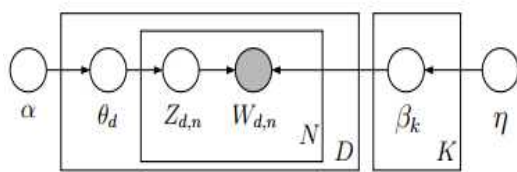


<그림 3.3> 텍스트 데이터 전처리 과정

다. 즉 문서를 이루고 있는 단어들을 바탕으로 문서에 숨겨져 있는 토픽들을 추정함으로써 해당 문서가 어떤 토픽들을 같이 다루고 있는지 찾아낼 수 있다.

문서의 토픽 분석을 위한 방법에는 Deerwester et al.(1990)의 LSI(Latent Semantic Indexing), Hofmann(1999)의 pLSA(Probabilistic Latent Semantic Analysis), Blei et al.(2003)의 LDA(Latent Dirichlet Allocation) 방법이 있다. 최근에는 LSI와 pLSA의 단점을 보완한 LDA 기법이 문서의 주제를 찾는 연구에서 많이 사용되고 있으며 본 연구에서는 토픽모델링의 LDA 기법을 이용한다(이원상 외, 2015).

LDA 알고리즘은 확률적 토픽 모델로서 문서 내의 숨겨져 있는 주제들을 찾아내는 알고리즘이다. 확률적 토픽 모델에서 각각의 문서는 다수의 주제가 존재하고 있는 것으로 간주한다. 그리고 각 주제별로 단어가 생성되는데, 문서 내의 주제들의 분포와 각 주제별로 단어가 생성될 확률을 모델링 하여 문서의 생성 확률을 도출한다(Blei et al., 2003). <그림 3.4>는 이러한 LDA의 알고리즘을 그래프로 표현한 것이다.



<그림 3.4> LDA의 그래프 모델

K : 토픽의 개수

α : 문서별 토픽 k 의 Dirichlet prior weight, θ 값을 결정하는 파라미터

η : 토픽별 단어 w 의 Dirichlet prior weight, β 값을 결정하는 파라미터

θ_d : 문서별 토픽의 비율

β_k : 토픽별 단어 w 의 생성확률

$Z_{d,n}$: 문서 d 의 n 번째 단어의 토픽(index)

$W_{d,n}$: 문서 d 의 n 번째 단어(문서에 관측되는 변수, index)

LDA 알고리즘은 문서 내에서 관찰되는 변수를 통해 보이지 않는 변수(hidden variable)를 추론하여 문서에 숨겨져 있는 주제를 발견하는 것으로 <그림 3.4>에서 관찰되어지는 변수는 단어($W_{d,n}$)이다. 그리고 hyperparameter인 α, η 와 단어를 추출하기 사용되는 hidden parameter인 β 가 있다. 마지막으로 문서 내에서 직접적으로 관찰할 없는 θ, z 의 hidden variable이 있다.

LDA 모델의 Z 는 문서별 토픽 비율인 θ 로부터 생성되고 θ 는 Dirichlet 분포를 따르는 값으로 α 값에 의해 형태가 정해진다. 마찬가지로 토픽별 단어 생성확률인 β 는 η 값에 의해 결정되어지는 값으로 η 에 따라 β 의 Dirichlet 분포가 결정된다. 각 단어의 토픽을 나타내는 값인 Z 와 토픽별 단어 비율인 β 값에 따라 단어 W 가 결정된다.

3.4.2 네트워크 분석

네트워크 분석은 개체(Node)와 관계(Relationship)로 구성된 네트워크 구조에서 노드간의 연관관계 따른 각 노드가 네트워크 구조에서 어떤 역할을 담당하거나 어떤 영향을 미치는지 분석하는 방법이다. 노드는 네트워크 구조를 구성하는 요소이며 Relationship은 노드 간의 링크로서 이를 활용하여 노드의 개별적 속성, 노드를 이어주는 연계(관계적 속성), 네트워크 구조의 전체적인 특징을 분석할 수 있다. 즉, 노드 간의 상호연관 관계를 통해 어떠한 요소들이 중요한지 파악할 수 있다(정근하, 2011).

네트워크의 구조에서 핵심 노드를 파악하기 위해 중심성 분석(Centrality Analysis)이 활용된다(Newman, 2010). 네트워크 중심성 분석은 각 요소를 표현하는 노드의 상호관계를 규명하는 것으로 네트워크에서 허브역할을 담당하는 노드와 중계의 역할을 하는 노드를 파악할 수 있다. 이러한 분석의 목적으로 노드들에 대한 역할, 위치, 특성, 영향력을 파악할 수 있는 중심성 지수가 많은 연구에서 활용되고 있다.

중심성(Centrality) 지수로는 특정 노드에 연결된 다른 노드의 수를 측정하는 연결중심성(Degree Centrality), 하나의 노드가 다른 노드들과의 네트워크를 구성하는데 있어서 중계 역할 및 다리 역할을 측정하는 매개 중심성(Betweenness Centrality), 하나의 노드가 다른 노드에 얼마만큼 가까이 있는지를 측정하는 근접 중심성(Closeness Centrality), 네트워크 구조에서 하나의 노드에 연결된 상대 노드들의 영향력 및 중요성에 가중치를 부여하여 특정 노드의 중요성을 측정하는 위세 중심성(Eigenvector Centrality)이 있다(Freeman, 1979).

본 연구에서는 토픽모델링을 통해 생성된 주제에 대해 네트워크 중심성 분석 및 주제 간의 관계를 시각적으로 표현하여 주제 간의 관계를 해석한다.

4. 산업경영공학 분야 연구 동향 분석

국내 산업경영공학 분야 8개 학술지의 영문 초록을 이용하여 키워드 분석과 토픽모델링을 수행하였다. 토픽모델링 기법을 통해 각 주제별 단어의 상위 출현 확률 및 상위 논문을 활용하여 주제를 정의하였고 각 주제들에 대해 연도별 분석과 상향 주제 및 하향 주제를 통해 추세를 분석하였다. 토픽모델링의 결과에 대한 분석 이전에 주요 키워드의 빈도 분석을 수행하였다.

또한, 토픽모델링 결과로 정의한 국내 산업경영공학의 연구 주제에 대하여 주제간 네트워크 분석을 수행하였다. 주제간 네트워크 분석을 통해 앞에서 정의한 20개 주제의 특성을 살펴보고자 한다.

4.1 키워드 빈도 분석

<표 4.1>은 8개 학술지의 초록에서 빈도수가 1000번 이상 나타난 키워드로서 단순키워드분석을 수행하였다. 주요키워드는 system, process, service, problem, performance, time, quality, product 등이다. 이는 국내 산업경영공학에서의 연구가 시스템, 프로세스, 서비스의 개선을 통한 성능 향상

<표 4.1> 주요 키워드 빈도표

단어	빈도	단어	빈도
system	3114	analysis	1167
process	2129	cost	1126
service	1950	company	1117
problem	1703	design	1095
performance	1691	customer	1093
information	1621	management	1070
time	1416	user	1041
quality	1403	order	1026
product	1366	factor	1003

및 시간 관리, 품질 문제를 다루고 있고 cost, customer, user, order 등의 단어가 등장한 것으로 보아 비용에 관한 경제성공학 연구, 고객관리에 관한 연구에 초점을 두고 있는 것으로 보여진다.

4.2 산업경영공학 분야 주제 정의

국내 산업경영공학 분야 8개 학술지의 총 3251개 초록을 이용하여 토픽모델링을 수행한 결과 분석이다. R program을 이용하여 전처리 작업을 수행한 초록 데이터에서 주제를 찾기 위해 Hornik와 Grün(2011)이 R Program의 Package로 배포한 “topicmodels”을 활용하였다.

토픽모델링을 활용하여 주제를 찾기 전에 몇 개의 주제를 찾을지를 먼저 결정하여야 한다. 토픽모델링에서 최적의 주제 수를 결정하는 값으로 language model 성능을 평가하는 척도인 perplexity value를 사용한다. 20개, 30개, 50개, 100개 등 주제의 수를 늘려가며 각 주제의 수에 대한 perplexity vlaue를 구하여 가장 낮은 값을 가지는 주제의 수를 선택한다(Blei et al., 2003).

그러나 perplexity value는 주제의 수가 증가할수록 감소하는 경향이 나타나기 때문에 효과적으로 해석할 수 있는 수준에서 주제의 수를 선정하였고 LDA 모델의 파라미터 추정을 위해 Gibbs Sampling 방식을 사용하였다.

토픽모델링의 실험으로 각 주제에 속하는 단어의 집합 및 상위 논문의 집합을 알 수 있다. 주제를 정의하기 위해서 <표 4.2>와 같이 각 주제별에 대해 단어 출현 확률이 상위로 나타난 단어와 <표 4.3>과 같이 각 주제에 대한 상위 논문들의 제목과 내용을 이용하였다. 이렇게 정의된 주제에 대해 국외 연구에서 정의한 주제와 비교를 하여 살펴보았다. Dastkhan와 Owlia(2009)가 정의한 국외의 산업공학 분야에서 연구되고 있는 주제는 <표 4.4>와 같다.

<표 4.5>는 20개의 주제와 각 주제에 대해 상위로 나타난 단어들로서 국내 산업경영공학에서 연구되고 있는 주제이다. 국내 산업경영공학에서 연구되는 20개로 정의된 주제는 공급망 관리, 헬스케어, 제조시스템 설계, 금융공학, 차세대 웹 기술, 최적화, 기업성과, 공정관리, 컨테이너 터미널, 간공학, 사용자 추천 시스템, 텍스트마이닝, 시스

<표 4.2> 각 주제별에 대한 상위 단어

Topic1	Topic2	Topic3	...	Topic20
cost	service	system	company	network
order	mobile	reliability	management	social
inventory	user	risk	business	research

<표 4.3> 각 주제별에 대한 상위 논문

Topic	Posterior	Title	Abstract	Year
Topic 1	0.977907	Inventory Model ~	This paper presents ~	2013
Topic 1	0.972464	Genetic Algorithm ~	Distribution centers ~	2009
Topic20	0.93875	Efficient Polling ~	A modified PCF ~	2012

토픽모델링을 이용한 국내 산업경영공학 연구동향 분석

HW/SW 신뢰성, DEA 효율성, 모바일 서비스, 인
템 모델링, 서비스 품질, 의사결정시스템, 품질경
영, 데이터 마이닝이다.

Dastkhan와 Owlia(2009)의 주요 주제와 비교를 통

해 국내와 국외의 연구되고 있는 주요 주제들이
명칭만 다를 뿐 전반적으로 비슷한 주제를 다루
고 있는 것을 확인할 수 있었고 헬스케어 및 텍
스트마이닝에 관한 연구가 새로이 나타나고 있는
것을 확인하였다.

<표 4.4> Dastkhan와 Owlia(2009)의 국외 산업공학 주요 연구 주제

Main Topic	Subjects
Production Management	Lean Production, Agile Production, Maintenance, Reliability, Total Productive Maintenance, Kanban System, Production Planning
Information Systems and Technology	Information Technology, Knowledge Management, Information Systems, MIS
Operations Research	Mathematical Programming, AHP, MCDM, Simulation, Genetic Algorithm, Tabu Search, Simulated Annealing, Heuristic Algorithms, Decision Support Systems, Chaos Theory, Constraints Theory
Project Management	Project Management, Project Control, Value Engineering and Management
Supply Chain Management	SCM, ERP, MRP, EOQ, MRPII
Total Quality management	TQM, 6 Sigma & Lean 6 Sigma, Quality Assurance, Quality Control, Quality Awards, BSC, DEA, Taguchi method, DOE, QFD, FMEA, CRM, Benchmarking, Kaizen
Advance Production Systems and Technology	Cellular manufacturing, FMS, CIM, GT, Reverse Engineering
Intelligent Systems and Methods	Neural Network, Artificial Intelligence, 퍼지 Logic, Data Mining- Expert Systems
Methods Engineering	Ergonomy and Human factors, Work Study, Time Study, Productivity, Facility Layout, BPR, Line Balancing
Other subjects	TRIZ, Innovation and Creativity, Problem Solving, Technology Management

<표 4.5> 국내 산업경영공학 분야의 주제 정의

[Topic 1]	[Topic 2]	[Topic 3]	[Topic 4]	[Topic 5]	[Topic 6]	[Topic 7]	[Topic 8]	[Topic 9]	[Topic 10]
공급망 관리	헬스케어/직무교육	제조시스템 설계	금융공학	차세대 웹 기술	최적화	기업성과	공정관리	컨테이너 터미널 시스템	HW/SW 신뢰성
cost	student	process	price	knowledge	problem	performance	process	container	system
supplychain	job	production	market	information	solution	factor	design	vehicle	reliability
product	university	manufacture	investment	web	time	effect	sample	system	time
inventory	medical	machine	financial	search	heuristic	relationship	chart	storage	cost
demand	education	line	stock	system	solve	influence	control	transportation	failure
order	hospital	assembly	strategy	agent	algorithm	positive	variable	location	distribution
plan	patient	time	economic	ontology	optimal	supplychain	controlchart	simulation	estimate
logistic	healthcare	system	cost	semantic	schedul	affect	performance	logistic	maintenance
optimal	train	part	option	user	genetic	research	quality	ship	parameter
distribution	worker	simulation	trade	framework	minimize	analysis	statistical	traffic	test
[Topic 11]	[Topic 12]	[Topic 13]	[Topic 14]	[Topic 15]	[Topic 16]	[Topic 17]	[Topic 18]	[Topic 19]	[Topic 20]
DEA 효율성 분석	모바일 기술	인간공학	사용자 추천 시스템	텍스트마이닝/소셜분석	시스템 모델링	서비스품질	의사결정시스템	품질경영	데이터마이닝
efficiency	user	image	customer	network	system	service	evaluation	company	prediction
dea	information	design	product	social	process	quality	risk	business	classification
analysis	service	measure	system	analysis	product	customer	project	industry	set
input	mobile	hand	recommendation	user	design	factor	decision	technology	technique
defense	technology	drive	market	information	management	satisfaction	criteria	management	performance
military	compute	performance	purchase	research	development	attribute	selection	research	mine
score	content	task	item	keyword	information	measure	process	strategy	cluster
measure	environment	time	user	news	business	product	factor	sigma	feature
output	rfid	experiment	preference	topic	standard	analysis	select	activity	network
productivity	security	human	shop	article	company	loyalty	analysis	enterprise	problem

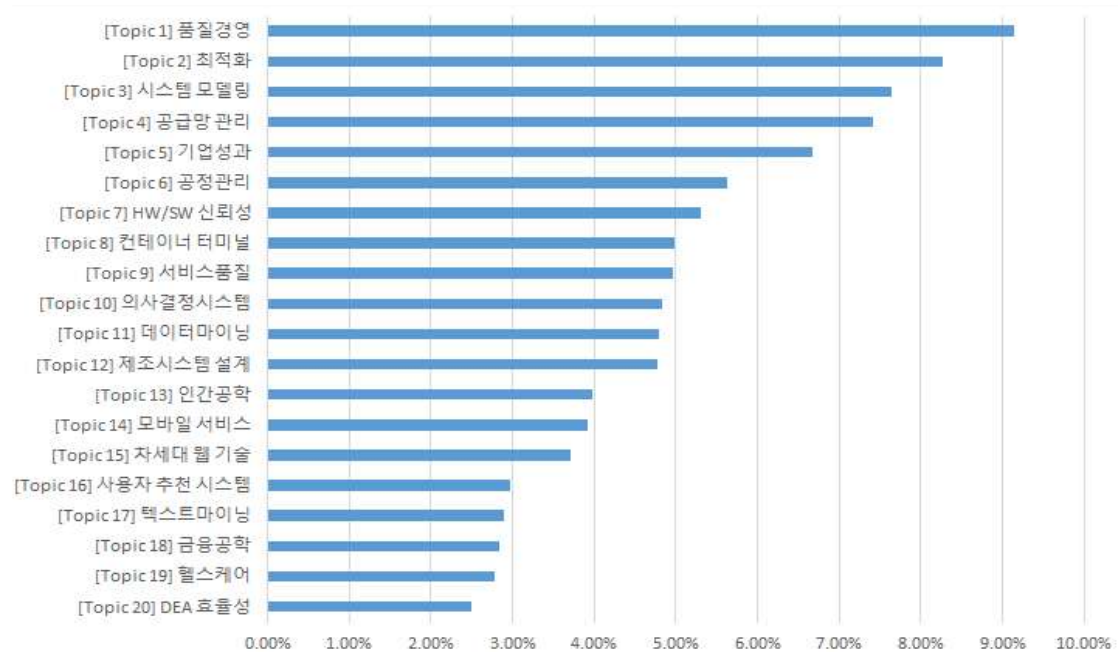
4.3 연도별 주제 추이 분석

토픽모델링의 실험 결과로 나타난 국내 산업경영공학 분야에서 연구되고 있는 주제에 대해 연도별 추이를 3가지 단계로 분석하였다. 첫째, 데이터를 수집한 전체 기간(2004 ~ 2015년)에 대해 어떤 주제가 많이 연구되었는지 분석하였다. 둘째, 전체 기간을 제 1구간(2004 ~ 2007년), 제 2구간(2008 ~ 2011년), 제 3구간(2012 ~ 2015년)으로 나누어 기간의 구간별 분석을 하였다. 셋째, 각 주제에 대해 선형회귀 분석을 통해 상향 주제(Hot Topic)와 하향 주제(Cold Topic)를 분석하였다.

분야에서 연구되고 있는 주제의 점유율은 <그림 4.1>과 같다. 품질경영, 최적화, 시스템 모델링, 공급망 관리, 기업성과, 공정관리에 대한 연구가 많이 수행된 반면 사용자 추천 시스템, 텍스트마이닝, 금융공학, 헬스케어, DEA 효율성에 관한 연구는 상대적으로 적게 수행된 것으로 나타났다.

4.3.2 기간별 주제 분석

전체 기간(2004 ~ 2015년)을 3개의 구간으로 나누어 주제 변화에 대한 분석을 수행한 결과는 <표 4.6>과 같고 이를 활용하여 기간별 변화를 살펴보았다.



<그림 4.1> 2004~ 2015년 주제별 점유율

4.3.1 전체 기간의 주제 분석

2004 ~ 2015년 기간 동안 국내 산업경영공학

<그림 4.3>과 같이 각 토픽에 대하여 3개 구간으로 나누어 구간별 변화를 그래프로 살펴보았다. 토픽 1 ~ 5를 살펴보면 공급망 관리에 관한 연구

<표 4.6> 국내 산업경영공학 분야의 기간별 주제 비중

순위	Period1 (2004~2007)		Period2 (2008~2011)		Period3 (2012~2015)	
	Topic Name	비중	Topic Name	비중	Topic Name	비중
1	[Topic 6] 최적화	9.473672	[Topic 6] 최적화	9.804514	[Topic 19] 품질경영	9.886752
2	[Topic 19] 품질경영	9.23735	[Topic 19] 품질경영	8.297122	[Topic 7] 기업성과	8.884477
3	[Topic 16] 시스템 모델링	9.149385	[Topic 1] 공급망 관리	7.760763	[Topic 1] 공급망 관리	7.535011
4	[Topic 1] 공급망 관리	6.926364	[Topic 16] 시스템 모델링	7.676058	[Topic 18] 의사결정시스템	6.142454
5	[Topic 8] 공정관리	6.769248	[Topic 7] 기업성과	6.086204	[Topic 16] 시스템 모델링	6.084063
6	[Topic 9] 컨테이너 터미널	6.105107	[Topic 8] 공정관리	5.436883	[Topic 6] 최적화	5.536091
7	[Topic 10] HW/SW 신뢰성	5.398049	[Topic 10] HW/SW 신뢰성	5.311044	[Topic 20] 데이터마이닝	5.331179
8	[Topic 3] 제조시스템 설계	5.299006	[Topic 9] 컨테이너 터미널	5.284985	[Topic 10] HW/SW 신뢰성	5.230732
9	[Topic 7] 기업성과	5.028223	[Topic 17] 서비스 품질	4.937283	[Topic 17] 서비스 품질	4.988439
10	[Topic 17] 서비스 품질	4.968572	[Topic 3] 제조시스템 설계	4.933293	[Topic 15] 텍스트마이닝	4.842852
11	[Topic 5] 차세대 웹 기술	4.367089	[Topic 20] 데이터마이닝	4.879716	[Topic 8] 공정관리	4.657117
12	[Topic 18] 의사결정시스템	4.355079	[Topic 12] 모바일 서비스	4.56465	[Topic 3] 제조시스템 설계	4.115861
13	[Topic 13] 인간공학	4.346164	[Topic 18] 의사결정시스템	3.986941	[Topic 4] 금융공학	3.90632
14	[Topic 20] 데이터마이닝	4.188607	[Topic 5] 차세대 웹 기술	3.941481	[Topic 13] 인간공학	3.88204
15	[Topic 12] 모바일 서비스	3.935851	[Topic 13] 인간공학	3.69549	[Topic 2] 헬스케어	3.627066
16	[Topic 14] 사용자 추천 시스템	3.218513	[Topic 11] DEA 효율성	2.817793	[Topic 9] 컨테이너 터미널	3.571553
17	[Topic 2] 헬스케어	2.017486	[Topic 14] 사용자 추천 시스템	2.744936	[Topic 12] 모바일 서비스	3.242974
18	[Topic 4] 금융공학	1.902954	[Topic 4] 금융공학	2.719819	[Topic 14] 사용자 추천 시스템	2.920593
19	[Topic 11] DEA 효율성	1.898419	[Topic 2] 헬스케어	2.690018	[Topic 5] 차세대 웹 기술	2.814294
20	[Topic 15] 텍스트마이닝	1.414861	[Topic 15] 텍스트마이닝	2.431007	[Topic 11] DEA 효율성	2.800131

는 지속적으로 수행되고 있고, 제조시스템 설계와 차세대 웹 기술에 관한 연구는 하락, 헬스케어와 금융공학의 관한 연구는 상승하는 것으로 나타났다.

토픽 6 ~ 10을 살펴보면 최적화에 관한 연구는 2012년 이후부터 연구가 큰 폭으로 하락하였고 기업성과는 2012년 이후 상승하는 것으로 나타났다

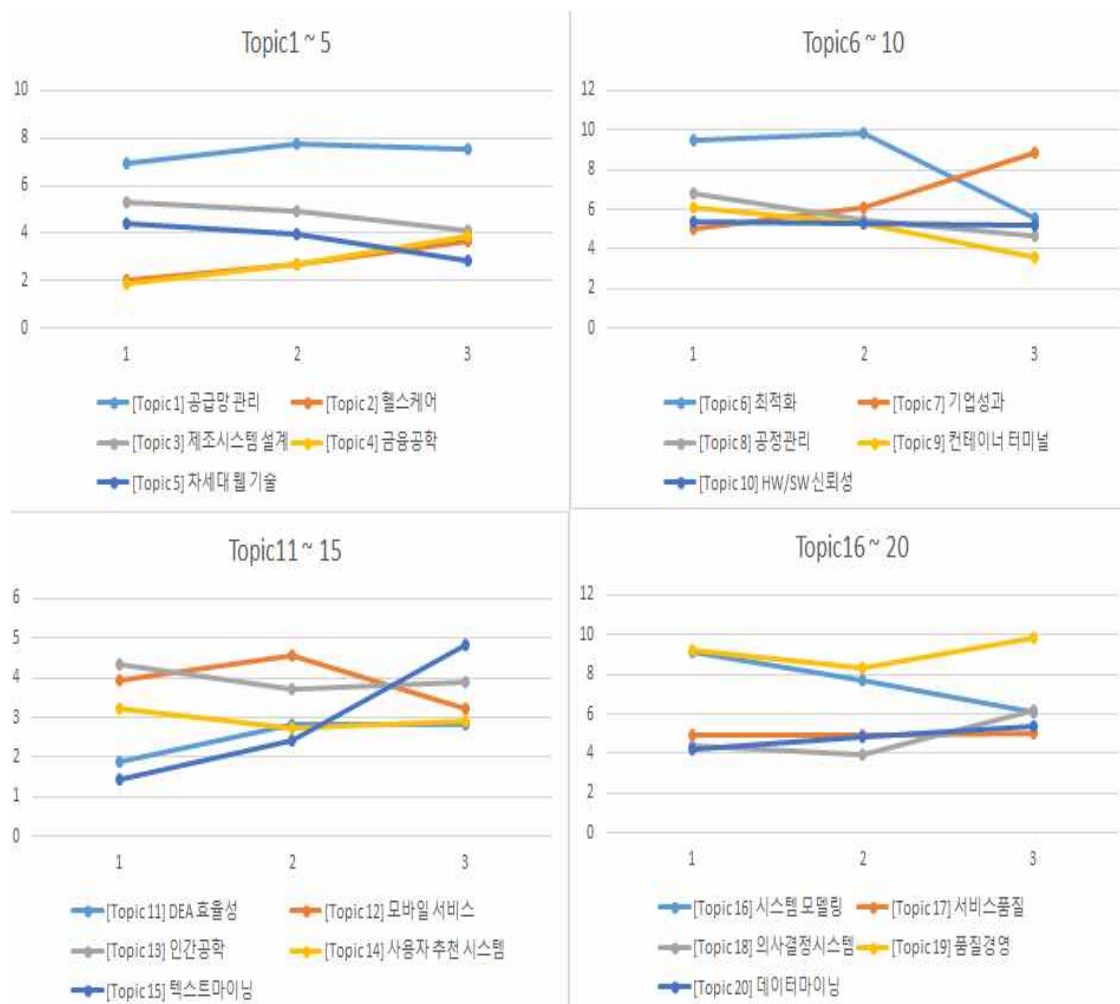
토픽모델링을 이용한 국내 산업경영공학 연구동향 분석

다. 또한 공정관리, 컨테이너 터미널 운영에 관한 연구는 하락하고 있으며 HW/SW 신뢰성에 관한 연구는 지속적으로 연구가 수행되는 것으로 나타났다.

토픽 11 ~ 15를 살펴보면 텍스트마이닝에 관한 연구는 지속적으로 상승하고 있는데 특히 2012년 이후 큰 폭으로 상승하는 것으로 나타났다.

모바일 서비스에 관한 연구는 2012년 이후 하락하는 것으로 나타났으며 DEA 효율성, 인간공학, 사용자 추천 시스템에 관한 연구는 지속적으로 수행되는 것으로 나타났다.

토픽 16 ~ 20을 살펴보면 시스템 모델링에 관한 연구는 하락하고 있으며 의사결정 시스템, 데이터



<그림 4.3> 토픽별 3구간 그래프

마이닝, 서비스 품질, 품질경영에 관한 연구는 지속적으로 수행되는 것으로 나타났다.

4.3.3 상향 주제(Hot Topic) 및 하향 주제(Cold Topic)

기간별 주제 분석을 통해 지속적으로 연구가 수행되고 있는 주제와 연구의 증가 및 감소하는 주제를 살펴보았다. 이는 기간별로 나타난 각 주제의 점유율을 산술평균하여 살펴본 것으로서 개별 주제에 대한 추세를 살펴보는 것에는 한계가 있다.

개별 주제에 대해 시간별로 추세를 살펴보는 방법으로 선행회귀모형을 구축하여 유의미한 회귀계수를 통해 상향 주제와 하향 주제를 확인할 수 있다(Griffiths & Steyvers, 2004). 독립변수는 2004 ~ 2015년에 대하여 1 ~ 12개의 구간이며 종속변수는 연도별에 따른 각 주제의 점유율이다.

유의수준 5%에서 유의한 주제를 선정하고 선정된 주제에 대하여 회귀계수 부호에 따라 (+)이면 상향 주제, (-)이면 하향 주제로 정하였다. <표 4.7>은 주제별 회귀계수와 유의수준을 나타낸 것이고 <그림 4.3>는 상향 주제(Hot Topic)와 하향 주제(Cold Topic)를 그래프로 나타낸 것이다.

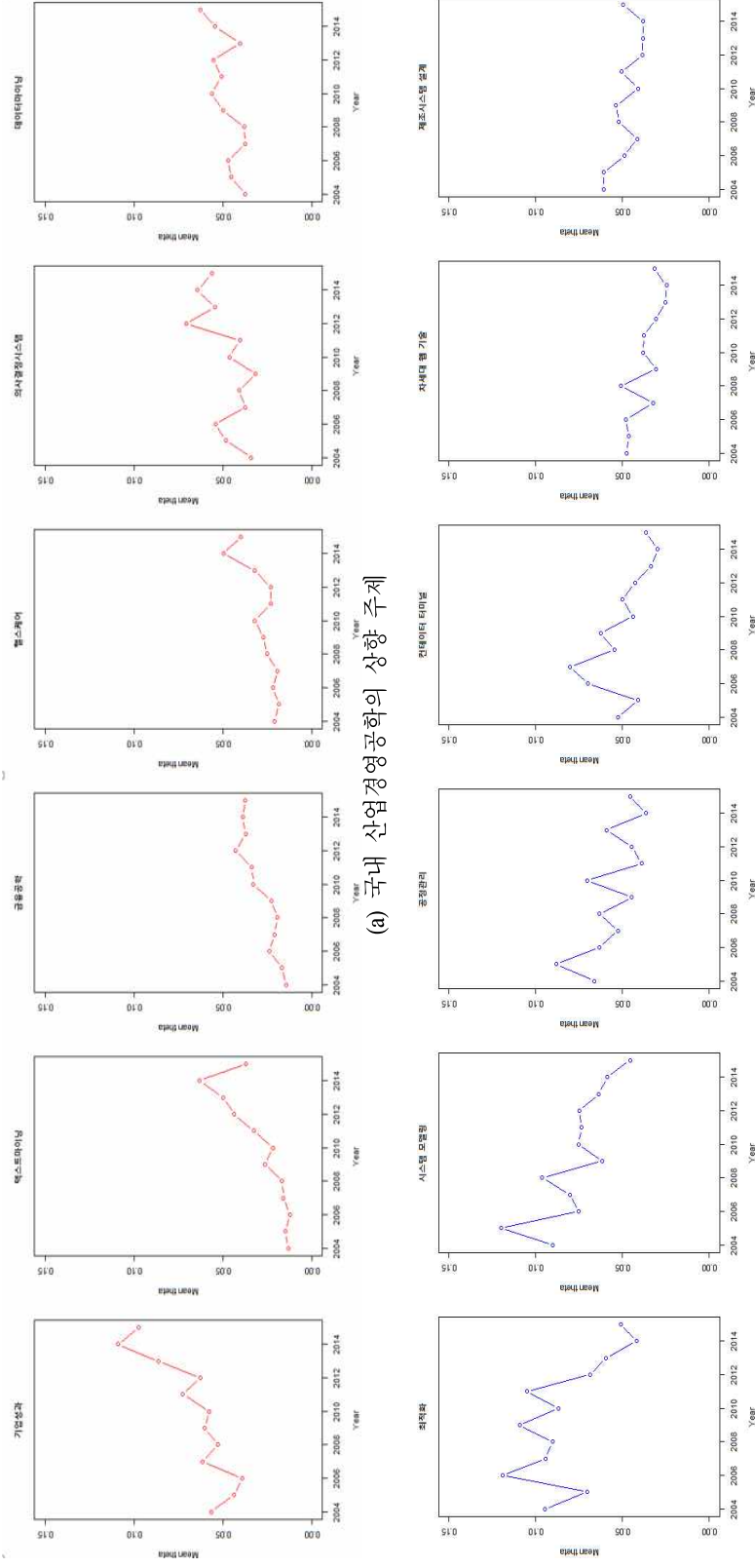
헬스케어, 금융공학, 기업성과, 텍스트마이닝, 의사결정시스템, 데이터 마이닝은 상향 주제(Hot Topic)로 나타났다. 기업성과와 금융공학은 지속적으로 상승하고 있는데 이는 국내 산업공학의 중요하게 다루는 분야인 것으로 보인다. 그리고 헬스케어, 텍스트마이닝, 데이터 마이닝의 연구는 빅데이터, IoT, 머신러닝 등 기술의 발전으로 상승 추세를 보이는 것으로 판단된다.

그러나 제조시스템 설계, 차세대 웹 기술, 최적화, 공정관리, 컨테이너 터미널 운영, 시스템 모델링의 연구는 하향 주제(Cold Topic)로 나타났다.

<표 4.7> 국내 산업경영공학 분야의 주제별 회귀 계수와 유의 수준

No	산업경영공학 주제	coefficient	P-value	Hot/Cold
1	[Topic 1] 공급망 관리	0.0004	0.642598	-
2	[Topic 2] 헬스케어	0.0020	0.002453	Hot
3	[Topic 3] 제조시스템 설계	-0.0015	0.020435	Cold
4	[Topic 4] 금융공학	0.0024	0.000297	Hot
5	[Topic 5] 차세대 웹 기술	-0.0019	0.002701	Cold
6	[Topic 6] 최적화	-0.0044	0.020907	Cold
7	[Topic 7] 기업성과	0.0049	0.000403	Hot
8	[Topic 8] 공정관리	-0.0028	0.016082	Cold
9	[Topic 9] 컨테이너 터미널	-0.0026	0.0295	Cold
10	[Topic 10] HW/SW 신뢰성	-0.000006	0.994349	-
11	[Topic 11] DEA 효율성	0.00101	0.101711	-
12	[Topic 12] 모바일 서비스	-0.0009	0.125428	-
13	[Topic 13] 인간공학	-0.0007	0.405472	-
14	[Topic 14] 사용자 추천 시스템	-0.0002	0.532074	-
15	[Topic 15] 텍스트마이닝	0.00398	0.000236	Hot
16	[Topic 16] 시스템 모델링	-0.004	0.002708	Cold
17	[Topic 17] 서비스품질	-0.0001	0.808984	-
18	[Topic 18] 의사결정시스템	0.00198	0.042725	Hot
19	[Topic 19] 품질경영	0.00119	0.310781	-
20	[Topic 20] 데이터마이닝	0.00156	0.017247	Hot

토픽모델링을 이용한 국내 산업경영공학 연구동향 분석



(b) 국내 산업경영공학의 하향 주제

<그림 4.3> Hot/Cold 주제에 대한 연도별 추세

4.4 산업경영공학의 주제 네트워크 분석

앞서 토픽모델링의 결과로 정의한 20개의 주제에 대하여 네트워크 분석을 수행하기 위해 주제 간 네트워크를 구축한 후 주제 간 네트워크 중심성 분석을 수행하였다.

4.4.1. 주제 간 네트워크 구축

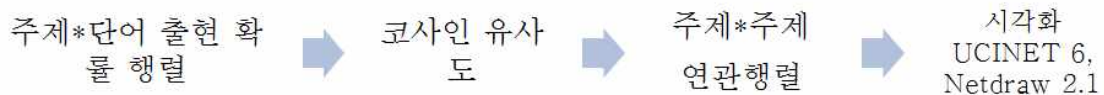
주제 간 네트워크 분석을 위한 순서는 <그림 4.4>과 같다. 20개의 주제에 대한 단어 출현 확률 행렬을 생성하고 생성된 행렬을 이용하여 주제 간 코사인 유사도를 산출하였다. 이를 통해 네트워크 시각화를 위한 주제 간 연관행렬을 생성하고 네트워크 시각화를 하였다. 주제 간의 연관행렬 데이터를 이용하여 시각화 프로그램 UCINET 6과 Netdraw 2.1을 활용하여 <그림 4.5>과 같은 주제 간 네트워크를 구축하였다. 노드 크기는 연결 중심성(Degree Centrality)이며 Cut-off는 0.19로 시각적으로 해석이 용이한 수준에서 정의하였다.

4.4.2. 주제 간 네트워크 중심성 분석

노드 크기가 연결 중심성으로 구축된 <그림 4.5>을 살펴보면 시스템 모델링, 공급망 관리, 공정관리, 기업성과, 데이터마이닝의 관한 연구가 연결 중심성이 높은 것을 알 수 있다. 이들 주제는 국내 산업경영공학 연구에서 허브역할(local hub)을 하고 있으며 핵심 주제라고 할 수 있다.

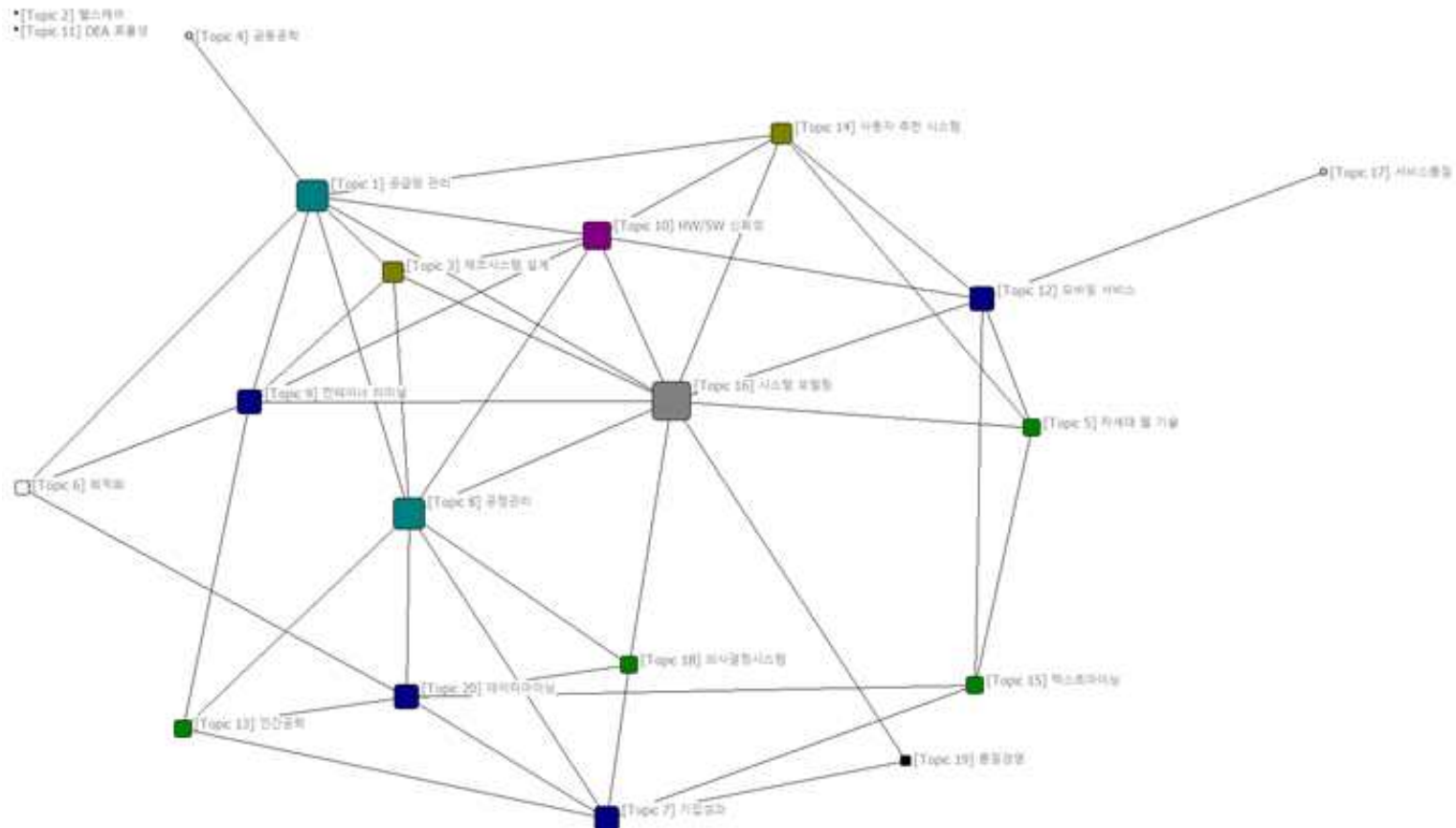
<표 4.8>은 정의된 주제에 대하여 연결 중심성의 근접 중심성, 위세 중심성, 매개 중심성 분석을 수행한 결과이다. 이들 4개의 중심성을 고려한 결과 시스템모델링, 공정관리, 공급망 관리, HW/SW 신뢰성, 제조시스템의 설계의 관한 연구가 전체적으로 중심성이 높게 나타났다. 이는 제조, 생산, 품질 등 전통적인 산업시스템 분야의 연구를 기반으로 하는 연구가 많이 수행되고 있는 것으로 보인다.

그리고 HW/SW 신뢰성 평가에 관한 연구는 특히 매개 중심성이 높게 나타났다. 이는 국내 산업공학에서 다루어지는 연구들 사이의 연결고리 역할을 하는 중요한 주제인 것으로 판단된다.



<그림 4.4> 주제 간 네트워크 분석 절차

토픽모델링을 이용한 국내 산업경영공학 연구동향 분석



<그림 4.5> 국내 산업경영공학 분야의 주제간 네트워크

김 상 결 · 장 성 용

<표 4.8> 국내 산업경영공학 분야의 주제 중심성

No	중심성					
	주제	연결중심성	주제	근접중심성	주제	위세중심성
1	[Topic 16] 시스템 모델링	0.526	[Topic 16] 시스템 모델링	0.559	[Topic 16] 시스템 모델링	[Topic 10] HW/SW 신뢰성
2	[Topic 1] 공급망 관리	0.421	[Topic 8] 공정관리	0.514	[Topic 10] HW/SW 신뢰성	[Topic 16] 시스템 모델링
3	[Topic 8] 공정관리	0.421	[Topic 1] 공급망 관리	0.5	[Topic 8] 공정관리	[Topic 1] 공급망 관리
4	[Topic 7] 기업성과	0.316	[Topic 10] HW/SW 신뢰성	0.495	[Topic 1] 공급망 관리	[Topic 12] 모바일 서비스
5	[Topic 9] 컨테이너 터미널	0.316	[Topic 9] 컨테이너 터미널	0.475	[Topic 3] 제조시스템 설계	[Topic 8] 공정관리
6	[Topic 12] 모바일 서비스	0.316	[Topic 3] 제조시스템 설계	0.463	[Topic 9] 컨테이너 터미널	[Topic 15] 텍스트마이닝
7	[Topic 20] 데이터마이닝	0.316	[Topic 7] 기업성과	0.463	[Topic 14] 사용자 추천 시스템	[Topic 20] 데이터마이닝
8	[Topic 3] 제조시스템 설계	0.263	[Topic 12] 모바일 서비스	0.463	[Topic 12] 모바일 서비스	[Topic 7] 기업성과
9	[Topic 14] 사용자 추천 시스템	0.263	[Topic 20] 데이터마이닝	0.463	[Topic 20] 데이터마이닝	[Topic 9] 컨테이너 터미널
10	[Topic 5] 차세대 웹 기술	0.211	[Topic 14] 사용자 추천 시스템	0.452	[Topic 18] 의사결정시스템	[Topic 14] 사용자 추천 시스템
11	[Topic 13] 인간공학	0.211	[Topic 18] 의사결정시스템	0.452	[Topic 7] 기업성과	[Topic 5] 차세대 웹 기술
12	[Topic 15] 텍스트마이닝	0.211	[Topic 5] 차세대 웹 기술	0.442	[Topic 5] 차세대 웹 기술	[Topic 6] 최적화
13	[Topic 18] 의사결정시스템	0.211	[Topic 15] 텍스트마이닝	0.422	[Topic 13] 인간공학	[Topic 13] 인간공학
14	[Topic 6] 최적화	0.158	[Topic 19] 품질경영	0.422	[Topic 6] 최적화	[Topic 18] 의사결정시스템
15	[Topic 19] 품질경영	0.105	[Topic 6] 최적화	0.413	[Topic 15] 텍스트마이닝	[Topic 19] 품질경영
16	[Topic 4] 금융공학	0.053	[Topic 13] 인간공학	0.413	[Topic 19] 품질경영	[Topic 3] 제조시스템 설계
17	[Topic 17] 서비스품질	0.053	[Topic 4] 금융공학	0.352	[Topic 4] 금융공학	[Topic 2] 웹스케이
18	[Topic 10] HW/SW 신뢰성	0.047	[Topic 17] 서비스품질	0.333	[Topic 17] 서비스품질	[Topic 4] 금융공학
19	[Topic 2] 웹스케이	0	[Topic 2] 웹스케이	0.2	[Topic 2] 웹스케이	[Topic 11] DEA 효율성
20	[Topic 11] DEA 효율성	0	[Topic 11] DEA 효율성	0.2	[Topic 11] DEA 효율성	[Topic 17] 서비스품질

5. 결 론

본 연구는 국내 산업경영공학 분야의 연구 동향을 규명하기 위해 한국한술지인용색인의 주제 분류 기준에 따라 산업공학으로 분류된 주요 학술지인, Industrial Engineering & Management Systems, 대한산업공학회지, 산업공학회지, 한국SCM학회지, 한국경영공학회지, 지능정보연구, 산업경영시스템학회지, 품질경영학회지의 2004 ~ 2015년까지 발표 논문의 영문 초록을 수집하여 토픽모델링을 수행하고 주제간 네트워크 분석을 수행하였다. 실험결과를 요약하면 다음과 같다.

첫째, 총 8개 학술지의 3251개의 초록을 분석하여 국내 산업공학 분야에서 연구되고 있는 20개 주제는 공급망 관리, 헬스케어, 제조시스템 설계, 금융공학, 차세대 웹 기술, 최적화, 기업성과, 공정관리, 컨테이너 터미널, HW/SW 신뢰성, DEA 효율성, 모바일 서비스, 인간공학, 사용자 추천 시스템, 텍스트 마이닝, 시스템 모델링, 서비스 품질, 의사결정시스템, 품질경영, 데이터 마이닝이다.

둘째, 토픽모델링의 실험 결과로 도출된 20개의 주제에 대하여 연도별 추이 및 점점 증가하고 있는 연구 주제와 하락하고 있는 연구 주제를 분석하였다. 2004년 ~ 2015년의 전체 기간 동안 품질경영, 최적화, 시스템 모델링, 공급망 관리, 기업성과는 많이 수행된 반면 사용자 추천 시스템, 텍스트 마이닝, 금융공학, 헬스케어, DEA 효율성 분석에 관한 연구는 적게 수행되었다.

또한, 기간을 3개의 구간으로 나누어 각 구간별 주제 비중을 분석하였다. 공급망 관리, HW/SW 신뢰성 분석, DEA 효율성, 인간공학, 사용자 추천 시스템, 의사결정시스템, 데이터 마이닝, 서비스 품질, 품질경영에 관한 연구는 지속적으로 수행되고 있는 것으로 나타났다. 그리고 헬스케어, 금융

공학, 기업성과, 텍스트마이닝의 관한 연구는 증가하고 있는 것으로 나타났으며 제조시스템 설계, 차세대 웹 기술, 공정관리, 컨테이너 터미널 운영, 모바일 서비스, 시스템 모델링의 관한 연구는 감소하는 것으로 나타났다.

개별 주제에 대해 회귀분석을 수행하여 유의미한 주제에 대한 상향 주제(Hot Topic)와 하향 주제(Cold Topic)를 분석하였다. Hot Topic은 헬스케어, 금융공학, 기업성과, 텍스트 마이닝, 의사결정시스템, 데이터마이닝에 관한 연구이고 Cold Topic은 제조시스템 설계, 차세대 웹 기술, 최적화, 공정관리, 컨테이너 터미널 운영, 시스템 모델링에 관한 연구이다. 전반적으로 기술의 발달에 따른 IT와 관련된 주제에 대한 연구가 증가 추세를 보이고 있는 것으로 판단된다.

셋째, 토픽모델링으로 도출된 20개의 주제들에 대하여 네트워크 분석을 수행하였다. 산업공학에서 허브 역할을 하고 있는 연구는 시스템모델링, 공급망 관리, 공정관리, 제조시스템 설계, HW/SW 신뢰성 분석에 대한 주제로 이들 주제는 연결 중심성이 높게 나타났다. 그리고 HW/SW 신뢰성 평가의 주제는 위세 중심성과 매개 중심성이 다른 주제들에 비해 높게 나타남으로서 산업공학의 연구에서 연결고리 역할을 하는 중요한 주제로 볼 수 있다.

본 연구는 기술의 발달로 기업환경이 급변하는 시점에서 기존의 산업경영공학의 연구동향 분석과는 다르게 국내 산업경영공학의 주요 학술지의 초록 데이터를 이용하여 토픽모델링과 네트워크 분석을 수행하였다. 이를 통해 산업경영공학의 연구주제에 대한 변화를 분석하여 새로운 정보를 제공했다는 것에 대해서 의의가 있다고 볼 수 있다.

본 연구의 결과로 도출한 토픽은 전체 연도(2004 ~ 2015)의 대하여 분석을 하였기 때문에 각 주제

에 대한 발전과정을 세부적으로 살펴보는 것에 한계가 있을 수 있다. 추후 연구로, 기간별 구간에 따라 토픽을 분석하여 각 주제에서 파생되는 세부적인 주제를 추출하여 발전과정을 살펴보는 연구가 필요하다.

참고문헌

- [1] 김형태. (2012). 융복합산업 및 기술시대에서의 산업공학의 역할. *ie 매거진*, 19(2), 13-19.
- [2] 박양병. (2006). 창의적이고 미래지향적인 산업공학 교육체계와 교과내용 개발. *ie 매거진*, 13(4), 14-15.
- [3] 이태억. (2003). 산업공학 교육의 혁신 - 미래의 교육방향. *ie 매거진*, 10(1), 20-25.
- [4] 정근하. (2011). 텍스트마이닝과 네트워크 분석을 활용한 미래예측 방법 연구. 한국과학기술기획평가원.
- [5] 최진영. (2014). 전국 산업공학 관련 학과 현황 분석. *ie 매거진*, 21(4), 11-15.
- [6] 한국교육개발원. (2015). 2015 학과(전공) 분류자료집.
- [7] 한상찬. (2012). 산업공학개론. 형설출판사.
- [8] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003), Latent dirichlet allocation, *Journal of Machine Learning Research*, 3, 993-1022.
- [9] Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- [10] Chakrabartij, S. (2002), *Mining The Web : Discovering Knowledge from Hypertext Data*, Morgan kaufmann publishers, 2-5, San Francisco, CA, USA.
- [11] Cho, S. G., Kim, S. B. (2012). Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining. *Korean Institute of Industrial Engineers*, 38(1), 67-73.
- [12] Cho, G. H., Lim, S. Y., Hur, S. (2014), An Analysis of the Research Methodologies and Techniques in the Industrial Engineering Using Text Mining. *Korean Institute of Industrial Engineers*, 40(1), 52-59.
- [13] Chung, D. K., Yun W. Y. (2015). A Study on the shop Management and Improvement Systems in Medium and Small-Sized Manufacturing Factories. *Journal of the Korea Management Engineers Society*, 20(3), 27-52.
- [14] Cha, Y. T. (2015). The Influence of Innovation Activities of Mainbiz Enterprise on Competitive Advantage and Management Performance. *Journal of the Korea Management Engineers Society*, 20(3), 87-103.
- [15] Choi, M. S., Han, K. S., Jeong J. A. (2011). A Study on how quality of education service impacts employee's job satisfaction and management performance. *Journal of the Korea Management Engineers Society*, 16(3), 159-181.
- [16] Dastkhan, H., Owlia, M. S. (2009). Study of trends and perspectives of Industrial Engineering research. *South African Journal of Industrial Engineering*, 20(1), 1-12.
- [17] Deerwester, S. C., S. T. Dumais, T. K. Landauer, G. W. Furnas, & R. A. Harshman. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science (JASIS)*, 41(6), 391-407.
- [18] Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks* 1(3), 215-239.
- [19] Griffiths, T. L., Steyvers, M. (2004). Finding scientific topics. *Proceeding of the National*

- Academy of Sciences, 101, 5288-5235.
- [20] Hofmann, T. (1999). Probabilistic latent semantic indexing. Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 50-57.
 - [21] Hornik, K., Grün, B. (2011). topicmodels : An R package for fitting topic models. Journal of Statistical Software, 40(13), 1-30.
 - [22] Lee, W. S., Song, S. Y. (2015). Topic Model Analysis of Research Trend on Spatial Big Data. Korean Institute of Industrial Engineers, 41(1), 64-73.
 - [23] Jeong, B.-K., Lee, H.-Y. (2014). Analyzing the Domestic Collaborative Research Network in Industrial Engineering. Journal of the Korean Institute of Industrial Engineers, 40(6), 618-627.
 - [24] Newman, M. (2010). Networks : An Introduction, Oxford University Press.
 - [25] Park, J. H., Song, M. (2013). A Study on the Research Trends in Library & Information Science in Korea using Topic Modeling. Journal of the Korean Society for information Management, 30(1), 7-32.
 - [26] Uys, J.W., Schutte, C.S.L., & Van Zyl, W. D. (2011). Trends in an international industrial engineering research journal: A textual information analysis perspective. International Conference on Computers & Industrial Engineering.
 - [27] Yoon, J., Kim, K. (2012). TrendPerceptor: A property - function based technology intelligence system for identifying technology trends from patents. Expert Systems with Applications, 39(3), 2927-2938.