Speech Research Trend Visualization

Xingyu Shen* Yueqian Lin* Zhixian Zhang* Xin Tong†

Duke Kunshan University

ABSTRACT

In our project, we employ visualization techniques to dissect the evolution and interdisciplinary impacts of deep learning in speech-related research. Leveraging data from academic papers, we illuminate author collaborations, cross-disciplinary influences, and emerging intelligent speech trends. Moreover, our examination of citation-publication correlations and publication timing patterns offers unique insights into academic publishing practices. The key of this study is to use intuitive visual representation to decode the complexities and patterns in the ever-evolving intelligent speech domain. The findings are set to bolster understanding and guide future explorations in this influential field. The demo[1] is available online.

Index Terms: Visualization—Visualization techniques—Human-centered computing—Network

1 INTRODUCTION

Research in speech-related areas, particularly with the combination of deep learning, has developed rapidly recently. However, the deluge of research papers published in this domain can be overwhelming for researchers. Traditional methods of literature review, such as using Google Scholar [1] and Arxiv [2], although reliable, can be laborious and time-consuming since researchers usually have to read dozens or even more papers. There are also some existing visualization tools trying to visualize trends. For example, [3] visualized the topic and citation of papers through a specially designed network graph. However, these tools usually focus only on a few specific aspects of papers, which inhibits researchers from acquiring a comprehensive understanding of the field's recent trends, collaborations, and intersections with other disciplines.

To tackle these problems, we propose a novel visualization tool that transforms the overwhelming pile of text information into an intelligible and interactive visual representation. Our tool aims to provide a holistic view of the field, fostering a deeper understanding of its various aspects, such as author collaborations, interdisciplinary studies, temporal trends, and publication practices.

Specifically, our visualization tool intends to highlight the connections between authors and research domains, identify trending topics within speech-related papers, analyze the correlation between citation and publication numbers, and uncover any prevailing trends in publication timing. This approach allows for a more intuitive and efficient exploration of complex research trends and patterns, facilitating the assimilation of diverse research aspects from a macroscopic perspective.

2 METHOD AND VISUALIZATION RESULTS

2.1 Data Collection, Cleaning, and Analysis

Our data were obtained from Semantic Scholars[2], a public tool for scientific exploration and discovery that provides access to over 200 million academic papers from a variety of sources. We utilized the Python package for Semantic Scholars, query keyword "speech" to gather 10,000 papers for each year from 2020 until now, according to the first 10,000 results given by the search, retrieving a total dataset of over 32,000 entries.

2.2 Results

We developed seven interactive visualizations that highlight different aspects of our research questions, providing clear representations of our data. Below, we explain each visualization and its relationship to our research questions.

2.2.1 Overall Co-Author Network

The overall co-author network (Figure 1) presents the network of authors and co-authors in the speech research domain over the past three years. The network core displays brighter and more closely interconnected authors, while the periphery, though less dense, still contains smaller clusters. Users can zoom in or out to examine specific regions and perform dynamic queries by entering an author's name. Notable authors in the research area include Junyu Li, Shinji Watanabe, Tara N. Sainath, Yu Zhang, Zhuo Chen, among others.

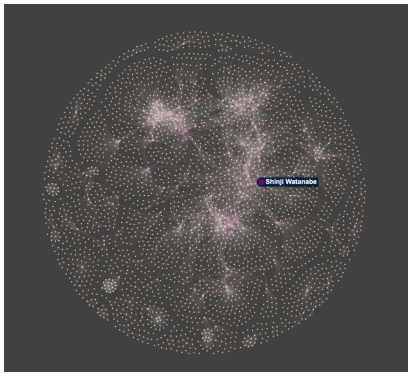

Figure 1: Overall Co-Author Network

2.2.2 Individual Co-Author Network

This visualization focuses on the individual networks of specific authors. For instance, as shown in Figure 2, Professor Watanabe has numerous collaborators, many of whom have multiple appearances, suggesting the existence of research groups under the same professor's leadership.

Figure 2: Individual Co-Author Network

2.2.3 Interactive Word Cloud in the Shape of a Microphone

The interactive word cloud, shaped like a microphone, displays the most frequently occurring words in the abstracts of the papers (Figure 3, left). The size of each word is proportional to its frequency, and the word count is displayed when hovering the mouse over each word. As expected, most of these words are closely associated with speech research, such as "speech," "recognition," and "language."

*Equal contribution

†Corresponding author: xin.tong@dukekunshan.edu.cn

[1]https://vis.yueqianlin.com

[2]https://www.semanticscholar.org/product/api


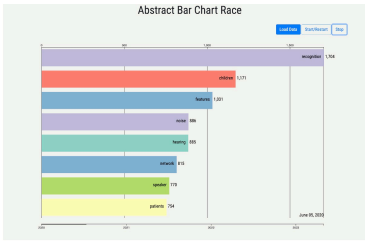

Figure 3: Left: Interactive Word Cloud in the Shape of a Microphone. Right: Bar Chart Race for Word Occurrences in Abstracts

2.2.4 Bar Chart Race for Word Occurrences in Abstracts

This interactive chart (Figure 3, right) shows how the frequency of the eight most common words has evolved. Apart from expected words like "recognition" and "features," we also observed frequent occurrences of "children" and "patient" in the abstract part of papers related to speech.

2.2.5 Chord Diagram for Interdisciplinary Correlations

The chord diagram (Figure 4) illustrates the relationships between different study fields, with the color of the ribbon indicating the major field in interdisciplinary studies. In 2020, computer science and medicine predominantly acted as major fields, while engineering, psychology, biology, mathematics, and physics primarily served as minor fields.

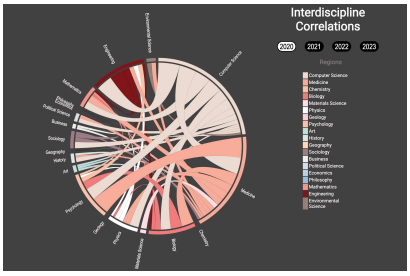

Figure 4: Chord Diagram for Interdisciplinary Correlations

2.2.6 Heatmap of Paper Publication Dates

The heatmap (Figure 5) illustrates the popularity of paper publication over the past three years, with darker blocks representing a higher number of published papers. Notably, the first day of every month shows a peak in publications, possibly due to some monthly journals scheduling their publications for the beginning of each month.

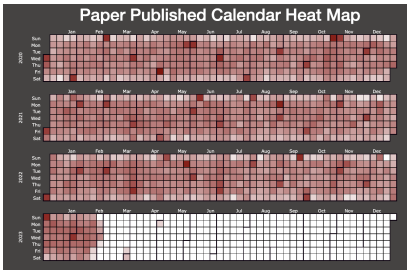

Figure 5: Heatmap of Paper Publication Dates

2.2.7 Quadrant Chart for Publication and Citation Count

The quadrant chart (Figure 6) shows the relationship between publication count and citation count. Authors are categorized into four groups based on the average publication counts and citation counts. Interestingly, most data points appear to be distributed linearly, aligning with our initial assumptions.

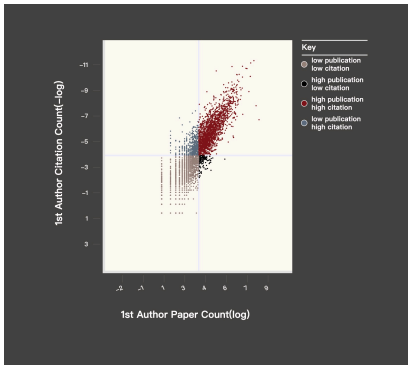

Figure 6: Quadrant Chart for Publication Count and Citation Count

2.3 Interactive Website Building

Our visualizations have been conveniently made accessible by publishing the entire website on GitHub Pages, utilizing HTML, CSS, and JS files. To simplify the process of merging and comparing different CSS files used on each page, we have embedded all individual visualizations in the portfolio using the ¡iframe¿ element. The website is accessible via vis.yueqianlin.com.

3 Discussion and Conclusion

Our work has created insightful visualizations of the development and trends within the domain of speech-related research papers. By focusing on author collaborations, interdisciplinary relationships, emerging topics, and publication metrics, we extracted meaningful patterns and insights that would not be readily apparent from perusing individual papers.

The interactive user interface of our visualizations facilitated easy exploration and comprehension of complex data. The co-authorship networks revealed clear evidence of research communities and collaboration patterns. The word frequency visualizations allowed for identifying key topics and their popularity over time. The citation-publication correlation graph and the publication heatmap provided valuable insights into publication patterns and trends.

Notably, our visualizations demonstrated that speech recognition research remains a prominent focus within this community, with a steadily increasing contribution from computer science disciplines. The publication patterns indicated a tendency for authors to publish their work at the beginning and end of the month.

While our work has achieved its primary objectives, there are areas for potential improvements and future work. Refining data processing techniques to handle name duplication, enhancing the stop-word list for more accurate word frequency analysis, and expanding the attributes of the quadrant chart could further enhance the usefulness of our visualizations. Additionally, investigating outliers in the quadrant chart could offer even more in-depth insights.

In conclusion, our study represents a significant step toward better understanding the landscape of speech-related research papers through interactive visualizations. We anticipate that these tools can be extended and adapted for other research domains, enabling researchers to explore and comprehend the vast sea of academic literature easily. The potential impact of our findings lies in providing researchers with valuable insights to inform their work and contribute to advancements in the field.

References

[1] Google scholar. `https://scholar.google.com/`. Accessed: Jun 26, 2023.

[2] arxiv. `https://arxiv.org/`. Accessed: June 26, 2023.

[3] Rina Nakazawa, Takayuki Itoh, and Takafumi Saito. A visualization of research papers based on the topics and citation network. pages 283–289, 07 2015.