One Venue, Two Conferences: The Separation of Chinese and American Citation Networks

Bingchen Zhao*
University of Edinburgh
bingchen.zhao@ed.ac.uk

Yuling Gu*
Allen Institute for Artificial Intelligence
yulingg@allenai.org

Jessica Zosa Forde
Brown University
jessica_forde@brown.edu

Naomi Saphra
New York University
nsaphra@nyu.edu

Abstract

At NeurIPS, American and Chinese institutions cite papers from each other's regions substantially less than they cite endogamously. We build a citation graph to quantify this divide, compare it to European connectivity, and discuss the causes and consequences of the separation.

1 Introduction

In recent years, the machine learning research landscape has been reshaped by the growth of Chinese AI research. China now consistently stands as the second-largest country in terms of total publications at NeurIPS, after the United States. In 2020, papers from Chinese institutions represented 13.6% of all NeurIPS publications [6]. The following year, this share increased to 17.5%, representing a relative increase of 28.7% [13].[1]

In spite of China's position as a national AI powerhouse, collaborations between Chinese and American institutions are rarer than those between, say, American and Western European institutions [11]. Anecdotally, they also form distinct social groups at machine learning conferences: Chinese researchers often converse and take meals in separate groups from European and North American researchers. This divide goes beyond only social interactions; one prominent non-Chinese professor in an applied area of machine learning drew attention on Twitter when they advised students to avoid talks by Chinese authors, arguing that the presentations would be difficult to understand or of poor quality.[2] While it is true that many non-native anglophones find publicly speaking in English to be a challenge, avoiding talks by Chinese researchers may limit a conference attendee's exposure to new topics and ideas.

In this work, we measure this divide between researchers in China and researchers in the United States. Using NeurIPS citation data, we analyze the impact of work from US-based and China-based institutions. We find that Chinese institutions under-cite work from the US and Europe, and both American and European institutions under-cite work from China.

2 Citation networks

How severe is this segregation between American and Chinese research communities? In order to measure the effect quantitatively, we collate citation data with institutions labeled by region. We consider the divisions between American, European, and Chinese researchers. We find that European and American institutions are more closely linked through citations, and Chinese research is less linked to American and European research.

---

[1]These numbers can be regarded as an underestimate of total share, as American multinational companies have been categorized as American regardless of the office location of the authors.

[2]Out of a respect for privacy, we have kept this anecdote anonymous.

2.1 Methods

We compiled a citation graph by joining the citation data of the NeurIPS papers available from SemanticScholar[3] with institutional information about authors from AMiner[4]. We first collected all paper titles from NeurIPS 2012 to 2021 from the NeurIPS website. We map the paper titles to their Semantic Scholar paper IDs using the Semantic Scholar Academic Graph (S2AG) API [15]. We manually searched for the unmatched papers, finding all but one in the Semantic Scholar database. For each of these papers, we then used the S2AG API to identify the authors, as well as authors of papers in its references.

We then used AMiner to identify institutional information for each author. The 9460 NeurIPS papers have 135,941 authors in total, of which we found institutions for 83,515 authors (61%). There are 4038 papers that have no authors on AMiner, so they are removed from the data. We then automatically marked institutes that included the name of a country, along with common cities and regions in China. We supplemented these automatic annotations with existing [5] regional matchings and manually added 364 additional rules for regional matching. Finally, we removed major multinational corporate labs (e.g., Google, Meta, Microsoft, Tencent, Alibaba, or Huawei), as they typically do not include information about the local office of the authors. Out of the remaining 5422 papers, we removed papers that weren't in one of the specified regions (China, US, Europe), or that included collaborators in multiple regions, leaving 1792 papers remaining. We then computed the average number and proportion of citations between papers from each region (Figure 1).

2.2 Results

Based on these numbers, we see the extent to which American and Chinese papers fail to cite each other. While American papers make up 60% of the dataset, they make up only 34% of Chinese citations. American citations of Chinese papers show a more dramatic effect: while Chinese papers make up 34% of the dataset, they are only 9% of American citations. The effect is clear when we compare these numbers to American citations of European papers: even though our dataset of NeurIPS papers has six times as many Chinese papers as European papers, American institutions cite Chinese papers *less* often than European papers. (Note that the cited non-NeurIPS papers can be older, and thus are less likely to be Chinese than the citing papers.)

We observe that each region tends to self-cite more frequently than they are otherwise cited: 21% for China, 41% for the USA, and 14% for Europe. However, the separation between American and Chinese research communities is more pronounced than one might expect from typical regional preference. The American and European research communities share similar citation behavior, with few citations to Chinese papers, whereas Chinese institutions cite American and European papers less often than either other region does.
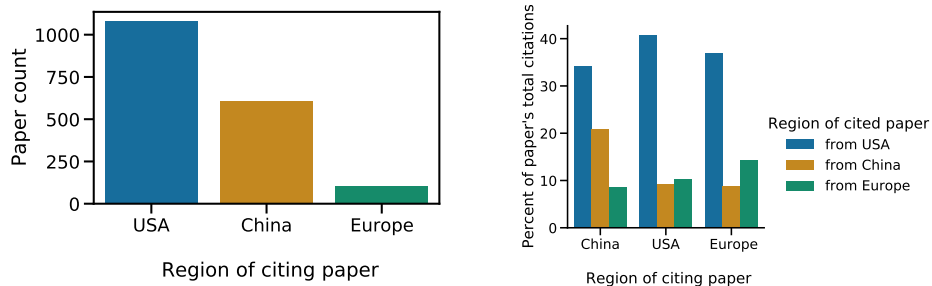

Figure 1: Papers from Europe and the US cite one another more often than they cite papers from China, and papers from China are least likely to cite institutions in Europe and the US.

2.3   Limitations

The strength of our claims is subject to the influence of a number of factors we have not accounted for in this work.

First, while we consider the work from any university situated in the United States to be American, US-based labs may still have close links to China, which may lead us to underestimate the division between AI research in the US and China. For example, there exist labs in the US that are made up largely or entirely of Chinese international students. Similarly, Chinese international students who return to their home country may bring international connections to their alma mater and other institutions. We have not measured the degree to which such graduates change their citation patters to focus on domestic papers, or whether they continue to widely cite American work. Additionally, our filter of multinational corporate labs may be incomplete. If these industry authors include companies with offices in both America and China, they may affect the reliability of our conclusions.

There is another limitation that affects the confidence of our conclusions: the number of papers that have been excluded on the basis of authors not appearing in AMiner, a Chinese company. This is likely to result in there being more Chinese papers in the dataset than if we had full information about every paper. Overall, we discarded 43% of papers due to a lack of author information, and this excluded set is likely to be a biased sample.

3   Consequences

While American and Chinese researchers publish in the same venues, they represent two parallel communities, each producing work with limited impact on the other. To some degree this divide can be attributed to an interest in different topics, as cultural norms inform research priorities. For example, multi-object tracking is an active area of research in China, with large scale benchmarks such as Wang et al. [16], which provides annotations of individual humans in each video. In North America, however, criticisms regarding of the misuse of biometric data [1] have led researchers to avoid related tasks and datasets. Similarly, the US tends to be heavily represented at fairness conferences such as FACCT, while representation from China remains limited [2].

However, even abstract topics or architectures that are popular in China may not catch on in other regions. PCANet [3], an image classification architecture from a Singaporean research lab, has 1200 citations which are largely from Chinese or East Asian institutions. Deep Forests [20], which were developed at Nanjing University, garnered over 600 citations that are mostly Chinese.

It is not only topics of research that are limited by a lack of exchange between regions. In recent years, the North American and European AI community has begun to engage in conversations and publishing research on the ethical considerations of AI. In accordance with these considerations, AI conferences have adopted systems for reviewers to flag ethical concerns and ask authors to write ethics statements or fill out checklists [1, 10]. Engagement with researchers from China on these topics, however, remains limited. An example of this disconnect is the Provisional Draft of the NeurIPS Code of Ethics [1]. At the time of initial publication, the authors' affiliations were with US universities, US-multinationals, and a university in Australia. None of the authors were based in Asia. This absence is despite the fact that observers have noted many similarities in ethics statements from Chinese based AI institutions [12, 9].

Yet despite these similarities in ethics statements from US and Chinese researchers, specific disagreements in research practice still exist. For example, Duke University stopped providing the Duke-MTMC dataset [8] for studying object tracking across multiple cameras, because the researchers did not obtain consent from the students they collected images from [14]. However, similar datasets like Market-1501 [18] from Tsinghua University are still being actively used [5, 4]. Furthermore, researchers in China have continued to use the Duke dataset even after its removal [17, 7, 19]. We can't know whether American researchers would have persuaded the authors with their arguments about the ethical hazards of such datasets, because they failed to engage the Chinese authors in these conversations.

The separation between these two communities has real impact on the individual researchers, the machine learning community as a whole, and potentially the societies affected by AI research. It is long past time that the AI community had a conversation about how to overcome this barrier.

References

[1] S. Bengio, A. Beygelzimer, K. Crawford, J. Fromer, I. Gabriel, A. Levendowski, , I. D. Raji, and M. Ranzato. Provisional draft of the neurips code of ethics, 2022.

[2] D. Card. Report from facct 2022, Jul 2022.

[3] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma. Pcanet: A simple deep learning baseline for image classification? *IEEE Transactions on Image Processing*, 24:5017–5032, 2015.

[4] D. Fu, D. Chen, H. Yang, J. Bao, L. Yuan, L. Zhang, H. Li, F. Wen, and D. Chen. Large-scale pre-training for person re-identification with noisy labels. In *The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)*, 2022.

[5] S. He, H. Luo, P. Wang, F. Wang, H. Li, and W. Jiang. Transreid: Transformer-based object re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 15013–15022, October 2021.

[6] S. Ivanov. NeurIPS 2020. Comprehensive analysis of authors, organizations, and countries., Oct. 2020.

[7] M. Li, X. Zhu, and S. Gong. Unsupervised tracklet person re-identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[8] E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In *European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking*, 2016.

[9] H. Roberts, J. Cowls, J. Morley, M. Taddeo, V. Wang, and L. Floridi. The chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & SOCIETY*, 36(1):59–77, Mar 2021.

[10] A. Rogers, T. Baldwin, and K. Leins. 'just what do you think you're doing, dave?' a checklist for responsible data use in NLP. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 4821–4833, Punta Cana, Dominican Republic, Nov. 2021. Association for Computational Linguistics.

[11] M. Rungta, J. Singh, S. M. Mohammad, and D. Yang. Geographic Citation Gaps in NLP Research, Oct. 2022. arXiv:2210.14424 [cs].

[12] M. Sheehan. Beijing's approach to trustworthy ai isn't so dissimilar from the world's, Aug 2021.

[13] SYNCED. NeurIPS 2021 Announces Its 6 Outstanding Paper Awards, 2 Datasets and Benchmarks Track Best Paper Awards, and the Test of Time Award | Synced, 2021.

[14] C. Tomasi. Letter: Video analysis research at Duke, 2019.

[15] A. D. Wade. The semantic scholar academic graph (s2ag). In *Companion Proceedings of the Web Conference 2022*, WWW '22, page 739, New York, NY, USA, 2022. Association for Computing Machinery.

[16] X. Wang, X. Zhang, Y. Zhu, Y. Guo, X. Yuan, L. Xiang, Z. Wang, G. Ding, D. Brady, Q. Dai, et al. Panda: A gigapixel-level human-centric video dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3268–3278, 2020.

[17] Y. Wu, Y. Lin, X. Dong, Y. Yan, W. Ouyang, and Y. Yang. Exploit the unknown gradually: One-shot video-based person re-identification by stepwise learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[18] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian. Scalable person re-identification: A benchmark. In *Computer Vision, IEEE International Conference on*, 2015.

[19] Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017.

[20] Z.-H. Zhou and J. Feng. Deep forest: Towards an alternative to deep neural networks. *ArXiv*, abs/1702.08835, 2017.