Disparate Citation Patterns Between Chinese and American Research Communities at a Unified Venue

Abstract

At NeurIPS, there is a tendency for American and Chinese institutions to cite papers from within their own regions substantially more often than they cite papers from the other region. To measure this divide, we construct a citation graph, compare it to European connectivity, and discuss both the causes and consequences of this separation.

1 Introduction

In recent years, the machine learning research community has been transformed by the rise of Chinese AI research. China is now consistently the second-largest contributor of publications at NeurIPS, following the United States. In 2020, 13.6% of all NeurIPS publications came from Chinese institutions. The next year, this increased to 17.5%, a relative increase of 28.7%.

Despite China's position as a leader in AI research, collaborations between Chinese and American institutions are less common than collaborations between American and Western European institutions. Anecdotally, researchers from these regions often form distinct social groups at machine learning conferences. This separation is not limited to just social interactions. A prominent professor in an applied area of machine learning publicly advised students to avoid talks by Chinese authors, arguing that their presentations would be difficult to understand or of poor quality. Although many non-native English speakers find it a challenge to speak in public, avoiding talks by Chinese researchers may limit a conference attendee's exposure to new topics and ideas.

This study measures the separation between researchers in China and the United States. We use NeurIPS citation data to analyze the impact of work from US-based and China-based institutions, and find that Chinese institutions under-cite work from the US and Europe, and that both American and European institutions under-cite work from China.

2 Citation Networks

2.1 Methods

To quantify the divide between the regions, we compiled a citation graph using NeurIPS paper citation data from SemanticScholar and institutional information about authors from AMiner. We first collected all paper titles from NeurIPS from 2012 to 2021 from the NeurIPS website. Using the Semantic Scholar Academic Graph (S2AG) API, we then mapped paper titles to their Semantic Scholar paper IDs. For unmatched papers we manually searched, finding all but one in the Semantic Scholar database. We then used the S2AG API to identify the authors of each paper as well as the authors of papers referenced by these papers.

We 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 (61%). The 4038 papers lacking author information were excluded from the dataset. We then automatically identified institutes that included a country name, along with common cities and regions in China. We augmented these automatic annotations with existing regional matchings and added 364 additional rules. Finally, we

.

removed major multinational corporate labs (e.g., Google, Meta, Microsoft, Tencent, Alibaba, or Huawei). Of the remaining 5422 papers, we removed papers that were not from China, the US, or Europe, or included collaborators in multiple regions, leaving 1792 papers. Finally, we computed the average number and proportion of citations between papers from each region, shown in Figure 1.

2.2 Results

We observed the extent to which American and Chinese papers fail to cite each other. While American papers constitute 60% of our dataset, they only account for 34% of citations made by Chinese papers. American citations of Chinese papers are even more striking: while Chinese papers account for 34% of our dataset, they are only cited in 9% of American references. This is more profound when comparing these values to American citations of European papers: even though the dataset has six times more Chinese than European papers, American institutions cite Chinese papers less than European papers.

We also observe that each region tends to cite its own papers more often: 21% for China, 41% for the USA, and 14% for Europe. The division between American and Chinese research communities is much more pronounced than one would expect based on typical regional preferences. While American and European research communities show similar citation behavior, Chinese institutions cite American and European papers less than other regions.

	USA	China	Europe
USA	41	9	12
China	34	21	6
Europe	15	9	14

Table 1: Proportion of papers from given regions citing other regions or endogenously. Values are in percentage.

3 Limitations

The conclusions we make in this paper are dependent on a few key choices we made during our data selection process. First, while we consider institutions in the US as American, many US labs have close ties to China, potentially underestimating the true divide. Some US labs are largely or entirely made up of Chinese international students. Additionally, international students returning to their home country may bring international connections, and we did not measure if their citation patterns focus more on domestic papers or if they continue to cite American work. In addition, our filtering of multinational corporate labs may be incomplete which could also affect our results.

Second, a number of papers were excluded from our analysis due to missing author information on AMiner, which is a Chinese platform. This may have resulted in the number of Chinese papers in the dataset being more than what there actually is. We discarded 43

4 Consequences

Though American and Chinese researchers publish in the same venues, they represent two parallel communities. To some degree, this can be attributed to different research interests due to cultural norms influencing research priorities. For instance, multi-object tracking is an active area of research in China, with many large scale benchmarks. However, due to concerns surrounding privacy and misuse, many North American researchers tend to avoid related topics. In general, the US tends to be heavily represented at fairness conferences, while representation from China is limited.

Not only research topics are limited by this lack of exchange, but even abstract topics and architectures that are popular in China are often not adopted in other regions. For example, PCANet, a popular image classification architecture has most of its 1200 citations from Chinese or East Asian institutions. Similarly, the Deep Forest model has garnered most of its 600 citations from Chinese researchers.

Recently, the North American and European AI communities have increasingly engaged in conversations regarding the ethical considerations of AI and have adopted review systems for ethical concerns and required authors to include ethics statements. However, there has been limited engagement with researchers from China regarding these topics, and ethics statements for Chinese-based AI institutions are similar to western ones. Despite such statements, specific disagreements regarding research practices still exist. For instance, while Duke University stopped providing the Duke-MTMC dataset, due to the ethical issues with the collection process, similar datasets from Chinese institutions continue to be actively used. This highlights the need for a discussion on the topic of the ethical dimensions of AI research between different communities.

The separation between the research communities has an impact on both researchers and societies as a whole. It is crucial that the AI community initiates a discussion to overcome this barrier.

Appendix A: Proof of Lemma 3

Appendix B: Sub-Gaussian Covering Numbers for ReLU Networks

C: Table 2

- **Name**: name of the attack
- **Threat Model**: the threat model used in the attack
  - 'aux' auxiliary information,
  - black - black box,
  - white - white box
- **Baseline**: method used to determine the performance of the attack.
  - 'A' - absolute, the proportion of correctly identified data points or some other metric of attack success
  - 'M' - mathematical privacy metrics (e.g., k-anonymity, DP)
  - 'R' - random
  - 'C' - a control baseline which is a subset of the real data that was not used for the training data
  - 'SL' - metrics from supervised learning such as precision and recall
- **Attack estimator**: The method used to estimate the success of an attack
  - 'IT' - information theory
  - 'NN' - nearest neighbor
  - 'ML' - machine learning
- **Attack Technique**: The technique of the attack.
  - 'VRD' - vulnerable record discovery through searching or sampling
  - 'SM' - shadow modeling
  - 'MIA' - membership inference attack
- **Attack type (WP29)** attack type based on WP29 specification.
  - 'S' - singling out
  - 'L' - linkage
  - 'I' - inference.

Model	Dataset	Clean	Evasion	Poisoning	Symbiotic
GCN	CiteSeer	0.68 ± 0.01	0.41 ± 0.01	0.4 ± 0.01	**0.38 ± 0.01**
	CiteSeer-J	0.68 ± 0.01	0.4 ± 0.01	0.4 ± 0.02	**0.38 ± 0.01**
	Cora	0.78 ± 0.01	0.37 ± 0.02	0.46 ± 0.02	**0.35 ± 0.01**
	Cora-J	0.74 ± 0.01	0.36 ± 0.01	0.43 ± 0.02	**0.36 ± 0.02**
	PubMed	0.78 ± 0.01	0.05 ± 0.01	0.12 ± 0.02	**0.03 ± 0.01**
	PubMed-J	0.77 ± 0.01	0.04 ± 0.01	0.11 ± 0.01	**0.02 ± 0.0**
GAT	CiteSeer	0.62 ± 0.02	0.3 ± 0.03	0.41 ± 0.02	**0.38 ± 0.02**
	CiteSeer-J	0.64 ± 0.01	0.3 ± 0.03	0.41 ± 0.03	**0.3 ± 0.03**
	Cora	0.69 ± 0.02	0.29 ± 0.02	0.48 ± 0.03	**0.32 ± 0.02**
	Cora-J	0.67 ± 0.01	0.28 ± 0.02	0.45 ± 0.02	**0.3 ± 0.03**
	PubMed	0.73 ± 0.01	0.24 ± 0.02	0.41 ± 0.01	**0.2 ± 0.03**
	PubMed-J	0.74 ± 0.01	0.27 ± 0.04	0.38 ± 0.04	**0.19 ± 0.02**
APPNP	CiteSeer	0.69 ± 0.01	0.47 ± 0.01	0.56 ± 0.01	**0.47 ± 0.01**
	CiteSeer-J	0.68 ± 0.01	0.45 ± 0.02	0.52 ± 0.02	**0.45 ± 0.02**
	Cora	0.82 ± 0.02	0.54 ± 0.02	0.64 ± 0.02	**0.51 ± 0.04**
	Cora-J	0.82 ± 0.01	0.57 ± 0.01	0.67 ± 0.01	**0.54 ± 0.01**
	PubMed	0.79 ± 0.0	0.09 ± 0.02	0.21 ± 0.02	**0.09 ± 0.01**
	PubMed-J	0.77 ± 0.01	0.1 ± 0.02	0.19 ± 0.03	**0.1 ± 0.02**
GPRGNN	CiteSeer	0.66 ± 0.01	0.34 ± 0.01	0.44 ± 0.02	**0.33 ± 0.01**
	CiteSeer-J	0.65 ± 0.01	0.35 ± 0.01	0.44 ± 0.01	**0.35 ± 0.01**
	Cora	0.82 ± 0.01	0.46 ± 0.01	0.53 ± 0.01	**0.4 ± 0.01**
	Cora-J	0.79 ± 0.01	0.42 ± 0.01	0.54 ± 0.01	**0.4 ± 0.01**
	PubMed	0.78 ± 0.01	0.08 ± 0.02	0.28 ± 0.03	**0.08 ± 0.02**
	PubMed-J	0.78 ± 0.01	0.16 ± 0.05	0.38 ± 0.04	**0.15 ± 0.04**
RGCN	CiteSeer	0.63 ± 0.01	0.39 ± 0.01	0.59 ± 0.02	**0.47 ± 0.01**
	Cora	0.74 ± 0.02	0.44 ± 0.01	0.74 ± 0.01	**0.52 ± 0.02**
	PubMed	0.77 ± 0.01	0.43 ± 0.01	0.42 ± 0.04	**0.15 ± 0.03**

Table 2: Perturbed accuracies ($\pm$ standard error) of the joint and sequential attacks under the symbiotic threat model with a 5% global budget. The -J suffix indicates the graph has been pre-processed with Jaccard purification.

Model	Dataset	Clean	Sequential	Joint
GCN	CiteSeer	0.68 ± 0.01	0.41 ± 0.01	$\mathbf{0.38 \pm 0.01}$
	CiteSeer-J	0.68 ± 0.01	0.4 ± 0.01	$\mathbf{0.38 \pm 0.01}$
	Cora	0.78 ± 0.01	0.37 ± 0.02	$\mathbf{0.35 \pm 0.01}$
	Cora-J	0.74 ± 0.01	0.36 ± 0.01	$\mathbf{0.36 \pm 0.02}$
	PubMed	0.78 ± 0.01	0.05 ± 0.01	$\mathbf{0.03 \pm 0.01}$
	PubMed-J	0.77 ± 0.01	0.04 ± 0.01	$\mathbf{0.02 \pm 0.0}$
GAT	CiteSeer	0.62 ± 0.02	$\mathbf{0.3 \pm 0.03}$	0.38 ± 0.02
	CiteSeer-J	0.64 ± 0.01	$\mathbf{0.3 \pm 0.03}$	0.36 ± 0.02
	Cora	0.69 ± 0.02	$\mathbf{0.29 \pm 0.02}$	0.32 ± 0.02
	Cora-J	0.67 ± 0.01	$\mathbf{0.28 \pm 0.02}$	0.3 ± 0.03
	PubMed	0.73 ± 0.01	0.24 ± 0.02	$\mathbf{0.2 \pm 0.03}$
	PubMed-J	0.74 ± 0.01	0.27 ± 0.04	$\mathbf{0.19 \pm 0.02}$
APPNP	CiteSeer	0.69 ± 0.01	$\mathbf{0.47 \pm 0.01}$	0.48 ± 0.01
	CiteSeer-J	0.68 ± 0.01	$\mathbf{0.45 \pm 0.02}$	0.45 ± 0.02
	Cora	0.82 ± 0.02	0.54 ± 0.02	$\mathbf{0.51 \pm 0.04}$
	Cora-J	0.82 ± 0.01	0.57 ± 0.01	$\mathbf{0.54 \pm 0.01}$
	PubMed	0.79 ± 0.0	$\mathbf{0.09 \pm 0.02}$	$\mathbf{0.09 \pm 0.01}$
	PubMed-J	0.77 ± 0.01	$\mathbf{0.1 \pm 0.02}$	0.12 ± 0.02
GPRGNN	CiteSeer	0.66 ± 0.01	0.34 ± 0.01	$\mathbf{0.33 \pm 0.01}$
	CiteSeer-J	0.65 ± 0.01	0.35 ± 0.01	$\mathbf{0.35 \pm 0.01}$
	Cora	0.82 ± 0.01	0.41 ± 0.01	$\mathbf{0.4 \pm 0.01}$
	Cora-J	0.79 ± 0.01	0.42 ± 0.01	$\mathbf{0.4 \pm 0.01}$
	PubMed	0.78 ± 0.01	0.08 ± 0.02	$\mathbf{0.11 \pm 0.03}$
	PubMed-J	0.78 ± 0.01	0.16 ± 0.05	$\mathbf{0.15 \pm 0.04}$
RGCN	CiteSeer	0.63 ± 0.01	0.47 ± 0.01	$\mathbf{0.47 \pm 0.01}$
	Cora	0.74 ± 0.02	0.56 ± 0.01	$\mathbf{0.52 \pm 0.02}$
	PubMed	0.77 ± 0.01	0.28 ± 0.04	$\mathbf{0.15 \pm 0.03}$