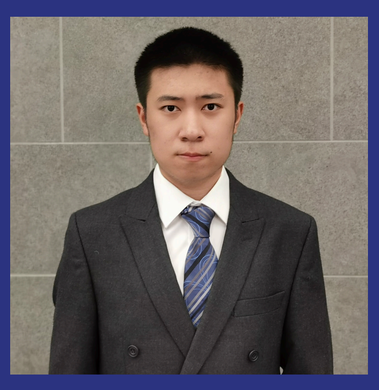
ANALYSIS OF MAPPING TECHNOLOGICAL DISTANCE IN ACADEMIA

A Study on the Faraday Institution



Zhengyang Li Imperial College London Supervisor: Dr Denis Cumming University of Sheffield

Motivation:

- Innovation and research are crucial factors that drive the development of technology and the growth of economy [1].
- In the field of academic research, it is argued that collaborations and knowledge transfers are instrumental for academic outputs [2].
- Therefore, it is necessary to investigate the factors that govern collaborations between academic groups, where technological distance takes a significant role.

Abstract:

- Technological distance, which measures the proximity of the technology positions, is an underlying factor governing collaboration and knowledge transfer in academia. In the literature, various methodologies of constructing the data frame and calculating the technological distance are raised.
- This project conducts a study of technological distance about different research projects of the Faraday Institution, using two sets of data categorizations and three methods of calculation. The results are evaluated comparing each categorization, and each method of calculation.
- It is shown that Euclidean distance and overlapping angle methods in the first set of data have an identical order, since Web of Science categories is a systematic categorization.
- However, across two data sets, min-complement distance manifests greatest consistency, for it projects non-overlapping sections to less dimensions, and possesses the property of Independence of Irrelevant Patent Classes.
- Thus, min-complement distance is likely to be the most appropriate method to characterize technological distance, and categories should be chosen according to research aims and characteristics of the data.

Methods:

- The published works of the principal investigator and project leaders of each project are sampled from the database Web of Science [3].
- Two sets of categories are established: Web of Science categories and common key words in the titles of publications.
- The data frame is then constructed as knowledge vectors with each component being the ratio between the number of published works in one category and the total number of published works of the project.
- Calculations of technological distance are performed using the method of Euclidean distance, overlapping angle, and min-complement distance.

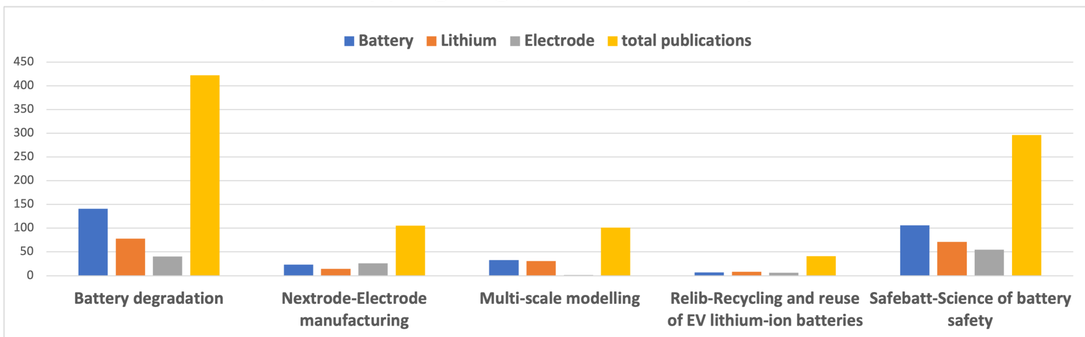


Figure 1: Publication summary by key words in titles

Results and analysis:

- All research projects possess a relatively large proportion of works related to Lithium-ion battery technologies, while their focus on the aspect of this topic may vary.
- For a systematic set of categories, Web of Science Categories, the results from the first two methods are consistent, but some deviations occur because of non-overlapping dimensions. For a more specialized set of categories, common key words in titles, there are discrepancies among all three methods because of the projection in higher dimensions and the larger amount of non-overlapping dimensions.
- However, between two sets of data, min-complement distance manifest greatest consistency, for it has a suitable dimension for calculation, a matching normalization method, and the property of Independence of Irrelevant Patent Classes.

Table 1 and 2: Results of calculation of technological distance

Order	Knowledge vector	Euclidean distance	Order	Knowledge vector	Euclidean distance
1	p1, p4	0.146674105	1	q2, q5	0.143467786
2	p2, p5	0.166316516	2	q1, q5	0.165719031
3	p3, p5	0.169571174	3	q1, q2	0.173557615
4	p1, p2	0.18375939	4	q3, q5	0.218387358
5	p1, p5	0.191124058	5	q1, q4	0.256938621
6	p2, p4	0.22182127	6	q1, q3	0.257866405
7	p4, p5	0.232798567	7	q4, q5	0.258106416
8	p1, p3	0.299138413	8	q2, q4	0.270248838
9	p2, p3	0.304806606	9	q2, q3	0.278637814
10	p3, p4	0.322422144	10	q3, q4	0.299607685

Order	Knowledge vector	Min-complement distance	Order	Knowledge vector	Min-complement distance
1	p1, p5	0.266725704	1	q1, q5	0.347470238
2	p2, p5	0.27994737	2	q2, q5	0.352573191
3	p1, p2	0.306507597	3	q1, q2	0.376108727
4	p3, p5	0.313520606	4	q3, q5	0.457953146
5	p1, p4	0.33141542	5	q1, q3	0.55643504
6	p2, p4	0.422733144	6	q2, q3	0.625533117
7	p4, p5	0.438193468	7	q1, q4	0.626176176
8	p1, p3	0.484549957	8	q3, q4	0.636871982
9	p2, p3	0.496273292	9	q4, q5	0.64711028
10	p3, p4	0.563841497	10	q2, q4	0.669817887

Conclusions/Impact:

- it is concluded that min-complement distance could be used as the priority method when calculating technological distance, and categories should be chosen according to the aim of the investigation and the characteristics of the data.
- The insights of the research above could also be extended to other stakeholders such as databases and publishers, that organizing a mathematically sensible category system to classify publications is necessary to fully convey the information of technological distance to a wider audience, and that contents of publication titles should also be discreet and appropriate.
- The visualizations of technological distances can facilitate the researchers' understandings about their positions in the knowledge space, and further promote the study of different hypotheses of research policy, improving collaborations and knowledge transfers.

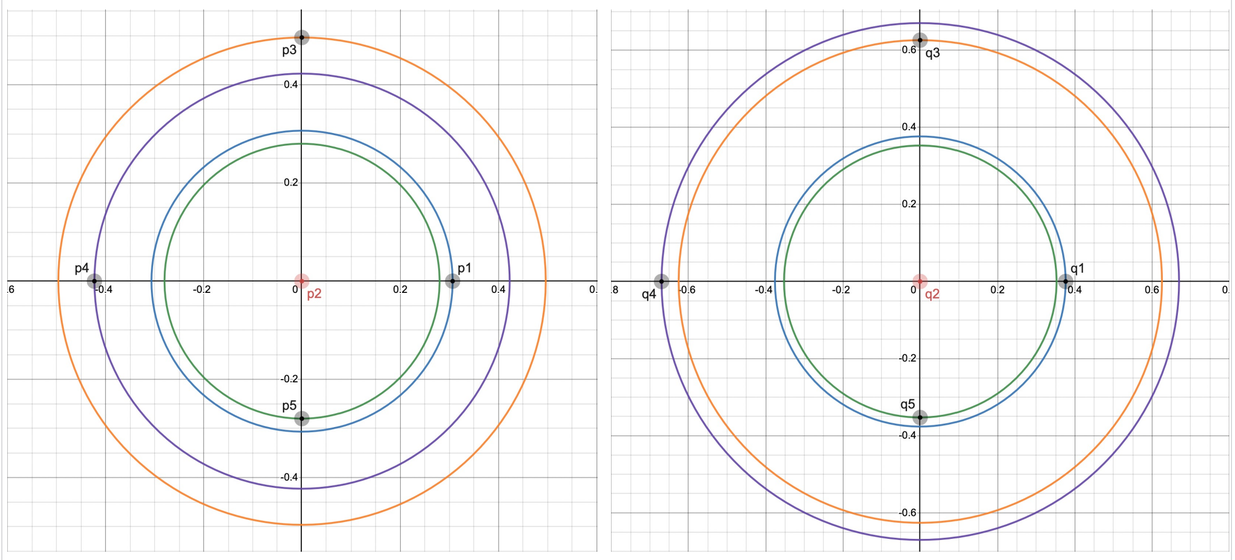


Figure 2 and 3: Visualization of the min-complement distance regarding "Nextrode-Electrode manufacturing"

Next steps:

- More advanced methods regarding the main text of publications are needed.
- Possible solutions may include using machine learning and Natural Language Processing Algorithm [4], where computational tools can be utilized to better understand the contents of publications.
- Investigations of establishing more comprehensive and mathematically sensible categories should also be made, improving the accuracy and sensitivity of the measurements.

References

[1] Drejer I, Jørgensen BH. The dynamic creation of knowledge: Analysing public–private collaborations. Technovation. 2005 Feb 1;25(2):83-94.

[2] Stein NV, Sick N. Technological distance in academic collaborations: Evidence from battery research. International Journal of Innovation Management. 2014 Dec 13;18(06):1440011.

[3] Reuters, T. Web of Science [Internet]. 2022 [cited 2022 Aug 23]. Available from: https://clarivate.com/webofsciencegroup/solutions/web-of-science/

[4] Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. Journal of the American Medical Informatics Association. 2011 Sep 1;18(5):544-51.

Intern bio

Zhengyang Li is a student from Imperial College London studying mathematics. He is interested in applying mathematics to scientific fields and solving real-world problems using mathematics.

THE FARADAY INSTITUTION
NEXTRODE
ELECTRODE MANUFACTURING
The University Of Sheffield.
Imperial College London
UK Research and Innovation