Five Decades of the ACM Special Interest Group on Data Communications (SIGCOMM): A Bibliometric Perspective

Waleed Iqbal
Information Technology University (ITU)-Punjab, Lahore, Pakistan
waleed.iqbal@itu.edu.pk

Junaid Qadir
Information Technology University (ITU)-Punjab, Lahore, Pakistan
junaid.qadir@itu.edu.pk

Saeed-Ul Hassan
Information Technology University (ITU)-Punjab, Lahore, Pakistan
saeed-ul-hassan@itu.edu.pk

Rana Tallal Javed
Information Technology University (ITU)-Punjab, Lahore, Pakistan
tallal.javed@itu.edu.pk

Adnan Noor Mian
Computer Laboratory, University of Cambridge, United Kingdom
anm42@cl.cam.ac.uk

Jon Crowcroft*
Computer Laboratory, University of Cambridge, United Kingdom
jon.crowcroft@cl.cam.ac.uk

Gareth Tyson*
Queen Mary University of London, United Kingdom
gareth.tyson@qmul.ac.uk

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.
The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

The ACM Special Interest Group on Data Communications (SIGCOMM) has been a major research forum for fifty years. This community has had a major impact on the history of the Internet, and therefore we argue its exploration may reveal fundamental insights into the evolution of networking technologies around the globe. Hence, on the 50^{th} anniversary of SIGCOMM, we take this opportunity to reflect upon its progress and achievements, through the lens of its various publication outlets, *e.g.*, the SIGCOMM conference, IMC, CoNEXT, HotNets. Our analysis takes several perspectives, looking at authors, countries, institutes and papers. We explore trends in co-authorship, country-based productivity, and knowledge flow to and from SIGCOMM venues using bibliometric techniques. We hope this study will serve as a valuable resource for the computer networking community.

CCS CONCEPTS

• **Applied computing** → **Digital libraries and archives**; • **General and reference** → *General conference proceedings*;

KEYWORDS

Co-authorship Patterns, Bibliometrics, Computer Networking, Social Network Analysis, Full-text

1 INTRODUCTION

The ACM's Special Interest Group on Data Communications (SIGCOMM) has performed a pivotal role in the development of computer networking. The research area has grown over decades, bridging work from three major domains: Computer Science, Electrical Engineering, and Computer Engineering. On the 50^{th} anniversary of SIGCOMM's foundation, we believe it is timely and worthwhile to explore its history and role, via the publication of cutting edge research. We approach this problem through a bibliometric analysis of SIGCOMM's various publication outlets, covering 50 years of accepted research articles (ranging from 1969 to 2018). These accepted papers are published in main proceedings, affiliated proceedings, and affiliated workshops of SIGCOMM events. Using our dataset, we explore bibliometric questions and examine publication behaviors. Through this study, we strive to reveal major contributors to all venues under the umbrella of SIGCOMM, as summarized in Table 1. Although a number of past bibliometric studies have been conducted in various fields ([1–8]), our work is the first to focus on the overall literature of SIGCOMM.

We start by explaining the details of our dataset in Section 2. We then discuss the results generated by our data, and highlight key observations in Section 3. Finally, we conclude by summarizing our analysis in Section 4. This paper intends to offer some initial insights and visualizations of the research activities within SIGCOMM venues. We do not, however, strive to provide comprehensive or deep coverage of *all* activities within SIGCOMM. Consequently, to facilitate further research, we have publicly shared the dataset used in this paper.[1]. We also have developed an interactive visualization

*Authors also affiliated with the Alan Turing Institute.

[1] https://github.com/waleediqbal411/CCR-paper-data2019

Table 1: Features of dataset extracted from the SIG venues during 1969–2018

Attribute Name	Type of Attribute	Count										
		SIGCOMM	IMC	CoNext	ICN	E-Energy	SenSys	SoSR	LANC	HotNets	ANRW	ANCS
Starting Year		1969	2001	2005	2011	2010	2003	2015	2003	2006	2016	2005
Number of Articles	Numerical	3480	779	773	210	479	989	121	103	239	60	353
Number [Name] of Authors	Numerical [String]	6182	361	357	394	301	334	419	298	337	258	501
Number [Name] of Institutes	Numerical [String]	159	121	160	161	161	160	123	83	160	52	160
Number of References	Numerical	27407	20314	9737	12146	7163	13907	3369	1764	6272	747	6556
Citations of Articles	Numerical	576534	34734	14702	5256	3087	16353	877	282	5071	98	4258
Number [Name] of Participating Countries	Numerical [String]	61	40	46	45	39	39	27	17	26	21	31

of our analysis which can be used to observe temporal and spatial trends in a more interactive manner.[2] We hope that this can be of benefit to the community, and trigger follow-up research into SIGCOMM's publication activities.

2 DATA PRELIMINARIES

To perform our analysis, we have used a collection of 7,586 accepted articles between 1969–2018 from the main proceedings and workshops of the flagship ACM SIGCOMM conference, as well as other affiliated proceedings of SIGCOMM.[3] For all other venues except SIGCOMM main proceedings, we exclusively include only main track papers in our analysis and exclude all poster and demo papers. Details of the venues are shown in Table 1. This dataset contains all indexed papers published in SIGCOMM affiliated venues obtained from different repositories, including Scopus[4] and the ACM Digital Library.[5]

The dataset contains bibliographic details for each paper, including title, keywords, references, publication year, as well as author affiliations. 103 incomplete or irrelevant entries were removed from the dataset: These entries include messages from editors, entries without references, and entries without relevant metadata such as author names, institute names and indexed keywords. Details of the features extracted from these articles are shown in Table 1. Among other things, the table shows that each venue has different characteristics and longevity. For example, ANRW only has 60 publications and LANC only ran 6 editions between 2001 and 2011. Hence, our later analysis should be tempered by this observation. We also gather citation counts using the Scopus digital repository. We choose Scopus because it contains a reliable, up-to-date and controlled set of citations, rather than open repositories (*e.g.*, Google Scholar) that crawl citations from any accessible site [9].

Note that the SIGCOMM conference proceedings include many forms of article, *e.g.*, main track, posters, workshops and Best of CCR. Therefore, when computing the top ranked entities (*e.g.*, authors, institutes, countries), we manually vet to *only* count SIGCOMM main track papers. Other analyses (*e.g.*, Openness to Emerging Authors) includes authors who have published any forms of article. That said, although we have taken great care in manually validating the dataset, we cannot discount minor errors in parsing the repository entries. This is because they contain a large number of variations and complexities across the year. As such, we make our dataset publicly available and welcome further validation efforts.

[2] https://charts-sigcomm.herokuapp.com/
[3] http://www.sigcomm.org/
[4] https://www.scopus.com
[5] https://dl.acm.org

Table 2: Top 5 authors in SIGCOMM venues (1969–2018). If a position is taken by multiple authors, we list them all.

Venue	Top Author
SIGCOMM	Scott Shenker, Dina Katabi, Ion Stoica, Jennifer Rexford, Nick Feamster, George Varghese
IMC	Vern Paxson, Anja Feldmann, Paul Barford, Konstantina Papagiannaki, Christo Wilson, Nick Feamster
CoNext	Jennifer Rexford, Christophe Diot, Konstantina Papagiannaki, Olivier Bonaventure, Domenico Giustiniano
ICN	Lixia Zhang, Luca Muscariello, Thomas C. Schmidt, Toru Hasegawa, Dario Rossi, Matthias Waehlisch, Giovanna Carofiglio
E-Energy	Srinivasan Keshav, Hermann de Meer, Sid Chi-Kin Chau, Vijay Arya, Krithi Ramamritham, Catherine Rosenberg
SenSys	Tian He, Prabal Dutta, John A. Stankovic, Mani B. Srivastava, Philip Levis, David E. Culler
SoSR	Jennifer Rexford, Laurent Vanbever, Robert Soulé, Theophilus Benson, Nate Foster, Nick Feamster, Changhoon Kim
LANC	Eduardo Cerqueira, Benjamín Barán, Pablo Belzarena, Antonio Jorge Gomes Abelém, Denis do Rosário, Héctor Cancela, Eduardo Grampín
HotNets	Scott Shenker, Hari Balakrishnan, Vyas Sekar, Aditya Akella, Sylvia Ratnasamy, Jennifer Rexford
ANRW	Georg Carle, Brian Trammell, Marco Chiesa, Marco Canini, Benoit Donnet, Mirja Kühlewind
ANCS	Patrick Crowley, Tilman Wolf, Laxmi N. Bhuyan, Bin Liu, Bill Lin, Jun Li, Andrew W. Moore, Jan Korenek

3 RESULTS AND FINDINGS

We now explore several features of our bibliometric dataset. We intentionally provide a broad brush overview of publication trends, and make our data publicly available for other researchers wishing to focus on any particular theme covered.

3.1 An Author Perspective

We begin by exploring trends pertaining to authors who regularly published in SIGCOMM affiliated events.

Author Paper Count. We first compute the top authors across each venue in an attempt to identify key players within the community. op author analysis for authors with most publications is manually vetted to include only authors for SIGCOMM main track papers. Say that the other analysis includes authors who have published any forms of article in the SIGCOMM conference, including Best of CCR papers, posters and workshops. Figure 1 presents the authors with the most publications across all venues. Unsurprisingly, a number of extremely prominent researchers can be observed in this top list. We see that the SIGCOMM main conference is prominent across all of these top authors, followed by HotNets, IMC, and CoNEXT. We also observe more specialist conferences dominating certain author's records; for example, Tian He has a significant number of publications in SenSys. Note that the size and longevity of each venue has a major impact on these results.

To give greater insight into the most prominent authors on a per-conference basis, Table 2 shows the top authors based on publication count in each of the major venue under the SIGCOMM banner. From Table 2 we observe that some of the authors are performing equally well in multiple top venues, *e.g.*, Scott Shenker and Jennifer Rexford are categorized as the top authors in both SIGCOMM and HotNets, and CoNext and SoS,R respectively. Further, both are the overall top two most published author across all venues.

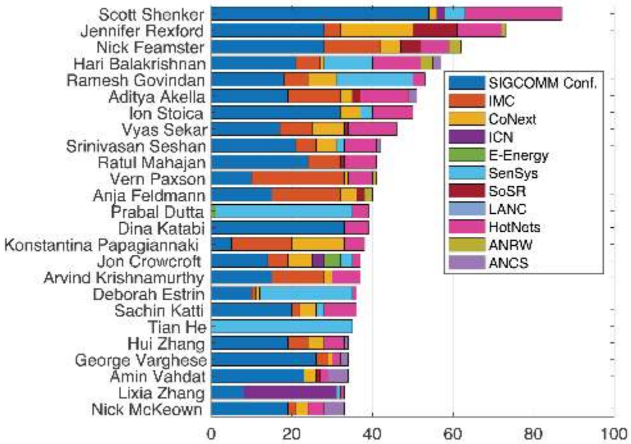

Figure 1: Top authors based on publication count during 1969–2018 in all SIGCOMM venues mentioned in our dataset. ***The flagship SIGCOMM conference dominates, but authors tend to have a mix of publications.***

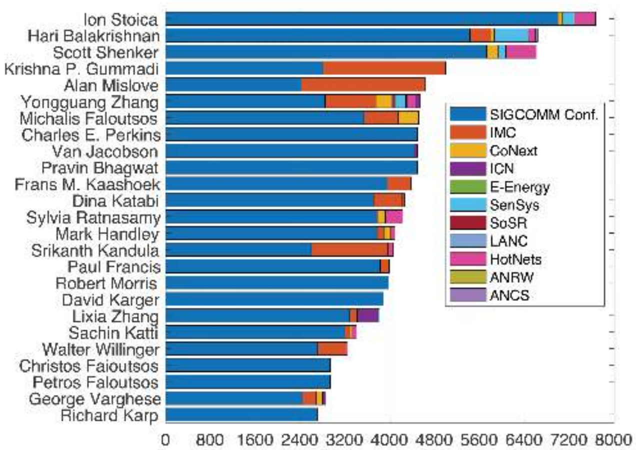

Figure 2: Most cited authors in SIGCOMM venues during 1969–2018, as defined by citation count. ***The majority of citations are accumulated from the flagship SIGCOMM conference papers***

Author Citation Rates. Of course, paper count alone does not necessarily provide insight into impact. Although a coarse measure, we turn to citation rates as a proxy of academic impact. Figure 2 shows the authors with the highest citation counts across their SIGCOMM sponsored publications. Interestingly, whereas Figure 2 reveals that many top authors publish in a number of venues, Figure 2 shows that the majority of citations come from papers published in the SIGCOMM main conference, followed by IMC. This highlights the importance of the SIGCOMM flagship conference, but also the importance of measurement research.

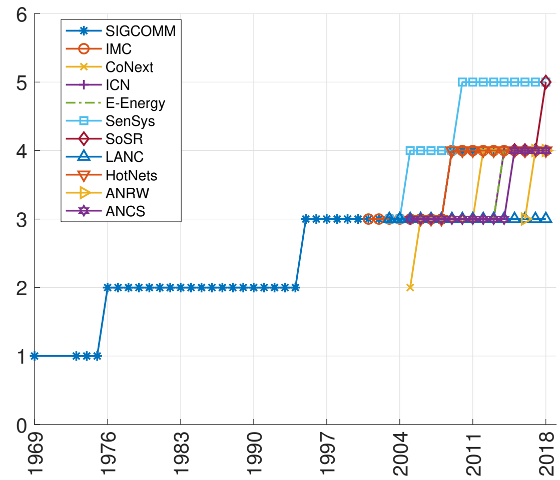

Figure 3: Median number of authors during 1969–2018 in SIGCOMM venues. ***Collaborative authorship is becoming more popular over time.***

Author Collaboration. A potential reason for the high productivity of certain authors is their ability to put together strong teams of collaborators. Hence, we proceed to explore the collaboration rates among well published authors. To begin, Figure 3 briefly presents the median number of authors in each year of SIGCOMM affiliated venues during 1969–2018. As expected, this shows that collaborative authorship trends are increasing across all venues. Whereas in the early years of SIGCOMM, papers tended to be authored by two people, it is now common to exceed four.

Of course, co-authorship counts alone are not sufficient to shed light on true collaborative practices, as it is also important to understand *who* collaborates. Figure 4 presents the co-authorship graph for all authors across SIGCOMM venues. To identify communities of collaborative networks, we compute modularity and colour nodes based on which cluster they belong to. We observe six major communities in the graph, although only four of them contain large numbers of top published authors. These groups are dominated by authors from universities such as UC Berkeley, MIT, USC, UCSD, and Princeton, which highlights the dominant role that US universities have historically played within the SIGCOMM community. For example, top authors like Nick Feamster, Jennifer Rexford and Scott Shenker have significantly co-authored articles. Similarly, Jia Wang and Soumya Sen have co-authored many papers. Of course, this in itself is not a novel observation, yet we argue it is useful to visualize these patterns.

As well as these dense clusters of collaborators, we also observe authors who interconnect the wider community; these are manifested as "bridges" or highly central nodes that connect important people within the co-authorship graph. To explore this, we compute the Eigenvector centrality [10] of all authors; Table 3 shows those with the highest values. There is a clear set of highly important

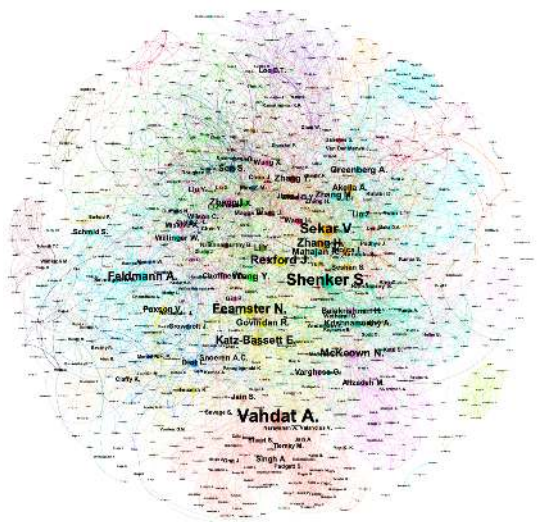

Figure 4: Co-authorship network in SIGCOMM. Node size indicates the number of links with other nodes in the co-authorship network and the node color represents cluster membership. ***Authors tend to form into collaborative communities.***

Table 3: Top authors with highest values of centrality in several leading venues sponsored by SIGCOMM

Venue	Top 5 Most Central Authors
SIGCOMM	Amin Vahdat, Scott Shenker, Jennifer Rexford, Yongguang Zhang, Ethan Katz-Bassett
IMC	Vern Paxson, Christo Wilson, Lixia Zhang, Anja Feldmann, David Choffnes
CoNext	Jon Crowcroft, Jennifer Rexford, Konstantina Papagiannaki, Yongguang Zhang, Chuanxiong Guo
ICN	Lixia Zhang, Alexander Afanasyev, Luca Muscariello, Jeff Burke, Beichuan Zhang
E-Energy	Vijay Arya, Deva P. Seetharam, Vikas Chandan, Tanuja Ganu, Harshad Khadilkar
SenSys	John Stankovic, Yongguang Zhang, Tarek Abdelzaher, Lixia Zhang, David Culler
SoSR	Nate Foster, Chang Kim, Jennifer Rexford, Robert Soule, Theo Jepsen
LANC	Eduardo Cerqueira, Augusto Neto, Antonio Abelem, Adalberto Melo, Denis do Rosario
HotNets	Srinivasan Seshan, Aditya Akella, Scott Shenker, Mohammad Alizadeh, Jennifer Rexford
ANRW	Godred Fairhurst, Felix Weinrank, Anna Brunstrom, Per Hurtig, Michael Tüxen
ANCS	Patrick Crowley, Michela Becchi, Jonathan Turner, John D. DeHart, Shakir James

"bridge" nodes; for example, in the SIGCOMM main conference, Amin Vahdat emerges as the most central author (whereas Scott Shenker has the highest publication count). Similarly, in CoNext, E-Energy, SenSys, and ANRW, Jon Crowcroft, Deva P. Seetharam, John Stankovic, and Godred (Gorry) Fairhurst emerge as key central nodes. These authors naturally play a vital role in the wider community.

Openness to Emerging Authors. From the above analysis it is evident that SIGCOMM sponsored events attract attention from significant researchers in the field. Hence, we posit that it may be difficult for new emerging scholars to publish in such venues. Indeed, anecdotally, this is often claimed. To identify emerging authors, we extract all papers with: (*i*) authors who have never published in the venue before; and (*ii*) authors who do not have any co-authors who have already published in the venue. Table 4 shows the distribution of emerging authors in SIGCOMM conferences during 1969–2018. Indeed, the majority of papers *do* contain authors who have previously published at the venue. Co-located workshops appear to play a critical role in providing opportunities to aspiring authors though. At the SIGCOMM flagship conference, 72.7% of emerging authors publish their manuscripts in SIGCOMM *workshops*,[6] leaving just 27.3% publishing in the main track. This suggests that, although it is feasible for new authors to access the SIGCOMM community more generally, it is much less regular to get papers published in the main track. It further highlights the importance of co-located workshops in opening the community to new entrants. Finally, Figure 5 presents the geo-distribution of these emerging authors. The US is ranked first in terms of new authors in SIGCOMM. Canada, China, UK, Germany, France also have top positions. This perhaps suggests that steps should be taken to better support new emerging authors coming from non-traditional academic powerhouses.

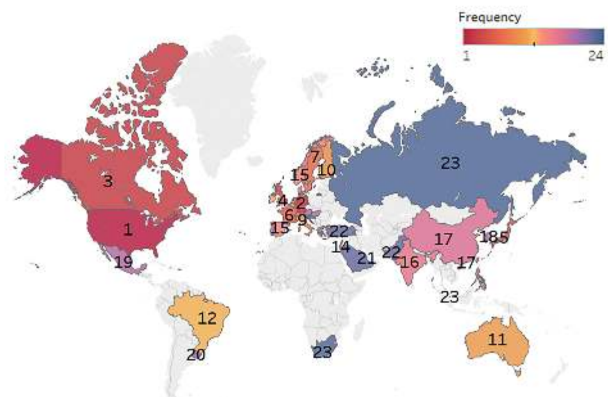

Figure 5: Rank of countries in SIGCOMM venues based on their publication count by emerging authors. ***There are significant numbers of global emerging authors, although the US still dominates in this regard.***

[6]In the ACM repository and Scopus, 58 SIGCOMM workshops are indexed (besides main proceedings)

Table 4: The distribution of emerging authors in SIGCOMM conference proceedings, workshops, and other SIGCOMM venues. *Workshop proceedings are more open towards new authors as compared to main proceedings.*

Venue	Sub-Venue	Percentage of New Authors	Percentage of Papers Published by New Authors
SIGCOMM	Main Proceedings	4.5	10.7
	Workshops	16.1	13.6
IMC	Main Proceedings	12.5	13.1
CoNext	Main Proceedings	14.2	15.9
ICN	Main Proceedings	17.3	23.8
E-Energy	Main Proceedings	12.1	17.2
SenSys	Main Proceedings	11.8	16.9
SoSR	Main Proceedings	12.8	19.8
LANC	Main Proceedings	15.9	21.3
HotNets	Main Proceedings	10.6	21.9
ANRW	Main Proceedings	14.7	22.7
ANCS	Main Proceedings	9.6	18.4

3.2 A Country Perspective

The above suggests that the country of origin may have an impact on an author's success. We next aggregate authors into their respective countries (as measured by home affiliations), and inspect country-based publishing trends.

Country Paper Count. Figure 6 presents the distribution of published articles across all conferences sponsored by SIGCOMM using a global heat map. As expected, the United States is in the highest position in terms of publication count. Other top countries include Canada, China, France, and the UK.

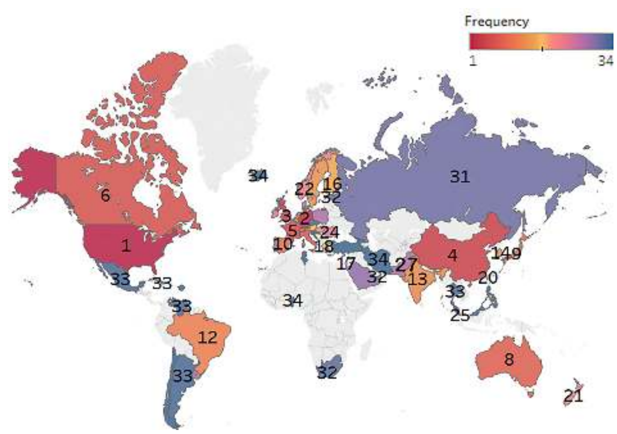

Figure 6: Publication count rank of different countries in SIGCOMM (the numbers represent the ranks). *The US, China and Western European countries have published the largest number of papers.*

Country Ranking. We proceed to compile a rudimentary ranking for each country, in terms of its productivity. Rather than solely relying on publication counts, we also include citation rates (taken from Scopus). The Normalized Rank Score (NRS) for each country can be calculated by using Equation 1 where P is publication count, C is citation count, hi is h-index of a country, P_{top} is maximum publication count, C_{top} is maximum citation count and hi_{top} is maximum h-index obtained by a country in a venue. We use h-index to avoid problems with raw or average citation counts [11]. Note that the h-index, mentioned here, is computed based on the publications and citations of a country in this paper's dataset.

$$NRS = \frac{1}{3} * \left(\frac{P}{P_{top}} + \frac{C}{C_{top}} + \frac{hi}{hi_{top}} \right) \quad (1)$$

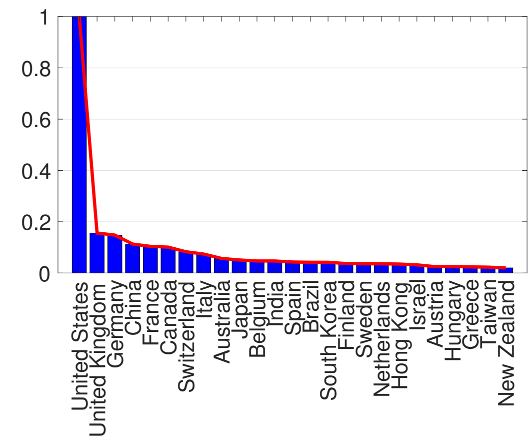

Figure 7: Rank of countries in SIGCOMM venues based on their publication count, citation count, and h-index. *The US is by far the top country by this measure.*

Figure 7 shows the ranking of countries, based on their scores. The US is ranked top across SIGCOMM venues during 1969–2018, with the UK, Germany, and China taking the subsequent positions. Although not depicted, we see that certain countries have been making dramatic improvements in their rankings too. For instance, both India and Brazil have increased their rankings by 3 and 5 positions over the last 10 years. India and Brazil are currently at the 12^{th} and 14^{th} positions in overall SIGCOMM venues based on productivity score (up from 15^{th} and 19^{th}, respectively). Whereas in Brazil this is primarily driven by LANC, India has also performed very well across multiple SIGCOMM venues. The former perhaps shows the importance of regional conferences in engaging countries.

3.3 An Institution Perspective

Although the previous section has explored authors on a regional basis, often individual countries contain a wide range of institutes. Therefore, we now aggregate authors by their home institutes and investigate the trends.

Institute Paper Count. Figure 8 shows the top institutes based on publication counts in SIGCOMM venues. We observe a clear dominance by a small set of major players. Most notably, prestigious US universities dominate the rankings; furthermore, universities from the UK, China, and Germany play a prominent role. We also note that research-based institutes have shown an impressive performance. For instance, AT&T Labs actually has had the most success in SIGCOMM venues, with Microsoft Research publishing heavily too.

Considering the prominence of these industrial research labs, we are curious to see how their involvement has evolved over time. Figure 9 shows the temporal development of publication counts at top research institutes. Bell Labs has been the longest contributor to SIGCOMM conferences, with a number of other labs starting to participate in the 1980s and 90s too. For example, HP, Intel, Facebook and Microsoft have emerged as rising stars and surpassed even AT&T in post-2003 publication counts. This effectively highlights the strong industry focus that SIGCOMM venues has had over the years.

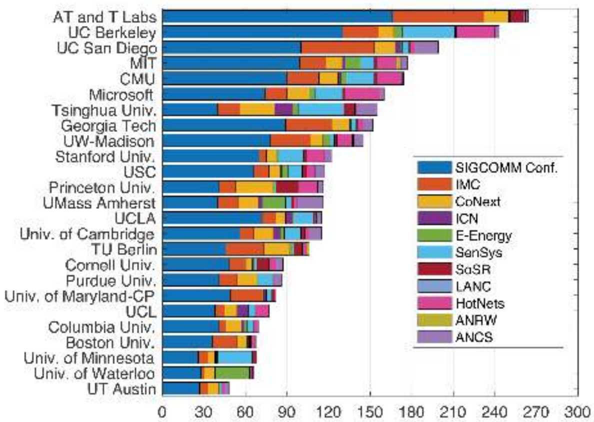

Figure 8: Top institutes in SIGCOMM based publication count during 1969–2018. ***Although universities are major contributors to SIGCOMM, research institutes are also top ranked.***

Institute Collaboration. As well as author-level collaboration, we are interested in exploring collaborative practices between institutes. Figure 10 presents a co-authorship graph between institutes; again, we compute modularity to identify communities (finding 15 clusters). Links are weighted by the number of publications co-authored by those nodes (institutes). The largest cluster contains four major research institutes and 9 key academic institutes: UC Berkeley, UCSD, Princeton, UIUC, USC, UW, ETH Zurich, Google, Facebook, Intel, and Microsoft. Similarly, AT&T Labs, Georgia Tech, UMich, and UW-Madison have also shown significant co-authorship patterns. Another major co-authorship pattern can be observed among UK based institutes, including the University of Cambridge, UCL, Lancaster University and several other European academic and research institutes. This confirms that geography plays a natural role in facilitating collaboration. That said, many Chinese institutes show more significant co-authorship patterns with US based institutes as compared to other Chinese institutes. This behavior may be because a large number of Chinese academics are alumni of US institutes.

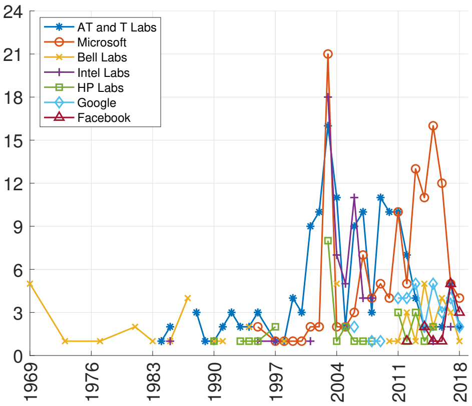

Figure 9: Top research institutes in all SIGCOMM venues based on publication count during 1969–2018 and their temporal development. Note that the line breaks where data is not available. ***AT&T is a major player in SIGCOMM venues and remaines the overall top contributor, but other research institutes (e.g., HP, Intel, Microsoft) have emerged as rising stars. Note that y-axis represents the number of publications of a research lab published across all venues from our dataset.***

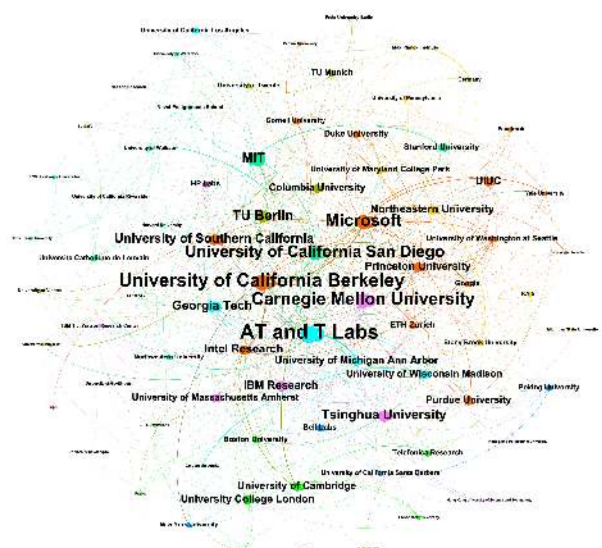

Figure 10: Co-authorship network among different institutes in SIGCOMM. Node size indicates the number of links with other nodes in the co-authorship network and the node color represents cluster membership. ***Key communities of collaborative institutes emerge, with geography playing a clear role.***

3.4 A Paper Perspective

Whereas the previous sections have focused on prominent authors and their affiliations, we next wish to inspect various attributes of the papers themselves.

Reference Patterns. First, we extract the references from all papers and create a citation graph, as we are curious to understand how SIGCOMM venues cite each other. Figure 11 is a Sankey diagram, showing the fraction of papers that SIGCOMM papers reference (left), as well as the other papers that in turn cite the SIGCOMM papers in our dataset (right). Note that this covers all SIGCOMM venues. Interesting patterns emerge from this analysis. Most noteworthy is the bias for citing papers from the same venue. For example, 26% of references in SIGCOMM papers are for other papers previously published in SIGCOMM. In contrast, a far more diverse body of papers list SIGCOMM publications in their references, particularly other conferences (57% of the papers in our dataset which cite SIGCOMM venues are actually conferences, rather than journals). Major citers of SIGCOMM papers include INFOCOM and LNCS (which subsumes many conference proceedings). This trend is perhaps intuitive as SIGCOMM is considered among the most prestigious outlets, and therefore it is unsurprising that a wide diversity of venues cite such papers.

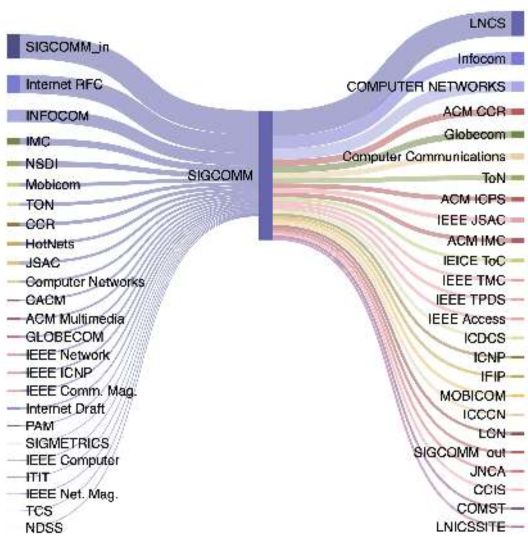

Figure 11: The distribution of references and citations in SIGCOMM. The left input shows the conferences that are referenced by SIGCOMM papers; the right output shows which papers cite SIGCOMM publications. *Major source of references and citations in SIGCOMM are from conferences.*

All that said, it is clear that a number of other publication venues feature heavily in the bibliographies of SIGCOMM papers, and these are dominated by conferences rather than journals. Even premier journals like IEEE/ACM Transactions on Networking are cited in SIGCOMM far less often than well known conferences like IMC, NSDI and INFOCOM: SIGCOMM contains 69.2% references from conferences, whereas it contains just 9.8% references from journals. Interesting SIGCOMM also contain 21% references from Internet RFCs, highlighting the strong historical links between the academic and standards communities.

Keyword Analysis. We are yet to touch upon the underlying topics of the papers being published by SIGCOMM venues. Keywords are a simple way to analyze changing dynamics of research interests. Table 5 presents the top occurring keywords in the various conference papers under-study. To date, it shows that the majority of proceedings have published research related to various aspects of networked systems. Naturally, there are subtle differences across the different venues though. For example, E-Energy discusses energy efficiency and optimization of systems, whereas SenSys discusses research related to embedded systems.

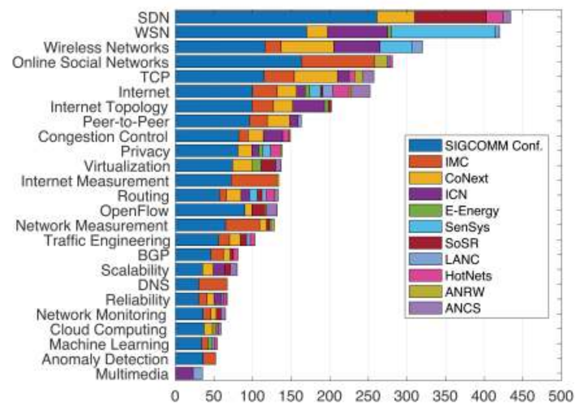

Figure 12: Top occurring keywords in SIGCOMM venues during 1969–2018. *The majority of the papers discuss networking protocols and techniques.*

Figure 12 shows the overall top keywords across all SIGCOMM venues during 1969–2018, as measured by simple frequency. Figure 14 then presents the keywords which receive the most citations.[7] The keywords are broadly similar across these two measures. In both cases, it is clear that the SIGCOMM main conference accumulates the most citations and papers, yet a number of specialist venues also stand out. For example, 32% of the 421 Sensor Network papers from SIGCOMM venues are published in SenSys, and 24% of the 431 SDN papers are published in SOSR. This highlights the importance of these more targeted conferences.

Table 6 further shows the evolving year-wise topics during 1969–2018. Those familiar with these conferences will likely recognize these trends. For instance, in the early years, authors published a lot of research related to architectures and network protocols. During 2000–2004, when cellular networks were growing in prominence, papers regularly discussed telecommunication, and more recently we have seen topics such as SDN coming to the fore. These trends suggest that authors are often well aligned with state-of-the-art technologies. We can also inspect how these topics are interrelated, as measured by their co-occurrence in papers. We compute a graph, consisting of all keywords (as nodes); two nodes are connected if

[7] Note that the conferences contributing to these keywords have also changed across the years.

Table 5: Top Occurring Keywords in SIGCOMM and its sponsored Proceedings during 1969–2018.

Venue	Top Keyword
SIGCOMM	Performance, Sensor networks, Security, Algorithms, QoS
IMC	Internet Protocols, Wireless Sensor Networks,Optimization, Telecom. Traffic, Social Networks
CoNext	Internet Protocols, WLAN, Mobile Telecommunication Systems, Network Architecture, TCP
ICN	Telecomm. Networks, Named Data Networking, QoS, Network Architecture, Information-centric Networking
E-Energy	Smart Power Grids, Energy Utilization, Energy Efficiency, Biomass, Electric Power Transmission Networks
SenSys	Embedded Systems, Wireless Sensor Networks, Internet Of Things, Energy Efficiency, Energy Harvesting
SoSR	Software Defined Networking, Network Function Virtualization, Data Planes, OpenFlow, Programmable Switches
LANC	QoS, Wireless Networks, Optimization, Network Routing, Bit Rates
HotNets	Data Centers, Design, Network Security, Network Management, Topology
ANRW	TCP, Computer Operating Systems, Internet Protocols, Network Measurement, Internet Of Things
ANCS	Network Architecture, Packet Networks, Packet Classification, Intrusion Detection, Field Programmable Gate Arrays (FPGA)

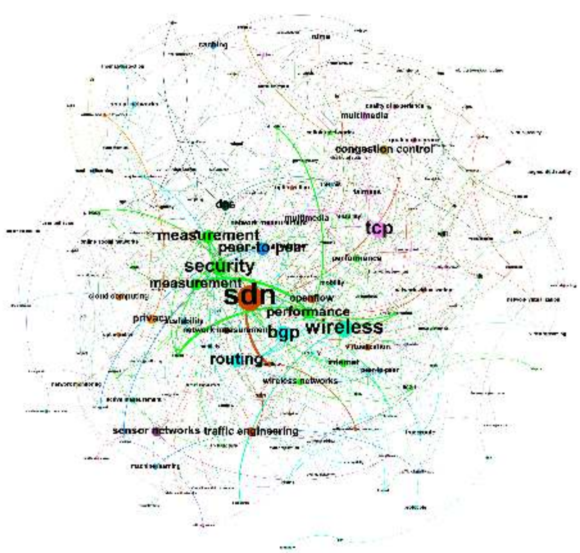

Figure 13: Graph of co-occurring keywords in SIGCOMM venues during 1969–2018. ***Intuitive keywords tend to cluster together, e.g., SDN and OpenFlow.***

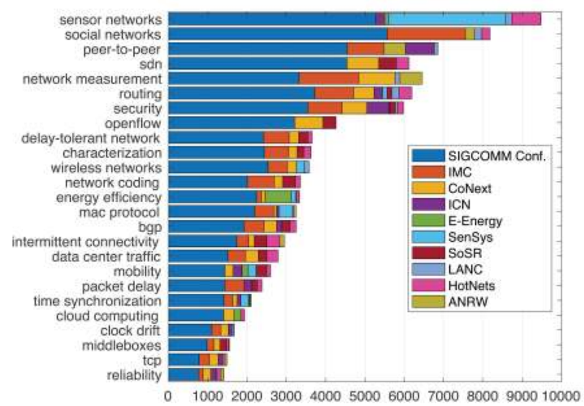

Figure 14: Most cited keywords in SIGCOMM venues during 1969–2018. ***The most frequently cited topic is Sensor Networks, with a mix of SIGCOMM and SenSys papers.***

they are associated with the same paper (and links are weighted by the number of papers with shared keywords). Figure 13 presents the graph, with links coloured by the community to which the nodes belong. We see that natural groupings emerge: *e.g.*, Quality of Service and Congestion Control are closely paired, whereas SDN is associated with keywords such as Scalability and OpenFlow.

4 CONCLUSIONS

We have presented a longitudinal study of publication trends across SIGCOMM venues. We have not limited ourselves to the flagship SIGCOMM conference but have accounted for all SIGCOMM-affiliated events, *e.g.*, IMC, CoNEXT, SenSys, HotNets. We have explored significant authors, institutes, and countries, and have inspected collaborative patterns among these different entities. Many of our results follow common intuition, *e.g.*, the US has outperformed all other countries in terms of productivity, and SIGCOMM venues experience assortativity, whereby they tend to cite themselves regularly. In-line with the conventional thinking, we observe that in computer networking, conferences (rather than journals) play a more critical role and tend to accumulate more citations. In addition, we have also observed some less intuitive findings: for instance, although the majority of papers *do* contain well-established authors, there is a surprising portion of emerging authors who have not previously published in SIGCOMM venues (or even published with other experienced authors before). We further hope that our analysis of important topics has provided useful insight to authors wishing to understand future directions. Our long term goal is to build a better understanding of the publishing culture in the data communications community. We have made our datasets publicly available and hope that others will find similar interest in this line of research.[8]

REFERENCES

[1] Gustavo Cattelan Nobre and Elaine Tavares. Scientific literature analysis on big data and internet of things applications on circular economy: a bibliometric study. *Scientometrics*, 111(1):463–492, 2017.

[2] João M Fernandes and Miguel P Monteiro. Evolution in the number of authors of computer science publications. *Scientometrics*, 110(2):529–539, 2017.

[8] https://github.com/waleediqbal411/CCR-paper-data2019

Table 6: Yearly top keywords in SIGCOMM venues

Year	Keywords
2018	Smart Power Grids, Embedded Systems, Multimedia Systems, Convolutional Codes, Internet Of Things
2017	Embedded Systems, Energy Efficiency, Smart Power Grids, Internet Of Things, Convolutional Codes
2016	Convolutional Codes, Embedded Systems, SDN, TCP, QoS
2015	Energy Efficiency, Smart Power Grids, Energy Utilization, Complex Networks, Wireless Sensor Networks
2014	Embedded Systems, Software-defined Networking, Complex Networks, Social Networking (online), Energy Utilization
2013	Wireless Sensor Networks, Complex Networks, SDN, Openflow, Optimization
2012	Network Architecture, Energy Efficiency, Data Centers, Social Networking (online), Openflow
2011	Wireless Sensor Networks, Embedded Systems, Energy Efficiency, Optimization, Distributed Computer Systems
2010	Embedded Systems, Wireless Sensor Networks, Network Protocols, Electric Network Topology, Computer Operating Systems
2009	Embedded Systems, Wireless Sensor Networks, Network Security, Social Networking (online), Peer To Peer Networks
2008	Embedded Systems, Convolutional Codes, Wireless Sensor Networks, Game Theory, Social Networks
2007	Wireless Sensor Networks, Internet Measurements, Semiconducting Intermetallics, Distributed Computer Systems, Web Services
2006	Wireless Sensor Networks, Distributed Computer Systems, Algorithms, Virtual Reality, Scalability
2005	Wireless Telecommunication Systems, Wireless Sensor Networks, Distributed Computer Systems, Delay Tolerant Networks, Testbeds
2004	Network Protocols, Telecommunication Traffic, Servers, Bandwidth, Mathematical Models
2003	Mathematical Models, Bandwidth, Telecommunication Traffic, QoS, Congestion Control (communication)
2002	Telecommunication Traffic, Digital Watermarking, Security Of Data, Topology, Servers
2001	Quality Of Service, Multimedia Systems, Algorithms, Bandwidth, Multicasting
2000	Multimedia Systems, Algorithms, Network Protocols, Telecommunication Services, User Interfaces
1999	Congestion Control (communication), Telecommunication Traffic, Mathematical Models, Multicasting, Algorithms
1998	Multimedia Systems, QoS, Semantics, Information Retrieval, Internet Protocols
1997	Multimedia Systems, Digital Storage, Indexing (of Information), Bandwidth, Content Based Retrieval
1996	Performance, Mobile Telecommunication Systems, Optimization, Congestion Control (communication), Asynchronous Transfer Mode
1995	Audio Systems, Computer Graphics, Bandwidth, Information Services, Telecommunication Services
1994	Memory Architecture, Network Routing, Topology, Address Space, Credit-based Flow Control
1993	Multimedia Computing, Video Signal Processing, Data Handling, Information Retrieval Systems, Synchronization
1992	Packet Switching, Telecommunication Control, Communication Protocols, High Speed Networks, Congestion Control
1991	Asynchronous Transfer Mode, Service Disciplines, Traffic Congestion, Graph Theory, Propagation Delays
1990	Packet Switching, Broadband Networks, Design Principles, Dual Bus, Gateways (computer Networks)
1989	ISDN, Graph Theory, Asynchronous Transfer Mode, Data Transmission, Open Systems Interconnection
1988	Topology, Bandwidth, Local Area Networks, Optical Communication, Congestion Avoidance
1987	Distributed Computer Systems, Natural Resources Management, Resource Allocation, Supercomputers, Back-bone Network
1986	Distributed Systems, Transport Protocols, Local Area Networks, Access Control, Application Programs
1985	Distributed Systems, Transport Protocols, Back-bone Network, Congestion Avoidance, Data Transmission
1984	Petri Nets, Data Transfer, Local Area Networks, Open Systems Interconnections, Reference Modeling
1983	Packet Switching, Access Control, Data Transfer, Gateways (computer Networks), Interconnection Networks (circuit Switching)
1982	Queueing Network Model, Data Link Control, Flow Control, Information Management, Performance Analysis
1981	Network Performance, Queueing Network Model, Complex Networks, Data Link Control, Flow Control
1980	Convolutional Codes, Packet Switching, Open Systems Interconnections, Reference Modeling, Authentication
1979	Data Handling, Information Management, Interconnection Networks (circuit Switching), Electronic Data Interchange, Gateways (computer Networks)
1978	Switching Networks, A-stable, Antenna Arrays, Asynchronous Data, Automatic Repeat Request
1977	Arpanet, Operating System Design, Queueing Theory, Satellite Broadcast, Slotted Aloha
1976	Average Delay, Band-width Utilization, Centralized Networks, Closed Loop Control Systems, Common Carriers
1975	Convolutional Codes, Data Communication Systems, Packet Switching, Packet Networks, Switching Networks
1974	Convolutional Codes, Switching Networks, Distributed Computer Systems, Digital Storage, Multiplexing Techniques
1973	Packet-switched, Adaptive Routing, Batch Data Processing, Branch Bound Method, Buffer Management
1969	Convolutional Codes, Switching Networks, Distributed Computer Systems, Digital Storage, Multiplexing Techniques

[3] Alexander Serenko, Nick Bontis, and Joshua Grant. A scientometric analysis of the proceedings of the McMaster world congress on the management of intellectual capital and innovation for the 1996-2008 period. *Journal of Intellectual Capital*, 10(1):8–21, 2009.

[4] Dah Ming Chiu and Tom ZJ Fu. Publish or perish in the internet age: a study of publication statistics in computer networking research. *ACM SIGCOMM Computer Communication Review*, 40(1):34–43, 2010.

[5] Periyaswamy Rajendran, R Jeyshankar, and B Elango. Scientometric analysis of contributions to journal of scientific and industrial research. *International Journal of Digital Library Services*, 1(2):79–89, 2011.

[6] S Nattar. Indian journal of physics: A scientometric analysis. *International Journal of Library and Information Science*, 1(4):043–61, 2009.

[7] Zhifeng Yin and Qiang Zhi. Dancing with the academic elite: a promotion or hindrance of research production? *Scientometrics*, 110(1):17–41, 2017.

[8] Waleed Iqbal, Junaid Qadir, Gareth Tyson, Adnan Noor Mian, Saeed-ul Hassan, and Jon Crowcroft. A bibliometric analysis of publications in computer networking research. *Scientometrics*, pages 1–35, 2019.

[9] Péter Jacsó. Google scholar: the pros and the cons. *Online information review*, 29 (2):208–214, 2005.

[10] Phillip Bonacich and Paulette Lloyd. Eigenvector-like measures of centrality for asymmetric relations. *Social networks*, 23(3):191–201, 2001.

[11] Per O Seglen. Why the impact factor of journals should not be used for evaluating research. *Bmj*, 314(7079):497, 1997.