

정형/비정형데이터 기반 반도체 시장 이슈 분석 및 예측 모형 개발

고동환/김민식/김봉진/임종호/김현태

2023. 12



정책자료 | 23-07

정형/비정형데이터 기반 반도체 시장 이슈 분석 및 예측 모형 개발

고동환/김민식/김봉진/임종호/김현태

2023. 12

목 차

요약문	11
제 1 장 서 론	17
1. 반도체 산업의 역사	17
제 2 장 정책 동향	20
제 1 절 국내 반도체 정책 현황	20
1. 우리나라의 반도체 및 인공지능 반도체 지원 정책 개요	20
2. 국내 반도체 산업 육성을 위한 주요 전략 및 내용	22
3. 인공지능 반도체 산업 육성을 위한 정책 현황	28
제 2 절 해외 반도체 관련 주요 정책 현황	36
1. 주요 기술 선도국의 반도체 관련 정책 개요(2021년~2023년)	36
2. 미국	37
3. 중국	38
4. EU	40
5. 일본	41
제 3 장 미국 등록 특허를 활용한 기술경쟁력 분석	43
제 1 절 개 요	43
1. 특허 기술 분류 및 데이터 수집	43
2. 분석 지표	43
3. 국가별 반도체 특허 동향	47

2 정책자료 23-07

4. 반도체 기술 분야별 특허 등록 추이	51
5. 반도체 기술 분야별 기술경쟁력 분석	57
6. 반도체 주요 기업별 등록 특허 분석	65
7. 주요 기업의 공동인용 분석	68
8. 공동협력 동향 분석	74
9. 기술 의존도(Backward Citation Analysis)	88
제2 절 주요국 반도체 스타트업의 투자유치 동향 분석	91
1. 데이터	91
2. 주요국 반도체 스타트업 투자유치 추이	92
3. 마중 갈등의 반도체 스타트업 투자 영향	95
4. 시사점	97
제4 장 반도체 시장 예측을 위한 모형 개발	99
제1 절 반도체 수출 전망 모형	99
1. 선행연구	99
제2 절 반도체 시장 불확실성 지수 개발	111
1. 추진 전략 및 방법	111
2. 선행연구	112
3. 지수 구축 방법	115
4. 반도체 시장 불확실성지수 계산	132
제3 절 반도체 시장 공급망 근사	145
1. 공급망 근사 방법	145
2. 반도체 공급망	147
3. 반도체 공급망 근사 네트워크	153
제5 장 결 론	169

제 1 절 시사점	169
제 2 절 한계 및 개선 방향	171
참고문헌	176
[부 록]	179

표 목 차

〈표 2-1〉 우리나라(인공지능) 반도체 산업 육성 정책 및 주요 내용	21
〈표 2-2〉 시스템 반도체 비전과 전략	22
〈표 2-3〉 시스템반도체 기술 혁신 지원	24
〈표 2-4〉 『K-반도체 전략』의 핵심 전략과 주요 내용	25
〈표 2-5〉 반도체 인력양성 추진 전략과 주요 내용	27
〈표 2-6〉 AI 기술경쟁력 확보를 위한 세부 정책	28
〈표 2-7〉 인공지능(AI) 반도체 산업 발전 전략의 12개 세부 과제	30
〈표 2-8〉 국산 인공지능 반도체 고도화 3단계	33
〈표 3-1〉 K-PEG 평가 항목	46
〈표 3-2〉 국가별 K-PEG 등급(%) 분포	51
〈표 3-3〉 기술 분류별 미국 등록 특허 건수 추이	52
〈표 3-4〉 국가별 기술별 피인용도 지수	58
〈표 3-5〉 국가별 기술별 기술 영향력 지수	59
〈표 3-6〉 국가별 FAB 공정 분야의 K-PEG 평가 결과의 등급별 분포 현황	61
〈표 3-7〉 국가별 설계 분야의 K-PEG 평가 결과의 등급별 분포 현황	62
〈표 3-8〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황	63
〈표 3-9〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황	64
〈표 3-10〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황	65
〈표 3-11〉 반도체 분야의 상위 20개 기업 등록 특허 현황	65
〈표 3-12〉 상위 20개 기업 특허 지표	67
〈표 3-13〉 국가별 공동특허 현황	74

〈표 3-14〉 국가 간 공동특허 건수	75
〈표 3-15〉 주요 기업의 공동특허	76
〈표 3-16〉 미국기업별 공동특허 현황	78
〈표 3-17〉 일본 기업들의 기업별 공동특허 현황	80
〈표 3-18〉 대만 기업들의 기업별 공동특허 현황	82
〈표 3-19〉 한국 기업들의 기업별 공동특허 현황	84
〈표 3-20〉 중국 기업들의 기업별 공동특허 현황	84
〈표 3-21〉 프랑스 기업들의 기업별 공동특허 현황	86
〈표 3-22〉 네덜란드 기업들의 기업별 공동특허 현황	87
〈표 3-23〉 영국 기업들의 기업별 공동특허 현황	87
〈표 3-24〉 반도체 분야 미국 등록 특허의 국가별 자기 인용 현황	89
〈표 3-25〉 반도체분야 미국등록특허의 기업별 자기인용 현황	90
〈표 3-26〉 반도체 분야 미국 등록 특허의 기업별 Backward Citation 현황	91
〈표 4-1〉 반도체 시장 불확실성 소스 및 주요 키워드	120
〈표 4-2〉 반도체 제품의 분류	148
〈표 4-3〉 반도체 생태계 분야별 특징 및 주요 기업	151
〈표 4-4〉 분석 대상 기업	155
〈표 4-5〉 V0 행렬: 후공정	156
〈표 4-6〉 거래 관계 규칙 단어	157
〈표 4-7〉 V1 행렬: 후공정	157
〈표 4-8〉 구조화된 데이터: 후공정	158
〈표 4-9〉 근사 결과 비교: 후공정	159
〈표 4-10〉 근사 결과 비교: 장비	161
〈표 4-11〉 근사 결과 비교: 소재	163
〈표 4-12〉 기간별 공급망 지도 성능 비교: 후공정	165

6 정책자료 23-07

<표 4-13> 기간별 공급망 지도 성능 비교: 장비 166
<표 4-14> 기간별 공급망 지도 성능 비교: 소재 167
<표 4-15> 확장된 거래 관계 규칙 단어 168
<표 5-1> 향후 과제 173

그림 목 차

[그림 2-1]	‘시스템반도체 비전과 전략’ 정책의 구조	23
[그림 2-2]	정책효과 전망	26
[그림 2-3]	인공지능(AI) 반도체 산업 발전 전략 목표 및 추진 전략	29
[그림 2-4]	인공지능(AI) 반도체 산업 성장 지원 대책 개요	32
[그림 2-5]	케이-클라우드 추진 방안	35
[그림 3-1]	반도체 분야 미국 등록 특허 추이	47
[그림 3-2]	국가별 기술 집중도 지수(RTA)	48
[그림 3-3]	국가별 피인용지수(CPP)	49
[그림 3-4]	국가별 / 구간별 기술순환주기	50
[그림 3-5]	국가별 FAB 공정 관련 등록 특허 추이	53
[그림 3-6]	국가별 설계 관련 등록 특허 추이	54
[그림 3-7]	국가별 소재 관련 등록 특허 추이	55
[그림 3-8]	국가별 장비 관련 등록 특허 추이	56
[그림 3-9]	국가별 차세대반도체 관련 등록 특허 추이	57
[그림 3-10]	국가별 기술별(로그) 기술력 지수	60
[그림 3-11]	FAB 분야 주요 기업의 Co-citation clustering 분석	69
[그림 3-12]	설계 분야 주요 기업의 Co-citation clustering 분석	70
[그림 3-13]	소재 분야 주요 기업의 Co-citation clustering 분석	71
[그림 3-14]	장비 분야 주요 기업의 Co-citation clustering 분석	72
[그림 3-15]	차세대반도체 분야 주요 기업의 Co-citation clustering 분석	73
[그림 3-16]	미국 기업 간 공동특허 현황	77

[그림 3-17] 일본 기업들의 전체 공동특허 현황 79

[그림 3-18] 대만 기업들의 전체 공동특허 현황 81

[그림 3-19] 한국 기업들의 전체 공동특허 현황 83

[그림 3-20] 중국 기업들의 기업별 공동특허 현황 85

[그림 3-21] 미국 반도체 스타트업 연도별 투자유치 금액 추이 93

[그림 3-22] 중국 반도체 스타트업 연도별 투자유치 금액 추이 94

[그림 3-23] 유럽 반도체 스타트업 연도별 투자유치 금액 추이 94

[그림 3-24] 한국 반도체 스타트업 연도별 투자유치 금액 추이 95

[그림 3-25] 주요국 반도체 스타트업 연도별 투자유치 합계 추이 96

[그림 3-26] 주요국 반도체 스타트업 연도별 투자유치 합계 추이 97

[그림 4-1] 비정형 텍스트 데이터를 활용한 시장위험지수 112

[그림 4-2] 반도체 시장 불확실성 지수 산출 과정 모식도 116

[그림 4-3] 예시문장을 통한 형태소 분석기 비교 118

[그림 4-4] 다양한 품사를 기반으로 구축한 키워드 119

[그림 4-5] Burstiness 활용한 키워드 도출 예시 122

[그림 4-6] 텍스트 기사를 수치형 feature 변수로 변환한 예시 123

[그림 4-7] 가우시안 혼합 모형(Gaussian Mixture Model) 126

[그림 4-8] 시간에 따른 뉴스기사 수와 리스크 정도의 변화 양상 131

[그림 4-9] 데이터 형식 예시 133

[그림 4-10] 기간별 반도체 관련 기사 개수 그래프 134

[그림 4-11] 위험 요인별 특성값의 히스토그램 135

[그림 4-12] 산업 하위 카테고리별 * 위험 요인별 * 월별 시장위험지수 .. 138

[그림 4-13] 산업 하위 카테고리별 * 월별 시장위험지수 139

[그림 4-14] 위험요인별 * 월별 시장위험지수 140

[그림 4-15] 월별 최종위험지수(EMA 비교) 141

[그림 4-16] 반도체 최종시장위험지수 계산 결과 142

[그림 4-17] 정책불확실성 요인 시장위험지수 계산 결과	143
[그림 4-18] 외교불확실성 요인 시장위험지수 계산 결과	144
[그림 4-19] 거시경제 불확실성 요인 시장위험지수 계산 결과	144
[그림 4-20] 외부충격 불확실성 요인 시장위험지수 계산 결과	145
[그림 4-21] 반도체 공급망 근사 프로세스	146
[그림 4-22] 반도체 산업의 수직통합과 수직분업	147
[그림 4-23] 반도체 산업의 수직통합과 수직분업	149
[그림 4-24] 반도체 8대 공정	151
[그림 4-25] 공급망 지도 시각화 결과: 정형 데이터-후공정	160
[그림 4-26] 공급망 지도 시각화 결과: 비정형 데이터-후공정	160
[그림 4-27] 공급망 지도 시각화 결과: 정형 데이터-장비	162
[그림 4-28] 공급망 지도 시각화 결과: 비정형 데이터-장비	162

요 약 문

1. 제 목

정형/비정형데이터 기반 반도체 시장 이슈 분석 및 예측모형 개발

2. 연구 목적 및 필요성

‘국가 첨단전략산업’ 중 하나인 반도체 산업에 대한 효과적인 지원을 위해서는 신뢰성과 속보성을 모두 만족할 수 있는 다양한 데이터를 구축하고, 이를 활용한 분석을 종합하여 정책을 지원하는 데이터 기반 의사결정 체계를 수립하는 것이 필요하다. 더불어 디지털 기술의 발달로 인해 데이터의 축적이 가속화하고 있으므로 다양한 형태의 데이터로부터 유용한 정보를 획득 및 활용한 정책 의사결정 체계가 필요한 시점이다. 정확도와 신뢰성이 높은 국가 승인 통계는 조사와 공표의 시차가 존재하여 속보성을 확보하기 어려우나, 온라인뉴스와 같은 텍스트 정보는 정확성 측면에서는 다소 떨어지지만 빠른 주기로 획득이 가능하므로 적극적으로 비정형데이터의 활용 방안을 검토할 필요가 있다. 본 연구의 목적은 정책당국 및 시장 참여자가 합리적 의사결정을 할 수 있도록 가용한 모든 데이터를 활용하여 반도체 시장의 이슈를 분석하고, 전망함으로써 그 결과를 종합하여 정부의 정책 수립을 위한 기초자료 제공에 있다.

3. 연구의 구성 및 범위

본 연구는 크게 네 가지 세부 과업으로 나눌 수 있다. ① 반도체 시장 관련 이슈

분석으로 1차연도에는 특허정보를 활용하여 반도체 분야 기술경쟁력을 분석했다. 또한 미·중 갈등의 반도체 분야 스타트업의 자금 조달에 대한 영향을 데이터에 기반하여 검토하는 것이다. 다음으로는 ② 계량경제학 기반 반도체 산업 전망 모형을 위해 다양한 방법론을 검토하는 것이다. 기존 반도체 시장 전망 관련 연구와 최신의 예측방법론을 검토하여 반도체 시장에 적용가능성을 검토하였다. ③ 세 번째로는 온라인뉴스 데이터를 활용하여 반도체 시장 관련 불확실성지수를 개발하는 것이다. ④ 마지막으로 텍스트 정보 기반 반도체 공급망 근사 네트워크를 구축하는 것이다.

이를 위하여 반도체 관련 온라인 텍스트를 수집할 수 있는 소스를 특정하고 2010년부터 2022년까지 텍스트 데이터를 수집하여 DB를 구축하였다. 각 모형은 이렇게 수집된 데이터를 활용하였다.

4. 연구 내용 및 결과

첫 번째로, 한국특허기술진흥원의 미국 등록 특허 DB를 활용하여, 반도체 기술 분류체계에 따른 특허자료를 수집하여, FAB 공정, 설계, 소재, 장비, 차세대반도체 5개 부문에 대해 국가별로 양적, 질적인 평가를 하고 분석했다. 구체적으로는 등록 특허 건수, 현시선호도, 피인용지수, 기술 영향력 지수, 기술력지수, K-PEG, 기술순환주기, 과학연계지수 등을 고려하였다. 반도체 관련 특허는 차세대반도체 분야를 제외하고는 등록 건수가 대체로 감소하는 추세이다. 차세대반도체 분야에서는 미국이 압도적인 성과를 보이며, 국가적 차원의 핵심 전략이 차세대반도체 분야에 있음을 추론할 수 있었다. 기초과학 연구개발이나 피인용 지수등 특허의 영향력 관점에서는 네덜란드가 미국보다 뛰어난 사례가 많이 관측되었다. 그러나 양적지표와 질적지표를 종합적으로 고려하여 도출한 기술력지수에서는 역시 미국이 가장 앞서 있음을 확인하였다. 특허의 질적인 측면에서 우리나라는 대만과 비교하였을 때 대부분의 기술 분류에서 상위 등급의 등록 특허를 더 많이 보유하고

있는 것으로 나타났다. 중국은 질적인 지표에서 주요 10개국 중 최하위 수준으로 아직 반도체 선진국과의 격차가 확연하게 있음을 다시 한번 확인하였다.

두 번째로 반도체 시장 전망을 위한 계량모형 개발을 위해 기존 연구와 최신 빅데이터 기반 예측방법론을 검토하였다. 반도체 시장은 주기성이 매우 뚜렷하고 변동성이 큰 시장이므로 목표변수를 무엇으로 하는지에 따라 다양한 접근방법을 검토할 수 있다. 단순 수출을 전망하기 위한 방법으로는 ARIMA, VCEM과 같은 시계열 모형, Local Projection과 같이 회귀분석 기반에 비선형, 비대칭 반응을 반영할 수 있는 모형이 있었다. 기존 반도체 시장을 전망하는 모형에서는 Regime Switching을 고려한 BVAR 계열의 연구가 많이 있었다. 최근에는 기계학습이나 딥러닝처럼 인공지능을 활용한 연구들이 활발히 진행되고 있으나, 반도체 시장에 적용한 사례는 제한적이었다. KISDI 연구에서는 Mixed DAat Sampling (MIDAS) 모형에 주기가 매우 짧은 온라인 지수를 활용하여 예측한 사례도 있다. 우리나라 반도체 수출은 주로 메모리 반도체로 비교적 짧은 주기의 수출데이터를 확보할 수 있으므로, 수출을 예측할 때 예측의 지평이나 주기를 결정하는 것이 매우 중요하다. 정책 수요를 검토하여 목표변수를 설정하고 모형을 결정하는 과정이 필요하다.

세 번째로는 온라인텍스트에 기반한 반도체 시장 불확실성 지수를 개발하는 것이다. 1차연도에는 네이버뉴스, 삼성전자 반도체룸, SK하이닉스 뉴스룸, 미국의 반도체협회로부터 수집한 텍스트를 활용하였다. 2010년부터 2022년까지 데이터를 수집하였다. 반도체 시장 불확실성지수의 체계는 먼저 소재와 장비, 설계, 파운드리, OSAT, IDM으로 반도체 기사 카테고리를 나누었다. 각 카테고리의 불확실성에 영향을 줄 수 있는 요인으로 5가지를 고려하였는데, 거시경제적 충격, 정책불확실성, 반도체 시장 내 충격(수요/공급 충격), 대외환경(외교) 충격, 자연재해 충격으로 구분하였다. 각 기사를 5가지 요인에 대해 충격으로 분류될 확률을 부여하고, 이를 평시와 위험도가 높은 상황과의 상대적 점수로 최종 지수를 산출하였다. 도출된 지수는 대체로 큰 불확실성 요인에 점프가 일어나는 것을 확인할 수 있었다. 그러나, 반도체 관련 주가지수와 반대로 움직이는 경향 또한 관측되었다. 반도체 주가는 반도체 기업의 실적이나 실물경제 보다 선행한다는 기존의 합

의된 관측 사실에 기초해 볼 때 지수의 부적합을 논의하기는 이르다고 판단했다. 향후 소스의 확대와 분류 모형의 고도화를 통해 지수를 개선함과 동시에 지수의 해석을 위해 전문가 풀을 활용할 계획이다.

네 번째로는 온라인텍스트를 활용하여 반도체 시장의 공급망을 근사하고자 했다. 반도체 시장의 공급망 이슈는 매우 중요한 정책적 관심의 대상이지만, 실제 깊이 있는 거래데이터를 확보하여 정확한 네트워크를 구성하기는 사실상 불가능하다. 이에 본 연구는 온라인텍스트를 활용하여 데이터 확보의 한계를 극복할 수 있는지 검토하였다. 우선 신용평가 데이터로부터 국내 주요 반도체 중소기업의 거래데이터를 확보하여 공정별로 네트워크를 구성하였다. 대기업의 1차 벤더는 비교적 쉽게 확인이 가능하기 때문에, 중소기업을 중심으로 대기업의 2차 벤더 이상의 네트워크를 구성해보고자 시도했다. 이를 위해 수집된 온라인 텍스트에서 명사만을 추출하여 반도체 기업을 식별하였다. 그리고 난 후 동시출현빈도를 통해 네트워크의 연결 강도와 노드를 설정하였다. 1차연도의 최종목표는 실제 거래데이터(신용평가데이터) 기반 네트워크를 온라인텍스트를 활용한 네트워크로 근사하는 것이다. 하지만 분석 결과 동시출현빈도로 구축한 네트워크와 실제 거래데이터 기반 네트워크에는 상당한 차이를 보였다. 그러나 2차연도에 데이터 소스가 확대되고, 활용할 수 있는 식별된 반도체 기업(labeled data)이 많아지면 충분히 네트워크를 구성할 수 있을 것으로 기대한다. 또한 온라인 뉴스기사 중에 실제 거래 관계를 유추할 수 있는 기사를 추려, 그 기사 안에서 동시출현빈도를 추출하여 활용한다면 충분히 1차연도의 한계를 극복할 수 있을 것으로 판단한다.

5. 기대효과

비정형 데이터를 활용하여 반도체 시장 이슈를 분석하고 전망하는 선행연구는 매우 제한적이다. 기술주기가 빠르고, 다양한 이슈에 영향을 받는 분야이므로, 대부분 정성적 연구와 전문가들의 의견 수렴을 통해 전망과 분석이 이루어졌다. 본

연구는 기존의 연구 방법을 데이터에 기반하여 보완하고, 훨씬 명확한 정책적 시사점을 도출할 수 있을 것으로 기대한다. 1차년도 연구과정에서 많은 한계점과 개선 가능한 점들이 식별되었으므로, 학술적으로 또 정책적으로 활용도가 매우 높은 모형을 개발할 수 있을 것이며, 이를 통해 빠르게 변화하는 반도체 시장에 대한 정량적 분석과 전망이 가능할 것으로 기대한다.

제1장 서론

집적 회로는 스마트폰, 컴퓨터, 그리고 다른 전자 제품들뿐만 아니라 셀 수 없이 많은 다른 제품들에도 필수적이다. 마이크로 칩은 자동차, 토스터, 로봇, 냉장고, 그리고 비행기에 동력을 공급한다. 게다가, 최첨단 칩은 국방에 중요한 인공지능과 다른 응용 프로그램들을 구동한다. 정책 입안자들이 반도체 장치의 안정적인 공급 확보에 집중하는 것은 놀랄 일이 아니다.

2017년, 미국의 전 대통령인 도널드 트럼프는 중국 반도체 산업에 대한 무역 전쟁을 시작했다. 트럼프 행정부는 중국 반도체 회사들에 대한 불투명한 보조금에 대한 불공정한 무역 조사를 시작했는데, 국가 안보를 이유로 나사를 조여서 미국 회사들이 화웨이, 반도체 제조 국제 공사(SMIC), 그리고 다른 중국 회사들에 판매하는 것을 금지하고, 이 회사들에 반도체를 공급하는 해외 회사들에 대한 미국 기술의 판매까지 제한했다.

트럼프의 조치로 반도체 산업에 불확실성이 높아진 상황에서 코로나19가 확산하면서 반도체 시장에 큰 충격을 주었다. 2020년 2월 부정적인 거시경제 전망으로 인해 대만에서는 반도체 주가가 30%, 한국과 일본에서는 더 하락했다. 봉쇄에 직면한 자동차 제조업체들은 칩 주문을 취소했고, 개인들이 재택근무를 시작하면서 정보통신기술(ICT)에 대한 지출이 급증하면서 칩 수요는 증가했다. 따라서 2020년 후반 자동차 수요가 회복되었을 때 반도체 제조업체는 자동차 제조업체에 판매할 칩이 부족했다. 단 2달러의 비용이 들 수 있는 이러한 반도체 장치의 부족으로 인해 자동차 공장은 생산을 중단하는 상황에 이르렀다.

1. 반도체 산업의 역사

1980년대 산업 표준이 CMOS로 바뀌었을 때 일본 기업들은 극의 위치에 있었습니다. Bown(2020)에 따르면, 1980년 매출액 기준 4대 반도체 기업 중 3개가

미국 기업이었고, 1990년에는 상위 4개 중 3개가 일본 기업이었다. 또한 일본 기업들이 DRAM 분야에서 도약하면서, 1978년에는 30% 미만이었던 시장 점유율이 1986년에는 75% 이상으로 높아졌다. 미국 제조업체들은 일본의 거대기업에 대해 보호무역주의 로비를 통해 대응했다. 그들은 미국 기업이 일본 시장에 접근할 수 없다고 주장하며 Section 301의 대일 소송을 제기했다. 그들의 불만 중 하나는 과거 일본 정부가 미국 반도체 기업을 일본 시장에서 공공연히 배제했다는 것이었다. 미국 정부는 일본 기업이 미국에서 칩을 공정가치보다 낮게 판매했다며 3건의 반덤핑 소송을 제기했다. 1986년 미국과 일본이 합의하면서 일본 반도체시장의 20%를 미국 기업에 할당하고, 일본 기업이 수출 물량을 제한하고 가격을 인상하기로 했다. 정부령으로 시장 점유율을 결정하는 것은 미국의 통상정책에 큰 변화가 있었음을 의미한다. 한편, 한국에서는 정부의 암묵적인 지원 덕에 삼성이 미국 마이크론사로부터 D램 기술을 사들여 사업을 시작하게 되었다. 일본 기업들이 미국에 더 높은 가격과 한정된 수량으로 수출해야 하는 상황에서, 한국 기업들은 양적인 제약 없이 더 높은 가격에 판매할 수 있었다. 1990년대 초 삼성은 DRAM의 선두 생산업체가 되었는데, 이는 2021년에도 여전히 유지되고 있다. 2021년 2위 생산업체인 SK하이닉스도 한국 기업이다. 한국이 D램을 비롯한 메모리 칩에서 우위를 점하는 상황에서 대만은 로직 칩에 집중했다. 1970년대 대만은 1973년 석유파동 이후 인플레이션에 의해 파괴되었고, 1974년에는 일본 자본재에 대한 접근이 끊겼습니다. 이러한 위기 환경에서 국가 생존을 위한 수단으로서의 경제개발로 전환했습니다. 미국의 많은 중국 연구자가 대만의 처지에 공감하고 자신의 전문성을 무료로 공유했다. 대만은 반도체 기술 개발을 감독하기 위해 산업 기술 연구소(ITRI)를 설립했고, 1979년 UMC로 분사했다. ITRI는 계속해서 기술 개발을 육성했고, 1987년에 대만 반도체 제조 회사인 TSMC도 설립했습니다. TSMC는 집적 회로 자체를 설계하는 것이 아니라 고객의 사양에 따라 만들었습니다.

중국은 2020년 현재 반도체의 세계 최대 소비국이며, 스마트폰의 90%, 스마트 텔레비전의 67%, 개인용 컴퓨터의 65%를 생산하며, 세계 반도체 수출의 20%를

차지하고 있다. 그런데, 수입은 최첨단 기기인 반면, 수출은 저가의 칩이다 보니, 중국 정부는 국내 반도체 산업을 촉진하고 따라서 자급자족을 추구하기 위해 적극적으로 지원하고 있다. 미국 정부는 이러한 보조금이 비시장적이고 국가 안보를 위협한다는 명분 아래 트럼프 행정부는 중국 반도체 산업을 상대로 무역 전쟁을 시작했다. 2017년 미국 정부는 불공정 무역 행위에 대한 Section 301 조사를 시작했다. 이에 따라, 중국산 칩 수입품에 25%의 관세가 부과되었다. 2019년 미국은 다시 화웨이를 겨냥했다. 화웨이는 5세대(5G) 통신망 장비 생산의 선두 주자였는데, 미국은 화웨이가 입수한 민감한 군이나 민간 데이터를 백도어를 통해 중국 정부에 제공해야 할 수 있다고 우려했다. 트럼프 행정부는 이후 미국산 칩의 화웨이 판매를 제한했지만, 화웨이는 대만, 한국 및 기타 출처로부터 충분한 칩을 확보할 수 있었다. 이에 미국은 화웨이에 공급한 외국 반도체 제조업체는 더 이상 미국의 도구와 기술에 접근할 수 없게 했다. 2020년에는 미국 기업이 정부의 명시적 승인 없이 SMIC에 판매하는 것을 금지하기도 했습니다. 이러한 환경 속에서 코로나19 사태가 반도체 부문을 강타했다.

제 2 장 정책 동향

제 1 절 국내 반도체 정책 현황

1. 우리나라의 반도체 및 인공지능 반도체 지원 정책 개요

정부는 오랜 기간 도안 반도체 분야의 산업경쟁력을 높이고, 선도적인 국가로 발돋움하기 위해 지속해 종합적·다각적인 전략을 수립 및 추진했다. 역대 정부는 기술개발(R&D) 지원, 전문인력 양성, 대·중소기업 동반성장, 생태계 육성, 펀드 조성, 세액공제, 통상환경 개선 등을 지원하는 종합계획을 범부처 차원에서 지속적으로 발표하고 세부 사업들을 추진하고 있다. 특히 국내 메모리반도체 분야는 고도화된 기술 기반의 과점산업 구조로 미세공정개발 R&D 투자 및 자본집약도가 큰 설비투자가 요구되는 종합반도체업체(IDM) 중심의 대기업이 주축이어서 정부는 세제지원 및 규제완화 등을 통해 지원하고 있다.

최근 반도체 산업에서 파운드리 영역이 차지하는 중요성이 높아지고 시장 규모도 확대되고 있어, 파운드리 경쟁력 제고를 위한 정부 차원의 노력도 이어지고 있다. 또한 국내 시스템반도체 산업은 팹리스, 테스트, 패키징 등 분야별 혁신기업과 중소기업들이 많이 있어 구축, R&D 지원, 기술사업화, 투자 연계, 기반 조성 등 다양한 성장 지원 및 생태계 조성을 위한 정책을 이어지고 있다.

특히 인공지능 기술 및 데이터 생태계 기반이 되는 인공지능 반도체의 기술력 및 기능이 통신, IoT, 자동차, 가전 등 다양한 산업경쟁력의 관건이 되어감에 따라, 정부는 인공지능 반도체 산업을 전략적으로 육성하기 위해 집중적인 투자를 시작하고 있다.

〈표 2-1〉 우리나라(인공지능) 반도체 산업 육성 정책 및 주요 내용

정책/부처발표시기	주요내용
시스템반도체 및 장비산업 육성전략 [산업통상자원부(2010. 7. 31)]	시스템반도체 및 반도체장비 산업을 체계적으로 육성하기 위해 수요기업-팹리스 연계, 중소·중견 시스템반도체·장비 육성, 반도체 클러스터 조성, 인력육성정책 추진
반도체산업 재도약 전략 [산업통상자원부(2013. 10. 23)]	반도체를 성장동력으로 지속 육성하기 위해, 메모리, 시스템반도체, 장비·소재, 인력·인프라 분야 집중 지원
지능형반도체 전문인력 양성사업 [산업통상자원부(2016. 9. 29)]	기업수요에 맞는 고용연계형 석사급 반도체 전문인력 양성하고, 바로 취업으로 이어질 수 있는 생태계 구축
반도체 희망펀드 조성 [산업통상자원부(2016. 10. 27)]	삼성전자, SK하이닉스, 정책금융기관 등이 2천억 규모의 반도체펀드를 조성하여 팹리스, 소재/장비기업에 투자
시스템반도체 산업경쟁력 강화방안 [산업통상자원부(2017. 3. 30)]	시스템반도체 선도국으로 도약, 시스템반도체 R&D 및 인재양성, 글로벌 수요연계 지원
시스템반도체 비전과 전략 [관계부처 합동(2019. 4. 30)]	시스템반도체 전반을 본격 육성하고 종합반도체 강국으로 도약, 팹리스 성장기반구축 및 파운드리 경쟁력 제고
인공지능 국가전략 [과학기술정보통신부(2019. 12. 17)]	IT강국을 넘어 인공지능(AI) 강국 달성을 위해 경제·사회 전반의 혁신을 위한 9대전략 제시. 특히 AI 반도체 및 신개념 반도체(PIM) 개발에 전략 투자
인공지능 반도체산업 발전전략 [관계부처 합동(2020. 10. 12)]	인공지능 반도체 관련 기술개발(R&D) 강화, 기술사업화, 인력양성, 금융지원, 기반구축, 생태계 구축 지원을 통해 인공지능·종합반도체 강국 실현
시스템반도체 기술혁신 지원 [관계부처 합동(2021. 2. 1)]	K-팹리스 성장지원, 유망시장 선점, 신시장도전(AI 반도체, PIM반도체), 예타준비사업(K-센서, 차세대 PIM 반도체) 추진
K-반도체전략 [대한민국 정책브리핑(2021. 5. 13)]	K-반도체 벨트를 구축해 세계 최대·최첨단 반도체 공급망을 구축하고, 2030년까지 510조원 이상의 민간투자자로 초격차 유지, 반도체 인력 3만 6000명 육성
인공지능(AI) 반도체 산업 성장 지원 대책 [과기정통부(2022. 6)]	인공지능반도체 첨단기술 연구에 향후 5년간 1조 200억원 투입·국산 인공지능반도체 초기수요 창출을 위한 대형 가늌터 구축 및 공공사업 적용 확대·차세대 반도체 상용화 위한 대기업과 산·학·연 간 첨단 공정 기술 협력 ·인공지능반도체 전문인력 7,000명 이상 양성

정책/부처발표시기	주요내용
반도체 관련 인재 양성방안 [관계부처 합동(2022. 7)]	향후 10년간 약 13만 명의 신규 인력수요에 대응해 기술발전을 선도할 반도체 혁신인재를 2031년까지 15만 명 이상 양성
국산 인공지능 반도체를 활용한 케이(K)-클라우드 추진방안 [과기정통부(2022.12)]	세계 최고 수준의 초고속·저전력 국산 인공지능반도체를 개발, 이를 데이터센터에 적용해 국내 인터넷기반자원공유 경쟁력을 강화하고, 국민들에게 향상된 인공지능 및 인터넷기반자원공유 서비스를 제공하기 위한 정책

2. 국내 반도체 산업 육성을 위한 주요 전략 및 내용

가. 시스템반도체 비전과 전략(2019. 4)

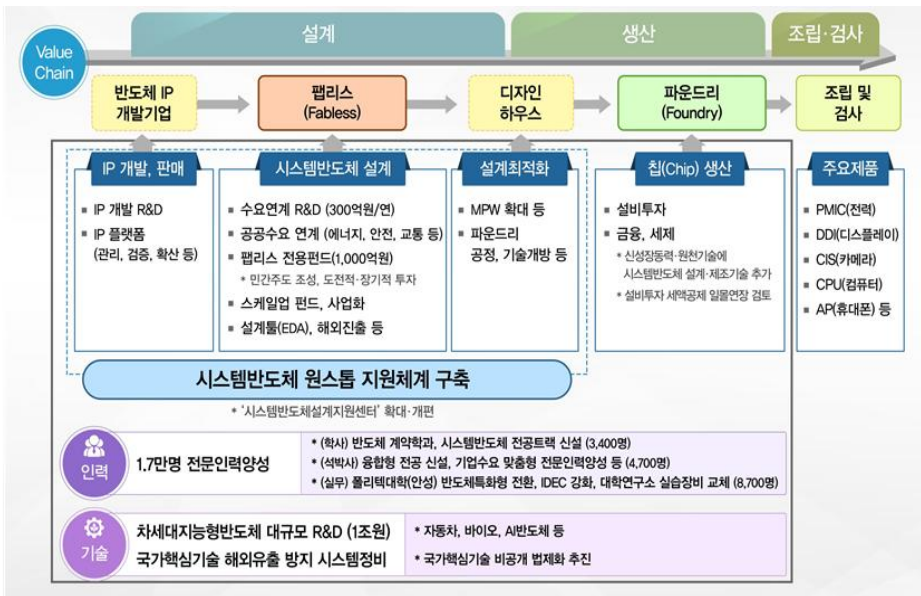
정부는 시스템반도체 산업을 육성하고 2030년까지 종합반도체 강국으로 도약하기 위한 5대 중점대책을 수립하고, 이를 통해 2030년까지 1) 파운드리 세계 1위, 2) 팹리스 시장 점유율 10% 달성, 3) 2.7만 명의 신규일자리를 창출할 계획을 발표했다. 정부는 구체적으로 시스템반도체 산업육성을 위해 ① 팹리스 지원, ② 파운드리 지원, ③ 생태계 구축, ④ 인력양성, ⑤ 기술지원 등 5개 분야에 대해 집중적인 중점대책을 수립하였다.

〈표 2-2〉 시스템 반도체 비전과 전략

비전	전략
팹리스 성장 기반 구축	<ul style="list-style-type: none"> • 대대적 수요창출(5대 시장 수요연계, 공공수요 2,600만 개·2,400억 원 이상 발굴, 5G와 시스템반도체 연계) • 전용펀드(1,000억 원) 신설, 스케일업 펀드 • 우수기업연구소 선정 등 스케일업 지원, 반도체 설계 톨 지원(추경 46 억원), • 창업부터 성장단계까지 원스톱 서비스 제공을 통해 글로벌 수준 성장 기반 마련
파운드리 지원	<ul style="list-style-type: none"> • 대표기업은 하이테크 첨단분야에서, 중견기업은 미들테크 틈새시장을 공략하고, 기업의 투자를 뒷받침할 수 있도록 세계·금융 지원 확대

비전	전략
생태계 구축	<ul style="list-style-type: none"> 디자인하우스 육성, 팹리스의 파운드리 공정 활용 확대 등 정부-업계 공동으로 상생협력 생태계 조성
인력양성	<ul style="list-style-type: none"> 계약학과 신설(연·고대, 2021~, 80명(기존정원 외) 및 단계적 확대 반도체 설계·공정 기술 R&D(정부, 기업 1:1 매칭) 지원 폴리텍대학(안성)의 반도체 특화형 전환, 대학 공정실습 Fab시설 확충(추경 100억)
기술개발	<ul style="list-style-type: none"> 산업의 패러다임을 바꾸는 차세대 반도체 기술을 확보 AI 반도체 등 차세대반도체 기술 개발에 향후 10년간 1조 원 투자 국가 핵심기술의 해외유출방지 시스템 정비

[그림 2-1] '시스템반도체 비전과 전략' 정책의 구조



자료: 과학기술정보통신부(2019. 12. 17)

나. 글로벌 K-팹리스 육성을 위한 '시스템반도체 기술 혁신 지원' 전략(2021. 2) 정부는 2021년 2월, 종합반도체 강국으로 도약을 위해 국내 팹리스 업체들의 경쟁력 제고가 시급하다고 판단하고, 국내 팹리스 기업의 성장 기반을 구축하고, 유망분야에 도전하기 위한 '시스템반도체 기술 혁신 지원' 전략을 발표하였다. 과

거 정부는 시스템반도체 비전과 전략(2019.4), 인공지능 반도체 산업 발전전략(2020.10)을 발표 이후 차세대반도체 경쟁력 제고를 위해 R&D 지원, 자금지원, 인력양성, 기술사업화 등 전방위적인 정책 지원을 이어오고 있다, 그런데도 국내 시스템반도체 업체(팹리스 등)들의 세계 경쟁력은 미미한 수준으로, 전력반도체 및 AI 반도체 등 차세대반도체를 해외수입에 의존하고 있다. 동 전략에서는 전방위적인 ‘팹리스의 성장 기반 구축’을 지원하고, 미래 수요 기반의 ‘유망시장 선점’을 지원하고, AI 반도체 등 ‘미래 신기술시장을 선점’하기 위한 정책 추진에 중점을 두고 있다.

〈표 2-3〉 시스템반도체 기술 혁신 지원

비전	전략
팹리스 성장 지원	팹리스의 지속 가능한 성장을 위해 국내·외 시장 맞춤형 R&D를 추진하고 수요 연계형 공동 R&D를 강화하는 정책 추진 <ul style="list-style-type: none"> • 글로벌 K-팹리스 육성을 위한 챌린지형 R&D • 설계, 소자, 공정 등 밸류체인 전반의 수요기업 기반 R&D를 통해 국내 수요-공급을 연계하는 대규모 프로젝트 추진 • 중소 팹리스의 창업·스케일업 촉진
유망시장 선점	모바일, 자동차, 헬스케어 등 미래 수요를 기반으로 성장 가능성이 높은 분야의 도전적 R&D를 추진하고, 특히 전력반도체, 센서 등 핵심 유망품목 적극 지원 <ul style="list-style-type: none"> • 미래 전력반도체 역량강화 및 기술선점 • 차세대 센서 R&D 신설 및 전주기 지원체계 마련
新시장 도전	AI 반도체 분야 초격차 기술력 확보를 위한 범부처 R&D를 추진하고, 원천기술 개발, 응용 기술 개발, 사업화·실증 연계 전주기 지원 <ul style="list-style-type: none"> • AI 반도체 플래그십 프로젝트(설계·소자·공정 기술 혁신 프로젝트) 추진 • 세계 최고 메모리 기반 신개념 PIM 반도체(메모리+프로세서 통합) 개발 • 기술·사업화 장벽 해소 및 실증 지원을 통해 초기 시장 창출 견인
예비타당성 준비 사업	동 전략을 원활하게 추진하기 위해 예타를 통해 대규모 사업추진 <ul style="list-style-type: none"> • 시장선도를 위한 K-Sensor 기술개발 사업 • 차세대 PIM반도체 기술개발사업

다. K-반도체 전략(2021. 5월)¹⁾

정부는 2021년 5월 ‘K-반도체 전략’을 발표하여, 2030년까지 세계 최고의 반도체 공급망을 구축하기 위해 K-반도체 벨트 조성, 인프라 지원 확대, 반도체 성장 기반 강화, 반도체 위기 대응력 제고 등 4대 추진 전략을 발표하였다.

〈표 2-4〉 『K-반도체 전략』의 핵심 전략과 주요 내용

	핵심전략	분야	주요 내용
전략 1	K-반도체 벨트조성	제조	첨단 메모리, 파운드리 제조시설이 집중 위치한 국내 최고의 선도형 반도체 단지 조성
		소부장	소부장 특화단지를 조성하고 50여개 협력사 유치, 메모리 양산 팹 연계 소부장 공급망 확보, 테스트~양산 원스톱 지원
		장비	화성, 용인, 천안은 글로벌 장비 기업과의 전략적 협업을 통해 첨단장비 연합기지 구축된 첨단 패키징 플랫폼 조성
		패키징	제품 소형화·고성능화 등 패키징 핵심역량 강화를 위해 실증·분석측정 장비가 구축된 첨단 패키징 플랫폼 조성
		설계	테크노밸리(제1판교)-설계지원센터(제2판교)-차세대 반도체 복합단지(제3판교) 등을 연결하는 팹리스 밸리 조성
전략 2	인프라 지원확대	세계	핵심전략기술 신설을 통해 R&D 및 시설투자를 통해 세계 인센티브 대폭 강화
		금융	국내 반도체 제조기업, 중소·중견 소부장 기업 등에 금융지원 확대. ‘반도체 등 설비투자 특별자금’을 신설
		규제	반도체 설비의 신속한 구축을 위해 고압가스안전관리법, 온실가스 배출권 규정, 화학물질관리법 등 규제 합리화
		기반	반도체 제조시설의 적기 건설과 원활한 운영을 위해 선제적 용수물량 확보, 전력 인프라 구축 및 공공폐수처리 지원
전략 3	반도체 성장기반 강화	인력	전주기 신규 전문인력 확보, 핵심인력의 해외 유출 방지를 위한 정책적·제도적 지원→ 반도체 산업인력 3.6만명 육성
		생태계	반도체 전방산업(수요-공급), 후방산업(소·부·장·소자)의 연대·협력 생태계 강화
		기술	차세대 전력 반도체, 인공지능 반도체, 첨단 센서, 소재·부품·장비 등 다각적인 반도체 신산업 기술역량 제고

1) 산업통상자원부(2021. 7. 1)

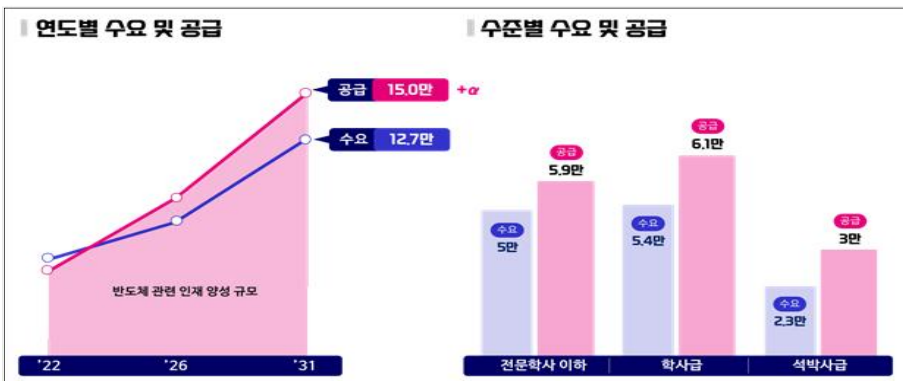
	핵심전략	분야	주요 내용
전략 4	반도체 위기 대응력 제고	특별법 제정	반도체업계, 반도체 특별법 제정 건의, 특별법 제정 여부 및 입법방향을 국회 및 관계부처와 논의
		차(車) 반도체	단기 수급 안정화를 적극 추진하고, 중장기적으로 미래차 핵심 반도체 개발을 위한 협력모델 발굴
		기술보호	국가핵심기술 지정 확대, 기술·인력 보안관리 강화, 기술안보 강화를 위한 범정부 협의체 구축 추진
		탄소중립	온실가스 감축을 위해 선제적 대응을 위한 R&D 지원 및 인증·실증 기반 강화

자료: 관계부처 합동(2021. 5. 13)

라. 반도체 인력양성 추진 전략(2022. 7)

정부는 향후 10년간 약 13만 명의 신규 인력 수요에 대응해 기술 발전을 선도할 반도체 혁신 인재를 2031년까지 15만 명 이상 양성하는 “반도체 관련 인재 양성방안”을 발표하였다. 본 반도체 인력양성 추진 정책 방안은 3대 분야 10대 과제로 구성되어 있다. (1) 과감한 규제 혁파와 지원으로 반도체 관련 정원 확대(양적 확대), (2) 고급 인력양성에 주력하면서 반도체 인재 양성 저변 확보(질적 제고), (3) 반도체 인재 양성 중장기 지원 기반 구축(기반 조성) 등으로 구성된 3대 분야와 더불어 3대 분야별 10과제가 매칭되어 제시되었다.

[그림 2-2] 정책효과 전망



자료: 관계부처합동(2022), 반도체 관련 인재 양성 방안

〈표 2-5〉 반도체 인력양성 추진 전략과 주요 내용

과제	내용
과감한 규제 혁파와 지원으로 반도체 관련 정원 확대(양적 확대)	산업 수요 대응을 위한 대학 정원 규제혁신 • 반도체 등 첨단분야 학과 신·증설 규제 완화, 계약정원제 도입·운영
	반도체 등 첨단산업 교원 및 대학 운영 규제 혁신 • 산업현장 전문가를 교수자원으로 적극 활용, 수업 운영요건 완화, 사내 대학 운영 활성화
	과감한 지원, 규제 특례를 통해 반도체 특성화 학과(전공) 육성 • 반도체 세부 분야별 특성화대학원을 지정·지원, 반도체 특성화대학(원)을 신규 지정, 반도체 분야 계약학과 확대, 직업계고 학과·교육과정 등 재구조화 지원
고급인력 양성에 주력하면서 반도체 인재 양성 저변 확보(질적 제고)	고급 전문인재: 반도체 기술 초격차를 이끌어낼 핵심 전문인재 육성, • 연구개발 투자 확대 및 반도체 세부분야 전문가 양성, 최고급 인재 유입·양성을 위한 연구개발·교육환경 개선, 산·학·연 프로젝트 및 인적 교류 활성화
	융복합인재: 초학문적 인재 육성을 위한 융합교육과정 활성화 • [단기] 반도체 분야 단기 집중과정 운영, [중기] 반도체 전공 트랙 운영 활성화
	실무인재: 산업현장에 즉시 투입가능한 맞춤형 실무인재 양성 • [단기] 직업계고 반도체 분야 단기 실무과정 확대 및 실습기반 확충, [중기] 현장 맞춤형 교육과정 개선
반도체 인재양성 증장기 지원기반 구축(기반 조성)	거점: 전국단위 대학·인프라 연계를 통한 연구·교육 집중 지원 • 대학 기반 전국 단위 반도체 연구·교육·실습 HUB 구축, 국가나노인프라 고도화를 통한 반도체 전문인력 양성 강화
	협업: 범부처·민관합동 인재양성 협업체계 구축 • 국가차원 인재양성 아젠다 발굴 및 정책 수립, 인적자원 관리·지원 체계 구축
	균형: 지역 및 중소기업 균형성장 및 상생 지원체계 • 지역 균형성장 지원체계 조성, 중소기업 균형성장·상생 지원체계 조성

3. 인공지능 반도체 산업 육성을 위한 정책 현황

가. 인공지능 국가전략(2019. 12)

정부는 'IT 강국을 넘어 AI 강국으로' 도약하기 위해 2030년까지 '디지털 경쟁력 세계 3위', 'AI를 통한 지능화 경제효과 455조 원 창출', '삶의 질 세계 10위' 등 3대 목표를 달성하기 위해, 3대 분야의 9대 전략과 100대 실행 과제를 마련하였다. 인공지능 국가전략의 비전 및 목표 달성을 위해 ① AI 인프라 확충, ② AI 기술경쟁력 확보(AI 반도체 관련), ③ 과감한 규제혁신 및 법제도 정비, ④ 글로벌을 지향하는 AI 스타트업 육성, ⑤ 세계 최고의 AI 인재 양성 및 전 국민 AI 교육, ⑥ 산업 전반의 AI 활용 전면화, ⑦ 최고의 디지털 정부 구축, ⑧ 포용적 일자리 안전망 구축과 역기능 방지와 AI 윤리 체계를 마련하는 등 9대 추진 전략과 방안을 발표하였다. 특히, 인공지능 국가전략(2019. 12)의 9대 전략 중 'AI 기술경쟁력 확보' 전략은 세계 최고의 메모리 반도체 경쟁력을 지렛대로 활용하여 AI 반도체 경쟁력 세계 1위를 목표로, AI 반도체 핵심기술 확보와 신개념 반도체(PIM) 개발에 대한 전략 투자를 강화하겠다는 내용을 담고 있다.

〈표 2-6〉 AI 기술경쟁력 확보를 위한 세부 정책

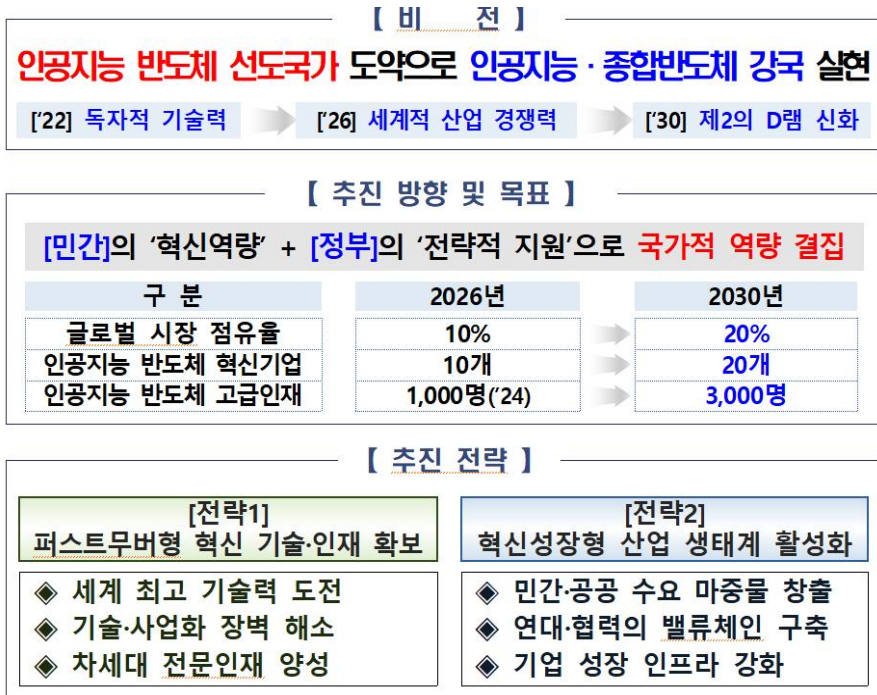
주제	세부내용
① 신개념 AI 반도체(PIM)개발	- AI 반도체 핵심기술*(설계·미래소자·장비 및 공정 등) 및 신개념 AI 반도체(PIM) 개발 등 AI 반도체 세계 1위 도약 * 차세대 지능형 반도체 1조 96억 원 투자('20~'29)
② 차세대 AI 연구개발	- 창의적·도전적 차세대 AI 연구개발* 선제 투자('20~) * 차세대 AI 개발사업 예타 추진 → '30년까지 핵심기술 5개 이상 확보
③ AI 기초연구 강화	- 지식표현 및 추론, 기계학습 알고리즘, 인지과학 등 AI 기초연구 강화('20. 예타 추진)
④ 혁신적 방식의 AI R&D 확대	- 선의의 경쟁과 창의적 도전을 촉진할 혁신적 방식의 AI R&D* 확대 * 서바이벌 방식의 '경쟁형', 국가 사회 현안을 AI로 해결하는 '챌린지형(경진대회)'

자료: 과학기술정보통신부(2019. 12. 17)

나. 인공지능 반도체 산업 발전 전략(2020. 10. 12)

정부는 2020년 10월 12일, 관계부처 합동으로 인공지능 반도체 선도 국가 도약으로 인공지능·종합반도체 강국 달성을 위한 ‘인공지능 반도체 산업 발전 전략’ 발표하였다. 정부는 인공지능 반도체의 독자적 기술력 확보(~2022), 세계적 산업 경쟁력 확보(~2026) 및 인공지능(AI) 반도체를 제2의 D램으로 육성(~2030)한다는 목표를 수립하고, 목표를 달성하기 위해 2대 전략과 12개 추진 과제를 마련하였다. 그리고 본 정책의 기대효과로 2026년까지 AI 반도체 시장 점유율 10%, 혁신기업 10개, 고급인재 1천명 달성, 2030년 AI 반도체 시장 점유율 20%, 혁신기업 20개, 고급인재 3천 명 달성 등을 이룩하고자 했다.

[그림 2-3] 인공지능(AI) 반도체 산업 발전 전략 목표 및 추진 전략



자료: 관계부처 합동(2020. 10. 12)

동 전략을 추진하기 위해 기술 개발(R&D) 강화, 기술사업화, 인력양성, 금융 지원, 기반 구축, 생태계 구축 등의 분야에 대해 2가지의 전략과 12개의 세부 추진 과제를 제시하였다.

〈표 2-7〉 인공지능(AI) 반도체 산업 발전 전략의 12개 세부 과제

‘제1전략: 퍼스트무버형 혁신 기술·인재 확보’	
	1. 세계 최고 기술력에 도전
	1-1. 인공지능 반도체 플래그십 프로젝트
	1-2. 세계 최고 메모리 기반 신개념 반도체 개발
	2. 초기수요와 연계한 기술·사업화 장벽 해소
	2-1. 국가 인공지능·데이터 인프라를 통한 시범 도입확산
	2-2. 맞춤형 Tech Jump-up 프로그램 지원
	3. 인공지능 반도체 차세대 전문인재 양성
	3-1. 인공지능 반도체 고급인재 확보
	3-2. 실무·융합인력 양성 및 저변 확산
‘제2전략: 혁신성장형 산업 생태계 활성화’	
	1. 민간·공공분야 초기 시장창출
	1-1. 혁신기업 육성을 위한 수요·공급
	1-2. D.N.A 서비스 혁신과 연계한 공공분야 초기시장 창출
	2. 연대·협력의 밸류체인 구축
	2-1. 설계 역량 강화를 위한 연대·협력 생태계 조성
	2-2. 첨단 파운드리 소·부·장 경쟁력
	3. 기업 성장 인프라 강화
	3-1. 혁신기업 Scale-up을 위한 대규모 성장자금 확보
	3-2. 인공지능 반도체 팹리스 집중 지원체계 구축

자료: 관계부처 합동(2020. 10. 12)

다. 인공지능(AI) 반도체 산업 성장 지원 대책(2022. 6)

정부는 인공지능 반도체 기업과의 간담회에서 제기된 반도체 업계의 정책 수요

를 바탕으로 산·학·연 논의를 거쳐 지원 대책을 마련하였다.

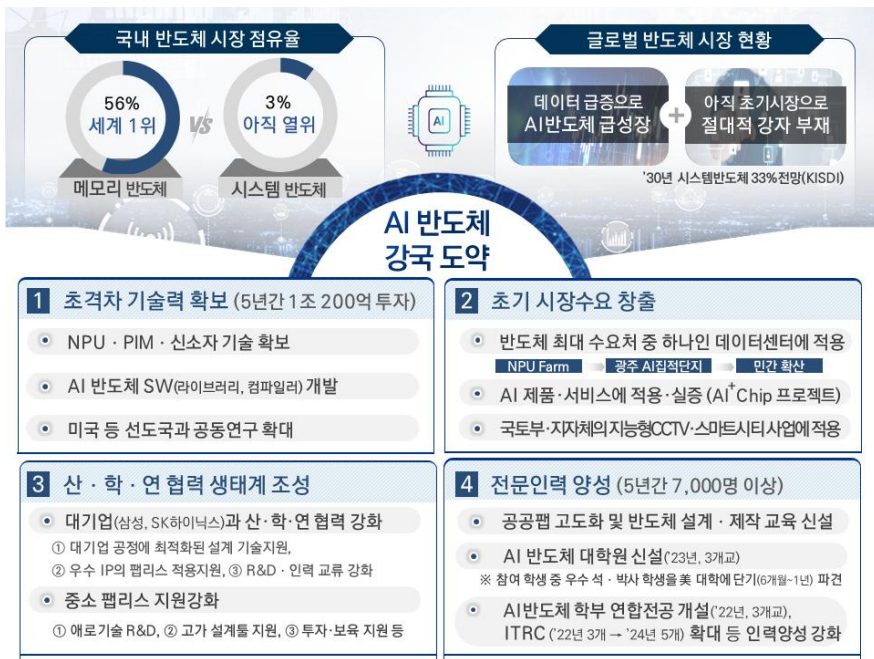
첫째, 인공지능 반도체 초격차 기술력을 확보하기 위하여, 인공지능 반도체 첨단 기술 연구개발(R&D)에 예비타당성 사업을 포함하여 향후 5년간 1조 200억 원을 투입하고 미국 등 선도국과 공동연구를 확대한다고 지원 대책을 발표했다. 주요 투자 분야로는 ① 신(新)소자와 설계기술을 융합한 차세대 NPU, ② 연산(프로세서)과 저장(메모리) 기능을 통합한 PIM 반도체(Processing In Memory), ③ 반도체 성능을 극대화하는 시스템소프트웨어, ④ NPU와 PIM의 장점을 결합하여 시스템 성능을 극대화하기 위한 초거대인공지능 시스템 등으로 제시했다.

둘째, 국산 인공지능 반도체 초기 시장수요를 창출한다. 반도체 최대 수요처 중 하나인 데이터 센터를 국산 인공지능 반도체로 구축하는 사업(NPU Farm 구축 및 실증)을 '23년 신설하고 인공지능 개발자에 컴퓨팅 파워를 무상 제공할 계획이다. 또한, 인공지능 제품·서비스 개발에 국산 인공지능 반도체를 활용하고 성능을 검증하는 '인공지능+ Chip 프로젝트'를 신규 추진한다. 아울러, 지능형 CCTV, 스마트시티 등 각 부처·지자체가 구축하는 공공사업에도 국산 칩이 적용·확산할 수 있도록 협의해 나갈 계획이다.

셋째, 대기업이 참여하는 산·학·연 협력 생태계를 조성한다. 대학·연구소가 첨단 상용 공정에 최적화된 반도체를 설계할 수 있도록 대기업과의 협력을 강화한다. PIM 반도체를 개발하는 정부 사업에 참여하는 연구기관에 대해 삼성전자와 SK하이닉스가 기술 자문을 제공하고, 성과가 우수한 연구 결과물의 반도체 생산 공정 적용을 검토할 예정이다. 또한 NPU를 개발하는 정부 사업의 연구 결과물 중 삼성전자 협력업체(디자인하우스)에서 검증하여 우수 설계기술(IP)로 평가된 경우, 삼성전자 파운드리 설계기술 데이터베이스(IP 풀)에 포함하고 다양한 팹리스 기업 제품에 적용될 수 있도록 지원하다, 특히, 정부 정보통신기술 연구개발 기획 과정에 삼성전자·SK하이닉스가 참여하여 유망기술에 대한 수요를 제기하고 기획 결과를 검증한다. 더불어 PIM 반도체 설계 연구 센터(PIM HUB)와 삼성전자·SK하이닉스 간 상호 인력파견 및 공동연구를 추진하는 등 연구개발·인력 교류도 지속 확대할 계획이다.

넷째, 인공지능 반도체 전문인력 7,000명을 양성한다. 인공지능 반도체 관련 다양한 학과(전기전자공학, 컴퓨터공학, 물리학 등)가 공동으로 교육과정을 구성·운영하는 ‘인공지능 반도체 연합전공(학부)’ 개설(3개교), 대학·연구소가 보유한 반도체 시험생산 설비의 고도화 및 이와 연계한 반도체 설계·제작 교육(학부생 대상) 신설 등을 추진한다. 특히, 연구 중심의 석·박사 고급인재 양성을 위해 ‘인공지능 반도체 대학원’을 ‘23년 신설(3개교)하고 참여 학생 중 우수 석·박사 학생을 해외 대학에 단기(6개월~1년) 파견하는 프로그램도 운영할 계획이다.

[그림 2-4] 인공지능(AI) 반도체 산업 성장 지원 대책 개요



자료: 과학기술정보통신부(2022. 6)

라. 국산 인공지능 반도체를 활용한 케이(K)-클라우드 추진 방안(2022. 12)

‘케이-클라우드’ 추진 방안은 세계 최고 수준의 초고속·저전력 국산 인공지능 반도체를 개발, 이를 데이터 센터에 적용해 국내 인터넷 기반 자원공유 경쟁력을

강화하고, 국민에게 향상된 인공지능 및 인터넷 기반 자원공유 서비스를 제공하기 위한 정책이다. 첫째, 국산 인공지능 반도체를 3단계에 걸쳐 고도화, 세계 최고 수준의 기술을 확보한다. 세계 최고 수준의 초고속·저전력 국산 인공지능 반도체 개발을 위해, 기존 인공지능 반도체 사업을 종합하고 체계화하여 '23년부터 '30년까지 총 8,262억 원(예정)을 투자하는 고도화 이행안을 마련했다.

〈표 2-8〉 국산 인공지능 반도체 고도화 3단계

단계	내용
1단계('23~'25)	상용화 초기 단계에 있는 국산 NPU를 지속적으로 고도화(추론→학습)하고 데이터 센터에 적용하여 성공 실적 확보 및 초기시장 창출을 지원
2단계('26~'28)	램 기반 상용 칩(PIM)과 국산 NPU를 접합(패키징)하여 세계적 수준의 연산 성능을 저전력으로 구현(DRAM 기반 PIM)
3단계('29~'30)	비휘발성 메모리(Non-Volatile Memory)를 활용, 아날로그 MAC 연산 기반의 NPU·PIM을 개발하여 극 저전력화를 달성

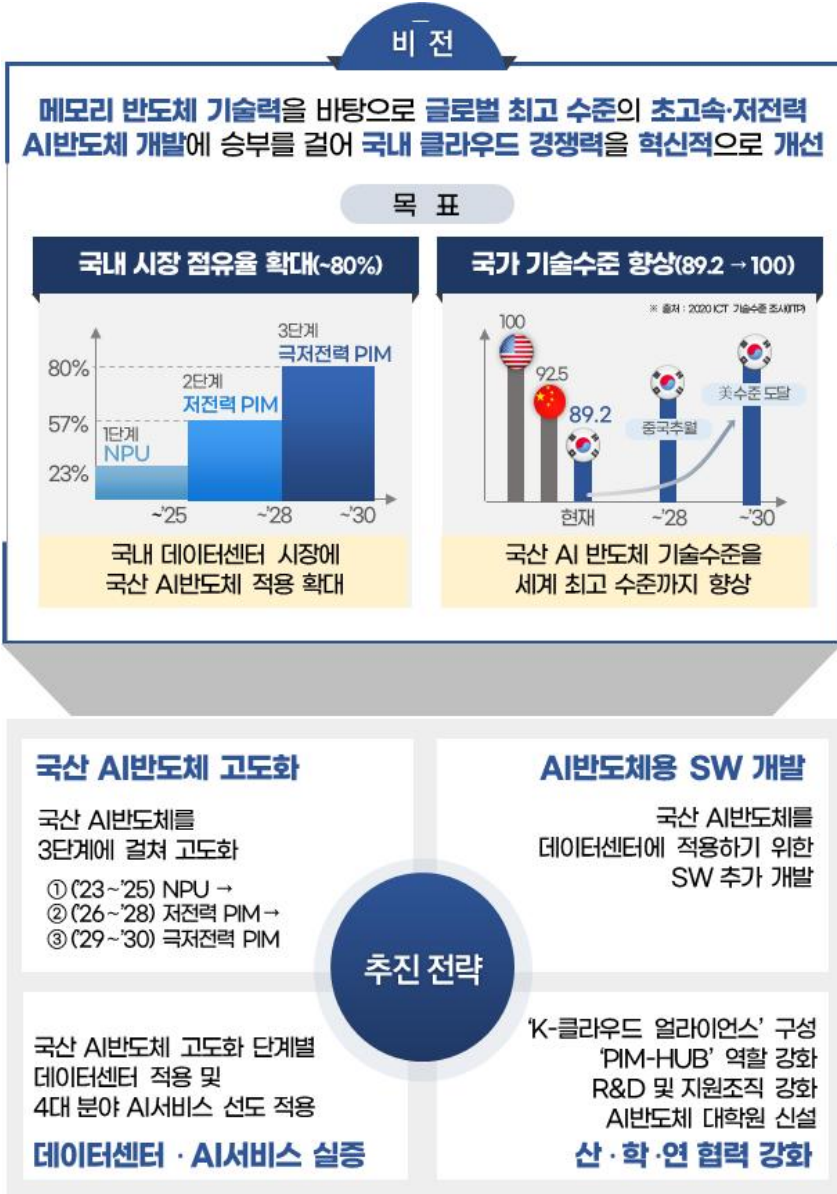
둘째, 국산 인공지능 반도체를 데이터 센터에 적용하기 위한 소프트웨어 예비타당성 사업을 추진한다. 인공지능 반도체 사업들은 프로세서(HW) 개발을 목표로 진행되고 있어, 엔피유·칩 등 단계별로 고도화한 국산 인공지능 반도체를 데이터 센터에 적용하기 위해서는 추가적인 소프트웨어 개발이 필요하다. 이에 신규 예비타당성 사업을 추진하여, 국산 인공지능 반도체에서 기저학습 등 인공지능 알고리즘을 초고속·극저전력으로 실행하는 컴파일러, 라이브러리, 인공지능모형 자동 병렬화 기술 등과, 이를 상용 인터넷 기반 자원공유에 적용하기 위한 가상머신(VM) 및 컨테이너, 가상 서버 클러스터 기술 등을 개발할 계획이다.

셋째, 단계별로 국산 인공지능 반도체의 데이터 센터 적용 및 인공지능·인터넷 기반 자원공유 서비스 제공을 추진한다. 인공지능 반도체 고도화 단계별로 이를 데이터 센터에 적용하고 인터넷 기반 자원공유 기반 인공지능서비스까지 제공하는 실증사업을 함께 추진한다. 2023부터 시작되는 1단계 실증사업으로는 국산 NPU 데이터 센터 구축 사업(신규)과 기존의 인공지능·인터넷 기반 자원공유 서

비스 개발 사업을 연계하여 '23년에 428억 원(안), '25년까지 3년간 약 1천억 원(잠정)을 투자한다. 실증사업에서의 인공지능·인터넷 기반 자원공유 서비스는 우선 사회·경제적 파급력과 수요가 높은 4개 분야(안전·보건·교육·국방 등)를 중심으로 추진하고, 이를 바탕으로 성장 가능성이 높은 주요 분야*까지 확대할 계획이다.

넷째, 산·학·연 협력 강화를 위한 추진체계를 마련한다. 케이-클라우드' 추진을 위한 민·관 협업 창구 마련 및 주요 과제 발굴을 위해, '국산 인공지능 반도체 기반의 케이-클라우드 얼라이언스'를 구성한다. 또한 인공지능 반도체 기업과 대학이 협력하여 교육과정을 개발하는 '인공지능 반도체 대학원'을 신설, 현장에서 요구되는 설계 역량을 갖춘 최고급 인재를 양성('23년~, 3개교 신설)한다. 그리고 2·3단계 펌(PIM) 고도화·실증을 위한 메모리 반도체 대기업과 산·학·연 간 기술 연계 등을 위해 '펌 허브(설계연구센터, 한국과기원)'의 역할을 강화하고, 관련 분야(인공지능 반도체·인터넷 기반 자원공유·인공지능 등) 협업 강화를 위해 연구개발 지원기관인 한국전자통신연구원(ETRI)과 정보통신기획평가원(IITP) 조직을 강화한다.

[그림 2-5] 케이-클라우드 추진 방안



자료: 과학기술정보통신부(2022. 12)

제 2 절 해외 반도체 관련 주요 정책 현황

1. 주요 기술 선도국의 반도체 관련 정책 개요(2021년~2023년)

주요 기술 선도국은 반도체 등의 첨단 산업에서 중국 의존도가 높아짐을 국가안보의 위협으로 인지하고, 자국 중심의 공급망 개편을 지속해 추진하고 있다. 아울러 디지털전환 등 대외환경 변화와 반도체에 관한 기술 안보 관점에서의 인식변화로 자국 중심의 공급망 확보와 기술 경쟁 우위 확보를 위한 지속적인 지원정책, 통상정책을 수립하여 추진하여 이를 강화 중이다.

미국, 중국, EU, 일본 등 주요 반도체 기술 선도국가는 반도체 육성 및 보호 관련 정책을 지속해서 발표하고 있으며, 반도체 가치사슬 변화 및 공급망 재편 과정에서 확고한 포지셔닝을 차지하기 위해 다양한 정책적인 대안들을 수립하고 있다.

반도체의 활용 분야의 빠른 확산으로, 기존 PC 부품에서 모바일, 엠티컴퓨팅, 모빌리티, 통신 네트워크, 가전 등으로 수요가 기하급수적으로 증가하고 있어, 반도체 분야는 기술 개발 및 무역·통상 측면에서 기술 안보적 전략자산 관점으로 급격한 인식변화가 이루어지고 있다. 따라서 국가적 차원의 핵심기술로 자리매김하고 있다.

하지만 이미 코로나19 팬데믹과 기후 변화 및 자연재해로 수요예측 변동성 및 공급망 불안 등으로 취약성이 높아졌다. 반도체 수요분야에 대한 모니터링과 더불어 반도체 공급망의 중요성이 증가함에 따라 주요 선도국가는 반도체 제조 기반 확보 및 초격차 기술 강화를 위한 다양한 정책을 마련하고 있다.

특히 주요 기술 선도국 사이에 반도체 관련 기술을 무기로 국가안보 문제를 해결하기 위해, 경쟁국에 대한 최첨단 반도체 기술 개발 및 제품 이전 가능성을 차단하는 동시에 동맹국의 참여를 끌어내는 등 반도체산업을 둘러싼 환경 변화가 진행되고 있다.

이에 따라 주요 기술 선진국은 기술주권 및 자국중심주의를 앞세워 자국 내 공급망 강화 및 확보와 더불어 반도체 관련 첨단기술을 선점하기 위한 정부 지원정

책, 통상정책을 강력하게 추진하고 있다.

2. 미국

글로벌 반도체 시장의 기술을 선도하며 기술 개발을 주도 중인 미국은 꾸준한 정부 및 민간투자와 경쟁국에 대한 지속적인 견제를 바탕으로 기술 지도력을 유지하고 있다. 최근 미국은 보조금 또는 세액공제 등을 통한 국내 R&D 및 제조시설 지원, 엄격한 수출관리를 통한 기술 통제, 동맹국을 중심으로 하는 글로벌 연대 등으로 반도체 정책을 강화하고 있다. 최근 반도체과학법안은 미국이 지난 반세기 수행한 산업정책 중에서 가장 큰 규모의 예산으로 진행되고 있다.

2022년 8월 바이든 대통령은 「반도체과학법(Chips and Science Act)」을 제정하여 자국 반도체 산업역량 강화를 위해 정부예산을 편성(반도체 기금)하여 투자를 지원하고 있다.²⁾

우선 반도체 R&D, 제조 및 인력양성에 약 520억 달러의 자금을 편성하였다. 이중 반도체 제조시설 지원에 390억 달러를 지원하는 등 인센티브(보조금)를 포함한다.³⁾

반도체과학법안에는 반도체 관련 투자기업에 25% 공제를 지원하는 내용도 포함되어 있으며, 이를 통해 국내 공급 확보 및 고속련 제조 일자리 창출 등 민간 투자 촉진을 유도하고 있다. 한편, 인센티브(보조금)에 대한 엄격한 보조금 신청요건과 더불어 2023년 상무부는 가드레일 세부 규정을 발표하여 이를 받은 기업이 우려 대상 국가(중국, 러시아, 이란, 북한 등)에 향후 10년간 반도체 시설에 대한 투자를 금지하는 가드레일 조항을 적용하는 등 주요 경쟁국 반도체 첨단기술 개발을 견제하고 있다.⁴⁾

이에 따르면 보조금 수혜기업은 중국 등 경쟁국 레거시(Legacy) 반도체 시설

2) KISTEP(2023), 2023 인공지능 반도체

3) KITA(2023), 미국과 EU의 반도체 산업 육성 전략과 시사점

4) 한국수출입은행(2023), 미중 반도체 전쟁에 따른 산업재편 및 영향

확장(material expansion)을 10년간 10% 미만 확장, 투자 10만 달러 미만(누적)으로 제한하였다. 또한 보조금 수혜기업은 우려국내 첨단반도체(레거시外 반도체) 시설 확장을 10년간 5%, 10만 달러 미만(누적)으로 제한하며, 보조금 수혜기업은 우려국 단체와 공동연구 및 기술 라이선싱이 제한된다.

2022년 8월 이후에는 첨단 AI·반도체 기술에 대한 동맹국 외 중국, 러시아 등의 경쟁국 접근을 차단하는 수출관리규정(통제정책)을 강화하였다. 2022년 8월, 상무부 산하 산업안보국은 반도체 관련 품목을 수출통제 리스트에 추가하였다. 2022년 10월에는 수출관리규정을 개정하여 중국의 고성능 첨단 반도체(AI 반도체, GPU, 슈퍼컴퓨팅) 구매 및 생산능력을 제한하기 위해 수출통제(반도체 장비 및 관련 소프트웨어·기술)를 강화하였으며, 아울러 2023년 10월 대중국 반도체 수출통제에 대한 확대보완 조치를 발표하였다.⁵⁾⁶⁾

이에 따르면 기존 수출통제의 주요 구성 요소였던 ① 첨단 반도체 제조 관련 품목과 ② 첨단 컴퓨팅 관련 반도체 제재가 모두 확대되었으며, 이와 함께 ③ Entity List에 13개 AI 반도체 중국기업이 추가되었다,

2023년 8월 바이든 행정부는 인공지능·반도체·양자기술 등 첨단 기술 분야의 중국 기업들에 대한 미국의 투자를 제한하는 투자제한 행정명령을 발표하였다.⁷⁾

이에 따라 2024년 1월부터 시행되는 행정명령에서 인공지능·반도체·양자기술 등 분야에서 미국 기업의 대중국 투자를 금지하였다.

3. 중국

중국은 국가 경쟁력 제고를 위한 국가차원 전략 정책에 반도체를 주요 중점분야로 지정하고 있다. 국가전략 정책인 중국제조 2025(2015년), 제13차 5개년 국가 과학기술 혁신 계획(2016년), 제13차 5개년 국가 전략적 신흥산업 발전계획(2016

5) KIEP(2023), 미국 반도체 수출통제 확대조치의 영향과 시사점
6) 한국수출입은행(2023), 미중 반도체 전쟁에 따른 산업재편 및 영향
7) KB금융지주경영연구소(2023), 미국의 對중 규제 강화와 한국 반도체 산업 영향

년), 제14차 5개년 계획(2021년) 등에서 전략육성 분야의 하나로 반도체를 선정하였다. 다양한 국가 지원정책(국가전략 수립, 펀드 운용, 세제지원 강화 등)을 통해 반도체 자급화를 지속하여 추진 중이나, 최근 미국 중심의 주요 기술 동맹국의 견제받고 있어 목표를 달성이 쉽지 않은 상황이다.

2021년 14차 5개년 계획·2035 중장기 목표의 7대 전략육성(인공지능, 양자정보기술, 반도체, 뇌과학, 유전자 바이오 기술, 임상의학, 심해·극지·우주 분야) 분야 중 하나로 반도체 분야를 선정하여 대대적인 지원을 추진 중이나 자립화율은 증가하지 못하고 있다.⁸⁾

이에 따르면 대규모 펀드와 각종 세제지원 등 전방위적인 자국 반도체 산업 육성정책을 추진 중이나, 여전히 자립화율은 10% 중반 수준으로 정체한 상황이다. 14년 5개년 계획은 중국제조 2025로 널리 알려진 13차 5개년 계획의 후속 국가전략으로 미국의 반도체 제재 분야(반도체 설계툴, 주요 소재·장비, 제조기술, 화합물 반도체 등)에 대한 전략적 지원·개발이 이루어질 것으로 전망된다. 반도체 기술에서는 반도체 설계 도구, 핵심 소재, 첨단제조기술, IGBT, MEMS, 첨단 메모리 기술, SiC·GaN 등 3세대 반도체 강조하고 있다.

2022년 10월, 미국의 제재 강화 이후 중국은 과학기술업무 컨트롤타워를 설치하고 187조 원 규모의 반도체산업 지원 패키지를 마련하였다.⁹⁾

이에 따르면 중국은 1조 위안(187조 원) 이상의 대규모 추가 보조금 지원을 발표했으며 지원 방향도 포괄적 대규모 지원에서 선별 기업에 대한 지원 확대로 변화하였다. 또한 미국의 기술 봉쇄에 대응하기 위해 신형 거국체제(新型舉國體制) 제시 및 공산당 산하에 과학기술 부문 정책을 주도할 중앙과학기술위원회를 신설하였다.

2022년 12월에는 미국의 수출 통제 강화 조치로 중국에 직접적인 피해가 발생하자, 중국은 미국의 대중국 수출 통제 및 관련 조치를 WTO에 제소하였다.¹⁰⁾

8) KISTEP(2023), 2023 인공지능 반도체

9) 한국수출입은행(2023), 미중 반도체 전쟁에 따른 산업재편 및 영향

10) KITA(2023), 미국과 EU의 반도체 산업 육성 전략과 시사점

이에 따르면 중국 상무부는 미국의 반도체 관련 통제 조치가 글로벌 공급망의 안정성을 위협하고 세계 경제에 혼란을 일으키고 있다고 주장하며 WTO에 분쟁 협의를 요청하였다.

2023년 6월에는 최첨단 반도체 장비의 수출통제와 관련해 미국·일본·네덜란드를 3국가를 조사해 줄 것을 세계무역기구(WTO)에 촉구하였다. 미국·일본·네덜란드의 대중국 반도체 수출 통제·규제가 세계무역기구의 형평성과 투명성 원칙을 위반하였다고 조사와 감시강화를 촉구했다.

2023년 8월에는 미국의 대중국 수출통제·규제에 맞서 갈륨(Ga) 및 저마늄(Ge) 등 소재의 수출통제·규제를 시행하며 원자재 무기화하고 있다.¹¹⁾

이에 따르면 중국은 전 세계 갈륨 및 저마늄 생산의 약 90% 이상을 차지하는 최대 생산국으로 수출통제 시 원자재 가격 상승으로 이어지며 반도체·디스플레이의 제조비용 상승이 우려되고 있다.

4. EU

2022년 EU는 반도체 산업의 육성을 통해 현재 10% 수준의 세계시장 점유율을 2030년까지 20%를 달성을 목표로, 민간·공공협력을 통해 총 430억 유로를 투자하는 유럽반도체법안(European Chips Act)에 제안하였다.¹²⁾

이에 따르면 유럽반도체법안은 반도체 제조시설에 대한 투자, EU 전지역의 반도체 허브 구축, 새로운 반도체 기술 시범 운영, 설계 및 기술 역량 개선, 기업의 자금 조달 및 자본에 대한 접근성 향상 등의 내용을 포함하고 있다. 주로 반도체 기술개발 역량 강화와 EU의 생산능력 확보, 공급망 모니터링 및 위기관리 체계 등의 도입을 주요 내용으로 삼고 있다.

유럽반도체법안은 공급망 위기에 대응하기 위한 권고(Recommendation: EU와 회원국 간 반도체 시장 상황 모니터링을 통한 조기경보 및 정보공유 체계 구

11) KB금융지주경영연구소(2023), 미국의 對중 규제 강화와 한국 반도체 산업 영향

12) KITA(2023), 미국과 EU의 반도체 산업 육성 전략과 시사점

축)와 첨단반도체 기술개발·생산확대를 주요 골자로 하는 중장기계획인 규정(Regulation: 연구·혁신, 설계·생산, 공급안정, 위기관리, 국제협력 등 주요 부문별 정책 방향 제안)으로 구성된다.¹³⁾

이처럼 반도체 공급난 대응과 EU 역내 생산역량 강화를 위해 유럽반도체법의 입법 활동을 전개하고 있다.

그리고 유럽반도체법안의 차질 없는 이행을 위해 반도체 이니셔티브(Chips for Europe Initiative)를 발족하고, 유럽 내 집적기술·설계 능력, 접근성 높은 파일럿 라인, 설계·생산·장비 기업 간 협력체계 등 확보를 통해 산업생태계 강화를 추진하고 있다.¹⁴⁾

이외에도 유럽공통관심 주요 프로젝트(Important Project of Common European Interest, IPCEI)를 통해 AI 프로세서, 엣지컴퓨팅, 모빌리티, 5G·6G 등 기술개발 프로젝트 지원하여 반도체 산업에 보조금을 지원한다.

5. 일본¹⁵⁾

2021년 6월 일본 경제산업성은 「반도체·디지털 산업 전략」제정을 통해 반도체 산업 부흥을 위한 단계별 로드맵을 설정했으며, 2023년 4월 「반도체·디지털 산업 전략(개정안)」발표로 로드맵을 구체화하고 있다. 지난 20여 년간 쇠퇴한 일본 반도체 산업의 재부흥과 디지털전환 등 대외환경 변화에 대응하기 위한 전략이다.

반도체를 디지털 인프라, 디지털 산업과 함께 3가지 핵심기술 요소로 설정하였으며, 국내 반도체 생산 기반을 강화하고(1단계), 차세대 반도체 설계기술을 확보하며(2단계), 나아가 미래 기술의 연구개발을 추진(3단계)하는 것이 목표이다.

1단계의 경우, 대만 및 미국과 연계하여, 일본 국내 반도체 생산 기반 강화하는 것이 목표로, 대만 TSMC와 일본 덴소, 소니 반도체 솔루션 합작회사인 JASM

13) KISTEP(2023), 2023 인공지능 반도체

14) KOTRA(2003), 유럽반도체법(European Chips Act) 관련 입법 동향

15) 「KOTRA(2003), 일본의 신 반도체 산업 전략과 글로벌 공급망 구축」 재구성

(Japan Advanced Semiconductor Manufacturing)는 일본의 반도체 산업전략 1단계의 핵심 프로젝트이다. JASM의 쿠마모토현 제1공장은 소니 그룹의 이미지 센서 사업에 들어갈 반도체 28nm 세대가 중심이고, 제2공장은 12nm 세대의 양산을 통해 자동차의 화상처리용 IC 등 용도에 대응할 것으로 예상된다.

2단계의 경우, 대만 및 벨기에와 연계하여, 차세대 반도체 설계기술을 개발하는 것이 목표이다. 차세대 반도체(Beyond 2nm) 설계 기술 개발을 위해 개방형 연구개발 거점 구축하고, 향후 차세대 반도체를 양산할 수 있는 거점 구축을 위해 반도체연구기관 LSTC와 Rapidus社 설립을 지원한다. Rapidus사는 도요타자동차, 소니, 키옥시아 등 일본 주요기업 8개사 출자로 출범하였다, 2027년까지 차세대 로직 반도체 개발을 목표로 하여 10년 간 5조 엔의 설비투자가 계획 중이며, 일본 정부는 700억 엔 지원을 결정하였다. Rapidus사는 미국 IBM과 2nm 노드 반도체의 공동 개발 파트너십을 체결했으며, 유럽 최고 수준의 반도체 연구개발 생태계를 형성하고 있는 벨기에 imec과 차세대 반도체 개발에 관한 MOC(협력각서)를 체결하였다.

3단계의 경우, 다양한 국가와 연계하여, 차세대 반도체 설계기술을 개발하는 것이 목표이다. 반도체·디지털 산업전략 3단계는 광전 융합이나 양자 등 기술의 연구개발을 목표로 하며, 2단계 차세대 로직 반도체 설계 이후의 단계이다. 2030년 이후 본격적으로 구현될 전망인 양자컴퓨터 등 새로운 사례 적용을 염두에 두고 장기적 연구개발을 진행하는 것이 3단계 전략으로 이를 위해 미국, EU, 벨기에, 네덜란드, 영국, 한국, 대만 등과 차세대 반도체 실용화 사례를 창출하고, 공동연구를 추진할 계획이다.

제 3 장 미국 등록 특허를 활용한 기술경쟁력 분석

제 1 절 개 요

1. 특허 기술 분류 및 데이터 수집

반도체 분야의 미국 등록 특허 분석을 위한 기술분류체계¹⁶⁾는 특허청 및 ‘한국 특허전략개발원’에서 발간한 「반도체 산업, 특허 메가트렌드 분석보고서 2021」의 기술 분류 기준을 따라 FAB 공정, 소재, 설계, 장비, 차세대반도체 등 5대 기술 분류로 구분하고 있다. 본 연구는 한국특허기술진흥원의 DB에서 2003년부터 2022년까지 각 분류체계에 해당하는 특허 정보를 추출하여 분석하였다.

2. 분석 지표

수집된 특허 정보는 특허권자를 기준으로 기술 분류별, 국가별로 구분하고 현시 기술우위지수, 특허당 피용수, 특허영향지수, 기술력 지수, 기술순환주기, 과학연계지수, 패밀리특허 분석, K-PEG 평가를 산출하여 경쟁력 분석을 하였다.

가. 현시기술우위 지수¹⁷⁾(Revealed Technological Advantage: RTA)

RTA 지수는 전체 특허에서 반도체 분야 특허의 상대적 비중을 기준으로 산출한 값이며, 해당 국가의 반도체 특허 정도를 나타내는 것으로 이해할 수 있다. 산출식은 아래와 같다.

16) 자세한 내용은 부록 표 참고

17) RTA 지수는 일반적으로는 특허활동지수(AI, Activity Index)로 알려져 있으며, 특화지수(Specialization Index), 기술 비교우위 지수(TCA, Technological Comparative Advantage), 기술 현시비교우위 지수(TRCA, Technological Revealed Comparative Advantage) 등 다양한 이름으로 사용된다.

$$RTA = \frac{(P_{ij} / \sum_i P_{ij})}{(\sum_j P_{ij} / \sum_i \sum_j P_{ij})}$$

국가 j의 특허에서 i 분야가 차지하는 비율을 의미하며, 분모는 전 분야의 특허에서 i 분야가 차지하는 비율을 의미한다. RTA 지수는 0부터 무한대까지의 범위에 분포하게 되며, 예를 들어, RTA가 클수록 대상 국가(또는 기업)가 해당 기술 분야에 특허된 정도가 높다는 것을 나타내며, 1보다 작으면 상대적으로 특허 수준이 낮고, 1이면 평균적인 수준으로 이해할 수 있다.

나. 특허당 피인용 수(Cite Per Patent; CPP)

CPP는 특정 연도(또는 기간)에 등록된 특허들이 이후 등록된 특허들에 의해 평균적으로 인용된 횟수로, 분석 대상(국가, 기업 등)의 특허가 이후의 기술 혁신 활동에 어느 정도의 영향을 미쳤는가를 보여주는 지표이다. 산출식은 아래와 같다.

$$CPP_t = \frac{\sum_{i=1}^{n_t} C_i}{N_t}$$

N_t 는 t연도에 등록된 특허 건수이며, C_i 는 특허 i의 피인용 건수를 나타낸다.

다. 특허 영향지수(Patent Impact Index; PII)

PII는 전체 특허 대비 특정 특허의 상대적 피인용 수를 나타낸 지표로, PII가 높을수록 특허의 질적 수준이 높은 것으로 이해할 수 있다. 상대적인 지표이므로 1인 경우 모든 특허의 평균적인 수준으로 인용이 된 것이며, 1보다 클 때는 다른 특허에 비해 상대적으로 인용이 많이 된 것이다. 산출식은 아래와 같다.

$$PII_t^a = \frac{CPP_t^a}{CPP_t} = \frac{\frac{C_t^a}{N_t^a}}{\frac{C_t}{N_t}}$$

여기서, C_t^a 는 t 시점에 a 가 보유한 특허의 피인용 수, N_t^a 는 a 가 보유한 특허 수, C_t 는 전체 특허의 피인용 수, N_t 는 전체 특허 수를 의미한다. 이 지수는 특정 기술 분야 내에서의 상대적인 피인용 값으로, 단순 평균으로 산출되는 CPP와 달리 해당 분야 경쟁 국가(혹은 기업)의 기술 수준이 고려된 “상대적인” 가치 정보를 나타낸다.

라. 기술력 지수(Technology Strength, ST)

ST는 PPI에 특허 등록 수를 곱하여 산출한 값으로 양적지표와 질적 지표를 모두 고려하여 해당 국가(혹은 기업)의 기술 역량을 나타내며 아래와 같이 산출할 수 있다.

$$TS_t^a = PPI_t^a \times N_t^a$$

PPI_t^a 는 t 시점의 a 의 특허 영향지수이며, N_t^a 는 a 의 등록 특허 건수이다.

마. 기술순환주기(Technology Cycle Time; TCT)

기술수명주기 지수 등으로도 알려진 TCT 지수는 인용된 특허들의 발행 연도와 해당 특허의 발행 연도와의 차이값의 중간값으로 산출하며, 값이 클수록 기술순환 주기가 긴 것으로 해석할 수 있다. 따라서, 등록된 특허가 더 긴 시간 동안 영향을 행사할 수 있음을 의미한다.

바. 과학연계지수(Science Linkage; SL)

SL은 해당 특허들이 인용한 과학기술 논문 개수의 평균값으로 산출하며, 값이 커질수록 기초 연구 또는 원천기술의 개발에 주력하고 있다고 해석할 수 있다.

$$SL_t = \frac{1}{N_t} \sum_{i=1}^{N_t} S_t^i$$

위에서 N_t 는 전체 등록 특허 건수이며, S_t^i 는 특허 i 의 과학기술 논문 인용 수이다.

사. 패밀리 특허(Patent Families)

삼극특허(Triadic Patent Families)는 같은 발명을 보호하기 위해 USPTO(미국특허청), EPO(유럽특허청), JPO(일본특허청)에 동시에 출원 및 등록된 특허이다. 국가별 홈어드밴티지(Home Advantage)의 영향을 최소화하여 국가별 특허의 질적 비교를 위해 OECD가 개발한 지표이다. IP5 Patent Families는 삼극특허의 특허청 이외에 KIPO(한국특허청), SIPO(중국특허청)에 동시에 출원 및 등록된 특허를 말한다.

아. K-PEG

K-PEG 평가시스템은 한국 특허 기술진흥원이 특허 평가모델로, 종합평가 등급은 해당 기간의 특허가 표준 정규분포를 따르게 정규화한 후 S~C2까지 9단계로 구분하고, 평균값 또는 중앙값에 상응하는 등급이 'B1' 등급으로 부여한다.¹⁸⁾ 지수를 산출하는 데 사용된 평가 요소는 목적별로 아래 3가지 형태로 구분할 수 있으며, 아래와 같이 정리할 수 있다.

〈표 3-1〉 K-PEG 평가 항목

평가지수	평가항목
권리성 지수	권리 범위의 광협, 권리의 견고성, 권리의 영향력, 등록 소요 기간, 권리의 연속성, 권리의 연동성
기술성 지수	기술 융합도, 기술의 영향력, 기술의 범용성, 기술의 협력도, 혁신 속도, 기술의 신규성, 과학 연계성
상업성 지수	지역 확장성, 실현 기여도, 특허의 포트폴리오, 특허권의 활동성, 국제협력도(권리)

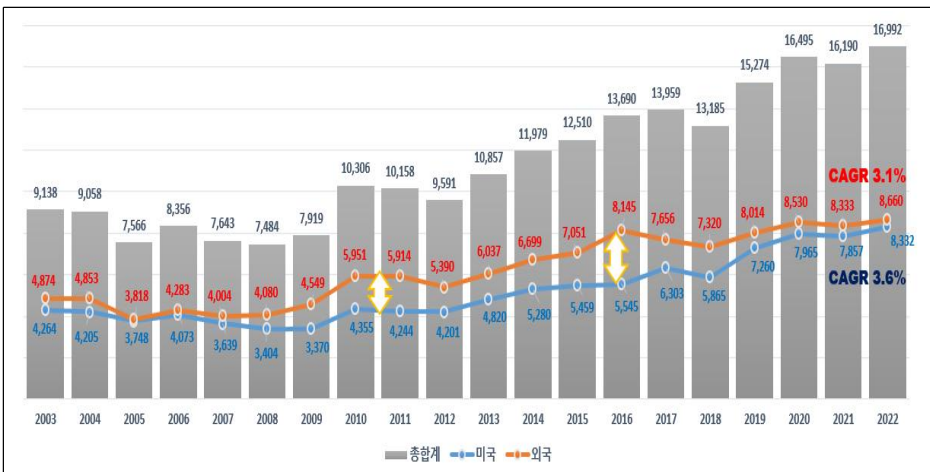
18) 생존확률이 높은 특허를 가치 있는 특허로 규정하고, 이와 관련된 평가 요소를 도출하고, 통계적인 방식으로 기술 분야별 평가모델을 설계하여, 유사 특허 군 내에서의 상대적인 평가가 가능하도록 설계.

3. 국가별 반도체 특허 동향

가. 연도별 동향

지난 20년간 반도체 분야의 미국 특허 등록 추이를 살펴보면, 2012년 이후 꾸준히 증가하고 있으며, 특허권자가 미국인 경우보다 외국인 경우가 더 많다는 점을 확인할 수 있다. 그러나 2003년부터 2022년까지 미국 특허권자의 등록 특허 증가율은 3.6%로 외국 특허권자의 증가율보다 0.5%p 더 높아 그 격차가 좁아지고 있다. 구체적으로는, 2007년 이후 등록 특허 수의 격차가 벌어지기 시작하여 2016년에 최대치를 기록하고, 점차 감소하여 2022년에는 몇백 건의 차이로 줄어들었다.

[그림 3-1] 반도체 분야 미국 등록 특허 추이



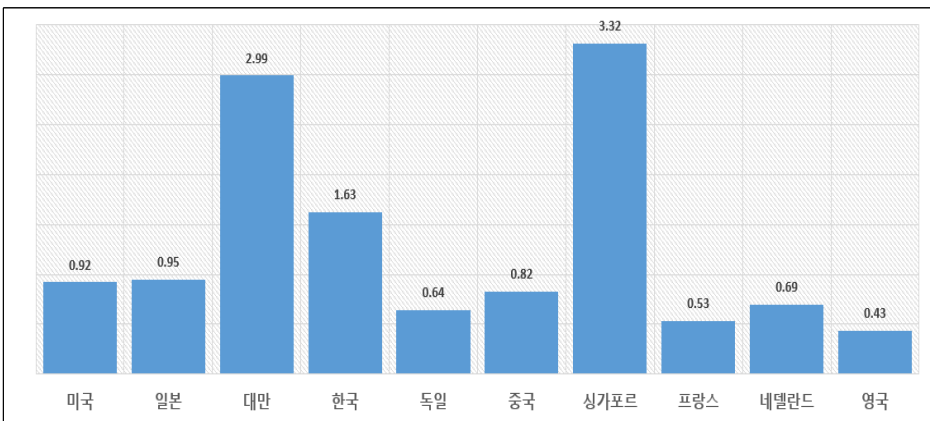
자료: 한국특허기술진흥원

주요 국가별로 살펴보면, 같은 기간 누적 특허 수로는 미국(104,189), 일본(39,483), 대만(24,418), 한국(21,954), 독일(7,326), 중국(6,062) 순으로 많았다. 그러나 증가율로 살펴보면 대만은 200% 이상 증가, 한국은 150% 이상 증가, 중국은 138배나 증가하였으나, 일본은 오히려 약 24% 감소하였다.

나. 국가별 RTA

지난 20년간 국가별 반도체 기술 집중도를 살펴보면, 싱가포르가 3.32로 가장 높고, 대만이 2.99, 한국이 1.63 순으로 나타났고, 모두 1보다 커, 상대적으로 각 국가에서 반도체 분야가 전략산업인 것으로 이해할 수 있다. 그러나, 미국과 일본은 각각 0.92와 0.95로, 양국에서 반도체 분야는 타 분야에 비해 전략적 우선순위에 떨어진다고 볼 수 있다. 눈여겨 볼만 한 사항은 미국의 반도체 분야 기술 집중도가 2016년 이후 지속 증가하여 2022년에는 1.02로 1을 넘어서 전략산업으로 전환하고 있다는 점이다. 또한, 한국과 대만은 2022년 기준으로 각각 1.48과 3.65로 반도체 기술의 중요도가 높지만, 추이를 보면 2003년 이후 한국은 집중도가 완화되는 반면 대만은 더욱 강화되고 있다는 점을 주목할 필요가 있다.

[그림 3-2] 국가별 기술 집중도 지수(RTA)

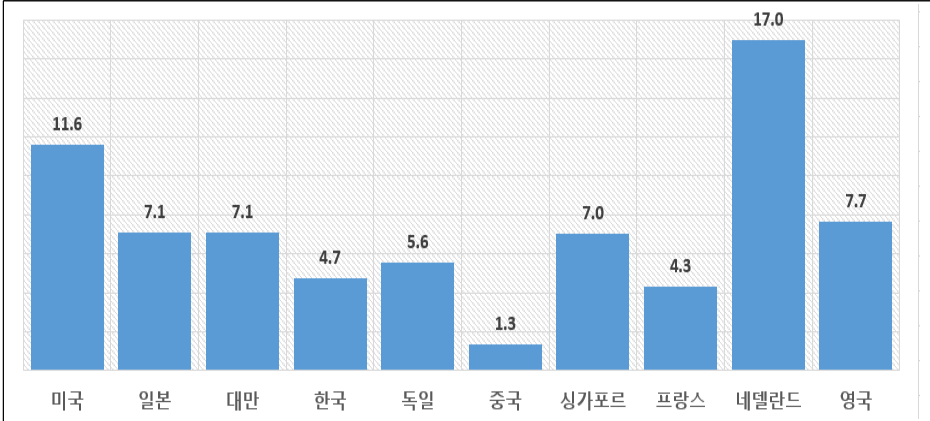


자료: 한국특허기술진흥원

다. 국가별 피인용지수

국가별 평균 피인용지수를 살펴보면, 한국은 4.7로 비교 대상 10개국 중 8위를 차지하고, 가장 높은 나라는 네덜란드로 17.0, 그다음으로 미국이 11.6, 영국이 7.7, 일본과 대만이 7.1로 집계되었다.

[그림 3-3] 국가별 피인용지수(CPP)



자료: 한국특허기술진흥원

라. 국가별 과학연계지수(SL)

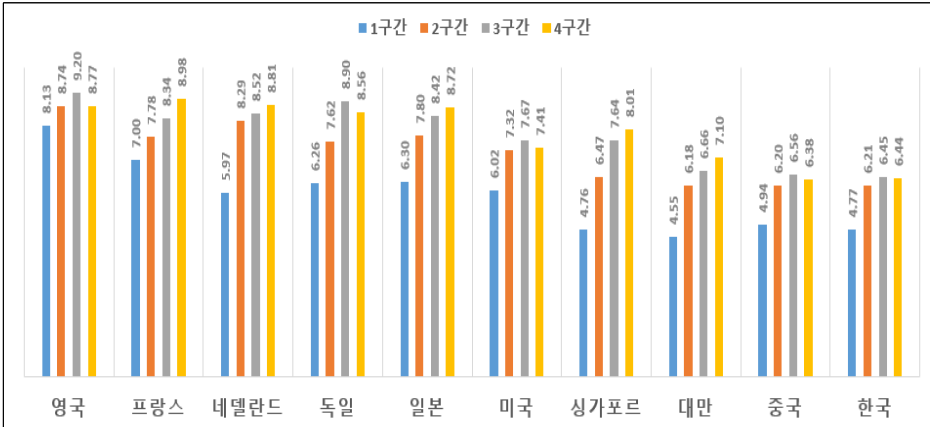
과학연계지수는 네델란드¹⁹⁾가 70.89로 가장 높고, 그다음으로 영국(12.42), 미국(11.46) 순으로 나타났다. 위의 3개국은 연구개발 활동 시 기초과학과의 밀접한 관계를 맺고 있다는 것을 의미한다. 반면, 반도체 분야 미국 등록 특허 건수가 많은 일본, 대만, 한국, 중국, 싱가포르 등은 과학연계지수가 낮아 대부분 국가가 응용 기술 중심의 연구개발 투자가 이루어지고 있음을 알 수 있었다.

마. 국가별 기술순환주기(TCT)

반도체 분야 미국 등록 특허의 기술순환주기는 영국, 프랑스, 네델란드, 독일 등 유럽 국가들에서 길고, 대만, 싱가포르, 중국, 한국 등 아시아 국가에서는 짧게 나타났다. 특히 한국의 TCT는 6.11년으로 주요 10개국 중에서 가장 짧았다. 추이를 보면, 대체로 2003년 대비 2022년에 TCT가 2~3년 정도 길어지고 있다.

19) 네델란드의 CPP와 SL이 높은 이유는 최근 5년 간 반도체 웨이퍼 처리장비의 설계, 제조, 판매 및 서비스를 전문으로 하는 다국적 기업인 ASM사의 CPP(61.8)와 SL(1020.9)의 영향으로 보이며, ASM을 제외한 나머지 특허권자만을 보면 미국보다 낮은 수준으로 나타남

[그림 3-4] 국가별 / 구간별 기술순환주기



자료: 한국특허기술진흥원

바. 국가별 기술 영향력 지수(PII)와 기술력 지수(TS)

주요 10개국 중에서 기술영향력지수가 1보다 큰 나라는 네덜란드와 미국뿐이며, 한국은 기술영향력지수가 0.5로 프랑스와 함께 여덟 번째로 상대적으로 기술 영향력이 낮은 국가로 분류되었다. PII에 특허 등록 건수를 곱하여 산출하는 TS는 미국, 일본, 대만, 한국, 독일, 중국 순으로 나타났다. 특히, 미국은 최근 10년간 등록 특허 수가 빠르게 증가하면서 기술력 지수도 같이 크게 상승하였다. 대만도 같은 기간 특허가 이전 기간에 비해서 월등히 많아지면서 기술력 지수가 많이 상승했다.

사. 국가별 K-PEG 등급

S등급의 특허 수는 미국이 5,641건으로 압도적으로 많은 특허를 보유하고 있는 것으로 나타났다. 한국은 398건의 S등급 특허를 보유하고 있어 미국의 1/14 수준이다.

〈표 3-2〉 국가별 K-PEG 등급(%) 분포

전체	S	A1	A2	A3	B1	B2	B3	C1	C2
미국	7.8	10.2	15.5	16.6	21.2	11.8	7.4	5.4	4.1
일본	3.9	7.7	7.2	12.6	25.0	17.6	12.2	8.8	4.9
한국	2.5	7.2	22.7	16.1	22.3	14.3	7.8	5.4	1.7
대만	1.6	5.0	5.3	32.7	34.2	10.2	6.0	3.7	1.2
독일	4.0	7.5	6.8	21.9	23.2	13.8	10.4	6.9	5.5
네덜란드	8.1	9.0	6.8	9.5	24.8	15.4	10.3	7.8	8.4
싱가포르	3.2	6.7	4.2	9.4	29.8	28.2	9.9	5.6	2.8
중국	1.4	2.5	2.1	4.7	25.1	25.7	21.4	12.0	5.0
프랑스	3.6	4.9	5.7	8.1	24.9	19.8	15.1	9.9	8.1
영국	4.3	8.1	7.5	11.4	21.9	17.9	13.0	8.5	7.6
표준	5.0	7.5	10.0	15.0	25.0	15.0	10.0	7.5	5.0

자료: 한국특허기술진흥원

각국의 평가대상 특허 대비 비중으로 구분하여 살펴보면, S등급의 경우 네덜란드(8.1%)와 미국(7.8%)이 정규분포 5%에 해당하는 S등급 특허를 상대적으로 더 많이 보유하고 있고, 나머지 국가들은 상대적으로 비중이 낮은 것으로 나타났다. A1 등급에서는 미국(10.2%)과 네덜란드(9.0%), 영국(8.1%), 일본(7.7%)이 표준인 7.5%보다 높은 비중의 A1 등급 특허를 보유하고 있다. 한국의 S등급 특허 비중은 2.5%, A1 등급은 7.2%로 최우수 등급의 특허 비중은 미국이나 다른 주요국에 비해 저조한 것으로 나타났다. 그러나 A3 등급 이상까지 한국의 특허 비중이 대만은 앞서는 것으로 나타났다.

4. 반도체 기술 분야별 특허 등록 추이

지난 20년간 반도체 기술 분야별로 특허를 구분하여 살펴보면, 전체적으로는 차세대반도체가 71,858건(31.5%)으로 가장 큰 비중을 차지했고, 다음으로는 FAB 공정이 55,950건(24.5%)으로 많았다. 소재와 장비 분야 특허는 점차 그 비중이 감소하여 최근 5년에는 비중이 각각 11.3%와 15.4%에 그쳤으며, 차세대반

도체는 점차 증가하다 최근 5년간 큰 폭으로 증가하여 47.8%까지 올랐다.

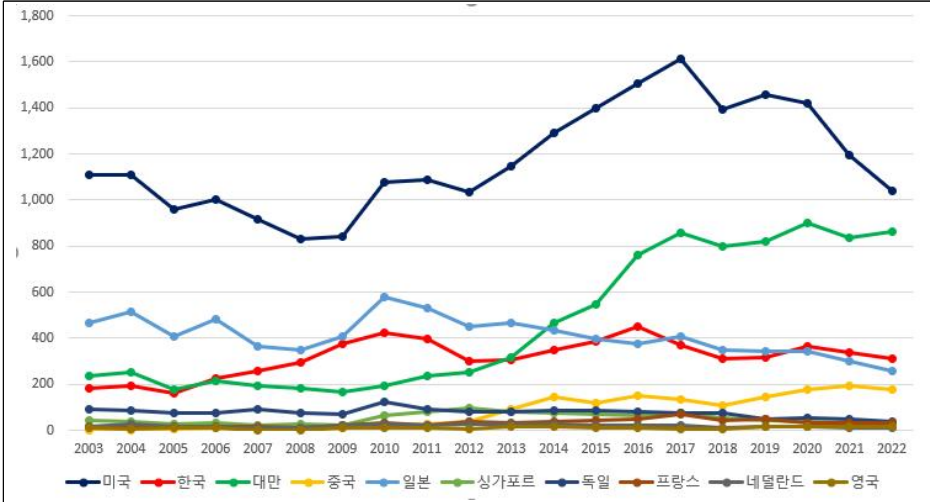
〈표 3-3〉 기술 분류별 미국 등록 특허 건수 추이

구분	2003-2007	2008-2012	2013-2017	2018-2022	총합계
FAB	10,738	11,742	17,056	16,414	55,950
설계	3,573	4,236	4,524	3,531	15,864
소재	8,671	8,265	10,747	8,792	36,475
장비	11,155	11,129	13,856	12,063	48,203
차세대반도체	7,624	10,086	16,812	37,336	71,858
총합계	41,761	45,458	62,995	78,136	228,350

가. FAB 공정 관련 특허

지난 20년간 등록된 FAB 공정 관련 특허의 44.6%가 특허권자가 미국이며 절반 가까이 차지하고 있으며, 대만, 일본, 한국의 특허는 각각 17.6%, 15.6%, 12%를 차지하고 있다. 나머지 국가는 5% 미만의 점유율을 보인다. 2017년 이후 모든 주요 국가에서의 FAB 공정 관련 특허 등록이 줄어들거나 성장세가 둔화된 가운데 대만과 중국이 더디지만, 계속 증가하고 있다. 2022년 순위로는 미국, 대만, 한국, 일본, 중국 순이다. 대만은 2009년부터 본격적으로 증가하기 시작하여 2020년 최고치(903건)를 기록하였고, 2022년 미국과 175건 차이로 그 격차를 좁혔다. 한국은 2016년 450건으로 최고를 기록한 후 서서히 감소하여 2022년에는 310건으로 대만의 36%, 미국의 30% 수준에 머무르고 있다. 표본기간 누적 특허 수가 두 번째로 많은 일본은 2010년 579건으로 최고치를 기록한 후 감소를 지속하여 2022년 257건으로 한국보다도 적은 특허를 등록하였다.

[그림 3-5] 국가별 FAB 공정 관련 등록 특허 추이

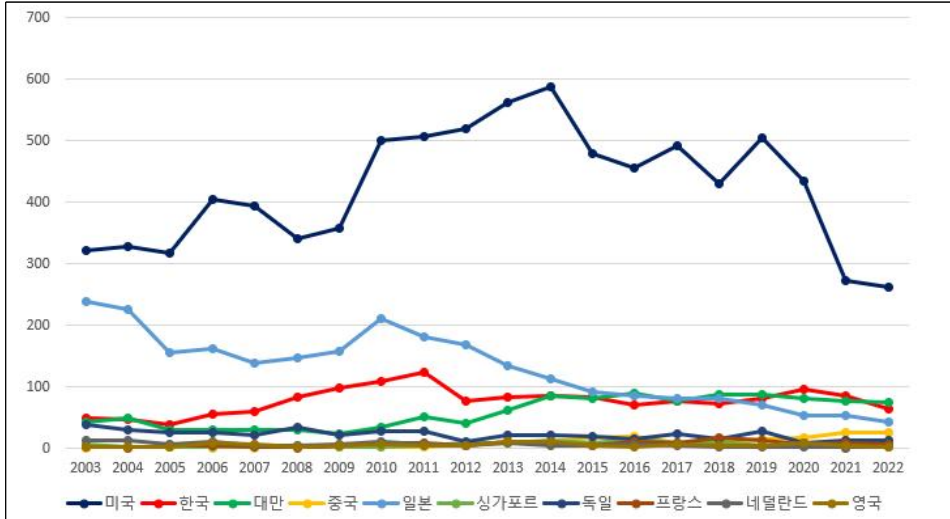


자료: 한국특허기술진흥원

나. 설계 관련 특허

미국은 반도체 관련 모든 기술 분야에서 특허가 가장 많지만, 각 기술 중에서도 설계 관련 특허의 비중이 가장 높고, 표본 기간 전체 점유율이 56.9%에 이른다. 그럼에도 불구하고 2014년(588건)을 기점을 빠르게 감소하여 2022년에는 그 절반에도 미치지 못하는 261건에 이르렀다. 설계 분야에서 두 번째로 등록 특허가 많은 국가는 일본으로 표본 기간 17.5%를 차지하고 있다. 일본은 미국보다도 더 빠른 2010년부터 감소하기 시작하여 2022년에는 대만(2위)과 한국(3위)에도 밀려 4위로 밀려났다. 2014년 이후 대만과 한국은 거의 비슷한 수준에서 등락을 반복하면서 추세적으로는 정체되어 있다. 2022년 기준으로 대만은 75건으로 한국보다 10건이 더 많다. 중국은 특허 수로는 그 비중이 1.2% 수준으로 미약하지만, 2017년 9건에서 2022년 26건으로 빠르게 증가하여, 주요 10개국 중 유일하게 증가하는 추세를 가진 국가이다.

[그림 3-6] 국가별 설계 관련 등록 특허 추이

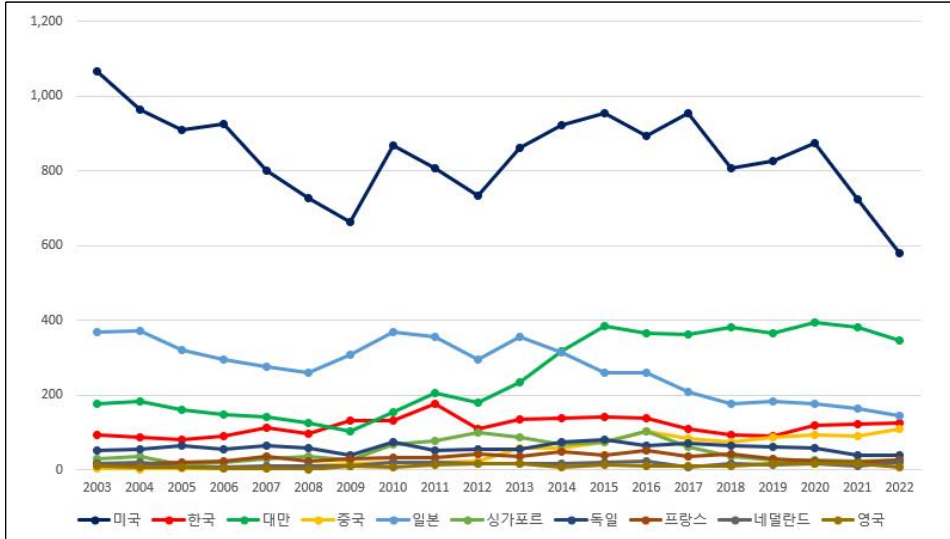


자료: 한국특허기술진흥원

다. 소재 관련 특허

소재 분야에서도 미국은 표본 기간 등록 특허의 49.7%를 차지하며 가장 특허 비중이 높은 국가이고, 다음으로 일본(16%), 대만(15%)이 뒤따르고 있다. 한국(6.8%)을 제외한 나머지 국가는 모두 5% 이하의 작은 비중을 차지하고 있다. 추세적으로는 대만과 중국이 빠른 상승세를 보이고 있는데, 대만은 2009년 103건에서 2022년 345건으로 미국의 60% 수준까지 따라왔고, 중국은 2011년 18건에서 2022년 110건으로 5배 이상 증가했다. 일본은 2014년 대만에 추월당하기 전까지 미국과는 상당한 차이가 있지만 다른 국가에 비해서는 압도적인 2위를 지키고 있었으나, 추세적으로 하락을 지속하다 2022년 143건으로 기록하면서 4위인 한국과 19건의 차이로 좁혀졌다. 한국은 2011년 176건으로 최고치를 기록한 후 등락을 거듭하다가 2019년 이후 3년 연속 증가하고 있다. 유럽에서는 독일이 표본 기간 3.5%로 가장 많은 특허를 등록하고 있으나, 2015년 이후 하락을 지속하고 있다.

[그림 3-7] 국가별 소재 관련 등록 특허 추이

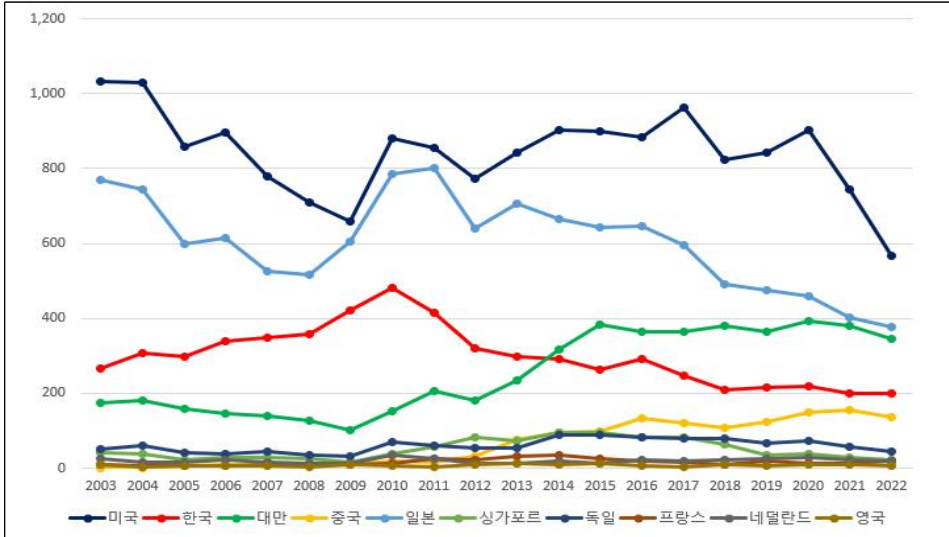


자료: 한국특허기술진흥원

라. 장비 관련 특허

반도체 장비 관련 특허는 절대적으로는 미국이 38%로 가장 많이 차지하고 있으나, 상대적으로는 한국(13.5%)과 일본(27.2%)에서 큰 비중을 차지하는 분야이다. 그러나 한국과 일본 모두 2010년을 기점으로 미국 등록 특허 수가 감소하고 있으며, 감소 속도는 일본이 더 빠르다. 2022년 기준 한국과 일본은 각각 4위와 2위를 기록했다. 소재 부문과 마찬가지로 다른 주요국과 달리 대만과 중국의 상승세가 뚜렷하게 나타나고 있다. 대만은 2009년 103건에서 2014년에 한국을 추월하고, 2018년 380건까지 증가하였다가 최근 2년간은 하락하며 2022년 345건을 기록하였다. 중국은 2009년 9건에서 2021년 156건을 무려 17배 이상 증가하였다. 2017년 962건까지 기록했던 미국은 빠르게 감소하여 2022년 568건에 이르러 2위인 일본과 190건 차이로 좁혀졌다.

[그림 3-8] 국가별 장비 관련 등록 특허 추이

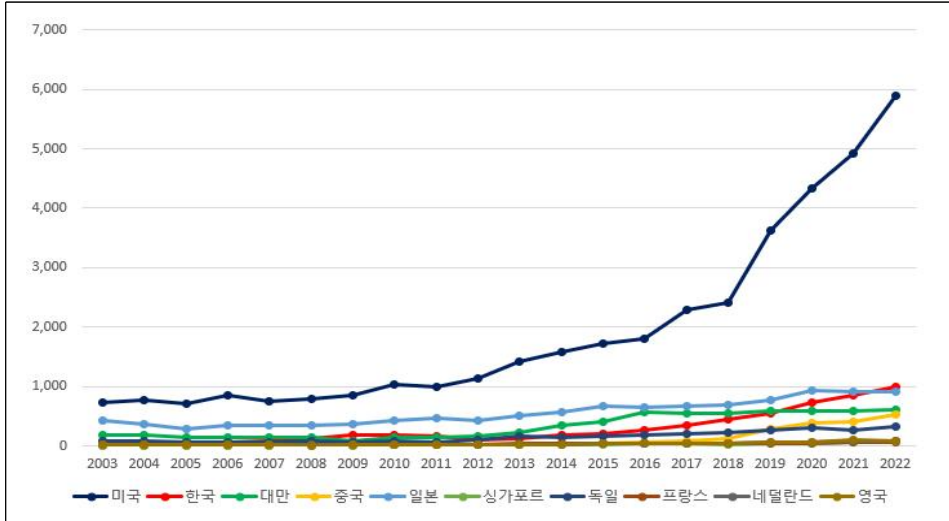


자료: 한국특허기술진흥원

마. 차세대반도체 관련 특허

차세대반도체 분야는 모든 국가에서 상승세를 보이는 분야이다. 특히 2018년 이후 미국의 증가는 다른 모든 국가를 합쳐도 비교가 안 될 정도이다. 미국은 2007년부 749건에서 서서히 증가하다가 2018년부터 급속히 증가하여 2022년 5,885건을 기록했다. 표본 기간 한국은 8.3%를 차지하며 일본(16%)과 대만(9.3%)에도 뒤져 있지만, 최근 빠른 속도로 등록 건수가 증가하면서 2022년 기준으로 등록 건수로는 2위를 차지했다. 한국의 차세대반도체 특허는 2012년 111건에서 2022년 987건으로 9배 가까이 증가했다. 차세대반도체 분야는 유럽의 독일, 네덜란드, 영국을 포함한 3개국의 상대적 비중이 가장 높은 분야이기도 하다. 미국이 다른 분야에서는 등록 특허가 감소하는 데 차세대반도체 분야에서 빠르게 증가하는 모습, 그리고 나머지 국가와의 차이가 빠르게 확대되고 있는 모습을 볼 때, 미국은 전략적으로 차세대반도체 분야에 연구 개발을 집중하는 것으로 추측할 수 있다.

[그림 3-9] 국가별 차세대반도체 관련 등록 특허 추이



자료: 한국특허기술진흥원

5. 반도체 기술 분야별 기술경쟁력 분석

가. 피인용지수

기술 분류별 평균적으로 피인용도가 높은 특허를 보유하고 있는 국가는 네덜란드이다. 네덜란드의 FAB 공정 분야 특허의 피인용 횟수는 평균 21.8회로 2위인 미국 11.8회 보다 두 배 가까이 많이 인용되고 있다. 소재와 장비 분야에서도 네덜란드 특허는 평균적으로 22회 인용되면서 2위인 미국의 약 13회보다 훨씬 많이 인용되었다. 특허 등록 건수 기준으로 FAB 분야에서 비교우위가 있었던 대만과 중국은 각각 7.4회와 1.4회로 네덜란드와 미국에 크게 못 미친다. 설계 분야 특허에 비교우위가 있었던 미국의 피인용 횟수는 11.4회로 네덜란드 13.6회보다 적지만 타 국가에 비교해서는 많은 인용되었다. 다른 분야에 비교해 차세대반도체 분야는 국가별 피인용지수의 편차가 작게 나타났다. 한국 특허 중 피인용지수가 가장 높은 분야는 소재 분야로 평균 6.3회이다. 그러나 소재 분야에서 한국은 네덜란드, 미국, 대만, 일본, 싱가포르 다음 6위를 차지하고 있다.

〈표 3-4〉 국가별 기술별 피인용도 지수

구분	FAB	설계	소재	장비	차세대반도체
미국	11.8	11.4	12.9	12.6	10.5
일본	6.8	8.0	8.6	7.0	6.5
대만	7.4	7.9	8.9	6.2	3.8
한국	5.0	5.3	6.3	5.4	2.9
독일	5.2	6.2	6.1	4.9	5.7
중국	1.4	1.8	1.3	1.2	1.3
싱가포르	7.5	4.4	8.4	7.2	3.5
프랑스	3.6	2.6	5.8	4.6	3.8
네덜란드	21.8	13.6	22.0	22.0	10.1
영국	5.7	8.3	4.7	5.9	9.5

자료: 한국특허기술진흥원

나. 기술 영향력 지수

기술 영향력 지수는 기술 분야에 크게 관계 없이 네덜란드와 미국이 가장 높다. 네덜란드는 FAB 공정, 설계, 소재, 장비 분야에서 모두 기술 영향력 지수가 가장 높은 나라이며, 차세대반도체 분야에서도 미국(1.25) 다음 1.21로 2위이다. 미국은 차세대반도체 분야에서 1위이며, 나머지 기술 분야에서는 모두 2위를 기록했다. 한국은 가장 낮은 차세대반도체 분야가 0.35이고, 가장 높은 장비 분야가 0.63로 나타났으며, 모든 분야에서 경쟁 국가로 볼 수 있는 대만보다 낮다. 특히, 한국의 차세대반도체 분야는 등록 특허 건수로는 볼륨과 증가 속도 측면에서 모두 강하게 나타났었는데 기술 영향력은 주요 10개국 중 중국 다음으로 낮게 나타났다. 한편, 중국은 모든 분야에서 기술 영향력이 가장 낮게 나타나 기술력에서는 다른 국가와 많은 차이를 보였다. 대만은 FAB 공정과 소재 부문에서 강세를 보이며 일본보다도 영향력이 크다. 싱가포르는 양적지표에서는 크게 두각을 나타내지 못하였으나, 기술 영향력 지수에서는 FAB과 소재, 장비에서 대만과 비슷하고 우리나라보다 뛰어난 것으로 나타났다.

〈표 3-5〉 국가별 기술별 기술 영향력 지수

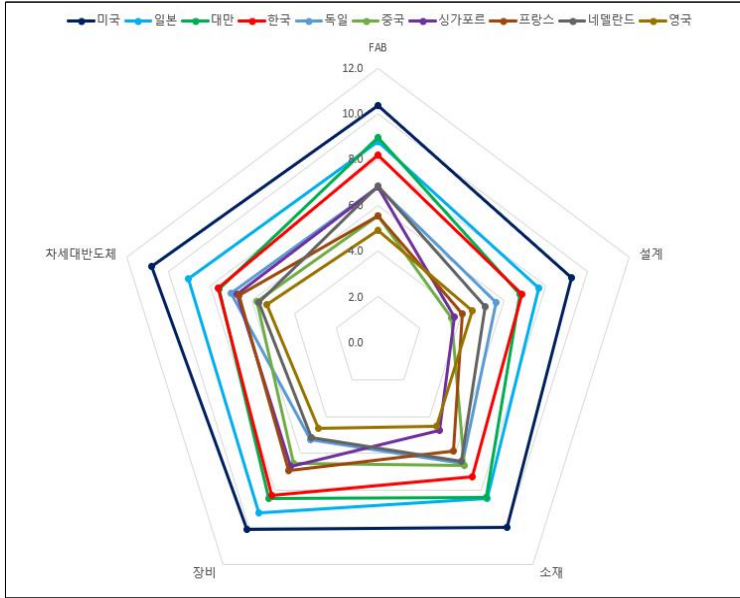
	FAB	설계	소재	장비	차세대반도체
미국	1.35	1.19	1.26	1.47	1.25
일본	0.78	0.84	0.83	0.82	0.77
대만	0.84	0.83	0.86	0.72	0.46
한국	0.57	0.56	0.61	0.63	0.35
독일	0.59	0.66	0.59	0.57	0.68
중국	0.16	0.19	0.13	0.14	0.16
싱가포르	0.85	0.46	0.82	0.84	0.42
프랑스	0.41	0.27	0.56	0.54	0.45
네덜란드	2.48	1.43	2.14	2.56	1.21
영국	0.65	0.88	0.45	0.68	1.13

자료: 한국특허기술진흥원

다. 기술력 지수

특허의 양적지표와 질적 지표를 모두 고려한 기술력 지수는 미국이 전 기술 분야에서 압도적으로 강한 것으로 나타났다. 일본은 양적지표와 질적 지표 각각에서는 약해지는 모습이 뚜렷했으나 표본 기간 전체를 보면 역시 다른 국가와 큰 차이로 모든 분야에서 2위를 지키고 있다. 우리나라는 소재 분야를 제외하고는 대만과 3위와 4위를 다투고 있으며, 소재에서는 대만에 크게 뒤진 4위이다. 기술 영향력에서 가장 낮았던 중국은 장비와 소재 분야에서는 중위권을 지키고 있으며, ASML 등 노광 장비로 유명한 네덜란드는 오히려 장비 분야에서 매우 낮게 나타났다.

[그림 3-10] 국가별 기술별(로그) 기술력 지수



자료: 한국특허기술진흥원

라. 기술 분류별 K-PEG 등급

1) FAB 공정 분야

FAB 공정 분야 특허에서 S등급은 네덜란드가 8.6%이고 미국이 6.8%로 평균(5%)보다 높았고, A1 등급에서도 네덜란드가 10.2%, 미국이 10.3%로 평균(7.5%)보다 높게 나타났다. 한국은 A2 등급의 특허가 29.2%로 평균(10%)보다 훨씬 크게 나타났다. FAB 공정 분야 전체 특허에서 각국의 우수 등급(S or A 등급) 비중을 살펴보면, 미국이 모든 우수 등급에서 표준보다 높게 나타났고, 대만은 A3 등급, 한국은 A2 등급, 네덜란드는 S와 A1 등급에서 표준보다 높게 나타났다. 특히 미국 특허는 FAB 공정 분야 S등급의 71.2%를 차지함으로써, 이 분야에서도 핵심 전략 기술은 미국이 보유하고 있을 것으로 추측된다.

〈표 3-6〉 국가별 FAB 공정 분야의 K-PEG 평가 결과의 등급별 분포 현황

구분	S	A1	A2	A3	B1	B2	B3	C1	C2
네덜란드	8.6	10.2	6.9	9.4	28.6	11.0	8.6	6.9	9.8
미국	6.8	10.3	20.9	20.6	20.4	9.8	5.2	3.6	2.4
영국	4.2	6.3	7.6	9.0	17.4	24.3	15.3	6.9	9.0
일본	3.3	6.8	5.8	14.6	25.1	19.6	12.6	7.9	4.3
독일	3.3	6.9	4.9	35.8	22.9	9.1	8.3	4.4	4.5
프랑스	2.2	5.1	4.7	7.1	25.4	23.0	17.1	6.7	8.7
한국	2.2	7.4	29.2	15.3	21.3	13.2	6.9	3.5	1.0
싱가포르	2.1	5.9	3.0	9.7	32.6	30.3	9.6	4.7	2.1
대만	1.3	4.7	4.8	34.0	37.6	8.5	5.0	3.1	0.9
중국	1.2	2.6	1.4	2.5	18.4	28.3	26.1	14.5	4.9

자료: 한국특허기술진흥원

2) 설계 분야

설계 분야의 S등급은 유일하게 미국이 6.2%로 평균(5%)보다 높았고, A1 등급에서는 네덜란드가 17.5%로 가장 높았고 그다음으로 미국이 9.8%로 한국이 9.5%, 일본이 9.4%, 대만이 9.1% 순으로 나타났다. 한국은 A1, A2 등급의 특허가 각각 평균보다 높은 것으로 나타났다. 설계 분야의 우수 등급 비중은 미국이 모든 등급에서 평균보다 높게 나타났고, 네덜란드, 한국, 일본은 A1 등급에서 평균보다 높은 것으로 나타났다. 한국의 설계 분야 우수 등급 특허 보유 비중은 55.4%로 대만(44.4%)이나 미국(50.3%)보다 더 많이 보유하고 있어, 특허의 질적 특성으로는 한국의 설계 분야 기술력이 우수한 것으로 나타났다. 반면, 중국은 17.9%에 불과해 아직은 기술격차가 크다는 점을 다시 한번 확인할 수 있다. 영국은 상대적으로 A3 등급을 많이 보유하면서 우수 등급 비중이 41.8%에 이르고, 네덜란드는 A1 등급을 17.5%나 보유하면서 우수 등급 비중이 35.1%를 차지했다.

〈표 3-7〉 국가별 설계 분야의 K-PEG 평가 결과의 등급별 분포 현황

구분	S	A1	A2	A3	B1	B2	B3	C1	C2
미국	6.2	9.8	16.1	18.2	21.5	11.2	6.7	4.6	5.7
영국	3.8	8.9	11.4	17.7	26.6	17.7	6.3	6.3	1.3
네덜란드	3.5	17.5	5.3	8.8	31.6	10.5	5.3	8.8	8.8
싱가포르	3.4	6.9	5.7	11.5	34.5	20.7	10.3	5.7	1.1
일본	3.1	9.4	8.5	15.1	27.4	17.3	11.1	5.4	2.8
독일	2.9	5.3	4.8	36.2	25.1	8.2	8.7	6.3	2.4
대만	2.8	9.1	8.3	24.2	34.9	11.6	4.2	3.4	1.5
한국	2.0	9.5	28.4	15.5	18.5	11.8	7.4	5.2	1.6
프랑스	1.9	3.7	7.5	12.1	23.4	22.4	17.8	7.5	3.7
중국	1.2	0.6	6.0	10.1	25.0	17.3	19.6	17.3	3.0

자료: 한국특허기술진흥원

3) 소재 분야

소재 분야의 S 등급 특허 수는 미국이 772건으로 일본의 143건의 약 7.6배의 많으며, A1 등급은 1,034건, A2 등급은 2,017건으로 나타났다. 한국은 S 등급 특허가 52건이고, A1 등급은 117건이며, A2 등급 특허는 373건으로 나타났다. S 등급 비중으로 보면, 네덜란드(11.2%)와 미국(7.7%)이 평균(5%)보다 높았고, A1 등급에서는 미국이 10.4%로 가장 높고, 그다음으로 네덜란드(8.6%), 일본(8.3%) 순으로 나타났다. 우수 등급 비중은 미국이 S, A1, A2 등급에서 평균보다 높게 나타났으며, 한국은 A2 등급에서만 평균보다 높게 나타났다. 대만은 A3 등급 비중이 35.2%를 기록하며, 우수 등급 비중이 48.1%를 차지했지만, S, A1, A2 등급 모두에서 한국보다 낮은 비중의 특허를 보유하고 있다.

〈표 3-8〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황

구분	S	A1	A2	A3	B1	B2	B3	C1	C2
네덜란드	11.2	8.6	3.6	7.1	24.9	16.2	9.6	6.6	12.2
미국	7.7	10.4	20.2	18.2	20.7	10.7	6.0	3.7	2.5
영국	5.2	8.2	2.2	12.7	26.9	12.7	10.4	7.5	14.2
일본	4.1	6.0	4.1	7.0	24.5	22.2	11.1	12.2	9.0
독일	3.9	7.7	4.1	23.9	19.2	12.6	10.8	8.8	9.0
프랑스	3.7	8.3	6.8	11.0	24.3	17.2	11.8	9.7	7.3
한국	3.4	8.0	5.2	10.1	27.6	31.1	7.2	5.8	1.6
싱가포르	3.2	7.3	23.2	14.4	22.7	13.3	7.9	4.7	3.3
대만	1.9	5.2	5.8	35.2	28.1	11.7	6.6	4.4	1.0
중국	1.4	2.7	1.4	2.2	21.2	26.9	24.7	13.7	5.9

자료: 한국특허기술진흥원

4) 장비 분야

소재 분야의 S 등급은 네덜란드(9.2%)와 미국(8.1%), 프랑스(5.6%)가 평균(5%)보다 높았고, A1 등급에서도 미국이 10.3%, 네덜란드가 9.2%로 평균(7.5%)보다 높게 나타났다. 한국은 A2 특허가 27.4%로 평균(10%)보다 크게 높은 것으로 나타났다. 우수 등급 특허의 비중은 미국이 S등급과 A1, A2 등급에서 평균보다 높은 것으로 나타났고, 네덜란드도 S등급과 A1 등급의 상대적 비중이 평균보다 높은 것으로 분석되었다. 한국은 A1, A2 등급이 평균보다 높게 나타났다. 대만은 장비 분야에서 A3 등급 비중이 36.2% 차지하며 우수 등급 특허 비중이 46.6%를 유지했으나, 한국의 51%보다는 낮은 수준이다. 유럽 국가 중에서는 독일이 49%로 우수한 특허를 가장 많이 보유한 것으로 나타났다.

〈표 3-9〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황

구분	S	A1	A2	A3	B1	B2	B3	C1	C2
네덜란드	9.2	9.2	4.8	8.1	18.8	18.1	10.7	8.9	12.2
미국	8.1	10.3	19.7	18.2	20.9	10.6	6.0	3.5	2.6
프랑스	5.6	2.8	6.0	10.3	24.2	14.7	15.1	10.7	10.7
일본	2.9	7.0	6.0	11.9	23.7	18.9	13.3	10.5	5.7
싱가포르	2.8	5.9	3.1	6.5	29.8	30.9	12.6	5.7	2.7
독일	2.4	6.6	4.3	35.7	24.6	9.9	6.3	4.4	5.9
한국	2.3	7.2	27.4	14.1	22.5	13.3	7.4	4.3	1.4
영국	2.1	4.3	3.2	7.4	22.3	12.8	18.1	11.7	18.1
대만	1.4	4.1	4.9	36.2	35.1	9.6	5.5	2.5	0.8
중국	1.1	2.1	0.8	2.4	22.0	28.4	25.0	13.7	4.5

자료: 한국특허기술진흥원

5) 차세대반도체 분야

차세대반도체 분야의 S 등급 미국이 8.6%이고 네덜란드가 6.6%, 싱가포르가 6.1%로 평균(5%)보다 높았고, A1 등급에서도 미국(10.0%), 영국(9.1%), 일본(8.4%) 순으로 평균(7.5%)보다 높게 나타났다. 한국은 A2 특허가 12.1%로 평균(10%)보다 높은 것으로 나타났다. 우수 등급 특허의 비중으로 보면, 미국은 S, A1, A2, A3 등급에서 모두 평균 이상의 특허를 보유하고 있으며, 한국은 A2와 A3, 대만은 A3 등급에서 1보다 크게 나타났고, 유럽 국가에서는 유일하게 영국이 A1 등급에서 1보다 크게 나타났다. 중국은 차세대반도체 분야에서 우수 등급 비중이 16.3%에 불과해 주요 10개국 중 가장 낮은 수준이며, 일본도 34.9% 불과해 미국, 대만, 한국에 비해 우수특허 보유 비중이 작게 나타났다.

〈표 3-10〉 국가별 소재 분야의 K-PEG 평가 결과의 등급별 분포 현황

구분	S	A1	A2	A3	B1	B2	B3	C1	C2
미국	8.6	10.0	9.9	13.4	21.9	13.5	9.5	7.5	5.7
네덜란드	6.6	7.5	9.0	11.1	25.4	16.3	11.6	7.9	4.5
싱가포르	6.1	7.6	6.8	12.7	27.4	15.2	10.0	6.8	7.3
일본	5.2	8.4	8.9	12.4	26.0	15.6	11.3	8.1	4.0
독일	5.0	8.1	9.3	10.3	23.8	17.6	12.7	8.2	4.9
영국	4.5	9.1	8.9	11.6	21.0	18.4	13.1	8.9	4.5
프랑스	3.8	5.0	7.6	8.0	25.5	16.6	16.2	11.0	6.2
한국	2.8	6.4	12.1	19.0	23.7	16.7	9.0	8.1	2.2
대만	2.2	6.6	5.7	17.2	30.5	15.0	11.0	8.3	3.5
중국	1.8	2.8	3.2	8.5	33.6	22.3	14.5	8.1	5.2

자료: 한국특허기술진흥원

6. 반도체 주요 기업별 등록 특허 분석

가. 반도체 상위 20개 기업 특허

반도체 분야 미국 등록 특허가 가장 많은 기업은 TSMC로 15,722건을 보유하고 있으며, 우리나라 기업으로는 삼성전자와 SK하이닉스가 각각 3위와 7위에 올라있다.

〈표 3-11〉 반도체 분야의 상위 20개 기업 등록 특허 현황

기업명	FAB	설계	소재	장비	차세대	합계
TSMC	6,414	705	3,217	4,584	802	15,722
IBM	4,889	1,499	2,971	2,929	2,431	14,719
삼성전자	3,676	1,077	1,168	2,838	2,333	11,092
마이크론	2,800	674	2,162	2,344	967	8,947
인텔	1,659	446	1,157	983	1,364	5,609
글로벌파운드리	1,663	229	886	1,079	88	3,945
SK하이닉스	1,508	291	296	1,407	438	3,940

기업명	FAB	설계	소재	장비	차세대	합계
도시바	1,524	509	424	1,037	395	3,889
인피니온	1,177	368	686	930	371	3,532
TI	915	268	642	740	379	2,944
SEL	478	103	253	978	862	2,674
퀄컴	411	149	182	160	1,468	2,370
르네사스	667	243	242	892	258	2,302
STM	584	145	516	387	578	2,210
소니	258	106	237	457	1,149	2,207
AMAT	598	28	419	705	399	2,149
AMD	676	120	418	586	176	1,976
SanDisk	813	259	251	242	34	1,599
UMC	719	66	213	472	66	1,536
총합계	31,429	7,285	16,340	23,750	14,558	93,362

자료: 한국특허기술진흥원

기술 분류별로 보더라도 TSMC는 FAB 공정 분야, 소재, 장비 분야에서 가장 많은 특허를 보유한 기업이며, IBM은 설계와 차세대반도체 분야에서 가장 많은 특허를 보유한 기업이다. 상위 20개 기업 중에서도 상위 4개의 기업(TSMC, IBM, 삼성, 마이크론)이 차지하는 비중이 54%를 넘어 상위기업에 특허가 집중되어 있다.

나. 반도체 상위 20개 기업 특허 지표

특허당 평균 피인용 수(CPP)는 AMT(어플라이드 머티리얼즈)가 17.75로 가장 높았고, 그다음으로 SEL(세미컨덕터 에너지 레보라토리)이 16.83, AMD(어드밴스드 마이크로 디바이스)가 16.79 순이었다. 한국의 삼성전자와 SK하이닉스는 각각 5.85와 3.44 회로 나타나 상위기업과 많은 격차가 있었다.

특허당 비특허 문헌의 평균 인용 수(SL)는 특허 등록 상위기업이 대체로 낮게 나타났는데, 이는 기초 연구보다는 상업화가 이루어지는 응용 기술 연구들이 이루어지기 때문으로 보인다. SL은 일본의 SEL이 22.83으로 가장 높고, SanDisk가 12.83, 마이크론이 10.95 순으로 그 뒤를 이었다.

상위 20개 기업의 특허 기술순환주기(TCT)는 대체로 약 6년 정도로 나타나는 데, SEL이 8.91년으로 가장 길고, 르네사스가 8.84년, 한국의 삼성전자는 6.21년 SK하이닉스는 5.30년으로 나타났다.

상위 20개 기업의 상대적 인용 비율을 나타낸 기술 영향력 지수(PII)에서 평균인 1을 넘어선 기업은 AMT(1.99), AMD(1.89), SEL(1.89), SanDisk(1.44), 마이크론(1.41) 다섯 개 기업이었다. 한국의 삼성전자는 0.66, SK하이닉스는 0.39로 기술 영향력에서 상위기업과의 차이가 크게 나타났다.

양적지표와 질적 지표를 모두 고려한 로그 기술력 지수(TS)는 TSMC가 9.5로 가장 높았고, IBM(9.46), 마이크론(9.45), 삼성전자(8.89), 인텔(8.59) 순으로 나타났다으며, SK하이닉스는 7.33으로 20개 기업 중 15위에 머물렀다.

〈표 3-12〉 상위 20개 기업 특허 지표

기업명	CPP	SL	TCT	PII	TS
TSMC	7.59	2.45	6.63	0.85	9.50
IBM	7.74	3.62	6.27	0.87	9.46
삼성전자	5.85	1.25	6.21	0.66	8.89
마이크론	12.58	10.95	7.06	1.41	9.45
인텔	8.56	6.43	6.83	0.96	8.59
글로벌파운드리	4.26	1.54	6.26	0.48	7.54
SK하이닉스	3.44	0.52	5.30	0.39	7.33
도시바	7.40	1.60	6.54	0.83	8.08
인피니온	5.02	1.69	7.43	0.56	7.60
TI	6.58	1.26	6.98	0.74	7.69
SEL	16.83	22.83	8.91	1.89	8.53
퀄컴	4.38	4.51	6.26	0.49	7.06
르네사스	2.13	1.42	8.84	0.24	6.31
STM	4.27	2.48	7.94	0.48	6.97
소니	5.22	1.81	8.12	0.59	7.17
AMAT	17.75	7.04	8.04	1.99	8.36
AMD	16.79	1.73	4.86	1.89	8.22

기업명	CPP	SL	TCT	PII	TS
SanDisk	12.84	12.83	5.40	1.44	7.74
UMC	3.32	0.75	6.48	0.37	6.35

자료: 한국특허기술진흥원

7. 주요 기업의 공동인용 분석

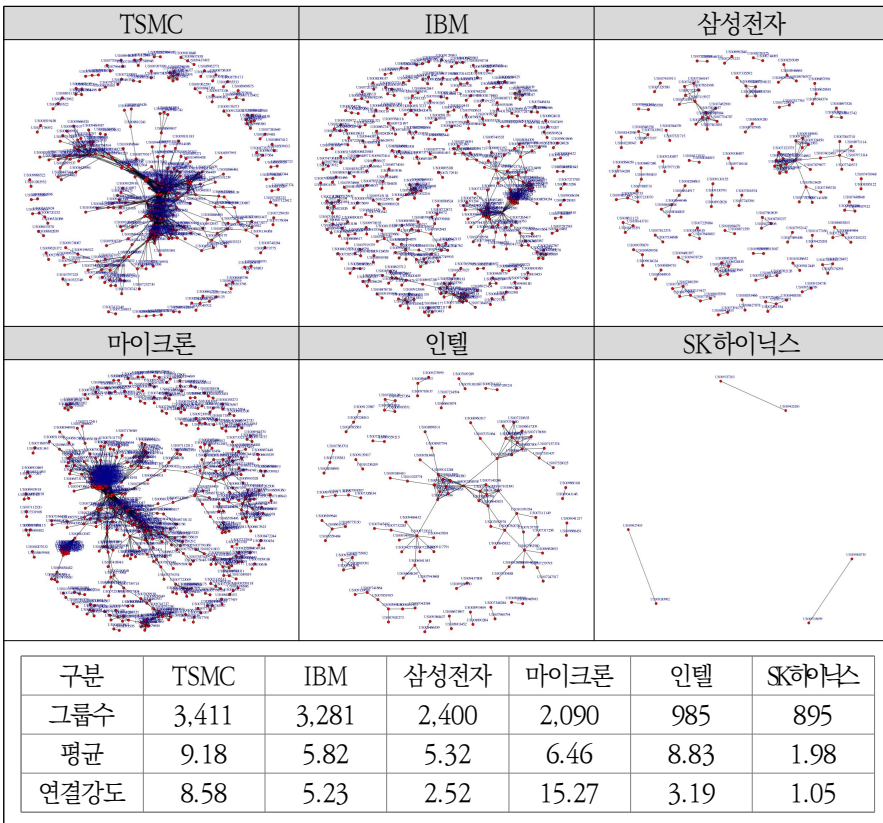
공동인용이란 두 개의 문서(논문, 특허 등)가 특허 등록 시 동시에 참조되는 경우를 일컬으며, 그 횟수가 많을수록 두 특허가 특정 기술에 더 밀접한 관련이 있다고 해석할 수 있다. 특정 기업이 보유한 특허가 공동인용 되는 경우가 많을 때, 그 기업의 핵심기술에 대한 특허 포트폴리오로 이해할 수 있다. 공동인용을 통해 형성된 특허 포트폴리오는 가치사슬의 관계를 형성하며 향후 라이선싱을 위한 특허 패키지로 활용될 수 있다. 본 분석에서는 반도체 분야 미국 등록 특허의 세부 기술 분야를 대상으로 공동인용을 분석하여 각 기업이 연구 개발을 수행할 때 기술 개발의 연속성을 가지는지, 어느 기술 분야에서 활발한지 등에 초점을 맞추어 살펴보았다. 특히 공동인용 링크의 강도는 노드의 연결 강도(Degree) 평균으로 측정하고 공동인용의 그룹 수 및 공동인용의 평균 공동인용 수를 통해 시각화된 정보와 관련한 수치를 제시하였다. 특허 등록 건수 기준 상위 5개 기업과 여기에 포함되지 않은 SK하이닉스까지 6개 기업에 대한 공동인용 분석하였고, 가중치가 5 이상인 공동인용 특허에 대해서만 그림으로 표현하였고, 공동인용 그룹 수와 평균 연결 강도는 모든 공동인용 특허를 대상으로 결과를 산출하였다.

가. FAB 공정 분야 공동인용 분석

TSMC와 마이크론의 FAB 공정 분야 포트폴리오가 다른 기업들과 비교해 체계적이고 강하게 연결된 것으로 나타났다. 공동인용 그룹 수를 살펴보면, TSMC가 3,411, IBM이 3,281, 삼성전자 2,400으로 나타났다. 기업별 미국 특허 등록 건수와 공동인용 그룹 수의 관계를 보면, TSMC는 6,414건에 3,411그룹, IBM 4,889건에 3,281그룹, 삼성전자는 3,676건에 2,400그룹, 마이크론은 2,800건에 2,090그룹이었다. 평균 공동인용 수를 살펴보면, TSMC가 9.18로 가장 높고, 인

텔이 8.83, 마이크론은 6.46으로 나타났다. 기업별 평균 연결 강도에서는 마이크론이 15.27로 가장 높았고, TSMC가 8.58, IBM이 5.23 순으로 나타났다. 한국 삼성전자의 평균 연결 강도는 2.52였고, SK하이닉스는 1.05로 분석되었다.

[그림 3-11] FAB 분야 주요 기업의 Co-citation clustering 분석

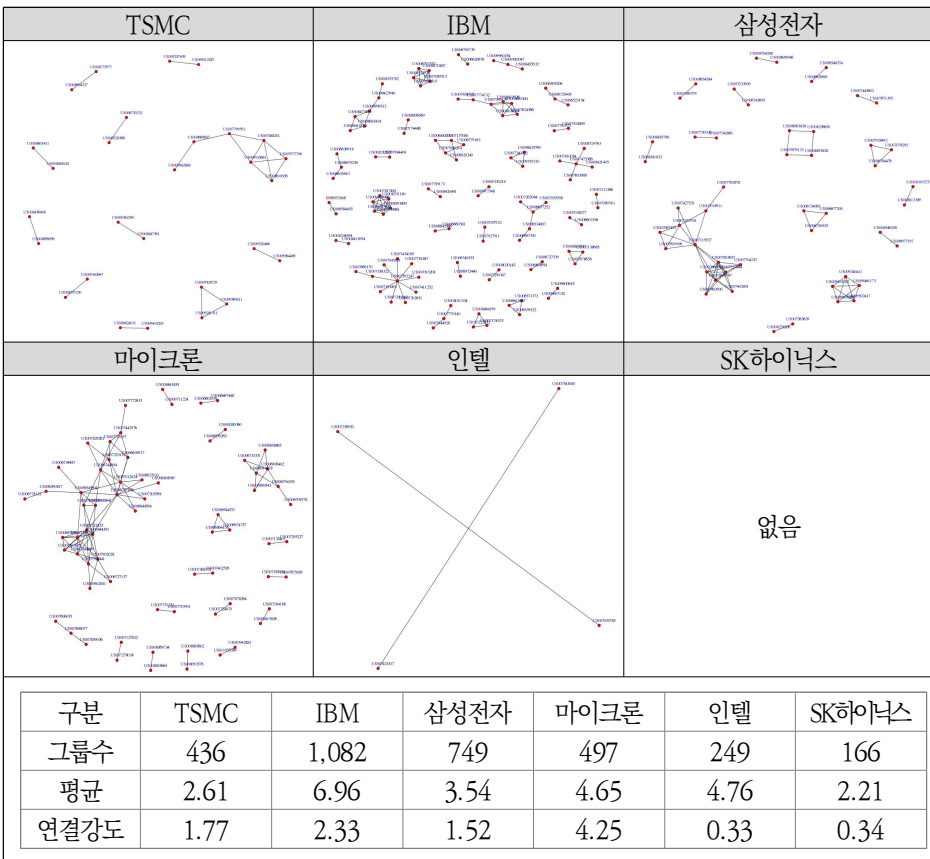


나. 설계 분야 공동인용 분석

설계 분야의 공동인용 그룹 수는 IBM이 1,082로 가장 많고 그다음으로 삼성전자가 749, 마이크론이 497 순으로 나타났다. 기업별 미국 특허 등록 건수와 공동인용 그룹 수의 관계를 보면, IBM 1,499건에 1,082그룹, 삼성전자는 1,077건에

749그룹, 마이크론이 674건에 497그룹이었다. 평균 공동인용 수를 살펴보면, IBM이 6.96이고 인텔이 4.76, 마이크론이 4.65, 삼성전자가 3.54로 나타났다. 기업별 평균 연결 강도에서는 마이크론이 4.25로 가장 높았고, IBM은 2.33, TSMC가 1.77 순으로 그다음을 이었다. 한국 삼성전자의 평균 연결강도는 1.52, SK하이닉스는 0.34로 분석되었다.

[그림 3-12] 설계 분야 주요 기업의 Co-citation clustering 분석

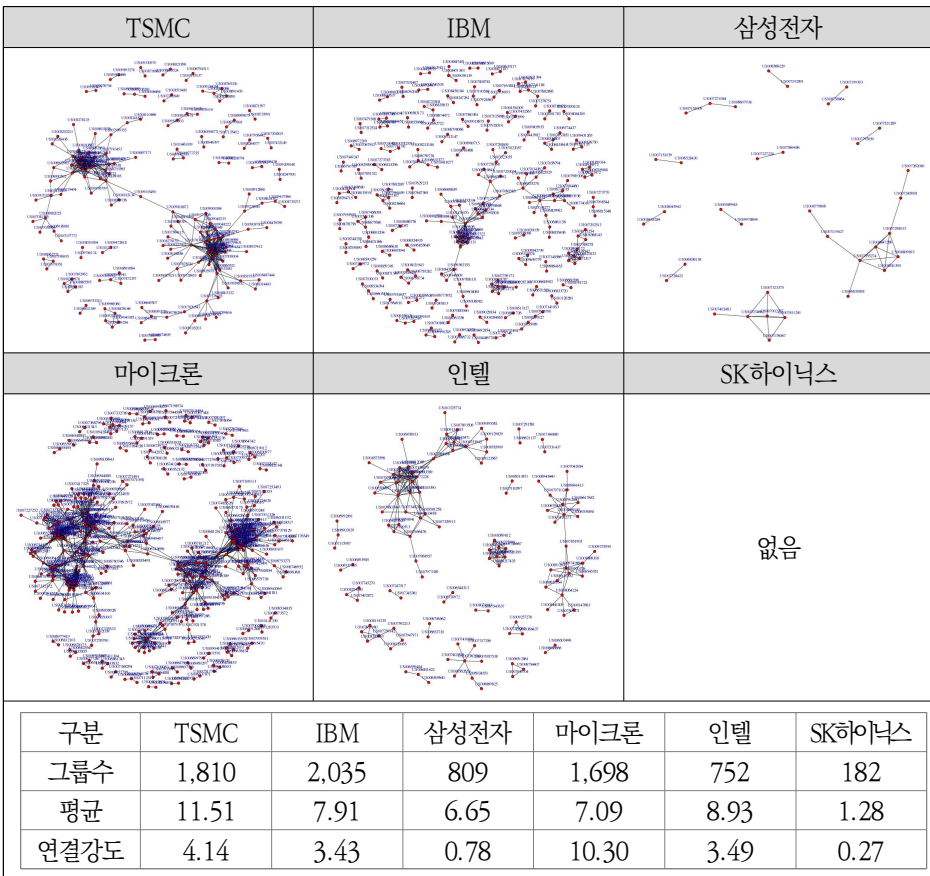


다. 소재 분야 공동인용 분석

소재 분야의 공동인용 분석 결과, 공동인용 그룹 수에서는 IBM가 2,035로 가

장 많고 그다음으로 TSMC가 1,810, 마이크론이 1,698 순으로 나타났다. 기업별 미국 특허 등록 건수와 공동인용 그룹 수의 관계를 보면, IBM 2,971건에 2,035 그룹, TSMC가 3,217건에 1,810그룹, 마이크론이 2,162건에 1,698그룹이었다. 평균 공동인용 수를 살펴보면, TSMC가 11.51로 가장 많고, 인텔이 8.93, IBM이 7.91, 마이크론이 7.09로 나타났다. 기업별 평균 연결 강도에서는 마이크론이 10.30으로 가장 높았고, TSMC가 4.14, 인텔이 3.49 순이었다. 한국 삼성전자의 평균 연결 강도는 0.78였고, SK하이닉스는 0.27로 분석되었다.

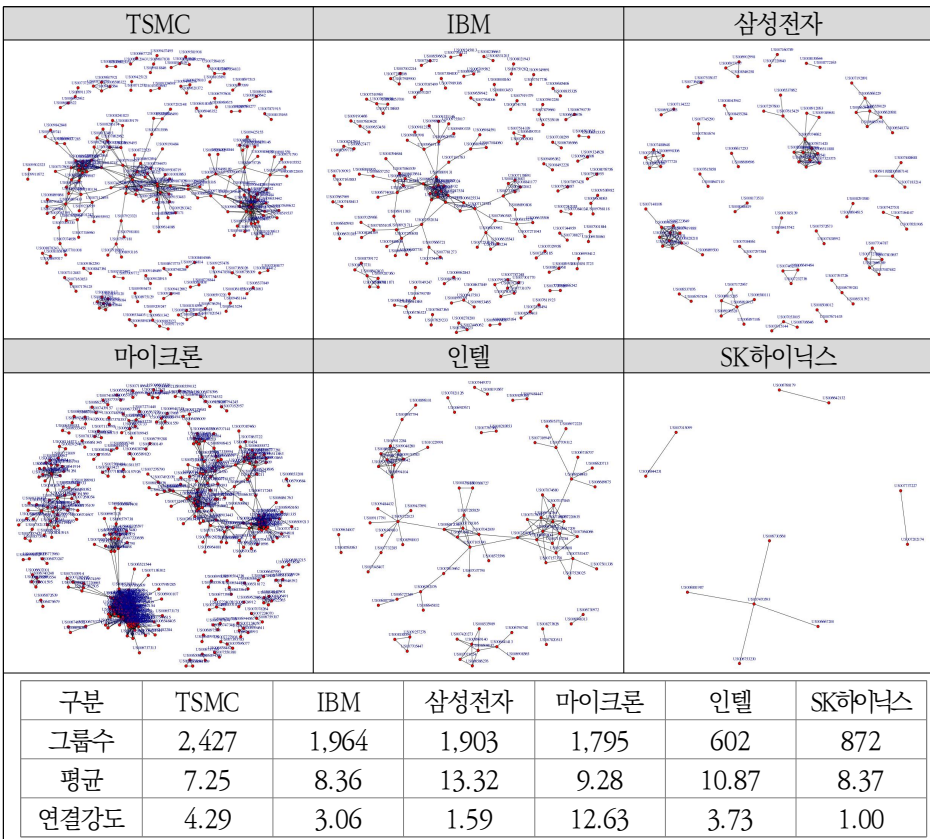
[그림 3-13] 소재 분야 주요 기업의 Co-citation clustering 분석



라. 장비 분야 공동인용 분석

장비 분야의 공동인용 그룹 수는 TSMC가 2,427로 가장 많고, IBM이 1,964, 삼성전자가 1,903 순으로 나타났다. 등록 건수와 공동인용 그룹 수의 관계를 보면, TSMC는 4,584건에 2,427그룹, IBM이 2,929건에 1,964그룹, 삼성전자가 2,838건에 1,903그룹이었다. 평균 공동인용 수는, 삼성전자가 13.32로 가장 많고, 인텔이 10.87, 마이크론이 9.28 순이었다. 기업별 평균 연결 강도에서는 마이크론이 12.63으로 가장 높았고, TSMC가 4.29, 인텔이 3.73 순이었다. 한국 삼성전자의 평균 연결 강도는 1.59였고, SK하이닉스는 1.00으로 분석되었다.

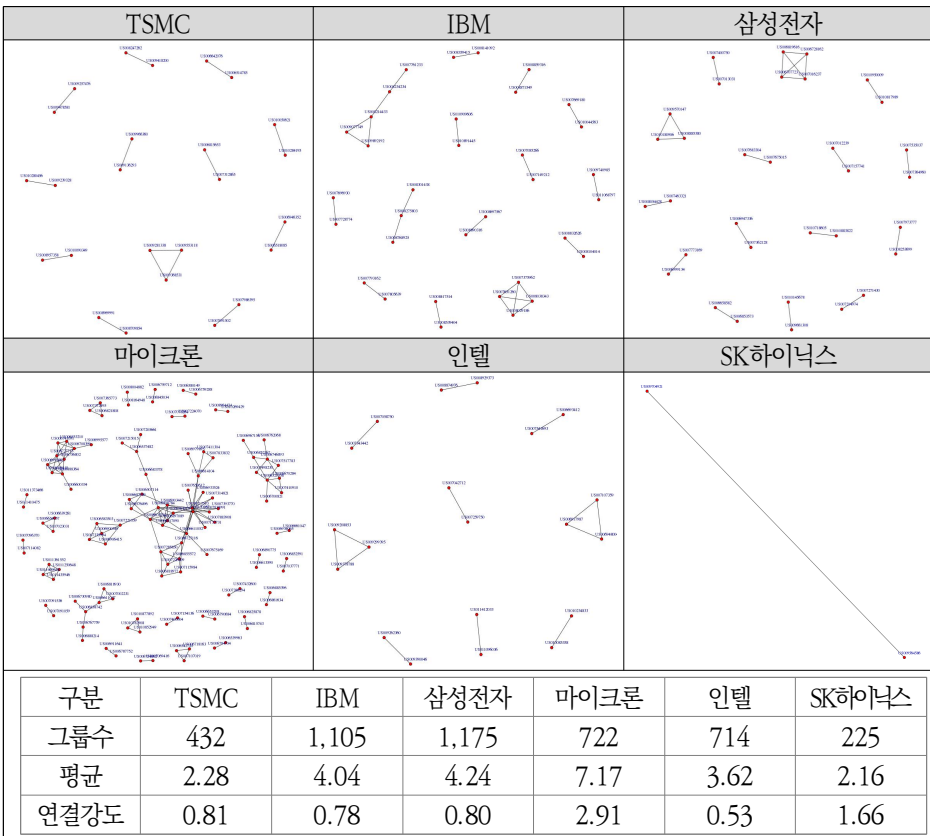
[그림 3-14] 장비 분야 주요 기업의 Co-citation clustering 분석



마. 차세대반도체 분야 공동인용 분석

차세대반도체 분야 공동인용 그룹 수는 삼성전자가 1,175, IBM이 1,105, 마이크로론이 722로 나타났다. 등록 건수와 공동인용 그룹 수의 관계는, 삼성전자가 2,333건에 1,175그룹, IBM이 2,431건에 1,105그룹, 마이크로론이 967건에 722 그룹이었다. 평균 공동인용 수는 마이크로론이 7.17로 가장 많고, 삼성전자가 4.24, IBM이 4.04 순으로 나타났다. 기업별 평균 연결 강도에서는 마이크로론이 2.91로 가장 높았고, SK하이닉스가 1.66, TSMC가 0.81, 삼성전자가 0.80으로 나타났다.

[그림 3-15] 차세대반도체 분야 주요 기업의 Co-citation clustering 분석



8. 공동협력 동향 분석

반도체 분야의 국가별, 기술 분야별 공동협력 특허 결과는 제1 특허권자의 국적과 기업명을 기준으로 분석을 수행하였다. 이를 통해 국가별 협력관계, 기업별 협력관계의 분석을 수행하였고, 협력관계의 양상을 보여주기 위해 Gephi를 활용하여 네트워크 분석을 하였다. 네트워크 분석을 수행하기 위한 노드 1에는 제1 특허권자의 국적과 기업명을, 노드 2에는 제2 특허권자의 국적과 기업명을 활용하였고 이러한 분석을 수행하기 위해 그룹과 계열사로 합쳐서 결과를 도출하였다.

가. 국가별 공동협력 동향

반도체 분야의 공동특허 현황을 살펴보면, 전체 미국 등록 특허 228,350건 중 6,510건의 공동특허를 나타내고 있어 2.9의 공동특허가 산출되는 것으로 나타났다. 미국이 1,812건으로 전체 공동특허의 27.8를 차지하고 있으며, 그다음으로 일본이 27.5, 한국이 12.1를 나타내고 있다. 국가별 특허 대비 공동특허 현황에서는 미국이 미국 국적 특허의 1.7가 공동특허이고, 일본은 4.5, 대만이 1.5, 한국은 3.6이다. 프랑스가 프랑스 국적의 미국 등록특허의 13.8의 특허가 공동특허로 나타나 공동특허의 발생 비중이 가장 높은 국가인 것으로 분석되었다.

〈표 3-13〉 국가별 공동특허 현황

구분	등록 특허 건수	공동협력 건수	비중(%)
미국	104,189	1,812	1.7
일본	39,483	1,788	4.5
대만	24,418	356	1.5
한국	21,954	788	3.6
독일	7,326	278	3.8
중국	6,062	525	8.7
싱가포르	3,588	81	2.3
프랑스	2,356	324	13.8
네덜란드	1,884	46	2.4

구분	등록 특허 건수	공동협력 건수	비중(%)
영국	1,342	25	1.9
전체	228,350	6,510	2.9

반도체 분야의 국가 간 공동특허 현황을 살펴보면, 대부분 국가가 자국 내 공동 특허를 많이 등록하고 있는 것으로 나타났다. 국가별로 살펴보면, 미국은 자국 내 공동특허가 882건이고, 일본과 292건, 카이만 제도와 230건의 공동특허가 발생하였다. 일본은 자국내 공동특허가 1,589건으로 대부분을 차지하고 있는 것으로 나타났고, 대만은 중국과의 공동특허가 많았다. 한국도 자국 내 공동특허가 697건으로 한국 내에서만 공동특허들이 많이 발생하였다.

〈표 3-14〉 국가 간 공동특허 건수

ISO	USA	JPN	TWN	KOR	DEU	CHN	FRA	SGP	NLD	UK
USA	882	129	36	36	95	56	36	27	13	4
JPN	292	1,589	3	24	6	4	4	-	-	7
TWN	24	6	196	-	6	134	-	-	-	-
KOR	56	5	-	697	-	-	-	3	-	-
DEU	106	4	3	13	104	4	22	2	1	-
CHN	41	7	107	1	-	321	-	1	-	1
FRA	30	12	-	1	12	-	226	9	7	2
SGP	29	3	3	2	6	-	1	33	-	-
NLD	17	-	-	-	3	-	9	-	7	1
UK	-	6	-	2	3	-	1	-	-	7
CAY	230	-	1	-	-	3	-	-	-	-
ITA	16	1	-	-	7	-	8	-	4	1
CAN	13	-	-	3	2	1	2	-	-	-
HK	7	5	-	-	-	-	-	-	-	-
Rest	69	21	7	9	34	2	15	6	14	2
합계	1,812	1,788	356	788	278	525	324	81	46	25

주: CAY는 케이맨제도, Rest는 기타 국가

나. 국가별 주요 기업의 공동협력

반도체 분야의 상위 특허권자들의 공동특허 현황을 살펴보면, STM이 259건으로 11.7의 공동특허를 발생하지만, 나머지 기업들은 공동특허의 비중이 매우 낮은 것으로 나타났다. STM도 타국의 STM의 계열사 간의 공동연구가 많아 실질적으로 다른 기업이나 대학 등과의 공동특허 건수는 상대적으로 작았다.

〈표 3-15〉 주요 기업의 공동특허

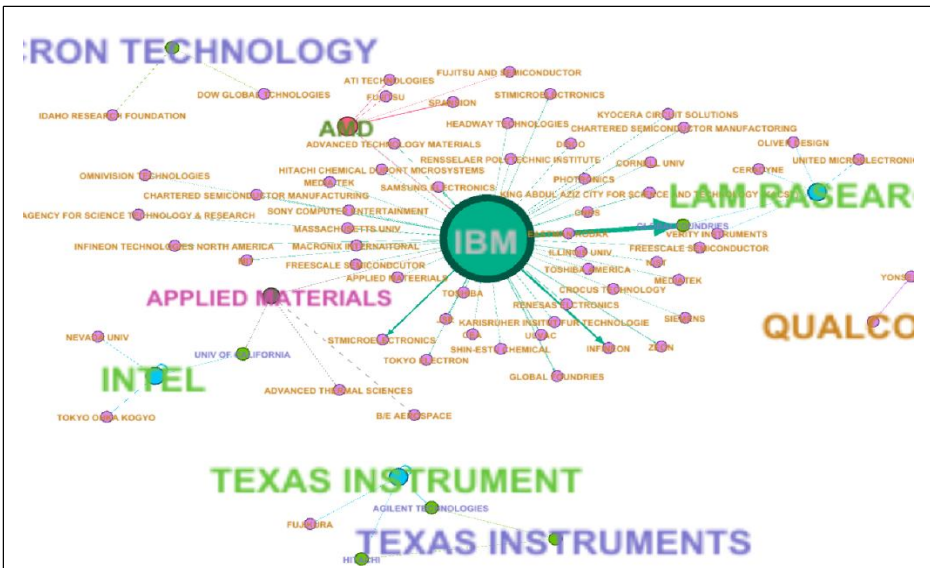
기업명	국가	등록건수	공동특허	비중
TSMC	대만	15,722	120	0.8
IBM	미국	14,719	453	3.1
삼성	한국	11,092	157	1.4
마이크론	미국	8,947	2	0.0
인텔	미국	5,609	7	0.1
글로벌파운드리	카이만제도	3,945	137	3.5
SK하이닉스	한국	3,940	46	1.2
도시바	일본	3,889	113	2.9
인피니온	프랑스	3,532	105	3.0
TI	미국	2,944	16	0.5
SEL	일본	2,674	10	0.4
퀄컴	미국	2,370	17	0.7
르네사스	일본	2,302	35	1.5
STM	스위스	2,210	259	11.7
소니	일본	2,207	84	3.8

국가별 주요 기업들의 공동특허 현황을 살펴보기 위해 SNA 분석 툴인 GEPHI를 활용하여 국가별 주요 기업들의 공동특허 관계를 살펴보았다. 제1 특허권자를 기준으로 분석을 수행하였으며, 제1 특허권자가 노드 1, 제2 특허권자가 노드 2를 가지도록 구성하였다.

1) 미국

미국 주요 기업의 공동특허 관계를 살펴보면, IBM이 가장 많은 노드가 발생하는 것을 볼 수 있고, AMT, 램리서치, AMD 등이 주변으로 분포된 형태를 보여주고 있다.

[그림 3-16] 미국 기업 간 공동특허 현황



미국 8개 주요기업의 공동특허 현황을 살펴보면, IBM이 가장 다양한 국가의 기업들과 공동특허가 발생하고 있는 것으로 나타났다. IBM의 공동특허 총 453건 중 글로벌파운드리와 189건의 공동특허가 산출되었고, 그다음으로 인피니온과 71건, STM과 49건, 삼성전자와 13건의 공동특허가 발생한 것으로 나타났다. AMD는 공동특허 45건 중 스펀션²⁰⁾과 26건의 공동특허가 산출되었고, TI는 애질런트테크놀로지²¹⁾사와 5개의 공동연구가 진행된 것으로 나타났다. 특이한 것은 켈

20) AMD와 후지쯔 합작회사

21) 1999년 휴렛팩커드에서 분사

컴의 공동특허 17건 모두 한국의 연세대학교와 공동연구가 진행된 것으로 나타났다. 미국의 주요 기업 8개의 공동특허 현황을 살펴볼 때, IBM만이 자국 또는 타국의 기업들과 공동특허들이 발생하지만, 나머지 기업들은 공동특허를 발생하는 기업들이 상호 겹치지는 않는 것으로 나타났다.

〈표 3-16〉 미국기업별 공동특허 현황

IBM 공동협력사	건수	AMD 공동협력사	건수
글로벌파운드리	189	스팬션	26
인피니온	71	ATI	8
STM	49	IBM	5
삼성전자	13	후지쯔	4
ZEON	13	후지쯔 반도체	2
JSR	12	총합계	45
AMD	12	AMT 공동협력사	건수
ULVAC	11	ADVANCED THERMAL SCIENCES	1
CSM	9	B/E AEROSPACE	1
KACST	8	IBM	2
기타	66	UNIVERSITY OF CALIFORNIA	2
총합계	453	총합계	6
TI 공동협력사	건수	램리서치 공동협력사	건수
애질런트 테크놀로지	5	CERADYNE	2
FUJIKURA	3	FRAUNHOFER	1
HITACHI	4	GLOBALFOUNDRIES	1
TI 계열사	4	LAM RESEARCH	1
총합계	16	OLIVER DESIGN	2
인텔 공동협력사	건수	UNITED MICROELECTRONICS	2
INTEL	2	VERITY INSTRUMENTS	1
NEVADA UNIVERSITY	2	총합계	10
TOKYO OHKA KOGYO	2	마이크론 공동협력사	건수
CALIFORNIA UNIVERSITY	1	다우 글로벌 테크놀로지	1
총합계	7	IRF(아이다호 연구소)	1
퀄컴 공동협력 현황	건수	총합계	2
연세대학교	17		
총합계	17		

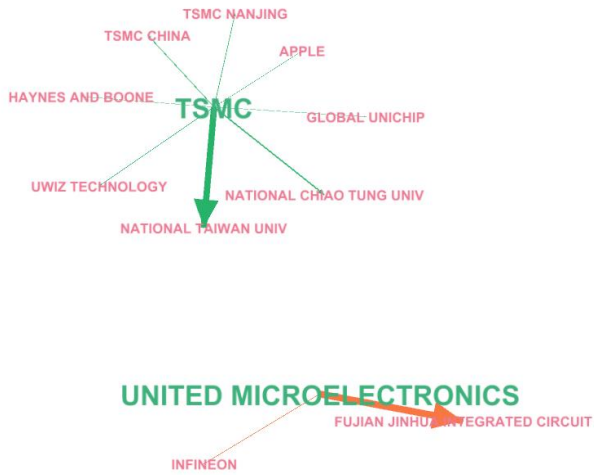
〈표 3-17〉 일본 기업들의 기업별 공동특허 현황

TOSHIBA 공동협력 현황	건수	TOKYO ELECTRON 공동협력사	건수
TOSHIBA 계열사	68	TOHOKU UNIVERSITY	8
EBARA	9	IBM	3
SanDisk 3D	6	IWATANI	3
JSR	5	TOHOKU UNIVERISTY	3
IBM	4	HORIBA	2
SANDISK TECHNOLOGIES	4	STANFORD UNIVERSITY	2
INFINEON	3	OSAKA UNIVERSITY	2
ROHM	3	TADAHIRO OHMI	2
DAI NIPPON PRINTING	2	TEXAS INSTRUMENTS	2
NSK	2	TOSHIBA	2
기타	15	기타	11
총합계	121	총합계	40
HITACHI 공동협력 현황	건수	르네사스전자 공동협력 현황	건수
HITACHI 계열사	150	HITACHI 계열사	15
RENESAS TECHNOLOGY	59	IBM	4
RENESAS ELCTRONICS	10	NEC	4
파나소닉 LCD	7	RENESAS NORTHERN JAPAN	3
NEC	7	HITACHI	2
ELPIDA MEMORY	6	ULVAC	2
AKITA ELECTRONICS	6	기타	5
L'AIR LIQUIDE	4	총합계	35
TEXAS INSTRUMENT	3	르네사스기술 공동협력사	건수
도쿄공업대학	3	HITACHI 계열사	63
기타	23	RENESAS 계열사	12
총합계	278	GRANDIS	4
SONY 공동협력사	건수	MATSUSHITA ELECTRIC	3
SONY 계열사	55	TOSHIBA	2
UNITED MEMORIES	7	SEIKO EPSON	2
SUMITOMO ELECTRIC IHDUSTRIES	5	기타	6
FUJIKURA	3	총합계	92
FZ JULICH	3	SEIKO EPSON 공동협력사	건수
TOSHIBA	3	OHMI, TADAHIRO	3
NORTH	2	TADAHIRO OHMI	3
기타	6	MITSUBISHI DENKI	2
총합계	84	PLASTIC LOGIC	2
마쓰시타(파나소닉) 공동협력사	건수	기타	4
SYMETRIX	5	총합계	14
KROSAKIHARIMA	2		
MITSUBISHI DENKI	2		
기타	6		
총합계	15		

3) 대만

대만 주요 기업 중 TSMC와 UNITED MICROELECTRONICS사의 공동특허 현황을 살펴보면, 두 회사의 공동특허 현황이 겹치지 않는 것으로 나타났다.

[그림 3-18] 대만 기업들의 전체 공동특허 현황



TSMC는 공동특허 120건 중 대만국립대와 86건으로 대부분의 공동특허가 발생하는 것으로 나타났고, UMC는 푸젠진화반도체사와 94건의 공동특허가 이루어지는 것으로 나타났다.

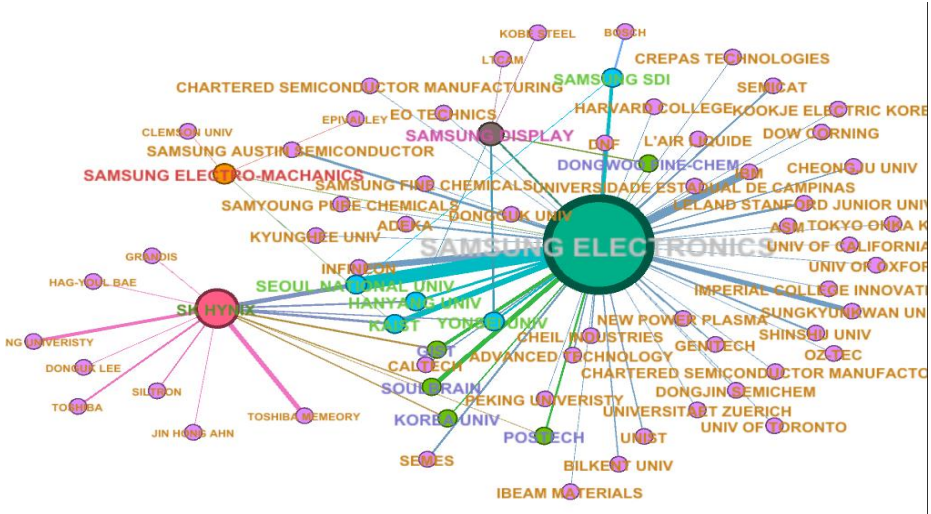
〈표 3-18〉 대만 기업들의 기업별 공동특허 현황

TSMC 공동협력사	건수	UMC 공동협력사	건수
대만국립대	86	푸젠진화반도체	94
국립자오통대학교	14	INFINEON	3
TSMC CHINA	11	총합계	97
UWIZ TECHNOLOGY	4		
APPLE	2		
COMPANY	1		
GLOBAL UNICHIP	1		
HAYNES AND BOONE	1		
총합계	120		

4) 한국

한국은 반도체 기업의 공동특허 현황을 살펴보면, 삼성전자가 공동특허 분포들이 다양하고 하이닉스는 작은 수의 연구 주체들과 공동특허가 발생하고 있는 것으로 나타났다.

[그림 3-19] 한국 기업들의 전체 공동특허 현황



삼성과 하이닉스의 공동특허에서 공동특허 발생 대상들이 대부분 겹치지 않는 독자적인 생태계를 구성하고 있는 것으로 나타났다. 공동특허 발생 대상이 겹치는 부분은 주로 대학이며, 삼성은 서울대학교와 22건, KAIST와 10건 등이고 하이닉스는 서울대학교와 7건, KAIST와 4건 등으로 나타났다. 대학 이외에 대상들은 서로 공동연구가 겹치지 않는 것으로 보인다. 삼성은 삼성 계열사와 14건, IBM과 13건, 인피니온과 10건 등의 공동특허를 보였고, 하이닉스는 도시바와 10건의 공동특허가 있었다.

〈표 3-19〉 한국 기업들의 기업별 공동특허 현황

삼성전자 공동협력사	건수	SK하이닉스 공동협력사	건수
서울대학교	22	도시바	10
삼성 계열사	14	서울대학교	7
IBM	13	서강대학교	5
인피니온	10	KAIST	4
KAIST	10	GIST	3
성균관대학교	8	한양대학교	3
SOULBRAIN	7	고려대학교	2
연세대학교	6	SILTRON	2
GIST	5	SOULBRAIN	2
한양대학교	5	연세대학교	2
기타	74	기타	6
총합계	174	총합계	46

5) 중국

〈표 3-20〉 중국 기업들의 기업별 공동특허 현황

SMIC 공동협력 현황	건수	징통팡(BOE tech) 공동협력사	건수
SMIC 계열사	18	BOE 계열사	76
온세미컨덕터	14	허페이신성광전자	23
NSI	11	오도스양센	7
SIEMENS	3	화남이공대학교	3
SHIRO SAKAI	1	K-TRONICS(SUZHOU) TECHNOLOGY	1
총합계	47	총합계	110
화웨이 공동협력 현황	건수	BAIDU 공동협력 현황	건수
BMW	2	BAIDU.COM TIMES TECHNOLOGY(BEIJING)	7
프라운호퍼(독일정출원)	1	KUNLUNXIN TECHNOLOGY(BEIJING)	6
푸단대학교	1	BAIDU US	1
맥길대학교	1	SHANGHAI XIAODU TECHNOLOGY	1
폭스바겐	1	총합계	15
총합계	6		

중국 기업들의 전체 공동특허 현황을 살펴보면, SMIC, BOE GROUP, 화웨이, 바이두 등의 협력네트워크가 구축되어 있으나, 기업마다 공동특허 주체가 연결되지 않고 서로 다른 생태계를 구성하고 있는 것으로 나타났다. SMIC는 47건의 공동특허 건수 중 SMIC 계열사와 18건의 공동특허 건수가 있었고, BOE GROUP도 110건의 공동특허 건수 중 BOE 계열사와 76건의 공동특허 건수를 보였다. 바이두 역시 15건의 공동특허 건수 중에서 바이두 계열사와 8건의 공동특허 건수가 있는 것으로 나타났다.

[그림 3-20] 중국 기업들의 기업별 공동특허 현황



6) 프랑스

프랑스 기업들의 공동특허 현황을 살펴보면, STM(FR)는 공동특허 건수 57건 중 STM과의 공동특허가 20건으로 나타났고, CEA와 18건, PHILIPS와 5건으로 조사되었다. SOITEC도 공동특허 건수 11건 중에서 CEA와 6건의 공동특허 건수를 보였다.

〈표 3-21〉 프랑스 기업들의 기업별 공동특허 현황

STM(FR) 공동협력사	건수	SOITEC 공동협력사	건수
STM계열사	20	CEA	6
CEA	18	에리조나대학교	3
PHILIPS	5	STM	2
NXP	3	총합계	11
SOCpra SCIENCE ET GENIE S.E.C.	2		
기타	5		
총합계	57		

7) 네덜란드

네덜란드 기업들의 공동특허 현황을 살펴보면, PHILIPS는 34건의 공동특허 건수 중에서 18건이 PHILIPS 계열사와 공동특허를 발생하였고, IMEC과 10건의 공동특허 건수를 보였다. NXP는 10건의 공동특허 건수 중에서 STM과 3건의 공동특허 건수가 있었다. ASML은 공동특허 건수가 1건 밖에 없어 공동특허를 창출하지 않는 것으로 나타났다.

〈표 3-22〉 네덜란드 기업들의 기업별 공동특허 현황

PHILIPS 공동협력 현황	건수	ASM IP HOLDING 공동협력 현황	건수
IMEC	10	IMEC	2
필립스전자	9	ASM IP HOLDING	1
필립스 루미레즈	9	CARL ZEISS	1
STM	2	코닌클리케 필립스	1
케이스 웨스턴 리저브 대학교	1	총합계	5
인피니온	1	NXP 공동협력 현황	건수
MAT-TECH	1	CEA	1
STIMICROELECTRONICS	1	FREESCALE SEMICONDUCTOR	2
총합계	34	IMEC	2
ASML 공동협력 현황	건수	NXP 계열사	2
CARL ZEISS	1	STM	3
총합계	1	총합계	10

8) 영국

영국 기업들의 공동특허 현황을 살펴보면, 공동특허 건수가 거의 발생하지 않는 것으로 조사되었다. CAMBRIDGE DISPLAY TECHNOLOGY의 7건의 공동특허 건수 중에서 PANASONIC과 5건의 공동특허 건수를 보였다.

〈표 3-23〉 영국 기업들의 기업별 공동특허 현황

캠브리지 디스플레이 테크놀로지 공동협력사	건수	ARM 공동협력사	건수
PANASONC	5	캠브리지대학교	1
CDT OXFORD	1	총합계	1
SUMITOMO CHEMICAL	1		
총합계	7		

9. 기술 의존도(Backward Citation Analysis)

미국 특허는 특허출원 시에 출원인이 이미 알고 있는 자신의 발명과 가장 유사한 문헌들에 대해 자진해서 그 정보를 미국특허상표청에 제출해야 하는 의무제인 IDS(Information Disclosure Statement)가 있다. 제출된 자료는 심사관의 심사 시 활용한 자료 중 좋은 자료와 함께 미국 특허 공보상에 US reference와 non-patent reference로 기재된다. 이렇게 기재된 US reference는 특허 분석에 있어서 Backward Citation으로 정의하고 분석 시 매우 중요한 의미가 있다. 특허를 생성할 때 어떤 기업의 어떤 기술을 활용하였는지를 통해 경쟁 관계를 파악할 수 있기 때문이다. 또한 Backward Citation을 두 개로 구분하여 자기 인용(Self Citation)과 타인 인용(Other Citation)으로 세분화할 수 있다. 자기 인용이란 기술 개발 시 기존 자사의 기술을 활용하여 연구함으로써 기술 개발의 연속성 유무를 파악할 수 있는 자료이다. 따라서 자기 인용 비율이 높으면 기술 개발의 연속성이 높다고 할 수 있다.

이번 분석에서는 기술 의존도 분석하기 위해 Backward Citation을 활용하여 자기 인용과 타인 인용으로 구분하여 분석을 수행하였다. 반도체 분야의 국가별 자기 인용(self citation)을 살펴보면, 전체 등록 건수의 평균은 2.15이다. 이는 특허 1건에 자기 인용되는 특허가 평균 2.15건을 의미하는 것이다. 네덜란드가 자기 인용 건수는 6.49로 가장 높고, 미국이 3.15, 일본이 2.17 순으로 나타났다. 네덜란드는 4구간에 자기 인용 건수가 크게 올라가서 자기 인용 건수가 가장 높게 나타난 것으로 각 구간으로 보면 자기 인용이 그렇게 높지는 않은 것으로 나타났다. 미국은 2구간에서 특허 1건당 평균 4.71 건의 자기 인용 특허 건수를 보였으나 4구간으로 오면서 2.40으로 조금 감소하고 있고, 일본 역시 3구간에서 2.51에서 4구간에는 1.86으로 감소한 것으로 나타났다. 자기 인용 건수가 1.0을 넘는 국가는 네덜란드, 미국, 일본, 대만이고 나머지 국가는 1.0 이하로 나타났다. 한국은 자기 인용 건수가 0.65로 조사되었다.

〈표 3-24〉 반도체 분야 미국 등록 특허의 국가별 자기 인용 현황

구분	2003-2007	2008-2012	2013-2017	2018-2022	총합계
미국	2.98	4.71	3.20	2.40	3.15
일본	1.96	2.30	2.51	1.86	2.17
대만	0.32	1.73	1.37	1.84	1.48
한국	0.35	0.58	0.72	0.80	0.65
독일	0.40	0.44	0.69	0.68	0.58
중국	0.18	0.54	0.13	0.19	0.19
싱가포르	0.30	0.24	0.48	0.24	0.34
프랑스	0.24	0.33	0.23	0.34	0.29
네덜란드	0.22	0.59	1.25	18.97	6.49
영국	0.21	0.33	0.16	0.45	0.32
총합계	1.99	2.81	2.11	1.89	2.15

상위 10개 기업의 자기 인용 현황을 살펴보면, 마이크론이 11.77로 가장 높고, IBM이 2.57, 인텔이 1.93 순으로 나타났다. 마이크론은 다른 기업에 비해서 자기 인용 건수가 매우 높은 것으로 나타났다. 이는 연구개발 시 또는 특허출원 시 자사가 보유한 기술에 대한 연속성이 매우 높음을 의미한다. 즉 보유한 기술을 기반으로 연속성을 가지고 연구 개발이 진행되고 있다는 것이다. 모든 구간에 걸쳐서 다른 기업들에 비해 자기 인용 건수가 매우 높은 것으로 나타났다. 그다음으로 IBM은 전 구간에 걸쳐 2.0 이상의 자기 인용 건수를 보여주고 있다. TSMC는 최근 구간으로 오면서 자기 인용 건수가 점점 많아지고 있다. 4구간에 자기 인용 건수는 마이크론에 이어 2번째로 많은 것으로 나타났다. 한국의 삼성전자는 자기 인용 건수가 1.01이고 SK하이닉스는 자기 인용 건수가 0.28로 나타났다.

반도체 분야 미국 등록 특허의 Backward Citation 데이터를 가지고 상위 10개 기업 간 어느 기업의 특허들을 많이 인용했는지를 살펴보면 다음과 같다. 대부분 기업이 자사의 특허를 인용한 비중이 가장 많은 것으로 나타났다. 자사 특허들의 인용 비중을 보면, TSMC가 33.76이고, 마이크론이 32.89, 도시바가 30.67순

이고, 삼성전자는 17.76, SK하이닉스는 12.7로 조사되었다. 각 기업이 타사의 특허를 얼마나 인용하였는지에 대한 비중을 살펴보면, TSMC는 IBM, 삼성전자, 인텔의 특허에 대한 인용 비중이 높은 것으로 나타났다. IBM 역시 TSMC 특허의 인용 비중이 가장 높았고, 그다음으로 인텔, 마이크론 순이었다. 삼성전자는 도시바 특허의 인용 비중이 5.51로 가장 높았고, 그다음으로 IBM, TSMC 순이었다.

〈표 3-25〉 반도체분야 미국등록특허의 기업별 자기인용 현황

구분	2003-2007	2008-2012	2013-2017	2018-2022	총합계
IBM	2.57	2.96	2.82	2.12	2.57
TSMC	0.35	0.50	1.41	2.38	1.71
삼성전자	0.58	0.89	1.19	1.17	1.01
마이크론	8.15	21.59	14.23	5.32	11.77
인텔	1.54	2.72	2.81	1.45	1.93
SK하이닉스	0.16	0.50	0.15	0.25	0.28
도시바	1.35	1.61	1.89	1.08	1.61
인피니온	0.29	0.46	0.77	0.76	0.58
TI	1.39	1.01	0.74	1.02	1.05
글로벌파운드리	-	0.04	0.17	0.43	0.23
총합계	1.99	2.81	2.11	1.89	2.15

마이크론, 인텔, 인피니온, TI 모두 IBM의 특허 인용 비중이 가장 높았고, 도시바는 삼성전자의 특허를 많이 인용하고 있는 것으로 나타났다. SK하이닉스는 자사 인용 비중보다 삼성전자 특허의 인용 비중이 더 커, 삼성전자의 특허에 많은 기술적 영향이 많은 것으로 나타났다.

〈표 3-26〉 반도체 분야 미국 등록 특허의 기업별 Backward Citation 현황

	TSMC	IBM	SE	MI	IN	SK	TO	IF	TI	GF
TSMC	33.76	5.09	4.91	1.51	5.28	4.14	1.50	2.69	3.71	10.95
IBM	8.58	28.90	4.97	6.09	11.80	4.35	4.03	7.11	6.75	23.57
SE	3.96	2.48	17.76	3.00	3.78	13.85	5.03	2.29	2.46	4.11
MI	2.29	2.60	3.63	32.89	2.47	6.36	2.80	3.04	2.70	3.10
IN	3.74	3.18	1.81	1.92	15.64	1.24	1.02	1.72	1.96	2.07
SK	0.63	0.53	2.84	0.77	0.60	12.70	1.08	0.40	0.37	0.84
TO	2.05	1.83	5.51	3.05	3.08	7.28	30.67	3.67	2.05	2.27
IF	0.70	0.96	0.63	0.68	0.69	1.02	0.65	12.26	1.44	1.28
TI	1.73	1.93	1.09	2.22	1.84	1.82	1.75	2.75	17.20	1.93
GF	1.61	2.46	1.00	0.04	0.72	0.27	0.12	0.13	0.38	7.84

주: SE, MI, IN, SK, TO, IF, TI, GF는 각각 삼성전자, 마이크론, 인텔, SK하이닉스, 도시바, INFINEON, TEXAS INSTRUMENT, Global Foundry 일컬음

제 2 절 주요국 반도체 스타트업의 투자유치 동향 분석

1. 데이터

본 절에서는 글로벌 스타트업 투자 관련 데이터 제공 플랫폼 ‘크런치베이스 (Crunchbase)’을 통해 미국, 중국, 유럽, 한국, 일본의 반도체 스타트업 기업을 식별하고, 투자유치 횟수, 투자유형, 투자자 등의 정보를 확보하였다.²²⁾ 유럽은 유럽연합 소속 국가만을 의미하지 않으며, 영국을 포함한다. 크런치베이스는 중국과

22) 크런치베이스가 공공이나 국가 혹은 공인된 써드파티 정보제공업체에서 받은 정보만 제공하는 것이 아니므로 정보 오염의 가능성에 유의해 분석 결과를 해석할 필요가 있다.

별개로 대만의 스타트업 투자 정보도 제공하나, 반도체 산업 스타트업의 투자 정보는 2건에 불과해 본 연구의 대상에서는 제외하였으며, 아시아, 중남미, 아프리카, 중동 지역의 국가는 현재 글로벌 반도체 산업에서 역할이 거의 없어 본 연구의 대상에서 제외하였다.

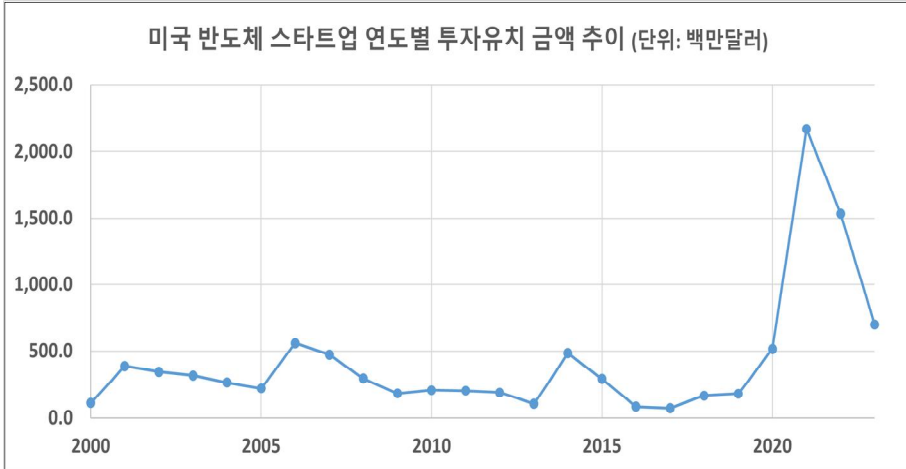
반도체 기업 식별은 반도체의 제조·생산 기업으로 제한하지 않고, 크런치베이스의 산업분류에서 과학과 공학 하부의 ‘Semiconductor’와 하드웨어 하부의 ‘Semiconductor’ 항목의 모든 기업의 투자 정보를 획득하였다. 분석 대상 5개국 및 지역별로 제공 데이터양이 균일한 ‘펀딩’ 항목 데이터를 주요 분석 대상으로 하였다.

2. 주요국 반도체 스타트업 투자유치 추이

가. 미국

미국 반도체 스타트업 생태계에는 '20년 이후 줄곧 연간 5억 달러 이내의 투자금이 유입되다가 트럼프 행정부가 중국 최대 반도체 파운드리 SMIC를 비롯한 중국 기업 10여 곳을 Entity list에 추가한 직후인 '21년에 약 21.7억 달러가 투자를 유치했다. 바이든 행정부의 반도체법 서명이 있는 '22년에도 약 15.4억 달러의 투자가 반도체 스타트업에 이루어졌으나, 전년도 대비해서는 약 29%가 감소한 수준이다. '23년은 10월 말 기준 약 7억 달러로 '20년까지 연간 투자유치 규모와 비교해 유의미한 차이가 없었다.

[그림 3-21] 미국 반도체 스타트업 연도별 투자유치 금액 추이

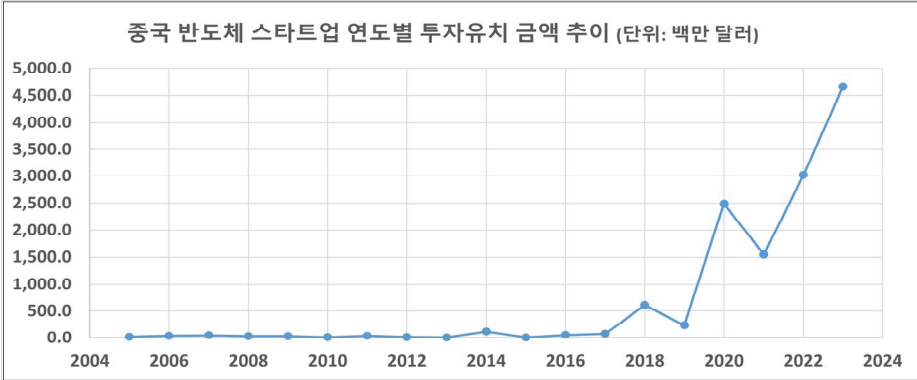


자료: Crunchbase 데이터 가공

나. 중국

중국 반도체 스타트업의 투자는 ‘집적회로산업발전촉진요강’이 발표되고 빅펀드가 조성된 ‘14년에 처음으로 1억 달러 이상의 투자가 스타트업에 유입된 후 거의 없다가, 트럼프 행정부의 대중 제재가 시작된 ’18년에 6억 달러, ASML 장비의 중국 반입 금지와 블랙리스트가 시작된 ’20년에 약 25억 달러가 유입되었다. 바이든 정부의 반도체법이 서명된 ’22년에 반도체 스타트업 투자는 30억 달러를 넘어섰으며, 올해 10월 말 기준으로 이미 46억 달러를 넘어섰다.

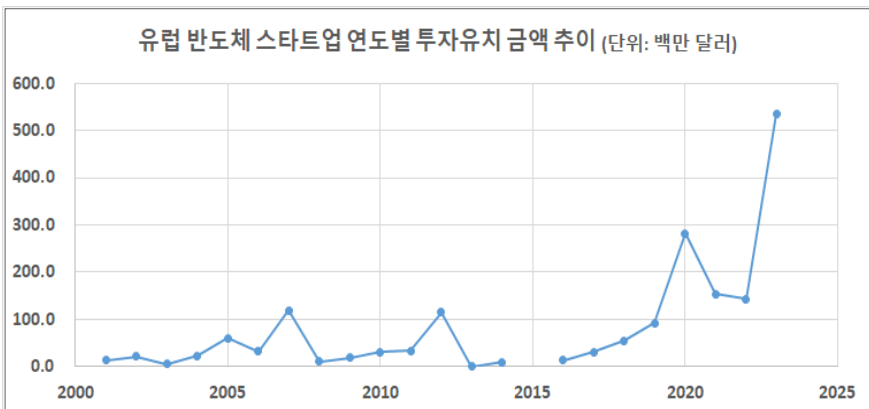
[그림 3-22] 중국 반도체 스타트업 연도별 투자유치 금액 추이



다. 유럽

유럽은 '20년 12월 EU 회원국들은 '프로세서 및 반도체 기술 관련 공동 성명'에 서명함으로써, 유럽의 반도체 설계 및 생산설비 제고를 위한 공동 대응의 필요성에 합의한 바 있다. 공동 성명 이후 '21년 EU 집행위원장은 유럽 반도체 산업 지원법을 제안했고, EU는 '23년 4월, 2030년까지 430억 유로를 투입해 유럽 반도체 산업을 지원하는 내용을 골자로 하는 '반도체법(Chips Acs)' 제정에 합의했다.

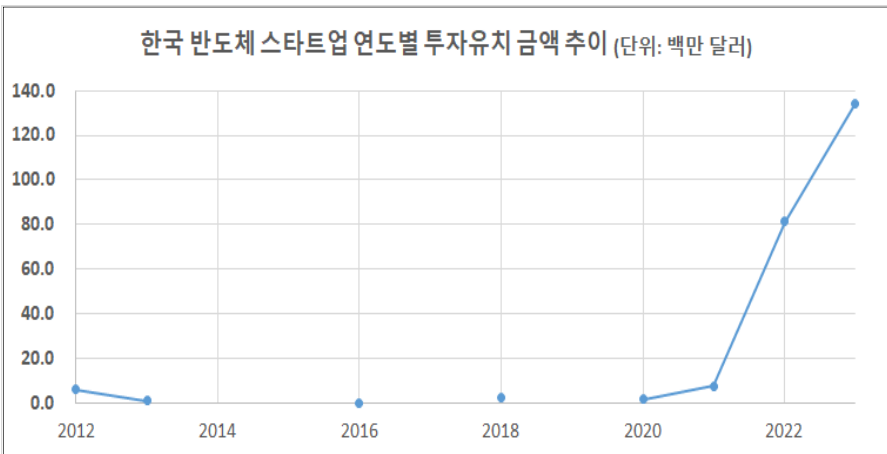
[그림 3-23] 유럽 반도체 스타트업 연도별 투자유치 금액 추이



라. 한국

한국 반도체 스타트업 생태계에 대한 투자는 아직 미미한 수준²³⁾으로, '22년에 유의미한 움직임이 시작되었으며, '23년에 처음으로 투자유입 금액이 1억 달러를 넘어섰다. 미국의 반도체법이 국내 반도체 산업에 미칠 파급에 대한 우려가 제기된 '22년부터 국내 스타트업에 대한 투자가 진행된 것으로 이해할 수 있다.

[그림 3-24] 한국 반도체 스타트업 연도별 투자유치 금액 추이

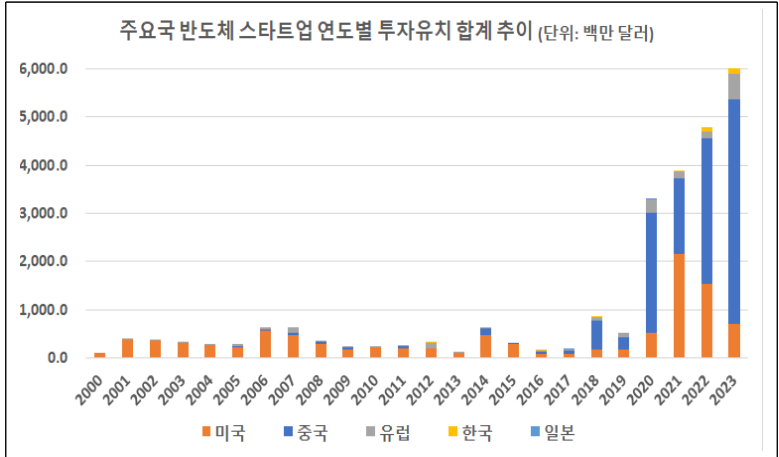


3. 미·중 갈등의 반도체 스타트업 투자 영향

미·중 간 무역분쟁이 불어졌던 2018년 이후 반도체 스타트업 기업들의 투자유치 추이를 보면 오히려 빠르게 증가했고, 특히 코로나19가 확대되었던 2020년 이후에는 연평균 증가율이 20%를 넘고 있다.

23) 크런치베이스의 통계에 잡히지 않았을 가능성이 있으나, 비교적 최근 연도의 펀딩 데이터만 보아도 한국이 글로벌 반도체 산업에서 갖는 위상에 비해 미미한 수준

[그림 3-25] 주요국 반도체 스타트업 연도별 투자유치 합계 추이

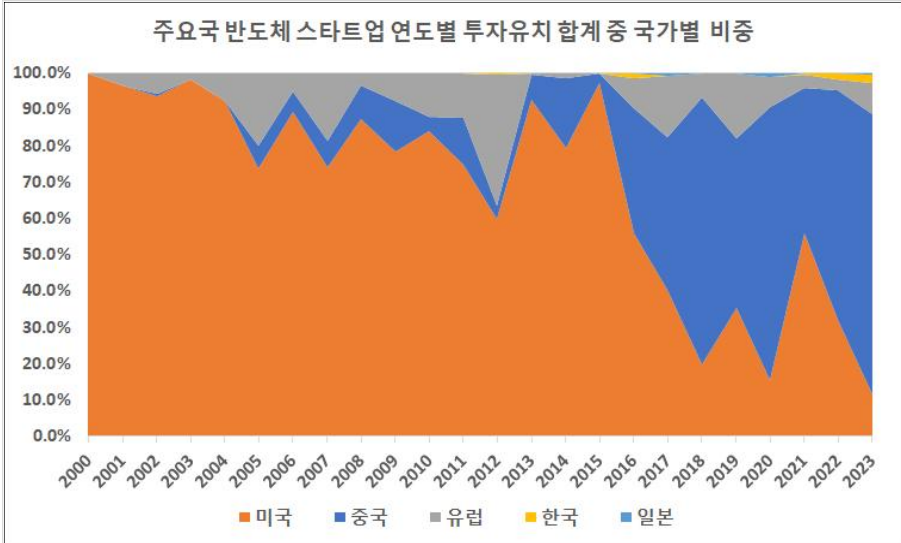


자료: Crunchbase 데이터 가공

미국, 중국, 유럽, 한국, 일본 등 주요 5개 국가·권역의 반도체 스타트업 연도별 투자유치 금액을 비교해 보면 '20년 이후 중국의 확연한 우위가 확연히 드러난다. 중국은 '21년에 투자가 전년 대비 감소하며 미국에 뒤처졌으나 '22년부터 더 큰 격차로 미국을 앞서고 있으며, 반면 미국은 '21년 이후 다시 급감하는 양상이다.

2017년에 중국의 점유율은 42.0%로 미국(40.3%)을 처음으로 넘어섰으며, '21년에 잠시 미국이 앞서기는 했으나 '22년부터 다시 미국을 크게 앞섰다. 2020년 이전까지는 글로벌 반도체 스타트업의 투자유치 규모가 작았기 때문에, 그 이후 글로벌 반도체 스타트업 투자유치에서 중국이 압도하고 있다고 해석할 수 있다.

[그림 3-26] 주요국 반도체 스타트업 연도별 투자유치 합계 추이



자료: Crunchbase 데이터 가공

4. 시사점

가. 자국 중심 반도체 스타트업 생태계

미국, 중국, 유럽, 한국, 일본 등 반도체 산업의 주요 국가 및 권역에서 모두 2020년 이후 반도체 스타트업에 대한 투자가 급증했으며, 이런 흐름은 당분간 지속 전망된다. 2020년 이전 연간 글로벌 반도체 스타트업의 투자유치 총액은 10억 달러를 넘지 못했으나, '20년에 30억 달러 이상으로 급증한 데 이어, '23년 10월 말 기준 60억 달러 초과한 상황이다. 2018년 이후 시작된 미국의 강도 높은 대중 반도체 제재 정책에도 불구하고 반도체 스타트업에 대한 투자가 급격히 증가했다는 점만 놓고 보면 독자 생태계 구축을 위한 기반 강화 투자는 지속될 것으로 전망된다.

유럽 역시 2023년 반도체법 제정 합의를 통해 미국과 중국에 대항하는 자체 반도체 생태계 강화를 표방했고, 실제로 반도체 스타트업 투자가 급증하는 흐름을

보이고 있다. 미국이 전체 반도체 밸류체인에 걸친 완전한 패권 추진하는 데 대한 반작용으로 중국과 유럽 역시 적극 대응에 나설 것으로 보여, 자국 중심 반도체 생태계 구축 경쟁이 격화될 것으로 보이며, 반도체 스타트업에 대한 투자 증가도 지속될 전망이다.

나. 국가 주도의 반도체 생태계 구축 경쟁

중국의 반도체 굴기 선언과 미국의 반도체 패권 선언은 개별 기업이 아닌 국가 차원의 비전이며, 따라서 반도체 생태계 경쟁은 기본적으로 국가 간 패권 경쟁으로 이해되고 있다. 반도체 스타트업에 대한 투자가 미국과 유럽의 경우 국가 기구 중심으로 투자가 이루어지고 있다는 점에서 이를 뒷받침하고 있다. 중국은 자국 VC 중심으로 투자가 이루어지고 있으나, 중국 정부 주도로 조성된 대규모의 빅펀드가 존재하고, 중국 VC에 대한 중국 정부의 영향력을 고려하면, 사실상 미·중·유럽 간 국가 주도의 자국 중심 반도체 생태계 강화 경쟁이 이루어지고 있는 상황으로 볼 수 있다.

다. 핵심 첨단 기술 산업으로서의 반도체

반도체 산업의 국제 분업 체계에 분열이 생기고, 각국 정부 주도의 자국 중심 반도체 생태계 구축 경쟁이 시작된 것은, 반도체 주도권이 미래의 국가 경쟁력에 핵심적 요소가 될 것임을 시사하고 있다. 글로벌 저성장 기조가 지속되는 가운데, 특이점을 넘어섰다는 평가를 받는 AI 기술을 구동하고 AI 기술 역량의 차이를 결정짓게 될 반도체 기술을 선점하고 독점하려는 국가 간 경쟁은 필연적이다. 2020년 이후 투자를 유치한 주요국 반도체 스타트업의 사업 범위에 공통으로 AI가 포함된 것도 이런 흐름을 반영하는 것이며, 반도체 스타트업에 대한 투자는 첨단기술산업에 대한 투자와 맞물리며 계속해서 증가할 것으로 전망된다.

제 4 장 반도체 시장 예측을 위한 모형 개발

제 1 절 반도체 수출 전망 모형

1. 선행연구

가. Aubry & Renou-Maissant(2014)

본 논문은 계량경제학적 모델링 방법을 활용하여 반도체 시장에 가장 적합한 예측 모형을 제시하였다. 다양한 지평(forecasting Horizon)에서 표본 외 예측 정확도(out-of-sample forecasting accuracy)를 평가하기 위해 다양한 기준으로 검정하였다. 분석 결과, VECM이 유용한 정보를 제공할 수 있다면, 베이지안 모형이 일변량 및 비제한적 VAR 모형에 비해 우수한 예측 성능을 보여주었다. 그러나 의사 결정자에게는 쉽게 적용할 수 있는 스펙트럼 방법(spectral method)이 유용한 도구가 될 수 있다. 또한 MS-AR 모형은 대규모 자본 및 연구 투자의 프로그래밍을 조정하기 위해 전환점을 예측하는 데 활용할 수 있음을 보였다.

첨단 기술 기반의 반도체 산업은 혁신의 속도가 빨라 관련 변수들의 경기 변동적 특성이 다른 산업과 많은 차이를 보인다. 기업들이 지속적으로 연구 개발 투자를 늘려온 결과, 점점 더 다양해지는 소비자의 수요를 충족하는 혁신을 이루어왔으며, 기존 제품은 빠르게 대체된다. 반도체 산업은 새로운 장비 및 공장에 대한 투자로 생산 여력이 증가하고 나면, 가격과 투자가 감소하는 경향을 보여왔다.

전 세계적인 반도체 산업이 뚜렷한 경기변동성을 보여주고 있음에도 불구하고, 이에 대한 모델링 및 매출 예측 연구는 거의 이루어지지 않았다(Chow and Choy, 2006; Liu, 2007; Liu and Chi, 2006; Liu et al., 2013). 그 이유는 시장의 구조가 복잡하고 데이터 접근이 어렵기 때문이다. 반도체 대금을 관리하는 확률적 과정을 예측하는 것은 기업 차원뿐만 아니라 거시경제 차원에서도 중요한

데, 실제로 이 하이테크 산업의 의사결정자들은 대규모 자본과 연구 투자의 프로그래밍을 조정하기 위해 미래 동향 평가나 전환점 예측이 필요하기 때문이다. 기업은 새로운 관측치를 발표하는 즉시 예측치를 신속하게 확보하고 업데이트해야 하며, 거시경제 차원에서 반도체가 세계 경제의 건전성을 나타내는 선행지표로 간주된다는 사실(Chow and Choy, 2006)은 정확한 예측을 제공할 필요성을 강조한다. 따라서 이러한 변수는 향후 세계 경제 활동을 예측하는 데 매우 좋은 지표입니다.

본 논문은 글로벌 칩 산업의 진화를 예측하는 최선의 방법에 초점을 두고 있습니다. 특정 유형의 예측 모델의 성공이 모든 시장에 적용될 수 없고 반도체 매출의 순환 행태가 특정 방법론을 필요로 할 수 있기 때문에 분야별 연구가 중요합니다. 예측 정확도를 높이기 위해 적절한 반도체 매출 모델링 기법을 찾는 이유입니다. 지수 평활화 및 ARMA 모델과 같은 표준 일변량 모델이 구현되었습니다. 그런 다음 다양한 복잡성 정도를 가진 다른 기술로 연구가 풍부해졌습니다. 마르코프 스위칭 접근 방식을 사용하여 전자 사이클을 식별하고 날짜를 지정할 수 있기 때문에 여기에서 마르코프 스위칭 자기회귀 모델이 예측을 생성하는 데 유용한 도구가 될 수 있는지 알아보려고 시도했습니다. 또한 예측을 생성하기 위해 스펙트럼 분석을 제안했습니다. 또한 다양한 다변량 모델을 개발하여 설명 요인과 변수 간의 가능한 상호 작용(벡터 자기 회귀 모델(VAR), 베이지안 벡터 자기 회귀 모델 및 벡터 오류 수정 모델)을 통합했습니다. 실제로 다양한 지표가 변동에 영향을 미칠 수 있습니다. 사이클은 산업 생산을 통한 거시 경제 변수에 의해 부분적으로 설명될 수 있습니다. 신규 주문, 투자 또는 재고와 같은 특정 산업 요인은 활동의 다양한 측면(수요, 공급 또는 가격)에 대한 정보를 제공합니다. 기업의 기대에 대한 대응 변수로서 재무 변수도 고려해야 합니다.

여기서 우리의 목표는 결정을 지원하는 데 사용될 수 있는 예측 모델을 제안하는 것이며, 투자자들을 설득하는 논거가 될 수 있습니다. “예측은 결정을 안내하는 데 사용되기 때문에 예측 정확도는 예측 사용자에게 분명히 중요합니다.”(Diebold

and Mariano, 1995). 우리가 특히 전환점 예측 능력에 초점을 맞춘 이유는 그 순간들이 투자 정책에 리듬을 주어야 하기 때문입니다. 우리는 다양한 예측 통계를 사용했습니다. 모형의 표본 외 예측 성능은 평균 제곱근 오차(RMSE), 평균 절대 백분율 오차(MAPE) 및 Theil의 U 통계와 같은 표준 예측 요약 통계에 의해 평가되었습니다. Diebold and Mariton 검정을 사용했습니다. 특히 이러한 움직임이 투자 결정에 관여하는 경우 시리지가 위아래로 움직일지 여부를 평가하는 것도 중요합니다. 그래서 Henriksson and Merton의 1981년 검정을 사용하여 시리즈의 변화 방향을 예측하는 예측 능력에 초점을 맞추어 예측의 정량적 평가에 국한되지 않는 정보를 얻을 수 있었습니다.

이 연구의 독창성은 연구 범위에 있습니다. 본 논문은 1996년 11월부터 2010년 6월까지 금융위기를 포함한 최근 데이터에 대해 반도체 산업에 대한 계량경제학적 예측모형의 성능을 광범위하게 비교하여 제시합니다. 이는 몇 가지 중요한 질문을 다루고 있습니다. 우선, 우리는 경기 사이클을 예측하기 위해 주로 개발된 일부 방법론들이 산업 사이클에 적응할 수 있는지를 검증하고자 노력했습니다. 두 번째로, 이들 모형의 순위는 매우 다양한 기준과 테스트를 기반으로 설정되었습니다. 세 번째로, 모형의 예측 성능입니다. 우리는 1개월에서 12개월까지의 예측 지평에 초점을 맞추었기 때문에 단기와 중기 모두에서 평가를 받았습니다. 논문의 구성은 다음과 같습니다. 2절의 반도체 산업 관련 문헌을 간단히 검토한 후 3절에서 사용된 데이터 세트의 속성을 검토합니다. 4절에서는 예측 비교에 사용된 모형의 사양에 대해 논의합니다. 5절에서는 광범위한 측정값을 사용하여 분석된 표본 외 예측 결과를 제시합니다. 6절에서는 결론을 제시합니다.

본 논문의 목적은 반도체 분야에서 다양한 대안적인 현대 계량 모형화 접근법의 예측 정확도를 검토하는 것이었습니다. 단변량 및 다변량 모형을 모두 사용하여 추정을 수행했습니다. 우리는 순환적 특성과 세계 경제에서 점점 더 중요한 역할을 할 핵심 시장의 요구를 고려했습니다. 우리가 아는 한, 이 연구는 지금까지 반도체 산업에 대한 가장 포괄적인 연구를 나타냅니다: 매우 다양한 모형이 추정되

었고, 1개월에서 12개월 사이의 다양한 예측 지평에서 모델의 예측 정확도를 비교하기 위해 많은 기준이 사용되었습니다. 또한 금융위기를 포함한 최근 데이터에 대한 연구를 진행했습니다.

마르코프 스위칭 모형은 단변량 모형과 관련하여 특히 단기적으로 그 가치를 입증해 왔습니다. 전환점을 예측하는 능력은 의사결정자들에게 귀중한 정보원이 됩니다. 위기의 시기에 사용하기 쉬운 스펙트럼 방법은 연구된 변수의 진화를 단계적으로 추적할 수 있게 합니다. 또한 다변량 접근 방식이 시장 행위자와 연구자에게 전자 사이클을 예측하는 데 도움이 되는 귀중한 정보를 제공한다는 것을 입증했습니다. 다변량 모형은 예측 시계가 증가할수록 더욱 효율적입니다. 따라서 변수 선택은 이들 모형의 예측 정확도에 의해 검증되었습니다. 표준 테스트는 위기 상황에서도 베이지안 VAR 모형을 사용하여 판매 진화를 예측함으로써 얻을 수 있는 이점이 있음을 보여줍니다. 베이지안 VAR 모형의 클래스에서 변수 간의 일반적인 상호 작용을 고려한 BVAR이 가장 정확한 예측을 생성합니다. 그러나 당사 모형을 사용하여 달성한 이득은 금융 위기로 특징지어지는 최근 기간의 임의 보행 프로세스와 비교할 때 크게 나타나지 않습니다. 금융 위기 이전보다 안정적인 기간에 평가된 모형의 예측 정확도는 BVAR 모형의 명확한 우월성을 보여줍니다. 정확도 테스트는 특히 위기 상황에서 주기적인 행동을 예측할 때 한계를 드러냅니다.

이 모델들은 이 하이테크 산업의 의사결정자들이 자본 및 연구 투자 일정을 조정하는 데 유용한 도구를 제공합니다. 이 작업은 산업 부문에 적용되는 산업 주기의 실증적 문헌에도 기여합니다. 우리는 비즈니스 사이클을 위해 개발된 방법론이 산업 사이클에 적용될 수 있고 적용되어야 한다는 점을 강조했습니다. 그러나 모델의 선택과 정확도를 테스트하는 데 사용되는 기준이 시장의 요구에 적합해야 한다는 점은 신중해야 합니다. 반도체 산업에서 행위자들은 단기적으로 시장의 압력에 영향을 받으며, 예를 들어 연구 개발과 관련된 장기적인 지분을 고려해야 합니다. 운영상 효과적인 분석을 위해서는 먼저 전문가가 행동할 수 있는 변수의 선

택에 주의를 기울여야 합니다. 둘째, 구현하기 쉬운 모델에 집중하는 것이 전문가의 반응성과 반응 속도를 보장하는 데 적절할 것으로 보입니다. 셋째, 의사결정 과정을 안내하기 위해서는 모델의 전환점을 예측하는 능력을 고려해야 합니다.

향후 연구에서 유망한 이슈는 전환점을 예측하는 능력에 관한 테스트를 개선하는 것입니다. 최근 일부 저자들은 가능한 불안정성이 존재하는 상황에서 두 경쟁 모델의 표본 외 예측 성능을 비교하기 위해 Diebold 및 Mariano 테스트에서 영감을 받은 새로운 방법을 제안했습니다(Giacomini and Rossi, 2010). 또한 베이저안 VAR 모델이 반도체 매출 진화를 예측할 수 있는 최적의 모델로 나타나면 해당 모델을 구축하기가 어렵습니다. 예를 들어, 최적의 표본 외 예측치를 얻기 위해서는 사전을 표본에 맞게 조정해야 합니다(Ribeiro Ramos, 2003). Doan and Litterman(1992)이 지적한 바와 같이 스펙트럼 분석은 적절한 데이터가 주어지면 빠르고 쉽게 좋은 예측을 할 수 있습니다. 사용의 단순성과 시장에 영향을 미치는 변수의 고려라는 두 가지 이점을 혼합할 수 있는 접근 방식은 의사 결정자에게 매우 유용할 수 있습니다.

나. Chien et al.(2020)

공급망에서 제3자 역할을 하는 반도체 유통업체는 다양한 공급업체로부터 다양한 부품을 구입하여 창고에서 공급업체가 관리하는 재고를 보유한 다수의 전자 제조업체에 재판매하는 한편 수요 불확실성으로 인한 공급 과잉과 부족의 위험을 모두 겪고 있습니다. 하지만 가전 시대의 제품 수명 주기 단축과 첨단 제조용 용량 확장을 위한 긴 리드 타임으로 수요 변동과 공급망 복잡성이 증가하고 있습니다. 본 연구는 반도체 부품 및 모듈에 대한 선도적인 총판의 현실적인 요구에 초점을 맞추어 산업 3.5에 대한 스마트 생산에 힘을 실어주기 위해 해당 수요 패턴을 가진 제품별 최적의 수요 예측 모델을 동적으로 선택하는 심층 강화 학습(RL)을 기반으로 UNISON 프레임워크를 구성하는 것을 목표로 합니다. 딥 러닝 아키텍처와 RL 알고리즘을 통합한 딥 RL은 역동적이고 복잡한 현실 세계에서 성공적인 정책을 학습할 수 있습니다. deep RL의 보상 함수 메커니즘은 수요 불확실성의 부정적인 영향을 줄일 수 있습니다. 제안된 접근 방식의 실제적인 실행 가능성

을 보여주는 검증을 위한 경험적 연구가 수행되었습니다. 실제로 개발된 솔루션은 실제 설정에 있습니다.

전 세계 공급망 네트워크의 기업들은 전 세계적으로 급증하는 고객 제품 수요, 특히 반도체 산업의 공급망 복잡성을 충족시켜야 하는 당면 과제에 직면해 있습니다. 반도체 업계는 웨이퍼 위에 제작되는 트랜지스터의 수가 12개월 또는 24개월마다 두 배씩 증가한다는 무어의 법칙(Moore 1965)에 따라 집적회로(IC)의 특징 크기를 축소하는 기술을 지속적으로 도입해 왔습니다. 자본 집약적이고 경쟁이 치열한 반도체 산업은 수직적으로 통합된 기업에서 집적회로 설계, 웨이퍼 파운드리, 조립 및 최종 테스트를 포함한 수평적 모듈러 파트너의 가상 협업으로 변화했습니다(Chien and Zheng 2012, Chien, Dou, Fu 2018, Chien, Liu, Chuang 2017). 반도체 유통업체의 경우 업스트림 산업에는 IC, 이산 장치, 광학 및 수동 부품 제조업체가 포함되며 다운스트림 고객은 전자 제품 및 주변 장비 제조업체입니다. 공급망의 기업들은 새로운 기술을 개발하고 고객의 요구를 충족시키며 전체 생태계의 성장을 유지하기 위해 협력적이고 경쟁적으로 협력합니다. 제품 수명 주기가 감소하고 공급망의 요구 민첩성이 증가함에 따라 공급망의 다양한 계층 간 공급 과잉 및 부족 위험을 줄이기 위해 수요 예측의 품질을 향상시키는 것이 중요합니다(Chien, Wu, Chien, Wu 2018, Uzsoy, Fowler, Mönch 2018). 실제로, 예측 업데이트는 생산 계획의 성능을 향상시킬 수 있습니다(Jamrus et al. 2015; Ziarnetzky, Mönch and Uzsoy 2018). 용량 계획 및 안전 재고 결정을 위한 결정은 정확한 수요 예측에 크게 의존하고 있습니다(Uzsoy, Fowler, Mönch 2018). 그러나 용량 확장을 위한 긴 리드타임과 다수의 플레이어가 있는 분산형 공급망의 비대칭 정보로 인해 수요 변동이 공급망의 상류에서 증폭되며, 이는 불협 효과(Jacobs, Chase, Lummus 2011, Lee, Padmanabhan, Wang 1997, Vosooghizaji, Taghipour, Canel-Depitre 2019)로 나타난다.

업스트림 공급업체와 다운스트림 구매자는 모두 공급망에서 과도한 재고 또는 고객 서비스 저하의 위험을 감수하기를 꺼리는 반면 중간 주체로서의 유통업체는 재고 보유 지역이 됩니다(스즈키, 마츠오카, 이시이 2017). 이들은 수명 주기가

짧은 제품을 포함하는 중개 중단의 추세에서 살아남기 위해 공급업체 관리 재고 서비스를 제공합니다(Hsu, Lin, Chien 2015; Mills and Cameck 2004). 디커플링 포인트(decoupling point)라고도 불리는 푸시 및 풀 경계 역할을 하는 분배기는 수요 변동성에 대응하고 허용 가능한 범위에서 적시의 서비스 수준을 유지하기 위해 완충재로서 재고와 공급망을 디커플링합니다(Ng and Chung 2008). 그러나 안전 재고를 추가로 보유하게 되면 유통업체, 특히 제품 수명 주기가 단축된 다양한 전자 부품을 보유하고 있는 반도체 유통업체의 경우 재고 위험이 증가하고 상당한 비용이 발생하게 됩니다. 실제로 수요의 매출 예측은 대부분 고객의 피드백에서 수집되며, 고객은 도메인 지식과 비즈니스 통찰력을 기반으로 영업팀을 조정합니다. 그러나 고객은 유통업체에 잠재적 사용에 필요한 충분한 재고를 준비하도록 요구하는 경향이 있기 때문에 신뢰할 수 없는 고객 예측은 유통업체의 수요 예측을 부정확하게 만듭니다(Chang and Yen 2017). 따라서, 유통업체는 재고 비용과 부족 위험을 줄이면서 수요 충족 및 전반적인 공급망 전략을 최적화하기 위한 수요 예측을 개발하는 것이 점점 더 중요해지고 있습니다(Syntetos et al. 2016).

반도체 산업의 짧은 제품 수명 주기를 고려하여 다세대 예측 문제에 대해 개선된 확산 모델(Chen and Chen 2018; Chen, Chen and Peng 2010; Fu and Chen 2019; Kim, Lee and Kim 2005)과 분석 방법(Chen, Yang and Hia 2008; Huang and Tzeng 2008)이 다수 연구되었습니다. 개별 모델로부터의 예측 편향을 방지하기 위해 다양한 하이브리드 접근법이 개발되었습니다(Chen and Wang 2011; Lee and Chiang 2016; Wang and Chen 2019; Zhang 2007). 다양한 수요 예측 모델이 제안되었지만 다음과 같은 몇 가지 제한 사항이 있습니다: (1) 다변량 모형의 모수를 각 제품에 대해 추정해야 합니다. 제품의 수가 증가하는 동안 시간이 많이 소요됩니다. 실제로 대부분의 기업은 필요한 모든 데이터를 수집할 수 있는 능력이 제한되어 있습니다. (2) 수요 데이터는 스마트 생산을 위한 공급망 민첩성을 충족할 수 없는 월 또는 분기 수준으로 집계되었습니다. (3) 유사한 제품 유형과 시계열 특성을 가진 소수의 제품에 대해 실증적 또

는 시뮬레이션적 연구가 수행되었습니다. 하지만 반도체 유통업체는 시장과 경제 상황에 비추어 매우 다른 패턴의 엄청나게 많은 제품을 관리하고 있습니다. 따라서 관련 접근 방식을 통합하고 해당 예측을 동적으로 조정하기 위한 포괄적인 프레임워크가 필요합니다.

강화 학습(RL)은 에이전트가 정교한 환경과의 상호 작용에서 보상을 극대화하여 행동을 선택하는 방법을 학습할 수 있도록 개발되었습니다(Sutton and Barto 2018). RL이 사용되고 있지만(Al-Rawi, Ng 및 Yau 2015; Kaelbling, Littman 및 Moore 1996; Wang and Usher 2005), 현재 문제의 고차원 상태는 RL이 환경의 효율적인 표현을 도출하기 어렵습니다. RL은 완전한 원시 입력을 통해 최적의 제어 정책을 학습하기 때문에 메모리 복잡성과 계산 복잡성은 RL의 한계가 됩니다(Aulkumaran et al. 2017). 컴퓨팅 성능이 향상됨에 따라 신경망 및 딥 러닝은 이러한 장애를 극복하기 위한 효과적인 솔루션을 제공할 수 있습니다. 심층 신경망은 언어 번역(Sutskever, Vinyals, Le 2014), 이미지 인식(He et al. 2016), 음성 이해(Collobert et al. 2011; LeCun, Bengio and Hinton 2015) 등 다양한 작업에서 기계 학습 기법을 획기적으로 개선하고 있습니다. 특히, DRL(Mnih et al. 2013)은 딥 러닝(deep learning)과 컨벌루션(convolutional) RL의 조합 에이전트로, 환경 상태를 통해 수요 불확실성 하에서 순차적인 의사 결정을 최적화하는 방법을 학습할 수 있습니다. DRL은 비디오 게임 플레이(Mnih et al. 2015; Silver et al. 2016), 분자 드 노보 디자인(Olivecrona et al. 2017), 자율 주행(Sallab et al. 2017), 빌딩 에너지 최적화(Mocanu et al. 2018) 및 스마트 생산을 위한 스케줄링(Hsieh, Chang and Chien 2014; Shi et al. 2020)과 같은 다양한 응용 분야에 채용되었습니다.

본 연구는 실제 설정에서 요구를 해결하기 위해 컨벌루션 시계열 예측 방법과 고급 기계 학습 기법을 DRL과 선택 메커니즘으로 통합하여 유통업체의 유연한 수요 계획을 가능하게 하여 인더스트리 3.5의 스마트 생산에 힘을 실어주는 프레임워크를 개발하는 것을 목표로 합니다. 수요 데이터의 높은 변동으로 인해 발생

하는 고차원 상태 문제를 고려하여 DRL 모델을 사용하여 최적의 예측 방법을 선택할 수 있습니다. 실제로 Fildes와 Petropoulos(2015)가 보여주는 것처럼 다양한 제품에 대해서는 적절한 예측 모델의 자동 선택이 필수적입니다. 인터스트리 3.0과 미래의 인터스트리 4.0의 하이브리드 전략으로 인터스트리 3.5를 제안한 것은 인공지능과 빅데이터 분석을 새로운 솔루션의 목표로 사용하면서 스마트 생산을 위한 기본 목표를 해결하기 위함입니다(Chien, Chou, Yu 2016, Jamrus, Wang, Chien 2020, Ku, Chien, Ma 2020, Hsu, Chen, Chen, Chien 2020). 검증을 위해 선도적인 반도체 부품 유통업체를 대상으로 실증적 연구를 수행하였습니다. 분석 및 결과는 예측 정확도를 향상시키고 강력한 예측 품질을 제공하기 위해 제안된 접근법의 실용적인 실행 가능성을 보여주었습니다. 개발된 솔루션은 이 회사의 의사결정 지원 시스템에 내장되어 있습니다.

본 논문의 나머지 부분은 다음과 같이 구성되어 있습니다. 2절에서는 관련 연구와 제안된 프레임워크의 기초를 검토합니다. 3절에서는 수요 예측 모형을 동태적으로 선택하기 위한 제안된 프레임워크를 소개합니다. 4절에서는 검증을 위한 반도체 부품 유통업체에 대한 실증 연구를 제시합니다. 5절에서는 기여와 향후 연구 방향에 대한 논의로 본 연구를 마무리합니다.

1) 모형 선택

현재 문제의 반도체 유통업체들은 급격한 수요 패턴 및 비즈니스 환경의 변화로 인해 선택된 예측 모델을 유지할 수 없기 때문에 다양한 제품을 대량으로 예측하는 것이 큰 과제입니다. 따라서 경쟁력을 유지하기 위해서는 다양한 예측 방법을 통합할 수 있는 예측 선택 메커니즘의 설계가 중요합니다(Montgomery, Jennings, Kulahci 2015). 예측 모델의 선택을 설계하고 개선하기 위한 방법을 조사하기 위해 많은 연구가 수행되었습니다(예: Tashman 2000; Poler and Mula 2011). 롤링 오리진 프로세스에 기반한 표본 외 테스트를 통해 성능을 평가하기 위한 선택 메커니즘을 개발했습니다. Adya et al.(2001)은 도메인 지식을 기반으로 자동으로 모델 선택을 가능하게 하고 규칙을 실행하는 6가지 기능을 식

별하는 방법을 개발했습니다. Wang, Smith-Miles 및 Hyndman(2009)은 메타 학습 절차로서 지도학습과 비지도학습을 모두 통합하여 선택규칙을 도출하였다. Petropoulos et al. 2014; Wang, Smith-Miles 및 Hyndman 2009 등은 시계열 특성을 측정하고 예측 정확도를 높이기 위한 방법을 개발하기 위해 많은 연구가 수행되었습니다. Villegas, Pedregal 및 Trapero(2018)는 Akaike 정보 기준과 Schwarz Bayesian 기준을 SVM과 통합한 자동화된 모델 선택 방법을 제안했습니다.

또한 예측 모델 조합을 위한 앙상블 방법은 모델링 불확실성을 최소화하기 위해 더 나은 성능을 가질 수 있습니다(Barrow and Kourentzes 2016, Fildes and Petropoulos 2015, Kourentzes, Barrow and Petropoulos 2019). 신경망(Wang et al. 2018), LSTM(Choi and Lee 2018), 다중 목적 접근법(Bacci et al. 2019) 등 다양한 기법에 대한 다수의 예측 조합 기법이 제시되었습니다(Pawlikowski and Chorowska 2019). 특히 단순 평균 앙상블 전략은 견고성과 관련하여 더 나은 성능을 갖는 것으로 입증되었습니다(Blanc and Setzer 2016; Brighton and Gigerenzer 2015; Grafe 2015; Green and Armstrong 2015). 그러나 예측 모델의 선정에는 한계가 있습니다. 특히 규칙 기반 접근법은 과적합되거나 정의가 불충분한 문제가 있을 수 있습니다. 특징 기반 선택은 특징 공식화 및 선택 단계에 따라 시간이 많이 소요되는 의존도가 높습니다. 반도체 디스트리뷰터의 제품 수량과 종류가 상당히 크기 때문에 필요한 모든 규칙을 정의하기 어려워 규칙 기반 방식을 적용하기에 적합하지 않습니다. 또한 반도체 산업에서 수요 변동성이 높은 만큼 다양한 특징으로 인해 특징 공식이 복잡하거나 부족하여 추가 예측 능력에 영향을 미칠 것입니다. 앙상블 모델의 경우 각 예측기에 할당된 가중치가 각 단계에서 업데이트되지 않기 때문에 제품 수명 주기가 단축된 첨단 제품과 같이 급변하는 시계열 패턴의 제품을 예측하는 데 한계가 있을 수 있습니다. 본 연구에서는 DRL 모델을 자동으로 선택하고 환경과의 상호작용을 통해 동적으로 예측 모델을 갱신할 수 있으므로 새로운 DRL 선택 프레임워크를 개발합니다.

2) 수요 예측

긴 리드타임 용량 확장, 제품 수명 주기 단축 및 반도체 공급망의 만연한 불확실성(Mönch, Uzsoy, Flower 2018)으로 인해 수요 예측은 스마트 생산에서 중요한 역할을 합니다. Norton and Bass(1987)는 세대 대체 및 반복 구매 매개변수를 고려하여 각 세대의 전체 라이프사이클 내에서 수요를 예측하는 다세대 확산 모델을 개발했습니다. 또한 소비자 행동(Kim, Lee, Kim 2005), 계절적 요인의 영향, 시장 성장률, 가격, 기술 대체성(Chien, Chen, Peng 2010), 도메인 전문가의 판단(Chen and Chien 2018) 등 다양한 요인을 고려하여 다수의 연구가 다세대 확산 모형을 수정하였습니다. 또한 Huang and Tzeng(2008)은 다세대 제품 수명 예측을 위한 2단계 퍼지 조각별 회귀 분석을 개발했습니다. Chen, Yang, and Hia(2008)는 다양한 기술 세대의 수요를 예측하고 예측 오차를 줄이기 위해 동적 가중 최소 제곱 추정을 집계 방법으로 개발했습니다.

단일 모형으로부터 예측 편향을 방지하기 위한 하이브리드 접근법이 개발되었습니다. 특히, Zhang(2007)은 두 개의 개별 접근법의 가중치 최적화를 통한 조합 예측 방법으로 벡터 일반화된 자기회귀 조건부 이분산성(GARCH) 모형을 개발했습니다. Chen과 Wang(2011)은 반도체 제조업에서 장부 대 청구서 비율을 예측하기 위한 하이브리드 퍼지 선형 회귀 및 역전파 신경 접근법을 제안했습니다. Lee와 Chiang(2016)은 용량-수요 불일치 문제의 손실을 최소화하기 위해 박막 트랜지스터 액정표시장치(TFT-LCD) 제품의 총 수요 예측과 용량 결정을 포함한 2단계 프레임워크를 제안했습니다. 본 연구에서는 최적의 용량계획을 수립하기 위해 선형회귀분석, 자기회귀분석, 신경망분석 등 세 가지 예측방법을 서로 다른 시나리오로 적용합니다. Wang and Chen(2019)은 추가적인 재무추정을 위해 재무정보와 통합된 수요-풀(demand-pull) 프레임워크를 제시했습니다. 자기회귀적 통합이동평균(ARIMA), 동태적 ARIMA, 벡터자기회귀(VAR), 다변량적응회귀 스플라인(MARS), 랜덤포레스트(RF), 지지벡터회귀(SVR) 등 시계열 및 회귀분석 방법을 통해 수요를 추정하였습니다. 매출액 추정은 관측을 통해 가장 적합한 모델을 선정하였습니다.

다수의 선행 연구에서 반도체 산업의 수요 충족 문제를 다루었지만 다음과 같은 단점과 한계로 어려움을 겪고 있습니다. 확산 모형의 경우 서로 다른 제품의 수요를 예측하기 위해 각 모수를 추정해야 합니다. 시간이 많이 걸리고 일부 회사는 다변량 예측에 필요한 모든 파라미터를 수집할 수 없습니다. 또한 기존 연구의 수요 데이터는 월 또는 분기 수준에서 다소 큰 시간 세분성으로 집계되었습니다. 입도(granularity)가 작을수록 더 정밀하고 세밀한 생산관리가 가능하도록 고려해야 합니다. 더욱이 위의 어느 것도 10개 이상의 제품에 대한 예측 문제를 다루지 않고 상품 유형 및 시계열 특성이 유사한 다소 적은 수의 제품에 대한 실증적 또는 모의적 연구만을 수행하였습니다. 하지만 반도체 유통업체의 경우 상당한 종류의 반도체 제품이 존재하기 때문에 다양한 종류의 제품에 대한 미래 수요를 예측할 수 있는 유연한 모델이 필요합니다. 위에서 언급한 연구 격차를 해결하고 더 높은 공급망 복원력을 달성하기 위해 주간 수준의 일변량 DRL 기반 프레임워크를 제안하여 서로 다른 시간 지평에서 각 제품에 대한 적절한 모델을 동적으로 결정합니다.

수요 예측 편향은 비효율적인 반도체 공급망을 야기할 것입니다. 반도체 부품 유통업체의 경우 상당히 많은 수의 제품에 대한 수요 예측 품질을 향상시키는 것이 중요합니다. 그러나 위탁에는 상당히 많은 수의 제품이 포함될 수 있으며 제품 시계열의 패턴은 수시로 다를 수 있습니다. 본 연구는 유연한 의사결정을 가능하게 하기 위해 RL 모델과 도메인 지식을 통합한 UNISON 예측 모델 선정 프레임워크를 개발하였습니다. 제안된 프레임워크를 검증하기 위해 대만의 선도적인 반도체 부품 유통업체에서 스마트 공급망의 수요예측을 위한 실증연구를 수행하였습니다. 모델의 분석 정보와 의사 결정자의 비즈니스 통찰력을 결합하여 결과는 예측 정확도를 크게 향상시키기 위해 강력한 의사 결정 품질로 예측 오류를 줄이기 위해 제안된 접근 방식의 실용적인 실행 가능성을 보여주었습니다. 실제로 개발된 솔루션은 이 회사에서 시행하는 의사결정 지원 시스템에 내장되어 있습니다.

향후 연구는 예측 결과의 정확성과 일반화 가능성을 향상시키기 위해 내부 및

외부 정보에 대한 더 많은 데이터를 가진 빅 데이터 분석을 포함한 다른 접근 방식을 개발하기 위해 수행될 수 있습니다. 더 많은 연구를 통해 비즈니스 통찰력과 도메인 지식을 통합하기 위해 조정 메커니즘을 자체 조정하는 대체 인공지능 알고리즘을 개발할 수 있습니다. 또한 도메인에 대한 다양한 수요 예측 방법론을 조사하여 다양한 맥락에서 산업 3.5에 대한 스마트 생산에 힘을 실어주기 위해 제안된 접근 방식을 구현함으로써 더 많은 통찰력을 얻을 수 있으며, 제안된 프레임워크가 다양한 상황에 비추어 예측 정확도를 향상시키기 위해 효과적으로 적용될 수 있는 상황을 식별할 수 있습니다.

제 2 절 반도체 시장 불확실성 지수 개발

1. 추진 전략 및 방법

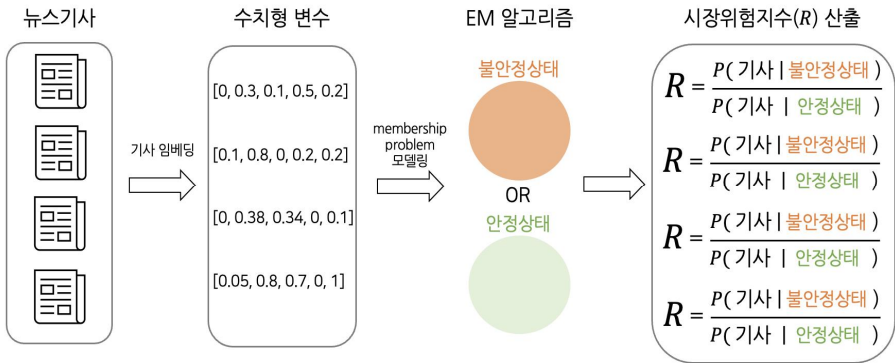
본절에서는 비정형 뉴스 기사 데이터를 활용하여 반도체 시장 위험(불확실성) 지수를 개발하는 것을 목표로 한다. 이는 반도체 시장의 위험이 증가하고 불확실성이 높을수록 관련 기사들이 시장의 상황을 반영한다는 가정에서 출발한다. 이를 위해 먼저 시장위험 및 불확실성과 관련된 뉴스를 어떻게 식별할지 그 기준을 정립하는 것이 필요하다.

본 연구는 반도체 시장의 위험이 크고 불확실한 상황을 불안정상태(abnormal period), 시장에 영향을 주는 별다른 이슈가 없는 상황을 안정상태(normal period)로 보았을 때, 불안정상태와 안정상태일 때 뉴스 기사에서 나오는 어휘 패턴이 다르다고 가정하였다. 예를 들어, 불안정상태 상황에서는 위험, 불안, 변동성과 같은 단어 혹은 부정적이거나 우려를 표현하는 예측성 어조가 많이 나온다면, 안정상태에서는 성장, 안정, 혹은 긍정적이거나 확신에 찬 어휘 패턴이 있는 뉴스 기사가 많이 나올 수 있는 것이다. 이를 토대로, 불안정상태에서 나오는 어휘 패턴들을 분석하여, 현재 상태가 불안정상태일 때 해당 어휘 패턴이 나올 확률과 안정상태일 때 해당 어휘 패턴이 나올 확률을 계산한다. 이후, 두 확률의 비를

지수화하여 현재의 상대적인 위험도를 파악하고자 한다.

뉴스 데이터와 같은 비정형 텍스트 데이터는 그 자체를 모형의 input으로 직접 사용할 수 없기에, 이를 수치형 데이터로 전환하는 작업이 선행되어야 한다. 이를 위해서 (i) 수집된 뉴스 기사를 전처리하는 작업, (ii) 반도체 시장 위험을 위험 요인(source)에 따라 분류하는 작업, (iii) 각 위험 요인에 매칭되는 어휘사전(word cluster)을 구성하는 작업, (iv) 기사가 해당 어휘사전의 단어를 어느정도 포함하고 있는지 확인하여 뉴스 기사를 수치형 데이터로 변환하는 작업이 순차적으로 필요하다. 이 과정을 바탕으로 (v) 해당 어휘 패턴 값을 가진 뉴스 기사가 안정상태 혹은 불안정상태에 속할 확률을 데이터 모형으로 표현하고, (vi) 이 모형의 결과를 활용하여 반도체 시장위험 지수를 산출한다. [그림 4-1] 은 이러한 일련의 작업을 도식화하여 보여주고 있다. 각 작업에 대한 자세한 내용은 이후 절에서 소개한다.

[그림 4-1] 비정형 텍스트 데이터를 활용한 시장위험지수



2. 선행연구

최근 텍스트 분석기술이 발전함에 따라 대규모 뉴스 데이터를 통해 경제 상태를 더 정교하게 분석하는 것이 가능해졌다. Chen et al.(2023)은 금융위기를 실시간

으로 탐지하기 위하여 텍스트 데이터와 기계학습을 사용하였다. 뉴스 데이터와 OECD 및 IMF의 공식 보고서를 활용하여 SVM(support vector machine) 모델과 랜덤포레스트(random forest) 모델을 통해 텍스트를 학습시키고, 이를 통해 금융위기의 심각성을 더 정교하게 식별하였다. 특히 연구 결과에 따르면 실시간 뉴스 데이터가 OECD 공식 보고서보다 예측에 있어 상대적으로 더욱 우수한 성능을 보였다. 이는 뉴스 데이터가 상대적으로 더 시장 상황을 효과적으로 반영할 수 있다는 점을 시사한다. 또한, Kalamara et al.(2022) 연구에서는 영국의 주요 신문 데이터에 뉴럴 네트워크와 리지 회귀와 같은 기계학습 모델을 적용하여 뉴스 데이터가 GDP, 인플레이션, 실업률을 포함한 주요 거시경제 변수에 대한 예측을 개선시킬 수 있다는 사실을 발견하였다. 이는 뉴스 데이터가 특히 위험성이 있는 시기를 예측하는 데 중요한 역할을 할 수 있음을 보여준다.

하지만 위와 같은 기계학습은 모두 지도학습(supervised learning)으로, 입력 데이터와 함께 그에 상응하는 정답(레이블)을 모델에 제공해야만 학습을 시킬 수 있다. 반면에 어휘사전을 이용한 비지도학습(unsupervised learning)의 경우 텍스트 데이터에서 특정 단어의 빈도나 존재 여부를 기반으로 분석을 수행하기 때문에 입력 데이터에 대한 명시적인 정답을 주지 않아도 모델이 데이터의 구조나 패턴을 자체적으로 찾아낼 수 있다. 모델링 소요 기간과 레이블링 작업의 비용을 고려하면, 지도학습 보다는 어휘사전을 활용한 비지도학습 형태로 반도체 시장 위험 지수를 개발하는 것이 효과적일 수 있다.

Tetlock et al.(2008)은 하버드에서 개발한 IV-4 어휘사전을 활용하여 뉴스 데이터를 통해 기업의 수익과 주식 가격을 예측하는 연구를 진행하였다. 이 연구에서는 각 기업에 대한 뉴스에서 부정적인 의미가 있는 단어들의 비율을 계산하여 기업의 경제 상황을 예측하였다. 그 결과로 기업과 관련한 뉴스에서 부정적인 단어가 많을수록 낮은 기업수익을 예측하는 것을 발견하였다.

그러나 Loughran et al.(2011)은 기존에 만들어진 어휘사전을 금융 및 회계 문서에 적용하는 것은 감정 표현을 측정하는 데 있어서 적합하지 않을 수 있음을 지적하였다. 이들은 금융의 맥락에서는 부정적인 의미가 있는 단어들

IV-4 단어 목록에서는 부정적으로 분류되지 않는다는 사실을 발견하였다. 예를 들어 ‘감가상각’, ‘부채’와 같은 용어들은 금융의 맥락에서는 부정적이지만, IV-4 단어 목록에 의하면 부정적으로 식별되지 않는다는 것이다. 따라서 특정 분야에 대한 텍스트를 분석할 때 그 분야에 맞는 단어 목록을 개발하는 것이 중요함을 강조하였다.

이에 최근에는 경제 분야에 특화된 새로운 어휘사전의 개발이 활발히 진행되고 있다. 일례로 Barbaglia et al.(2022)에서는 수백만 개의 뉴스 기사와 유럽중앙은행 문서를 포함하는 대규모 말뭉치를 기반으로 경제 및 금융 분야의 텍스트 분석에 적합한 새로운 어휘사전인 EL(economic lexicon)을 제안하였다.

반면에 자신의 연구에 맞는 어휘사전을 직접 개발하여 사용하는 연구 역시 활발하다. 대표적으로, Baker et al.(2016) 연구에서는 뉴스 데이터를 통해 경제정책 불확실성(EPU, economic policy uncertainty) 지수를 개발하기 위해 ‘경제’, ‘정책’ 그리고 ‘불확실성’에 대한 어휘사전을 새롭게 구축하여 각 사전 속 단어를 포함하는 신문 기사의 빈도수를 바탕으로 경제정책불확실성을 측정했다. 이 EPU 지표는 전 세계적으로 사용되고 있으며, 다양한 국가와 지역의 경제정책불확실성의 변화를 실시간으로 모니터링하고 분석하는 데 중요한 역할을 하고 있다. 최근 Wu(2023)의 연구에서는 기업의 분기별 수익 컨퍼런스에서의 토론 내용을 바탕으로 공급망 관련 어휘사전을 새롭게 구축하고, 단어 빈도 접근 방식을 사용하여 공급망 위험 노출에 대한 위험 수준의 측정치인 SCRisk를 도출하기도 하였다.

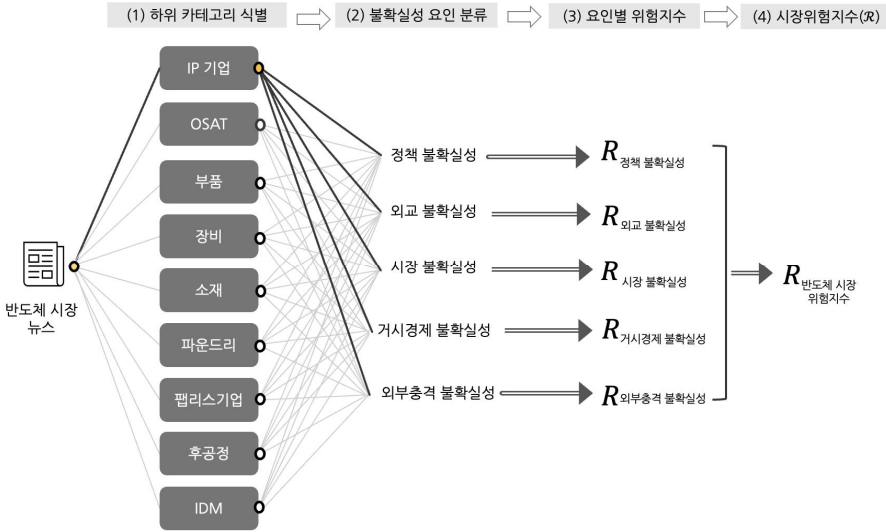
본 연구에서는 연구 기간과 레이블링 비용을 고려하여 어휘사전에 기반한 비지도 학습으로 반도체 시장 위험지수를 개발한다. 반도체 시장에 특화된 어휘사전을 구축한 뒤 이를 통해 반도체 시장의 불확실성을 분석하고자 한다. 이는 기존 연구에서 다루지 않은 새로운 영역을 탐색하는 것으로, 반도체 산업의 변동성을 새로운 시각으로 이해하는 데 이바지할 것으로 기대한다.

3. 지수 구축 방법

본 연구에서는 반도체 시장위험 및 불확실성이 안정상태와 불안정상태의 선형 결합으로 표현될 수 있다고 가정하였다. 이를 토대로, 큰 이슈가 없어 반도체 시장위험 및 불확실성이 크지 않은 상태(안정상태)와 미·중 무역 갈등처럼 시장에 큰 영향을 줄 수 있는 이벤트가 발생하여 반도체 시장의 전망이 불확실한 상태(불안정상태)가 결합되어 있는 반도체 시장에서의 위험 수준을 평가할 수 있는 방법론을 연구하였다.

반도체 시장과 관련된 뉴스 기사의 어휘 분포가 안정상태와 불안정상태에서 서로 다르다는 점에 착안하여, 이를 토대로 반도체 시장의 불확실성(위험) 수준을 평가하고자 한다. 이를 위해, 안정상태와 불안정상태를 구별하는 방안으로 비정형 데이터를 활용한 비지도학습 모델을 개발하였다. 비지도학습이란 미리 정해진 정답 없이 데이터를 학습하여 데이터 내에 숨겨진 특성이나 구조를 발견하는 방식이다. 반도체 시장과 관련된 뉴스 데이터는 다양한 의미를 내포하고 있지만, 이러한 뉴스가 시장의 불확실성과 관련되어 있는지에 대한 명확한 정답은 존재하지 않는다. 따라서, 뉴스 데이터가 시장의 불확실성을 나타내는지 여부를 판단하기 위해서는 누군가가 라벨링을 해야 하지만, 명확한 기준이 없는 상황에서는 비지도 학습 방법이 효과적이다. 이 연구에서는 이러한 명시적인 정답이 없는 상황에서 텍스트 데이터가 가진 내재된 특성을 이해하고 활용하기 위해 비지도학습 방식을 채택했다.

[그림 4-2] 반도체 시장 불확실성 지수 산출 과정 모식도



구체적인 반도체 시장위험 지수와 산출 과정은 [그림 4-2]에 요약된 형태를 기본으로 한다. 첫째, 반도체 시장 위험지수를 산업 카테고리 단위에서 산출하기 위하여, 개별 반도체 시장 뉴스가 어떤 카테고리에 속하는지 IP 기업, 부품, 소재, IDM 등으로 분류하는 과정이 필요하다. 이는 ㈜와이즈넷에서 진행하였으며, 하위 카테고리는 IP 기업, OSAT, 부품, 소재, 장비, 파운드리, 팹리스 기업, 후공정, IDM으로 총 9가지로 분류하였다.

두 번째로, 개별 뉴스를 반도체 시장 불확실성 요인별로 요약하는 작업이 필요하다. 본 연구에서는 반도체 시장의 불확실성을 그 요인에 따라 정책·외교·시장·거시경제·외부 충격으로 구분하였다. 예를 들면, 반도체 시장의 위험을 정책적인 요인이 반도체 시장의 불확실성에 미치는 경우, 외교적인 요인이 반도체 시장의 불확실성에 미치는 경우 등으로 구분한 것이다. 이를 구분하기 위해서는 각각의 뉴스가 다섯 가지 불확실성 요인별 정보를 얼마나 포함하고 있는지를 계량화하는 작업이 필요하다. 이는 해당 뉴스 기사가 불확실성 요인별로 어휘사전의 단어를 얼마나 많이 포함하고 있는지를 정규화하여 표현하였다.

최종적으로, 불확실성 요인별로 계량화된 기사에 EM 알고리즘을 적용하여 상대 불확실성(위험) 수준을 계산한다. 상대 불확실성 수준은 반도체 시장이 안정적일 때 나올 수 있는 어휘 패턴에 대한 확률과 불안정할 때 나올 수 있는 어휘 패턴에 대한 확률의 상대적인 비로 표현하였다. 반도체 시장 위험지수 산출과 관련된 세부적인 내용은 아래에 순차적으로 기술하였다.

본 연구의 최종 목표인 반도체 시장 위험지수 산출을 위해 하기 2가지 사항을 가정하여 연구를 진행한다.

(A1) 반도체 시장위험 정도가 높을 때(H , 불안정상태)와 낮을 때(L , 안정상태) 작성되는 온라인 뉴스 기사(w , 어휘 벡터)의 패턴이 달라진다.

$$P(w | H) \neq P(w | L) \quad [\text{식 4-1}]$$

(A2) 반도체 시장위험 수준에 따라서 발현되는 어휘 벡터의 패턴은 시점에 따라 변화하지 않는다. 즉, 모든 t 에 대하여 [식 2-2]가 성립한다.

$$\begin{aligned} P(w | H_t) &= P(w | H_0) \\ P(w | L_t) &= P(w | L_0) \end{aligned} \quad [\text{식 4-2}]$$

위 가정은 바탕으로 하기의 알고리즘을 순차적으로 진행하여 최종적으로 반도체 시장 위험지수를 산출한다.

가. 기사 임베딩

1) 비정형데이터 전처리

본 연구는 가정(A1)과 같이 온라인뉴스의 어휘 패턴이 시장위험 정도에 따라 다르게 나타난다는 점에 착안하여 진행하므로, 뉴스 기사의 핵심 키워드 및 어휘 패턴을 분석하는 것이 중요하다. 이에, 해석이 쉽고 직관적이지만 간단한 어휘 기반 접근 방식을 채택했다. 어휘 기반의 접근을 위한 뉴스 기사의 형태소 분석을 위해서 Kiwi 형태소 분석기를 사용하였다.²⁴⁾ [그림 4-3]은 간단한 예시 문장에 명사 추출기(noun extractor)를 적용하여 각 형태소 분석기의 품질을 확인한 것이다.

[그림 4-3] 예시문장을 통한 형태소 분석기 비교

예시 문장: 이번에 개발한 계속기로 3 나노미터 두세의 이산화규소 박막시면을 0.082 나노미터 정밀도로 측정하는 데 성공했다.

	time	nouns
Mecab	0.00045 sec	[이번, 개발, 계속, 기로, 나노미터, 세, 이산화규소, 박막, 시면, 나노미터, 정밀도, 측정, 데, 성공]
Komoran	0.00210 sec	[이번, 개발, 계속, 기로, 나노미터, 세, 이산, 화, 규소, 박막, 시, 나노미터, 정밀도, 측정, 데, 성공]
Kkma	0.01389 sec	[이번, 개발, 계속, 계속기로, 기로, 3, 나노, 나노미터, 미터, 세의, 이산화, 이산화규소, 규소, 박막, 박막시면을, 시, 면을, 0.082, 정밀, 정밀도로, 도로, 측정, 데, 성공]
Okt	0.00743 sec	[이번, 개발, 계속, 기, 나노미터, 이산화규소, 박막, 면, 나노미터, 정밀도, 측정, 데]
Hannanum	0.86967 sec	[이번, 개발, 계속기, 3, 나노미터, 두세, 이산화규소, 박막시, 0.082, 나노미터, 정밀도, 측정, 데, 성공]
Kiwi	0.00155 sec	[이번, 개발, 계속, 기, 이산화, 규소, 박막, 면, 정밀도, 측정, 성공]

그 결과, 중복된 단어가 없으며, 분석 시간이 상대적으로 짧고, 위의 실험 결과를 참고했을 때 가장 효율적이라고 판단되는 Kiwi 형태소 분석기의 명사 추출기를 사용하기로 결정했다. 형태소 분석에 있어 텍스트 데이터를 명사만으로 추출할 때보다 다른 품사를 추가 했을 때 더 정확한 어휘사전을 구축할 수 있을지 확인하기 위해 형용사, 부사, 동사 등의 다양한 품사를 추가하여 키워드를 추출했다. 그러나, 실제로 형용사, 부사 등의 품사를 추가하여 진행하였을 때 키워드로 보기 어려운 “속보호주의”, “화무역”, “성상황무역” 등의 단어들이 [그림 4-4]와 같이 상당수 추출되는 것을 확인할 수 있었다. 이에 핵심 키워드를 선택하는 데 있어 다른 품사를 추가하는 것보다는 명사만으로 진행하는 것이 가장 적합할 것으로 판단하고 Kiwi의 명사 추출기를 사용하여 진행하였다.

24) Kiwi는 텍스트 분석 속도, 모호성 해소의 정확도, 문장분리의 정확도 등에서 다른 형태소 분석기보다 우수한 성능을 보임

[그림 4-4] 다양한 품사를 기반으로 구축한 키워드

```

1 model.wv.most_similar('무역정책변화',topn=10)
✓ 0.5s

(['무역순위역대', 0.8065393567085266),
 ('속보호주의', 0.781361997127533),
 ('화무역', 0.7751139998435974),
 ('성상황무역', 0.7650607824325562),
 ('미래무역분야', 0.7600175142288208),
 ('무역시장다변', 0.7571958303451538),
 ('경신무역', 0.7530136704444885),
 ('무역블랙리스트', 0.74796462059021),
 ('성무역', 0.7461166381835938),
 ('확대기대바이든', 0.745688796043396)]

```

추가로, 뉴스 기사의 내용과 관련 없는 언론 매체명이나 기사 내에서 불필요하게 반복되는 문장의 단어들은 불용어로 처리하였다. 불용어의 예시로는 [“무단”, “배포”, “연합”, “데일리”, “구독”, “기자”, “특파원”, “뉴스”, “신문”, “기사”, “연합뉴스”, “한경”, “헤럴드”, “세계일보”, “한국경제”, “아시아 경제”, “뉴시스”, “투데이”, “핫뉴스”, “한국 경제”] 등이 포함되어 있다. 또한, 의미 있는 단어를 찾기 위해 두글자 이상인 명사 단어만을 선정하는 기준을 적용했으며, 키워드가 하나일 때와 비교해 맥락을 더 잘 드러내는 키워드를 생성하기 위해 bigram과 trigram 또한 단어사전에 포함시켰다. bigram은 2개의 연속적인 단어의 나열, trigram은 3개의 연속적인 단어의 나열을 의미하는데, 뉴스 기사에서 bigram, trigram 속의 각각의 단어들이 개별적으로 나오는 것을 볼 때보다, 연속적으로 함께 나오는 키워드를 파악할 때 전체적인 맥락이 잘 표현되는 것으로 판단하였기 때문이다.

2) 불확실성 요인별 어휘사전 구축

본 연구에서는 <표 4-1>처럼 5개의 불확실성 요인을 선정하고, 각각의 불확실성 요인과 관련한 seed word를 선택하였다.

〈표 4-1〉 반도체 시장 불확실성 소스 및 주요 키워드

불확실성 요인	seed word
정책 불확실성	금리조절, 규제변화, 규제강화, 정책불확실, 거래규제, 산업규제, 긴축정책
외교 불확실성	무역갈등, 수출규제, 무역불확실성, 양국신경전, 관세부과, 보복조치, 외교분쟁, 외교갈등, 외교긴장, 외교문제, 수출 규제, 양국갈등, 줄타기외교
반도체 시장 불확실성	공급타격, 반도체부족, 부족예상, 생산부족, 생산위축, 생산타격, 수급불안, 수급타격, 수요증가, 시장동향
거시경제 불확실성	경기침체, 소비위축, 인플레이션, 유가변동, 산업위축, 고금리, 불황, 실업증가, 금융위기
외부충격 불확실성	전쟁, 자연재해, 감염병, 기술유출, 사고발생, 탄소중립, 인력부족, 산재

(가) Burstiness를 활용한 어휘사전의 seed word 선정

소스별 seed word를 선정하기 위해 이머징 이슈(emerging issue) 탐색 방법을 사용했다. 우선 불확실성 키워드인 "불안", "우려", "리스크", "불확실", "위험", "위기"를 포함하는 기사 선별하고, 해당 기사들을 구성하는 어휘사전의 단어 중에서 Burstiness 시계열이 특정 임계값을 초과하는 키워드를 추출했다. 이후, 계층적 군집화를 통해 유사한 키워드끼리 클러스터링 한 후, 해당 클러스터의 단어를 바탕으로 seed word를 구성하는 방식으로 진행하였다.

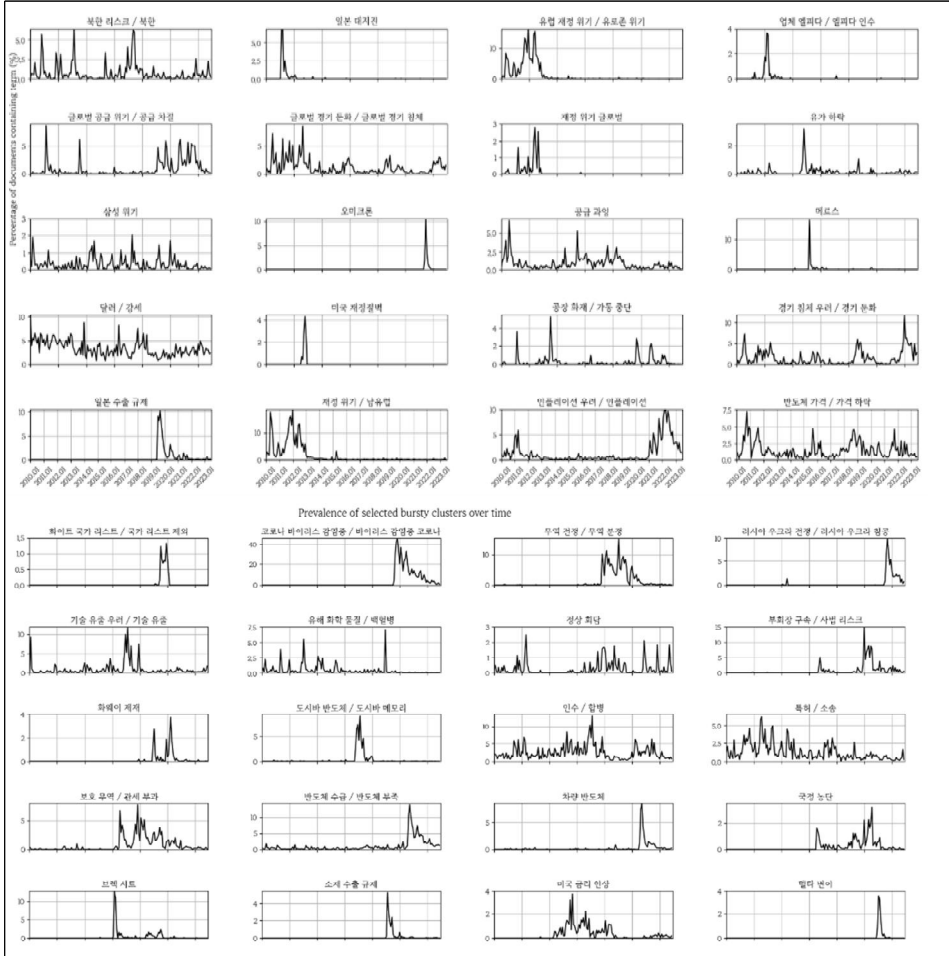
본 연구에서 사용한 Burstiness는 장기적 추세 대비 단기간에 특정 단어가 얼마나 자주 언급되는지를 나타내는 지표이다. 단순 단어 빈도(word frequency)는 특정 단어가 얼마나 등장했는지를 나타내는 반면, Burstiness는 일반적인 상황을 고려했을 때 얼마나 그 단어 빈도가 이상치인지를 나타내는 지표이다. 이러한 특성 때문에 Burstiness는 이머징 이슈를 탐지할 때 주로 사용된다. 본 연구에서는 불확실한 상황을 이머징 이슈로 인한 상태로 가정하였기에, 이를 잘 식별해내기 위하여 Burstiness 방법을 활용하는 것이 적절하다고 판단하였다. Burstiness를 구하는 식과 관련된 자세한 내용은 He and Parker(2010)을 참조하길 바란다.

Burstiness 지표를 활용하여 이머징 이슈를 감지하고, 반도체 시장의 리스크에 영향을 미치는 주요 키워드들을 확인하였다. 참고한 키워드 목록 중 40개를 예시로 들면 [그림 4-5]와 같다. 이러한 키워드 결과를 참여 연구원이 필터링하여 반도체 시장의 불확실성 요인을 총 5가지로 분류하고, 각 요인과 관련된 seed word를 선정하였다.

(나) Word2Vec 알고리즘을 통한 어휘사전 구축

위 Burstiness 지표를 활용하여 seed word가 될 수 있는 단어들을 각 요인별로 선정한 후, 모든 기사에 Word2Vec 알고리즘을 적용하여 어휘사전에 속한 모든 단어에 대한 임베딩을 수행하였다. Word2Vec 이란 텍스트에서 단어를 벡터 형태로 변환하는 임베딩 기법으로, 단어간 유사도를 반영할 수 있도록 단어의 의미를 벡터화하는 방법이다. 임베딩을 마친 후, 위에서 선정한 seed word에 대해 코사인 유사도(cosine similarity)가 높은 상위 20개의 단어를 선택하였다. 다만, 이렇게 추출된 단어 중에는 keyword로 특정하기 어려운 단어가 존재하였으므로, 참여 연구원이 추출된 단어 중 관련 요인별 유사도가 없는 것으로 판단한 단어는 제거한 후 각 요인별로 총 80개의 단어를 최종 선정하여 각 위험요인별 어휘사전을 구축하였다.

[그림 4-5] Burstiness 활용한 키워드 도출 예시



3) 텍스트 기사를 수치형 특성 변수로 요약

전처리 과정을 거친 뉴스 기사가 불확실성 요인별 어휘사전에 있는 키워드를 얼마나 포함하고 있는지를 확인하여, 이를 바탕으로 수치화한 특성 변수를 생성하였다. 특성 변수는 기사 내의 불확실성과 관련된 단어의 개수를 전체 기사의 길이로 정규화하기 위해서, 기사의 단어 중 어휘사전 내에 포함되어 있는 단어의 수를 해당 기사의 전체 단어 수로 나누는 방식으로 생성하였다. 이러한 정규화 과정을 통

해 기사의 길이가 단순히 긴 이유로 어휘사전의 단어 개수를 많이 포함하고 있는 경우를 잡아내고자 하였다. 이렇게 생성된 feature 변수는 0~1사이의 연속된 값을 갖게 된다. feature 변수의 값이 0인 경우, 해당 기사 내에 어휘사전에 포함된 단어가 존재하지 않는다는 의미이며, feature 변수의 값이 1인 경우 해당 기사의 모든 단어가 어휘사전에 포함되어 있다는 의미이다.

본 연구에서는 불확실성 소스별로 총 5가지 요인별 반도체 시장 위험지수를 구하고자 하므로, 개별 기사를 5개의 불확실성 요인에 맞게 5차원의 수치형 자료로 요약하여 feature 변수를 생성하였다. 이렇게 생성된 특성 변수는 총 5차원의 벡터값이며, 그 예시는 [그림 4-6]과 같다.

[그림 4-6] 텍스트 기사를 수치형 feature 변수로 변환한 예시

news_id	기사일자	tokenized_sentence	feature_vector (5 dimension)
000176eaa1	2019.9.24	['알파', '알파홀딩스', '알파홀딩스자회사', '홀딩']	[0, 0, 0024509803921568627, 0, 0]
0001a46f0c8	2023.2.27	['내용', '내용요약', '내용요약지자체', '요약', '!']	[0, 0, 0, 0, 0]
00021353f3f	2022.12.29	['제양자', '제양자기술', '양자', '양자기술', '양']	[0, 0, 0, 0, 0]
00025a7bc5f	2018.7.22	['한겨레', '한겨레이재용', '한겨레이재용부회장']	[0, 0, 0, 0, 007882882882882882]
0002874259f	2020.7.16	['이재용', '이재용정의선', '이재용정의선현대차']	[0, 0, 0, 0, 0]
00028b3b85f	2012.12.28	['장투자', '장투자전략', '투자', '투자전략', '투']	[0, 0, 0005142710208279763, 0005142710208279763, 0]
000292ea9af	2012.5.10	['서울', '서울민속영', '서울민속영반도체', '민속']	[0, 0, 0, 0, 010830324909747292]
0002c6a1e7f	2022.5.18	['미국', '미국뉴욕', '미국뉴욕증시', '뉴욕', '뉴']	[0, 0, 0, 0, 005016722408026756, 0]
0003058fbc4	2017.10.18	['박성욱', '박성욱SK하이닉스', '성욱', '성욱SK']	[0, 0, 0058309037900874635, 0, 0]
00030b01cfd	2018.4.4	['작년', '작년주요', '작년주요대', '주요', '주요']	[0, 0, 0, 0, 0]
0003314a12f	2013.5.13	['머니투데이', '머니투데이송', '머니투데이송선']	[0, 0, 0, 0, 0019723865877712033]
0003b0b4fac	2019.6.5	['삼성전자', '삼성전자기흥', '삼성전자기흥공정']	[0, 0, 0, 0, 0273224043715847]
0003e8c33f1	2023.1.27	['인텔', '인텔지난해', '인텔지난해분기', '지난해']	[0, 0, 0, 0, 0]
0003fb8120f	2016.10.30	['노컷뉴스', '노컷뉴스김학일', '노컷뉴스김학일']	[0, 0, 0, 009746588693957114, 0]
00041882aaf	2015.11.17	['유진', '유진테크', '유진테크삼성전자', '테크']	[0, 0, 0, 0, 0]
000424354cf	2012.11.5	['아시아', '아시아경제', '아시아경제명진', '경제']	[0, 0, 0, 0, 0]

5차원 특성 벡터의 각 차원이 나타내는 순서는 정책 불확실성, 외교 불확실성, 반도체 시장 불확실성, 거시경제 불확실성, 외부 충격 불확실성 순이다. [그림 4-6]에서 첫 번째 기사를 보면, 이는 정책, 외교, 거시경제, 외부 충격 불확실성의 특성값은 0이고 반도체 시장 불확실성(3번째)의 특성값은 0.0024509... 임을 의

미하는 것이다.

나. EM 알고리즘을 통한 조건부 확률의 모수 추정

본 연구에서 정의한 불확실성 지수인 R 은 현재 상태가 불안정상태일 때 해당 어휘 패턴이 나올 확률과 안정상태일 때 해당 어휘 패턴이 나올 확률의 비를 의미한다. 즉, 조건부 확률의 비로 표현되는 것인데, 조건부 확률은 직접적으로 관측되는 확률이 아니기 때문에 모델링을 통하여 이를 추정하고자 하였다.

본 연구에서는 반도체 시장위험이 안정상태와 불안정상태의 선형 결합으로 이루어져 있다고 가정한다. 그러나, 실제로 안정상태일 때의 어휘 패턴과 불안정상태일 때의 어휘 패턴이 명시적으로 라벨링 되어있지 않기 때문에 명시적인 라벨링 없이 모델링하는 비지도학습 형태로 모형을 적합시켰고, 시장의 상태가 기사들의 어휘 패턴을 통해 어느 시장에 속할지를 확률의 형태로 추정하는 과정에서 EM 알고리즘을 활용하였다.

1) EM 알고리즘의 특징

EM 알고리즘은 하기의 2단계를 반복적으로 수행하여 모수 추정값을 구하는 알고리즘이다. ① E-Step(Expectation): 이 단계에서 현재 모델의 모수(parameter) 추정치를 사용하여 결측된 변수에 대한 조건부 기댓값을 계산한다. E-step의 기본 정의는 [식 4-3]과 같다. $\eta^{(t)}$ 가 η 의 현재 모수 추정치 값이라고 가정하면, E-step에서는 아래 [식 2-3]을 계산하여 현재 모수 추정치가 주어졌을 때 η 를 구하고자 한다. [식 4-3]의 우변에서 볼 수 있듯이 이는 모든 관측된 부분이 주어졌을 때 로그 질량 함수(log-density function)의 기댓값으로 계산하여 구할 수 있다.

$$Q(\eta \mid \eta^{(t)}) = E\{\ln f(y, \delta; \eta) \mid y_{obs}, \delta, \eta^{(t)}\} \quad [\text{식 4-3}]$$

② M-step(Maximization): E-step에서 계산된 기댓값을 활용해 모수를 최적화하는 단계이다. 이는 모델의 로그 가능도 $Q(\eta \mid \eta^{(t)})$ 를 최대화하는 새로운 모수 $\eta^{(t+1)}$ 을 찾는 과정이다.

$$\eta^{(t+1)} = \text{Argmax}_{\eta} Q(\eta \mid \eta^{(t)}) \quad [\text{식 4-4}]$$

2) EM 알고리즘의 적용

본 연구에서는 각 기사를 수치화하여 요약한 feature 변수를 EM 알고리즘을 통해 모델링하였다. 특정 키워드 패턴(feature vector)이 주어지면, 해당 기사가 불안정상태(시장 위험이 높은 불확실한 상태) 또는 안정상태에 속할 확률을 모델 파라미터로 추정하는 것이 목적이다.

EM 알고리즘을 진행하기 위해서 변수의 모집단에 대한 분포 가정을 선행해야 한다. 본 연구에서는 변수 형태를 이산형(0 또는 1)으로만 표현하면 기사에 포함된 수많은 정보를 과도하게 축약시키게 되므로 이 또한 정확한 지수산출에 영향을 미칠 수 있다고 판단했다. 이에 변수 형태를 이산형 변수가 아닌 연속형 변수로 모델링하였다. 뉴스 기사가 어휘사전 내의 단어 개수를 많이 포함할수록, 해당 불확실성 요인에 대해 더 많이 나타내고 있는 것이므로, 뉴스 기사가 각 불확실성 요인별 어휘사전의 단어를 얼마나 많이 포함하고 있는지를 정규화하여 변수로 사용하기로 하였다. 변수를 연속형으로 설정하였으므로 연속형 분포를 다룰 때 가장 널리 사용되는 가우시안 혼합 모형(Gaussian mixture model)을 가정해서 EM 알고리즘 모델링을 진행하였다.

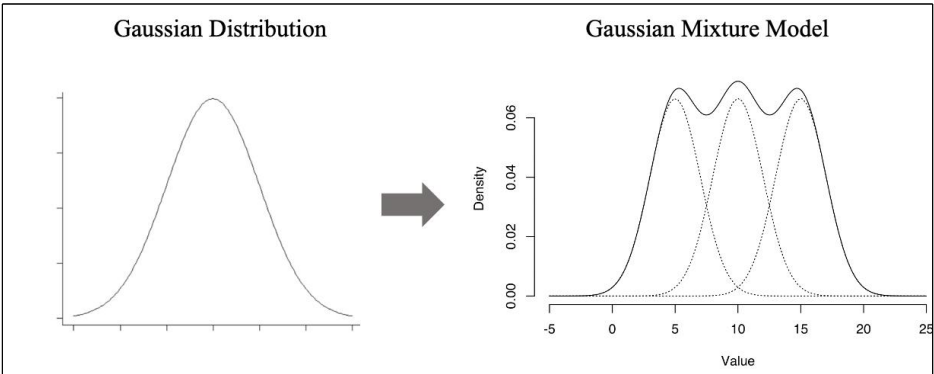
가우시안 혼합 모형은 가우시안 분포를 여러개 혼합한 알고리즘으로, 현실의 복잡한 형태의 확률 분포를 [그림 11]²⁵⁾과 같이 K개의 가우시안 분포를 혼합하여 표현하는 것이다. 이를 EM 알고리즘을 통해 모델링하면, 주어진 데이터가 K개의 가우시안 분포 중 어떤 가우시안 분포에서 생성되었는지를 추정할 수 있게 된다. 데이터(x_1, \dots, x_n)가 있을 때, x_n 이라는 데이터가 주어지면, EM 알고리즘을 통해 이 데이터가 K개의 가우시안 분포에 대해 각 분포에서 생성되었을 확률이 총 K개 구해지고, 이 중 확률이 가장 큰 가우시안 분포에서 x_n 이 생성되었다고 추정하

25) S.Gaussian-mixture-example. 재인용

<https://commons.wikimedia.org/wiki/File:Gaussian-mixture-example.svg>

게 되는 것이다.

[그림 4-7] 가우시안 혼합 모형(Gaussian Mixture Model)



가우시안 혼합 모형을 통해 본 연구에서 EM 알고리즘을 적용한 과정은 아래와 같다. 뉴스 기사를 $x = \{x_1, x_2, \dots, x_N\}$ 으로 정의하면, 주어진 x_i (기사)에 대해 x_i 가 발생할 확률은 [식 4-5]와 같이 2개의 가우시안 분포의 합으로 표현할 수 있다. 반도체 시장이 불안정상태와 안정상태 2개의 선형결합으로 표현된다고 가정하였기 때문에 2개의 가우시안 분포의 합으로 정의한 것이다. (μ_1, σ_1^2) 을 모수로 갖는 분포를 $H_{cluster}$ (불안정상상태 분포), (μ_2, σ_2^2) 을 모수로 갖는 분포를 $L_{cluster}$ (안정상상태 분포)로 가정하였다. 이때, 잠재변수 $z_{ik} = (z_{i1}, z_{i2})$ 를 정의하는데, z_{ik} 는 x_i 가 k 번째 분포에 속하면 1, 아니면 0값을 갖는 지시변수(indicator variable)이다 ($k=1, 2$). $P(z_{ik}=1) = \pi_k$ 는 두 분포를 혼합하여 주는 모수로 k 번째 가우시안 분포가 선택될 확률을 나타낸다: $\pi = (\pi_1, \pi_2)$. 따라서, $\sum_{k=1}^2 \pi_k = 1$ 이며 $0 \leq \pi_k \leq 1$ 이다.

$$\begin{aligned}
 p(x_i) &= \pi_{i,1} N(x_i | \mu_1, \sigma_1^2) + \pi_{i,2} N(x_i | \mu_2, \sigma_2^2) \\
 &= \pi_{i,1} H_{cluster} + \pi_{i,2} L_{cluster}
 \end{aligned}
 \tag{식 4-5}$$

결국 본 모델링을 통해 구하고자 하는 것은 기사 x_i 가 주어졌을 때 $z_{ik} = 1$ 일 각

확률($P(z_{ik}=1) = \pi_k$)인데, 이는 베이즈정리를 이용하면 [식 2-6]과 같이 나타낼 수 있다.

$$\begin{aligned} p(z_{ik}=1 | x_i) &= \frac{p(z_{ik}=1)p(x_i | z_{ik}=1)}{\sum_{j=1}^2 p(z_{ij}=1)p(x_i | z_{ij}=1)} \\ &= \frac{\pi_k f(x_i | \mu_k, \sigma_k^2)}{\pi_1 f(x_i | \mu_1, \sigma_1^2) + \pi_2 f(x_i | \mu_2, \sigma_2^2)} \end{aligned} \quad \text{[식 4-6]}$$

이후 E 단계와 M 단계를 반복적으로 수행하게 된다.

E-step: 현재 모수를 [식 2-6]에 대입하여 숨겨진 변수 z_{ik} 의 조건부 기댓값을 추정한다. 본 연구에서 observed log-likelihood(L_{obs})는 [식 2-7]과 같고, complete-data log-likelihood(L_c)는 [식 2-8]로 표현할 수 있다.

$$\begin{aligned} \log L_{obs}(\pi) &= \sum_{i=1}^n \log f(x_{(i)}; \pi_k, \mu_k, \sigma_k^2) \\ &= \sum_{i=1}^n \log \left\{ \sum_{k=1}^2 \pi_k f_k(x_{(i)}; \pi_k, \mu_k, \sigma_k^2) \right\} \end{aligned} \quad \text{[식 4-7]}$$

$$\log L_c(\pi) = \sum_{k=1}^2 \sum_{i=1}^n z_{ik} \log \pi_k + \sum_{k=1}^2 \sum_{i=1}^n z_{ik} \log f_k(x_i; \mu_k, \sigma_k^2) \quad \text{[식 4-8]}$$

[식 4-8]을 [식 4-3]에 대입하여 풀면 $z_{ik}^{(t)} \equiv E(z_{ik} | x_i, \pi^{(t)})$ 가 되어 조건부 기댓값을 추정할 수 있게 된다([식 4-9]).

$$\begin{aligned} E\{\ln f(x, \delta; \eta) | x_{obs}, \delta, \eta^{(t)}\} &= E\{\ln L_c(\pi) | x_{obs}, \pi^{(t)}\} \\ &= \sum_{k=1}^2 \sum_{i=1}^n E(z_{ik} | x_i, \pi^{(t)}) \log \pi_k + \\ &\quad \sum_{k=1}^2 \sum_{i=1}^n E(z_{ik} | x_i, \pi^{(t)}) \log f_k(x_i; \mu_k, \sigma_k^2) \end{aligned} \quad \text{[식 4-9]}$$

$$\text{where } E(z_{ik} | x_i, \pi^{(t)}) = \frac{\pi_k^{(t)} f_k(x_i)}{f(x_i; \pi^{(t)})}$$

M-step: M 단계에서는 위 조건부 기댓값을 최대화하는 모수를 찾는 단계이다. 가우시안 혼합 분포에서는 하기와 같이 모수 추정치를 업데이트할 수 있는 것으로 알려져 있다.

$$\begin{aligned} \pi_k^{(t+1)} &= \frac{\sum_{i=1}^n z_{ik}^{(t)}}{n} \\ \mu_k^{(t+1)} &= \frac{\sum_{i=1}^n z_{ik}^{(t)} x_{(i)}}{\sum_{i=1}^n z_{ik}^{(t)}} \\ \sigma^{2(t+1)} &= \frac{\sum_{k=1}^2 \sum_{i=1}^n z_{ik}^{(t)} (x_i - \mu_k^{(t+1)})^2}{n} \end{aligned} \quad \text{[식 4-10]}$$

위와 같이 E-step 과 M-step을 수렴할 때까지 반복함으로써 두 개의 가우시안 분포에 대한 모수(μ_k, σ_k^2, π_k)를 추정할 수 있다. 이를 통해 모든 모수값 (μ_k, σ_k^2, π_k)이 추정되면, [식 4-6]을 통해 기사가 주어졌을 때 2개의 분포 중 어떠한 분포에서 생성되었는지 그 확률을 계산할 수 있게 된다.

위의 EM 알고리즘을 적용하는 부분은 본 연구에서는 하나의 반도체 산업 하위 카테고리의 한가지 불확실성 요인에 대한 조건부 확률을 구하는 과정이다. 따라서, 위에서 정의한 π_k 모수는 본 연구에서는 반도체 산업 카테고리(i)와 불확실성 요인(j) 조합의 개수만큼 총 $i \times j$ 개 추정되게 된다.

다. 지수산출 대상 기사 추출

본 연구의 최종 목표는 뉴스 데이터를 지수화하는 작업이기 때문에 강건한 추정량을 만드는 것은 매우 중요하다. 수집한 모든 뉴스 데이터를 사용한다면 사용하는 데이터는 많아질 수 있으나, 특정 패턴에 지나치게 과적합될 수 있다. 따라서 수집한 모든 뉴스 데이터를 사용하는 대신 일정 크기의 표본을 선정하여 지수를 산출하였다. 반도체 산업을 하위 카테고리 단계로 나누어 월별 지수화를 진행하기 때문에, 하위 카테고리 지수 간의 강건성(robustness)을 보장하는 것이 중요하다

고 판단하였다. 이에 매월 i 번째 산업 하위 카테고리의 시장 불확실성지수를 산출하는 데 사용되는 기사 수(m_i)를 동일하게 추출하기로 하였다. 이는 표본 크기를 통제함으로써, 반도체 산업 하위 카테고리 간 비교할 때 신뢰성을 높이고자 한 것이다. 또한, 월별로도 동일한 기사의 개수를 표본 추출하는 것이 지수의 시계열적인 특성 상 분석의 검정력을 높일 수 있다고 판단하였다. 이에, 결과적으로 매월 각 산업 하위 카테고리에서 추출하는 표본 기사의 개수는 m 개로 설정하기로 하였다.

라. 반도체 시장 위험지수 산출

지수산출 대상으로 추출된 기사에 EM 알고리즘을 통해 수렴된 파라미터 값을 적용하여 반도체시장 위험지수(R) 값을 구하였다.

$$R = \frac{P(x \mid x \in H)}{P(x \mid x \in L)} \quad [\text{식 4-11}]$$

R 은 특정 기사(x)가 불안정상태(H)에 속할 확률과 안정상태(L)에 속할 확률의 비율을 나타낸다. 분자가 불안정상태에 속할 확률이고, 분모가 안정상태에 속할 확률이므로, R 이 클수록 특정 기사(x)가 불안정상태에 속할 확률이 안정상태에 속할 확률보다 높다는 것을 의미하는 것이다. 이는 결국 R 값이 클수록 기사가 불안정상태에 속할 확률이 높다는 것이며, 이는 해당 기사가 반도체 시장이 위험할 때 나온 기사일 확률이 높다는 의미가 된다.

본 연구에서는 반도체 산업 하위 카테고리별, 불확실성 요인별로 R 을 산출하는 것이 목표이므로, 지수산출 대상으로 추출된 모든 기사에 대해서 $R_{ijt(k)}$ 를 계산하였다. 이 때 i 는 반도체 산업 하위 카테고리 ($i=1, 2, \dots, 9$), j 는 불확실성 요인 ($j=1, \dots, 5$), t 는 월단위를 나타내는 월($t=1, 2, \dots, T$), 그리고 k 는 각 월별 뉴스 기사 수($k=1, 2, \dots, m$)를 나타낸다.

마. 정규화 적용

지수산출의 강건성을 높이기 위해 지수산출의 대상이 되는 모든 기사의 $R_{ijt(k)}$

값에 정규화를 적용하였다. 정규화를 통해 $R_{ijt(k)}$ 값의 범위를 일정하게 맞춰줌으로써, 지수작성 기간별, 산업 카테고리별, 불확실성 위험 요인별로 $R_{ijt(k)}$ 값의 스케일이 다르게 산출된 각 값을 동일한 범위 내에서 비교하고자 한 것이다. 또한, 정규화 과정을 통해서 스케일을 맞추는 작업뿐만 아니라, 극단적인 값을 갖는 기사가 포함되어 있더라도 지수의 전체적인 추세나 패턴이 이 기사에 의해 과도하게 왜곡되는 것을 방지하고자 하였다.

정규화 방법으로는 일반적으로 활용되는 표준화 방법(Standardization)을 사용하였다. 표준화 방법은 데이터가 평균 0, 표준편차 1의 표준정규분포의 형태로 조정하는 것이다. 표준화를 적용하는 식은 [식 4-12]와 같다.

$$x_{new} = \frac{x - mean}{std} \quad \text{[식 4-12]}$$

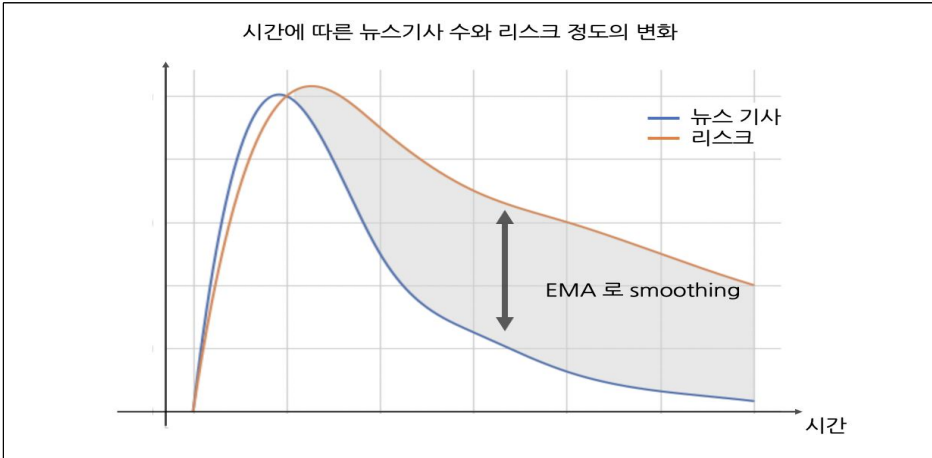
바. 지수평활법 적용

본 연구에서는 앞서 계산한 기사별 $R_{ijt(k)}$ 을 월별로 결합하여, 산업 하위 카테고리별, 위험요인별 월별 위험지수를 산출하였다. 월별로 모든 기사에 동일한 가중치를 두는 방식으로 이를 결합하였다([식 2-13]).

$$R_{ijt} = \frac{\sum_k R_{ijt(k)}}{m} \quad \text{[식 4-13]}$$

월별 위험지수 R_{ijt} 를 산출한 후에 R_{ijt} 값에 지수평활법을 적용하여 최종적인 R_{ijt}^{EMA} 값을 도출했다. 뉴스 기사의 경우, 특정 사건이 발생하면 관련 기사의 수가 급격히 증가하는 경향이 있지만, 시간이 지나면서 해당 사건과 관련된 기사의 수는 감소한다. 그러나, 기사 수의 감소가 반드시 해당 사건으로 인한 리스크 감소를 의미하지는 않는다. 관련 기사의 수가 줄어들었다 하더라도, 그 사건으로 인한 리스크는 여전히 일정 수준에서 지속될 수 있기 때문이다.

[그림 4-8] 시간에 따른 뉴스기사 수와 리스크 정도의 변화 양상



이러한 뉴스 데이터의 특성을 고려하여, 본 연구에서는 단기적인 변동성을 안정화시키기 위해 지수평활법(EMA, exponential smoothing)을 적용하였다. 평활화 기법은 데이터 내의 무작위적 변화로 인한 불규칙적인 효과를 제거하기 위한 근사 함수를 만드는 방법이다. 이를 통해 단기적인 변동성은 완화하고, 장기적인 추세를 강조하는 것이 목표이다. 평활화 기법에는 여러 방식이 있으며, 이 중에서 본 연구는 지수평활법을 사용했다.

지수평활법은 시계열 데이터 분석에 있어 시점마다 다른 가중치를 두어 분석하는 방법인데, 특히 최근 데이터에 더 많은 가중치를 부여해서 분석하는 것이다. 구체적으로 구하는 수식은 [식 4-14]와 같다. 이때, x_t 는 t 시점에서 x 값을 의미하고, 본 연구에서는 5차원의 수치형으로 변환한 feature vector를 의미하는 것이다. α 는 평활화 요소(smoothing factor)로써, 이 값이 클수록 현재 시점(t)을 더 많이 반영하게 된다. n 은 평활화하는 데 반영하는 기간의 길이를 의미하는 것으로, n 이 클수록 더 예전의 기간의 값까지도 반영된다.

$$\begin{aligned}
 EMA(n)[x]_t &= \alpha x_t + (1 - \alpha) EMA(n-1)[x]_{t-1} \\
 &= \sum_{k=0}^n \alpha (1 - \alpha)^k x_{t-k}
 \end{aligned}
 \tag{식 4-14}$$

여러 방법 중에서도 본 연구에서 지수평활법을 사용하는 이유는 하기와 같다. 이 방법은 최근 데이터에 더 큰 가중치를 부여하기에 이는 뉴스 데이터가 반영하는 시장 분위기를 가장 잘 포착할 것으로 예상된다. 또한, 오래된 데이터의 영향력을 점진적으로 감소시키므로, 특성 사건의 발생으로 리스크가 급증한 후 감소하는 동안에도 뉴스 기사에 따라 함께 급감하는 것이 아니라 리스크는 일정 수준으로 지속될 것으로 가정하는 본 연구의 가설을 잘 반영한다. 따라서, 우리는 월별로 도출한 시장 위험지수에 최종적으로 지수평활법을 적용한 값을 최종 월별 시장 위험지수로 사용하였다.

4. 반도체 시장 불확실성지수 계산

가. 데이터

본 연구에서는 ㈜와이즈넷으로부터 제공받은 반도체 관련 네이버 뉴스 기사 데이터를 활용하였다. 데이터는 작성 날짜, 뉴스 ID, 주차(date-week), 미디어, 섹션, 기사 제목, 기사 내 문장 순서, 기사 내 문장 개수, 문장, 기사 URL, 반도체 관련 키워드, news_id별 기업, 하위 카테고리 분류(부품, 소재, 장비 등) 등으로 구성 되어있다. 2010년 1월 1일부터 2023년 7월 1일까지 수집된 기사 데이터로, 총 11,729,257개의 기사 문장이 포함되어 있다.

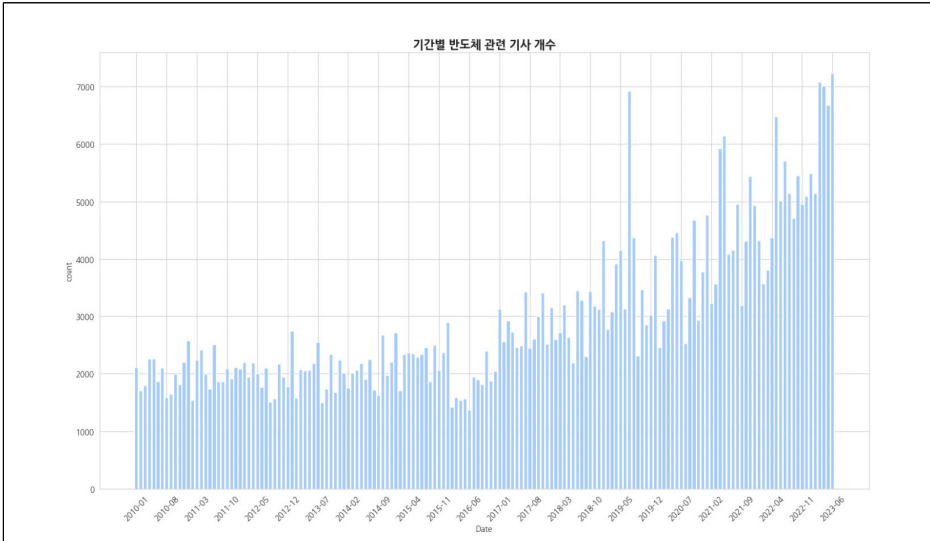
본 연구에서는 기사 단위로 변수를 생성하게 되므로, 이를 위해 뉴스 ID(news_id)를 식별자로 하여 각 문장 단위로 된 데이터를 뉴스 ID별로 결합하였다. 식별자가 되는 기사(news_id)의 개수는 총 478,205개이다. 데이터의 형식은 [그림 4-9]과 같다.

[그림 4-9] 데이터 형식 예시

필드명	데이터 샘플	설명
yearmonth	2011-02-23	연도+월+일
news_id	448ff0368d5638e05e482b20801c58e4	뉴스 기사 식별자
media	네이버 뉴스	언론사
section	경제	네이버뉴스 카테고리 ID
title	온세미컨덕터 한스 스투크 박사를 수석 부사장 임명	뉴스 제목
sentence_cnt	12	기사 내 문장의 수
keyword	IBM	기사수집 시 query 로 제공한 keyword
domain 분류	IP기업	0
	OSAT	0
	부품	0
	소재	0
	장비	0
	파운드리	0
	팹리스기업	0
	후공정	0
IDM	1	더미 변수 (각 domain 에 해당하면 1, 아니면 0)

본 연구에서 사용된 반도체 관련 기사는 [그림 14]에서 확인할 수 있듯이 최근으로 올수록 기사의 수 자체가 늘어나는 추세를 보인다. 이는 미디어 환경 변화에 따라서 특정 이벤트나 이슈에 따라서 생성되는 기사의 평균적인 수가 증가한 것으로 보인다.

[그림 4-10] 기간별 반도체 관련 기사 개수 그래프



나. 반도체 시장 위험지수 산출 과정

1) 비정형데이터 전처리 및 word2vec

뉴스 기사 데이터를 Kiwi 형태소 분석기를 이용하여 품사를 태깅하고, 이 태깅된 품사 중 명사에 해당하는 “NNG(일반 명사)”, “NNP(고유 명사)”, “XPN(체언 접두사)”, “XSN(명사 파생 접미사)”, “XR(어근)” 품사를 가진 단어들만 구별하였다. 이후, bigram이 가능한 경우와 trigram이 가능한 경우를 파악하여 추가하였다. Bigram과 trigram을 적용하다보니 띄어쓰기로 인하여 같은 단어 조합이 다른 단어로 잡히는 경우가 있어, 띄어쓰기는 삭제하는 과정을 거쳐 전처리를 진행하였다.

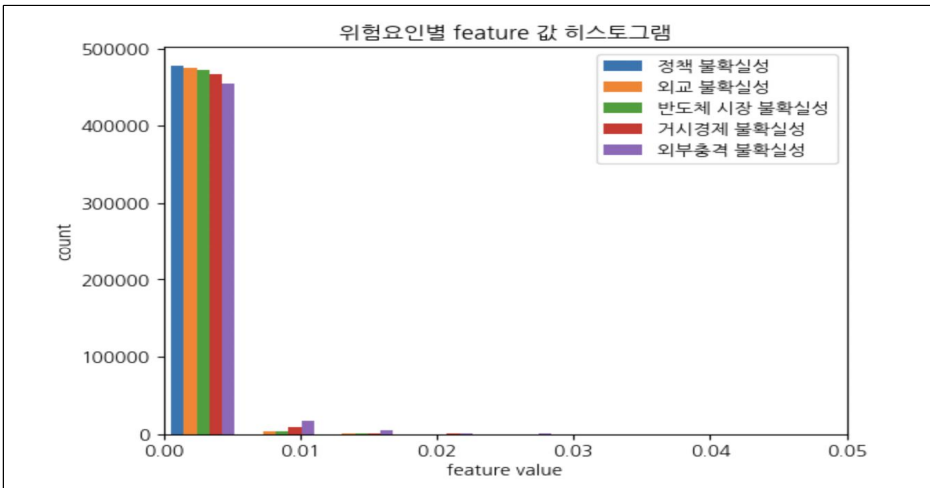
2) 불확실성 요인별 어휘사전 수립

Berstiness를 기준으로 하여 각 불확실성 요인을 정의하고 요인별 seed word가 될 수 있는 단어를 식별하였다. 이후, 해당 seed word를 word2vec 모델에 적용하여 유사성이 높은 10개씩의 단어를 추가적으로 확인하였다. 이러한 단어 중 seed word로 설정할 수 있을 것으로 참여 연구원들의 의견이 합치된 단어를

추가적인 seed word로 설정하였다. 이 과정을 통해 선정한 seed word를 기준으로 코사인 유사도가 높은 상위 20개의 단어를 확인하고, 이를 참여 연구원이 검토하여 관련성이 없거나 매우 적다고 판단하는 단어는 삭제한 후 각 불확실성 위험 요인별 총 80개의 단어를 선정하여 불확실성 위험 요인별 어휘사전(word cluster)을 작성하였다. 각 위험 요인별 어휘사전은 [부록]에 첨부했다.

3) 특성 변수 생성

[그림 4-11] 위험 요인별 특성값의 히스토그램



feature 변수는 기사에 포함된 요인별 어휘사전 내의 단어 개수를 해당 기사의 전체 단어 수로 나누는 방식으로 생성하였다. 요인별로 각 기사의 특성값에 대한 히스토그램은 [그림 4-11]와 같다. 즉 특성 변수는 대부분의 값이 0임을 확인할 수 있고, 이 중에서도 정책불확실성 요인에서 0이 가장 많음을 확인할 수 있다. 이는 기사가 정책불확실성 요인 어휘사전 내의 단어를 포함하는 경우가 가장 적다는 것을 의미한다.

4) EM 알고리즘의 적용

획득한 기사별 특성 변수에 대해서 가우시안 혼합 모형을 가정하여 EM 알고리즘을 적용하였다. 모델은 python의 사이킷런 패키지에 내재된 있는 “Gaussian Mixture” 모듈을 사용하였다. 본 연구는 기사가 불안정상태와 안정상태, 총 2개의 cluster 각각에 속할 확률을 확인하는 것이기 때문에 cluster의 개수를 의미하는 n_components는 2로 설정하였다. 추가로, 본 연구에서 사용한 데이터의 특성 및 우리의 가정 상 변수에서 0 값이 0 이외의 값보다 훨씬 많이 나오게 된다. 시장위험이 큰 불안정상태는 특별한 사례이고, 일반적으로는 시장이 안정상태에 있다고 보기 때문이다. 따라서, EM 알고리즘을 통해 어떤 cluster에 속할 확률 값이 총 2개가 나오면, 이 중 높은 확률을 지닌 cluster를 정상 상태로 설정하여 이 방법이 가질 수 있는 identifiable 이슈를 해결하고자 하였다.

다. 반도체 시장 위험지수 산출 결과

비지도학습 방법론을 적용하여 각 산업 하위 카테고리별로 월별 총 80개의 개별 뉴스를 단순 무작위 표본추출(simple random sampling) 방식으로 샘플링한 후 월별 지수를 산출하기로 하였다. 다만, 월별로 기사 개수를 분류했을 때, OSAT, 부품 카테고리에서는 한 달에 있는 기사의 총개수가 1개밖에 없는 경우도 있고, OSAT 카테고리의 경우 한 월에 있는 최대 기사 수가 73개, 부품 카테고리의 경우 한 달의 최대 기사 수가 38개였다. 따라서, 지수를 정규화하여 산출하기 위해서는 OSAT와 부품 카테고리는 표본이 부족하다고 판단하여 이번 지수산출 시에는 산업 하위 카테고리에서 제외하고 계산하였다.

표본 추출한 모든 기사에 대해 $R_{ijt(k)}$ 을 계산하고, 이 값에 정규화 방법을 적용한 뒤, 각 산업 하위 카테고리별, 불확실성 위험 요인별로 월별 기사의 $R_{ijt(k)}$ 값을 결합하였다. 월별로 결합 시 단순 기중 합의 방법(즉, 평균)으로 계산하였으며, 월별로 결합한 R_{ijt} 값에 수평활법을 적용하여 R_{ijt}^{EMA} 한 값으로 최종 산업 하위 카테고리별, 위험 요인별 월별 시장위험지수를 산출하였다.

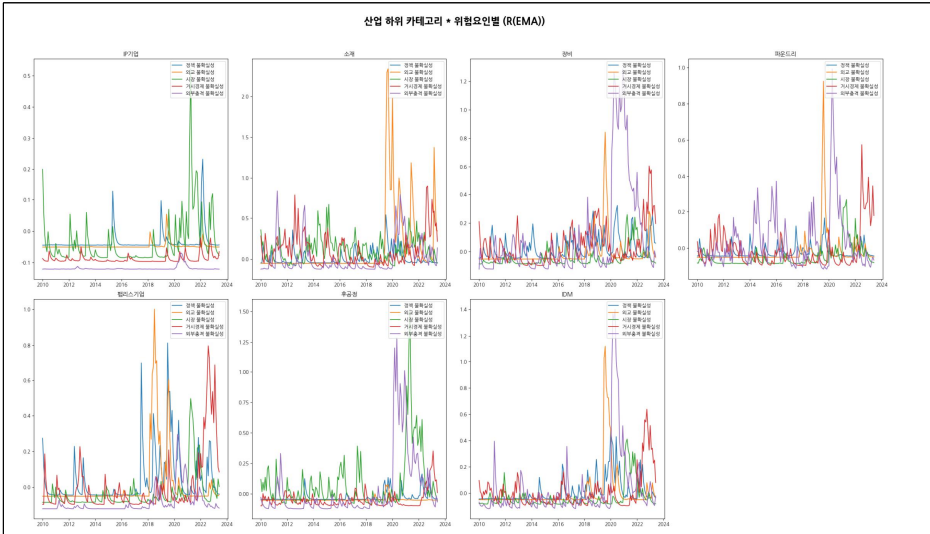
1) 반도체 시장 카테고리 및 위험 요인별 불확실성 지수

각 기사의 $R_{i,jt(k)}$ 를 월별로 평균하여 반도체 시장 산업 하위 카테고리별, 위험 요인별로 확인한 결과는 [그림 4-12]과 같다. [그림 4-12]에서 IP기업과 후공정 카테고리에서 다른 카테고리에 비해 외교 불확실성요인 지수의 변화량이 크지 않은 것을 확인할 수 있다. IP 기업의 경우 IP를 제공하고 사용에 따른 라이선스를 받는 구조이므로 물리적인 제조 공정에 직접적으로 관여하지 않는다. 이에, 다른 산업 카테고리에 비해 외교적 이슈에 다소 덜 민감하게 반응한 것으로 파악된다. 또한, 후공정 카테고리의 경우 완성된 반도체 제품의 패키징 및 테스트와 관련된 카테고리이므로, 글로벌 공급망에 영향을 많이 받는 소재, 팹리스기업에 비해 외교적 이슈에 크게 민감하지 않은 것으로 파악된다. 소재 카테고리의 경우 다른 불확실성 요인에 비해 외교 불확실성 지수가 급증하는 부분이 많이 보인다. 반도체 소재 카테고리는 해외 의존도가 굉장히 높은 카테고리이기 때문에²⁶⁾ 외교 불확실성 요인이 다른 불확실성 요인에 비교하여 큰 영향을 미쳤음을 확인할 수 있다.

또한, IP 기업 카테고리에서 다른 요인들에 비하여 수요, 공급과 관련된 시장 불확실성 지수가 높음을 확인할 수 있다. IP 기업의 경우 시장의 수요를 보고 어떤 IP가 필요할지 미리 예상하여 개발하는 역할을 한다. 그리고 이와 같이 개발한 IP를 반도체 제조의 가장 초기 단계에서 제공하는 역할을 하기 때문에 다른 불확실성 요인에 비해 수요, 공급과 관련된 시장불확실성 지수에 민감하게 반응하는 것으로 판단하였다.

26) 대외경제정책연구원, 한국 반도체 산업의 공급망 리스크와 대응방안, 2021. 11. 23.

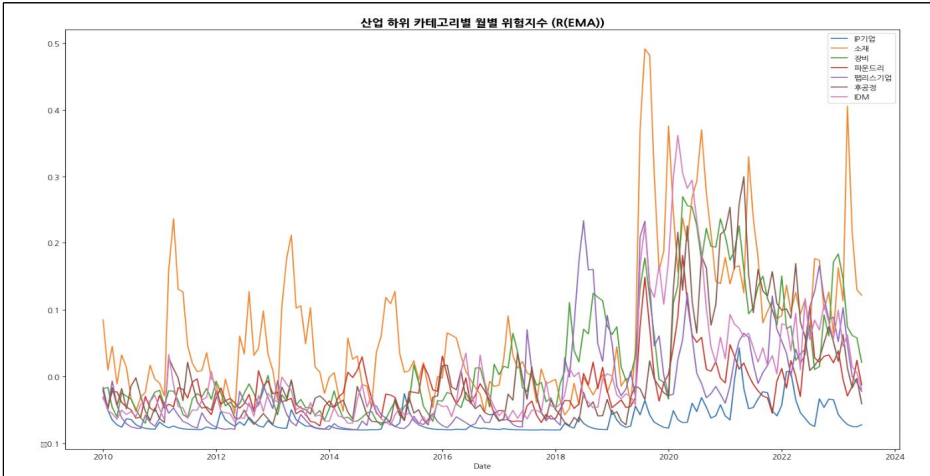
[그림 4-12] 산업 하위 카테고리별 * 위험 요인별 * 월별 시장위험지수



2) 반도체 시장 하위 카테고리별 월별 불확실성지수

각 기업의 R_{ijt}^{EMA} 를 반도체 시장 하위 카테고리별로 월별 평균하여 산출한 $R_{i,t}^{EMA}$ 은 [그림 4-13]과 같다. [그림 4-13]에서 다른 카테고리에 비해 2018년도에서 2019년도 사이에 팹리스 기업 카테고리의 시장위험지수가 급격하게 높아지는 것을 확인할 수 있다. 이는 다른 카테고리에 비해 해당 시기에 팹리스 기업의 시장 위험도가 높았음을 의미한다. 또한, 이전 시기에 비해 2020년도 이후에 모든 반도체 산업 하위 카테고리의 불확실성 지수가 높아졌음을 확인할 수 있다. 이는 반도체 산업 자체가 성장함에 따라, 여러 가지 요인에 영향을 많이 받게 됨으로써 전 분야의 불확실성 지수가 증가한 것으로 해석하였다. 예를 들어, 기업이 창립 초기에는 불확실성이 상대적으로 낮거나 거의 없을 수 있으나, 기업이 성장하고 발전함에 따라 다양한 요인들에 영향을 받게되어 갑작스럽게 하락할 가능성이 증가하면서 불확실성 수준도 함께 상승하게 되는 것과 같은 맥락이다.

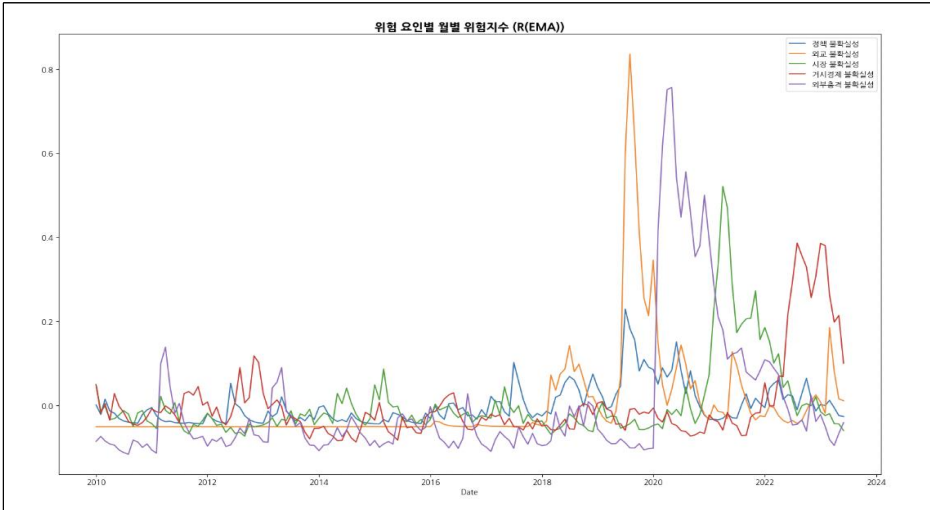
[그림 4-13] 산업 하위 카테고리별 * 월별 시장위험지수



3) 위험 요인별 월별 시장위험지수

각 기사의 R_{ijt}^{EMA} 를 위험 요인별 월별 평균하여 산출한 R_{jt}^{EMA} 은 하기의 [그림 18]과 같다. [그림 4-14]에서는 2020년 이후 외부충격 불확실성 요인으로 인한 시장위험지수가 급증하는 것을 확인할 수 있다. 이는 코로나-19의 발병이 2019년 말부터 시작되었음에 따라, 해당 이슈가 반도체 시장 하위 카테고리 전반에 큰 영향을 주었음을 시사한다. 추가로, 2019년도에 외교 불확실성 지수가 급격하게 증가하였음을 확인할 수 있으며, 2022년도 하반기 이후부터 거시경제 불확실성 지수가 이전 시점과 대비하여 증가한 뒤 지속적으로 높은 지수를 유지하고 있음을 확인할 수 있다.

[그림 4-14] 위험요인별 * 월별 시장위험지수



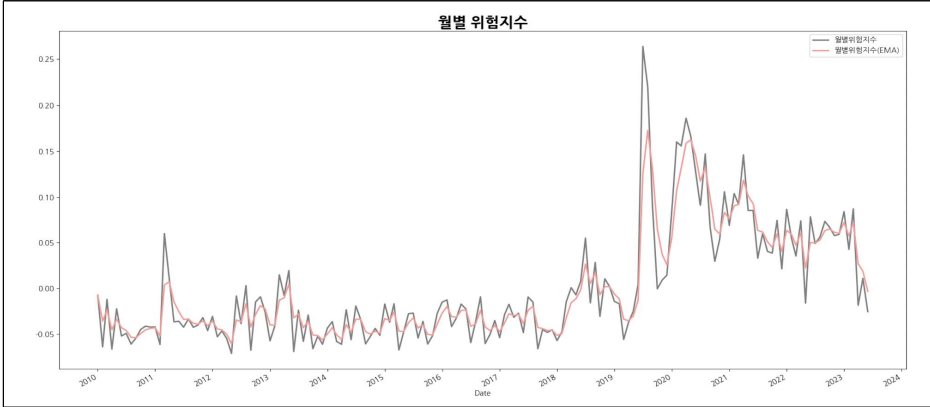
4) 월별 최종시장위험지수

전체 반도체 시장에 대한 최종시장위험지수는 하기와 같이 산출되었다. 지수로서의 정합성을 판단하기 위해 지수평활법(EMA)을 적용한 월별 최종시장위험지수와 지수평활법을 적용하지 않은 월별 최종시장위험지수를 비교 평가해보았다.

지수는 전체 시장이나 경제 상황의 진정한 상태를 반영해야 하기에, 비정상적인 변동에 지나치게 영향을 받는 것은 바람직하지 않다. 지수의 변동이 심하면 이를 분석하여 활용하는 시장의 전반적인 변동성과 불안정성이 증가할 수 있기 때문이다. 이는 도리어 지수가 시장 리스크를 식별하는 것이 아니라 반대로 시장 리스크를 증가시킬 수도 있음을 시사하게 된다. 이에 갑작스러운 변동성이 생겨도 이를 직전의 전반적인 추세와 비교해서 평탄화하여 반영하는 지수평활법을 적용한 지수가 본 연구의 목표에 더 적합하다고 판단하였다.

월별 최종시장위험지수 이외에도, 반도체 시장 하위 카테고리별, 위험 요인별 월별 시장위험지수를 지수평활법 적용한 결과와 적용하지 않은 결과 또한 부록에 첨부하였다.

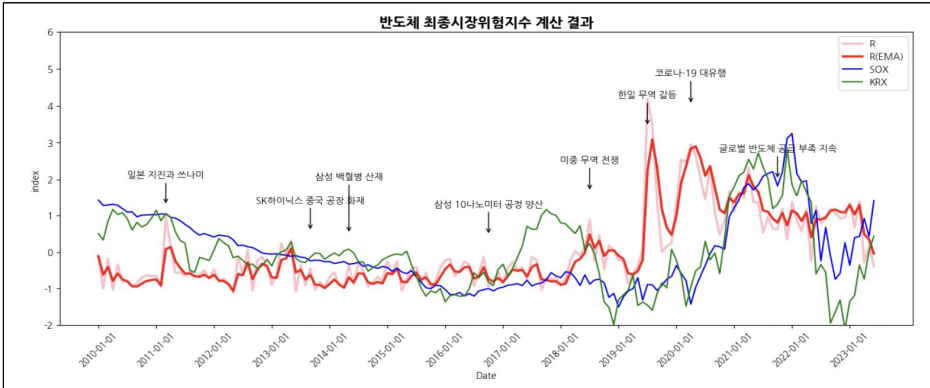
[그림 4-15] 월별 최종위험지수(EMA 비교)



5) 월별 반도체 최종시장위험지수 계산 결과

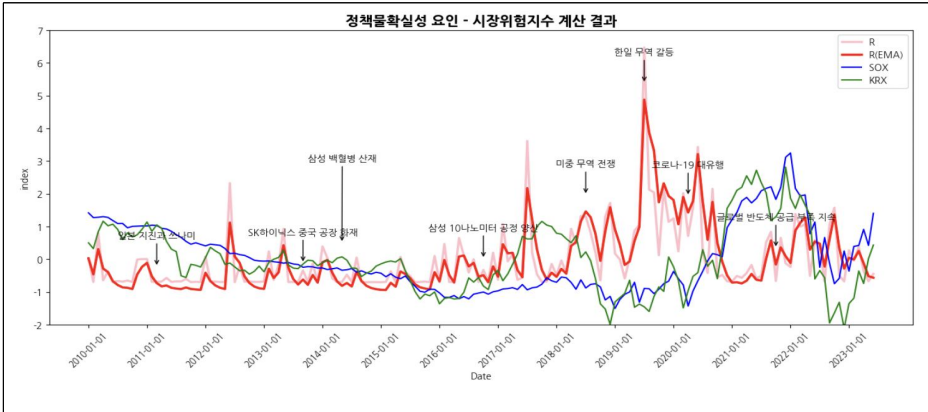
본 연구에서 도출한 월별 반도체 최종시장위험지수가 그 효용성이 있을지 확인해보고자 통상 활용되고 있는 지표와의 관계성을 파악해보았다. 통상 활용되고 있는 지표로는 필라델피아 반도체지수(SOX)와 KRX 반도체 지수를 사용하였다. 필라델피아 반도체지수는 반도체 설계, 공급, 제조 판매에 관련된 기업들의 주가를 지수화한 것이고, KRX 반도체 지수는 국내 반도체 산업을 대표하는 기업들을 시가총액 가중방식으로 구성한 지수이다. 해당 지수들의 경우 선형 회귀분석을 수행하여 선형 회귀식을 데이터에서 뺀 값으로 나타내었다. 선형 회귀식을 데이터에서 빼는 것은, 시간에 따라 자연스럽게 발생하는 변화를 제거하고, 순수한 가격 변동성이나 지수의 움직임을 더 명확하게 파악하기 위함이다. 본 연구에서 최종 도출한 월별 반도체 최종시장위험지수를 필라델피아 반도체지수(SOX), KRX 반도체 지수와 함께 비교해본 결과는 하기 [그림 4-16] ~ [그림 4-22]와 같다. 이때, 시장위험지수($R_{i,t}^{EMA}$)도 시간에 따라 발생하는 변화를 제거하기 위하여 $R_{i,t}^{EMA}$ 시계열 값들의 평균을 빼고 표준편차를 나누는 방식으로 표준화하여 분석하였다.

[그림 4-16] 반도체 최종시장위험지수 계산 결과



[그림 4-16]을 2018년 이후에 전반적인 리스크가 올라간 것을 확인할 수 있으며, 한일 무역 갈등이나 코로나19 대유행처럼 기사화가 많이 된 이벤트에 리스크가 민감하게 반응한 것을 확인할 수 있다. 이는 온라인 뉴스의 특성이 반영된 것으로 향후에 미디어 효과(media effect)를 제어하여 최대한 객관적인 리스크 크기를 측정해야 하는 과정이 필요함을 시사한다. 방법의 특성에 따라서 지수평활화 기법이 적용된 지수 R(EMA)가 적용 전 지수 R 보다는 smoothing 형태로 계산된 것을 확인할 수 있다. 평활화 정도는 지수 고도화 과정에서 다시 확인해야 될 것으로 사료된다. 한편, 최종시장위험지수에서 확인할 수 있듯이, 지수값의 증가 혹은 감소세는 SOX, KRX 지표와는 반대로 이동하는 경향을 보인다. 즉, 본 연구에서 구한 반도체 시장 위험지수값이 증가하는 구간에서 SOX, KRX는 감소하는 추세를 보이며, 반도체 시장 위험지수값이 감소하는 시기에는 SOX, KRX는 증가하는 경향을 보이는 것이다. SOX, KRX는 반도체 기업 주가와 관련된 지수이기 때문에 주가가 높을 때 같이 높아지며, 주가는 일반적으로 시장 상황이 좋을 때 상승하는 경향이 있다. 따라서 본 연구에서 도출한 반도체 시장위험지수와 SOX, KRX 지수가 반대로 이동하는 경향성은 본 연구의 결과가 반도체 시장의 위험성을 식별하는 타당한 지표가 될 수 있는 가능성을 시사한다.

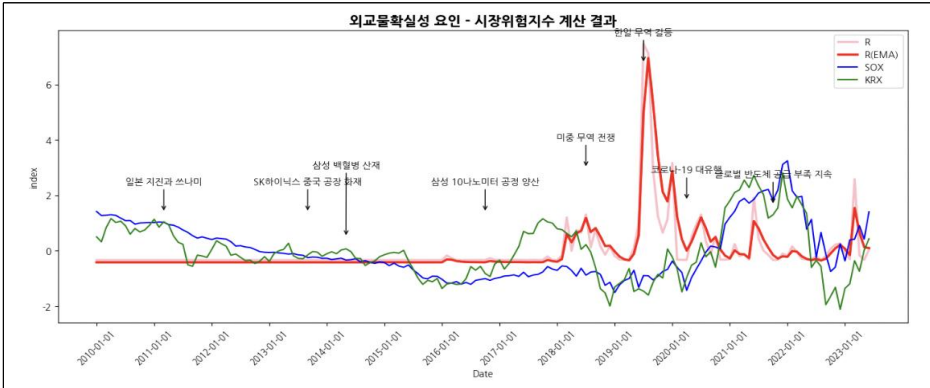
[그림 4-17] 정책불확실성 요인 시장위험지수 계산 결과



[그림 4-17]은 정책불확실성에 따른 시장위험 지수를 산출한 결과이다. 대체적으로 반도체 시장 전체의 시장위험 지수와 유사한 형태를 보이거나 전체 시장 지수에 비하여 코로나19 대유행에서 상대적으로 지수값이 작은 것을 확인할 수 있다. 반면에 한일 무역 갈등에서는 높은 지수값을 보이고 있는 점을 고려하면, 제안 방법론이 정책에 따른 불확실성과 그 외 불확실성간의 차이가 잘 구별되고 있음을 확인할 수 있다.

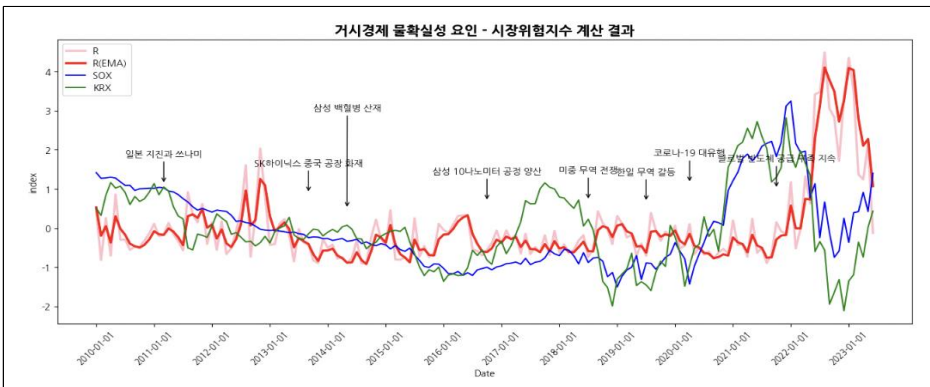
[그림 4-18]에서 외교불확실성 요인의 위험지수(R_2)는 미중 무역 전쟁, 한일 무역 갈등 시점에 높아지지만, 다른 이슈들이 있을 때는 꽤 완만한 상태를 보이는 것을 확인할 수 있다. 2010년에서 2018년 사이에서는 외교문제와 관련된 기사들이 거의 나오지 않았기에 플랫폼한 형태로 지수가 산출되고 있음을 확인할 수 있다. [그림 4-19]에서 반도체시장 불확실성 요인(R_3)은 반도체 수요, 공급과 관련있는 어휘사전으로 구성했기에 글로벌 반도체 공급 부족 지속이 있던 시기에 지수 값이 증가하는 것을 확인할 수 있다.

[그림 4-18] 외교불확실성 요인 시장위험지수 계산 결과

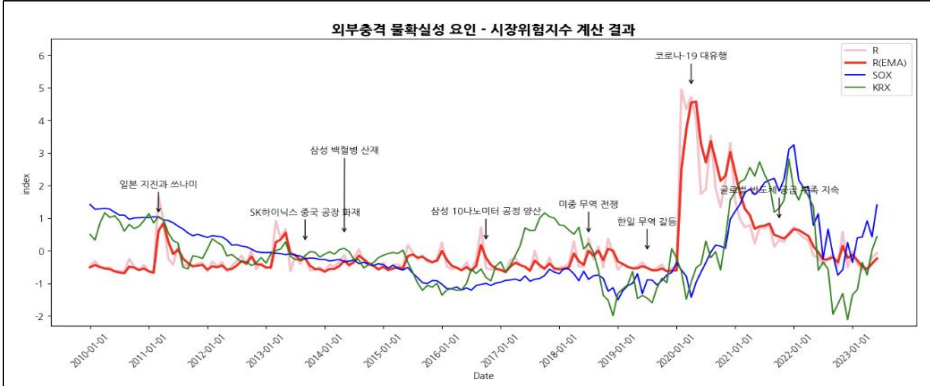


한편, 거시경제 불확실성 요인 위험지수(R_4)의 경우, 2022년 하반기부터 상당히 높아짐을 확인할 수 있는데, 이는 앞에서 기술하였던 2022년 하반기 이후 국내 경제 성장 둔화와 관계가 있는 것으로 사료된다. 즉, 반도체 시장 위험 및 불확실성을 거시경제 상황과 연관되어 작성된 기사들의 수와 해당 기사들의 부정적인 톤이 2022년 하반기부터 본격화되기 시작한 것으로 예상할 수 있다.

[그림 4-19] 거시경제 불확실성 요인 시장위험지수 계산 결과



[그림 4-20] 외부충격 불확실성 요인 시장위험지수 계산 결과



마지막으로 외부충격 불확실성 요인 지수(R_t)는 일본 지진과 쓰나미, 코로나19 대유행 때 높은 값을 보이는 것을 확인할 수 있으며 다른 불확실성 요인과 대비하여 외부충격에 대한 불확실성을 식별하여 반영하고 있음을 확인할 수 있다.

[그림 16]-[그림 20]의 결과를 종합하면, 본 연구진이 제안한 지수가 불확실성 요인별로 시장위험지수를 제대로 구분하여 표현하고 있으며 개별 위험들이 종합 지수 형태로 잘 요약된 것을 확인할 수 있다.

제 3 절 반도체 시장 공급망 근사

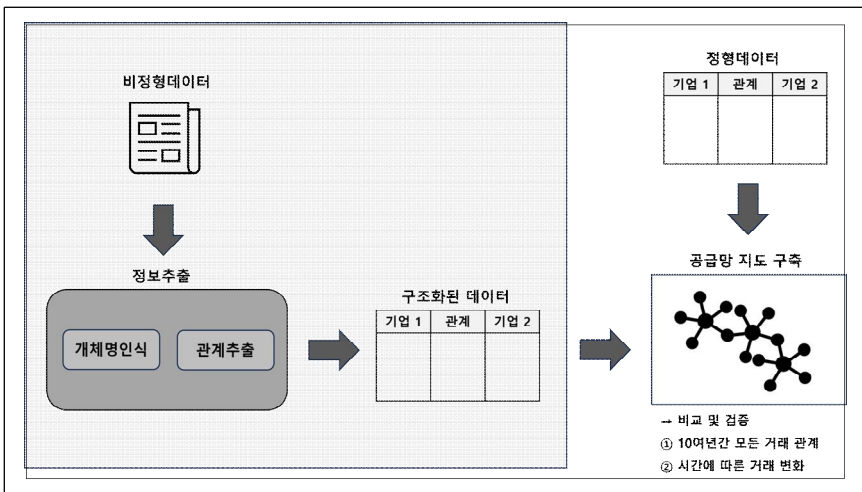
1. 공급망 근사 방법

반도체와 같은 전략산업의 공급망 지도에 대한 요구는 높으나 조사나 정형데이터에 기반한 공급망 지도를 구축하는 것은 현실적으로 한계가 있다. 고품질의 조사데이터와 속보성이 있는 정형데이터를 확보하여 반도체 산업의 현황을 파악할 수 있는 공급망 지도를 구축하는 것은 매우 어렵기 때문이다. 이러한 한계를 극복하기 위하여 본 연구에서는 비정형 뉴스데이터를 활용하여 반도체 산업의 공급망 지도를 근사하는 것을 목표로 한다.

이를 위해 개별 기업을 노드(node)로 기업 간 거래(구매/판매) 관계를 엣지(edge)로 정의하여 베이스 공급망 지도를 근사하는 작업을 실시하였다. 공급망 지도 구축은 반도체 기업 리스트를 만들고, 리스트 내 기업들을 선택한 후 온라인 뉴스를 활용하여 해당 기업들과 거래 관계에 있는 기업들을 식별하여 그래프로 표현하는 방법론을 제안하였다.

제안 방법론의 성능 검증은 기업 수가 많지 않은 칩리스(chipless)와 전방산업과의 긴밀한 관계 때문에 반도체 산업 내 관계 포착이 어려운 팹리스(fabless)는 제외하고 정형데이터 수집이 가능했던 후공정, 소재, 장비 분야의 기업을 6개씩 선별하여 진행하였다. 거래처 및 판매처 데이터²⁷⁾가 포함된 정형데이터를 활용하여 그린 공급망 지도와 2장에서 사용된 네이버 뉴스를 활용하여 작성한 공급망 지도를 비교하였다. 이러한 모의실험 과정을 통하여 비정형데이터를 활용한 공급망 지도 근사의 가능성과 한계를 정리하였으며, 전반적인 작업 모식도는 [그림 4-21]에 요약하였다.

[그림 4-21] 반도체 공급망 근사 프로세스



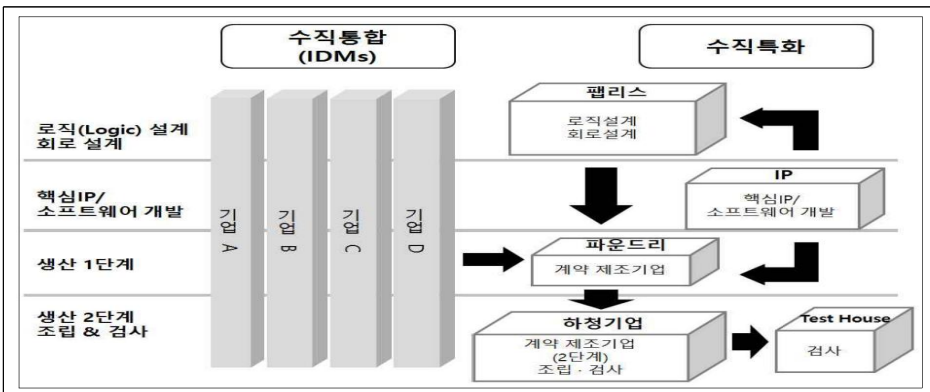
27) ㈜와이즈넷이 제공한 기업 간 거래 정보(부록 참고)

[그림 4-21]에 소개된 작업을 요약하면, (i) 반도체 기업 리스트 만들기, (ii) 문장 단위에서 동시에 등장한 기업쌍을 식별하기, (iii) 어휘사전을 활용하여 식별된 기업의 비즈니스 관계(판매/구매)를 구분하기로 분류할 수 있다. 제안된 알고리즘을 작성된 공급망은 정형데이터를 활용하여 그린 공급망과 비교하는 작업을 최종적으로 진행하였다.

2. 반도체 공급망

반도체 산업의 초창기인 1960년대에는 반도체는 하나의 기업에서 제품 생산을 위한 모든 공정 과정을 진행하는 수직통합구조로 생산되었다. 그러나 1980년대 말에서 1990년대 초, 첨단 공정의 발전과 함께 세계 반도체 산업의 중심이 기존의 D램에서 더 복잡한 반도체 제품으로 확산되며 수직분업구조가 탄생하였다([그림 4-22]). 공급망의 특정 공정을 전문으로 하는 기업의 수가 증가함에 따라 공급망 관리 문제는 더욱 중요해지고 있다. 특히 현재는 주요 국가들이 반도체 개발과 관련된 서로 다른 공정 단계를 국제적으로 분업화하고 있기에 특정 지역에서 반도체를 자급자족하는 것은 어려워지며 기업 간 협력뿐만 아니라 국가 간 협력이 경쟁력의 매우 중요한 요인으로 부각되고 있다.

[그림 4-22] 반도체 산업의 수직통합과 수직분업



반도체 생산 구조는 반도체 제품 유형에 따라 달라지기 때문에 제품의 분류 먼저 살펴볼 필요가 있다. 먼저 반도체 제품은 크게 집적 회로(integrated circuit, IC)와 광·개별소자로 분류된다. 집적 회로란 하나의 판 위에 반도체 개별소자를 전선으로 연결한 회로를 말하며 많은 경우 반도체는 집적 회로만을 의미하기도 한다. 이는 집적 회로와 광·개별소자를 생산하는 프로세스가 다르기 때문이며 본 연구에서도 반도체의 범위는 집적 회로에 한정된다. 집적 회로는 용도에 따라 정보를 저장하는 메모리 반도체와 정보를 처리하는 시스템반도체로 구분된다. 메모리 반도체는 신호를 저장·기억하는 반도체이며 D램, S램, 플래시 메모리 등이 있다. 시스템반도체는 메모리 반도체를 제외한 집적 회로를 말하며, 신호를 정류·증폭·변환·전환·제어하는 등의 다양한 기능을 수행한다.

〈표 4-2〉 반도체 제품의 분류

	대분류	중분류	세분류	설명
집적 회로	메모리 반도체	휘발성 (RAM)	D램	기억 장치의 내용을 일정 시간마다 재생시켜야 되는 동적 RAM. PC의 주메모리.
			S램	전원 공급이 계속되는 한 저장된 내용을 계속 유지하는 정적 RAM. 속도가 빠르나 가격이 비싸 캐시 메모리 등 속도가 중요한 부분에 제한적 사용.
		비휘발성 (ROM)	플래시 메모리	전원이 꺼져도 데이터가 사라지지 않는 롬의 장점과 데이터를 쓰고 지울 수 있는 램의 장점을 가짐.
	시스템 반도체	마이크로 컴포넌트		주로 제어, 연산 기능을 하는 초소형 반도체.
		로직		NOT, OR, AND 등 논리회로로 구성된 반도체.
		아날로그		각종 아날로그 신호를 컴퓨터가 인식할 수 있는 디지털 신호로 바꿔주는 반도체.
광·개별소자	광소자		전자 소자들을 광소자들과 함께 단일 기판 위에 집적시켜 광신호를 전기신호로 변환.	
	개별소자		제품 내에서 단일한 기능을 하는 소형 전자 반도체 부품.	

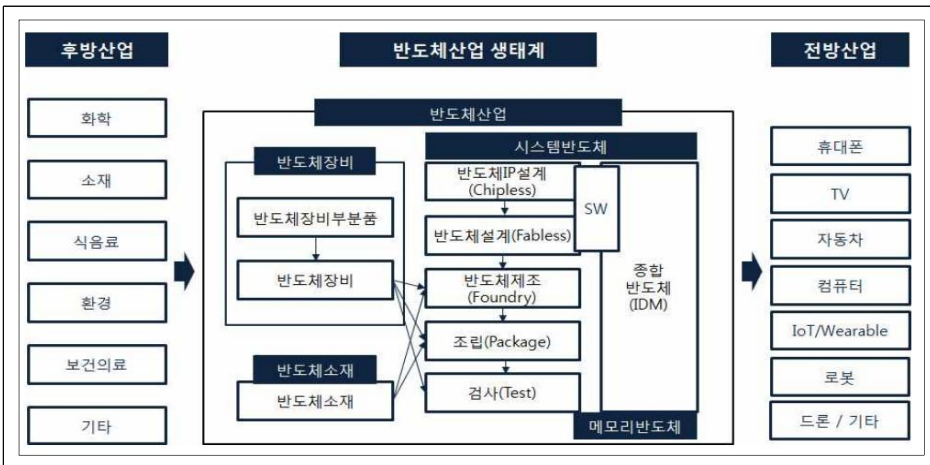
자료: 저자 작성

반도체 생산 프로세스는 [그림 4-23]과 같이 크게 설계, 제조, 조립, 검사하는 기업과 제조시설 구축에 필요한 장비를 공급하는 기업, 제조과정에서 필요한 재료를 생산하는 기업으로 나누어 볼 수 있으며, 메모리 반도체와 시스템 반도체에 따라 생산 구조가 구분된다.

메모리 반도체는 대부분 반도체 설계와 제조를 동시에 하는 IDM(integrated device manufacturer, 종합반도체 기업)에서 소품종 다량 생산된다. IDM은 일부 생산 능력이 부족한 경우에 위탁을 맡기는 경우를 제외하고는 설계, 생산, 유통에 이르는 전 과정을 독자적으로 수행한다. 인텔, 삼성전자, SK하이닉스가 이에 속하며, 우리나라는 메모리 산업에서 세계적인 경쟁력을 가지고 있다.

반면 시스템 반도체는 개별제품에 적합한 기능을 설계하여 상당히 큰 규모로 생산되는 CPU와 GPU를 제외하고는 수직분업구조로 다품종 소량 생산된다.

[그림 4-23] 반도체 산업의 수직통합과 수직분업



자료: 반도체산업협회

반도체 산업 내 기업을 좀 더 세부적으로 살펴보면 [표 4-11]과 같다. 먼저 칩리스는 반도체 설계에 사용되는 기능 블록인 IP를 설계하는 기업이다. 반도체 회로가 복잡해짐에 따라 기능별 설계도(IP) 형태의 거래가 확대되고 있고 고유 기능

은 자체 설계, 나머지는 기존 IP를 활용하는 전략이 일반적으로 활용되고 있다. 비슷하게 팹리스 또한 반도체 설계 전문기업으로 반도체 회로도를 설계하여 파운드리에서 위탁 생산한다. 통상 칩리스를 팹리스의 범주 안에 포함시키기도 하나 칩리스는 설계 라이선스만 제공할 뿐, 자사의 브랜드로 판매하지는 않는 반면 팹리스는 칩을 설계해서 판매와 유통까지 한다는 차이가 있다.

파운드리(foundry)는 제조시설을 구축하여 반도체 제조만을 전문으로 하는 기업이며, 기업규모나 투자 금액이 크다. 또한 팹리스 회사와의 전략적 협력관계가 매우 중요하다.

조립(package)은 가공된 웨이퍼를 절단하여 전자제품에 부착하기 쉽도록 하는 공정이며 검사는 제조가 완료된 웨이퍼 또는 칩의 불량여부를 판단한다. 두 공정을 담당하는 기업을 통틀어 패키징 및 테스트 전담 기업(outsourced semiconductor assembly and test, OSAT) 기업 또는 후공정 기업이라고 부르기도 한다.

반도체 장비 산업은 반도체 제조시설 구축에 필요한 장비를 개발하고 생산하는 산업이다. 반도체 장비에는 웨이퍼 공정에 사용되는 전공정 장비, 반도체 검사장비, 반도체를 조립하는 데 사용되는 조립 장비, 대부분의 장비에 공통적으로 사용되는 기반 장비로 구분된다. 또한, 소재 산업은 웨이퍼, 포토 레지스트, 마스크, 가스 등 반도체 제조에 필요한 재료를 개발하고 생산하는 산업이다.

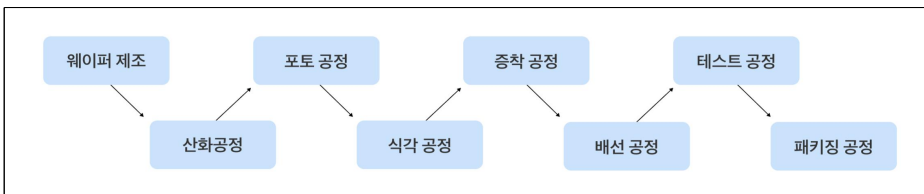
전 세계적으로 설비가 분산되고 특정 단계를 전문으로 하는 기업의 수가 증가함에 따라 반도체 공급망 관리의 중요성은 더욱 커지고 있다. 그러나 역설적으로 체계적인 분업화는 산업 구조를 복잡하게 하여 가시성을 제한하였다.

다른 산업과 마찬가지로 반도체 공급 시스템의 가장 기본적인 지도 역시 공정 기반 지도이다. 흔히 반도체 공정을 [그림 29]와 같이 8대 공정(웨이퍼 제조, 산화, 포토, 식각, 증착, 금속 배선, 테스트, 패키징)으로 분류한다. 다만 현대의 반도체 공정이 극도로 고도화되면서 이 8대 공정의 분류법이 부적절하다는 지적도 있다.

〈표 4-3〉 반도체 생태계 분야별 특징 및 주요 기업

분야	사업 특성	주요 기업
종합반도체 (IDM)	- 칩 설계에서 제조 및 테스트까지 일괄 공정 체제 구축하여 직접 수행 - 기술력과 규모의 경제를 통한 경쟁 확보 - 거대 투자의 고위험 고수익 형태	Intel, 삼성전자, SK 하이닉스, Micron
IP전문 (Chipless)	- 설계기술 R&D 전문 - IDM이나 팹리스에 IP 제공 - 칩 설계용 Architecture IP, 칩 생산용 Physical IP, System IP, 일반기술 IP 등	SRM, Synopsys, Mentor, 칩스앤미디어
설계전문 (Fabless)	- 칩의 설계만 전문으로 하는 업체 - 고위험의 거대투자를 회피할 수 있으나 위탁제조의 비용 부담 필요 - 고도의 시장예측이 필요하며 생산의 최소물량수준 예측 필요	Qualcom, 미디어텍, 실리콘웍스
제조전문 (Foundry)	- 주문방식에 의해 칩 생산만 전문 - 설계전문업체로부터 위탁제조	TSMC, UMC, SMIC, DB하이텍
조립 및 검사 (OSAT)	- 일반적으로 메모리 제조는 자체적으로 조립하지만, 시스템 분야는 다양한 제품을 모두 패키징 처리할 수 없어 외부 위탁 - 반도체 테스트는 전수검사이므로 자체적으로 조립업체에서 하지만, 검사장비가 고가이므로 다양한 칩을 모두 할 수 없어 테스트 전문업체에 위탁	ASE Test, AMkor
장비	- 반도체 제조에 필요한 장비 - 생산공정 기술개발도 주도	TEL, ASML, Lamresearch, 세메스
소재	- 반도체 제조에 사용되는 소재의 개발 및 생산	동진세미켄, 실트론

[그림 4-24] 반도체 8대 공정



자료: TIGER ETF(2023)

공정 기반 지도를 바탕으로 기업 운영에서 효율적인 반도체 공급 시스템을 관리하기 위한 연구가 진행됐다. 대표적으로 Callarman et al.(2004)은 설문지를 기반으로 반도체 공급망 통합을 위해 고려해야 하는 연구 의제를 개발하였으며, 수요 관리·공급망 설계·계획 및 실행이 중요한 영역으로 나타났다. 또한, Sun et al.(2010)은 고객 리드타임(lead time)²⁸⁾과 정시증권 납품의 중요성에 기반하여 공급망 설계 시 언제 재고를 보충해야 하는지 결정하는 계층적 의사결정 지원 프레임워크를 제안했다.

반면, 반도체 산업에서 공급 시스템 지도의 상위 단계인 공급사슬 지도나 공급망 지도의 구축이나 관련 연구는 제한적이다. 이는 앞서 언급된 바와 같이 기업 내부데이터인 거래 관계 데이터에 접근이 어려우며, 기업 내부에서도 직접적으로 거래하고 있는 1차 업체까지만 파악하고 2차 이상은 모르는 경우가 다수이기 때문이다.

이러한 반도체 공급망의 가시성 부족 때문에 많은 공급망 보고서는 분야별 대표적인 기업을 선정하여 분석하고 있다. 예를 들어 글로벌 반도체 공급망 재편의 대표적인 연구보고서였던 「공급망 100일 조사 보고서(100-Day Reviews under Executive Order 14017)」에서도 해당 분야의 전문가들이 미국 내 반도체 공급망에서 취약 지점을 정성적으로 평가하였다. 또한, 국내 이베스트투자증권 리서치센터의 산업보고서²⁹⁾는 8대 공정보다 세분화한 반도체 공정 기반 지도에 대표 기업을 연결하여 각 단계에서 영향력을 가지는 반도체 기업을 한눈에 볼 수 있었다. 그러나 이는 전반적인 산업 흐름만 보여줄 뿐, 기업 간 관계는 식별할 수 없어 공급망 내 취약 지점을 도출할 수 없다.

최근 논의되고 있는 글로벌 공급망 재편에 대응하기 위해서는 반도체 산업에 종사하는 전체 기업 간 관계를 조망할 수 있으며, 즉각적으로 산업의 변화를 반영하는 지도를 구축해야 한다. 따라서 본 연구는 기존의 반도체 공급망 지도를 보완하

28) 고객의 주문부터 납품까지의 소요시간

29) https://msg.ebestsec.co.kr/eum/20200914_B4510_31475_40.pdf

여 (i) 지속적으로 자동화할 수 있고 (ii) 기업 간 관계성과 공급망 내 취약 지점을 도출할 수 있는 반도체 공급망 네트워크 구축 방법을 찾고자 한다.

3. 반도체 공급망 근사 네트워크

가. 관계 데이터 추출

1) 개체명(기업명) 인식

본 분석의 목적은 지속적으로 산업의 변화를 반영할 수 있는 국내 반도체 산업의 공급망 구축에 있으며 이를 위해 비정형데이터에서 정보추출을 통해 국내 반도체 기업 간 관계 데이터를 확보하고자 한다. 정보추출은 엔티티와 엔티티 간의 관계를 식별하고 추출하는 것이 목표로 하위 작업 중 개체명 인식(named entity recognition)과 관계 추출(relation extraction)을 중점으로 연구가 진행된다. 개체명 인식은 사람, 장소, 기관, 날짜 등 명명된 개체를 텍스트로 식별하는 작업이다. 본 분석은 기업과 기업 간의 관계를 추출하는 것이 목표이므로 문장 내 기업이라는 엔티티 인식 시스템을 구축하는 과정이 필요하다. 다만 앞서 언급한 바와 같이 이번 1차 연도 연구에서는 정형 데이터를 벤치마크로 활용함으로써 거래 정보 수집이 가능한 후공정, 소재, 장비 기업으로 분석 대상 기업이 한정되었다.

또한 Raw 데이터인 뉴스에서는 기업명이 다양한 방식으로 표현될 수 있다. 예를 들어서 미국 반도체 제조 기업인 'BROADCOM'의 경우 법인 형태를 포함하여 'BROADCOM Inc.'로도 쓸 수 있으며, 한글로 브로드컴이라고 표현할 수도 있다. 본 분석에서는 단순화를 위해 기업명은 종목명만 고려하였으며, 종목명이 존재하지 않는 비상장회사나 해외 기업의 경우 기업명에서 법인 형태를 제외하는 방식으로 통일하였다. 분석 대상 기업은 <표 4-12>와 같다.

2) 관계 추출

관계 추출의 목표는 각각의 엔티티 둘 사이의 관계 유형을 추출하여 구조화된 데이터로 저장하는 것이다. 관계 추출은 특정한 방법론이 정해져 있지 않고 코퍼

스(corpus)에 맞는 방법론을 사용한다. 그 중 많이 사용되는 방법인 규칙 기반 방법은 사람이 직접 규칙을 생성하며, 비정형 텍스트에서 언어적인 분석을 통해 표현 형식을 찾아 관계 규칙을 정의하고, 정의된 규칙을 사용하여 관계 튜플을 추출한다. 사람이 직접 정의하는 만큼 규칙의 질은 높을 것으로 기대할 수 있지만, 실수로 인해 놓치거나 생각하지 못한 패턴이 있을 수 있다.

본 분석에서는 규칙에 기반 관계 규칙을 정의하기 위해 동시 발생행렬(co-occurrence matrix) 접근법을 사용하였다. 동시 발생행렬은 단순히 같은 문장에서 두 단어가 동시에 발생한 횟수를 나타낸다. 동시 발생행렬을 이용한 접근은 구문적 또는 의미론적 역할을 참고하지 않고 텍스트에서 관련 키워드의 존재만을 추구하기 때문에 문장 구조를 중심으로 접근하는 것보다 훨씬 간단하다.

먼저 두 기업 간 거래 관계를 나타내는 규칙을 파악하기 위해 897만여 개의 전체 문장에서 <표 4-12>의(중심 기업, 거래처 기업)의 조합이 언급된 문장의 횟수로 동시 발생행렬이 생성되었다. 이때 정형 데이터는 한 기업(중심 기업)에 대한 거래처 데이터이므로 그와 형식을 맞추기 위해 동시 발생행렬 또한 '6×N' 행렬로 구축되었다. 행 개수인 6은 중심 기업의 개수, 열의 개수인 N은 해당 중심 기업의 거래처 후보 개수이다. 또한 행렬은 3가지 기업 유형(후공정/장비/소재)에 따라 각각 만들어졌다.

〈표 4-4〉 분석 대상 기업

후공정	
중심 기업(6개)	거래처 기업(41개)
하나마이크론, 네패스, 시그네틱스, SFA 반도체, 한양디지텍, LB세미콘	삼성전기, 대덕전자, 엠케이전자, 삼성전자, 텔레칩스, SK하이닉스, LX세미콘, 심텍, 드림텍, 다이얼로그 세미컨덕터 코리아, BKE&T, 유니트론텍, 엘티메탈, FUJITSU, MXIC, 네패스이앤씨, 에스앤에프씨, 제일E&S, LG디스플레이, 한덕화학, TAMA CHEMICAL, 씨앤씨테크, 제이에스알 마이크로 코리아, SONY, 메탈로코팅스 코리아, 지엠텍, 매그나칩반도체, DB하이텍, N.E. CHEMCAT, EEJA, 유창금속공업, 실리콘마이터스, Unimicron, Tripod, APCB Electronics, 에스케이브로드밴드, 코리아씨키트, LG헬로비전, BROADCOM, 아나패스
장비	
중심기업(6개)	거래처 기업(41개)
원익IPS, 원익홀딩스, 제우스, AP시스템, 피에스케이홀딩스, 이오테크닉스	삼성전기, 대덕전자, 엠케이전자, 삼성전자, 텔레칩스, SK하이닉스, LX세미콘, 심텍, 드림텍, 테크엘, 다이얼로그세미컨덕터코리아, BKE&T, 유니트론텍, 엘티메탈(구.희성금속), FUJITSU, MXIC, 네패스이앤씨, 에스앤에프씨, 제일E&S, LG디스플레이, 한덕화학, TAMA CHEMICAL, 씨앤씨테크, 제이에스알마이크로코리아, SONY, 메탈로코팅스코리아, 지엠텍, 매그나칩반도체, DB하이텍, N.E. CHEMCAT, EEJA, 유창금속공업, 실리콘마이터스, Unimicron, Tripod, APCB Electronics, 에스케이브로드밴드, 코리아씨키트, LG헬로비전, BROADCOM, 아나패스
소재	
중심기업(6개)	거래처 기업(40개)
동진썬미캠, 엠케이전자, 한솔케미칼, 원익QnC, 이엔에프테크놀로지, 후성	삼성전기, 대덕전자, 삼성전자, 텔레칩스, SK하이닉스, LX세미콘, 심텍, 드림텍, 테크엘, 다이얼로그세미컨덕터코리아, BKE&T, 유니트론텍, 엘티메탈, FUJITSU, MXIC, 네패스이앤씨, 에스앤에프씨, 제일E&S, LG디스플레이, 한덕화학, TAMA CHEMICAL, 씨앤씨테크, 제이에스알마이크로코리아, SONY, 메탈로코팅스코리아, 지엠텍, 매그나칩반도체, DB하이텍, N.E. CHEMCAT, EEJA, 유창금속공업, 실리콘마이터스, Unimicron, Tripod, APCB Electronics, 에스케이브로드밴드, 코리아씨키트, LG헬로비전, BROADCOM, 아나패스

만약 한 문장 내에 ‘하나마이크론’과 ‘삼성전자’가 동시에 언급되면 해당하는 셀에 1이 더해진다. 따라서 구축된 동시 발생행렬의 각 셀은 해당 조합이 언급된 빈도로 구성된다. 그러나 이러한 해당 조합의 빈도는 기업 간 관계성의 강도를 나타내는 것은 아니므로 한 개 이상의 언급된 문장이 있는 경우 1, 언급된 문장이 아예 없는 경우 0으로 변환되었다(V0). <표 4-5>는 V0 행렬의 예시로 후공정 기업의 결과이다.

<표 4-5> V0 행렬: 후공정

	삼성전자	SK하이닉스	삼성전기	대덕전자	엠케이전자	...	테크셀
하나마이크론	1	1	1	1	1	...	1
네패스	1	1	1	1	1		0
시그네틱스	1	1	1	1	1		0
SFA반도체	1	1	1	1	1		1
한양디지텍	1	1	1	1	1		0
LB세미콘	1	1	0	0	0		0

또한, 거래 관계 규칙을 나타내는 후보 단어를 선정하기 위해 [표 5](중심 기업, 거래처 기업) 조합이 언급된 문장에서 기업 간 거래 관계를 나타내는 것으로 보이는 단어가 선택되었다. 그러나 앞서 언급된 바와 같이 사람이 직접 규칙을 정의하는 방식은 놓치는 패턴이 많으므로 더 많은 패턴을 정의하기 위해 데이터 분석을 통해 보강한다. 실제로 본 분석에서도 단순 문장을 보고 선정한 규칙만으로는 많은 관계성이 포착되지 않았다. 따라서 조합된 언급이 문장을 형태소 단위로 토큰화하여 빈도순으로 나열 후 기업 간 거래 관계를 나타내는 것으로 보이는 규칙 단어를 추가로 선정하였다.

〈표 4-6〉 거래 관계 규칙 단어

선정 방식	거래 관계 규칙 단어
문장에서 직접	'공급', '계약', '맺었다', '체결', '성사', '발주', '납품', '판매', '출고'
빈도분석	'거래처', '합작', '맺고',

중심 기업과 거래처 기업 간 거래 관계는 해당 조합이 언급된 문장에서 [표 4-14]에서 선정한 거래 관계 규칙 단어가 두 가지 이상 포함된 경우로 정의되었다. 이는 한국어가 어간에 접사가 붙는 것만으로 단어나 문장의 의미가 쉽게 변하는 교착어이므로 한 가지 단어의 포함 여부만으로는 관계를 정의하기는 어렵기 때문이다. 정의한 규칙을 적용하여 중심 기업과 거래 기업 간 거래 관계를 포착하는 동시에 발생행렬(이하 V1)을 〈표 4-7〉과 같이 구축하였다. V1의 셀이 1이라는 것은 해당(중심 기업, 거래처 기업) 조합이 언급된 문장 중 거래 관계 규칙 단어가 두 가지 이상 포함된 경우가 한 개 이상이라는 뜻이다.

〈표 4-7〉 V1 행렬: 후공정

	삼성전자	SK하이닉스	삼성전기	대덕전자	엠케이전자	...	테크셀
하나마이크론	1	1	1	0*	0*	...	0*
네패스	1	1	1	1	1		0
시그네틱스	1	1	0*	1	0*		0
SFA반도체	1	1	1	0*	0*		0*
한양디지텍	1	1	0*	1	0*		0
LB세미콘	0*	0*	0	0	0		0

* V0에서는 포착되었으나 V1에서는 포착되지 않음

마지막으로 '6×N'로 구축된 V1 행렬을 엔티티 간의 의미적 관계를 나타내는 관계형 튜플 형태를 통해 구조화된 데이터로 나타냈다. 관계형 튜플 표현은 엔티티 1과 엔티티 2에 대한 관계는(엔티티 1, 관계, 엔티티 2)로 표현된다. [표 8]은 후공정의 V1을(중심 기업, 관계, 거래처 기업)의 구조화된 데이터로 나타낸 결과의

일부이다.

〈표 4-8〉 구조화된 데이터: 후공정

중심 기업	관계	거래처 기업	중심 기업	관계	거래처 기업
하나마이크론	거래	삼성전자	네패스	거래	삼성전자
하나마이크론	거래	SK하이닉스	네패스	거래	텔레칩스
하나마이크론	거래	심텍	네패스	거래	아나패스

나. 공급망 네트워크

본 연구에서는 네트워크 분석을 활용하여 국내 반도체 공급망 현황을 지도로 도식화한다. 네트워크 분석은 특정 사회 주체들의 관계를 도식화하는 분석기법으로 다양하고 복잡하게 얽혀 있는 사회관계의 핵심을 일목요연하게 보여준다.

여기서는 비정형데이터로부터 얻은 근사 네트워크를 정형 데이터로부터 얻은 네트워크 결과와 비교하여 근사가 잘 이루어졌는지 검증한다. 또한, 분석은 2010년과 2023년까지의 전체 기간에서 포착된 거래 관계와 시간에 따른 거래 관계 변화를 포착하는 두 가지 방법으로 진행되었다. 기업 간 공급망을 도식화하기 위해 R 소프트웨어(version 4.3.1)의 igraph 패키지가 사용되었다. 본 패키지는 사회적 네트워크 분석을 목적으로 개발된 패키지이다.

네트워크 그림에서 동그라미, 즉 노드는 각 기업을 의미하며, 각 노드를 연결하는 선, 즉 엣지는 기업 간에 직접적인 거래 관계가 있음을 뜻한다. 주요 기업은 붉은색으로 처리하였으며, 반도체를 위탁생산하는 파운드리 기업인 삼성전자, SK하이닉스, DB하이텍은 노란색으로 구분하였다. 또한, 정형 데이터로부터의 네트워크 결과는 거래처와 판매처를 구분하는 방향이 있으므로 진입(in-degree) 엣지와 진출(out-degree)엣지로 구분될 수 있다. 반면 비정형데이터에서 얻은 결과는 거래처와 판매처를 구분하는 거래 방향이 없으며 두 기업 간 거래 관계의 가능성 유무만 포착된다. 따라서 방향이 존재하지 않으며 기업 간 관계는 오직 실선으로 표현된다.

1) 전체 기간

2010년부터 2023년까지 전체 기간에서 포착된 기업 간 거래 관계를 공급망 지도로 시각화하였다. [그림 4-24]은 후공정 기업에 대해 정형데이터로부터 얻은 공급망 지도 결과이고, [그림 4-25]은 비정형 데이터로부터 얻은 공급망 지도 결과이다. 근사 성능을 파악하기 위해 <표 4-17>에 후공정 기업에 대해 비정형 데이터로부터 거래 관계가 있다고 나타난 기업의 개수를 정형 데이터와 비교하였다. 성능은 기업에 따라 크게 다르게 나타나는데 가장 높은 시그네틱스의 경우 67%의 거래 관계(6개 중 4개)를 포착하였으나 LB세미콘의 경우 어떤 거래 관계도 포착하지 못했다.

<표 4-9> 근사 결과 비교: 후공정

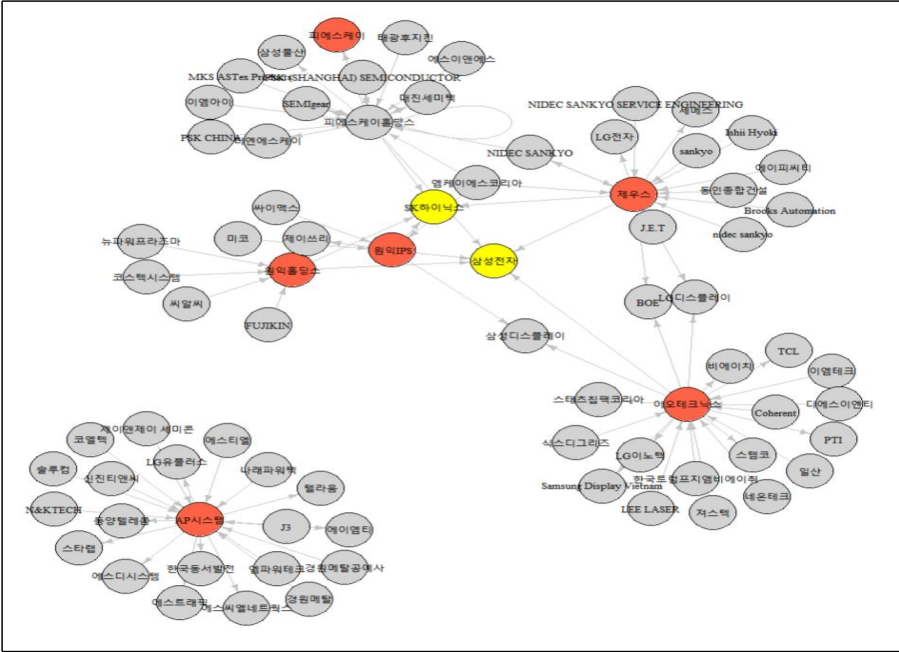
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
하나마이크론	4	11	0.36
네패스	3	11	0.27
시그네틱스	4	6	0.67
SFA반도체	2	7	0.29
한양디지텍	2	9	0.22
LB세미콘	0	11	0.00

다음으로 [그림 4-26]는 장비 기업에 대해 정형 데이터로부터 얻은 공급망 지도 결과이고, [그림 4-27]은 비정형 데이터로부터 얻은 공급망 지도 결과이다. 근사 성능을 파악하기 위해 [표 4-18]에 장비 기업에 대해 비정형 데이터로부터 거래 관계가 있다고 나타난 기업의 개수를 정형 데이터와 비교하였다. 성능은 전반적으로 후공정보다 떨어졌는데 가장 높은 원익홀딩스의 경우 33%의 거래 관계(6개 중 2개)를 포착하였으나 AP시스템의 경우 어떤 거래 관계도 포착하지 못했다.

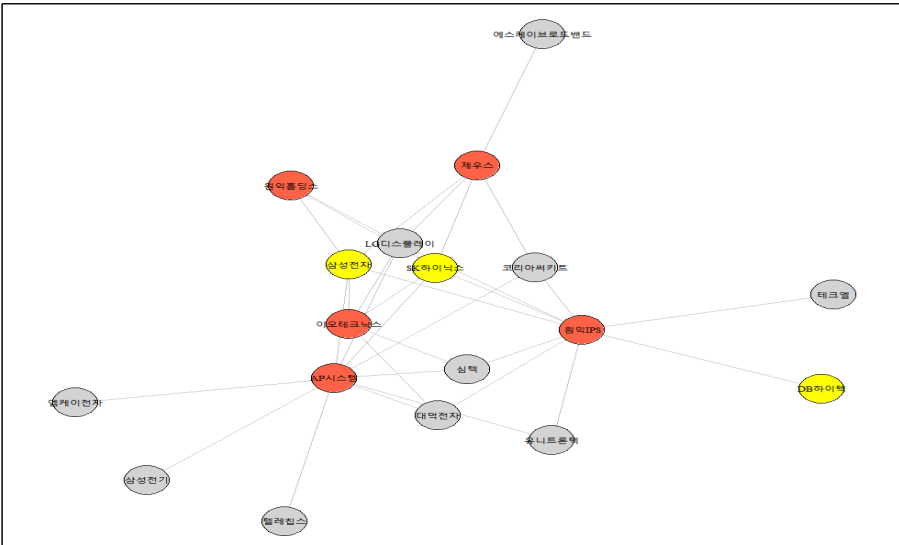
〈표 4-10〉 근사 결과 비교: 장비

기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
AP시스템	0	20	0.00
원익IPS	2	7	0.29
원익홀딩스	2	6	0.33
이오테크닉스	2	20	0.10
제우스	3	16	0.19
피에스케이홀딩스	2	16	0.13

[그림 4-27] 공급망 지도 시각화 결과: 정형 데이터-장비



[그림 4-28] 공급망 지도 시각화 결과: 비정형 데이터-장비



〈표 4-11〉 근사 결과 비교: 소재

기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
동진씨미캡	3	10	0.30
엠케이전자	1	3	0.33
원익QnC	2	14	0.14
이엔에프테크놀로지	3	12	0.25
한솔케미칼	0	9	0.00
후성	2	16	0.13

마지막으로 [그림 4-28]는 소재 기업에 대해 정형 데이터로부터 얻은 공급망 지도 결과이고, [그림 4-29]는 비정형 데이터로부터 얻은 공급망 지도 결과이다. 근사 성능을 파악하기 위해 [표 11]에 소재 기업에 대해 비정형 데이터로부터 거래 관계가 있다고 나타난 기업의 개수를 정형데이터와 비교하였다. 성능은 전반적으로 후공정보다 떨어졌는데 가장 높은 엠케이전자의 경우 33%의 거래 관계(3개 중 1개)를 포착하였으나 한솔케미칼의 경우 어떤 거래 관계도 포착하지 못했다.

2) 기간별 공급망 네트워크

반도체 공급망의 작성 기간에 따른 변화를 포착하기 위해 2010년부터 2023년을 3개의 기간(2010~2014/2015~2019/2020~2023)으로 나누어 공급망 지도를 작성하였다. 정형 데이터에 12년, 18년, 22년에 보고된 데이터들이 대다수를 차지하였고, 각 기간을 나누어 살펴보기 위해 위와 같이 기간을 나누었다. 각 기간에 대해 후공정, 장비 소재별로 공급망 지도를 구축한 그림은 [부록]에 첨부하였다.

〈표 4-20〉, 〈표 4-21〉, 〈표 4-22〉는 각각 후공정, 장비, 소재 기업에서 기간에 따라 비정형 데이터로부터의 근사 성능이 어떻게 나타나는지 정리한 것이다. 전반적으로 공급망 작성 기간에 따라 거래 관계 네트워크는 유동적으로 변화하고 있는 것으로 보이며, 근사 네트워크 또한 관계 포착률이 높은 기업 조합에 대해서는 시간에 따른 변화도 잘 추적하고 있는 것으로 보인다.

근사 성능은 관계 포착률 최대 10~20%로 전체 기간에 대해 진행했을 때보다

더 떨어져 보일 수도 있으나 전체 기간에 대해서는 3개의 기간 중 한 번이라도 거래 관계를 포착하면 해당 기업 조합의 관계를 파악한 것으로 계산되기 때문에 당연한 결과이다.

공급망 작성 기간에 따라 네트워크가 유동적으로 변화하는 반면, 관계 포착률은 상대적으로 높은 기업이 계속 높고 낮은 기업이 계속 낮은 경향을 보인다. 이는 관계 포착률이 낮은 이유가 개체명 식별이 미흡했거나 뉴스 노출 빈도가 기업별 특성에 기인하기 때문인 것으로 사료된다. 즉, 문장에서 관계를 추출하기 어려운 특정 기업명이 있을 수 있으며, 매출 규모 혹은 기업의 마케팅 노력에 따라서 뉴스에 노출되는 정도가 달라지는 것이다. 이 외에도 실제로는 거래 관계가 있지만 근사 결과에서 포착되지 않는 경우에 대한 분석이 필요한 것으로 사료된다.

〈표 4-12〉 기간별 공급망 지도 성능 비교: 후공정

2010~2014			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
하나마이크론	3	6	0.50
네패스	1	5	0.20
시그네틱스	3	6	0.50
SFA반도체	0	6	0.00
한양디지텍	1	5	0.20
LB세미콘		미등장	
2015~2019			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
하나마이크론	2	8	0.25
네패스		미등장	
시그네틱스		미등장	
SFA반도체		미등장	
한양디지텍	1	6	0.17
LB세미콘	0	6	0.00
2020~2023			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
하나마이크론	2	7	0.29
네패스	3	7	0.43
시그네틱스		미등장	
SFA반도체	0	1	0.00
한양디지텍		미등장	
LB세미콘	0	7	0.00

〈표 4-13〉 기간별 공급망 지도 성능 비교: 장비

2010~2014			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
AP시스템	0	8	0.00
원익IPS		미등장	
원익홀딩스	0	6	0.00
이오테크닉스	1	8	0.13
제우스	2	6	0.33
피에스케이홀딩스	0	5	0.00
2015~2019			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
AP시스템	0	9	0.00
원익IPS	2	7	0.29
원익홀딩스		미등장	
이오테크닉스	1	11	0.09
제우스	1	8	0.13
피에스케이홀딩스	1	8	0.13
2020~2023			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
AP시스템	0	8	0.00
원익IPS		미등장	
원익홀딩스		미등장	
이오테크닉스	1	7	0.14
제우스	2	7	0.29
피에스케이홀딩스	0	5	0.00

〈표 4-14〉 기간별 공급망 지도 성능 비교: 소재

2010~2014			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
동진썬미켄	2	8	0.25
엠케이전자	1	3	0.33
원익QnC	2	9	0.22
이엔에프테크놀로지	0	5	0.00
한솔케미칼	0	6	0.00
후성	1	7	0.14
2015~2019			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
동진썬미켄	2	7	0.29
엠케이전자		미등장	
원익QnC	0	5	0.00
이엔에프테크놀로지	2	9	0.22
한솔케미칼	0	5	0.00
후성	2	7	0.29
2020~2023			
기업명	V1에 잡힌 기업수(A)	정형데이터 기준 기업수(B)	비율(A/B)
동진썬미켄	3	7	0.43
엠케이전자		미등장	
원익QnC		미등장	
이엔에프테크놀로지	2	6	0.33
한솔케미칼	0	5	0.00
후성	2	7	0.29

3) 기업 간 관계 포착률이 차이 나는 이유

전체 기간에 대한 결과와 시간에 따른 변화에 대한 결과를 종합해보면 어떤 중심 기업에 대해서는 50% 이상의 실제 거래 관계를 포착하는 반면, 어떤 기업에 대해서는 0%로 관계를 아예 포착하지 못하는 등 기업마다 큰 차이를 보였다. 이런 결과가 나타난 이유를 V0(기업명을 포함한 문장)에서는 있었으나 V1(규칙 단어를 포함한 문장)에서는 사라진 문장을 정성적으로 분석하여 정리하고자 한다.

첫째로 기업명이 외래어거나 영문이 포함되는 경우 등 기업명이 통일되지 않음

면 관계 포착률이 더 낮다. 예를 들어 ‘네패스’가 뉴스 문장 상에서 ‘네페스’로 표기되어 포착되지 않는 경우가 있었으며, ‘원익QnC’도 ‘원익큐엔씨’, ‘원익qnc’ 등 다양하게 표기된 문장은 고려하지 못했다. 네패스는 전체 기간 기준 후공정에서 27%, 원익Qnc는 소재에서 14%의 낮은 관계 포착률을 보였다.

둘째로 포착률은 회사 규모에 편향될 수 있다. 영향력이 큰 기업일수록 더 많은 관련 뉴스가 작성되며 비즈니스 관계에 대한 정보가 더 많이 수집될 수 있다. 또한 ‘하나마이크론과 SFA 반도체, 시그네틱스 등이 삼성전자로부터 후공정 물량을 받아와 사업을 수행한다.’라는 문장처럼 규모가 작은 기업의 경우 기타로 포함되어 문장에서 누락 될 수도 있다.

본 분석 결과를 보완하여 실제 거래 관계가 있었으나 포착이 안 된 경우의 문장을 중심으로 규칙 단어를 <표 4-15>와 같이 확장하였다. 이는 추후 연구에서 반영할 수 있을 것으로 사료된다. 또한, 후속 연구에서는 어휘사전 기반 방법외에 머신러닝이나 딥러닝 모형을 통하여 기업 간 관계를 식별하는 방법을 활용할 수 있을 것이다.

<표 4-15> 확장된 거래 관계 규칙 단어

규칙 단어	예시 문장
고객사	네패스는 반도체 부품 및 터치패널 등을 생산하는 업체로 삼성전자와 LG디스플레이 등이 주요 고객사다.
고객	하나마이크론은 반도체 후공정 패키징 업체로 주 고객은 삼성전자, SK하이닉스 등이다.
수주	하나마이크론은 지난해 11월 SK하이닉스에 D램, 낸드 메모리 패키징, 테스트 후공정 물량을 수주했다.
협력사	삼성전자 협력사인 SFA반도체, 네패스, 이오테크닉스, 하나마이크론과 SK하이닉스 협력사인 에이티세미콘, 하이셀, 윈팩 등은 모두 올해 실적이 대폭 개선됐다.
매출처	WLP는 비메모리 반도체의 주요 패키징 기술로서 네패스는 삼성전자와 미국 브로드컴을 주요 매출처로 확보하고 있다
주거래처	LB세미콘은 이후 구 대표의 역량에 힘입어 LG디스플레이를 주거래처로 삼고 본격적인 사세 확장에 나섰다.

제5장 결론

제1절 시사점

반도체는 세계 경제에서 필수적인 요소이다. 비록 반도체는 미국에서 개발되었지만, 그들의 제조업은 한국, 대만 및 기타 아시아 국가로 이동했다. 오늘날 미국과 그 서방 동맹국들은 그렇게 중요한 부품을 생산할 수 있는 위험성을 먼 곳에 두고 인식하고 있다. 지진이나 화재, 또는 전쟁으로 인해 서구의 생산자들은 이러한 중요한 투입물에서 벗어날 수 있다. 따라서 미국과 EU는 국내 반도체 제조업 육성을 모색하고 있다. 2001년부터 2011년까지 구글의 CEO를 역임한 에릭 슈미트가 인터뷰에서 언급한 바와 같이, 그 문제에 돈을 퍼붓는 것만으로는 성공할 수 없다.

과거 일본, 한국, 대만 기업의 성공은 몇 가지 교훈을 제공합니다. 첫째, 수익성이 좋은 정부 계약은 세계 시장에서 경쟁하는 것보다 강력한 산업을 창출할 가능성이 적습니다. 미국의 전자 회사들은 제2차 세계대전 이후 트랜지스터, CMOS 반도체, LCD 디스플레이 등 수많은 기술 혁신을 이루었습니다. 이들 기업은 종종 방산 계약에 시달렸기 때문에 이러한 기술을 시장성 있는 제품으로 전환해야 한다는 압박을 거의 받지 않았습니다. 반면 아시아 기업들은 소비자 시장에서 수출과 경쟁에 집중했습니다. 이러한 규율은 그들로 하여금 기술을 신중하게 선택하여 수익성 있는 제품을 만들도록 조정하도록 했습니다.

두 번째 교훈은 기업가들이 필수적이라는 것입니다. Tadashi Sasaki는 계산기를 소형화하기 위해 집적 회로를 사용하는 비전을 가지고 있었습니다. 이로 인해 수억 대의 계산기가 판매되었습니다. 이병철은 DRAM 칩을 생산하기 위해 위험을 감수했습니다. 이 분야에서 그의 성공은 삼성을 2021년 시가총액이 4,000억 달러에 달하는 회사로 만드는 데 도움을 주었습니다. 1960년대와 1990년대 사이

에 동아시아 기적에 기여한 좋은 정책과 교육을 받은 노동자들과 같은 요소들을 되짚으며 요시토미(2003)는 기업가들이 가장 중요하다고 결론지었습니다. 위험을 감수하고 새로운 아이디어를 생산적인 활동으로 전환함으로써 그들은 기술적 진보를 지속하고 경제 성장을 이끌었습니다.

세 번째 교훈은 시민들이 국가의 생존이 위험에 처하고 따라서 국가의 이익을 위해 단결할 때 산업 정책이 성공할 가능성이 높다는 것입니다. 이웃의 침략 위협에 직면한 한국과 대만의 경우가 그러했을 수 있습니다. 노동자, 기업가, 정부 관계자, 외부 자문기들이 협력하여 이를 달성하는 것이 경제 발전이 필수적이라고 여겼습니다. 말레이시아의 경우처럼 재분배가 정부의 주요 목표일 때 산업 정책은 지대 추구의 낭비를 초래할 수 있고 경쟁력 있는 산업을 생산하지 않을 수 있습니다. 시사점은 국가들이 생존이 걸린 상황을 만들어야 한다는 것이 아니라 실존적 위협이 없는 국가들은 산업 정책이 지대 추구의 손실을 발생시킬 수 있다는 것을 인식해야 한다는 것입니다.

네 번째 교훈은 교육과 기술 이전의 중요성입니다. 사사키 타다시는 교육을 잘 받았고, 자신의 회사가 투자해야 할 기술들을 재빨리 감지했습니다. 한국과 대만의 엔지니어들은 비교적 교육을 잘 받았고, 미국의 엔지니어들로부터 노하우를 흡수했습니다. 이것은 그들이 필요한 제조 경험을 얻을 때까지 지식의 차이를 메웠습니다.

다섯 번째 교훈은 기업들에게 목표하고 교정된 인센티브를 제공할 필요성입니다. 말레이시아는 Bumiputera 기업들의 성과가 좋지 않을 때도 종종 실패하도록 내버려 두지 않았습니다. 반면, Hausmann and Rodrik(2003)은 한국 정부가 수출에 성공하지 못한 기업들에게 혜택을 제공하는 것을 중단했음을 보여주었습니다. 적절한 인센티브는 산업정책이 목표를 달성하는 데 도움이 될 수 있습니다.

여섯 번째 교훈은 보호무역주의가 역효과를 낼 수 있다는 것입니다. 오토네틱스와 같은 회사들이 일본에 반도체를 계속 판매하는 것을 일본이 거부하는 것과 일본의 반도체 회사들에 대한 미국의 조치는 결국 양국의 반도체 산업을 약화시켰습니다.

일곱 번째 교훈은 학계, 산업계, 정부 연구기관 간의 협력이 가치 있을 수 있다는 것입니다. 대만의 TAC에는 일류 교수들이 포함되어 있을 뿐만 아니라 대만의 과학단지는 대학, 기업, 정부 간의 비옥한 상호작용을 가능하게 했습니다.

서구의 많은 나라들은 활기찬 반도체 제조업을 육성하고자 합니다. 전쟁이나 자연재해가 발생하면 동아시아의 반도체 공급과 단절될 수 있다는 점을 인식하고 있습니다. 반도체 제조 능력을 키우려면 세심한 고민과 계획이 필요합니다. 단순히 한국과 대만 정부가 반도체에 투자한 것을 보고 서구 정부가 따라야 한다고 선언하는 것만으로는 부족할 것입니다.

코로나19 경험의 또 다른 교훈은 반도체를 사용하는 기업들이 제때 재고를 관리하는 수준을 넘어서야 한다는 것입니다. 팬데믹이 발생하면서 자동차 수요는 빠르게 감소했고 비슷한 속도로 회복되었습니다. 자동차 제조업체들은 반도체 수요가 어떻게 바뀔지 예측하기 어려웠습니다. 예방적 이유로 과도한 재고를 유지하는 것이 현명할 것입니다. 또한 공급망과 공급망 고유의 취약성을 더 잘 이해하면 기업들이 많은 비용을 절약할 수 있습니다. 팬데믹 기간 동안 4군 공급업체가 어려움을 겪을 수 있다는 것이 분명해졌고, 이는 때로는 공급망 전체를 되돌려서 여러 번의 지연을 초래하기도 했습니다. 따라서 기업들이 전쟁, 자연 재해, 팬데믹 및 기타 혼란 상황에서도 복원력을 유지하기 위한 세심한 계획과 적절한 준비가 필요하다는 것은 분명합니다.

제 2 절 한계 및 개선 방향

이번 1차 연도 분석에서 고려하지 못한 몇 가지 한계점들이 있었다. 여기서는 그 한계점을 정리하고 추후 연구에서 고도화할 수 있는 방안을 찾고자 한다.

먼저 비정형 데이터를 처리하는 방법에서의 한계점이다. 첫째로 기업명 식별 단계에서 단순화를 위해 종목명 기준으로 기업을 추출하면서 그 외의 다양한 기업명으로 뉴스에 작성된 경우는 고려되지 못했다. 비슷하게 'A기업은 반도체 장비를

생산하는 기업이다. 이 기업은 B기업과 공급계약을 체결했다.’와 같이 지시대명사로 앞뒤 문장이 이어지는 경우 또한 고려되지 못했다.

둘째로 비정형 데이터에서 공급 관계가 있다고 포착했지만 실제로 관계가 없는 경우도 있었다. 예를 들어 ‘주 거래처는 A기업, B기업 등이며, 검사 장비를 A기업에 공급하고 있다.’라는 문장처럼 주어가 생략되고 기업 두 개가 포함된 경우, A기업과 B기업이 공급관계에 있다고 판단하는 연결 오류 문제가 발생했다.

셋째로 아직 실제 거래 관계가 아닌 추측 단계임에도 불구하고 거래 관계로 판단하는 경우도 있었다. 예를 들어 ‘A기업과 B기업이 공급 계약을 체결한다는 소문에 최근 조희공시 답변은 검토 중이나 현재까지 결정된 사항은 없다고 답했다.’라는 문장을 들 수 있다. 그러나 이 경우 충분히 계약 관계로 이어질 가능성이 있는 의미 있는 표현이기 때문에 규칙 단어의 중요도를 고려하는 접근이 필요할 것으로 보인다.

한편, 공급망 근사 네트워크를 도식화함에 있어 이번 연구는 기업 간 거래 관계를 나타내는 베이스 네트워크 구축에 한정되었다. 반도체 공급망 내에서 나타나는 다양한 위험을 감지하기 위해서는 각각의 기업이 네트워크상에서 어떤 위치에 있는지 파악할 수 있어야 한다.

이는 [표 3]의 반도체 생태계 내 기업 분류나 [그림 29]의 8대 제조공정보다 더 세분화된 분류를 필요로 한다. 예를 들어 세정은 반도체 생산의 모든 공정 단계에서 불순물을 제거하는 필수적인 과정으로 세정 장비의 공급 중단은 반도체 생산에 큰 영향을 줄 수 있다. 그러나 8대 공정에서는 세정 단계는 포함되지 않는다. 또한 포토 공정 장비는 사용하는 광원에 따라 KrF, ArF, EUV 장비로 나눌 수 있다. 이중 EUV 장비는 네덜란드의 장비 기업인 ASML이 독점으로 생산하고 있으며, 첨단 반도체 생산을 위해서는 대체 불가능한 필수장비이다. 단순히 포토 공정으로 분류하면 EUV 장비 수급의 어려움으로 인한 첨단 반도체 생산 중단 위험을 감지할 수 없다.

따라서 본 연구에서는 더 효과적인 반도체 공급의 위험관리를 위해 공급망 지도

의 하위 구조인 공정 기반 지도를 새로운 분류기준으로 재구축하였다([부록] 참고). 이러한 분류체계(taxonomy)를 베이스 네트워크 기억 각각에 덧씌움으로써 반도체 산업의 이슈 및 현상을 설명할 수 있는 공급망 근사 네트워크를 구축할 수 있다. 이는 후속 연구 과제로 남겨두고자 한다.

본 연구 결과의 시사점 및 한계 등을 고려하여 <표 5-1>에 향후 과제를 연구 주제별로 요약하였다.

<표 5-1> 향후 과제

주제	연번	향후 과제
위험지수 개발	1	불확실성 요인별 어휘사전의 구성 단어 정교화
	2	최종시장위험지수 산출을 위한 요인별 가중치 정교화
	3	지도학습 혹은 고도화된 데이터 모형 방법론 검토
공급망 근사 네트워크 구축	1	뉴스 기사 데이터 전처리 과정에서 문장 분할 개선
	2	거래 관계 추출 프로세스 고도화
	3	분류체계를 반영한 공급망 근사 네트워크 구축

본 연구를 통해 도출한 반도체 시장 위험지수가 지수로서 적합한 역할을 하기 위해서는, 지수 값으로의 객관적인 비교분석이 가능해야 한다. 다만, 현재 연구 단계에서는 값을 객관적으로 비교 분석하기에 아직 부족한 부분이 있다. 즉, 값이 높은 지점이 여러 곳 있을 때, 가장 높은 지점이 실제로 다른 지점들에 비해 가장 불안정한 상태인지를 객관적으로 해석할 수 있는지에 대한 여부이다. 이 부분을 해결하기 위해서 본 연구에서는 기사 표본 추출시 표본의 개수를 통제하여 모델링하였다. [그림 20]에서 확인할 수 있듯이 반도체 시장 관련 기사의 수는 최신으로 올수록 많아지고 있다. 표본추출을 하지 않고 전체 뉴스를 모두 사용하게 되면, 실제 동일한 위험정도의 사건이더라도 최신으로 올수록 기사수가 많아지므로 기사 수 자체가 지수 산출시 영향을 주어 최신 이벤트로 인한 값을 높이는 결과를 초래할 수 있기 때문이다.

그러나, 본 연구의 결과, 표본의 개수를 통제하더라도 각 요인별 어휘사전 내 단어의 구성이 시장위험지수 값에 영향을 줄 수 있다고 판단되었다. 예를 들어, 같은 외부충격 불확실성 요인 어휘사전 내에서도 코로나 관련 단어의 개수가 지진과 관련한 단어의 개수보다 많다면 코로나 관련 기사들이 지진 관련 기사들보다 어휘사전 내의 단어를 더 많이 포함하고 있다고 판단되어 값이 높아질 수 있기 때문이다. 이에, 해당 부분에 대한 추가 연구를 통해 어휘사전의 구성 단어를 정교화하고자 한다. 이를 통해 반도체 시장의 위험도를 더욱 정확하고 객관적으로 표현할 수 있는 반도체 시장 위험지수를 개발할 수 있을 것으로 기대한다.

추가로, 현재는 최종 시장위험지수를 산출하기 위해 다섯 가지 정책 불확실성 요인을 모두 동일한 가중치로 합산하였다. 그러나, 각 요인이 반도체 시장에 미치는 실질적인 영향은 상이할 수 있다. 이에 따라, 요인별 영향력을 반영한 차별화된 가중치를 적용하는 것은 본 시장위험지수의 정밀도를 향상시킬 수 있을 것으로 생각한다. 이에, 향후에는 이론적 근거에 기반한 방법론을 개발하여, 불확실성 요인별로 보다 세밀하게 가중치를 조정하는 방안을 연구하고자 한다.

마지막으로 본 연구에서는 시간과 비용의 제약 조건으로 인하여 EM 알고리즘 기반의 비지도학습 모형을 활용하였지만 반도체 시장과 관련된 전문가 지식 및 의견을 반영할 수 있는 지도학습 모형을 대안으로 고려할 수 있을 것이다. 본 연구에서 제안한 비지도학습 모형은 온라인뉴스의 어휘 패턴에 기반하기에 실제 반도체 시장위험이나 불확실성과 매칭되지 않은 부분이 발생할 수 있다. 이러한 겅을 줄이고 보다 객관적이고 정교한 지수산출 모형을 구축하기 위해서는 지도학습 기반의 방법론을 검토할 필요가 있다.

반도체 공급망 근사 네트워크 구축에 있어 관계 추출의 성능을 높이기 위해 몇 가지 개선이 필요한 부분이 있다. 첫째로 전처리 과정에서 뉴스 기사 데이터의 문장 분할(sentence segmentation)이 부정확하여 문장이 아닌 긴 문단으로 남는 경우가 있었다. 이는 분석 결과에 큰 영향을 주므로 보다 정확하게 문장 단위로 구분할 수 있도록 개선이 필요하다.

둘째로 비정형 데이터로부터 거래 관계 데이터를 추출하는 프로세스를 고도화할 필요가 있다. 이번 연구에서는 사람이 직접 거래 관계를 표현하는 형식을 찾아 규칙을 정의하고 관계 튜플을 추출했다. 그러나 이는 앞서 논의된 바와 같이 많은 패턴을 놓칠 수 있다는 한계가 있다. 따라서 추후 연구에서는 딥러닝을 활용하여 직접 패턴을 발견하고 학습할 수 있도록 고도화할 수 있을 것이다.

셋째로 이번 연구에서 진행한 베이스 네트워크는 단순한 기업명의 연결이므로 반도체 산업의 이슈 및 현상에 대한 통찰을 제공하는 데에는 한계가 있다. 따라서 향후 과제에서는 기업 각각에 반도체 산업 분류체계를 덧씌워 이슈의 증가 및 감소에 따라 공급망 근사를 비교하는 작업이 필요하다. 예를 들어, [부록]에 소개되어 있는 제조공정 분류 결과를 베이스로 하여 공급망을 근사하는 한다면 반도체 공급망의 활용을 증대시킬 수 있을 것으로 예상된다.

참 고 문 헌

[국내 문헌]

특허청 한국특허전략개발원, 「반도체 산업 특허 메가트렌드 분석 보고서 2021」, , 산책길.

[해외문헌]

Aubry, M., & Renou-Maissant, P.(2014). Semiconductor industry cycles: Explanatory factors and forecasting. *Economic Modelling*, 39, 221-231.

Albey, E., Norouzi, A., Kempf, K. G., & Uzsoy, R.(2015). Demand modeling with forecast evolution: an application to production planning. *IEEE Transactions on Semiconductor Manufacturing*, 28(3), 374-384.

Chen, Y. J., & Chien, C. F.(2018). An empirical study of demand forecasting of non-volatile memory for smart production of semiconductor manufacturing. *International Journal of Production Research*, 56(13), 4629-4643.

Chien, C. F., Lin, Y. S., & Lin, S. K.(2020). Deep reinforcement learning for selecting demand forecast models to empower Industry 3.5 and an empirical study for a semiconductor component distributor. *International Journal of Production Research*, 58(9), 2784-2804.

Chien, C. F., Hsu, C. Y., & Hsiao, C. W.(2012). Manufacturing intelligence to forecast and reduce semiconductor cycle time. *Journal of*

- Intelligent Manufacturing, 23, 2281-2294.
- Chien, C. F., Chen, Y. J., & Peng, J. T.(2010). Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics*, 128(2), 496-509.
- Chow, H. K., & Choy, K. M.(2006). Forecasting the global electronics cycle with leading indicators: A bayesian var approach. *International Journal of Forecasting*, 22(2), 301-315.
- Dequeant, K., Vialletelle, P., Lemaire, P., & Espinouse, M. L.(2016, December). A literature review on variability in semiconductor manufacturing: The next forward leap to Industry 4.0. In 2016 Winter Simulation Conference(WSC)(pp. 2598-2609). IEEE.
- Fu, W., Chien, C. F., & Lin, Z. H.(2018). A hybrid forecasting framework with neural network and time-series method for intermittent demand in semiconductor supply chain. In *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0: IFIP WG 5.7 International Conference, APMS 2018, Seoul, Korea, August 26-30, 2018, Proceedings, Part II*(pp. 65-72). Springer International Publishing.
- Jou, Y. T., Wee, H. M., Chen, H. C., Hsieh, Y. H., & Wang, L.(2009). A neural network forecasting model for consumable parts in semiconductor manufacturing. *Journal of Manufacturing Technology Management*, 20(3), 404-412.
- Liu, W. H.(2007). Forecasting the semiconductor industry cycles by bootstrap prediction intervals. *Applied Economics*, 39(13), 1731-1742.
- Tan, H., & Mathews, J. A.(2010). Cyclical industrial dynamics: The case of the global semiconductor industry. *Technological Forecasting*

- and Social Change, 77(2), 344-353.
- Terwiesch, C., Ren, Z. J., Ho, T. H., & Cohen, M. A.(2005). An empirical analysis of forecast sharing in the semiconductor equipment supply chain. *Management science*, 51(2), 208-220.
- Wang, C. H., & Chen, J. Y.(2019). Demand forecasting and financial estimation considering the interactive dynamics of semiconductor supply-chain companies. *Computers & Industrial Engineering*, 138, 106104.
- Wang, J., & Zhang, J.(2016). Big data analytics for forecasting cycle time in semiconductor wafer fabrication system. *International Journal of Production Research*, 54(23), 7231-7244.
- Xu, Q., & Sharma, V.(2017). Ensemble sales forecasting study in semiconductor industry. In *Advances in Data Mining. Applications and Theoretical Aspects: 17th Industrial Conference, ICDM 2017*, New York, NY, USA, July 12-13, 2017, Proceedings 17(pp. 31-44). Springer International Publishing.

[부 록]

〈부록 표 1〉 반도체 분야의 기술분류체계

분류	세부내용	분류	세부내용
FAB 공정	다층금속 / 차세대 배선 기술	소재	CMP용 소재
	다층배선 절연막 기술		Gas(Dry 공정)
	평탄화 / Channel 공정 기술		Hard Mask / Mask
	Gate Stack 공정 기술		Precursor
	Isolation		SOD(spin On Dielectric) 소재
	Mask(less) Lithography		Target 소재
	Shallow Source/Drain 형성기술		도금 용액
	Analog 및 Mixed signal 공정		세정&식각액(wet 공정)
	FD-SOI 공정		Anti-reflection Coating(ARC)
	FinFET 공정		resist
	전력반도체 공정		Si
	3D stack 기술		SOI(silicon on insulator)
	Flip chip package		화합물 반도체
	DRAM / PRAM / SRAM		Bonding / Bump 소재
	NAND Flash		Dicing 소재
	ReRAM / STT-MRAM		Heat Sink 소재
	공정 모니터링 기술		TSV용 소재
	박막 및 소자 분석기술		기판
설계	SoC 설계 및 검증	차세대 반도체	비메모리(시스템) 반도체
	아날로그		모션센서 기술
	프로세서		이미지센서 기술
	휘발성 / 비휘발성 메모리		물리/화학 특성 감지
장비	세정 장비		지문감지
	평탄화 장비		운전 지원/모니터링
	증착 / 식각 / 노광 장비		차량용 통신 / 차량용 전력 제어
	가스제어 장치		전력 제어
	기타 장치		에너지 변환
	열제어 장치		개인/정보보안
	Dicing(절단) 장비		네트워크 지원
	공정 모니터링 장비		디스플레이 SoC
	패키징 장비		실감영상 SoC / 웨어러블 SoC
	후공정 검사 장비		로봇/드론/바이오헬스케어 반도체

[부록 그림 1] 미국 기업의 기업별 공동특허 현황

1. AMT(어플라이드 머티리얼즈)



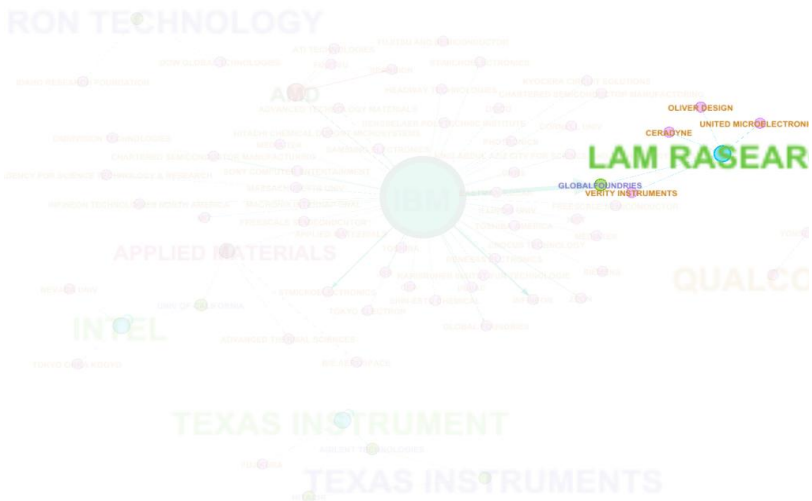
2. 인텔



3. AMD



4. 램리서치



5. 마이크론



6. 퀄컴



7. 텍사스 인스트루먼트





● 저 자 소 개 ●

고 동 환

- 중앙대 경제학과 졸업
- 일리노이대 경제학과 석사
- 럿거스대 경제학과 박사
- 현 정보통신정책연구원 연구위원

김 민 식

- 고려대 경영학과 졸업
- KAIST 경영학과 석사
- 고려대 과학기술학 박사
- 현 정보통신정책연구원 부연구위원

김 봉 진

- 명지대학교 전자공학과 졸업
- 현 한국특허정보기술진흥원 수석

김 현 태

- 워털루 대학교 보험통계학 박사
- 현 연세대 통계데이터사이언스학과 교수

임 중 호

- 연세대 경제학과 졸업
- 연세대 통계학 석사
- 아이오와 주립대 통계학 박사
- 현 연세대 응용통계학 부교수

정책자료 23-07

정형/비정형데이터 기반 반도체 시장 이슈 분석 및 예측
모형 개발

2023년 12월 일 인쇄

2023년 12월 일 발행

발행인 배 경 울

발행처 정보통신정책연구원

충청북도 진천군 덕산읍 정통로 18

TEL: 043-531-4114 FAX: 043-535-4695~6

인쇄 인성문화

ISBN 979-11-7000-383-0 93320

〈비매품〉