

동적토픽모델링과 네트워크 분석 기반의 기술융합 동향분석 : ICT-Agritech 융합 사례 연구

Analysis of Technology Convergence by an Integrated Use of Dynamic Topic Modeling and Network Analysis : ICT-Agritech Case

저자 (Authors)	박상현, 최영재, 이성주, 설현주 Sanghyun Park, Yongjae Choi, Sungjoo Lee, Hyeonju Seol
출처 (Source)	대한산업공학회지 46(3) , 2020.6, 211-221(11 pages) Journal of the Korean Institute of Industrial Engineers 46(3) , 2020.6, 211-221(11 pages)
발행처 (Publisher)	대한산업공학회 Korean Institute Of Industrial Engineers
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09356257
APA Style	박상현, 최영재, 이성주, 설현주 (2020). 동적토픽모델링과 네트워크 분석 기반의 기술융합 동향분석 : ICT-Agritech 융합 사례 연구. 대한산업공학회지 , 46(3), 211-221
이용정보 (Accessed)	부산대학교 164.125.8.*** 2021/08/20 17:11 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

동적토픽모델링과 네트워크 분석 기반의 기술융합 동향분석 : ICT-Agritech 융합 사례 연구

박상현¹ · 최영재¹ · 이성주¹ · 설현주^{2*}

¹아주대학교 산업공학과 / ²충남대학교 국가안보융합학부

Analysis of Technology Convergence by an Integrated Use of Dynamic Topic Modeling and Network Analysis : ICT-Agritech Case

Sanghyun Park¹ · Yongjae Choi¹ · Sungjoo Lee¹ · Hyeonju Seol²

¹Department of Industrial Engineering, Ajou University

²School of Intergrated National Security, Chungnam National University

Since the 21st century, the ICT industry has contributed to the growth and sophistication of other industries as a driving force for the growth of the world economy, and has pioneered new technologies and industries by fusing with other technologies. Recently, ICT technology is facing a major change, such as rapid development of artificial intelligence related technology and commercialization of 5G low latency mobile communication technology, and the integration trend of ICT technology is also changing every moment. Research that analyzes convergence trends between technologies in a quantitative manner has mainly been focused on patent analysis. In particular, network analysis using patent bibliographic information to derive trends and at the patent keyword level are typical. This research also analyzes convergence trends at the level of patent keywords, but focuses on how the detailed topics of convergence technologies and the relationship between the topics change from a time-series perspective. For this purpose, we selected agricultural technology, which is one of the next-generation core technology fields, selected the integration with ICT-Agritech technology as a case study subject, and applied the dynamic topic modeling method of LDA-based analysis method. We tried to grasp the convergence trend by combining network analysis with the time-series change of topic keywords. Research from a theoretical perspective is expected to contribute to explain the changes in ICT-Agritech technology convergence since the Fourth Industrial Revolution. In terms of methodology, it is expected to present an approach that can support strategy development of convergence technologies by combining network analysis with the dynamic topic modeling method, which is an LDA-based methodology considering time series.

Keywords: Technology Convergence, Patent Analysis, Dynamic Topic Modelling, Network Analysis

제8회 산업융합 활성화 방안 및 사례연구 논문공모전 수상논문.

이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2018S1A5A2A01037180).

* 연락저자 : 설현주 교수, 34134 대전광역시 유성구 대학로 99 충남대학교 국가안보융합학부, Tel : 042-821-8531, Fax : 042-821-8868,

E-mail : hjseol@cnu.ac.kr

2020년 1월 2일 접수; 2020년 6월 3일 수정본 접수; 2020년 6월 9일 게재 확정.

1. 서론

오늘날 하나의 기술 분야 내에서의 단선적인 기술 개발 및 혁신은 이미 포화상태에 이른 것으로 인식되고 있다. 따라서 최근 미국, 유럽 국가 등 세계 각국에서는 국가와 기업을 대상으로 신성장동력을 창출함과 동시에 글로벌 경쟁력 확보를 위해 정부 주도의 광범위한 기술융합을 추진하고 있다(Kim and Jung, 2013; Kim *et al.*, 2018; Park *et al.*, 2012). 기술융합이란 일반적으로 산업 간에 명확하게 정의되고 확립된 경계에서 혁신이 일어날 때 발생하는 현상을 일컫는다(Heo *et al.*, 2019). 기술융합은 혁신 분야에서 기술 개발의 본질과 경향을 설명할 수 있는 중요한 키워드이기 때문에 기술융합의 가능성을 분석하고 발굴하는 것은 필수적이라고 할 수 있다(Kim and Lee, 2017).

특히 다양한 종류의 데이터 중에서 접근성과 활용성 측면에서 적용 범위가 광범위함에 따라 기술과 관련된 연구에 있어 오랫동안 유용한 데이터의 원천으로 여겨져 왔다(Ernst, 2003; Pilkington *et al.*, 2009; Zhang, 2011). 이러한 특성 때문에 특허 정보를 이용한 많은 연구가 이루어져 왔는데, 특허 분석을 통한 기술융합의 트렌드 도출(Caviggioli, 2016; Lee *et al.*, 2015; Jiang and Luan, 2018), 특허의 인용정보를 활용한 연구(Geum *et al.*, 2012; Karvonen and Käsä, 2011; Kho *et al.*, 2013; No and Park, 2010), 특허 분류 코드를 활용한 연구(Curran and Leker, 2011; Marki *et al.*, 2010; Meyer, 2007), 특허의 비정형 데이터를 활용한 연구(Moon, 2017; Song, 2019) 등이 있으며, 이외에도 특허 정보의 다양한 요소들의 혼합 및 다른 지표와 함께 사용한 연구들이 수행되어 왔다.

특히 최근에는 특허의 텍스트 정보를 기반으로 tf-idf, word2vec, 토픽모델링 기법 등과 같은 정량적인 방법을 활용하여 기술융합을 분석하는 연구가 많이 진행되고 있다. 이러한 기법들 중 토픽모델링은 기계 학습 및 자연어 처리 분야에서 문서 집합에 포함된 추상적인 주제를 발견하기 위한 통계적 모델 중 하나로, 문서 내에 잠재되어있는 의미구조를 발견하기 위해 사용된다. 토픽모델링 기법들 중 일반적으로 가장 많이 활용되는 알고리즘은 잠재디리클레할당(Latent Dirichlet allocation, LDA)으로, 주어진 문서에 대하여 단어의 패턴을 바탕으로 어떠한 주제들이 존재하는지를 파악하는 확률 모형이다. 최근에 이와 같은 LDA 기법을 사용하여 기술융합을 분석하는 다양한 연구(Lee *et al.*, 2017; Jeong *et al.*, 2016; Oh, 2015)들이 진행되고 있다.

하지만 LDA 기법을 활용한 연구들의 의미 있는 발견에도 불구하고, 기존 연구는 몇 가지 한계가 있다. 첫째, LDA 기법의 경우 개별 토픽이 전체 분석 기간에 걸쳐 변화 없이 유지된다는 전제하에 진행되기 때문에 대규모 말뭉치의 분석이 가능하지만, 토픽들이 상호 간에 영향을 주며 생성되고 사라지는 시계열적인 변화과정을 나타내지는 못하는 단점이 있다(Jin *et al.*, 2013). 또한 전체 분석기간에 자주 등장한 단어만이 핵심 키워드로 나타나기 때문에, 세부적인 관점에서 볼 때 핵심 키워드의 의미가 희미해지는 현상이 발생할 위험이 있다(Lim and Kim,

2014). 둘째, 기술융합은 기술 지식이 지속적으로 교류되기 때문에 기술융합의 동향을 파악하기 위해서는 정량적이고 체계적인 분석 방법이 필요하다(Seo, 2013). 하지만 기존 연구들은 이러한 기술융합의 동향을 세부적으로 파악하기 위한 이중 기술 간의 연결에 대한 정의가 명확하게 이루어지지 못하였다.

따라서 본 논문에서는 선행 연구들의 이러한 한계점을 극복하기 위하여 기존의 LDA 기법을 확장시킨 동적토픽모델링(Dynamic Topic Modeling, DTM) 기법을 사용하였다. 동적토픽모델링 기법을 사용함으로써 데이터의 시계열적인 흐름을 반영하였으며, 동적토픽모델에서 키워드 수준의 연결 정도를 정량화하고 이를 기반으로 동적 네트워크를 생성함으로써 이중 기술간의 기술융합의 동향을 살펴볼 수 있도록 하였다. 또한 제안한 방법의 유용성을 검증하기 위하여 이를 스마트 농업 부문에 적용하였으며, 거시적 수준에서의 전반적인 융합 동향 분석과 함께 미시적 차원에서의 상세한 기술융합 동향분석을 수행하였다. 분석 결과 ICT 기술과 농업 기술 분야와의 융합이 많이 이루어지는 기간에 네트워크가 활발히 확장되는 것을 발견하였는데, 이러한 확장은 특정 노드 중심으로 이루어지고 있음을 알 수 있었다.

본 연구에서 제안한 방법론은 특허 분석의 영역에 동적토픽모델링 기법과 네트워크 분석 기법의 결합을 기반으로 기술융합의 동향을 살펴보기 위한 새로운 시도로서 그 의미가 크다고 할 수 있으며, 실제로 ICT-Agritech 분야의 사례 연구를 통해 기술융합 관계와 동향을 정량적이고 다양한 관점에서 파악할 수 있음을 보여 주었다는 관점에서 의의가 있다고 생각된다.

이후 본 논문의 구성은 다음과 같다. 제 2장에서는 특허 데이터를 활용하여 기술융합 연구를 수행한 기존의 접근 방식을 검토하고, 동적 토픽모델링과 네트워크 분석에 대한 이론적 배경에 대해 살펴보았다. 제 3장에서는 제안한 접근법을 연구 아이디어 및 프로세스 측면에서 설명하였다. 제 4장에서는 제안한 접근 방법에 대한 사례 연구 및 이에 대한 결과를 제시하였다. 마지막으로 제 5장에서는 분석결과에 대한 논의와 더불어 연구의 한계점 및 향후 연구 방향에 대하여 언급하였다.

2. 이론적 배경

2.1 특허 데이터를 활용한 기술융합 연구

최근 특허 데이터를 활용하여 정량적인 연구에 초점을 맞춘 산업·기술융합 연구들이 많이 나타나고 있다(An *et al.*, 2016; Kim *et al.*, 2018). 특허 데이터를 활용한 기술융합 연구 방법 중 주로 활용되는 방법은 특허 인용 분석과 특허 동시분류분석 그리고 텍스트 분석이 있으며, 각각의 방법들을 사용한 연구들이 다수 진행되었다. 예를 들어, Karvonen and Käsä(2013)는 특허 인용 정보를 사용하여 초기 산업 영역 내에서의 융합의 가능성을 포착하고자 하는 연구를 진행하였으며, Yoo *et al.* (2019)은 자율주행 트럭 운송 산업의 기술융합을 식별하고 융합

기술 변화의 특성을 포착하기 위해 특허 인용정보를 사용하였다. Curran and Leker(2011)는 특허 IPC 코드를 활용한 동시분류분석을 통해 기술융합의 정도를 파악하였다. Noh and Ji (2019)는 PCT 특허의 동시분류분석을 실시하여 주요 국가별 고분자 소재의 기술융합 형태와 기술 수준을 분석하고자 하였다. 이외에도, 특허 내의 비정형 텍스트 데이터를 활용하여 기술 동향을 확인하거나 기술융합의 기회를 포착하려는 연구들은 Kim and Lee(2016), Guo *et al.*(2016), Yoon and Kim(2012) 등이 있으며, 특히 Nam and Choi(2018) 이나 Lim and Yoo (2018)의 경우 LDA 기법을 사용하여 기술 및 산업의 융합 가능성과 추세를 분석하고자 하였다. 또한, Chen *et al.*(2017)은 대규모 특허 데이터에 LDA 기법을 적용하여 특허의 콘텐츠 분석 및 트렌드 예측을 진행하고자 하였고, Bruens and Moehrl (2018)은 블록체인 기술의 특허 청구항 데이터에 tf-lag-idf 방식을 적용하여 특허 관점에서 기술의 확산 경로를 관찰하였고, Suominen *et al.*(2017)은 LDA 기법과 네트워크 분석을 특허 데이터에 적용하여 기업이 기술 인텔리전스를 관리하기 위해 빅데이터를 활용하는 접근 방식을 제안하였다. 하지만 앞의 논문들은 시간의 변화를 고려하지 못한 LDA 기법 기반의 토픽 모델을 사용하였으며, 이를 보완하기 위하여 시간의 흐름을 고려하여 기술의 동향을 분석한 연구들이 진행되었다. Ranaei and Suominen(2017)은 자동차 관련 대규모 특허 데이터에 LDA 토픽모델링과 DTM 기법을 각각 적용하고 두 모델을 사용한 결과의 차이를 도출 하였으며, DTM 기법을 사용했을 때, LDA와는 다른 결과를 얻을 수 있음을 확인하였다. Denter *et al.*(2019)은 RFID 기술 분야에 DTM 기법을 적용하여 기술의 확산 동향과 방향을 측정하고, 추가적으로 기존 토픽모델링과의 분석 결과를 비교하여, DTM이 기존의 방법보다 응용 분야로의 기술 확산을 측정하는데 보다 더 적합하다고 주장하였다. 마지막으로 Huang *et al.*(2019)은 5백만 개가 넘는 미국의 생명 의학 분야 특허 문서를 DTM 기법을 활용해 분석하였고, 시간이 지남에 따른 특허 문헌의 질병 범위와 잠재 토픽을 조사하였다.

정리해보면 기존 특허 데이터를 활용한 기술융합 연구는 거시적인 관점에서 융합의 추세나 가능성을 파악하는 데는 유용하나, 미시적인 관점과 시간의 흐름을 고려하지 못한 융합 기술 분석이라는 한계점이 존재한다. 따라서 본 연구에서는 시간의 흐름을 고려한 텍스트 마이닝 기법 중 하나인 동적 토픽 모델링 기법을 활용하여 거시적인 관점과 미시적인 관점에서 기술융합의 동향을 분석하고 시사점을 도출하는데 초점을 맞춘 방법론을 제안하고자 한다.

2.2 동적 토픽모델링

텍스트 마이닝은 구조화되지 않은 비정형 데이터인 텍스트 데이터에서 의미 있는 정보를 찾아내기 위한 기법으로(Tan, 1999), 최근 사회과학분야를 포함하여 과학기술분야에서도 자주

사용되는 기법 중 하나이다. 다양한 텍스트 마이닝 기법 중 LDA 기반의 토픽모델링은 문서 내 핵심 토픽을 찾아주고, 구조화되지 않은 비정형의 방대한 문서 집합에서 핵심 토픽을 쉽게 확인할 수 있는 확률 기반의 알고리즘이다(Wang and Blei, 2011).

본 연구에서는 LDA를 시계열적 토픽 분석에 적합하게 확장한 DTM 기법을 활용하였다(Blei *et al.*, 2010). DTM 기법은 연도, 월 등과 같이 시계열적인 흐름으로 데이터가 분할될 수 있다고 가정하며, 토픽과 토픽의 단어 비율 분포를 함께 묶음으로써 순차적인 토픽 모형의 연결이 가능하다(Bhadury *et al.*, 2016).

2.3 네트워크 분석

네트워크는 많은 양의 데이터를 체계적으로 구성하고, 분석 단위 간의 관계를 시각화하고, 이러한 관계 및 한 단위가 다른 단위에 미치는 영향을 정량화 하는 방법으로 널리 사용되었다(Borgatti and Cross, 2003; Borgatti *et al.*, 2009; Otte and Rousseau, 2002). 네트워크는 기본적으로 노드(node)와 엣지(edge)로 구성되며, 노드는 키워드, 논문 및 저자를 포함하여 특허 문서에 대해 다양한 수준으로 정의 할 수 있는 분석 단위이다(Noyons, 2001). 엣지는 노드 간의 관계를 나타내며 공동 인용, 공동 저자 및 공동 단어와 같은 여러 분석 지표를 사용하여 정의할 수 있다(Chen *et al.*, 2011; Lee and Jeong, 2008; White and McCain, 1997). 네트워크 분석에는 네트워크 안에서 엣지로 연결되어 있는 패턴을 통해 네트워크의 구조를 파악하는 속성 분석과 특정 노드가 전체 네트워크 구조 내 중심에 위치하는 정도를 의미하는 중심성 분석의 두 가지 분석 방법이 있다. 여러 연구에서는 네트워크 구조 내에서 영향력이 가장 큰 노드를 찾기 위한 중심성 지표를 사용하고 있으며, 중심성 지표에는 연결 중심성, 매개 중심성, 고유벡터 중심성, 근접 중심성 등이 있다. 본 연구에서는 DTM 기법을 통해 도출된 결과를 메타데이터로 활용하여 네트워크를 구성하였으며, 이때 엣지는 방향성을 가지지 않으나 개별로 가중치가 있어, 비방향성 가중치 네트워크를 구성하였다.

3. 연구절차

본 연구는 네트워크를 활용한 기술융합의 시계열적 동향 식별을 위한 방법론으로 제안하는 프로세스는 다음과 같다. [Phase 1] 동적 기술 네트워크 모델 생성 및 [Phase 2] 이전 결과물을 통해 융합 동향을 분석으로 이루어진다. [Phase 1]에서 특허 데이터에 동적토픽모델링을 적용하여 연도별 기술용어로 이루어진 네트워크의 메타데이터 매트릭스를 추출하고, 이를 통해 동적 기술용어 네트워크 모델을 생성한다. [Phase 2]에서 분석 수준 및 목적에 따라 메타데이터를 개별로 분석하는 Micro 수준의 동향 분석 또는 메타데이터를 합쳐서 분석하는 Macro 수준의 동향 분석을 진행한다.

Phase 1. 동적 기술용어 네트워크 모델 생성

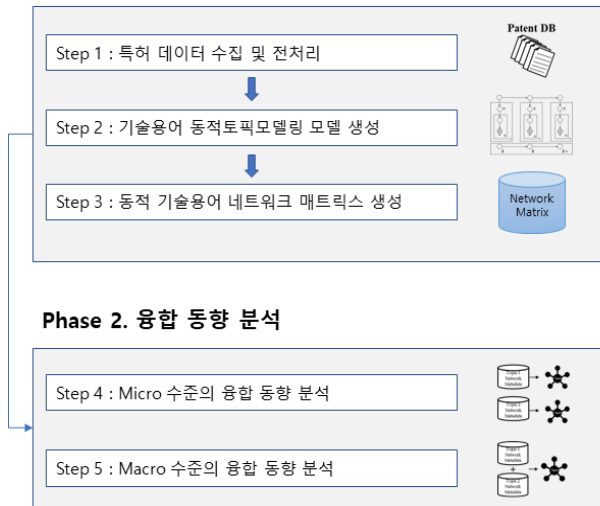


Figure 1. Dynamic Technical Term Network Model Creation and Convergence Trend Analysis

3.1 동적 기술용어 네트워크 모델 생성

[Phase 1]은 1) 특허 데이터 수집 및 전처리 2) 네트워크 메타데이터 생성을 위한 동적토픽모델링 적합 3) 동적 기술용어 네트워크 매트릭스 생성 3개의 step로 이루어진다.

(1) 특허 데이터 수집 및 전처리

먼저 융합 기술 특허 데이터를 수집하고 분석에 필요한 특허 정보 변인을 추출한다. 이후 주 분석대상인 텍스트에서 기술용어와 관련한 키워드를 추출하고 분석이 가능한 형태로 전처리를 수행한다.

(2) 기술용어 동적토픽모델 생성

특허 데이터 내 변인인 키워드와 시간 데이터를 통해 동적토픽모델링을 통해 모델을 생성한다. 토픽모델링 결과로 토픽수만큼의 연도별 확률로 이루어진 키워드 매트릭스가 생성된다. 본 방법론을 이중기술 간 융합에 초점이 맞추어져 있기 때문에 생성된 토픽에 할당된 특허들의 CPC 비율 및 대표 키워드를 통해, 이중 기술을 구분하여 상위 토픽 카테고리 라벨링을 수행한다. 예를 들어 토픽의 수를 20개로 설정하면, 20개의 토픽 별 키워드 연도별 확률변화를 나타낸 매트릭스를 얻을 수 있다. 해당 키워드 및 할당된 특허의 CPC 비율 분포를

확인한 결과를 통해 ICT 관련 기술, Nano 관련 기술 등 이중기술로 분류를 수행할 수 있다.

(3) 동적 기술용어 네트워크 매트릭스 생성

[Step 2]에서 생성한 네트워크 메타데이터 매트릭스에서 키워드로 노드(Node)를 생성하고, 노드간의 연결인 엣지(Edge)의 경우 두 키워드의 토픽 별 동시출현의 모든 조합을 고려하되 식 (1)과 같이 토픽 별 기술용어 확률 통계량의 기하평균을 계산하여 연관성을 추정한다.

이전 연구에서는 동일 토픽에서 용어가 상위에 출현한 경우 해당 토픽에서 두 용어의 관계가 있음을 착안하여 연결 혹은 연관성을 식별하였다. 하지만 연결/비연결의 이분법적 접근에는 한계가 있다. 이는 연구자가 토픽모델링에서 상위 용어를 몇 개까지 선택하는가에 따라 결과가 변할 수 있으며, 용어 간 상대적 연관성 차이를 정량적으로 반영하지 못한다.

상대적 연관성은 LDA에서 잠재토픽을 추정하는 과정에서 예시인 <Figure 3>에서 아이디어를 얻었다. 4차원 용어 공간에서 토픽 1과 가장 가까운 용어는 A이다. 따라서 용어 A는 토픽 1에서 가장 높은 용어 확률 통계량을 갖는다. 용어 A는 토픽 1과 가장 떨어진 용어 D와 거리가 멀다. 즉 특정 토픽에 가까운 (토픽 내 확률이 높은) 용어들 간의 거리는 잠재토픽을 기준으로 중심에 위치하기 때문에 상대적으로 가깝고, 특정 토픽으로부터 먼(토픽 내 확률이 낮은) 용어들 간의 거리는 잠재토픽을 기준으로 바깥부분에 위치하기 때문에 상대적으로 멀다.

따라서 용어 간의 연관성을 특정 토픽에서 확률을 통해 다음과 같이 계산하였다. 식 (1)에서 연관성은 특정 토픽 내 동일 기간에서 두 용어의 확률 통계량에 비례하며 두 용어 모두 토픽에서 핵심 용어일 때, 잠재토픽으로 인한 연관성이 강해지고 확률 통계량이 낮은 비핵심 용어일 때, 잠재토픽으로 인한 연관성은 낮아진다.

정리하면 <Figure 3(a)>에서 특정 토픽을 중심으로 가까운 용어들 간 거리의 평균은 잠재 토픽과 가까이 위치한 Term A가 Term B에 비해 낮기 때문에 용어 간 잠재토픽에 의한 연관성은 평균적으로 더 높을 것이라 추정할 수 있다.

$$y = \sqrt{P(\text{term}A)_t \times P(\text{term}B)_t} \quad (1)$$

y = Correlation of term A and term B of specific topic in period t

$P(\text{term}A)_t$ = Probability of term A of specific topic in period t

$P(\text{term}B)_t$ = Probability of term B of specific topic in period t

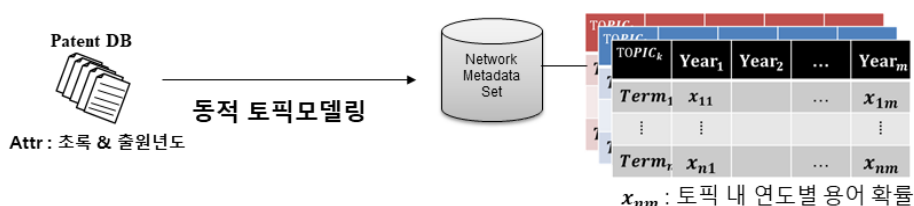


Figure 2. Network Metadata Generation Process

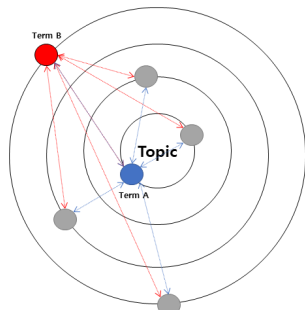
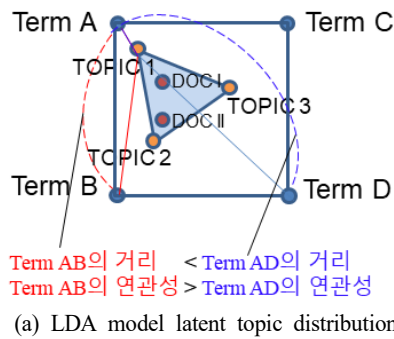


Figure 3. Example of Distance Between Terms

3.2 융합 동향 분석

[Phase 2]에서는 [Phase 1]에서 생성한 네트워크 데이터를 통해 [Step 4] Micro 수준의 모형과 [Step 5] Macro 수준의 모형을 생성하여 동향 분석을 실시한다. 두 단계는 개별적으로 수행되며 분석 범위 및 목적에 맞추어 필요에 따라 선택하거나 함께 진행한다.

(1) Micro 수준의 융합 동향 분석

네트워크 데이터를 개별 토픽 단위로 생성하여 개별 토픽들의 동적 기술용어 네트워크 분석을 진행하여 보다 세부적인 동향을 파악할 수 있다. 해당 모형은 특정 토픽에서 기술 용어의 평균성장률(Compound annual growth rate : CAGR) 혹은 평균 확률(토픽 내 중요도)과 같은 정보를 전반의 수준이 아닌 개별 수준으로 시각화하여 분석할 수 있는 장점이 있다.

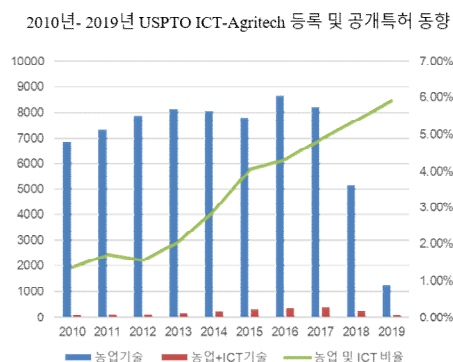


Figure 5. Current Status of ICT-Agritech Patents 2010~2019

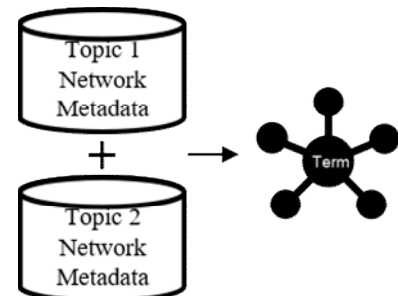
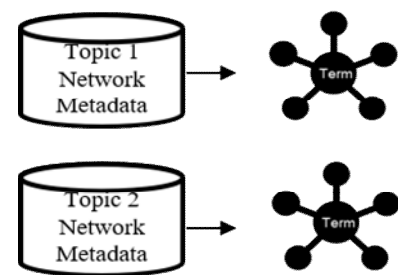


Figure 4. Micro and Macro Trend Analysis Database

(2) Macro 수준의 융합 동향 분석

두 개 이상의 토픽으로 구성된 네트워크 데이터를 결합하여 분석의 범위를 확대시킨 네트워크 데이터를 생성한다. 전반적인 동향 분석이나 토픽 간의 연결여부를 확인할 수 있지만, 네트워크가 보다 복잡해지고, 토픽 간의 기술용어가 공유되기 때문에 특정토픽 내 기술용어의 성장률과 같은 정보는 얻을 수 없다. 기술 융합의 관점에서는 해당 단계에서 이중 기술 분야로 구분된 토픽 간 키워드 연결의 동적 변화를 통해 융합 궤적을 식별할 수 있다.

4. 사례연구

본 장에서는, 최근 급격한 성장을 보이는 ICT 기술과 마찬가지로 핵심 기술분야 중 하나인 농업기술 분야(Agritech)를 융합 기술로 선정하여 사례연구를 진행하였다.

Agritech 대비 ICT-Agritech 특허 증가율

연도	Agritech	ICT-Agritech
2010	0.00%	0.00%
2011	6.88%	26.76%
2012	7.48%	-9.17%
2013	2.98%	32.63%
2014	-0.78%	40.98%
2015	-3.14%	39.27%
2016	10.62%	6.37%
2017	-4.81%	12.66%
2018	-37.26%	10.38%
2019	-75.36%	9.90%

1기간
2기간
3기간

4.1 Process 1 : ICT-Agritech 동적 기술용어 네트워크 모델 생성

미국 특허청(USPTO)에서 농업관련(A01), ICT관련(G06/H04) CPC를 가진 2010년부터 2019년까지의 2,181건의 ICT-Agritech 융합기술 특허를 수집하였다. 탐색적 분석 결과 전체 농업 기술 특허는 2010년부터 2019년까지 큰 변화가 없는 것에 비해 ICT-Agritech 융합기술의 비중은 높아지고 있었고, 해당 증가율을 토대로 기간을 3기간으로 나누어 분석하였다. 융합기술의 비중은 2기간(2013-2015)에 급격히 증가하여 2016년부터는 증가율이 점차 완만해진다. 이후 전처리를 수행하고 python gensim패키지를 활용하여 동적토픽모델링을 진행하였다. DTM 모형의 최적 토픽 수를 추정하기 위해 토픽의 수를 변경해가며 여러 LDA 모형을 먼저 생성하고, LDA 모형의 주제 일관성(Topic coherence) 지표에 대한 비교분석을 진행하여 최적토픽 개수를 15개로 결

정하였다. 동적토픽모델링에서 토픽 내 키워드의 동적변화를 결정하는 파라미터인 체인 파라미터(chain parameter)는 0.05로 설정하여 모델을 생성하였다.

이후 모든 특허를 가장 확률 통계량이 높은 토픽으로 분류하여 <Figure 6>과 같이 CPC 비율을 계산한다. 이를 참조하여 토픽을 명명하고 3가지 토픽인 Image, Signal, Animal 토픽을 타깃 토픽으로 삼아 분석을 진행하였다. 동적토픽모델링 결과, <Table 1>과 같이 연도별 용어 확률통계량의 동적 변화를 식별할 수 있다. <Table 1>과 같은 형태의 메타데이터를 바탕으로 기술용어 자체를 기반으로 노드 데이터셋, 기술용어 조합과 상관성을 기반으로 엣지 데이터셋을 작성하여 네트워크 모델을 생성할 수 있다. 결과적으로 Signal 토픽의 경우 총 노드의 수 93개, 총 엣지의 수는 375개를 가진 복잡한 네트워크를 형성하였다.

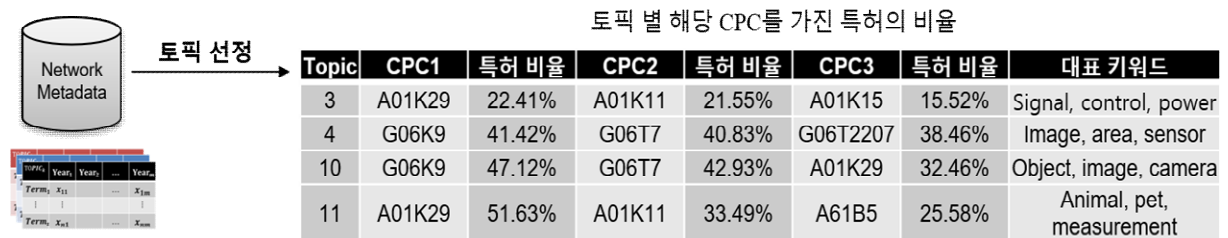


Figure 6. Representative Keywords and CPC for Selecting Topics

Table 1. Probability of the Top 20 Terms of Signal Topic

Term	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Avg	CAGR
signal	0.078	0.047	0.029	0.043	0.022	0.019	0.016	0.013	0.012	0.012	0.029	-16%
tag	0.040	0.034	0.032	0.060	0.032	0.019	0.016	0.009	0.009	0.007	0.026	-15%
power	0.017	0.017	0.024	0.020	0.017	0.018	0.027	0.032	0.024	0.017	0.021	0%
dog	0.017	0.020	0.018	0.008	0.006	0.006	0.000	0.000	0.000	0.000	0.007	-25%
control	0.015	0.018	0.031	0.024	0.026	0.018	0.013	0.016	0.027	0.039	0.023	9%
enclosure	0.015	0.010	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	-24%
source	0.014	0.011	0.010	0.010	0.010	0.011	0.011	0.008	0.006	0.005	0.010	-9%
sequence	0.013	0.011	0.008	0.006	0.007	0.004	0.000	0.000	0.000	0.000	0.005	-23%
detection	0.013	0.013	0.009	0.008	0.006	0.005	0.004	0.000	0.000	0.000	0.006	-23%
speed	0.011	0.009	0.009	0.011	0.008	0.006	0.004	0.000	0.000	0.000	0.006	-22%
location	0.010	0.011	0.012	0.011	0.009	0.008	0.007	0.007	0.007	0.007	0.009	-4%
distance	0.010	0.007	0.006	0.005	0.005	0.005	0.004	0.005	0.007	0.010	0.006	0%
detecting	0.009	0.010	0.009	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.003	-21%
output	0.009	0.008	0.007	0.007	0.006	0.005	0.004	0.004	0.004	0.000	0.005	-20%
supply	0.009	0.009	0.008	0.007	0.006	0.005	0.005	0.000	0.000	0.000	0.005	-20%
reading	0.008	0.008	0.007	0.006	0.005	0.004	0.000	0.000	0.004	0.000	0.004	-20%
connected	0.008	0.007	0.007	0.007	0.009	0.008	0.010	0.011	0.007	0.006	0.008	-3%
surface	0.008	0.009	0.009	0.008	0.007	0.009	0.007	0.007	0.006	0.006	0.008	-2%
pick	0.008	0.009	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	-20%
body	0.008	0.010	0.007	0.007	0.009	0.008	0.012	0.013	0.011	0.008	0.009	0%

Table 2. Edge Database of Signal Topic

source	target	weight	topic	year
signal	tag	0.056	3	2010
signal	power	0.037	3	2010
signal	dog	0.037	3	2010
signal	control	0.035	3	2010
signal	enclosure	0.034	3	2010
signal	source	0.033	3	2010
signal	sequence	0.032	3	2010
signal	detection	0.031	3	2010
signal	speed	0.029	3	2010
signal	location	0.028	3	2010
tag	power	0.026	3	2010
tag	dog	0.026	3	2010
tag	control	0.025	3	2010
tag	enclosure	0.024	3	2010
tag	source	0.024	3	2010
...
machine	heading	0.019	3	2019

4.2 Process 2 : ICT-Agritech 융합동향 분석

전체 노드 및 엣지 데이터 네트워크를 생성할 경우, 복잡성으로 인해 해석이 불가하여 Node의 Cut-off는 0.005 Edge의 Cut-off는 0.02로 설정하여 분석을 진행하였다.

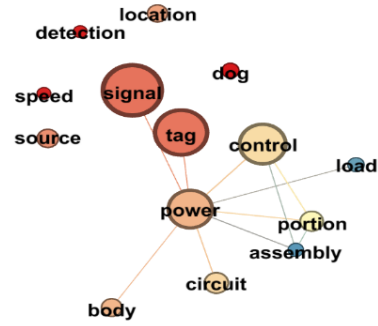
(1) Micro 수준의 융합동향 분석

Micro 수준의 동향 분석을 진행한 결과 대표적으로 Signal 토픽의 경우, 1기간<Figure 7(a)>에 ‘signal’, ‘tag’ 노드를 중심으로 네트워크가 형성되어 2기간<Figure 7-(b)>에 ‘detection’, ‘source’, ‘dog’ 노드의 확률 통계량 감소로 인해 네트워크의 연결이 끊어졌다가 3기간<Figure 7(c)>에 ‘power’ 노드를 중심으로 새로운 네트워크가 형성되는 것을 확인할 수 있다.

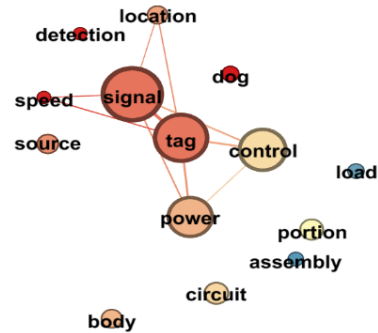
기간 별 네트워크 확장의 심화가 이루어지는 토픽과 노드를 정리하면 2기간에는 Image 토픽이 ‘animal’, ‘field’ 노드를 중심으로 결합하여 평균 Degree가 증가하였다. 3기간에는 Animal 토픽이 ‘pet’, ‘health’, ‘monitoring’ 노드를 중심으로 결합하여 평균 Degree가 증가하였다. 또한 Agritech-ICT간의 융합이 활발히 이루어지는 2기간에 Weighted degree가 전반적으로 높아지는 것을 확인하여, 이중 분야의 융합이 활발히 이루어지면서 기술용어간의 일관성 있는 연결 또한 강화되어 네트워크가 확장되는 것으로 추측할 수 있다.

(2) Macro 수준의 융합동향 분석

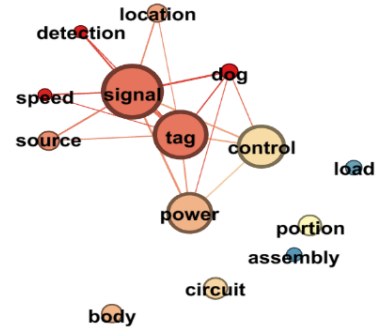
이후 Macro 수준의 동향 분석을 진행한 결과 1기간<Figure 9(a)>에 ‘object’, ‘camera’ 노드가 매개가 되어 Image 토픽과 Animal 토픽간의 연결을 식별하였고, 2기간<Figure 9(b)>에는



(a) Signal-topic micro network period 1(2010~2012)



(b) Signal-topic micro network period 2(2013~2015)



(c) Signal-topic micro network period 3(2016-2018)

Figure 7. Signal-Topic Dynamic Micro Network

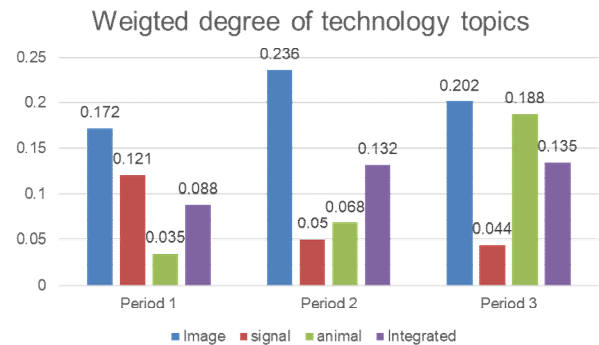
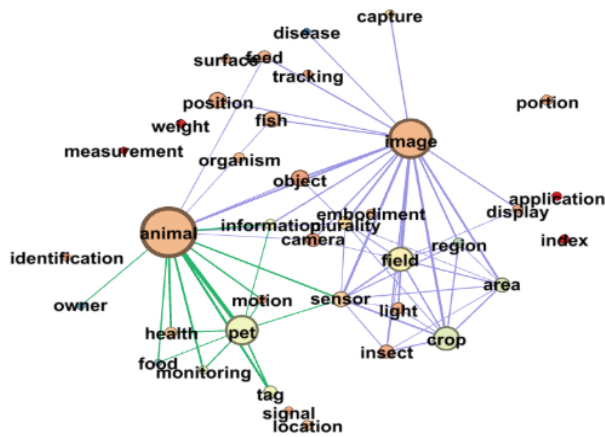
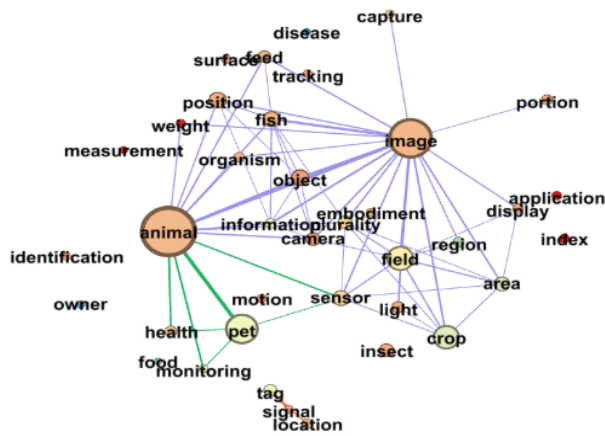


Figure 8. Average Weighted Degree by Period

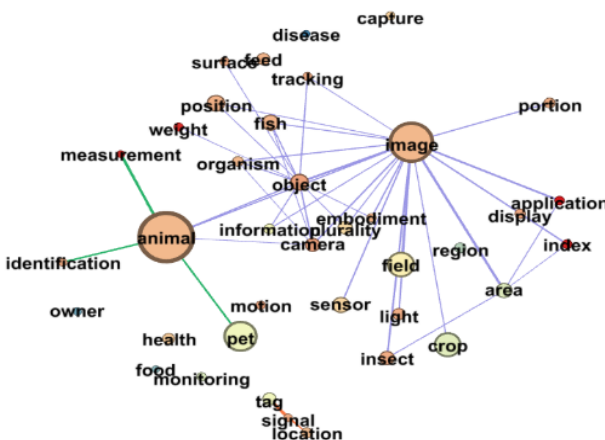
‘fish’, ‘organism’ 노드가 매개가 되어 두 토픽간의 연결을 식별하였고 마지막으로 3기간<Figure 9(c)>에 ‘pet’, ‘tag’ 노드가 매개가 되어 Signal 토픽과 Animal 토픽간의 연결성을 식별할 수 있다.



(a) Macro network period 1(2010~2012)



(b) Macro network period 2(2013~2015)



(c) Macro network period 3(2016~2018)

Figure 9. ICT-Agritech Dynamic Macro Network

Macro 수준의 융합 모델을 통해 세 토픽간의 연결을 정리하면 <Figure 10>과 같이 나타낼 수 있는데, Image와 Animal 토픽의 경우 1기간에 'camera', 'object' 노드로 연결되어 이후 2기간에 'fish' 및 'organism' 노드로 연결이 확대되고 3기간에는 Signal과 Animal 토픽이 'pet', 'tag' 노드로 인해 연결이 생기는 것을 확인하였다.

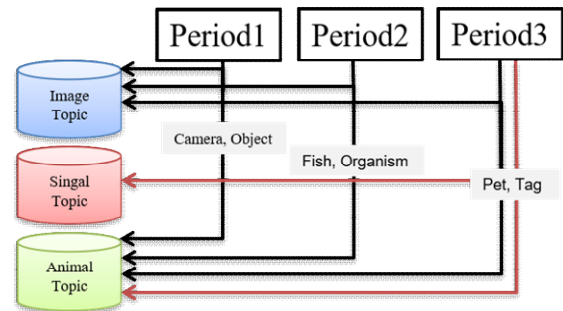


Figure 10. ICT-Agritech Dynamic Topic Link

Macro 수준의 융합 모델을 통해 세 토픽간의 연결을 정리하면 <Figure 10>과 같이 나타낼 수 있다. Image와 Animal 토픽의 경우 1기간에 'camera', 'object' 노드로 연결되었다. 1기간 내 특허의 명칭을 확인한 결과, 'Feeder Cam', 'Vision system for robotic attacher', 'Real-time insect monitoring device' 등 물체인식을 활용하여 동식물, 가축을 통제하는 기술들이 대거 출현하였다. 2기간에는 Image와 Animal 토픽의 연결이 'fish' 및 'organism' 노드로 연결이 확대되었다. 2기간 내 특허의 명칭을 확인한 결과, 'Watercraft automation and aquatic effort data utilization', 'Systems and methods for cultivating and distributing aquatic organisms' 등 동물관련 물체인식 기술이 육상기술 위주에서 수중 기술로 확대되었다. 3기간에는 Signal과 Animal 토픽이 'pet', 'tag' 노드로 인해 연결이 생기는 것을 확인하였다. 3기간 내 특허의 명칭을 확인한 결과, 'Transportable pet care center', 'Programmable training system for pets', 'Mobile Phone Clip and Pet Treat Holder' 등 기존 동식물 관련 기술이 기존 가축위주의 기술에서 반려동물 복지 관련 기술로 확대되는 양상을 확인하였다. Agritech와 ICT분야 내 융합 관점에서 정리하면, 이중분야 토픽인 이미지 토픽 내 키워드와 동물 토픽 내 키워드 네트워크는 2010년대 초반부터 중반에 주로 확대되었다. 2010년대 초반에는 가축생산성을 위한 이미지 인식과 같은 기술들이 대거 등장하였고, 이후 2010년대 중반 해당 기술은 수중 관련 기술까지 확대되었다. 또한 신호 토픽 내 키워드와 동물 토픽 내 키워드 네트워크는 최근(3기간: 2017~2019)에 형성되어 특정 신호를 감지하여 반려동물을 케어하는 기술이 해당 분야에서 주로 나타나고 있는 것을 확인하였다. 이는 최근 세계적으로 급증하는 반려동물 시장의 규모를 반영하며 매개 키워드인 'signal' 및 'tag'와 연결된 키워드가 'location', 'monitoring', 'health'인 것을 고려하면 반려동물에 'tag'를 통해서 발생한 'signal'의 응용 범위가 주로 'location', 'monitoring', 'health' 쪽임을 알 수 있다.

5. 결론

본 연구는 기술융합의 가능성을 분석하고 발굴하기 위한 하나의 방법론을 제시하였다. 본 방법론은 특허 분석의 일환으로 비정형 텍스트 데이터인 초록에 네트워크 분석을 적용하기 위해

동적토픽모델링 기법을 통해 메타 정보를 추출하는 과정을 거친다. 본 방법론은 동적토픽모델링과 네트워크 분석을 결합하여 기술융합분석의 새로운 패러다임을 제시하였으며 연구의 기여는 다음과 같이 정리할 수 있다. 첫째 동적토픽모델링을 적용하여 기존 융합 동향 연구에서 사용되었던 전통적 LDA 기법이 지닌 개별 토픽이 전체 분석기간에 걸쳐 변화 없이 유지된다는 전제로 인한 한계(Jin *et al.*, 2013)를 극복하여 토픽간의 상호 영향을 주는 시계열적인 변화과정을 나타낼 수 있는 모형을 제시하였다. 또한 네트워크 분석에서 연결에 대한 정의를 명확하게 한 연구가 부족한 현황(Seo, 2013)에서 기존 LDA를 활용한 네트워크 분석에 상관성 지표를 정의하여 보다 정량적이고 체계적인 방식으로 접근한 방법론을 제시하였다. 둘째, lda 모델에서 생성된 토픽을 파생된 키워드와 CPC 분포를 통해 이중 기술로 구분하고 이들 간의 동적 연결성을 통해 기술 궤적을 추정하여 융합의 매개체가 되는 기술용어를 동적으로 식별할 수 있는 방법론을 제시하였다. 셋째, 실무적으로는 분석가의 목적에 따라 Cut-off를 조정하여 다양한 시사점을 줄 수 있고 거시적, 미시적 관점의 접근 방법론을 제시하여 이를 적절히 선택하여 활용한다면 시간적, 경제적인 비용을 줄일 수 있을 것으로 예상된다.

본 연구는 비정형 텍스트 기반의 동적 네트워크 분석으로 텍스트가 지닌 비정형적 속성으로 인해 다양한 어려움이 존재하였다. 첫째, 기존 키워드 수준에서의 분석과 마찬가지로 기술용어에 대한 명확한 정의가 존재하지 않았기 때문에, 정확한 결과를 얻기 위해서는 전문가의 자문이 필수불가결해 보이며 전문가를 통해 기술 체계도를 구성하고 본 네트워크를 활용하여 기술 로드맵을 구축한다면 보다 정확한 기술 트렌드 분석이 가능할 것이라 예상된다. 또한 전문가의 도움으로 기술용어 동의어의 통일과 기술용어를 활용한 해석에 있어 신뢰성 있는 결과물을 만들 것이라 예상된다. 둘째, 특허에는 이러한 비정형 텍스트뿐만 아니라, 분류 체계, 인용정보, 출원인과 같은 중요한 정보가 있는데 본 연구에서는 제한적으로 활용하여 시사점을 주는 것에 한계가 있다.

본 연구는 방법론을 제안하는 것에 중점을 두어 동향을 분석하는데 그쳤지만, 후속연구로 이중분야 기술 매개 키워드를 중점으로 보다 정성적으로 접근하여 진행하게 된다면 이중분야에 대한 융합 가능성 탐색이 가능할 것으로 예상된다. 또한 본 연구는 기술의 동향을 분석하기 위한 특허를 활용한 방법론을 제안하였지만, 데이터의 범위를 확대하여 마찬가지로 시계열 정보를 가진 비정형 텍스트인 논문이나 기사와 같은 데이터를 함께 활용한다면 다양한 관점에서 기술융합에 대한 시사점을 줄 수 있을 것으로 예상된다.

참고문헌

- An, J., Kim, K., Noh, H., and Lee, S. (2016), Identifying Converging Technologies in the ICT Industry : Analysis of Patents Published by Incumbents and Entrants, *Journal of the Korean Institute of Industrial Engineers*, **42**(3), 209-221.
- Bhadury, A., Chen, J., Zhu, J., and Liu, S. (2016), Scaling up dynamic topic models, In Proceedings of the 25th International Conference on World Wide Web, *International World Wide Web Conferences Steering Committee*, 381-390.
- Blei, D., Carin, L., and Dunson, D. (2010), Probabilistic Topic Models : A focus on graphical model design and applications to document and image analysis, *IEEE Signal Processing Magazine*, **27**(6), 55.
- Borgatti, S. P. and Cross, R. (2003), A relational view of information seeking and learning in social networks, *Management Science*, **49**(4), 432-445.
- Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009), Network analysis in the social sciences, *Science*, **323**(5916), 892-895.
- Cavaggioli, F. (2016), Technology fusion : Identification and analysis of the drivers of technology convergence using patent data, *Technovation*, **55**, 22-32.
- Chen, D. Z., Huang, M. H., Hsieh, H. C., and Lin, C. P. (2011), Identifying missing relevant patent citation links by using bibliographic coupling in LEDilluminating technology, *Journal of Informetrics*, **5**(3), 400-412.
- Curran, C. S. and Leker, J. (2011), Patent indicators for monitoring convergence-examples from NFF and ICT, *Technological Forecasting and Social Change*, **78**(2), 256-273.
- Ernst, H. (2003), Patent information for strategic technology management, *World Patent Information*, **25**(3), 233-242.
- Geum, Y., Kim, C., Lee, S., and Kim, M. (2012), Technological convergence of IT and BT : evidence from patent analysis, *Etri Journal*, **34**(3), 439-449.
- Guo, J., Wang, X., Li, Q., and Zhu, D. (2016), Subject-action-object-based morphology analysis for determining the direction of technological change, *Technological Forecasting and Social Change*, **105**, 27-40.
- Heo, P. S. and Lee, D. (2019), Evolution patterns and network structural characteristics of industry convergence, *Structural Change and Economic Dynamics*, **51**, 405-426.
- Jeong, B., Kim, J., and Yoon, J. (2016), A Semantic Patent Analysis Approach to Identifying Trends of Convergence Technology : Application of Topic Modeling and Cross-impact Analysis, *The Journal of Intellectual Property*, **11**(4), 211-240.
- Jiang, Q. and Luan, C. (2018), Diffusion, convergence and influence of pharmaceutical innovations : a comparative study of Chinese and US patents, *Globalization and Health*, **14**(1), 92.
- Jin, S., Heo, G., Jeong, Y., and Song, M. (2013), Topic-network based topic shift detection on twitter, *Journal of the Korean Society for information Management*, **30**(1), 285-302.
- Karvonen, M. and Kassi, T. (2011), Patent analysis for analysing technological convergence, *Foresight*, **13**(5), 34-50.
- Karvonen, M. and Kassi, T. (2013), Patent citations as a tool for analysing the early stages of convergence, *Technological Forecasting and Social Change*, **80**(6), 1094-1107.
- Kho, J., Cho, K., and Cho, Y. (2013), A study on recent research trend in management of technology using keywords network analysis, *Journal of Intelligence and Information Systems*, **19**(2), 101-123.
- Kim, J. and Lee, S. (2016), Issues for Technology-Humanities Convergence : Patent Keyword Analysis, *Journal of Korean Institute of Industrial Engineers*, **42**(3), 231-240.
- Kim, J. and Lee, S. (2017), Forecasting and identifying multi-technology convergence based on patent data : the case of IT and BT industries

- in 2020, *Scientometrics*, **111**(1), 47-65.
- Kim, K., An, J., Yun S., and Lee, S. (2018), Deriving Convergence Opportunities between CT-IT Industry through Technical Structure Network Analysis, *Journal of Korean Institute of Industrial Engineers*, **44**(6), 404-414.
- Kim, K. and Jung, J. (2013), A Typology of Industry Convergences Based on Sources for Convergence Industries and Analysis of Critical Success Factors, *Journal of Korean Institute of Industrial Engineers*, **39**(3), 204-211.
- Lee, B. and Jeong, Y. I. (2008), Mapping Korea's national R&D domain of robot technology by using the co-word analysis, *Scientometrics*, **77**(1), 3-19.
- Lee, J., Lee, I., Jung, K., Chae, B., and Lee, J. (2017), Patents and Papers Trends of Solar-Photovoltaic(PV) Technology using LDA Algorithm, *Journal of Digital Convergence*, **15**(9), 231-239.
- Lee, W., Han, E., and Sohn, S. (2015), Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents, *Technological Forecasting and Social Change*, **100**, 317-329.
- Lim, M. and Kim, N. (2014), Analyzing the Issue Life Cycle by Mapping Inter-Period Issues, *Journal of Intelligence and Information Systems*, **20**(4), 25-41.
- Makri, M., Hitt, M., and Lane, P. (2010), Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions, *Strategic Management Journal*, **31**(6), 602-628.
- Meyer, M. (2007), What do we know about innovation in nanotechnology? Some propositions about an emerging field between hype and path-dependency, *Scientometrics*, **70**(3), 779-810.
- Moon, J., Gwon, U., and Geum, Y. (2017), Analyzing Technological Convergence for IoT Business Using Patent Co-classification Analysis and Text-mining, *Journal of Technology Innovation*, **25**(3), 1-24.
- No, H. and Park, Y. (2010), Trajectory patterns of technology fusion : Trend analysis and taxonomical grouping in nanobiotechnology, *Technological Forecasting and Social Change*, **77**(1), 63-75.
- Noh, J. S. and Ji, I. (2019), A Comparative Analysis of Convergence Types and Technology Levels of Polymer Technologies in Korea and Other Advanced Countries : Utilizing Patent Information, *Journal of the Korea Convergence Society*, **10**(3), 185-192.
- Noyons, E. (2001), Bibliometric mapping of science in a policy context, *Scientometrics*, **50**(1), 83-98.
- Oh, J. (2015), Identifying Research Opportunities in the Convergence of Transportation and ICT Using Text Mining Techniques, *Journal of Transport Research*, **22**(4), 93-110.
- Otte, E. and Rousseau, R. (2002), Social network analysis : A powerful strategy, also for the information sciences, *Journal of Information Science*, **28**(6), 441-453.
- Park, H., Seo, W., and Yoon J. (2012), Identifying Interdisciplinarity of Korean National RandD Using Patent CoIPC Network Analysis, *Journal of the Korean Society for Library and Information Science*, **46**(4), 99-117.
- Pilkington, A., Lee, L. L., Chan, C. K., and Ramakrishna, S. (2009), Defining key inventors : A comparison of fuel cell and nanotechnology industries, *Technological Forecasting and Social Change*, **76**(1), 118-127.
- Song, B. and Suh, Y. (2019), Identifying convergence fields and technologies for industrial safety : LDA-based network analysis, *Technological Forecasting and Social Change*, **138**, 115-126.
- Tan, A. (1999), Text mining : The state of the art and the challenges, *In Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, **8**, 65-70.
- Wang, C. and Blei, D. (2011), Collaborative topic modeling for recommending scientific articles, *In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 448-456.
- White, H. D. and McCain, K. W. (1997), Visualization of literatures, *Annual Review of Information Science and Technology*, **32**, 99-168.
- Yoo, D. H., Lee, B. K., and Sohn, S. Y. (2019), Analysis of patent citation network for identifying development trends of convergence technologies of self-driving truck industry, *Journal of the Korean Institute of Industrial Engineers*, **45**(1), 40-52.
- Yoon, J. and Kim, K. (2011) Detecting signals of new technological opportunities using semantic patent analysis and outlier detection, *Scientometrics*, **90**(2), 445-461.
- Zhang, L. (2011), Identifying key technologies in Saskatchewan, Canada : Evidence from patent information, *World Patent Information*, **33**(4), 364-370.

저자소개

박상현 : 아주대학교 산업공학과에서 2019년 학사를 취득하고 아주대학교에서 산업공학과 석사과정에 재학 중이다. 연구분야는 Patent analysis, User community analysis, Text-mining이다.

최영재 : 아주대학교 산업공학과에서 2019년 학사를 취득하고 아주대학교에서 산업공학과 석사과정에 재학 중이다. 연구분야는 Patent analysis, User innovation analysis, Machine-learning이다.

이성주 : 서울대학교 산업공학과에서 2002년 학사, 2007년 박사 학위를 취득하였다. 2007년 UCIC에서 선임연구원을 역임하고 2009년부터 아주대학교 산업공학과 교수로 재직하고 있다. 연구분야는 Technology roadmapping, Discovering new product/Service opportunities, Technology/Business forecasting, ICT and network technology management이다.

설현주 : 공군사관학교 및 서울대학교에서 각각 산업공학 학사 학위를 취득하였으며, 서울대학교 대학원에서 산업공학 전공으로 석사 및 박사학위를 취득하였다. 공군사관학교 교수를 거쳐 현재 충남대학교에 재직 중이며, 주요 연구 관심 분야는 Technology management, Process management, Industrial engineering applications in national defense areas 등이다.

<APPENDIX 1> 최적 토픽 수 추정

Table A1. 최적 토픽 수 추정을 위한 Topic coherence 그래프

