

비즈니스 기회 발굴을 위한 문제-해결방법 기반의 특허분석 방법*

정재민** · 정병기*** · 윤장혁****

I. 서론

II. 이론적 배경

1. 문제-해결방법 기반의 특허분석
2. 협업 필터링
3. 연관관계 분석

III. 분석모형

1. 특허 데이터 수집
2. 출원인별 포트폴리오 구축

3. 포트폴리오 기반 요소 추천

4. 요소 연관관계 파악

IV. 사례연구

1. 데이터 수집 및 전처리
2. 분석대상 기업 선정 및 포트폴리오
구축
3. 분석 결과

V. 결론 및 추후 연구

* 본 논문은 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 「대학혁신지원사업」의 연구결과임, 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2018R1D1A1B07045768).

** 건국대학교 산업공학과 석사과정생.

*** 건국대학교 산업공학과 석박사통합과정생.

**** 건국대학교 산업공학과 부교수, 교신저자.

초 록

최근 발명의 기술적 내용과 활용 분야를 구분하여 분석하는 문제-해결 방법 기반의 특허분석에 관한 관심이 증가하고 있다. 하지만 문제-해결방법 관점을 적용한 기존의 방법들은 구체적인 비즈니스 전략 수립을 지원하지 못하였고, 특허가 가지는 풍부한 정보를 충분히 활용하지 못하였다. 또한 일부 연구는 특정 국가의 특허 출원 양식을 활용하였기 때문에 타 국가의 특허에 적용이 어렵다는 한계점이 존재한다. 본 연구는 이러한 선행연구들의 한계점을 극복하기 위해 특허가 가지는 서지정보와 텍스트를 모두 활용하여 문제-해결방법 관점에서 비즈니스 기회를 발굴하는 분석 방법을 제시한다. 본 연구가 제시하는 방법은 1) 특허 데이터를 수집하고, 2) 출원인이 보유한 문제 및 기술 포트폴리오를 구축한 뒤, 3) 기업의 보유 역량을 기반으로 문제 및 기술요소를 추천한 다음, 4) 요소들의 연관관계를 통하여 비즈니스 및 기술 기회를 식별한다. 본 연구에서는 제시하는 방법의 실효성을 검증하고자 제약 산업의 대상 기업에 대한 사례연구를 진행하여 녹내장, C형 간염, 조현병, 허혈성 뇌졸중에 대한 비즈니스 기회를 식별하였다. 본 연구가 제시하는 방법은 객관적인 특허 데이터를 바탕으로 기업이 보유한 기술을 통하여 실현 가능성이 높은 비즈니스 기회를 발굴하며, 단순히 발굴에서 그치지 않고 문제요소와 기술요소의 연관관계를 파악하여 직접적인 비즈니스 전략 수립을 지원한다. 본 연구에서 제시하는 방법은 특정 산업의 기술요소 및 문제요소 모니터링을 위한 도구로써 활용될 수 있다.

주제어

특허지능, 기술기회 발굴, 비즈니스 개발, 문제-해결방법, 협업 필터링, 연관관계 분석

I. 서론

특허지능은 특허 정보의 획득, 저장, 결합, 분석을 통하여 최종적으로 전문가의 의사결정을 지원할 수 있는 착안점을 도출하는 전반적인 프로세스를 의미한다.¹⁾ 특허지능을 통하여 기술 모니터링, 기술 평가, 기술 예측과 같은 형태로 다양한 의사결정 정보를 획득할 수 있다.²⁾ 특히 최근의 비즈니스 환경이 초경쟁 시대로 변모함에 따라 기술경쟁 역시 심화되고 기술개발의 리스크가 급등하였기 때문에, 기업의 기술기획 활동에 있어 특허지능은 필수적인 것으로 여겨진다.³⁾

특허는 연구개발 결과를 담은 신뢰할 수 있는 최신의 기술 문서이며, 발명가, 출원인, 인용 등의 서지정보와 초록, 청구항 등의 텍스트 정보를 포함하고 있는 데이터이기 때문에, 다양한 관점에서의 기술 분석이 가능하다. 특허지능과 관련된 선행연구들은 텍스트 마이닝, 소셜 네트워크 분석, 추천 시스템 등을 활용하여 기술의 변화 동향을 파악하거나,⁴⁾⁵⁾ 향후 유망할 기술을 예측하거나,⁶⁾⁷⁾ 기업에 적합한 기술 또는 제품을 추천하는 방법론을 제시하

1) Lichenthaler, E., "Technology intelligence processes in leading European and North American multinationals", *R&D Management*, Vol.34 No.2(2004), pp.121-135.

2) Lichenthaler, E., "Technological change and the technology intelligence process: a case study", *Journal of Engineering and technology Management*, Vol.21 No.4(2004), pp.331-348.

3) 신영수 외 2인, "R&D 투자, 기술경영능력, 기업성과간의 관계", 『경영학연구』, 제38권 제1호(2009), 105-132면.

4) 정병욱 외 2인, "융합기술의 동향분석을 위한 의미론적 특허분석 접근방법: 토픽모델링과 교차영향분석의 활용", 『지식재산연구』, 제11권 제4호(2016), 211-240면.

5) Choe, H. et al., "Structural properties and inter-organizational knowledge flows of patent citation network: The case of organic solar cells", *Renewable and Sustainable Energy Reviews*, Vol.55(2016), pp.361-370.

6) Lee, W. S. et al., "Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents", *Technological Forecasting and Social Change*, Vol.100(2015), pp.317-329.

7) Érdi, P. et al., "Prediction of emerging technologies based on analysis of the US patent citation network", *Scientometrics*, Vol.95 No.1(2013), pp.225-242..

였다.⁸⁾⁹⁾

특히 최근의 특허지능 연구들에 있어, 문제요소와 기술요소를 분리하여 분석하는 연구들이 수행되고 있다. 특허가 가지는 문제요소는 출원인이 발명을 통해 해결하고자 하는 문제를 의미하며, 기술의 적용 대상 또는 출원인의 목표 비즈니스 영역으로 볼 수 있다.¹⁰⁾ 기술요소는 특허가 문제를 해결하기 위해 활용한 해결방법을 나타내며, 출원인의 보유 기술을 의미한다.¹¹⁾ 특허가 가지는 문제요소와 기술요소를 추출하여 문제-해결방법 관점에서 특허를 분석하여, 발명에 나타난 기술과 기술의 목적 및 활용 분야에 관한 추가적인 정보를 획득할 수 있으며, 기존기술을 응용할 수 있는 새로운 영역이나, 기존 영역에서의 위치를 공고히 할 새로운 기회를 발굴할 수 있다는 장점이 존재한다.

Yoon, J. & Kim, K.¹²⁾는 특허가 가지는 기술적 요소인 SAO(Subject-Action-Object) 구조를 추출하고, SAO 구조의 유사성을 통하여 경쟁기업의 기술적 능력을 식별하였다. SAO 구조는 “어떤 객체의 특징을 변화시키는 작용”으로 정의되는 기능정보를 표현하는 방법이다.¹³⁾ 특히 다양한 SAO 구조 중, AO가 기술적 문제를 설명하고 S가 해결방법을 나타내는 경우, 특허에 SAO 구조를 문제-해결방법 관계로 해석하여 기술에 대한 분석이 가능하

8) Yoon, J. et al., “Technology opportunity discovery (TOD) from existing technologies and products: A function-based TOD framework”, *Technological Forecasting and Social Change*, Vol.100(2015), pp.153-167.

9) Seo, W. et al., “Product opportunity identification based on internal capabilities using text mining and association rule mining”, *Technological Forecasting and Social Change*, Vol.105(2016), pp.94-104.

10) 최성철 외 2인, “기술로드맵핑을 위한 특허정보의 SAO 기반 텍스트 마이닝 접근 방법”, 『기술혁신연구』, 제20권 제1호(2012), 199-234면.

11) 이지호 외 4인, “특허의 Problem-Solution 텍스트 마이닝을 활용한 기술경쟁정보 분석 방법”, 『지식재산연구』, 제13권 제3호(2018), 171-204면.

12) Yoon, J. & Kim, K., “Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks”, *Scientometrics*, Vol.88 No.1(2011), pp.213-218.

13) Savaransky, S. D., *Engineering of creativity: Introduction to TRIZ methodology of inventive problem solving*, CRC Press, 2000.

다.¹⁴⁾ 이지호 외 4인¹⁵⁾은 한국, 중국, 미국, 일본, 유럽 등 IP5 국가들이 협약한 Common Application Formation(CAF) 항목들에 텍스트마이닝 기법을 적용하여 경쟁 동향분석 방법을 제안하였다. CAF의 항목 중, ‘배경 기술’과 ‘해결하려는 과제’ 항목을 문제로, ‘과제의 해결 수단’ 항목을 해결방법으로 정의하여 문제-해결방법 관점의 특허분석을 실시하였다.

이러한 연구들은 문제-해결방법 관점을 적용하여 특허를 분석한 초기연구라는 점에서 의의가 있지만, 여전히 일부 한계점이 존재한다. 첫째, SAO를 이용한 연구는 기업의 비즈니스 또는 제품들과 관계없이, 단순히 기술의 관점에서만 기회를 발굴하였다. 둘째, CAF 항목을 분석한 연구는 현실적으로 적용 가능한 특허의 범위가 제한된다. 예를 들어 한국 특허의 경우 CAF의 모든 항목을 필수적으로 채워야 하지만, 타 국가의 특허는 CAF 항목들을 선택적으로 작성하는 경우가 많으므로 ‘배경 기술’, ‘해결하려는 과제’, ‘과제의 해결 수단’ 항목만을 추출하여 분석하기가 어렵다. 마지막으로, 위 연구들은 텍스트 기반의 연구이기 때문에, 텍스트 데이터 외에 특허가 보유한 다른 서지정보를 활용하지 못하였다. 특허가 보유한 텍스트 데이터뿐만 아니라 인용 정보, 분류코드 등의 다양한 서지정보를 결합하여 분석에 사용한다면, 더욱 풍부한 기술 해석이 가능하다.¹⁶⁾

따라서 본 연구는 특허 텍스트에서 특허가 해결하고자 하는 문제요소를 추출하고, 특허의 기술 분류코드를 특허가 나타내는 기술요소로 정의하여 기업의 내부 역량에 적합한 비즈니스 및 기술 기회 발굴 방법을 제시한다. 본 연구에서 제시하는 방법은 1) 특허 데이터를 수집한 후 2) 출원인별 문제 포트폴리오와 기술 포트폴리오를 정의한 뒤, 3) 대상 기업에 적합한 문제요

14) Moehrle, M. G. et al., “Patent-based inventor profiles as a basis for human resource decisions in research and development”, *R&D Management*, Vol.35 No.5(2005), pp.513-524.

15) 이지호 외 4인, “특허의 Problem-Solution 텍스트 마이닝을 활용한 기술경쟁정보 분석 방법”, 『지식재산연구』, 제13권 제3호(2018), 171-204면.

16) 박영진 외 2인, “보유특허 기반의 기술기회탐색을 위한 특허추천방법: 3D 프린팅 산업을 중심으로”, 『지식재산연구』, 제10권 제1호(2015), 169-200면.

소 및 기술요소를 추천하고, 4) 대상 기업이 보유한 요소와 추천된 요소들을 통해 도출된 연관관계를 기반으로 기업에 적합한 문제-기술요소 조합을 식별하는 순서로 진행된다.

본 방법을 통해 도출된 문제-기술요소 조합은 보유한 문제요소와 새로운 기술요소 조합, 새로운 문제요소와 보유한 기술요소 조합, 새로운 문제요소와 새로운 기술요소 조합으로 나타나며, 각각의 조합에 따라 상이한 R&D (Research and Development) 및 비즈니스 전략 기획이 가능하다. 본 연구는 문제-해결방법 관점을 통하여 비즈니스 기회와 기술 기회를 동시에 식별한다는 점에서 의의가 있다. 본 연구에서는 기업의 보유 기술 및 비즈니스를 기반으로 기업이 접하지 않은 새로운 기술과 확장 가능한 시장을 식별하고, 이들의 연관관계를 파악하여 구체적인 비즈니스 기회를 정의한다. 또한 본 연구는 특정 국가의 정책과 관련이 없는 범용적 데이터인 텍스트 데이터와 기술 분류코드를 활용하여 특허의 문제요소와 기술요소를 정의하기 때문에, 다양한 국가의 특허를 대상으로 적용할 수 있으며, 텍스트 데이터만을 기반으로 하는 방법보다 합리적으로 특허의 기술적 내용을 정의할 수 있다는 장점이 있다.

본 논문은 2장에서 본 연구에 활용되는 이론적 연구들을 소개하고, 3장에서 연구의 구체적인 절차에 관하여 서술한 뒤, 4장에서 제약 특허를 대상으로 한 사례연구를 실시한다. 마지막으로 5장에서는 연구의 결론 및 토의점에 대하여 서술한다.

II. 이론적 배경

1. 문제-해결방법 기반의 특허분석

특허는 기업의 기술 연구개발의 최종 산출물이며, 발명에 대한 구체적인 기술 및 과학 정보를 담고 있는 신뢰할 수 있는 문서이다.¹⁷⁾ 기업이나 발명

가가 보유한 특허를 분석한다면 그들이 가지고 있는 기술적 역량에 대한 정의가 가능하며, 이를 통하여 기술 모니터링, 기술 평가, 기술 예측과 같은 형태로 다양한 의사결정 정보를 획득할 수 있다.¹⁸⁾ 특허를 기반으로 의사결정 정보를 찾아내는 특허지능 연구 중, 최근에는 특허에 나타난 문제요소와 기술요소를 분리하여 분석하는 연구들이 등장하였다. 특허의 문제요소는 발명이 기술적으로 해결하고자 하는 문제를 의미하며, 기술요소는 그 문제를 해결하기 위하여 사용된 해결방법을 의미한다. 문제-해결방법의 관점을 기반으로 특허를 분석한다면, 발명의 목적 및 활용 분야에 대한 정보를 획득할 수 있으며, 다수의 특허가 가지는 문제-해결방법 집합을 기반으로 기업의 기존 보유 기술을 응용할 수 있는 새로운 비즈니스 영역이나 기존 영역을 강화할 수 있는 새로운 기술 기회를 발굴할 수 있다.

문제-해결방법 관점에서 특허를 분석한 연구로, SAO 구조 기반의 특허분석 연구가 존재한다. SAO 구조는 문장 내에서 주제(명사구), 작용(동사구), 객체(명사구)의 순서를 가지는 구조로, 기술 텍스트에 나타나는 구성요소 간의 관계를 명확하게 제공한다.¹⁹⁾ 특히 AO가 기술적 문제를 나타내고, S가 해결방법을 나타내는 경우 위 SAO 구조는 문제-해결방법 관계로 해석할 수 있다.²⁰⁾ 이러한 SAO 구조의 개념을 통하여, Yoon, J. & Kim, K.²¹⁾는 특허에 나타난 SAO 구조의 유사도를 기반으로 한 의미론적 특허 네트워크를 제시

17) 노현정 & 임효정, “특허정보를 활용한 나노바이오 융합기술의 특성 분석”, 『지식재산 연구』, 제4권 제3호(2009), 109-129면.

18) Lichenthaler, E., “Technological change and the technology intelligence process: a case study”, *Journal of Engineering and technology Management*, Vol.21 No.4(2004), pp.331-348.

19) Yoon, J. & Kim, K., “Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks”, *Scientometrics*, Vol.88 No.1(2011), pp.213-218.

20) Moehrle, M. G. et al., “Patent-based inventor profiles as a basis for human resource decisions in research and development”, *R&D Management*, Vol.35 No.5(2005), pp.513-524.

21) Yoon, J. & Kim, K., “Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks”, *Scientometrics*, Vol.88 No.1(2011), pp.213-218.

하였다. 그들은 특허가 가지는 SAO 구조의 유사성을 통하여 특허 네트워크를 도출하고, 네트워크를 통하여 특허의 기술적 중요성, 특허 클러스터의 특성 및 경쟁기업의 기술적 능력을 식별하기 위한 새로운 지표를 제안하였다. 이 외에도 Choi, S. et al.²²⁾은 SAO 구조를 기반으로 기술의 사용 및 목적에 대한 정보를 나타내는 기능(function)을 정의하여 정량적인 기술 로드맵 구축 방법을 제시하였으며, Wang, X. et al.²³⁾은 문제-해결방법을 나타내는 SAO 구조를 통해 기업들의 의미론적 유사성을 기반으로 적합한 R&D 파트너를 선택하는 방법을 제시하였다. 문제-해결방법 관점을 기반으로 특허를 분석한 연구로, 특허의 CAF 항목을 활용한 연구도 존재한다. CAF는 한국, 미국, 일본, 유럽, 중국 등 5개국의 합의를 통하여 특허 문서의 작성방법을 통일화한 서식으로, 2010년부터 도입되었다. 이지호 외 4인²⁴⁾은 CAF 항목을 통하여 문제-해결방법 기반의 경쟁 동향분석 방법을 제시하였다. 그들은 다양한 CAF 항목 중 ‘배경 기술’ 항목과 ‘해결하려는 과제’ 항목을 문제로 정의하고, ‘과제의 해결 수단’ 항목을 해결방법으로 정의하였으며, 문제와 해결방법에 토픽모델링과 대응일치분석을 적용하여 기업들의 경쟁 현황을 파악하였다.

이와 같은 다양한 연구에도 불구하고, 문제-해결방법 기반의 비즈니스 및 기술 기회를 식별하기 위해서는 몇 가지 한계점이 존재한다. 우선, SAO 구조 기반의 연구들은 기업이 보유한 비즈니스 또는 제품을 고려하지 않고, 단순히 기술의 관점에서만 특허를 분석하였다. 선행연구들이 식별한 일반적인 기회는 특정 기업에 실제로 적용되지 않을 수 있으며, 이는 기업에 적합한 기회라고 보기 어렵다.²⁵⁾ 둘째, CAF 항목을 이용한 연구는 다양한 국가의

22) Choi, S. et al., “An SAO-based text-mining approach for technology roadmapping using patent information”, *R&D Management*, Vol.43 No.1(2013), pp.52-74.

23) Wang, X. et al., “Identifying R&D partners for dye-sensitized solar cells: a multi-level patent portfolio-based approach”, *Technology Analysis & Strategic Management*, Vol.31 No.3(2019), pp.356-370.

24) 이지호 외 4인, “특허의 Problem-Solution 텍스트 마이닝을 활용한 기술경쟁정보 분석 방법”, 『지식재산연구』, 제13권 제3호(2018), 171-204면.

25) Yoon, J. et al., “Identifying product opportunities using collaborative filtering-based patent analysis”, *Computers & Industrial Engineering*, Vol.107(2017), pp.376-387.

특허를 대상으로 분석 방법을 적용하기가 어렵다. 한국의 경우 특허 출원 시 CAF 항목을 모두 작성해야 하며 특허 문서에서 해당 항목들을 모두 구분하여 제공하지만, 미국 특허의 경우 모든 CAF 항목을 필수적으로 작성해야 한다는 제약이 없으며 항목을 개별적으로 구분하지 않는다. 따라서 한국이 아닌 타 국가의 특허를 분석하는 경우, CAF 항목을 모두 얻을 수 없을 가능성이 있을뿐더러, 원하는 항목만을 따로 추출하기 어렵다는 특징을 갖기 때문에, 다양한 국가의 특허에 대한 적용이 어렵다. 마지막으로, 위 연구들은 텍스트 데이터만을 기반으로 한 분석 방법이므로, 특허가 가지는 다양한 정보를 활용하지 못하였다. 특허는 텍스트 정보뿐만 아니라 인용, 분류코드 등의 다양한 정보를 포함하고 있으므로, 특허의 다양한 정보들을 통합하여 사용하면 더욱 합리적이고 포괄적인 기술 해석이 가능하다. 따라서 본 연구는 특허의 청구항에서 추출한 문제 개체를 발명이 가지는 문제요소로, 특허의 기술 분류코드를 발명의 기술요소로 정의하고, 문제요소와 기술요소를 기반으로 문제-해결방법 관점의 기회 발굴 방법을 제시한다. 선행연구들과 비교하였을 때, 본 연구에서 제시하는 방법은 기업의 보유 포트폴리오를 기반으로 기회를 식별하고, 특허 출원의 필수 항목을 활용하므로 특허의 다양한 정보를 활용한 범용성 있는 기회 도출 방법이라는 특징을 가진다.

2. 협업 필터링

협업 필터링은 비슷한 선호도를 보인 사용자들의 과거 선호도를 기반으로 한 추천 시스템이다.²⁶⁾ 협업 필터링은 추천 시스템의 사용 목적에 따라 사용자 기반의 협업 필터링 또는 아이템 기반의 협업 필터링으로 나눌 수 있으며, 사용한 알고리즘에 따라 모델 기반의 협업 필터링과 메모리 기반의 협업 필터링으로 나눌 수 있다.²⁷⁾ 사용자 기반의 협업 필터링은 사용자들의 과거

26) Herlocker, J. L. et al., "Explaining collaborative filtering recommendations", ACM, 2000

27) Li, W. et al., "A collaborative filtering recommendation method based on discrete

선호도 유사도를 기반으로, 특정 아이템에 대한 대상 사용자의 선호도를 예측하는 방법이다. 반면 아이템 기반의 협업 필터링은 아이템들이 가지는 선호도의 유사도를 기반으로 대상 사용자의 선호도를 예측한다. 모델 기반의 협업 필터링은 데이터 마이닝 기법을 통한 학습으로 사용자의 선호도를 예측하는 반면, 메모리 기반의 협업 필터링은 두 사용자 또는 두 아이템의 선호도 유사도를 통하여 특정 아이템에 대한 사용자의 선호도를 예측한다.

협업 필터링은 사용자들의 선호도를 기반으로 대상 사용자가 접해 보지 못한 아이템에 대한 선호도를 예측한다. 선호도 기반의 추천이라는 특성 때문에, 협업 필터링은 아이템의 특성을 정량적으로 정의하지 않고도 사용자의 선호도를 예측한다는 특징을 가진다.²⁸⁾ 또한 메모리 기반의 협업 필터링은 타 예측 모델과는 달리, 복잡한 데이터 마이닝 알고리즘 없이 유사도만으로 예측할 수 있다는 장점이 존재한다.²⁹⁾ 기술 데이터를 활용한 연구에서도 협업 필터링이 다수 사용되었는데, Yoon, J. et al.³⁰⁾은 특허에 나타난 제품 키워드를 통하여 기업의 내부 역량에 적합한 제품기회를 식별하였으며, Park, Y. & Yoon, J.³¹⁾는 기업이 보유한 특허의 IPC 코드를 추천하여 기업이 시도해 볼 만한 기술 기회를 탐색하였다.

본 연구에서는 문제 및 기술요소를 추천하기 위하여 사용자 기반, 메모리 기반의 협업 필터링을 사용한다. 일반적으로, 특허에 나타난 문제요소 또는 기술요소의 특징을 정량적으로 정의하기 위하여 해당 기술 분야에 대한 전

quantum-inspired shuffled frog leaping algorithms in social networks”, *Future Generation Computer Systems*, Vol.88(2018), pp.262-270.

28) Good, N. et al., “Combining collaborative filtering with personal agents for better recommendations”, AAAI/IAAI, 1999.

29) Sarwar, B. M. et al., “Item-based collaborative filtering recommendation algorithms”, *Proceedings of the 10th international conference on World Wide Web*, 2001.

30) Yoon, J. et al., “Identifying product opportunities using collaborative filtering-based patent analysis”, *Computers & Industrial Engineering*, Vol.107(2017), pp.376-387.

31) Park, Y. & Yoon, J., “Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering”, *Technological Forecasting and Social Change*, Vol.118(2017), pp.170-183.

문지식이 필수적이다. 하지만 협업 필터링은 요소에 대한 사용자의 과거 선호도만을 사용하기 때문에, 요소의 특성을 정의하기 위한 전문가의 개입 없이 정량적으로 문제요소와 기술요소의 추천이 가능하다. 또한 사용자 기반의 협업 필터링은 사용자가 이미 보유하고 있는 선호도를 통해 요소를 추천하므로, 기업의 내부 보유 역량을 기반으로 시도해보지 않은 비즈니스 기회를 식별할 수 있다.

3. 연관관계 분석

연관관계 분석은 하나의 거래 정보에 같이 등장한 아이템들은 어떠한 관계가 있다는 가정을 기반으로, 거대한 데이터 안에 존재하는 개체들의 숨겨진 관계를 찾아내는 데이터 마이닝 기법이다.³²⁾ 연관관계 분석에서는 개체들의 관계 정도를 계량화하기 위해 Support와 Confidence, Lift라는 세 가지 측정치를 사용한다. Support를 통하여 해당 규칙이 빈도가 많은 좋은 규칙 인지를 판단할 수 있으며, Confidence를 통하여 해당 규칙이 유용한 규칙인지를 파악할 수 있고, Lift를 통하여 규칙을 이루는 아이템들의 상관관계 여부를 파악할 수 있다.³³⁾ $X \rightarrow Y$ 규칙을 예로 들면, 연관관계 분석을 위한 세 가지 측정치는 〈식 1, 2, 3〉을 통하여 구할 수 있다.

$$Support = P(X \cap Y) \quad \langle \text{식 1} \rangle$$

$$Confidence = P(Y|X) \quad \langle \text{식 2} \rangle$$

$$Lift = P(Y|X)/P(Y) \quad \langle \text{식 3} \rangle$$

연관관계 분석은 개체들의 가능한 모든 조합에 대하여 세 측정치를 구해

32) Shih, M. -J. et al., "Discovering competitive intelligence by mining changes in patent trends", *Expert Systems with Applications*, Vol.37 No.4(2010), pp.2882-2890.

33) Seo, W. et al., "Product opportunity identification based on internal capabilities using text mining and association rule mining", *Technological Forecasting and Social Change*, Vol.105(2016), pp.94-104.

야 하므로 연산량이 매우 크다.³⁴⁾ 이러한 문제점을 해결하기 위하여, 일반적으로 Apriori 알고리즘이 사용된다.³⁵⁾ 연관관계 분석에서 Apriori 알고리즘은 가능한 모든 조합의 규칙을 생성한 뒤, Support 수치가 미리 정한 threshold를 넘지 않는 규칙을 제거하고, 남은 규칙 중 Confidence 수치가 미리 정한 threshold를 넘지 않는 규칙을 제거하는 순서로 진행된다.

연관관계 분석은 눈으로 파악하기 어려운 개체들의 관계를 찾아낼 수 있으므로 데이터를 기반으로 한 다양한 연구에서 활용되었다. 고객의 구매 데이터를 통한 고객 행동 조사,³⁶⁾ 헬스케어 산업의 데이터 품질 평가,³⁷⁾ 제품 키워드의 연관관계를 통한 기업 내부 역량 기반의 제품기획 식별³⁸⁾ 등의 연구에서 연관관계 분석을 사용하였는데, 이는 연관관계 분석이 분야에 상관 없이 데이터상에 나타난 개체들의 관계를 찾는 데 유용함을 의미한다. 본 연구에서는 기업이 보유한 요소들과 추천된 요소들을 기반으로, 문제요소에 적합한 기술요소 또는 기술요소에 적합한 문제요소를 찾아내기 위하여 연관관계 분석을 사용한다. 모든 문제에는 그에 적합한 해결방법이 존재하며, 부적절한 해결방법을 사용한다면 문제에 대한 해결 효과가 전혀 없을 수 있다. 본 연구에서는 연관관계 분석을 통하여 특히 데이터에 나타난 문제요소와 기술요소의 유의미한 규칙을 생성하고, 생성된 규칙을 통하여 문제요소에

34) Chiclana, F. et al., "ARM-AMO: an efficient association rule mining algorithm based on animal migration optimization", *Knowledge-Based Systems*, Vol.154(2018), pp.68-80.

35) Abaya, S. A., "Association rule mining based on Apriori algorithm in minimizing candidate generation", *International Journal of Scientific & Engineering Research*, Vol.3 No.7(2012), pp.1-4.

36) Kaur, M. & Kang, S., "Market Basket Analysis: Identify the changing trends of market data using association rule mining", *Procedia computer science*, Vol.85 (2016), pp.78-85.

37) Peng, M. et al., "Exploration of association rule mining for coding consistency and completeness assessment in inpatient administrative health data", *Journal of biomedical informatics*, Vol.79(2018), pp.41-47.

38) Seo, W. et al., "Product opportunity identification based on internal capabilities using text mining and association rule mining", *Technological Forecasting and Social Change*, Vol.105(2016), pp.94-104.

적합한 기술요소를 식별한다. 식별한 문제-기술요소 조합은 그 자체로 기술 기회를 의미하며, 기술 기회를 실행으로 옮기기 위한 R&D 및 비즈니스 방향으로 해석할 수 있다.

III. 분석모형

본 연구가 제시하는 분석모형은 〈그림 1〉과 같은 방법을 따른다. 1) 우선, 분석에 필요한 특허 데이터를 수집하고, 2) 문제요소와 기술요소를 정의하여 출원인들의 문제 포트폴리오와 기술 포트폴리오를 구축한다. 3) 이후 협업 필터링을 통하여 보유 포트폴리오를 기반으로 문제요소와 기술요소를 추천한 뒤, 4) 보유 요소와 추천 요소를 통합하여 요소들의 연관관계를 생성하고, 생성된 규칙을 통해 대상 기업의 내부 역량을 기반으로 실현 가능성이 충분한 비즈니스 및 기술 기회를 식별한다. 본 장에서는 위에 기술한 분석 방법을 상세하게 설명한다.

[그림 1] 전체 분석절차



1. 특허 데이터 수집

본 연구가 제안하는 방법은 분석하고자 하는 특허 데이터를 수집하는 것 으로부터 시작된다. 특허 데이터 수집 단계는 분석대상이 되는 특허에 대한 검색조건을 설정하는 과정을 포함하는데, 검색조건은 일반적으로 원하는 분야의 기술적 키워드나 관련 기술의 분류코드를 기반으로 생성된다. 이후 분석에 필요한 특허 데이터는 PatentView(<http://www.patentsview.org>), KIPRIS (<http://www.kipris.or.kr>) 등과 같은 공공 서비스, 또는 WIPS-ON(<http://www.wipson.com>)과 같은 민간 서비스를 통하여 수집할 수 있다. 수집한 특허 데이터는 마이크로소프트 엑셀이나 텍스트 파일 등의 형태로 저장할 수 있으며, 저장된 데이터는 청구항, 출원인, 분류코드 등의 정보를 포함한다.

2. 출원인별 포트폴리오 구축

본 단계에서는 앞서 수집한 특허를 이용하여 출원인의 내부 역량을 정의 하기 위하여 문제 포트폴리오와 기술 포트폴리오를 구축한다. 본 방법에서는 특허 청구항에서 추출한 문제 개체를 통하여 문제요소를 정의한다. 특허 의 청구항은 해당 발명이 보호받고자 하는 기술적 내용과 범위에 대한 항목 이며, 해당 발명이 풀고자 하는 문제와 발명의 효과 및 적용 분야에 대한 포 괄적인 내용을 담고 있다.³⁹⁾⁴⁰⁾ 문제요소는 분석대상이 되는 특허 집단에 따 라 달라질 수 있는데, 제약 특허에서는 발명을 통하여 치료하고자 하는 질병 을, 기계 및 전기 특허에서는 발명이 적용될 수 있는 제품을, 식품 화학 특허 에서는 발명을 통해 개발 가능한 음식을 특허의 문제요소로 볼 수 있다. 따 라서 청구항에서 추출한 문제 개체는 해당 발명이 해결하고자 하는 문제점,

39) Marco, A. C. et al., "Patent claims and patent scope", *Research Policy*, Vol.48 No.9(2019).

40) Seo, W. et al., "Product opportunity identification based on internal capabilities using text mining and association rule mining", *Technological Forecasting and Social Change*, Vol.105(2016), pp.94-104.

또는 발명의 적용 가능한 비즈니스 및 제품 등과 관련된 의미를 가지기 때문에, 본 방법에서는 이를 통하여 특허의 문제요소를 정의한다. 특허가 다루는 기술요소를 정의하기 위하여, 본 방법에서는 특허가 가지고 있는 기술 분류코드를 사용한다. 기술을 기반으로 특허를 분류하는 목적으로 개발된 기술 분류코드는 특허의 기술적 내용 및 특성을 정의하기 위하여 다양한 연구에서 활용되었으며, 다른 특허 데이터보다 특허의 기술적 내용을 해석하는 데 편리하다는 이점을 갖는다.⁴¹⁾ 또한 기술 분류코드는 발명에 사용된 기술들의 특성을 나타내는데, 복수의 기술 분류코드가 조합되는 경우 개개의 기술 분류코드가 나타내는 기술 영역의 교집합에 위치한 구체적인 기술을 나타낼 수 있다.⁴²⁾ 따라서 본 방법에서는 특허가 보유한 기술 분류코드를 통하여 해당 특허가 문제요소를 해결하기 위해 개발한 발명의 기술적 내용인 기술요소를 정의한다. 문제-해결방법 관점에서, 특허는 기술적으로 어떠한 문제를 해결하고자 하며, 기술 분류코드는 특허의 기술적 특징이나 내용을 분류하고 있다. 따라서 기술 분류코드는 특허가 기술적으로 해결하고자 하는 문제에 대한 기술요소로 해석할 수 있다.

문제 포트폴리오는 출원인이 기술적으로 해결하고자 하는 문제요소의 집합을 의미한다. 문제요소는 출원인이 보유한 특허들이 해결하고자 하는 문제들을 의미하므로, 문제 포트폴리오는 해당 출원인이 기술적으로 해결하고자 하는 비즈니스 영역을 의미한다. 앞선 단계에서 수집한 특허 데이터로부터 특허가 다루는 핵심 문제요소를 추출하기 위하여 키워드 추출 알고리즘, 또는 개체명 인식 알고리즘 등을 활용할 수 있다. 추출된 문제요소들의 집합에는 발생빈도가 너무 낮거나 의미가 모호한 요소들이 포함되어 있다. 이와 같은 노이즈들을 제거한 뒤, <식 4>과 같이 출원인이 보유한 특허 중

41) Foglia, P., "Patentability search strategies and the reformed IPC: A patent office perspective", *World Patent Information*, Vol.29 No.1(2007), pp.33-53.

42) Yoon, J. et al. "Tracing the evolving trends in electronic skin (e-skin) technology using growth curve and technology position-based patent bibliometrics", *IEEE Access*, Vol.6(2018), pp.26530-26542.

특정 문제 개체가 등장한 특허의 개수를 통하여 출원인별 문제 포트폴리오 PP_i 를 구축할 수 있다. <식 4>에서, n 은 출원인 i 가 보유한 전체 특허의 개수를 의미하며, $PC_{i,j,k}$ 는 출원인 i 의 특허 k 에서 문제요소 j 의 등장 여부를 의미한다.

$$PP_i = [\sum_{k=1}^n PC_{i,1,k}, \sum_{k=1}^n PC_{i,2,k}, \dots, \sum_{k=1}^n PC_{i,j,k}] \quad \langle \text{식 4} \rangle$$

기술 포트폴리오는 출원인이 목표하는 문제요소를 해결하기 위해 개발한 기술의 집합을 의미한다. 기술요소는 출원인이 보유한 특허에 나타난 기술 분류코드이므로, 출원인은 해당 기술요소를 다루고 응용할 수 있는 능력을 보유한다.⁴³⁾ 따라서, 출원인이 보유한 기술요소를 통합한다면 기업이 보유한 기술적 역량을 나타낼 수 있다. 특허의 기술적 내용을 정의하기 위하여 International Patent Classification(IPC), Cooperative Patent Classification(CPC) 등 다양한 분류코드를 사용할 수 있다. <식 5>와 같이 출원인이 보유한 특허 중 특정 기술 분류코드가 등장한 특허의 개수를 통하여 출원인별 기술 포트폴리오 TP_i 를 구축할 수 있다. 문제 포트폴리오를 구축하는 방법과 유사하게, <식 5>에서 n 은 출원인 i 가 보유한 전체 특허의 개수를, $TC_{i,j,k}$ 는 출원인 i 의 특허 k 에서 기술 분류코드 j 의 등장 여부를 의미한다.

$$TP_i = [\sum_{k=1}^n TC_{i,1,k}, \sum_{k=1}^n TC_{i,2,k}, \dots, \sum_{k=1}^n TC_{i,j,k}] \quad \langle \text{식 5} \rangle$$

43) Park, Y. & Yoon, J., "Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering", *Technological Forecasting and Social Change*, Vol.118(2017), pp.170-183.

3. 포트폴리오 기반 요소 추천

본 단계에서는 앞서 구축한 문제 포트폴리오와 기술 포트폴리오에 협업 필터링 방법을 적용하여 대상 기업의 내부 역량을 기반으로 기업이 접하지 않은 문제요소와 기술요소를 추천한다. 협업 필터링을 통해 요소를 추천하기 위하여 타 출원인들과의 포트폴리오 유사도를 도출하고, 대상 기업이 시도하지 않은 요소에 대한 점수를 산출하는 두 단계를 거쳐야 한다. 첫 번째 단계에서, 포트폴리오 유사도를 도출하기 위해 본 연구에서는 코사인 유사도를 기반으로 대상 기업과 타 출원인의 포트폴리오 유사도 $PS(TF, i)$ 를 산출한다(식 6). 두 번째 단계에서는 타 출원인의 포트폴리오와 타 출원인과의 유사도를 기반으로 대상 기업이 시도하지 않은 요소에 대한 점수 $Score_{TF, j}$ 를 산출할 수 있다(식 7). <식 6, 7>에서, $P_{i, j}$ 는 출원인 i 의 포트폴리오상에 요소 j 의 값을 의미한다.

$$PS(TF, i) = \frac{\sum_{j=1} P_{TF, j} \times P_{i, j}}{\sqrt{\sum_{j=1} P_{TF, j}^2} \times \sqrt{\sum_{j=1} P_{i, j}^2}} \quad \langle \text{식 6} \rangle$$

$$Score_{TF, j} = \frac{\sum_{i=1} PS(TF, i) \cdot P_{i, j}}{\sum_{i=1} PS(TF, i)} \quad \langle \text{식 7} \rangle$$

위와 같은 방법을 통해 도출한 점수는 대상 기업이 해당 요소로 진출 가능한 정도를 의미한다.⁴⁴⁾ 예를 들어, 대상 기업이 시도하지 않았던 문제요소 중, 가장 점수가 높은 문제요소는 대상 기업과 유사한 기업이 다수 진출한 비즈니스 영역이므로, 대상 기업이 확장할 수 있는 시장을 의미한다. 마찬가지로

44) Park, Y. & Yoon, J., "Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering", *Technological Forecasting and Social Change*, Vol.118(2017), pp.170-183.

지로, 가장 점수가 높은 기술요소는 대상 기업과 유사한 기업이 다수 보유하고 있는 기술이므로, 대상 기업이 개발할 가치가 있는 기술을 의미한다.

4. 요소 연관관계 파악

본 단계에서는 발굴된 기회를 이용하여 구체적인 R&D 및 비즈니스 전략을 도출하기 위하여 연관관계 분석을 실시한다. 앞선 단계의 결과물로 얻은 문제요소와 기술요소는 단순히 대상 기업과 타 출원인의 유사도를 기반으로 추천되었다. 따라서 대상 기업이 보유한 기술적 능력으로 추천된 시장에 진출할 수 있는지, 또는 추천된 기술이 대상 기업의 목표 시장에 대한 문제 해결 효과가 있는지를 확인할 수 없다. 문제-해결방법 관점을 적용하여 기존에 등록된 특허들을 기반으로 연관관계 분석을 수행한다면, 비즈니스 시장과 기술에 어떠한 관계가 존재하는지를 파악함으로써 위와 같은 한계를 극복할 수 있다. 본 연구에서는 대상 기업의 보유 요소와 추천 요소들이 전체 특허 데이터에서 동시 출현한 빈도를 기반으로 규칙을 생성한다. 따라서 본 연구에서 아이템은 보유 요소와 추천 요소를 통합한 요소 집합이며, 거래 정보는 요소 집합의 원소를 포함하고 있는 모든 특허 데이터이다.

발굴된 규칙 중에서, 비즈니스 전략 수립에 활용 가능한 유의미한 규칙들을 식별하기 위해 다음과 같은 후처리(Post-Processing) 과정을 거친다. 1) 우선, 조건부 요소에 문제요소와 기술요소가 섞인 규칙을 제거하고, 2) 조건부 요소와 결과 요소가 모두 문제요소이거나, 모두 기술요소인 규칙들을 제거한 뒤, 3) 대상 기업이 보유한 요소끼리 생성된 규칙을 제거한다. 결과적으로 도출된 유의미한 규칙들은 조건부 요소와 결과 요소의 조합에 따라 크게 세 가지 유형으로 구분할 수 있으며, 각 유형에 따라 상이한 R&D 및 비즈니스 전략 도출이 가능하다(표 1). 첫 번째 유형은 추천 문제와 추천 기술의 조합이다. 이 유형으로 구분된 규칙들은 기업이 보유한 기술적 역량과 기존에 집중하고 있던 시장을 기반으로 새로이 추천된 기술 및 시장의 관계를 의미한다. 따라서 기업은 확장 가능한 시장과 개발할 가치가 있는 기술을 통하여

기업이 전혀 접해 보지 않은 새로운 시장으로 비즈니스를 확장할 수 있다. 두 번째 유형은 추천 문제와 보유 기술의 조합이다. 위 조합의 규칙을 통하여 기업은 보유한 기술적 역량을 기반으로 진출할 수 있는 새로운 시장에 대한 착안점을 얻을 수 있다. 마지막 유형은 보유 문제와 추천 기술의 조합이다. 마지막 유형의 조합을 통하여 대상 기업은 새로운 기술을 개발함으로써 기존 시장에서의 경쟁우위를 강화할 수 있다.

[표 1] 추천된 문제-기술요소 조합에 따른 R&D 및 비즈니스 전략

		기술요소의 타입	
		추천	보유
문제요소의 타입	추천	새로운 기술을 개발하여 비즈니스 영역 확장	보유 기술을 기반으로 진출 가능한 새로운 시장 탐색
	보유	새로운 기술개발을 통한 기존 시장의 경쟁력 강화	

IV. 사례연구

본 장에서는 제약 산업의 특허를 대상으로 앞서 제시한 연구절차를 적용하는 사례연구를 진행한다. 제약 특허의 경우, 분자식이나 병명과 같은 개체의 분석이 난해하기 때문에, 특허를 단순히 텍스트로만 분석하는 기존 연구방법의 활용이 어렵다는 특징이 존재한다. 하지만 제약 특허들은 공통적으로 질병을 치료하는 목적으로 개발되므로, 제약 특허의 텍스트상에 나타난 질병 개체를 활용하여 특허가 해결하고자 하는 문제요소의 정의가 용이하다는 장점을 가진다. 따라서 제약 특허는 문제-해결방법 관점에서 비즈니스 및 기술 기회를 발굴하는 본 연구방법의 사례연구 대상으로 적절하다.

1. 데이터 수집 및 전처리

본 연구에서 사례연구로 진행하는 제약 특허는 World Intellectual Prop-

erty Organization(WIPO)의 IPC and Technology Concordance Table을 이용하여 수집되었다. 위 자료에서 IPC 코드들은 기술 분야에 연결되어 있으므로, 제약 기술과 연결된 IPC 코드를 통하여 사례연구에서 사용할 특허를 검색할 수 있다. 제약 산업과 관련된 IPC 코드는 A61K와 A61P이며, 따라서 본 연구에서는 USPTO에서 제공하는 PatentView를 이용하여 2009년부터 2018년까지 총 10년간 등록된 미국 특허 중 A61K와 A61P를 보유한 특허 총 142,075건을 수집하였다. 수집된 특허는 특허 등록번호, 특허 등록일, 출원인, 청구항, CPC 분류코드 등의 데이터를 포함하고 있다.

문제요소로 정의된 질병 개체를 추출하기 위하여, 본 사례연구에서는 생화학 자연어 처리 툴인 scispaCy를 활용하였다⁴⁵⁾. scispaCy는 텍스트 데이터에서 DNA, RNA, 단백질, 분자식, 질병 등 생화학 기술 관련 명사구를 추출할 수 있는 패키지이다. 본 연구에서는 59,184개의 질병 개체를 추출하였으며, 두 가지 조건을 기준으로 불용어를 제거하였다. 첫 번째 조건은 문서 빈도(Document Frequency)가 1인 개체로, 해당 질병 개체를 포함하는 문서의 개수가 1개인 경우이다. 두 번째 조건은 질병과 직접적인 관련이 없는 개체로, 개체의 길이가 너무 짧거나, 질병을 의미하지 않는 너무 일반적인 개체인 경우이다. 위 두 조건을 기반으로 불용어를 제거한 결과, 질병 개체 18,492개와 해당 개체를 보유한 제약 특허 38,261건, 그리고 해당 특허들을 보유한 9,485개의 출원인이 본 연구의 사례연구 대상으로 선별되었다. 기술 요소의 정의를 위해 사용한 CPC 코드의 경우, 이전의 분류코드보다 기술을 정교하고 세밀하게 분류한다는 특징을 가지고 있다.⁴⁶⁾ 또한 서브그룹 수준의 CPC 코드는 CPC 코드 내 최소 검색 단위로, 기술 전체의 범위를 포괄적으로 설명한다. 따라서 서브그룹 수준의 CPC 코드는 출원인이 보유하고 있는 기술적 역량을 설명하기 위하여 사용하기에 적합하다. 전체 대상 특허가

45) Neumann, M. et al., "ScispaCy: Fast and Robust Models for Biomedical Natural Language Processing", arXiv preprint arXiv:1902.07669, 2019.

46) Mueller, S. C. et al., "Monitoring innovation in electrochemical energy storage technologies: A patent-based approach", *Applied Energy*, Vol.137(2015), pp.537-544.

보유한 서브그룹 수준의 CPC 코드는 총 15,402개로, 별도의 처리 과정 없이 사용할 수 있다.

2. 분석대상 기업 선정 및 포트폴리오 구축

본 연구에서 제안하는 방법은 과거의 데이터를 기반으로 기업에 적합한 비즈니스 및 기술 기회를 식별하는 방법이다. 또한 본 방법을 통하여 기업의 내부 역량을 기반으로 실패확률이 적은 기회를 식별할 수 있다. 따라서 본 연구에서 제안하는 방법은 인적, 물적 자원이 한정적인 중소기업에 적합하다. Pharnext는 2007년 설립되어 신경퇴행성 질환에 대한 새로운 치료제를 개발하는 바이오 제약 회사로, 2013년 첫 특허 등록 이후 현재 총 32개의 등록특허를 보유하고 있다. 이 중 질병 개체를 포함하고 있는 제약 특허는 총 24개인데, 최근 5년간 23개의 특허를, 최근 3년간 17개의 특허를 등록하였다. 이는 Pharnext가 관련 기술을 개발하기 위한 충분한 역량과 잠재성, 성장 가능성을 보유하고 있다는 점을 보여 준다. 따라서 Pharnext는 적절한 수의 특허를 보유한 중소기업임과 동시에 해당 분야의 R&D 능력을 보유하고 있으므로, 본 연구에서 진행하는 사례연구의 대상 기업으로 적절하다.

앞서 추출한 질병 개체와 서브그룹 수준의 CPC 코드를 기반으로, 앞서 제시한 〈식 4〉와 〈식 5〉를 통하여 전체 대상 출원인의 문제 포트폴리오와 기술 포트폴리오를 구축할 수 있다. 대상 기업 Pharnext는 67개의 문제요소와 59개의 기술요소를 보유하고 있다. 상위 발생빈도를 보인 문제요소로는 신경변성 질환으로 분류되어 아직은 치료법이 존재하지 않는 *Alzheimer's disease*, 말초신경을 형성하는 유전자 중복으로 인해 손, 발의 말초신경 발달이 저하되는 *Charcot-Marie-Tooth disease*, 수의근을 제어하는 신경세포의 소멸로 인한 병으로, 흔히 루게릭 병이라 불리는 *Amyotrophic lateral sclerosis*, 불완전한 도파민의 생성 및 작용으로 인하여 운동신경의 자극이 감소되는 *Parkinson's disease* 등 직접적인 병명 요소와, *Toxicity*, *Bradykinesia*, *Neuropathy* 등 질병의 원인 또는 증상을 나타내는 요소 등이 존재한다. 위

요소들을 통하여 대상 기업이 신경 관련 질병에 대한 치료법을 개발하고 있다는 것을 파악할 수 있다. 상위 발생빈도를 보인 기술요소 중, 대상 기업 Pharnext는 ‘유기 활성 성분을 함유하는 의약품 제제’를 뜻하는 메인그룹 A61K 31/00의 하위에 존재하는 기술요소를 대다수 보유하고 있다. ‘산; 산무수물, 산할로젠화물 또는 그들의 염’을 의미하는 서브그룹 A61K 31/185는 대상 기업이 보유한 기술요소 중 발생빈도가 두 번째로 높는데, 위 서브그룹의 하위 계층에 속한 A61K 31/197, A61K 31/195, A61K 31/198, A61K 31/192 요소 역시 높은 발생빈도를 보이고 있다. 또한, 상위 발생빈도를 보인 A61K 31/44, A61K 31/485, A61K 31/445, A61K 31/42, A61K 31/4164, A61K 31/436 등의 요소들은 서브그룹 A61K 31/33의 하위 서브그룹 A61K 31/395에 포함된 요소로, ‘이종 환식화합물’ 중, ‘이종 환원자로서 질소가 있는 것’에 해당하는 기술요소이다. 이처럼 기술요소와 요소들의 계층구조를 통하여 대상 기업이 보유하고 있는 기술적 역량과 그 분야를 파악할 수 있다.

3. 분석결과

대상 기업의 보유 포트폴리오를 기반으로 확장 가능한 질병 분야와 개발 가치가 있는 기술 분야를 추천하기 위해 협업 필터링을 활용하여 문제요소와 기술요소를 추천하였다. 본 연구에서는 문서빈도가 높은 상위 20개의 문제요소와 기술요소를 제거한 포트폴리오를 협업 필터링의 입력 데이터로 사용하였다. 본 연구에서 사용하는 협업 필터링은 대상 기업과 유사한 출원인이 많이 보유하고 있는 요소를 추천하기 때문에, 일반적으로 문서빈도가 높은 요소가 다수 추천된다. 따라서, 문서빈도가 높은 요소들을 제거한 뒤 협업 필터링을 수행하면 대상 기업의 내부 역량에 적합하며, 너무 흔하지 않은 추천 결과를 얻을 수 있다.

협업 필터링을 위하여, 우선 대상 기업과 타 출원인들의 보유 포트폴리오 유사도를 산출하였다. <식 6>을 이용하여 산출하였으며, 대상 기업과 포트폴리오 유사도가 높은 상위 출원인은 <표 2>와 같다. 대상 기업과 문제 포트

폴리오 유사도가 높은 출원인은 대상 기업과 *Toxicity, Cognitive impairment, Memory impairment, Neuronal cell death, Dementia* 등의 문제요소를 공유하고 있으며, 기술 포트폴리오 유사도가 높은 출원인은 *A61K 31/197, A61K 31/137, A61K 31/4164, A61K 31/165* 등의 기술요소를 대상 기업과 공유한다. 문제 포트폴리오가 유사한 출원인은 대상 기업과 비슷한 시장을 목표로 상이한 기술요소를 보유하고 있는 경쟁 출원인으로 볼 수 있고, 기술 포트폴리오가 유사한 출원인은 비슷한 기술을 기반으로 상이한 시장을 목표로 하는 출원인들로 볼 수 있다.

[표 2] 포트폴리오별 상위 유사도 출원인

출원인	문제 유사도	출원인	기술 유사도
Honeywell International Inc.	0.2449	Saol International Limited	0.3515
Dart NeuroScience LLC	0.2108	Omeros Corporation	0.3497
Neuraltus Pharmaceuticals, Inc.	0.2025	Orexigen Therapeutics, Inc.	0.3234
Neuron Biopharma, S.A.	0.1952	Signum Biosciences, Inc.	0.3197
Tetra Discovery Partners, LLC	0.1925	BioVentures, LLC	0.2695

타 출원인과의 포트폴리오 유사도를 기반으로 대상 기업이 확장 가능한 문제요소와 기술요소를 추천할 수 있다. <식 7>을 통한 추천 결과 중, 상위 점수를 받은 요소들은 <표 3>, <표 4>와 같다. 점수가 높은 문제요소는 대상 기업과 비슷한 문제 포트폴리오를 가진 출원인들이 목표하고 있는 질병을 의미하며, 높은 점수의 기술요소는 비슷한 기술 포트폴리오를 가진 출원인들이 다수 보유하고 있는 기술을 의미한다. 높은 점수를 받은 문제요소 중 *Dry-eye disease*와 *Ocular hypertension*은 눈 관련 질병으로, 대상 기업이 시도하지 않은 새로운 분야이다. 이 외에도 *Platelet-growth, VLDL-C, HCV infection* 등의 요소가 상위에 등장하였으므로, 심장 및 심혈관 질환, C형 간염 등은 안구 관련 질병들과 더불어 대상 기업이 접하지 않았지만, 진출 가능한 질병으로 볼 수 있다. 가장 높은 점수를 받은 기술요소 *A61Q 5/10*은 모발 영구 염색제 관련 기술 분류코드이다. 서브클래스 *A61Q*는 대상 기업

이 보유하지 않은 화장품 또는 유사 화장품 제제의 특정 용도와 관련된 기술이다. 또한 A61K 38/00, C07K 16/00, C07D 405/00 등의 기술요소들 역시 상위에 등장하였다. 따라서 모발 염색 화장품, 펩타이드 함유 의약품, 면역 글로불린 항체, 산소 또는 질소만을 갖는 이중원자 고리 화합물과 관련된 기술은 대상 기업이 개발할 가치가 있는 기술 분야로 볼 수 있다.

[표 3] 추천된 문제요소 (상위 10개)

문제요소	점수	설명
Platetet-growth	3.5005	혈소판 성장 인자
VLDL-C	3.3325	초저밀도지단백 콜레스테롤
Dry-eye disease	3.3289	안구건조증
Ocular hypertension	3.2903	고안압증
Anastomotic	3.1475	혈관, 장기 등의 문합
Renal osteodystrophy	3.1095	신장성 골 형성 장애
HCV infection	3.0307	C형 간염 감염
Bile duct obstruction	2.9514	담관폐색증
PEP-inhibitors	2.9366	PEP 억제제
HIV nucleotide	2.9042	인간면역결핍바이러스 뉴클레오타이드

[표 4] 추천된 Solution Technology 요소 (상위 10개)

기술요소	점수	설명
A61Q 5/10	4.7689	모발 영구 염색제
A61K 38/066	4.3155	펩티드를 함유하는 의약품 제제 중 TRH, 티로리베린, 갑상선 자극 호르몬 방출 호르몬
C07K 16/2833	4.0323	면역 글로불린 항체로, MHC-분자에 관한 것
C07K 16/3007	3.9879	면역 글로불린 항체로, 카르시노 배아 항원
C07D 405/14	3.8020	이중원자로서 질소 원자만을 갖는 3개 이상의 이중원자 고리를 함유하는 것
A61K 8/27	3.7452	아연과 그 화합물의 화장품
C07K 14/605	3.6716	21 이상의 아미노산을 갖는 펩티드로 글루카곤
C07D 417/14	3.6544	적어도 1개가 이중원자로서 질소와 황만을 갖는 3개 이상의 이중원자 고리를 함유하는 것
A61K 38/26	3.5872	펩티드를 함유하는 의약품 제제 중 글루카곤
C07K 2317/92	3.5853	면역 글로불린의 친화성(KD), 연관율(KA), 분리율(Kd) 또는 EC50 값의 특성

구체적인 비즈니스 전략 수립을 위하여, 앞서 추천된 요소들과 기업이 보유한 포트폴리오를 결합하여 전체 특허를 대상으로 연관관계 분석을 수행하였다. 대상 기업이 보유한 요소 중, 문서빈도가 높은 요소들을 제외한 62개의 문제요소와 55개의 기술요소에 추천된 문제요소와 기술요소 각 50개를 통합하여 연관관계 분석에 사용되는 아이템을 정의하였고, 해당 요소들을 포함하고 있는 2,552개의 특허를 연관관계 분석에서의 거래 정보로 정의하였다. 본 연구에서는 Support의 threshold를 0.01로, Confidence의 threshold를 0.2로 설정하였고, III.4장에서 정의한 후처리 과정을 거쳐 총 20개의 규칙을 도출하였다(표 5). Lift가 가장 높은 규칙인 *Bowel syndrome*과 *A61K 38/26* 요소 조합의 규칙들은 동일 출원인에 의한 연속출원과 분할출원으로 인한 결과이다. 해당 규칙들은 대상 기업에게 유의미한 연관규칙으로 볼 수 없으므로, 결과 해석 대상에서 제외한다.

결과적으로 도출된 18개 규칙은 요소 조합 유형에 따라 상이한 결과 해석이 가능하다. <표 5>에서, 요소 타입은 문제요소와 기술요소가 추천된 요소 인지(R), 대상 기업이 보유한 요소인지(E)를 나타낸다. 전체 18개 규칙 중, 5개의 규칙이 추천 문제와 추천 기술 조합의 유형이며, 2개의 규칙이 추천 문제와 보유 기술 조합 유형, 나머지 11개 규칙이 보유 문제와 추천 기술의 조합 유형이다. 각 유형에 따라 새로운 기술을 기반으로 한 신규 질병 시장, 보유한 기술을 기반으로 한 새로운 질병 시장, 기존 질병 시장에 적용 가능한 새로운 기술로 해석할 수 있다.

추천 문제와 추천 기술 조합의 유형은 확장 가능한 질병 분야와 개발 가치가 있는 기술을 통하여 진출할 수 있는 새로운 비즈니스를 의미한다. 도출된 규칙 중, *Glaucoma*와 *A61K 31/5575*, *Hepatitis C*와 *A61K 39/12*가 위 조합에 해당한다. 녹내장은 시신경 이상으로 인하여 시야 결손이 발생하는 질환이다. 녹내장 발병의 가장 주된 위험 요인은 안압 상승인데, 앞선 협업 필터링 결과에서 고안압증은 4번째로 높은 점수를 받은 확장 가능성이 큰 질병 분야로 도출되었다. 실제로, *Glaucoma*와 *A61K 31/5575*가 등장한 특허들을 살펴보면, 프로스타글란딘, 라타노프로스트 등을 포함하는 치료제를 통하여

녹내장, 안압 상승과 같은 안과 질환의 치료를 위한 발명이 대다수를 이룬다. 따라서 대상 기업은 A61K 31/5575을 포함한 녹내장 및 고안압 관련 기술을 개발하여 안구 관련 질병 치료 시장으로의 신규 진출이 가능할 것이다. C형 간염은 앞선 협업 필터링을 통하여 확장 가능한 질병 분야로 추천되었으며, Hepatitis C와 A61K 39/12의 조합은 C형 간염 바이러스 입자, 백신, 게놈 복제, 돌연변이 등과 관련된 발명에서 다양하게 등장하였다. 따라서 추천 요소 A61K 39/12를 기반으로 C형 간염 관련 질병 치료 분야로 비즈니스 확장이 가능할 것으로 예상된다.

[표 5] 생성된 연관관계 규칙

요소 타입 (문제, 기술)	조건부 요소	결과 요소	Support	Confidence	Lift
(R, R)	Bowel syndrome	A61K 38/26	0,0118	0,9091	28,6420
(R, R)	A61K 38/26	Bowel syndrome	0,0118	0,3704	28,6420
(R, R)	A61K 31/5575	Glaucoma	0,0192	0,8750	11,4513
(R, R)	Glaucoma	A61K 31/5575	0,0192	0,2513	11,4513
(R, R)	A61K 39/12	Hepatitis C	0,0153	0,6190	10,6744
(R, R)	Hepatitis C	A61K 39/12	0,0153	0,2635	10,6744
(R, E)	A61K 31/4985	Schizophrenia	0,0121	0,3196	3,1612
(R, R)	C07D 519/00	Schizophrenia	0,0106	0,2647	2,6183
(R, E)	A61K 31/55	Schizophrenia	0,0102	0,2167	2,1432
(E, R)	C07D 405/14, C07D 413/14	Atherosclerosis	0,0110	0,2800	2,0016
(E, R)	C07D 417/14, C07D 413/14	Atherosclerosis	0,0121	0,2696	1,9270
(E, R)	C07D 413/14	Atherosclerosis	0,0235	0,2620	1,8730
(E, R)	C07D 413/12	Atherosclerosis	0,0157	0,2247	1,6064
(E, R)	C07D 401/04	Atherosclerosis	0,0168	0,2205	1,5763
(E, R)	C07D 417/14	Atherosclerosis	0,0165	0,2199	1,5719
(E, R)	C07D 403/14	Atherosclerosis	0,0125	0,2192	1,5668
(E, R)	C07D 401/04	Stroke	0,0157	0,2051	1,5442
(E, R)	C07D 405/14	Atherosclerosis	0,0165	0,2059	1,4717
(E, R)	C07D 403/12	Atherosclerosis	0,0165	0,2049	1,4646
(E, R)	C07D 405/12	Atherosclerosis	0,0145	0,2044	1,4613

추천 문제와 보유 기술 조합은 보유한 기술을 기반으로 확장 가능한 사업 영역에 대한 착안점을 줄 수 있는 유형이다. 본 연구에서 시행한 사례 연구의 결과로, 기업이 접하지 않은 질병 *Schizophrenia*가 기업의 보유 기술 *A61K 31/4985*, *A61K 31/55*와 규칙을 이루었다. 결과 요소로 등장한 *Schizophrenia*는 정신질환의 일종으로, 일부 현실을 제대로 인식하지 못하여 망상, 환각, 환청, 환영 등의 증상과 감정표현 및 발언이 감소하거나, 타인의 시각에 무관심해지고 대인관계에 필요 이상의 긴장감을 느끼는 행동을 보이는 조현병을 의미한다. 조현병은 일반적으로 도파민과 글루탐산 신경전달물질의 이상으로 인해 발병하는 정신증이다. 조건부 요소로 등장한 보유 기술 *A61K 31/4985*와 *A61K 31/55*를 추천 기술 *C07D 519/00*은 PDE 효소 억제제를 통하여 도파민 수용체를 치료하거나, 인지기능 증진 및 기억형성 향상에 유용한 선택적 HDAC3 억제제를 개발하여 신경 퇴행성 장애 및 신경계 장애를 치료하기 위한 특허에 사용된 바 있다. 대상 기업은 보유 기술 *A61K 31/4985*와 *A61K 31/55*를 활용하여 인지장애, 기억장애, 치매, 운동불능 등의 신경성 장애를 치료하기 위한 발명을 한 경험을 다수 보유하고 있다. 따라서 대상 기업은 보유한 기술적 역량을 기반으로 하여 조현병 치료 시장으로의 진출이 가능할 것으로 예상되며, 이를 위하여 새로운 기술 *C07D 519/00*에 대한 개발 역시 고려해볼 수 있을 것이다.

보유 문제와 추천 기술 조합의 유형은 기존 사업 분야의 경쟁우위를 확보하기 위한 새로운 기술의 도입으로 해석할 수 있다. 도출된 규칙 중, *Atherosclerosis*와 *Stroke*를 결과 요소로 가지는 규칙들이 위 유형에 해당한다. 동맥경화증은 동맥의 탄력 하락으로 인하여 혈전이 생기는 등의 이유로 동맥이 좁아지는 질병이며, 뇌졸중은 뇌 혈류 이상으로 인하여 국소적인 신경 결손 증상을 이르는 질병이다. 뇌졸중은 허혈성과 출혈성으로 나뉘는데, 허혈성 뇌졸중의 경우 뇌에서 발생한 동맥경화로 인해 뇌로 향하는 혈류량이 감소하여 발생한다. *Atherosclerosis*, *Stroke*에 대한 조건부 요소들은 뇌 허혈과 관련된 심혈관계 질환뿐만 아니라 암, 관절염, 천식, 비만 등 다양한

질병을 치료하기 위한 기술로 널리 활용되는 기초적인 기술이다. 위 요소들은 모두 이중원자 고리의 유기 화합물 관련 기술 분류코드인 서브클래스 C07D에 포함되어 있다. 이들은 이중원자 고리에서 고리를 구성하는 질소, 산소, 황 등의 원소 함유 여부에 따라 구분되며, 사슬결합에 의해 결합되거나, 3개 이상의 이중원자 고리를 함유하는 등의 공통적인 특징을 가진다. 대상 기업은 생체의 비정상 또는 병적상태 예방, 생리적 신체기능 조절, 생리적 상태의 진단을 위한 의약품, 치과용, 화장용 제제와 관련된 기술인 서브클래스 A61K를 중심으로 기술 포트폴리오가 구성되어 있으며, C07D 관련 조건부 요소들과 유사한 기술인 이중원자 고리 관련 기술 역시 소수 보유하고 있다. 따라서 대상 기업의 보유 역량을 기반으로 한 R&D를 통하여 서브클래스 C07D에 포함된 조건부 요소 관련 기술적 역량을 강화한다면, 기존에 진출하였던 뇌허혈 치료 분야에서 더욱 경쟁력을 갖출 수 있을 것이다. 또한 심혈관계 질환뿐만 아니라, 대상 기업이 현재 집중하고 있는 알츠하이머 병이나 파킨슨 병, 비만 등에 대한 더 나은 치료법을 찾는 데 큰 도움이 될 것으로 예상된다.

V. 결론 및 추후 연구

본 연구는 기술의 문제-해결방법 관점에서 특허를 분석하여, 기업에 확장 가능한 비즈니스 영역과 개발 가치가 있는 기술을 추천하고, 연관관계를 기반으로 구체적인 비즈니스 전략 수립을 지원하는 방법을 제시하였다. 특허는 어떠한 문제를 해결하기 위하여 개발된 기술이므로, 특허를 분석함에 있어 특허가 해결하고자 하는 문제와 해결방법을 나누어 분석할 필요가 있다. 위 두 가지 사항을 분리하여 분석할 경우, 문제와 해결방법에 따라 기술의 목적이 달라지며, 기술의 목적에 따라 도출되는 R&D 및 비즈니스 전략이 상이하다. 본 연구에서 제시하는 방법은 기업의 보유 기술과 목표 시장에 적합

하며 실현 가능성이 큰 문제요소와 기술요소를 추천하기 위하여 협업 필터링과 연관관계 분석을 사용하였다. 우선, 특허 데이터를 수집하여 문제 포트폴리오와 기술 포트폴리오를 구축하였다. 그 다음, 협업 필터링을 통하여 기업의 보유 기술 및 비즈니스를 기반으로, 확장 가능성이 존재하는 문제요소와 개발 가치가 있는 기술요소를 추천하였다. 마지막으로 기업이 보유한 포트폴리오와 추천된 요소들을 통합하여 연관관계 분석을 실시하였고, 생성된 규칙을 구분하여 대상 기업에 대한 R&D 및 비즈니스 전략을 도출하였다. 본 연구에서는 2009년부터 2018년 사이에 등록된 제약 특허를 대상으로 사례 연구를 진행하였다. 대상 기업으로 선정된 Pharnext가 보유하고 있는 신경성 질환 관련 문제 포트폴리오와 기술 포트폴리오를 기반으로 세 가지 전략을 제시하였다. Pharnext는 추천된 기술 A61K 31/5575를 개발하여 녹내장 치료 시장으로, A61K 39/12를 개발하여 C형 간염 치료 시장으로 비즈니스 영역 확장이 가능하며, 기존에 보유한 기술인 A61K 31/4985와 A61K 31/55를 기반으로 조현병 치료 시장으로 신규 진출이 가능하다. 또한 C07D 405/14, C07D 417/14, C07D 413/14, C07D 401/04 등의 기술을 개발하여 기존 비즈니스 시장인 허혈성 뇌졸중 치료 시장에서의 경쟁우위를 강화할 수 있을 것이다.

본 연구는 다음과 같은 기여점을 가진다. 첫째, 본 연구는 특허에 문제-해결방법 관점을 적용하여 보유 역량을 기반으로 비즈니스 및 기술 기회를 식별하였다. 기존의 연구들은 특허를 단순히 기술의 관점에서만 분석하였으므로, 보유 기술 및 비즈니스에 대한 고려가 부족하였다. 본 연구에서는 기업이 보유하고 있는 기술과 목표하는 시장을 기반으로 실현 가능성이 큰 기회를 식별하였다. 따라서 본 연구에서 제안하는 방법을 통하여 기업은 보유 역량을 기반으로 확장 가능한 시장을 탐색하거나, 목표 시장에서 경쟁력을 높이는 데 필요한 기술을 식별할 수 있다. 둘째, 본 연구가 제안하는 방법을 통하여 세계 특허에 대하여 명확한 문제요소와 기술요소의 정의가 가능하다. 선행연구들은 문제-해결방법 기반의 특허지능 연구를 진행함

에 있어 특허를 대표하는 문제요소와 기술요소를 선정하는 기준이 불분명하거나, 특정 국가의 특허만을 대상으로 활용할 수 있다는 한계점으로 인하여 문제-해결방법 개념의 사용이 제한적이다. 본 연구에서는 특허의 청구항에서 추출한 문제 개체를 문제요소로, 특허의 기술 분류코드를 기술요소로 정의하였으므로, 문제-해결방법 개념의 활용에 범용성 및 활용성이 높다.

이러한 기여에도 불구하고, 추후 연구를 통하여 다음과 같은 한계점을 개선할 필요가 있다. 첫째, 문제 포트폴리오 정의를 위하여 문제 개체 추출에 관한 연구가 필요하다. 본 연구에서 진행한 제약 산업에 관한 사례연구의 경우, 질병의 치료라는 제약 특허만의 특징을 기반으로 문제요소를 정의할 수 있다. 본 연구가 제안하는 방법에 대한 활용을 위하여, 특정 텍스트 데이터를 사용하거나, 특정 분야의 문제 개체를 추출하는 개체명 인식 알고리즘을 개발하는 등 타 기술 분야에 적합한 문제 개체 추출 방법이 개발될 필요가 있다. 또한 문제 개체의 선별에 있어 개체 선정 기준에 대한 객관성 부여 방법 역시 추후 연구가 필요할 것으로 생각된다. 둘째, 문제 유사도가 높은 출원인에 대한 추가적인 선별작업이 필요하다. 문제 유사도가 높은 출원인은 실제로 비슷한 비즈니스 시장을 목표할 수도 있으며, 출원인의 비즈니스 다각화로 인하여 문제 포트폴리오의 유사도가 높게 도출된 것일 수도 있다. 이러한 문제는 연관관계 분석의 결과에 노이즈로 작용할 수 있다. 따라서 문제요소 및 기술요소의 보완관계나 대체관계, 상하관계 등을 파악하거나, 분석대상이 되는 기술 분야와 출원인의 주요 도메인을 비교하는 등의 방법을 통하여 분석에 활용되는 출원인을 선별하는 작업에 대한 후속연구가 필요하다. 셋째, 연관관계 분석에 사용되는 threshold에 대한 연구가 필요하다. 본 연구에서는 Support의 최솟값을 0.01로, Confidence의 최솟값을 0.2로 설정하였다. 하지만 연관관계 분석에서 높은 측정치를 보이는 규칙은 이미 어느 정도 특허화된 발명을 의미하므로, 도출된 기회의 독창성을 보장하지는 못한다. 따라서 분석에 있어 threshold의

상한값을 선정하는 등의 추가 연구가 필요하다. 마지막으로, 본 연구에서는 기업이 보유한 역량을 기반으로 비즈니스 및 기술 기회를 발굴하였으므로, 발굴된 기회의 경쟁 정도나 시장성 등을 파악하는 후속 연구가 필요하다.

참고문헌

〈단행본(서양)〉

Savaransky, S. D., *Engineering of creativity: Introduction to TRIZ methodology of inventive problem solving*, CRC Press, 2000.

〈학술지(국내 및 동양)〉

노현정 & 임효정, “특허정보를 활용한 나노바이오 융합기술의 특성 분석”, 『지식재산연구』, 제4권 제3호(2009).

박영진 외 2인, “보유특허 기반의 기술기회탐색을 위한 특허추천방법: 3D 프린팅 산업을 중심으로”, 『지식재산연구』, 제10권 제1호(2015).

신영수 외 2인, “R&D 투자, 기술경영능력, 기업성과간의 관계”, 『경영학연구』, 제38권 제1호(2009).

이지호 외 4인, “특허의 Problem-Solution 텍스트 마이닝을 활용한 기술경쟁정보 분석 방법”, 『지식재산연구』, 제13권 제3호(2018).

정병욱 외 2인, “융합기술의 동향분석을 위한 의미론적 특허분석 접근방법: 토픽모델링과 교차영향분석의 활용”, 『지식재산연구』, 제11권 제4호(2016).

최성철 외 2인, “기술로드맵핑을 위한 특허정보의 SAO 기반 텍스트 마이닝 접근 방법”, 『기술혁신연구』, 제20권 제1호(2012).

〈학술지(서양)〉

Abaya, S. A., “Association rule mining based on Apriori algorithm in minimizing candidate generation”, *International Journal of Scientific & Engineering Research*, Vol.3 No.7(2012).

Albercht, M. A. et al., “Quality assurance in the EPO patent information resource”, *World Patent Information*, Vol.32 No.4(2010).

Chiclana, F. et al., “ARM-AMO: an efficient association rule mining algorithm based on animal migration optimization”, *Knowledge-Based Systems*, Vol.154(2018).

Choe, H. et al., “Structural properties and inter-organizational knowledge flows of patent citation network: The case of organic solar cells”, *Renewable and Sustainable Energy Reviews*, Vol.55(2016).

Choi, S. et al., “An SAO-based text-mining approach for technology roadmapping using patent information”, *R&D Management*, Vol.43 No.1(2013).

- Érdi, P. et al., “Prediction of emerging technologies based on analysis of the US patent citation network”, *Scientometrics*, Vol.95 No.1(2013).
- Foglia, P., “Patentability search strategies and the reformed IPC: A patent office perspective”, *World Patent Information*, Vol.29 No.1(2007).
- Good, N. et al., “Combining collaborative filtering with personal agents for better recommendations”, *AAAI/IAAI*, 1999.
- Kaur, M. & Kang, S., “Market Basket Analysis: Identify the changing trends of market data using association rule mining”, *Procedia computer science*, Vol.85(2016).
- Li, W. et al., “A collaborative filtering recommendation method based on discrete quantum-inspired shuffled frog leaping algorithms in social networks”, *Future Generation Computer Systems*, Vol.88(2018).
- Lichenthaler, E., “Technological change and the technology intelligence process: a case study”, *Journal of Engineering and technology Management*, Vol.21 No.4(2004).
- Lichenthaler, E., “Technology intelligence processes in leading European and North American multinationals”, *R&D Management*, Vol.34 No.2(2004).
- Marco, A. C. et al., “Patent claims and patent scope”, *Research Policy*, Vol.48 No.9(2019).
- Moehrle, M. G. et al., “Patent-based inventor profiles as a basis for human resource decisions in research and development”, *R&D Management*, Vol.35 No.5(2005).
- Park, H. et al., “A patent intelligence system for strategic technology planning”, *Expert Systems with Applications*, Vol.40 No.7(2013).
- Park, Y. & Yoon, J., “Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering”, *Technological Forecasting and Social Change*, Vol.118(2017).
- Peng, M. et al., “Exploration of association rule mining for coding consistency and completeness assessment in inpatient administrative health data”, *Journal of biomedical informatics*, Vol.79(2018).
- Sarwar, B. M. et al., “Item-based collaborative filtering recommendation algorithms”, *Proceedings of the 10th international conference on World Wide Web*, 2001.
- Seo, W. et al., “Product opportunity identification based on internal capabilities using text mining and association rule mining”, *Technological Forecasting and*

Social Change, Vol.105(2016).

Shih, M. -J. et al., "Discovering competitive intelligence by mining changes in patent trends", *Expert Systems with Applications*, Vol.37 No.4(2010).

Wang, X. et al., "Identifying R&D partners for dye-sensitized solar cells: a multi-level patent portfolio-based approach", *Technology Analysis & Strategic Management*, Vol.31 No.3(2019).

Yoon, J. & K. Kim, "Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks", *Scientometrics*, Vol.88 No.1(2011).

Yoon, J. et al., "Identifying product opportunities using collaborative filtering-based patent analysis", *Computers & Industrial Engineering*, Vol.107 (2017).

Yoon, J. et al., "Technology opportunity discovery(TOD) from existing technologies and products: A function-based TOD framework", *Technological Forecasting and Social Change*, Vol.100(2015).

Yoon, J. et al., "Tracing the evolving trends in electronic skin (e-skin) technology using growth curve and technology position-based patent bibliometrics", *IEEE Access*, Vol6(2018).

〈기타자료〉

Herlocker, J. L. et al., "Explaining collaborative filtering recommendations", ACM, 2000.

Neumann, M. et al., "ScispaCy: Fast and Robust Models for Biomedical Natural Language Processing", arXiv preprint arXiv:1902.07669, 2019.

A Problem-Solution based Patent Analysis Approach for Business Opportunity Discovery

Chung Jaemin, Jeong Byeongki & Yoon Janghyeok

Identifying business and technology opportunities based on patent intelligence is essential for a firm's sustainable growth. Therefore, studies that analyze patents as technologies developed to solve certain problems has emerged. Despite these studies, they have limitations in identifying firm-specific business and technology strategies, application of patents in various countries and making full use of the patent bibliographic information. In this regard, this paper suggests an approach that identifies specific business strategies by recommending a business area and technology to a firm by defining patent keywords as problems and classification codes as solutions. We first collect patent data, and construct firms' problem and technology portfolios. Then, we recommend problem and technology entities based on the capabilities of the target firm. Finally, we analyze association rules of entities to identify business and technology opportunities. We expect that our approach will contribute to discover feasible opportunities based on the firm's internal capabilities and will be used as a Problem-Solution based patent intelligence tool for patents in various countries using text and bibliographic information.

Keyword

Patent intelligence, Technology opportunity discovery, Business development, Problem-Solution, Collaborative filtering, Association rule mining