ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Paper ID #35832

Thematic and Authorship Analysis of ASEE Gulf-Southwest Conference Papers

Anna Huang, Texas A&M University

Anna Huang is a sophomore at Texas A&M University studying computer science. Currently, she is an undergraduate researcher at the Institute for Engineering Education and Innovation. Huang is interested in artificial intelligence, data science, and machine learning and is passionate about advancing engineering education practices.

Amanda Jane Latham, Texas A&M University
Erika Yao
Dr. Tracy Anne Hammond, Texas A&M University

Dr. Hammond is Director of the Texas A&M University Institute for Engineering Education & Innovation and also the chair of the Engineering Education Faculty. She is also Director of the Sketch Recognition Lab and Professor in the Department of Computer Science & Engineering. She is a member of the Center for Population and Aging, the Center for Remote Health Technologies & Systems as well as the Institute for Data Science. Hammond is a PI for over 13 million in funded research, from NSF, DARPA, Google, Microsoft, and others. Hammond holds a Ph.D. in Computer Science and FTO (Finance Technology Option) from the Massachusetts Institute of Technology, and four degrees from Columbia University: an M.S in Anthropology, an M.S. in Computer Science, a B.A. in Mathematics, and a B.S. in Applied Mathematics and Physics. Hammond advised 17 UG theses, 29 MS theses, and 10 Ph.D. dissertations. Hammond is the 2020 recipient of the TEES Faculty Fellows Award and the 2011 recipient of the Charles H. Barclay, Jr. '45 Faculty Fellow Award. Hammond has been featured on the Discovery Channel and other news sources. Hammond is dedicated to diversity and equity, which is reflected in her publications, research, teaching, service, and mentoring. More at http://srl.tamu.edu and http://ieei.tamu.edu.

Session XXXX

Thematic and Authorship Analysis of ASEE Gulf-Southwest Conference Papers

Dr. Tracy Hammond, Erika Yao, Anna Huang, Amanda Latham
Computer Science Department
Texas A&M University

Abstract

Throughout the history of the American Society of Engineering Education Gulf-Southwest (ASEE GSW) annual conference, hundreds of papers have been submitted and presented over a wide range of topics in the field of Engineering Education. By systematically scouring the ASEE Papers on Engineering Education Repository, the ASEE-GSW Index of Past Proceedings, and Google Scholar's document database, papers submitted to the annual conference from 2009 to 2021 were collected and compiled into a highly detailed and comprehensive database. This process enabled the application of data analysis techniques to identify trends that reside in those papers, e.g., statistical analysis of publication frequency and identifying common research topics after data preprocessing. We focused on discovering connections between various institutions and their submitted papers, including notable variances in the number of papers, number of authors per paper, and number of early career researchers submitting papers to the conference. With the collected data, overall paper submission counts by affiliation were produced, and interesting connections could then be drawn between the levels of enrollment at a given institution versus the number of papers submitted by that institution. This analysis showed that the top submitting affiliations tend to be in Texas and have varying levels of engineering enrollment. In fact, all but one of the top eight do not fall within the Top 50 in engineering enrollment, and the one that does have the highest engineering enrollment in the country. This analysis also revealed a core community of researchers from several different schools, often those with high paper contribution counts, who continually submit papers to the conference across numerous years. Following authorship analysis, paper titles were processed to remove stopwords and allow for the visualization of word frequencies, giving a surface level exploration of the general topics and methodologies covered by all papers. Interesting words that came from this analysis include "design", "course", and "system", indicating a focus on curriculum development and improvement. These insights can aid future authors in the presentation of their work for the ASEE-GSW community and will improve understanding of the current compositions and interests of the Engineering Education research community as a whole.

Introduction

Over the past thirteen years, i.e., from 2009 to 2021, 618 papers from 135 different institutions have been submitted to the ASEE-GSW Annual Conference. These papers cover a wide variety of topics aimed towards furthering education in engineering and engineering technology. Our team has been a long-time contributor to the ASEE-GSW Annual Conference, so we were interested in finding whether any identifiable trends within the works submitted exist over the many years of the conference's history. Given the data we would be collecting in order to complete that analysis, it

became clear that this research would be able to highlight patterns in both the topics, as initially intended, and the authors of the papers themselves.

Methodology and Results

Historically, papers submitted to the conference have been stored in various repositories, and their method of storage and documentation has slowly evolved over the years. In order to facilitate our research and analysis, the information available regarding these papers needed to be standardized. We gathered available details about previous papers from existing ASEE-GSW document repositories, including title, author, affiliation, and PDF copies when possible. Everything we were able to gather was then organized into several Google Sheets, which was then processed using Python code compiled with Google Colab. The usage of Google Colab not only allowed us to all work in a single platform, but also allowed us to quickly update the results of our various analyses any time a new paper was added to one of our Google Sheet documents.

Several noticeable gaps popped in the data we collected, most notable where the repository or documentation switched formats. This included years 2011, 2014, and 2015, where going through the existing ASEE-GSW repositories did not suffice, and we had to search for what papers were available using other tools like Google Scholar. This often resulted in fewer data points – mainly for keywords, but occasionally in author lists, affiliation, and other fields as well – as many of the results from Google Scholar failed to yield a PDF copy or physical copies of the original paper.

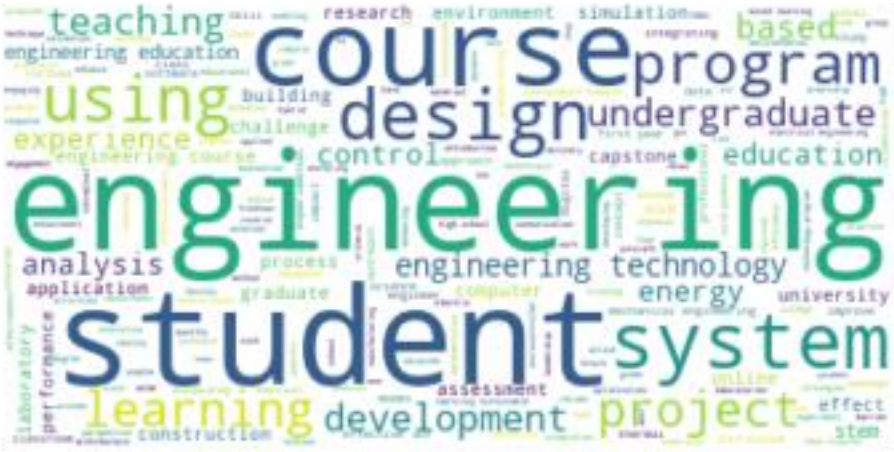
	Author	**Affiliation**
1	Dr. Amir Karimi	University of Texas at San Antonio (3rd)
2	Dr. Randall D. Manteufel	University of Texas at San Antonio (3rd)
3	Dr. Mequanint Moges	University of Houston (1st)
4	Dr. Tariq Khraishi	University of New Mexico (4th)
5	Dr. Kenneth W. Van Treuren	Baylor University (6th)
5	Dr. Ronald E. Barr	University of Texas at Austin (9th)
5	Dr. Jerry K. Keska	University of Louisiana at Lafayette (12th)
8	Dr. John Attia	Prairie View A&M University (7th)
8	Dr. Dani Fadda	University of Texas at Dallas (10th)
10	Dr. Farrokh Attarzadeh	University of Houston (1st)
10	Dr. Walter W. Buchanan	Texas A&M University (2nd)

Table 1: Top 11 ASEE-GSW authors and their affiliation, as well as their affiliation's ranking in terms of number of papers within the collected data (2009-2021)

Having compiled information across individual years, we appended each of our spreadsheets into a single Pandas dataframe. From there, we could easily parse for patterns in the list of authors and affiliations between papers. Unsurprisingly, the authors who had submitted the highest quantities of papers to the ASEE-GSW Annual Conference had done so over most, if not all, of the years of our analysis. These frequent contributors are generally composed of a couple of representatives from each of the affiliations with the highest overall paper counts. Table 1 displays the top eleven

members of ASEE-GSW's core community of contributors, as well as their affiliations and how their affiliations compare in overall submissions.

In the course of our analyses regarding the affiliations of various papers, we investigated whether there was any correlation between enrollment at these universities and the number of papers that were submitted. For example, in 2020, a study revealed Texas A&M University to be the institution with the highest engineering enrollment, closely followed by Arizona State University (Roy, 2021). However, this ranking is not very well reflected in the affiliation data that we collected. Over the thirteen years represented in the data we collected, the University of Houston, which is ranked outside the top fifty institutions in engineering enrollment, had the highest number of authors that submitted papers to the conference. Meanwhile, Texas A&M University was only second in this ranking of paper submissions. The third ranked institution from our collected data, University of Texas at San Antonio, was also ranked outside of the top fifty institutions in terms of engineering enrollment. The next highest ranked institutions in our data to be ranked in the top fifty are only ranked 9th and 10th in paper submissions, despite having the 20th highest and 16th highest engineering enrollments nationally, respectively.



We were also able to utilize the dataframe to find commonalities in the titles of papers. Using the WordCloud library of Python, Figure 1 illustrates the frequencies of various papers in titles across all the papers we managed to collect. Aside from words like "engineering" and "student", which are to be expected in papers about engineering education, some of the most notable recurrences are words like "design", "course", and "system", which suggest continuous development of various teaching methods year after year. Specifics on these can be identified by some other repeated wordings. For example, the prevalence of "undergraduate" over "graduate" indicates a focus towards the former in these educational developments.

Conclusion and Future Goals

Through our analysis of the hundreds of papers that have been submitted to the ASEE-GSW Annual Conference, we identified a few interesting trends. Firstly, a core community of frequent

contributors to the conference was identified, hailing from several different schools in the region. Some unexpected results came out of our affiliation analysis, as enrollment does not necessarily correlate to the quantity of papers submitted. Finally, by bringing together all the titles of the papers we collected, we are able to see the various keywords that repeatedly appear across the numerous years of our analysis.

As mentioned previously, there were several years where the list of papers we managed to collect was likely incomplete. Since we were limited to public resources to attempt to find these papers, we were also unable to locate many PDF copies of papers. It would make for a much more thorough analysis if we were able to contact some of the original authors, especially some members of the core community revealed through our analysis, and obtain more information and copies of older papers.

Due to time and resource constraints, there were also a number of forms of analysis that we initially planned to complete but were unable to follow through with. In our author analysis, we originally planned on doing additional analysis through web scraping of Google Scholar's h-index data. Especially for the core community of authors we identified from the base analysis, this would be an interesting statistic to look at. We had intended to calculate the TF-IDF of the extracted keywords from the PDF copies of papers we collected to discover any interesting trends in the papers but this analysis is left for future work. Keywords would likely be a clearer summary of the overall content of a paper than just the titles and correspondingly would allow our analysis to produce results that would be more representative of all the papers that were submitted.

By finding a list of most frequent authors and their affiliations, we highlight trends in research topics that contributing universities focus on. Through this analysis, universities may utilize trends and findings to their advantage in improving engineering education research as a whole.

References

1. Roy, J. (2021). (rep.). *Engineering & Engineering Technology By The Numbers* (pp. 20–24). Washington, DC: American Society for Engineering Education.

2. Index of Past Proceedings. (2010, August 31). Retrieved 2021, from http://www.aseegsw.com/past%20Proceedings/.

3. ASEE PEER Document Repository. (n.d.). Retrieved 2021, from https://peer.asee.org/.