

핵심연구자 선별을 위한 분석방법 연구

The study on the analysis method to find core researchers

이재민, 김강희, 고병열, 이호신, 박영욱, 권오진
한국과학기술정보연구원

Lee Jae-Min, Kim Kang-hoe, Koh Byoung-Youl,
Lee Ho-Shin, Park Youngwook, Kwon Oh-Jin
Korea Institute of Science and Technology
Information

요약

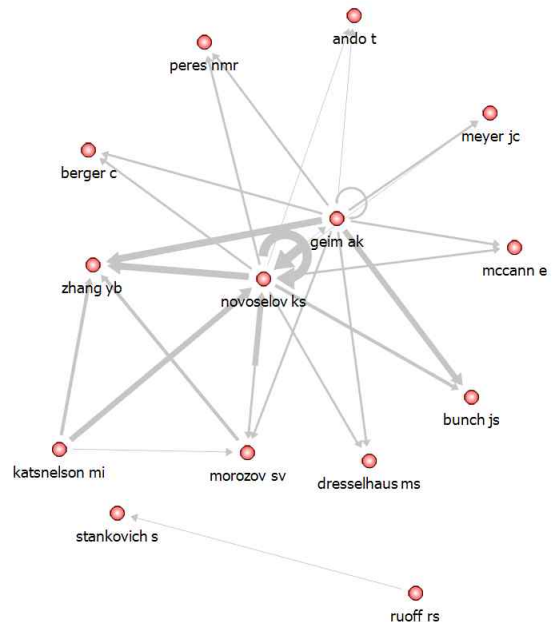
본 연구는 분야별 핵심 연구자를 선별하기 위한 방법론을 다루고자 한다. 핵심 연구자를 선별할 때 가장 흔히 사용하는 방법 중에 하나는 높은 피인용회수(Timces Cited)를 갖는 연구자를 찾는 것이지만 피인용회수가 곧 연구 성과의 순위를 대변한다고 볼 수 없다. 본 연구에서는 특정 학문 분야에서의 연구 성과에 대한 분석을 토대로 핵심연구자를 선별하기 위한 요인들을 수립하고 검증하고자 한다.

I. 서론

흔히 사용되는 Impact Factor의 경우, 저널 논문의 평균적인 파급정도를 나타내는 것으로 개별논문의 성과를 측정하기에는 부족한 지표라고 할 수 있다. 이에 반해 피인용회수는 개별 논문의 질 혹은 파급정도를 계량하는데 가장 보편적이고 효과적인 지표라고 할 수 있으며, 노벨상 등 과학기술 부분 수상의 평가기준으로 언급되는 것이 대표적인 활용의 예이다. 하지만 최근에 피인용회수의 취약점을 보완하기 위한 몇 가지 시도들이 이루어지고 있다. 피인용회수는 동일 분야 연구자들의 평가(peer-review)가 간접적으로 포함되어 있는 것으로 여겨지지만 모든 인용의 영향력을 동등하게 인정하는 문제점을 갖고 있다. 이런 문제 해결을 위해 몇몇 시도가 있었는데, Sergi et al.는 인용하는(citing) 논문의 질에 따라 가중치를 주는 방법으로 google의 pagerank 알고리즘을 이용하여 핵심연구자를 선별하고자 하였다¹⁾. 또 다른 시도로는 논문의 인용-피인용 관계를 사회네트워크의 틀로 이해하고 계량하려는 시도가 있었다²⁾³⁾. 본 연구에서는 고피인용 회수를 갖는 논문들을 바탕으로 네트워크 분석을 수행하여 핵심연구자를 선별하고 그 사이의 관계를 분석하고 하였다.

본 연구에서는 최근 응집물질물리 분야에서 가장 활발히 논의되고 있는 Graphene 관련 논문들의 인용-피인용 관계를 분석하였다. Web of Science에서 Graphene 관련 논문 총 10276 편을 수집하였고, 그 중 피인용회수가 100위권인 논문의 author와 cited author의 관계를 통해 핵심 연구자를 선별하고자 하였다. 데이터는 WoS DB를 통해 수집한 후 Knowledematrix 를 통해 준비하였고, 네트워크 분석은 Netminer 3.0을 사용하여 수행

하였다. (cited author field의 경우 각피인용 논문당 제 1저자 1인만 기록되어 있어 추가 연구를 진행하고 있음.)



▶▶ 그림 1. Graphene 분야 Author-Cited Author 네트워크 분석

그림 1의 네트워크를 보면 Novoselov 과 Geim 을 중심으로 citation 관계가 형성되어 있음을 알 수 있다. Novoselov 의 경우 in-degree 의 중심에 있으며 Geim 의 경우는 Out-degree의 중심에 있어 두 사람을 중심으로 핵심 연구가 진행되었음을 의미한다. 아래의 표에는 이번 분석의 결과로 얻어진 저자별 연결정도(Degree)중

심성과 근접(Closeness)중심성, 그리고 저자별 총피인용 회수를 상위 10명에 대해 정리하였다.

표 1. 저자별 연결정도(Degree)중심성, 근접(Closeness)중심성, Total Times Cited,

Author	In-Degree Centrality	Out-Degree Centrality
novoselov ks	0.130673	0.22793
zhang yb	0.115212	0.015461
berger c	0.092768	0.02793
geim ak	0.076309	0.240898
dresselhaus ms	0.073317	0.118204
stankovich s	0.064838	0.050374
ferrari ac	0.056359	0.069825
ijijima s	0.047382	0.017955
morozov sv	0.038404	0.068329

Author	In -Closeness	Out -Closeness
novoselov ks	0.144511	0.405796
zhang yb	0.137215	0.313322
berger c	0.127845	0.29749
dresselhaus ms	0.124262	0.383114
stankovich s	0.117176	0.333175
ferrari ac	0.113544	0.352979
geim ak	0.112784	0.414589
saito r	0.110709	0.291881
ijijima s	0.10843	0.248813

Cited Authors	Times Cited
novoselov ks	8341
geim ak	3376
zhang yb	3009
berger c	1881
castroneto ah	1334
stankovich s	2114
katsnelson mi	1540
ferrari ac	1377
saito r	1694
son yw	1430

연결정도중심성은 한 노드에 직접적으로 연결된 노드들의 합으로 얻어지지만 근접중심성(Closeness centrality)은 직접적으로 연결될 노드와 간접적으로 연결된 모든 노드와의 거리를 고려한다. 근접중심성은 두 번째 표에 정리되어 있는데 연결정도중심성과 비슷하게 in-closeness 수치는 Novoselov가 out-closeness는 Geim이 높게 나타난다. Zhang 의 경우 연결정도중심성,

근접 중심성 모두에서 2위권을 유지하고 있다. 전체 인용-피인용 네트워크는 직, 간접적으로 Novoselov, Geim 그리고 Zhang을 중심으로 연결되어 있다. 이번 분석에서 한 가지 간과된 점은 교신저자의 중요성이다. Zhang의 주요업적의 교신저자인 김필립은 학계에서 큰 비중을 차지하고 있으나 이번 네트워크 분석에서는 주요 노드로 인식되지 못하였다. 이는 제1저자 위주로 네트워크 노드를 설정하기 때문인데, 이에 대한 보완 연구를 진행하고 있다.

마지막으로 Dresselhaus 의 경우, 총피인용 회수로 보면 12위 권에 머무르고 있지만, 이번 네트워크 분석 결과에서 주요한 노드로서 연결정도중심성 5위, 근접중심성 4위로 랭크되어 있다. 이는 Dresselhaus의 연구성과가 상대적으로 전체 연구자보다는 주요연구자들에게 더 많이 인용되고 높은 인정을 받고 있는 것으로 추정해볼 수 있다.

이번 네트워크 분석은 피인용 회수 기반의 업적 평가를 보완하는 방법을 찾기 위해 시도되어졌으며, 인용회수에서는 볼 수 없었던 저자들 간의 연결관계를 inward(cited), outward(citing) 방향에 대해서 그리고 직접, 간접연결 정도 관점에서 고찰할 수 있었고 그 관계를 시각화하여 쉽게 이해할 수 있도록 하였다.

■ 참고 문헌 ■

- [1] Sergi Maslov, Sidney Redner, "Promise and Pitfalls of Extending Google's pageRank Algorithm to Citation Networks, The Journal of Neuroscience 28(44) 11103-11105(2008)
- [2] 이재윤, "계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구" 한국문헌정보학회지 제40권 제3호(2006)
- [3] 권오진, 노경란, 서진이, 장태중, 문영호, "특허 인용 정보를 이용한 중요 특허 탐색 방법에 관한 연구" 제 10회 한국과학기술정보 인프라워크숍 584-592 (2005)