

# 융합기술전문가의 공동연구에 대한 사회적 연결망 분석

## Social Network Analysis on Interdisciplinary Collaboration of Convergence Technologies Specialists

이중만\*, 최민석\*\*

호서대학교\*, 한국과학기술연구원(KAIST)\*\*

Jungmann Lee(mann@hoseo.edu)\*, Minseok Choi(cooldenny@business.kaist.co.kr)\*\*

### 요약

본 연구는 국내 융합기술 분야 연구자들 간의 연구 협력 행태를 알아보기 위하여 사회적 연결망 분석(social network analysis)을 수행하였다. 융합기술분야 연구자 1,095명으로 구성된 국내 융합기술 연구자 네트워크의 특성은 다음과 같다. 첫째, 융합기술 연구자들은 다른 과학기술분야에 비해 인당 논문 편수가 많아, 융합기술분야에서 새로운 지식 창출이 활발하게 이루어지고 있음을 보여주고 있다. 둘째, 높은 생산성에 비해 융합기술연구자 그룹의 상당수가 논문의 공동저작을 활발하게 진행하고 있지 않고 제한된 협력 관계만을 보여주고 있다. 마지막으로, 네트워크 구성원들 간의 거리가 다른 과학기술분야의 그것에 비해 가깝지만, 삼각형 모양 관계의 공동 연구하는 형태로 발전하지 않고 양자 간의 공동 연구로만 머물러 있는 경우가 많다. 오직 소수의 연구자들의 허브 역할로 연구자들 간의 연결고리가 유지되고 있다. 즉, 국내 융합기술 연구 분야의 협력 관계는 분권형 이라기보다는 집중형의 특성을 보여주고 있다. 국내 융합기술 연구자들 간의 협력이 보다 활발하게 진행되기 위한 기반조성 정책 프로그램이 필요하다.

■ 중심어 : | 기술융합 | 융합기술 | 공동연구 행태 | 사회적 연결망 분석 |

### Abstract

Converging technologies have become a major issue in science policy. This paper describes the current state of scientific collaboration for convergent technologies among researchers in South Korea, by conducting Social Network Analysis (SNA) with the data set of 1,095 researchers who have involved in the development of the convergent technologies. It is found that the researchers in convergent technology are more productive than the researchers in other technology domains. However, the researchers in convergent technologies have small number of collaborators, compared with their productivity. Only a few researchers have a role of the hub in the collaboration networks, meaning that the structure network is closer to than the core than the peripheral. The scientific collaboration network of the convergent technology researchers shows that the members of the network are close to each other, but there is small number of cliques.

■ keyword : | Technological Convergence | Fusion Technology | Collaboration Research | Social Network Analysis |

\* 이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(기초연구지원 기초과학(공동연구), D00001)

접수번호 : #100309-002

접수일자 : 2010년 03월 09일

심사완료일 : 2010년 05월 19일

교신처자 : 최민석, e-mail : cooldenny@business.kaist.ac.kr

## I. 서론

21세기의 기술개발에서는 분야 간의 융합이 가장 큰 화두로 떠오르고 있다. 최근 나노(Nano)-바이오(Bio)-정보기술(IT) 간의 융합(이하 NBI 융합기술)에 과학계의 관심이 집중되고 있으며, 어느덧 가시적인 초기 성과를 보여 주고 있다. 정보기술은 20세기 말의 두드러진 사회 변화를 이끌었던 것처럼 기술개발과정의 전영역에 걸쳐 많은 영향을 끼치고 있는데, 예를 들면 인간 유전자 지도를 작성하고자 시작되었던 게놈 프로젝트(Genome Project)는 바이오기술과 정보기술이 융합된 대표적인 사례라고 할 수 있다. 최근에는 정보기술의 한 영역으로 다루어졌던 정보보안 분야에서 인간의 생물학적 특성을 이용하게 되면서 홍채인식보안시스템 등이 만들어지게 되었다. 또한 나노 단위에서의 물질 특성을 이해하는 나노기술 분야가 발전하면서 과학 분야 전반의 이해 수준에 혁명적 변화가 예상되고 있다. 반도체 집적회로는 나노 단위에서 제조됨으로써 물리적 한계를 극복하는 수준까지 발전하게 되었으며, 의료 현장은 극소형 수술용 로봇이 사용됨으로써 빠르게 변화하고 있다. 앞으로 육안으로는 식별할 수 없는 크기의 재료나 기기를 제작할 수 있게 되면 인간의 인식 변화까지 초래할 것으로 보인다. 이러한 NBI 융합기술은 바이오 인포메틱스(bio-informatics) 등과 같은 새로운 기술 영역을 등장시키면서 빠르게 발전하고 있다. 앞으로도 기술 융합을 통해 미처 상상할 수 없을 정도로 많은 성과를 보여 줄 것으로 기대된다. 기술 융합은 21세기의 새로운 패러다임으로 자리 잡고 있다.

세계 각국은 융합기술 시대의 기술 경쟁에서 살아남기 위해 새 패러다임을 적극적으로 받아들이고 있는데, 기술 융합의 시대의 기술 경쟁력은 시대에 적합한 인력을 어떻게 확보하고 어떻게 조직화하는 것인가에 의해 좌우될 것으로 보인다. 대부분의 지식은 개인에 내재되어 있다. 비록 인쇄물이나 컴퓨터 시스템에 저장된 코딩화된 지식과 조직 문화와 같은 개인 간의 관계 형태로 존재하는 지식도 있지만, 대부분의 과학 기술은 개인 두뇌에 저장되어 있다. 따라서 기술 융합 시대에도 연구 인력의 지적 능력을 향상시키는 것이 무엇보다 중

요하다. 그리고 융합기술개발은 본질적 특성상 이종 분야의 지식의 통합을 필요로 한다. 만약 한 연구자가 연구에 필요한 여러 분야의 지식에 모두 정통하다면 융합 기술 연구를 홀로 수행할 수 있겠지만, 일반적으로 이런 경우는 매우 드물다. 융합기술개발에서는 여러 분야에서 활동하는 연구자들이 보유하고 있는 지식을 교환하고 통합하기 위해서 연구자들 간의 협력이 필연적으로 요구된다.

본 연구에서는 기술 융합 시대에 요구되는 연구자들의 역량 중에서 연구자들의 협력이 초점을 두고자 한다. 구체적으로 국내 NBI 융합기술 연구자들이 서로 어떻게 협력하고 있는지를 살펴봄으로써 국가적 차원에서 기술 융합 시대를 준비하는데 도움이 될 정책적 시사점을 찾고자 한다. 과학 분야 연구자들 간의 협력 행태를 연구한 대표적인 분야가 연구자들 간의 공동 집필 분석(co-authorship analysis)인데, 특히 최근에는 사회적 연결망 분석(social network analysis, 이하 SNA)을 이용한 공동 저자 분석이 활발하게 이루어지고 있다. 본 연구에서는 SNA 방법을 이용한 공동 집필 분석을 통해 국내 융합기술 분야 연구자들 간의 연구 협력 관계를 이해하고자 한다.

이 논문은 다음과 같이 구성되어 있다. II장에서는 연구자들의 과학적 협력을 이해할 수 있는 이론적 배경을 제시할 것이다. 이어서 SNA에 대한 일반적 설명과 함께, 지금까지 과학적 연구협력을 SNA로 분석한 기존의 연구결과도 함께 제시할 것이다. III장은 본 연구에서 어떻게 데이터를 수집하여 분석을 위해 정제하였는지를 설명할 것이며, IV장에서는 국내 융합기술분야에서 공동연구자들이 어떻게 협력하고 있는지에 대한 분석 결과를 보여줄 것이다. 그리고 마지막 장인 V 장에서는 결론 및 연구의 한계점을 설명하고자 한다.

## II. 이론적 배경

### 1. 연구 협력의 증가

연구자들 간의 협력은 여러 연구자들이 과학적 지식을 만들려는 공동의 목적을 달성하기 위해 함께 일하는

것이다. 그런데 연구자들이 협력하는 정도에 따라 다양한 경우가 가능하다. 예를 들면, 모든 연구자들이 전체 연구 과정에 모두 참여하고 연구 결과에 직접적으로 기여할 수 있다. 반면에 고에너지 물리학에서는 수백 명의 과학자들이 서로 일면식이 없더라도 공동의 과학적 목적을 위해 실험에 참여하는 것과 같이 연구과정의 부분 기여 등을 연구 협력<sup>1)</sup>으로 간주할 수도 있다[1].

본 논문에서는 위의 연구 협력에 대한 범주를 참고하여 하나의 논문을 공동으로 저작한 경우에 두 연구자가 연구 협력의 관계에 있다고 가정하고자 한다. 즉, 만약 한 논문에 두 명 이상의 저자가 등장한다면, 등장하는 저자들 중 임의의 두 저자는 공동 집필의 관계에 있다고 할 수 있다. 그런데 [2]의 연구에 따르면, 과학자들 간의 협력 중 단지 5%만이 논문의 공동 저작으로 이어지지 않는 경우라고 보여 주고 있다[2]. 따라서 논문의 공동 저작을 과학자들 간의 연구 협력의 측정 지표로 사용하는 것은 타당하다. 실제로 과학 분야 연구자들 간의 협력을 설명하기 위해서 공동 집필 분석(co-authorship analysis)을 많이 이용하고 있다.

과학자들은 동료 연구자와 협력하여 연구를 수행하기도 한다. 최근에는 이런 과학 분야 협력이 많이 증가하고 있다. [3]에 따르면, 시간이 지나면서 전체 발표 논문 중에서 두 명 이상이 공동 집필한 논문의 비중이 증가하고 있다는 사실을 확인할 수 있다. 수학 분야의 대표 저널 중 하나인 Mathematical Review(MR)를 살펴보면, 1인 저자 논문의 비율이 지속적으로 감소한데 반해, 2인 이상의 저자가 작성한 논문의 비율이 지속적으로 증가하여 왔다는 것을 알 수 있다. 1940년에는 1인 저자의 논문이 90%를 상회했던 것이 1993년에는 60% 이하로 감소하였다. 반면에 동 기간에 두 명의 연구자가 함께 작성한 논문의 비율이 10% 미만에서 30%를 넘게 되었다[3].

이러한 과학 분야 연구자들 간의 협력이 증가하는 이유를 연구 수행과정에서 발생하는 비용의 급격한 증가, 특히 첨단 분야에서의 장비를 완비하는 데 소요되는 비용의 증가를 하나의 원인으로 보고 있다[1]. 즉, 연구자들 각자가 연구에 필요한 장비를 모두 확보할 수 없기 때문에 각자 보유하고 있는 장비를 공동으로 사용하는 것이 연구 생산성 측면에서 유리하기 때문에 연구자들 간의 협력이 촉진된다는 설명이다. 장비 구입비용과 함께, 연구자들이 습득하고 있는 지식과 기술이 점차 세분화되어 가는 점 또한 연구자들 간의 지식교환을 통한 협력을 촉진시키고 있다. 그리고 이러한 연구자들 간의 지식 교환에 소요되는 비용도 감소하고 있다. 여행비용과 통신비용의 감소 등과 같은 커뮤니케이션(communication)에 소요되는 비용의 감소가 연구자들 간의 왕래를 증가시키고 의견교환 확대시켜서 연구자들 간의 공동 연구가 증가하고 있다. 무엇보다도 고급 지식의 창조가 연구자들 간의 사회적 상호작용에 의해서 발생할 가능성이 높기 때문에, 연구자들은 연구 성과를 극대화하기 위해서 다른 연구자들과의 협력을 희망한다[1]. 이런 맥락에서 최근에는 NBI 융합기술과 같은 다학제적인 분야의 중요성이 대두되면서 연구자들 간의 협력이 더욱 중요한 연구 주제로 부상하고 있다.

또한 연구자들 간의 협력이 과학적 성과 향상에 기여하고 있는지에 대해서, 과학자들의 연구 생산성에 관한 기념비적인 [4]의 연구에 따르면 특정 개수의 논문을 발표한 과학자들의 수는 그 논문 수의 제곱에 반비례하는 관계에 있다는 것을 확인할 수 있다. 그런데, 그는 이러한 연구자의 연구 생산성이 차이는 연구자들 간의 연구 협력에 의해 좌우된다고 주장하고 있다[4]. 실제로 다른 연구자들과의 협력이 논문 발표 수와 직접적으로 연관 있다는 것을 실증적으로 보여 주고 있다[5]. 그런데 협력하는 연구자가 누구인가에 관계없이 모든 종류의 연구 협력이 생산성을 향상시키지는 않는다. 생산성이 높은 연구자와의 협력은 생산성의 향상을 이끌지만, 반대로 생산성이 낮은 연구자와의 협력은 생산성을 저하시킨다. 그리고 저자 수가 많은 논문일수록 논문 심사를 통과하는 비율이 높다는 것을 보였는데, 이것은 논문심사자가 공동연구자가 많은 논문을 품질이 높은

1) 연구 과정의 전 기간 또는 그에 상응할 정도의 많은 기간 동안 직접 참여하거나 결과도출에 결정적으로 기여한 경우, 최초 연구 제안서를 함께 작성한 경우, 연구의 실증적 검증 부분(예: 실험설계, 실험장비 개발, 실험 수행, 실험 데이터의 분석 및 해석, 연구 보고서의 공동 작성)을 담당한 경우, 연구의 이론적 부분(예: 원 아이디어 창안, 가설 도출, 이론적 해석)을 담당한 경우, 원 프로젝트 제안자나 자금 지원자 등도 연구협력으로 볼 수 있다.

논문으로 인식할 가능성이 높다는 것으로도 해석 가능하다[6]. [6]의 연구결과는 [7]의 연구에서 보는 바와 같이, 저자의 수가 많은 논문에 보다 높은 신뢰를 부여하는 것에서 그 원인을 찾을 수 있다. 그리고 이러한 신뢰는 다 저자 논문이 보다 많이 인용되는 현상과도 연관이 있다. 특히, 다양한 국가나 기관에서 참여한 공동 저작의 논문이 연구 분야에서 큰 영향력을 가지고 있는 것으로 밝혀졌다[7]. 따라서 NBI 융합기술 연구자들이 협력하는 패턴을 분석하는 것은 궁극적으로 연구 생산성 향상 측면에서도 중요하다.

## 2. 사회적 네트워크로서의 공저자망 연구

최근에는 네트워크 이론이 사회적 현상을 설명하는 데 적용되면서, 과학자들의 연구 협력도 네트워크 형태로 이해되고 분석되고 있다. 네트워크(network)란 노드(node)라고도 일컬어지는 꼭지점(vertex)과 이들 꼭지점을 연결하는 연결선(edge)의 집합을 의미한다. 각 꼭지점을 2개 이상의 그룹(예: 남자와 여자)으로 분류해서 다양화하는 경우, 이를 다모드 네트워크라고 한다. 예를 들어, 남자와 여자, 두 가지로 구성된 네트워크는 2-모드 네트워크(2-mode network)이다. 한편, 모드가 하나인 네트워크를 단모드 네트워크(1-mode network)라고 일컫는다. 마찬가지로, 연결선들도 다양한 형태로 세분해서 표현할 수 있다. 이 외에도 연결선의 연결강도(예: 연결 강함, 연결 약함)를 다르게 표현하거나 방향성(예: 꼭지점 A에서 꼭지점 B로 연결)을 부여할 수도 있다. 연결선의 강도라 다를 때를 값 부여 네트워크(valued network)라 부르고 방향성이 있을 때를 방향성 네트워크(directed network)라 한다. 한편, 연결선의 강도나 방향성이 없는 네트워크를 단순 네트워크(simple network)라고 한다[8].

이러한 네트워크 중 하나가 사회적 연결망(social network 이하 SN)이다.<sup>2)</sup> SN에서는 사람이 꼭지점이

되고 사람들 간의 상호작용이 연결선을 형성한다. 지금까지 사회적 연결망 시각을 바탕으로 한 사회 현상 분석은 다양한 분야에 적용되어 왔는데, 이것을 사회적 연결망 분석(social network analysis)다. 지금까지의 연구를 구분해 보면 다음과 같다. 첫째, SNA의 초기 단계로서 소규모 그룹을 대상으로 한 연구들이다. 이들 연구에는 데이터 확보의 어려움으로 인해 친구 관계나 동업 관계, 결혼 관계 등과 같이 제한된 규모의 그룹을 대상으로 분석한 것들이 있다. 일명 '여섯 단계(six degree)'나 '작은 세계(small world)'라는 문구로 널리 알려진 Milgram의 실험이 있다. 그는 실험을 통해 세상의 사람들은 평균적으로 여섯 단계를 거치면 지구상의 그 누구와도 알게 된다는 것을 보여 주었다[9]. 그러나 비록 Milgram의 방법이 당시로서는 획기적인 방법이 었지만, SNA는 여전히 제한된 규모의 데이터만을 이용했다. 둘째, 위와 같은 데이터 부족의 문제를 극복해서 대규모의 사회적 연결망을 분석한 연구들이 등장하였다. 가장 대표적인 사례로는 영화 배우들 간의 공동 출연 관계를 네트워크로 분석한 것이다. 인터넷의 발달과 함께 등장한 인터넷 영화 데이터베이스(Internet Movie Database)가 갖추어짐으로써 가능한 연구였다. 이 외에도 기업 이사회 임원들의 사회적 관계를 분석한 연구가 있으며, 본 연구의 주제인 과학자들 간의 연구 협력에 관한 연구들도 여기에 속한다.<sup>3)</sup>

[16]은 국내 과학기술 연구자들의 연구 협력 관계를 사회연결망 측면에서 분석하였다. 2002년도에 정부가 지원한 연구과제에 참여한 대학 소속 연구자들을 대상으로 하고 있는데, 총 2,785명으로 구성 연구자 네트워크는 한 연구과제에 참여한 연구원들이 상호 협력하고 있다고 가정하고 있다[16]. 이 연구에서는 국내 과학자들 간에는 서구의 과학자 네트워크에서 드러나는 "작은 세계"의 구조임을 보여 주고 있다[16]. 즉, 과학자들은 평균적으로 작은 수의 과학자간의 연결만으로도 모든

2) 소셜 네트워크 외에도 첫째, 정보 네트워크(또는 지식 네트워크)가 있다. 위의 소셜 네트워크의 꼭지점이 사람에서 정보나 지식 등으로 대체된 네트워크이다. 여기에 속한 대표적인 연구는 연구 논문이나 특허 등의 인용 분석이 있다. 이 외에도 단어의 유사어 관계나 개인이 특정 책이나 영화를 선호하는 것을 네트워크로 표현해서 분석한 사례가 있다. 인터넷상의 웹문서간의 관계를 분석하는 것도 여기에

해당된다. 둘째, 기술적 네트워크가 있는데, 대표적으로 전기망이나 통신망을 분석하거나 도로나 철도교통, 항공로 등을 분석하는데 활용되고 있다. 마지막으로 생물학적 네트워크가 있다. 신진대사과정이나 생태계의 먹이사슬, 인간의 신경계가 여기에 해당된다.

3) 이 외에도 커뮤니케이션 기록을 SNA로 분석한 연구들이 있다. 대표적인 것으로는 이메일이나 휴대폰 전화 통화 기록을 분석하는 것이다. 이러한 연구에서는 연결선이 방향성을 가지는 것이 일반적이다.

과학자들과 연결될 수 있다는 것이다. 특히, 학연과 지연 등의 사회적 관계로 인해 소그룹 위주로 네트워크가 분절화된 네트워크가 형성될 것으로 기대했던 것과 다르게 연구자들이 가깝게 연결되어 있다는 점은 매우 흥미롭다. 또한 연구자들이 연구 협력 네트워크 내에서 차지하는 구조적 위치가 연구성과에 영향을 준다는 것도 보여 주고 있다[16].

또한 연구과제를 정보기술(IT)와 생명공학기술(BT), 나노기술(NT), 환경기술(ET), 청정기술(CT), 우주항공기술(ST)에 관한 것들로 분석하였으며, 연구분야간의 상호협력에서는 IT 분야의 전공자들이 타 분야와의 협력을 많이 하고 있지만, BT 분야의 전문가들이 모든 분야의 협력에서 중심적 위치를 차지하고 있음을 보여 주고 있다[16].

3. 공저자망 연구와 네트워크의 구조적 특성 변수

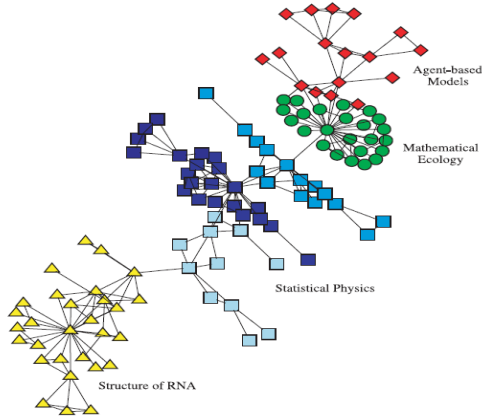


그림 1. SNA를 이용한 공저자망 연구 ([10], p5201, Fig. 1)

[그림 1]에서 보는 바와 같이, SNA를 이용한 공저자망 연구(co-authorship)에서는 꼭지점에 해당하는 것이 과학자이고 연결선은 과학자들 간에 공동으로 논문을 집필했을 경우에 형성된다. 가장 단순한 형태의 연구는 특정 연구자를 중심으로 과학자들이 어떻게 협력하는 것을 보는 것으로, 이런 형태의 네트워크를 자아 중심 네트워크(ego-centric network)라고 한다[12]. 가장 대표적으로 수학 분야의 대가인 폴 에르되스(Paul Erdős)와 함께 논문을 작성한 저자들의 관계를 표현한

에르되스 수를 분석한 연구가 여기에 해당된다. [3]의 연구에서는 1936년부터 연구 결과를 발표한 이 위대한 수학자와 공동으로 논문을 집필한 연구자들에게 숫자 1을 부여하고, 숫자 1을 부여 받은 연구자들과 공동으로 논문을 집필한 연구자들에게 숫자 2를 부여하는 방식으로 에르되스 수를 부여하고 이것을 분석한 것이다. 현재까지 폴 에르되스 숫자 1을 부여 받는 연구자는 총 485명이고 숫자 2에 해당하는 연구자는 5,337명이다. 대표적으로 알버트 아인슈타인(Albert Einstein)은 숫자 2를 부여 받았다. 그의 프린스턴 연구소의 동료인 Ernst G. Straus와 2개의 논문을 발표했는데, 그가 폴 에르되스와 20편의 논문을 공동 집필했기 때문이다. 한편 2차 대전 말기에 미국의 원자폭탄 개발의 주역이었던 J. Robert Oppenheimer는 숫자 4를 부여 받았다. 즉, 그는 직접 폴 에르되스와 논문을 집필하지 않았지만 4 단계를 거쳐 논문을 공동 집필 네트워크를 통해 폴 에르되스와 연결된다[3].

표 1. 분야별 연구자 네트워크 특성 ([10], p5201, Table 1, [13], p405, Table 1)

구분	MED	LA	SPIRES	NCSTRL	MATH
저자 수	1,520,251	52,909	56,627	11,994	253,339
논문 수	2,163,923	98,502	66,652	13,169	-
저자당 논문 수	6.4	5.1	11.6	2.55	6.9
논문당 저자 수	3.75	2.53	8.96	2.22	1.45
평균 공동 저자 수	18.1	9.7	173	3.59	3.9
최대 하위 네트워크	92%	85%	88.7%	57.2%	82%
평균 거리	4.6	5.9	4.0	9.7	7.6
최대 거리	24	20	19	31	27
군집화 정도	0.066	0.43	0.726	0.496	0.15
계층화 정도	0.13	0.36	-	-	0.12

폴 에르되스 수가 한 연구자들 중심으로 나뉘어 가지 형태로 형성된 네트워크를 분석한 것이라면 복잡하게 연결되어 있는 진정한 의미에서의 연구자 네트워크를 분석한 것들이 있다. 작은 규모로는 특정 저널들에서 공저자들의 네트워크가 어떻게 구성되어 있는가를 연구한 논문들이 있다. [17]은 MIS(Management Information System) 분야의 대표적인 학술지인 MIS

Quarterly에서 1996년부터 2004년까지 발표된 242편의 논문의 저자들의 공저자 네트워크를 분석하였는데, 독립된 세 개의 주요 하위네트워크가 형성되어 각각의 하위네트워크는 연구주제에서도 차이를 있다는 것을 보여 주었다[17].

특정 저널에 국한하지 않고 분야 전체를 대상으로 한 연구들도 있다. [10]은 1995년부터 1999까지의 의학 분야와 이론 물리학 분야, 고에너지 물리학 분야, 컴퓨터 과학 분야의 논문을 대상으로 분석하였다. 분석에 이용한 데이터베이스는 MEDLINE(의학 분야, 이하 MED)와 Los Alamos e-Print Archive(이론 물리학 분야, 이하 LA), SPIRES(고에너지 물리학 분야), NCSTRL(컴퓨터 과학 분야)이다. 결과는 각 분야마다 다른 특성을 보이고 있다. 고에너지 물리학 분야에서의 과학자들 간의 연구 협력은 다른 분야에 비해 논문 당 저자 수에서 월등히 많다. 이것은 이 분야의 연구가 주로 대규모의 공동 실험에 의해 이루어지는 특성이 반영되어 있기 때문이다. 그리고 이 분야의 과학자들의 대다수인 88.7%의 과학자들이 하나의 네트워크로 연결되어 있으며 짧은 평균 거리와 높은 군집화 정도에 비추어 볼 때 네트워크 구성원들 간의 연결의 밀도가 높은 것을 알 수 있다. 반면에 컴퓨터 과학 분야는 전체 과학자의 57.2%만이 최대 규모의 네트워크에 연결되어 있고 거의 절반에 가까운 나머지는 주류 네트워크와 연결되어 있지 않다. 그리고 최대 규모의 하위 네트워크에서도 다른 분야에 비해서 상대적으로 평균 거리와 최대거리가 높은 것으로 볼 때, 네트워크 구성원들 간의 연결이 조밀하지 않다는 것을 알 수 있다[13].

또한 2001년도에 분석한 의학 및 이론 물리학 분야 [12]와 수학 분야의 논문[13]을 대상으로 분석한 [표 1]과 같이, 저자 당 논문 수는 5.1에서 6.9으로 대략 비슷하다. 따라서 분석하는 세 분야(의학, 이론 물리학, 수학)의 저자 수의 차이에 따른 효과는 미미할 것으로 본다. 논문 당 저자 수의 결과를 보면, 생물학 분야에서 보다 활발하게 공동으로 논문을 작성하는데 반해 수학 분야에서는 그 수치가 매우 낮다. 생물학 분야가 수학 분야의 두 배가 조금 넘는다. 이러한 논문 당 저자 수의 영향으로 평균 공동 저자 수 역시 생물학 분야에서는

매우 높는데 반해 수학 분야는 매우 낮다. 두 차이가 대략 4.5배가 조금 넘는다. 그리고 한 번 이상의 논문 공동 집필을 통해 연결되어 있는 과학자 네트워크 중 가장 큰 네트워크 역시 생물학 분야는 전체 1,520,251명의 과학자 중 92%를 포함하는데 반해 물리학 분야와 수학 분야는 각각 85%와 82%를 포함한다. 즉, 뒤의 두 분야에서는 20%에 가까운 과학자들이 다수로 구성된 네트워크에 연결되어 있지 않다는 것이다[10].

평균 거리는 최대 하위 네트워크에 포함된 모든 과학자들 간의 거리를 평균한 것인데, 생물학 분야는 네트워크가 다른 분야 비해 매우 큼에도 불구하고 4.6으로 평균적으로 5단계 이하에서 다른 과학자와 연결된다는 것이다. 반면에 수학 분야는 평균적으로 7단계 이상에서 서로 연결되어 있다[10]. 최대 거리, 일명 네트워크의 지름(diameter)은 최대 하위 네트워크에 포함된 두 과학자간의 거리 중 최대 거리를 의미한다[8].

군집화 정도(clustering coefficient)는 한 명의 과학자가 다른 두 명의 과학자와 각각 공동 저작했을 때, 두 명의 공동 저자 사이에도 공동 저작이 있는 정도를 의미한다. 즉, 네트워크 모형에서 삼각형(triangle)이 나타나는 정도를 의미하고, 이것은 곧 네트워크 구성원들 사이의 연결 밀도가 높다는 것을 의미한다. 군집화 정도를 계산하는 식은 아래와 같다.

$$C_i = \frac{\text{꼭지점 } i \text{와 연결되어 있는 삼각형의 수}}{\text{꼭지점 } i \text{가 중심인 } V \text{자형의 수}}$$

이때, 꼭지점  $i$ 가 중심인  $V$ 형이란 삼각형이 아니라 두 변만 연결된 열린 형태의 도형인데 이때 중심이 꼭지점  $i$ 인 것을 말한다[8]. 그런데 특이하게도 생물학 분야가 매우 낮은 6.6% 정도이고 물리학 분야가 43%로 매우 높다. 마지막으로 계층화 정도(associativity)에서는 물리학 연구자들이 0.36으로 가장 높다. 즉, 물리학 분야의 연구자들은 다른 두 분야의 연구자들에 비해 많은 공동 저자를 가지고 있는 연구자들끼리 서로 협력하는 경우가 많다는 것이다[10].

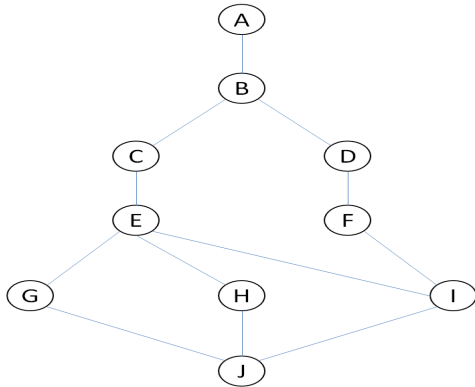


그림 2. 네트워크의 집중도 설명을 위한 예시

앞서 살펴 본 네트워크의 특성을 나타내는 지표들 외에도 집중도(centrality), 즉 한 꼭지점이 자신이 속한 네트워크에서 차지하는 중요한 정도를 나타내는 지표가 있다. 집중도를 나타내는 방법에는 일반적으로 세 가지가 있는데, 이미 앞에서 언급한 공동 저자수, 즉 연결성 기준의 집중도 지표가 그 중 하나이다. 공저자 네트워크에서 연결성기준 집중도가 높다는 것은 공동으로 논문을 작성한 동료 연구자가 많다는 것을 의미한다[8].

다른 두 가지 지표에는 중간성(betweenness)과 근접성(closeness)을 기준으로 산정하는 지표들이 있다. 위의 두 지표를 이해하기 위해서는 두 꼭지점 간의 최단 경로(shortest path)라는 개념을 이해해야 한다[8]. 그림 2에서 보는 바와 같이, 꼭지점 A에서 꼭지점 J에 이르는 가장 짧은 경에는 중간에 4개의 꼭지점을 거치는 총 4가지 방법(A-B-C-E-G-J, A-B-C-E-H-J, A-B-C-E-I-J, A-B-D-F-I-J)이 있다.

중간성 기준 집중도( $C_B(v)$ )는 아래의 식을 이용해서 구할 수 있다. 즉, 네트워크 상의 임의의 두 꼭지점 s와 t의 최단경로의 개수( $\sigma_{st}$ ) 중 꼭지점 v를 통과하는 최단경로 개수( $\sigma_{st}(v)$ )의 비율을 꼭지점 v를 제외한 모든 꼭지점의 쌍에 대해 합산한 것이다. 그리고 네트워크의 크기로부터의 영향을 상쇄시키기 위해서 아래의 식에서 얻은 값을 전체 꼭지점의 개수 n을 이용하여  $(n-1)(n-2)$ 로 나누어 정규화한다[8]. 예를 들면, 그림 2에서 꼭지점 A를 제외한 B부터 J까지의 꼭지점 두 임

의의 두 점 사이의 최단경로 중 A를 통과하는 최단경로는 없기 때문에 꼭지점 A의 중간성 기준 집중도는 0이다. 꼭지점 J의 경우에는 G와 I 사이의 유일한 최단경로 G-J-I 에서 중간에 J가 위치하고 있기 때문에 꼭지점 J의 중간성 기준 집중도는 식  $\frac{1}{(10-1)(10-2)}$ 에 의해 약 0.014를 얻을 수 있다. 즉, 중간성 기준 집중도에서는 꼭지점 J가 꼭지점 A보다 중요하다. 중간성기준 집중도가 높은 연구자란 연구자 네트워크에서 임의의 두 연구자가 직접 공동으로 연구하지 않고 제 3의 연구자를 매개로 공동협력이 이루어질 때, 교량 역할 측면에서 얼마나 중요한가를 나타낸다.

$$C_B(v) = \frac{1}{(n-1)(n-2)} \sum_{s \neq t \neq v \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

근접성 기준 집중도( $C_C(v)$ )는 아래의 식에 의해서 구할 수 있다. 즉, 꼭지점 v의 근접성 기준 집중도는 네트워크 상의 나머지 꼭지점에 이르는 최단거리( $d_G(v,t)$ ) 합의 역수이다[8]. 예를 들면, 그림 2에서 꼭지점 A의 경우, A에서 B까지의 최단거리는 1, C와 D까지는 각각 2, E와 F까지는 각각 3, G와 H, I까지는 각각 4, 그리고 마지막으로 J까지는 5이므로 식  $\frac{1}{(1+2+2+3+3+4+4+4+5)}$ 에 의해 0.036이다. 꼭지점 B의 경우에는  $\frac{1}{(2+1+2+1+3+2+2+2+3)}$ 에 의해 0.056이 된다. 다시 말해서, 꼭지점 B가 꼭지점 A에 비해 네트워크상의 다른 꼭지점들에 접근하기 쉽다는 것을 알 수 있다. 공저자 네트워크에서의 근접성 기준 집중도는 전체 연구자들과 얼마나 가깝게 연결되어 있는가를 판단할 수 있게 해 준다.

$$C_C(v) = \frac{1}{\sum_{t \in \sqrt{v}} d_G(v,t)}$$

[16]에 따르면, 네트워크에서의 위치 중요성을 나타내는 세 가지 집중도 변수인 연결성과 근접성, 중간성 기준 집중도가 모두 연구자의 연구성과와 연구비 수주

에 영향을 준다. 따라서 본 연구에서도 융합기술 연구자 네트워크에서 각 연구자의 집중도 변수를 살펴 볼 것이다. 비록 본 연구가 개별 연구자의 성과를 탐구하는 것은 아니지만 전체 네트워크 차원에서 연구자들이 집중도 측면에서 어떤 차이가 있는 지 살펴봄으로써 융합기술 연구자들의 공저자망에 대한 이해를 높일 수 있기 때문이다.

본 연구는 이전 연구와 다음과 같은 차별성을 가진다. 첫째, 본 연구는 융합기술, 구체적으로 NBI 융합기술 연구자들을 대상으로 분석하고 있다. 기존에 연구자들 사이의 연구 협력관계를 분석한 논문들은 특정 저널에 기고한 연구자이거나 특정 분야의 연구자들을 대상으로 하고 있는데 반해, 본 논문에서는 이질적인 분야의 연구자들 사이에서 어떻게 협력이 이루어지고 있는지를 분석하고 있다. [16]의 연구에서도 전공 분야 사이의 연구 협력에 대한 분석이 이루어지고 있으나 융합기술을 전문적으로 연구하는 집단의 특성을 직접 분석하지 않았다. 따라서 본 연구가 국내 NBI 융합기술 연구자 집단의 네트워크 특성을 이해할 수 있게 해 준다는 점에서 차별성이 있다.

둘째, 본 연구는 NBI 융합기술 연구자들의 연구 협력을 이해함에 있어 이전의 연구보다 효과적이고 실제 상황에 가깝게 분석하고 있다. [16]의 연구는 국내 과학자들의 연구 협력을 공동 연구 차원에서 분석하였는데, 본 연구에서는 연구자 사이의 협력 관계를 가장 잘 보여 주는 공저자 네트워크를 대상으로 하고 있다. 그리고 [16]의 연구에서는 2002년 한 해만을 대상으로 하고 있는데 반해, 본 연구는 그리고 55년간 축적된 논문 정보를 이용하고 있다는 점에서 연구 협력 관계를 보다 실제와 유사하게 이해하는데 기여할 것이다.

### III. 연구 방법

본 연구에 활용한 데이터는 두 가지다. 하나는 국내의 융합기술 연구자들 목록이다. 이 목록은 본 연구의 일환으로 진행된 융합기술 분야 전문가들과의 인터뷰를 통해서 도출된 것이다. 전문가들은 산업계를 비롯해서 학계와 국가연구소의 연구원들로 구성되어 있다. 총

1,440명의 국내 융합기술 연구자를 도출하였다. 다른 하나는 1954년부터 2009년 5월까지 국내에서 발표된 논문을 모아 놓은 한국학술정보(주)의 데이터베이스에서 공학 및 자연과학, 그리고 의학 분야의 총 307,606편의 논문만 추출한 데이터다. 여기에는 논문의 제목과 저자, 게재한 저널명, 그리고 출판 연도 등의 정보가 포함되어 있다. 이 데이터베이스 전체의 연구자 수는 135,948명인데 평균적으로 연구자당 6.09편의 논문을 작성했고 논문당 평균적으로 3.06명의 저자를 가지고 있다. 연구자당 논문 편수나 논문당 저자 수는 이전의 다른 연구 결과([표 2] 참고)와 비슷한 점으로 미루어볼 때 실증 분석에 사용하기에 무리가 없는 데이터베이스라고 볼 수 있다. 그런데, 확보한 융합기술 연구자 중에서 확보한 논문 데이터베이스의 공학 및 자연과학, 그리고 의학 분야의 국내 저널에 논문을 게재한 연구자는 총 1,095명이었다. 누락된 345명은 산업계에 종사해서 논문을 발표하지 않거나 공학 및 자연과학, 의학 분야가 아닌 다른 분야의 전문가들인 경우가 대부분이었다.

논문 정보를 이용하여 과학자들간의 연구 협력을 분석한 논문에서는 대체로 분석하고자 하는 분야를 먼저 정의한다. 예를 들면, [10]의 연구에서는 특정 연구 분야의 논문만을 수집하여 제공하는 데이터베이스 서비스(MEDLINE, SPIRES, NCSTRL 등)를 이용하였다. 따라서 그는 각 데이터베이스에서 특정 기간에 출간된 논문들을 대상으로 집필에 참여한 연구자 정보를 추출하여 분석하였다[12]. 그런데, 나노기술 분야와 같이 범위 정의가 모호한 경우에는 검색어를 이용하여 먼저 나노기술 분야의 논문을 검색하고 이들 논문에 등장하는 연구자들을 대상으로 네트워크 분석을 수행하였다. 이때 나노기술에 사용한 검색어 중에서 “nano-”로 검색된 논문이 전체의 90% 이상을 차지하였다[14]. 그러나 본 논문에서는 위와 같은 전형적인 방법을 사용할 수 없는 제약이 있었다. 첫째, 융합기술 분야를 전담하여 논문 정보를 수집하는 데이터베이스가 없었다. 둘째, 특정 키워드를 이용하여 충분한 양의 논문을 확보할 수 없었기 때문이다. 예를 들면, “융합-이 나”나 “나노-”로 검색할 경우 100편 이하의 매우 적은 양의 논문이 도출될 뿐이었다. 이러한 현실적인 제약 때문에 본 논문에서는



융합기술 전문가를 먼저 도출하고 이들이 참여한 논문을 융합기술 분야의 논문이라고 정의하는 방식을 선택하였다.

한편, 대용량의 논문 데이터를 분석에 이용할 때는 연구자를 어떻게 식별할 것인가에 대한 이슈가 발생한다. 분석 데이터가 연구자별로 정확하게 구분되어 있어야 분석의 신뢰도는 높아짐에도 불구하고 연구자 아이디(ID)가 없는 경우가 대부분이기 때문에 대부분의 논문 데이터를 이용한 연구에서는 항상 연구자 식별 오류의 가능성이 내재되어 있다. 연구자를 올바르게 식별 인식하는 대표적인 사례에는 동명이인의 오류와 이명동인의 오류가 있다. 전자는 동일인이 아님에도 불구하고 단지 이름이 같아서 동일인으로 간주하는 경우이고, 후자는 동일인인데도 불구하고 이름이 다르게 표기되어서 동일인이 아닌 것으로 간주하는 경우이다. 동명이인의 오류는 아시아계 연구자가 많이 포함된 논문 데이터에서 보다 많이 발생하고 있으며, 이명동인의 오류는 연구자 이름이 복수 언어로 표기된 논문 데이터에서 드러난다. 그런데 지금까지 발표된 관련 연구들, 특히 영어로 작성된 논문을 대상으로 분석한 연구들에서는 동명이인과 이명동인의 인식 오류를 심각하게 고려하지 않고 있다. 그것은 서구인들의 이름에서 동명이인이 존재하는 비율이 극히 미미하고, 영어 하나로 이름을 표기해도 이명동인의 비율도 매우 낮을 것이라는 가정이 합리적이라고 받아들여지고 있기 때문이다. 비록 아시아계 연구자들이 분석하고자 하는 논문 데이터에 포함되어 있다고 하더라도 그 비율이 낮으면 연구 결과에 큰 영향을 미치지 않는다고 가정하는 경우도 있다[15].

그러나 본 논문에서 분석하고자 하는 데이터는 국내에서 발표된 논문들을 대상으로 하고 있기 때문에 동명이인의 문제를 고려하지 않을 수 없다. 한국인의 이름에서는 동명이인이 쉽게 발견된다. 한국의 다수 성씨인 ‘김’, ‘이’, ‘박’, ‘최’, ‘정’이 전체 인구의 절반 이상을 차지하고 있고 일명 돌림자를 이용하여 이름을 짓는 관습 때문에 동명이인이 많은 아시아 국가들 중에서도 우리나라가 독보적으로 동명이인이 많다. 그리고 국내에서 발표된 논문에 대한 데이터에서는 이명동인의 오류에 대해서도 고려해야 한다. 국내 저널 중에서는 영문 논

문만을 게재하는 경우가 다수 있어서 논문 저자의 이름이 한글 없이 영어로만 표기된 경우가 있는데, 이때 동일한 저자가 한글 저널에 한글로 논문을 발표할 때 표기된 저자의 한글 이름과 어떻게 연결할 것인가에 대한 문제가 발생한다. 만약 동일인임에도 불구하고 한글 이름과 영문 이름을 서로 연결시키지 않으면 두 명의 각각 다른 연구자로 인식할 수밖에 없기 때문이다. 따라서 본 연구에서는 데이터를 이용하기 전에 동명이인과 이명동인의 문제가 전체 연구 결과에 어떤 영향을 미칠 수 있는 지 사전에 가늠해 볼 필요가 있으며, 해결 가능하다면 사전에 오류를 최소화시켜야 한다.

먼저 동명이인의 오류를 해결책에 대해서 생각해 보자. 예를 들면, 실제로 홍길동이라는 이름을 가진 연구자가 두 명 이상 존재한다고 하자. 이때 주어진 데이터가 연구자의 이름뿐이라면 홍길동이란 동명의 각각 다른 여러 연구자들이 발표한 논문들을 연구자별로 구분하는 것은 불가능하다. 이러한 동명이인 오류를 해결하기 위한 방법 중 하나는 연구자들의 소속기관 정보를 이용하는 것이다. 연구자들의 이름과 소속기관을 함께 고려함으로써 이름이 같다고 하더라도 소속기관이 다른 경우에는 각각 다른 연구자로 인식하는 것이다. 그러나 이 방법 역시 연구자들이 소속기관을 변경하는 경우가 존재하기 때문에, 완벽한 해결책이 되지 않는다. 다른 해결 방법으로는 연구 분야 정보를 연구자 이름과 동시에 고려하여 연구자를 구별하는 방안이 있다. 만약 논문 데이터에 연구 분야 정보가 포함되어 있다면 분야가 다른 논문에 등장하는 동일 이름의 저자를 각각 다른 연구자로 인식하는 것이다. 또 다른 해결방법으로 연구자의 생년 또는 생년월일을 함께 고려할 수 있는데, 생년월일은 개인의 프라이버시 침해 논란을 초래하기 때문에 현실적으로 이용할 수 없는 경우가 대부분이다. 위에서 언급한 세 가지 방법, 즉 소속기관과 연구 분야, 생년월일의 정보를 동시에 이용할 수 있다면 동명이인의 연구자들을 거의 완벽하게 식별할 수 있을 것이다. 그러나 대부분의 국내 논문 데이터베이스는 위의 세 가지 중 하나도 포함하지 않고 있다. 본 논문에서 사용한 논문 데이터에서도 소속기관이나 연구 분야, 생년월일 등의 연구자를 상세하게 구별하는데 필요한 추가

정보가 없기 때문에 동일철자의 저자는 동일인으로 간주한다. 그런데 본 연구의 활용 데이터 중 하나인 전문가 인터뷰를 통해 확보한 1,440명의 융합기술 연구자를 전수로 확인해 보니, 같은 이름의 연구자가 등장하지 않았다. 그리고 분석에 사용한 논문 데이터베이스에서는 연구 분야에 대한 정보가 포함되어 있어서 연구 분야가 다른 동명이인은 다른 연구자로 간주하고 있다.

동명이인을 구분하지 못하는 문제와 함께 동일인이 다르게 표기됨으로써 다른 사람으로 인식하는 것도 심각한 문제이다. 국내 논문의 경우, 저자를 한글이름 또는 영문이름만을 이용하여 표기하기도 하고, 한글이름과 영문이름을 동시에 제공한다[16][17]. 이때 가장 큰 문제가 되는 것은 영어 이름만 제공하는 경우와 한글이름만 제공하는 경우이다. 동일인이더라도 한글이름과 영문이름을 사용하는 경우를 의미적으로 동일하다고 인식하는 것은 매우 어렵다.

본 연구에서는 논문 데이터베이스에 등장한 한글이름과 영문이름은 다르게 취급하였으며, 동일한 철자의 이름에 대해서는 동일인으로 간주하였다. 한글이름과 영문이름을 다르게 취급하면 동일인이 여러 명으로 인식되기 때문에 연구자 수가 증가할 것이며, 반면에 한글이건 영문이건 동일한 철자의 이름에 대해서 동일인으로 간주하게 됨으로써 연구자 수는 감소할 것이다. 따라서 증가와 감소가 서로 상쇄되어서 전체 연구자 수에 일반적으로 증가하거나 감소하는 영향을 미치지 않을 것으로 기대된다.

위에서 밝힌 논문 정보가 30만 건 이상의 논문을 포함하는 대용량 데이터이고 공동저자가 텍스트 형태로 서로 연결되어 있기 때문에 각각의 독립된 저자로 추출하기 위해서 MS-SQL 서버를 이용하여 데이터를 정제하였다. 이렇게 정제된 데이터는 네트워크 분석용 소프트웨어인 Pajek을 이용하여 1,095명을 꼭지점으로 연결하는 네트워크를 구성하여 네트워크의 특성을 나타내는 값을 얻었다. 그리고 얻어진 국내 융합 기술 연구자들의 네트워크 특성을 기존에 발표된 다른 분야(의학, 이론 물리학, 컴퓨터 공학, 수학 등)와 비교 분석하였다. 마지막으로 국내 융합기술 연구자들을 산업계와 학계, 국가연구소로 구분하여 각각의 특성을 살펴보았다.

#### IV. 국내 융합기술분야 연구자 네트워크 분석 결과

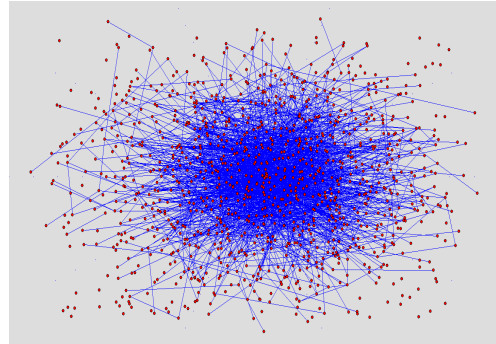


그림 3. 융합기술 연구자의 공저자 네트워크

그림 3은 융합기술 연구자들이 중앙부를 중심으로 매우 밀집된 형태로 서로 연결되어 있으므로 알 수 있다. 즉, 중앙부에 위치한 연구자들이 융합기술 관련 논문 작성에서 중요한 역할을 담당하고 있다는 것을 확인할 수 있다. 한편, 주변부에는 중심부에 위치한 최대 서브 집단과 연관은 되어 있지만 이들 상호간의 연결이 거의 없는 것을 확인할 수 있다. 따라서 국내 융합기술 네트워크는 중심과 주변부의 연결이 중심이 되는 집중형의 네트워크라고 말할 수 있으며, 일부 연구자들은 최대 하위 집단과 연결되어 있지 않다.

표 2. 본 연구결과와 기존 연구결과 비교

구분	NBI	MED	LA	SPIRES	NCSTRL	MATH
저자 수	1,095	1,520,251	52,909	56,627	11,994	253,339
논문 수	47,895	2,163,923	98,502	66,652	13,169	-
저자당 논문 수	43.74	6.4	5.1	11.6	2.55	6.9
논문당 저자 수	-	3.75	2.53	8.96	2.22	1.45
평균 공동저자수	5.15	18.1	9.7	173	3.59	3.9
최대 하위 네트워크	-	92%	85%	88.7%	57.2%	82%
평균 거리	3.9	4.6	5.9	4.0	9.7	7.6
최대 거리	9	24	20	19	31	27
군집화 정도	0.108	0.066	0.43	0.726	0.496	0.15
계층화 정도	-	0.13	0.36	-	-	0.12

본 연구의 결과를 기존 연구결과와 비교하면 [표 2]과 같다. 분석에 포함된 연구자들은 평균적으로 43.74편의 논문을 작성하여 발표하였는데, 매우 높은 연구생산성을 보여 주고 있다. 본 분석에서 선정한 1,095명의 연구자들이 인터뷰를 통해 도출된 전문가 집단인 점에 미루어 볼 때 이들이 매우 생산성이 높은 연구자 집단이라는 점은 당연하다. 높은 생산성에 비해 연구자들은 평균적으로 5.15명의 다른 연구자들과 공동으로 논문을 발표하였다. 앞서 살펴 본 [10]과 [13]의 연구결과와 비교하면, 고에너지 물리학(173명)과 의학(18.1명), 이론물리학(9.7명)의 연구자들에 비해 낮은 것을 확인할 수 있다. 반면 수학(3.9명)과 컴퓨터 과학(3.59명)보다는 조금 높다.

최대 하위 네트워크에서 평균 거리가 3.9이고 최대 거리가 9인 것으로 미루어볼 때, 이 네트워크의 구성원들 간의 거리가 다른 과학분야에 비해 가까운 것을 알 수 있다. 그러나 군집화 정도에는 0.108로 매우 낮아 네트워크 구성원들 사이의 연결 밀도가 낮다는 것을 의미한다.

연구결과를 자세히 살펴보면, [표 3]의 공동 저자수 분포(4)에서 알 수 있듯이 전체의 70% 가량이 5명 이하의 공동 연구자를 가지고 있다. 그 중에서 3명의 연구자는 공동 저자 없이 단독으로 연구 결과를 발표하고 있다. 공동 저자수가 가장 많은 연구자는 총 42명의 공동 연구자와 협력하고 있으며 그 다음으로는 33명의 공동 저자를 가지고 있는 연구자이다. 공동 저자수에 따른 꼭지점의 집중도를 나타내는 지표인 연결성 기준 집중도(degree centrality)에서 보면, 전체 연구자의 절반 가까이가 2명 이하의 공동 연구자를 두고 연구하고 있음을 알 수 있다.

표 3. 융합기술 연구자 네트워크의 인당 공동 저자수 분포

구 분	빈도수	비중	누적 빈도수	누적 비중
0	3	0.27	3	0.27
1	281	25.67	284	25.94
2	196	17.90	480	43.84
3	120	10.96	600	54.79
4	101	9.22	701	64.02
5	68	6.21	769	70.06
6	42	3.84	811	74.06
7	35	3.20	846	77.26
8	39	3.56	885	80.82
9	34	3.11	919	83.93
10	28	2.56	947	86.48
11	20	1.83	967	88.31
12	13	1.19	980	89.50
13	16	1.46	996	90.96
14	15	1.37	1011	92.33
15	10	0.91	1021	93.24
16	15	1.37	1036	94.61
17	14	1.28	1050	95.89
18	5	0.46	1055	96.35
19	4	0.37	1059	96.71
20	4	0.37	1063	97.08
21	7	0.64	1070	97.72
22	3	0.27	1073	97.99
23	4	0.37	1077	98.36
24	5	0.46	1082	98.81
25	2	0.18	1084	99.00
26	3	0.27	1087	99.27
27	3	0.27	1090	99.54
28	0	0.00	1090	99.54
29	1	0.09	1091	99.63
30	2	0.18	1093	99.82
31	0	0.00	1093	99.82
32	0	0.00	1093	99.82
33	1	0.09	1094	99.91
34	0	0.00	1094	99.91
35	0	0.00	1094	99.91
36	0	0.00	1094	99.91
37	0	0.00	1094	99.91
38	0	0.00	1094	99.91
39	0	0.00	1094	99.91
40	0	0.00	1094	99.91
41	0	0.00	1094	99.91
42	1	0.09	1095	100.00

표 4. 융합기술 연구자 네트워크의 중간성기준 집중도 분포

구 분	빈도수	비중	누적 빈도수	누적 비중
0.0000	538	49.13	538	49.13
0.0000 ~ 0.0134	540	49.32	1078	98.45
0.0134 ~ 0.0268	15	1.37	1093	99.82
0.0268 ~ 0.0402	2	0.18	1095	100.00

4) 추가적으로 본 연구에서 분석을 위해 설정한 1,095명의 융합기술개발의 전문가들과 공동으로 일회 이상의 논문을 작성한 모든 연구자는 데이터베이스 전체의 연구자인 135,948명 중에서 29,143명인 것을 확인했다. 이들을 대상으로 인당 공동 저자수를 산출하였더니, 평균적으로 3.78명이었고 최대 737명의 공동 연구자를 갖고 있는 연구자가 있었으며 상위 20위의 연구자들은 최소 365명 이상과 공동으로 논문을 발표하였다. 위에서 보는 바와 같이, 29,143명을 기준으로 하면 1,095명을 기준으로 했을 때보다 평균 공동 저자수는 감소하지만 상위 연구자는 매우 많은 공동 저자와 협력하는 것을 알 수 있다.

표 5. 융합기술 연구자 네트워크의 근접성 기준 집중도 분포

구 분	빈도수	비중	누적빈도수	누적 비중
0.0000	283	25.84	283	25.84
0.0000 ~ 0.0882	16	1.46	299	27.31
0.0882 ~ 0.1764	228	20.82	527	48.13
0.1764 ~ 0.2646	568	51.87	1095	100

그런데 네트워크 상의 꼭지점의 집중도, 즉 네트워크 상에서 해당 꼭지점이 얼마나 중요한 위치에 있는가를 측정하는 기준에는 연결성 기준 집중도 외에도 중간성 기준 집중도(betweenness centrality)와 근접성기준 집중도(closeness centrality)가 있다. 본 연구에서 이용하고 있는 데이터를 통해 위의 두 가지 추가적인 집중도 지표를 산출하면 평균적으로 각각 0.001과 0.139이다. 그리고 중간성 기준 집중도에서는 두 개의 연구자가 융합기술 전문가들 사이의 최단경로를 형성하는데 중요한 역할을 하고 있음을 알 수 있다. 즉, 위에서 언급한 두 개의 꼭지점을 다른 연구자들과 연결에서 다리(bridge) 역할을 담당하고 있기 때문에 전체 네트워크에서의 협력관계를 신속하게 유지시키는데 중요하다. 만약 이 두 명의 연구자가 없다면 연구자간의 평균거리를 급속하게 증가하거나 아예 연결이 끊길 수도 있다. 반면에 근접 기준의 집중도에서는 다수의 꼭지점이 다른 꼭지점들에 비슷한 거리만큼 떨어져 있음을 알 수 있다.

표 6. 융합기술 연구자 네트워크의 집중도 기준 중요 꼭지점

꼭지점 식별번호	공동 저자수 기준	중간성 기준	근접성 기준
9	42	0.040	0.265
13	33	0.029	0.263
27	30	0.018	0.251
32	30	0.021	0.254
18	29	0.015	-
30	27	0.017	0.250
59	27	0.016	0.247
17	27	0.016	0.248
12	26	-	0.250
38	26	-	0.246
33	-	0.016	-
36	-	0.015	-
34	-	-	0.249

융합기술 연구 협력 네트워크의 집중도 측면에서 중요한 꼭지점을 정리하자면 [표 6]과 같다. 꼭지점 식별번호 9번의 연구자를 비롯한 총 9명의 연구자(세 가지

기준 모두에서 상위 10위 안에 든 경우)가 1,095명의 연구자 네트워크에서 매우 중요한 역할을 하고 있음을 알 수 있다. 특히, 2명의 연구자(식별번호 9번과 13번)는 전체 연구 협력 네트워크에서 매우 중요한 역할을 담당하고 있다. 따라서 이들은 국내 융합기술 연구 개발에 있어서 리더 역할을 한다고도 볼 수 있다.

### V. 결어

융합기술 연구자에 대한 연구의 특성 및 네트워크에 대한 분석을 통해 다음 세 가지를 확인할 수 있었다. 첫째, 융합기술 연구자들은 인당 논문 편수가 다른 과학기술 분야에 비해 매우 높다. 즉, 높은 연구 생산성을 보여주고 있으며, 융합기술 분야에서 새로운 지식 창조가 활발하게 진행되고 있음을 알 수 있다.

둘째, 높은 생산성에 비해 인당 공동 저자 수는 매우 적었다<sup>5)</sup>. 이것은 곧 융합기술연구자 그룹의 상당수가 논문의 공동저자를 활발하게 진행하고 있지 않다는 것을 의미한다. 융합기술 과학자들이 공동으로 논문을 작성하는 비중이 낮은 것은 여러 측면에서 추측해 볼 수 있다. 그 중 하나는 융합기술에 대한 연구에서 다양한 분야의 연구자들이 참여하지 못하고 제한된 협력 관계만을 지향한다는 것을 의미한다고 볼 수 있다. 즉, 새로운 전문들과 연구를 진행하기 보다는 이전에 연구를 함께 수행한 동료 연구자와 협력이 지속되어서 공동 연구자 수가 증가하지 않는 것일 수도 있다. 더구나 이종 분야 전문가들이 공동으로 논문을 작성할 때는 각 분야로 특화된 논문 작성 방식이나 데이터 분석 방식으로 인해 공동 집필하는 것에 어려움을 느낄 가능성도 있다.

셋째, 융합기술연구자들은 네트워크의 구성원들 간의 거리가 다른 과학 분야에 비해 가깝게 나타났지만, 삼각형 형태의 삼자연구 등 서로 공동으로 연구하는 형태로 발전하지 않고 양자 간의 공동 연구로만 머물러 있는 경우가 많고, 소수의 연구자가 네트워크 전체에

5) 융합기술 연구자들은 평균적으로 5.15명의 연구자들과 공동으로 논문을 발표하고 있어 의학(18.1명)과 이론물리학(9.7명), 고에너지 물리학(17.3명) 분야에 비해 낮아 공동저자를 상대적으로 활발하게 진행하고 있지 못하다.

큰 영향력을 행사하는 형태이다. 양자 위주의 공동 연구는 우리나라 특유의 사회적 특성이 반영된 것이라고 볼 수도 있다. 연구자들 간에 연령이나 소속 조직, 그리고 학연과 지연 등에 의해서 연구 협력의 기회가 제한되기 때문일 것으로 추측된다. 그런데 이러한 사회적 현상이 연구 협력에까지 영향을 준다면 연구 협력을 통한 연구 생산성 향상을 위해서 정부는 융합기술 분야에 대한 연구자 네트워크를 활성화할 수 있는 기반조성을 해야 할 것이다.

그리고 마지막으로, 융합연구가 활발하게 공동연구가 이루어지기 위해서 고려해야 될 사항으로 연구자간 인력이동(Mobility)에 대한 확대를 생각해 볼 수 있다. 지식의 전파는 교육과정을 통해서도 이루어지지만, 연구인력 자체의 이동을 통해서도 이루어진다. 그러나 국내 연구자들의 인적 이동은 쉽지 않다. 만약 융합기술 개발 연구를 수행하던 교수가 다른 대학으로 자리를 옮겼을 때, 기존에 수행하던 연구과제를 새로운 대학이나 연구기관에서 수행할 수 없게 제한하는 경우가 대부분이며, 심지어 함께 연구를 수행하던 전임연구인력이나 학위과정학생이 함께 움직일 수 없다. 지식의 확대 재생산은 다양한 집단의 연구자들 간 통합을 통해 이루어져야 한다. 특히, 융합기술개발의 시대에는 종적인 지식의 전달보다 횡적인 지식의 융합이 더 중요하다. 또한 융합기술 분야에서는 타 분야 연구자와의 의사소통의 질적 수준이 연구성과에 중대한 영향을 미치기 때문에 그 중요성이 더욱 커진다. 최근에는 R&D 국제화의 진전으로 인해 외국 연구자와의 의사소통(Globalization)이 매우 중요한 이슈가 되고 있다. 영어의 중요성은 말할 것도 없고, 외국어로 된 융합기술 분야의 전문용어를 한국어로 체계적으로 변환하는 것도 필요하다. 이를 위해 언어학자, 지식정보시스템 전문가, 각 기술분야의 전문가 등이 함께 참여하는 공동연구(Collaboration)가 필요하다.

본 연구의 한계점에 있어서, 융합기술 분야 연구자에 대한 공동연구 네트워크분석은 매우 중요한 주제임에도 불구하고 지금까지 이에 대한 분석 및 연구는 매우 미흡하다. 본 연구가 SNA 연구임에도 불구하고 1,095개의 꼭지점을 이용했다는 점에서 실증 분석의 의미를

계약할 수 있어, 조금 더 큰 용량의 데이터를 이용할 수 있는 방안을 찾아야 할 것이다. 무엇보다도 보다 많은 국내 융합기술연구자의 명단을 확보하는 것이 하나의 방안이 될 것이다. 동일철자의 이름을 동일인으로 간주하고 영어이름만 존재할 경우 별개의 저자로 인식하게 됨으로써 현실을 정확하게 반영하지 않을 가능성이 있다. 비록 동일철자에 의한 저자 수의 감소와 영어이름 별개 인식에 의한 저자 수의 증가가 서로 어느 정도 상쇄될 것이라고 하더라도 연구자들이 작성한 논문이 제대로 반영되지 않은 가능성이 존재한다. 본 연구에 사용된 융합기술 연구자들의 논문이 모두 융합기술과 관련되었다기보다는 연구자들이 초기에는 각자의 전문영역을 연구하다가 융합기술 분야로 영역을 전환하거나 확장하였을 가능성이 높다. 그런데 분석에 사용된 논문을 융합 분야에 관한 것과 그렇지 않은 것을 구분하고 있지 않기 때문에 연구결과를 해석하는데 제약이 따른다. 또한, 해외의 사례 연구에서도 드러난 것처럼 동명이인 또는 이인동명의 문제를 쉽게 해결할 수 있는 방안이 강구되어야 한다. 만약 위와 같은 한계가 모두 해결된다면 연구자들이 지식을 생성하기 위해서 협력하는 현황을 네트워크 모형으로 구현하게 되는 것이므로 융합기술 분야 연구자들이 실제 어떻게 협력하고 있는지를 보다 세밀하게 이해할 수 있을 것으로 기대된다.

#### 참 고 문 헌

- [1] J. S. Katz and B. R. Martin, "What is research collaboration?," *Research Policy*, Vol.26, pp.1-18, 1997.
- [2] G. Melin and O. Persson, "Studying research collaboration using co-authorships," *Scientometrics*, Vol.36, No.3, pp.363-377, 1996.
- [3] J. W. Grossman and P. D. F. Ion, "On a portion of the well-known collaboration graph," *Congress Numerantium*, Vol.108, pp.129-131, 1995.
- [4] A. J. Lotka, "The frequency distribution of

scientific productivity," Journal of Washington Academy of Science, Vol.16, pp.317-323, 1926.

[5] N. Pravdic and V. Oluc-Vukovic, "Dual approach to multiple authorship in the study of collaboration/ scientific output relationship," Scientometrics, Vol.10, pp.259-280, 1986.

[6] M. D. Gordon, "A critical reassessment of inferred relations between multiple authorship, scientific collaboration, the production of papers and their acceptance for publication," Scientometrics, Vol.2, pp.193-210, 1980.

[7] A. E. Nudelman and C. E. Landers, "The failure of 100 divided by 3 to equal  $33\frac{1}{3}$ ," The American Sociologist, Vol.7, pp.9-11, 1972.

[8] M. E. J. Newman, "The structure and function of complex networks," SIAM Review, Vol.45, pp.167-256, 2003.

[9] S. Milgram, "The small-world problem," Psychology Today, Vol.1, pp.62-67, 1967.

[10] M. Newman, "Coauthorship networks and patterns of scientific collaboration," Proceedings of the National Academy of Sciences, Vol.101, Suppl.1, pp.5200-5205, 2004.

[11] M. E. J. Newman, "Ego-centered networks and the ripple effect," Social Networks, Vol.25, pp.83-95, 2003.

[12] M. Newman, "The structure of scientific collaboration network," Proceedings of the National Academy of Sciences, Vol.98, No.2, pp.404-409, 2001.

[13] J. W. Grossman, "The evolution of the mathematical research collaboration graph," Congress Numerantium, Vol.158, pp.202-212, 2002.

[14] X. Li, Y. Lin, H. Chen, and M. C. Roco, "Worldwide nanotechnology development: a comparative study of USPTO, EPO, and JPO patents (1976-2004)," Journal of Nanoparticle

Research, Vol.9, No.6, pp.977-1002, 2007.

[15] X. Li, H. Chen, Y. Dang, Y. Lin, C. A. Larson, and M. C. Roco, "A longitudinal analysis of nanotechnology literature: 1976-2004," Journal of Nanoparticle Research, Vol.10, pp.3-22, 2008.

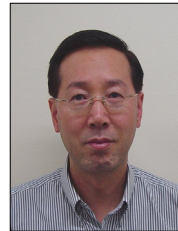
[16] 김용학, 윤정로, 조혜선, 김영진, "과학기술 공동 연구의 연결망 구조: 좁은 세상과 위치 효과", 한국사회학, 제41집, 제4호, pp.68-103, 2007.

[17] 황명호, 안중호, 장정주, "협력 네트워크 패턴에 관한 연구: MIS Quarterly 공저자 분석을 중심으로", 한국전자거래학회지, 제13권, 제4호, pp.193-207, 2008.

#### 저자 소개

##### 이 중 만(Jungmann Lee)

정회원



- 1986년: 고려대학교 경영학과 (경영학사)
- 1997년: New York시립대학교 (경제학박사)
- 2001년: ETRI 기술정책연구팀 선임연구원

▪ 현재: 호서대학교 디지털 비즈니스학부 조교수  
디지털 기술경영(MOT) 사업단장, 산학협력 중심대학 사업 부센터장

<관심 분야> : 과학기술 기획 및 인력정책, 산학협업, 콘텐츠 기술경영

##### 최 민 석(Minseok Choi)

정회원



- 1999년: 연세대 산업시스템공학과 졸업
- 2001년: ETRI 기술정책연구팀 연구원
- 현재: KAIST 테크노경영대학원(경영공학 박사과정)

<관심 분야> : 기술정책, 과학적 협력, 지식경영, 디지털 사회