UC Santa Cruz
UC Santa Cruz Previously Published Works

Title

Invited: Mapping Two Decades of Innovation: Lessons from 25 Years of ISPD Research

Permalink

https://escholarship.org/uc/item/8kx3k212

Authors

Rahmaniani, Gona
Guthaus, Matthew
Behjat, Laleh

Publication Date

2025-03-16

DOI

10.1145/3698364.3709130

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Invited: Mapping Two Decades of Innovation: Lessons from 25 Years of ISPD Research

Gona Rahmaniani
gona.rahmaniani@ucalgary.ca
University of Calgary
Calgary, Alberta, Canada

Matthew Guthaus
mrg@ucsc.edu
University of California Santa Cruz
Santa Cruz, CA, USA

Laleh Behjat
laleh@ucalgary.ca
University of Calgary
Calgary, Alberta, Canada

Abstract

The design automation research community has driven the evolution of integrated circuits from a handful of transistors in the 1960s to billions today. The International Symposium on Physical Design (ISPD) has been instrumental in tackling challenges like scaling complexities, hardware security, and the exponential growth in transistor counts. This study conducts a comprehensive bibliometric analysis of ISPD publications using Natural Language Processing, machine learning, and network analysis. It explores research themes, collaboration dynamics, and global contributions through citation networks, co-authorship graphs, geographical and spatial mapping, and topic modeling. Key areas of focus include Physical Design Optimization, Power Efficiency, and Emerging Technologies, with prominent topics such as placement, routing, clock skew, lithography, machine learning, and hardware security. The analysis highlights the evolution of foundational techniques like placement and routing while identifying emerging trends such as AI-driven design automation. These insights provide a roadmap for sustaining innovation in physical design over the next 25 years.

CCS Concepts

• **Hardware → Best practices for EDA**; **Methodologies for EDA**.

Keywords

Bibliometrics, Design Automation, Physical Design, Natural Language Processing, Community

ACM Reference Format:
Gona Rahmaniani, Matthew Guthaus, and Laleh Behjat. 2025. Invited: Mapping Two Decades of Innovation: Lessons from 25 Years of ISPD Research. In *Proceedings of the 2025 International Symposium on Physical Design (ISPD '25), March 16–19, 2025, Austin, TX, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3698364.3709130

1 Introduction

The field of physical design in Very Large-Scale Integration (VLSI) has become increasingly complex as advances in transistor scaling and the sheer number of transistors in integrated circuits (ICs) continue to push the boundaries of design automation [1]. These advances bring critical challenges to achieving optimal placement, routing, timing closure, and power efficiency [2]. Beyond technical challenges, issues such as research reproducibility, the adoption of open source tools, and fostering global collaboration remain underexplored yet pivotal to the advancement of physical design methodologies [3]. However, as transistor sizes are approaching atomic limits and their numbers have passed trillions, there is a need to accelerate innovation in the area.

In this paper, we look at the contributions published in the ISPD [4] over the past 25 years to see what lessons we can learn to accelerate future innovations and how to make the research ecosystem more productive. The data analyzed includes 682 research papers authored by 1,322 unique contributors from 484 institutions. The detailed research questions we looked at are:

- What catalyzes research in physical design?
- How can research impact be shown within the ISPD community?
- What are the impacts of open-source tools and benchmarks?
- What is the interplay between different research teams?
- Does collaboration between geographically separated research teams result in impactful research?

As designs become more complex, physical design and its different stages such as placement, routing, and timing closure become harder to manage [2, 5]. Overcoming these challenges will require a collaborative ecosystem across academia, industry, open source communities, as well as other research communities to make research accessible and inclusive [6, 7]. Without this ecosystem, innovation cannot happen. However, to make such an ecosystem, it is essential to identify past successes and address existing gaps. Over 10% or 71 out of 682 of the articles published in ISPD state that the work they are producing is novel. However, not all of these works have catalyzed other works in the field. The reasons for this can be the lack of sufficient methodological details limiting validation and broader adoption [8], i.e. reproducibility; lack of access to tools and compatibility issues [9, 10]; lack of access to the research ecosystem ristricting the diversity and inclusivity of contributions [11–13]; and, finally, scalability of new techniques like machine learning due to the lack of high-quality datasets [5, 14].

Despite advances in VLSI design automation, significant gaps persist. Limited adoption of open-source tools and uneven global contributions hinder collaboration and innovation [15]. Many studies lack reproducible methodologies and high-quality datasets, restricting validation and scalability [16]. Although machine learning is increasingly integrated into workflows, its application in complex real-world scenarios remains underexplored [17]. Additionally, advanced NLP techniques, such as Latent Dirichlet Allocation (LDA) and BERTopic [18, 19] offer promising avenues for analyzing thematic trends in VLSI research, yet remain underutilized in

This work is licensed under a Creative Commons Attribution 4.0 International License.
ISPD '25, Austin, TX, USA
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1293-7/25/03
https://doi.org/10.1145/3698364.3709130

bibliometric studies. These tools enable a nuanced understanding of research topics, which this paper explores in depth.

In this paper, we have used network analysis, natural language processing, and geospatial mapping to find the answers to the aforementioned problems. We have used network analysis to map citation patterns and co-authorship networks, uncovering influential research clusters and collaboration dynamics [20, 21]. Using natural language processing (NLP) techniques, we extracted core themes from abstracts [22]. Finally, we used geospatial mapping to provide insights into the evolution of global research hubs and highlight the shifting innovation centers in VLSI research [23]. The dataset analyzed includes research papers from ISPD conference proceedings, citation networks, author affiliations, and references.

Our analysis shows that there were several key topics, such as placement, routing, and clock networks that continue to be present. However, research on these topics was fueled by the ISPD contests. The study also shows that over time, as the complexity of the designs grew, the separation between some topics was reduced. For example, there are more papers that work on routing and placement as a whole now than a decade ago. There were several seminal papers with results so dramatic that the research in the topic has nearly stopped, for example, hMetis [24] for partitioning, and Flute [25] for Steiner tree generation. From the geospatial mapping, we can see that the center of research has shifted over the years, but at the same time, collaboration between teams have increased. Finally, new mathematical advancements such as advances in convex optimization and machine learning have had a lasting impact on the research.

The remainder of the paper is organized as follows: Section 2 reviews the related literature. Section 3 outlines the techniques for research analysis. In Section 4, our findings and insights are presented, potential solutions are discussed, and some recommendations are provided. Section 5 concludes with future directions.

2 Literature Review

In this section, we will first give a brief overview of the ISPD and its contests in Section 2.1. Then, in Sections 2.2, 2.3, we will discuss the background material for the tools used in the research. We will conclude the section by providing research on the trend we saw while analyzing the ISPD papers including reproducibility.

2.1 Physical Design Ecosystem Review

2.1.1 ISPD Conference and Community. The International Symposium on Physical Design (ISPD) was established in 1987 as a dedicated forum for addressing challenges in physical design automation within electronic design automation (EDA). Its history is marked by influential contests, benchmarks, and a lasting impact on the evolution of EDA tools and techniques.

2.1.2 ISPD Contests. ISPD contests have played an important role in the advancement of methodologies for placement, routing, and clock network optimization [28]. These contests serve as benchmarks for innovation, providing real-world challenges that encourage participants to push the boundaries of existing approaches [4, 28]. Recent studies, such as NVIDIA Research's GPU/ML-enhanced routing contests, highlight the integration of machine learning into design automation workflows, further advancing the field [27].

2.1.3 Open-Source Benchmarks and Tools. Open-source initiatives at ISPD have become pivotal in driving transparency, collaboration, and reproducibility in scientific research. ISPD Contests also provide sets of benchmarks for physical design including placement, global routing, detailed routing, clock synthesis, gate sizing, and security [28].

Recently, the OpenROAD Project [29] emphasizes the role of open-source tools in democratizing access to design automation workflows, enabling reproducible results and scalable solutions. Many of the physical design tools within OpenROAD started as projects as either ISPD papers or in ISPD contests. Tools like OpenMPL have addressed specific challenges in lithography, showcasing the potential of open source frameworks in advancing VLSI design [30]. Platforms such as VLSI Web provide a comprehensive overview of open source tools and their role in accelerating research and innovation [31]. Tools like iEDA further emphasize the role of open-source frameworks in bridging the gap between academia and industry [32].

2.2 Modeling Techniques Used in Analysis

2.2.1 Community Detection Algorithms. To reduce bias in clustering the research papers, we used several different community detection algorithms. These algorithms are integral to the analysis of collaborative networks. Key methods used in this research include:

- **Louvain Method**: A modularity optimization algorithm known for its scalability and efficiency in large networks, though limited in detecting smaller communities [33].
- **Girvan-Newman Algorithm**: Based on edge-betweenness centrality, this method excels in small networks but is computationally expensive for larger graphs [34].
- **Label Propagation Algorithm (LPA)**: A fast and scalable method for large-scale networks, though results can vary due to its non-deterministic nature [35].
- **Spectral Clustering**: Grounded in graph theory, it effectively identifies dense clusters but is computationally intensive for large datasets [36].
- **Graph Neural Networks (GNNs)**: These state-of-the-art methods leverage graph embeddings for community detection, excelling in complex and heterogeneous networks [37, 38].
- **Infomap**: An information-theoretic approach that optimizes flow-based community detection, suitable for directed networks [39].

These community detection algorithms have been widely applied across domains. A few examples are as follows: Louvain and GNNs have successfully analyzed large-scale networks like Twitter and Facebook, uncovering meaningful communities [33, 37, 38]. Citation Networks such as Girvan-Newman and spectral clustering have been used to explore collaboration patterns and academic influence [40, 41]. Biological Networks such as info-map and spectral clustering have demonstrated utility in understanding protein interactions and ecological systems [36, 42].

2.3 Application of NLP in Bibliometric Analysis

The integration of Natural Language Processing (NLP) techniques has revolutionized bibliometric studies by enabling the extraction of meaningful insights from large corpora of textual data. In this study, multiple NLP methods were applied to ISPD research publications to analyze research trends, collaboration patterns, and global contributions.

TF-IDF (Term Frequency-Inverse Document Frequency) [43] was employed to preprocess text data, enabling efficient keyword extraction and clustering of research topics. Additionally, Latent Dirichlet Allocation (LDA) [44] was used for topic modeling, revealing thematic clusters within the corpus.

To enhance the semantic understanding of textual data, transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) [22] and Sentence-BERT [45] were utilized. These methods provided contextual embeddings that allowed for semantic clustering of paper abstracts and identification of nuanced relationships between research topics.

Furthermore, Named Entity Recognition (NER) [46] was applied to extract key entities such as institutions and geographic locations from author affiliations. This facilitated the geospatial analysis of research contributions. For summarizing research abstracts and extracting dominant keywords, TextRank [47], a graph-based ranking algorithm, was employed.

Clustering algorithms such as k-means [48] and DBSCAN [49] were used to group research topics and collaboration patterns. These techniques provided insights into thematic clustering and the dynamics of team-based research. The study also leveraged advanced topic modeling techniques like BERTopic [50], which combines BERT embeddings with clustering algorithms to uncover dynamic and interpretable topics. This dual use of LDA for high-level trends and BERTopic for semantic analysis ensured a comprehensive exploration of ISPD publications.

By combining classical methods like TF-IDF and LDA with modern approaches such as BERT and BERTopic, this study demonstrated the potential of NLP in bibliometric analysis, offering robust insights into the evolution and impact of research within the ISPD community.

3 Analytical Techniques and Methodology

This study employs a multidisciplinary approach to analyze the relationships between research reproducibility, open-source tools, and their impact. By combining network analysis, natural language processing (NLP), and geospatial mapping, the methodology provides a comprehensive framework for understanding trends and contributions in the field.

The flowchart shown in Figure 1 show the flow used for the analysis. At the first stage, *Data Collection and Preprocessing*, data from ISPD was preprocessed to be used in the analysis. In the second stage, *Analysis*, three types of analysis were performed on the data. These include topic identification, building a citation network and geographic collaboration mapping. Finally, in the last stage, several metrics are designed, and research recommendations are made.

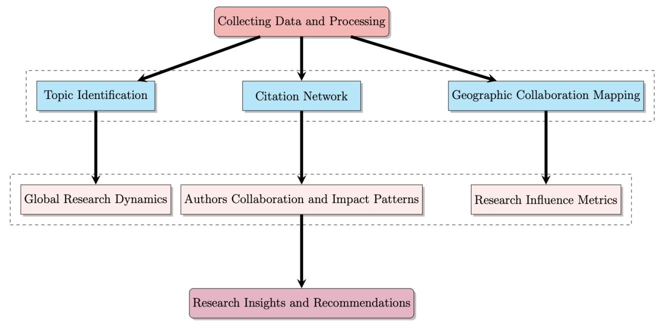

Figure 1: Flowchart illustrating the process from data collection to advanced analysis, using network metrics, NLP, citation network, and geospatial mapping.

3.1 Data Collection and Preprocessing

To conduct a robust bibliometric analysis, a curated dataset of ISPD conference proceedings spanning the past 25 years was utilized. The key steps in data preparation are as follows:

- **Data Collection:** Over 1000 research articles that were published in ISPD were collected. The metadata includes titles, abstracts, publication dates, author names, affiliations, and citation relationships. Citation data was used to construct a directed citation graph for thematic clustering and influence analysis.
- **Data Cleaning and Standardization:**
  - Duplicate entries and incomplete records were removed.
  - Variations in author names and institutional affiliations were standardized using automated scripts.
  - Non-research articles (e.g., editorials) were excluded.
  - Text data was processed with the Natural Language Toolkit (NLTK) to remove stopwords, punctuation, and non-standard characters [51].
- **Feature Extraction:** Additional features were derived, including:
  - *Temporal Features:* Trends over time were analyzed using publication dates.
  - *Citation Features:* Citation relationships were mapped in a directed graph to highlight influence and collaboration patterns.
  - *Geospatial Features:* Geographic data from affiliations was geocoded to visualize research hubs and international collaborations.
  - *Reproducibility Indicators:* Mentions of open-source tools, datasets, and reproducibility-related keywords were identified through semantic analysis.
- **Graph Construction and Community Detection:** Directed citation and co-authorship graphs were constructed using NetworkX [55]. The Louvain method [33] was applied for community detection, identifying thematic clusters and influential research groups.

- **Geocoding and Mapping:** Author affiliation data was geocoded with GeoPandas [53], facilitating visualization of global research hubs and collaboration networks. Results were validated and cached for efficiency.

3.2 Analysis

The second stage involved applying advanced techniques to extract insights from the data:

- **Natural Language Processing (NLP):**
  - *Topic Modeling:* Latent Dirichlet Allocation (LDA) [44] and BERTopic [50] identified dominant themes such as placement, routing, reproducibility, and open-source tools.
  - *Semantic Analysis:* Indicators of reproducibility (e.g., datasets, benchmarks) were extracted to assess their prevalence and impact.
  - *Trend Analysis:* Temporal patterns in research themes were analyzed using word embeddings [52].
- **Community Detection:** The Louvain method [33] was used to analyze the citation graph, identifying clusters of closely related studies. Metrics such as modularity, centrality, and cluster size highlighted influential communities and key research topics.
- **Geospatial Mapping and Collaboration Analysis:**
  - *Author Affiliations:* Geographic data was used to map global research hubs and analyze collaboration patterns.
  - *Temporal Shifts:* Changes in collaborative networks and emerging research centers were visualized over time [53].
  - *Co-Authorship Networks:* Metrics such as degree centrality and clustering coefficients identified prominent researchers and tightly connected groups.

3.3 Metrics and Evaluation

Key metrics were employed to evaluate research collaboration and impact:

- **Citation Analysis:** Statistical tests, such as the t-test [54], were applied to examine the relationship between citation counts and specific research indicators (e.g., datasets, benchmarks).
- **Community Metrics:** Metrics like modularity and betweenness centrality assessed the structure and cohesion of research communities, identifying influential clusters and key connections.
- **Collaboration Patterns:** Co-authorship networks were evaluated using degree centrality and clustering coefficients to understand the density and diversity of collaborations.
- **Visualization Tools:**
  - *Matplotlib:* Used for static visualizations like bar charts and network diagrams [56].
  - *Plotly:* Enabled interactive visualizations, including dynamic graphs and maps [57].

This structured methodology integrates advanced techniques and tools, providing actionable insights into research reproducibility, collaboration, and the role of open-source tools in ISPD contributions.

4 Key Finding

Over the last quarter century, there has been 682 articles published in ISPD. In Figure 2 the annual publication trends at ISPD from 2000 to 2024 is shown. The early growth phase (2000–2010) established ISPD as a key platform for physical design, with notable peaks in 2001, 2005, and 2010. Between 2011 and 2019, publication counts stabilized, averaging around 42 papers annually. In recent years (2020–2024), variability is evident, including a dip in 2021 likely due to the COVID-19 pandemic, followed by recovery. These trends highlight that even though the circuits themselves have scaled exponentially in the number of transistors and their complexity has increased due to lower technology nodes, the number of articles in ISPD has, in fact, decreased.

Table 1: Counts of Reproducibility, Open Source, and Contest Related Mentions

Category	Count
Total Articles	682
Reproducibility	127
Open Source	178
Contest Related	40

Table 1 highlights the connection between reproducibility, open-source tools, and research impact in VLSI design. Among 772 articles, 132 focus on reproducibility, 127 on open-source tools, and 59 on contests. These counts were derived from analyzing the abstracts of the articles, where specific keywords and patterns were flagged using predefined regular expressions. Citation counts were calculated by matching article titles referenced within the dataset, providing insights into the influence and recognition of specific papers.

4.1 Research Topics

The first analysis was done on the different research topics from ISPD. In Table 2, the results of such analysis using BERTopic Modelling are given. In this table, the research articles are clustered into 10 different topics with each topic cluster having a series of 4 words to describe it.

In Figure 3, the evolution of each one of these research topics as obtained by Bertopic modeling from 2000 to 2024 is shown. These figures highlight key catalysts for physical design research. There are consistent topics such as *topic 1: circuit design*, *Topic 2: placement algorithms*, and *Topic 10: analog design* constant stream of papers, as shown in subplots 1, 2 and 10. On the other hand, emerging topics include Topic 2 shown in light green. This topic is related to research that uses machine learning. The trends in Figure 3 also show a decline in publication in areas such as *floorplanning* and *power grid analysis.* This lack of popularity can be attributed to the lack of contests in the area, which results in a lack of suitable benchmarks and/or testing mechanisms to show the efficacy of the proposed solutions. The LDA topic modelling showed similar patterns in topics.

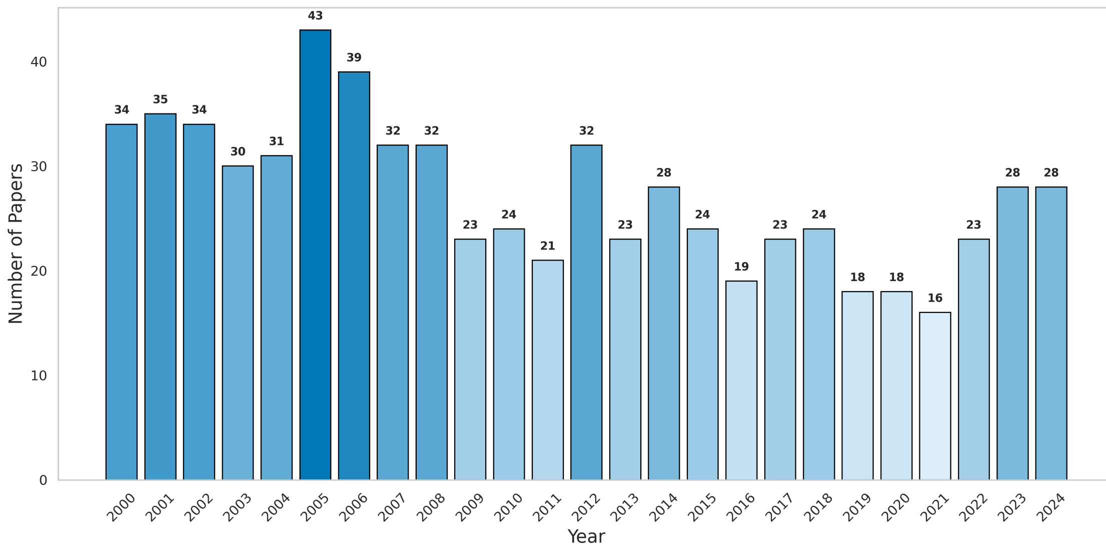

Figure 2: Number of Papers Published Each Year

Table 2: Bertopic (BT) Modeling and LDA Summary of the papers published in ISPD between 2000 to 2024

BT	word 1	word 2	word 3	word 4	Count
1	design	physic	circuit	technolog	107
2	placement	cell	wirelength	placer	99
3	design	layout	learn	pattern	97
4	rout	algorithm	tree	steiner	82
5	interconnect	thermal	delay	wire	69
6	clock	tree	skew	power	67
7	FPGA	design	ASIC	placement	57
8	floorplan	block	algorithm	represent	39
9	power	grid	current	design	38
10	analog	circuit	cluster	design	27
LDA	**word 1**	**word 2**	**word 3**	**word 4**	**#**
1	physic	optim	time	design	195
2	global	rout	algorithm	placement	91
3	perform	chip	power	design	84
4	problem	pin	design	rout	81
5	wirelength	algorithm	cell	placement	60
6	placement	benchmark	power	floorplan	43
7	method	densiti	thermal	circuit	38
8	technolog	lithographi	design	layout	32
9	variat	skew	tree	clock	32
10	placement	FPGA	algorithm	cluster	26

4.2 ISPD as a Network of Research

Another analysis was done to design a graph of the connections between papers. This is called finding communities. The communities show connections between papers through citations. In essence, the communities show which papers are catalysts for research.

This analysis of citation networks in ISPD provides a comprehensive view of the research landscape. Larger research communities focus on established areas like placement, routing, and timing, while smaller communities highlight emerging topics such as 3D IC integration and AI-driven methodologies.

Temporal trends reveal increasing research activity, with notable spikes aligned with technological advancements and benchmarking initiatives like ISPD contests. ISPD contests are emphasized as pivotal for driving impactful research through realistic benchmarks, fostering academia-industry collaboration, and producing foundational works.

4.3 Collaborations

In our last analysis, we are showing the geographic locations of the authors who have published in ISPD based on their affiliation in Figure 5. The map illustrates how the centers of innovation and research for ISPD.

In addition, if the figures are shown over the year, there is a evolving landscape of the field where more and more research is performed in East Asia, while more collaborations are made internationally. The shifts reflect changes in research funding, institutional focus, and international collaborations, suggesting a more distributed and globally collaborative landscape for VLSI research.

ISPD contests have notably influenced Electronic Design Automation (EDA) by providing benchmarks and challenges that foster innovation. For instance, the 2019 ISPD contest on detailed routing attracted 33 teams from nine countries, emphasizing its global significance. Contests like these not only address real-world challenges but also guide research directions, as evidenced by the high citation counts of contest-related papers [60].

5 Conclusions

In this work, we conducted an extensive analysis of 25 years of research presented at ISPD. This longitudinal study highlights the

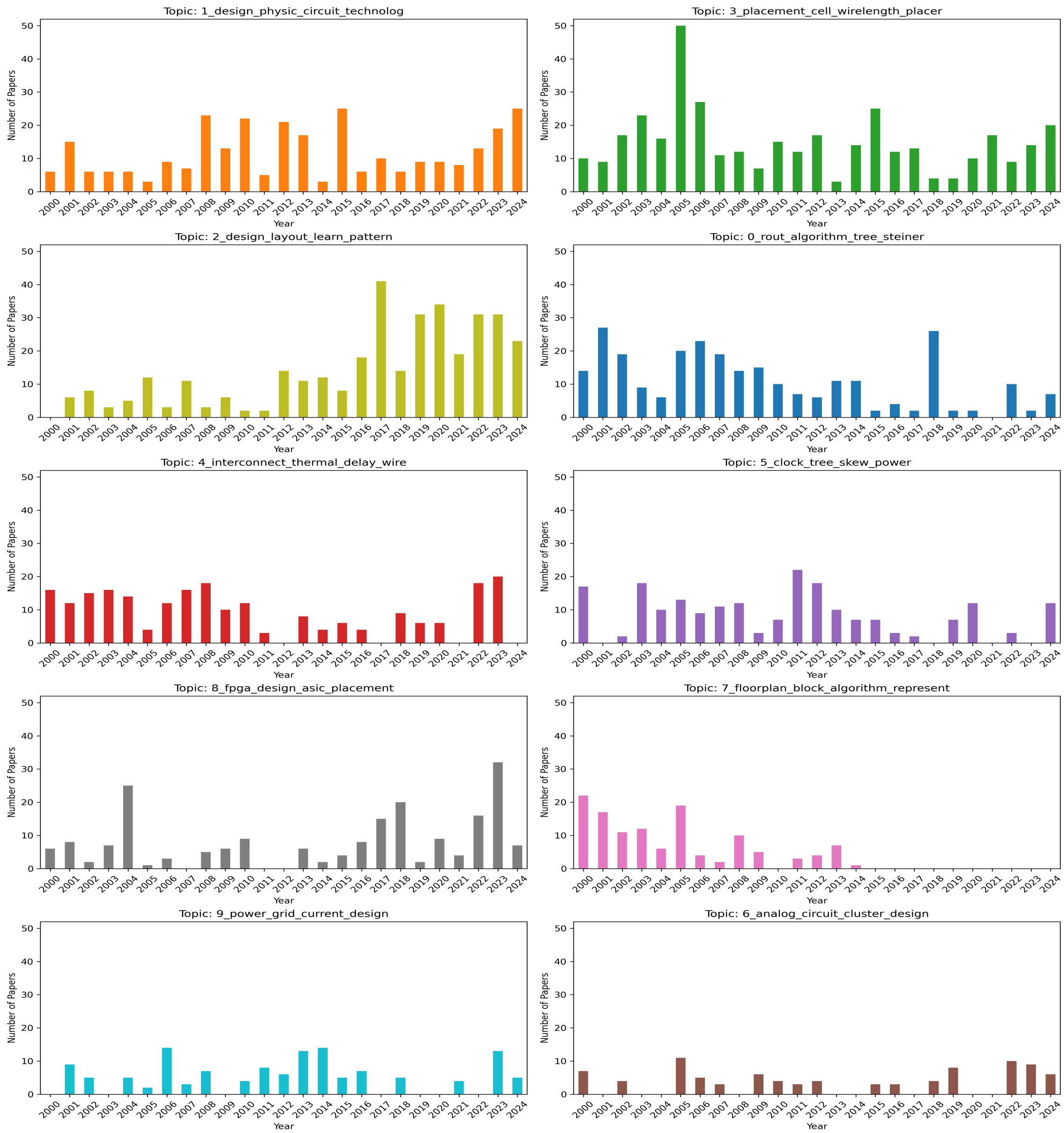

Figure 3: Year-by-Year Analysis of the Most Frequent Topics in ISPD Research by Bertopic

evolving trends, key contributors, and collaborative networks shaping the field of VLSI Physical Design. The findings emphasize the growing importance of open-source tools, reproducibility, and competitive problem-solving as catalysts for impactful research and innovation.

Our analysis reveals that reproducibility and accessibility of resources, such as datasets and benchmarks, significantly enhance research visibility and adoption. Additionally, ISPD contests have emerged as a driving force for innovation by addressing practical, industry-relevant challenges and fostering academia-industry collaboration. The geographic analysis indicates a gradual shift in research leadership, suggesting the emergence of new regional powerhouses contributing to the global research ecosystem.

References

[1] Sutherland, I., Sproull, R., and Harris, D. (2017). *Logical Effort: Designing Fast CMOS Circuits*. Morgan Kaufmann.

[2] Kahng, A. B. (2018). "VLSI Physical Design: From Graph Partitioning to Timing Closure." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 37(9), 1649-1666.

[3] Bjørner, D., and Havelund, K. (2014). "40 years of formal methods: Some obstacles and some possibilities." *Science of Computer Programming*, 78(10), 1132–1145. DOI: 10.1016/j.scico.2013.03.012.

[4] International Symposium on Physical Design (ISPD). "ISPD Conference Archives." Available at: https://ispd.cc.

(a) Research community 2000 to 2005

(b) Research community 2000 to 2010

(c) Research community 2000 to 2015

(d) Research community 2000 to 2020

Figure 4: Comprehensive citation graph from 2000 to 2024. In these figures, each node is a paper published in ISPD. The color of the node is the representation of the community it belongs to. The size of the node is the illustration of the number of ISPD citations it has had. The directed edges are illustrations of the citation.

[5] Pan, J., Chang, C.-C., Xie, Z., Chen, Y. (2023). "EDALearn: A Comprehensive RTL-to-Signoff EDA Benchmark for Democratized and Reproducible ML for EDA Research." *arXiv preprint*, arXiv:2312.01674.

[6] Baker, M. (2016). "Reproducibility Crisis: Blame it on the Computer." *Nature*, 533, 452-454.

[7] Sauter, P., Benz, T., Scheffler, P., Gürkaynak, F. K., Benini, L. (2024). "Insights from Basilisk: Assessing Open-Source EDA Tools for Large-Scale SoC Design." arXiv preprint arXiv:2405.04257.

[8] Collberg, C., and Proebsting, T. A. (2016). "Repeatability in Computer Systems Research." *Communications of the ACM*, 59(3), 62-69.

[9] Patel, M., and Roy, S. (2021). "Open-Source Electronic Design Automation (EDA) Tools." *IEEE Design Test*, 38(2), 7–16. DOI: 10.1109/MDAT.2021.3050345.

[10] Shah, J., and Koufos, K. (2020). "Open-source EDA: Challenges and Opportunities in VLSI Design Automation." *Proceedings of the Design Automation Conference (DAC)*, 1-6. DOI: 10.1145/3394782.3430624.

[11] Leslie, J. (1993). "Competition and Cooperation: A Comparative Analysis of SEMATECH and the VLSI Project." *California Management Review*, 35(4), 37-56.

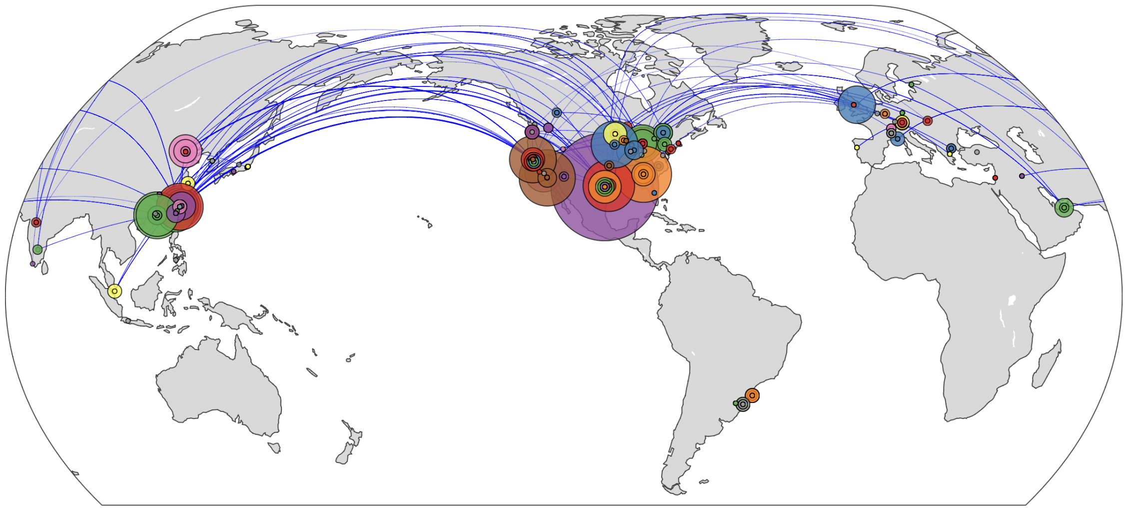

Figure 5: Illustration of the geographic location of the authors of ISPD. The lines in the figure show the collaborations between authors.

https://www.jstor.org/stable/41166745

[12] Lamsal, R. R., Devkota, A., and Bhusal, M. S. (2023). "Navigating Global Challenges: The Crucial Role of Semiconductors in Advancing Globalization." *Journal of The Institution of Engineers (India): Series B*, 104, 1389–1399. https://doi.org/10.1007/s40031-023-00938-4.

[13] Ren, Y., Yang, Y., Wang, Y., and Liu, Y. (2023). "Dynamics of the Global Semiconductor Trade and Its Dependencies." *Journal of Geographical Sciences*, 33, 1141–1160. https://doi.org/10.1007/s11442-023-2123-9.

[14] Srivastava, P., Kumar, P., Abbas, Z. (2023). "Qualitative Data Augmentation for Performance Prediction in VLSI Circuits." *arXiv preprint*, arXiv:2302.07566.

[15] Ajayi, O., & Vidya, S. (2022). "Open source VLSI design automation: An RTL to GDSII flow using OpenLane." *AIP Conference Proceedings*. DOI: 10.1063/5.0078532.

[16] Chai, Z., Yu, X., Zhou, J., Guo, C., & Chang, C.-K. (2022). "CircuitNet: An Open-Source Dataset for Machine Learning Applications in Electronic Design Automation (EDA)." *arXiv preprint arXiv:2201.01234*.

[17] Khetarpal, R., Singh, A., & Sharma, N. (2024). "Machine Learning Techniques for VLSI Circuit Design: A Review." *Springer Nature: Advances in Intelligent Systems and Computing*. DOI: 10.1007/978-3-031-30989-4_22.

[18] Kaur, A., & Wallace, J. R. (2024). "Moving Beyond LDA: A Comparison of Unsupervised Topic Modelling Techniques for Qualitative Data Analysis of Online Communities." *arXiv preprint arXiv:2412.14486*.

[19] Turan, S. C., Yildiz, K., & Büyüktanir, B. (2024). "Comparison of LDA, NMF and BERTopic Topic Modeling Techniques on Amazon Product Review Dataset: A Case Study." In *Computing, Internet of Things and Data Analytics* (pp. 23–31). Springer. DOI: 10.1007/978-3-031-53717-2-3.

[20] Liu, J., Tang, T., Wang, W., Xu, B., Kong, X., and Xia, F. (2018). "A Survey of Scholarly Data Visualization." *IEEE Access*, 6, 19205–19221. https://doi.org/10.1109/ACCESS.2018.2815030.

[21] Ding, Y., and Yan, E. (2014). "Scholarly Networks Analysis." In *Encyclopedia of Social Network Analysis and Mining* (pp. 1715–1726). Springer. https://doi.org/10.1007/978-1-4614-6170-8-249.

[22] Devlin, J., Chang, M., Lee, K., and Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *arXiv preprint arXiv:1810.04805*.

[23] Authors Unknown. (2023). "A Geospatial Data and Model Hub for Online Geospatial Analysis." *Proceedings of the IEEE International Conference on Cloud Computing and Geospatial Applications*, pp. 1-8. DOI: 10.1109/Cloud-GeoApp2023.10233547.

[24] Karypis, George, and Vipin Kumar. "A hypergraph partitioning package." Army HPC Research Center, Department of Computer Science Engineering, University of Minnesota (1998).

[25] Chu, Chris, and Yao-Wen Chang. "FLUTE: Fast lookup table based rectilinear Steiner minimal tree algorithm for VLSI design." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 27.1 (2008): 70-83. DOI: 10.1109/TCAD.2007.907073.

[26] Eslami, M., Knechtel, J., Sinanoglu, O., Karri, R., and Pagliarini, S. (2023). "Benchmarking Advanced Security Closure of Physical Layouts: ISPD 2023 Contest." *Proceedings of the 2023 International Symposium on Physical Design*, pp. 256-264. DOI: 10.1145/3569052.3578924.

[27] NVIDIA Research. (2024). "GPU/ML-Enhanced Large-Scale Global Routing Contest." Presented at the International Symposium on Physical Design (ISPD).

[28] ISPD Previous Contests, https://ispd.cc/ispd2025/index.php?page=contests, 2005-2025.

[29] The OpenROAD Project. "An Open-Source, Autonomous, and Scalable RTL-to-GDSII Toolchain for VLSI Design."

[30] W. Li, Y. Ma, Q. Sun, Y. Lin, I. H.-R. Jiang, B. Yu, and D. Z. Pan, "OpenMPL: An Open Source Layout Decomposer," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 5, pp. 903–916, May 2021. DOI: 10.1109/TCAD.2020.3042175.

[31] VLSI Web. "Open-Source Tools and Resources for VLSI Design." Available: https://www.vlsiweb.com.

[32] X. Li et al., "iEDA: An Open-Source Intelligent Physical Implementation Toolkit and Library." *arXiv preprint arXiv:2308.01857*, 2023.

[33] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment*, 2008.

[34] M. Girvan and M. E. Newman, "Community Structure in Social and Biological Networks." *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821-7826, 2002.

[35] U. N. Raghavan, R. Albert, and S. Kumara, "Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks." *Physical Review E*, vol. 76, no. 3, 2007.

[36] U. von Luxburg, "A Tutorial on Spectral Clustering." *Statistics and Computing*, vol. 17, no. 4, pp. 395-416, 2007.

[37] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks." *arXiv preprint arXiv:1609.02907*, 2016.

[38] Chen, Z., Li, L., Bruna, J. (2019). "Supervised Community Detection with Line Graph Neural Networks." *International Conference on Learning Representations (ICLR)*.

[39] M. Rosvall and C. T. Bergstrom, "Maps of Random Walks on Complex Networks Reveal Community Structure." *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118-1123, 2008.

[40] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and Identifying Communities in Networks." *Proceedings of the National Academy of Sciences*, vol. 101, no. 9, pp. 2658-2663, 2004.

[41] L. Donetti and M. A. Muñoz, "Detecting Network Communities: A New Systematic and Efficient Algorithm." *Journal of Statistical Mechanics: Theory and Experiment*, 2005.

[42] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The Map Equation." *The European Physical Journal Special Topics*, vol. 178, pp. 13-23, 2011.

[43] Salton, G., Wong, A., & Yang, C. S. (1988). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613-620. DOI: 10.1145/361696.361698.

[44] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*, 993-1022.

[45] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3982-3992. DOI: 10.18653/v1/D19-1410.

[46] Tjong Kim Sang, E. F., & De Meulder, F. (2003). Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition. *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, 142-147.

[47] Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing Order into Texts. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 404-411.

[48] MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics, 281-297.

[49] Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 226-231.

[50] Grootendorst, M. (2022). "BERTopic: Leveraging BERT and c-TF-IDF to Create Easily Interpretable Topics." *arXiv preprint arXiv:2203.05794*.

[51] Bird, S., Klein, E., Loper, E. (2009). "Natural Language Processing with Python." *O'Reilly Media Inc*. Available: https://www.nltk.org.

[52] Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). "Efficient Estimation of Word Representations in Vector Space." *arXiv preprint arXiv:1301.3781*.

[53] Jordahl, K., Den Bossche, J. V., Fleischmann, M., et al. (2021). "GeoPandas: Python tools for geographic data." Available: https://geopandas.org.

[54] Student. (1908). "The Probable Error of a Mean." *Biometrika*, vol. 6, no. 1, pp. 1–25. DOI: 10.2307/2331554.

[55] Hagberg, A. A., Schult, D. A., Swart, P. J. (2008). "Exploring Network Structure, Dynamics, and Function Using NetworkX." *Proceedings of the 7th Python in Science Conference (SciPy)*, pp. 11–15.

[56] Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment." *Computing in Science Engineering*, vol. 9, no. 3, pp. 90–95.

[57] Plotly Technologies Inc. (2015). "Collaborative data science." Available: https://plotly.com.

[58] M.-C. Kim, D.-J. Lee, and I. L. Markov, "SimPL: An Effective Placement Algorithm," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 1, pp. 50–60, Jan. 2012.

[59] I. Kuon and J. Rose, "FPGA Architecture: Survey and Challenges," *Foundations and Trends® in Electronic Design Automation*, vol. 2, no. 2, pp. 135–253, 2007.

[60] J. Hu and S. S. Sapatnekar, *ISPD 2005 Placement Contest and Benchmark Suite*, International Symposium on Physical Design, 2005.