Studying the Publication Pattern of Canadian Computer Scientists / Étude des pratiques de publication des scientifiques canadiens en informatique

Li Zhang

Canadian Journal of Information and Library Science, Volume 39, Number 1, March/mars 2015, pp. 60-78 (Article)

Published by University of Toronto Press
DOI: *https://doi.org/10.1353/ils.2015.0002*

Studying the Publication Pattern of Canadian Computer Scientists

Étude des pratiques de publication des scientifiques canadiens en informatique

Li Zhang
Science and Engineering Libraries, University Library, University of Saskatchewan
li.zhang@usask.ca

Abstract: This study explored the publication pattern of Canadian computer scientists and compared the impact of conference papers and journal articles published by these researchers. It was found that conference proceedings are the preferred venue for scholarly communication, but the impact of conference papers is not comparable to that of journal articles. The potential reasons for the lower impact of conference papers are discussed, and possible approaches to improve the current scholarly communication system in computer science are proposed.

Keywords: conference proceeding, journal article, computer science, Canada, bibliometric analysis

Résumé : Cette recherche explore les pratiques de publication des scientifiques canadiens en informatique et compare l'impact des présentations à des colloques et des articles de revues savantes publiées par ces chercheurs. Nous avons constaté que les actes de colloque sont le lieu privilégié pour la communication scientifique, mais que l'impact des présentations à des colloques n'est pas comparable à celui des articles de revues. Nous discutons des raisons possibles de l'impact plus faible des présentations à des colloques, et nous proposons des approches possibles pour améliorer le système actuel de communication savante en informatique.

Mots-clés : actes de colloque, articles de revue savante, science informatique, Canada, analyse bibliométrique

Introduction

In the field of computer science, conference proceedings are an important venue for scholarly communication, and many researchers even prefer to publish their research findings in conference proceedings rather than in traditional journals (Eckmann, Rocha, and Wainer 2012). Several factors have contributed to the development of this unique scholarly communication practice. Computer science is one of the fastest-developing fields, and thus the speed of dissemination of research findings is important for researchers. As stated by Wainer, Przibisczki de Oliveira, and Anido (2011, 135), conference proceedings are usually published faster than traditional journals, and the increased speed allows researchers to be competitive in this changing field. The in-person interactions at

conference venues enable researchers to directly reach other colleagues in the same field from different geographical locations and receive immediate feedback on their research, thus further improving the research and increasing the opportunities for collaboration (Laplante et al. 2009, 184). Another factor that contributes to the popularity of conference proceedings is the nature of computer science research. Many of the computer science research studies, such as studies on interactive video, are more suited for a conference presentation instead of a static journal publication, as indicated in an article by Drott (1995).

Although conference proceedings represent a large part of the literature in computer science, the quality of conference papers has been controversial. In this present study, *conference proceeding* refers to a collection of conference papers, and *conference paper* to the individual document included in a conference proceeding. In the last decade, several studies have explored the impact of conference proceedings in computer science from different perspectives as compared to traditional journals. Many of these studies used bibliometric indicators to assess the quality of conference papers, though this method for research evaluation has its inherent limitations (Weingart 2005). Commonly used bibliometric indicators include publication counts, citation counts, *h*-index, and journal impact factor. Below is a review of the recent findings on this topic.

Chen and Konstan (2010) compared the impact of journal articles and conference papers published in the Association for Computing Machinery (ACM) Digital Library as of 2007. They found that about 58% of the total publications were conference papers, and 42% were journal articles. When the citation counts of the two types of documents were compared, it was found that, based on citation data in ACM Digital Library, conference papers received an average of 2.15 cites per document in two years, compared to 1.53 cites per journal article. However, when four-year citation counts were used, journal articles outperformed conference papers (citation counts were 4.03 per document for journal articles and 3.16 for conference papers). The study results implied that while conference papers had a higher immediate impact, journal papers had a higher long-term impact.

Instead of studying the impact at the individual-document level, several studies have compared the quality of the two types of publications at the collective level: comparing the impact factor of computer science journals and conference proceedings. For example, Freyne et al. (2010) compared the Google Scholar impact factors of 15 journals and 15 conference proceedings in the artificial intelligence and machine learning sub-fields. It was found that while the top ranks of this 30-publication collection were mostly occupied by journals, the leading conference proceedings achieved comparable impact factors to the mid-ranking journals. Another study of publications in the sub-field of database research found that while there were many conference papers with relatively little scientific impact, the Google Scholar impact factors of the two top conference proceedings were significantly higher than those of the two top journals (Rahm 2008).

The above studies seem to suggest that conference papers had comparable scientific impact to journal articles, at least for top-ranking conferences. However, conflicting results were also reported in studies using different samples and methods.

Based on data from Web of Science with Conference Proceedings Citation Indexes, Bar-Ilan (2010) studied the publication pattern of highly cited computer scientists. It was found that on average only 39% of the total publications were conference papers, whereas 52% were journal articles. This result is different from Chen and Konstan's (2010) findings, indicating that these highly cited researchers actually preferred to publish in traditional journals. This study used a relatively new citation indicator, *h*-core, to compare the impact of the two types of documents. The *h*-core is the collection of papers included in the calculation of the *h*-index (Burrell 2007), and it can be viewed as the most influential works of any given author. For example, if a researcher has an *h*-index of 10, those 10 papers are called his *h*-core. The results of Bar-Ilan's (2010) study showed that, within the average *h*-core of 19.4, only 12.7% of the *h*-core documents were conference papers, meaning that most of the highly influential papers of these computer scientists were journal articles, and only a small proportion were conference papers.

Franceschet (2010) studied the publications on the 10 most popular computer science topics from Web of Science. He found that 78% of the publications were conference papers, and only 22% were journal articles. However, journal articles had a much higher impact as measured by citations per document (5.41 citations per document for journal articles and 0.71 citations for conference papers). The higher *h*-index of the journal article collection also confirmed this (the *h*-index was 27 for the journal article collection, 16 for the conference paper collection).

In his later research studying the citation distribution of computer science publications, Franceschet (2011) compared the citations of articles published in journals in the computer science categories of Journal Citation Reports with those of conference papers from 10 research-intensive countries. It was found that journal articles received an average of 11.69 citations, and an *h*-index of 106, but those numbers for conference papers were 4.71 and 65, respectively. However, Franceschet found that the citation distribution of computer science publications was heavily skewed and concluded that the mean would not be the best measurement for citation counts. The citation data were also from the Web of Science.

In addition to citation indicators, conference paper acceptance rate is another parameter often used as a proxy for the quality of conference papers. It is commonly assumed in the field of computer science that a paper presented at a conference with a lower acceptance rate would be of higher quality. However, studies testing this assumption also yielded inconsistent results. Based on the citation data from the ACM Digital Library, Chen and Konstan (2010) found that conference papers presented at highly selective conferences (30% acceptance rate or less) received significantly more citations than journal articles and concluded that papers published in these conference proceedings should be treated

as high-quality research. However, another study based on citation data from Google Scholar found that lower acceptance rates did not necessarily translate into higher impact of the individual papers and that the correlation between conference paper acceptance rate and citation impact was weak, arguing that using the conference acceptance rate as a proxy for conference paper quality was questionable (Freyne et al. 2010).

Objectives

The above discussion indicates that the quality of conference papers has been of concern to the computer science community. Although different data sources, samples, and measurements (bibliometric indicators and conference paper acceptance rate) have been used, no consensus regarding the quality of conference papers has been reached. Different countries may have different merit systems and research evaluation criteria, and these factors will influence researchers' decisions on where and how to publish their research findings (Wainer, Goldenstein, and Billa 2011, 146). Therefore, it is necessary that samples from different countries be studied to achieve a more complete understanding of the culture of scholarly communication in computer science.

To the best of our knowledge, none of the existing studies has focused on the publications of Canadian computer scientists. Therefore, the present study will explore the impact of conference papers by Canadian computer scientists. Specifically, the following research questions will be investigated:

- What is the proportion of conference papers and journal articles in computer science in Canadian production?
- What are the publishing trends over the years, if any?
- Are there differences between proceeding papers and journal articles in their document structure?
- Are there differences between the two types of publications in their impact?

Methods

For this study 30 of the 496 faculty members in the computer science departments, schools, or the equivalent from 15 Canadian Medical Doctoral universities from Maclean's University Rankings (Maclean's 2012) were randomly selected by random number generation. There are many university rankings available, and the rankings of universities are slightly different in each system. Maclean's University Rankings is probably the best-known university ranking in Canada; therefore, we decided to focus our study on these 15 research universities. These universities are generally considered the most research-intensive universities in Canada, and thus it is believed that the publications of the faculty in the computer science departments in these universities will provide a snapshot of research activities in this field in Canada. A detailed description of the selection process is described elsewhere (Zhang 2014).

The Scopus database was used as the data source to identify publications and citations by each of the 30 sample faculty. When deciding on the data

source for this study, we mainly considered three sources: Scopus, Web of Science, and Google Scholar. Previous studies have demonstrated that Scopus includes more documents than Web of Science in the computer science field; therefore, searching Scopus would allow us to create a more comprehensive list of the publications of this sample of researchers (Wainer, Goldenstein, and Billa 2011; Zhang 2014). Although Google Scholar has become a popular source for scientific information, there are concerns about using it as a reliable source for citation analysis. Some of the major concerns with Google Scholar include its susceptibility to citation manipulation (Delgado López-Cózar, Robinson-García, and Torres-Salinas 2014; Jacsó 2011), lack of quality control of the citations (Aguillo 2012), lack of transparency of its source base (Wouters and Costas 2012), and the work required to clean data from Google Scholar (Jacsó 2006). Another reason to use Scopus is because of its "Author Search" function. This function automatically searches for name variations and assigns each author an author ID, which makes the process of identifying publications by a specific researcher relatively easy. However, 5 of the 30 researchers were listed under different author IDs. These authors were further examined, and the publications under different IDs were combined if it was confirmed that these IDs were for the same author.

The publication list of each faculty member was downloaded for detailed analysis. The number of publications by document type was recorded as per Scopus document type: conference paper, journal article, editorial, review, letter, and so on. Scopus does not assign a document type for articles in press; therefore, the current author examined these documents and assigned a document type based on the author's judgement.

The quality and impact of the conference papers and journal articles were compared from two perspectives:

Publication structure

We adopted the method used by González-Albo and Bordons (2011) to evaluate the structure of the publications. The publication structure of each document includes the number of authors, the number of author affiliations, the number of references, and the number of pages. A growing demand for research quality has led to an increase in research complexity, for example, broader research scope and more in-depth analysis of research results. Complex studies usually require knowledge and skills from different areas; thus, collaboration among authors and institutions with different sets of expertise is almost necessary to accomplish these projects (Bordons and Gómez 2000). On the other hand, a higher number of references and higher page counts signal greater scholarly efforts and completeness in the investigation, and these papers are likely to achieve greater scientific impact (McVeigh and Mann 2009, 1107–8; Seglen 1997, 1054). Therefore, in this method, the publication structure is used as an indirect indicator of the complexity and the completeness of publications.

Scopus did not directly provide article structure data. A computer command program was created to calculate these data from the bibliometric information of

Table 1: The institutions of the 30 sample faculty members

Institution	No. of researchers selected
Dalhousie University	4
Laval University	1
McGill University	2
McMaster University	1
Queen's University	1
University of Alberta	1
University of British Columbia	4
University of Calgary	3
University of Montreal	4
University of Ottawa	3
University of Saskatchewan	3
University of Toronto	2
University of Western Ontario	1
Total	30

each record. In cases where Scopus did not provide page information or reference information for a document, the author searched a variety of other sources to identify the missing information. If the information for a record could not be found, it was excluded from the calculation.

Citation indicators

The number of citations each document received and the *h*-index of each researcher were recorded. The composition of the *h*-core of each researcher by document type was also recorded. The definition of *h*-core can be found in the introduction above.

Data was collected during January–April 2014.

Results

Table 1 lists the number of researchers from each institution that were randomly selected. Note that although there are 15 research universities in Canada, the random process selected the 30 sample researchers from only 13 universities. Of the 30 sample researchers, 1 was an assistant professor, 8 were associate professors, and 18 were full professors. The position titles of 3 faculty members were unknown, but it was believed that they were in the professoriate rank.

Composition of the publications

A total of 2,402 publications were identified for the 30 computer science researchers, of which 1,446 (60%) were conference papers, 841 (35%) were journal articles, and 115 (5%) were other types, including editorials, reviews, letters, and so on (see Figure 1). The average number of publications of the 30 faculty was 80, the average number of conference papers of these researchers was 48, and the average number of journal articles was 28. The number of publications of each researcher by document type is listed in Table 2. To protect the confidentiality of the researchers, the name of each faculty member is represented by

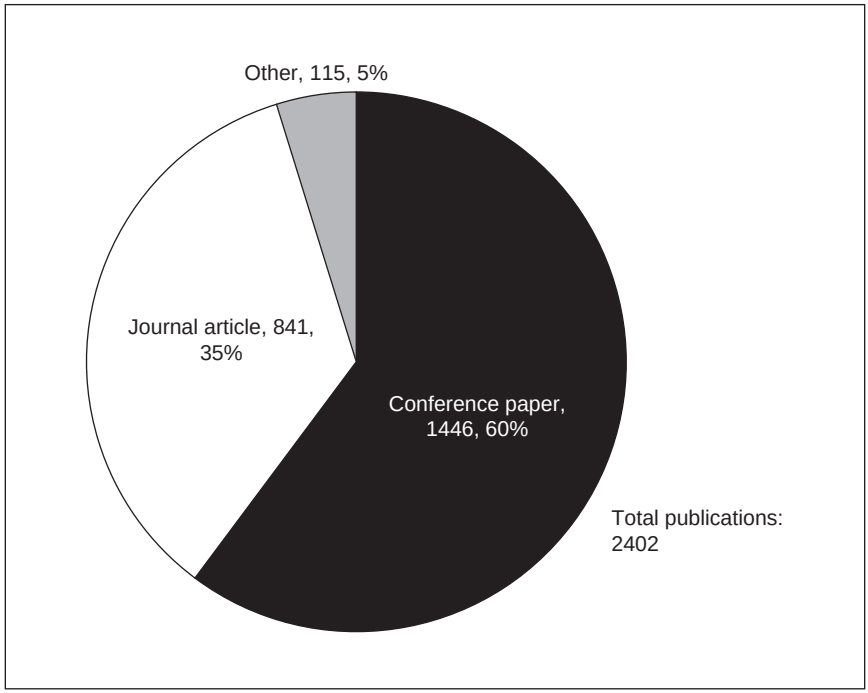

Figure 1. Composition of publications by document type

a researcher ID. The majority of the researchers published more in conference proceedings than in journals, with the exception of only 5 researchers (Researchers 5,7, 8, 12, and 26). Because of the small representation of assistant professors in the sample, one might think that they have been acculturated to publishing predominantly in conference proceedings. However, a closer look at those 5 researchers who published more in journals found that the only assistant professor (Researcher 5) turned out to be in this category, along with 2 associate professors and 2 full professors.

Publication trends

Figure 2 shows the publication trends over the years. The first journal article by these Canadian computer scientists was published in 1974, and the first conference paper was published in 1982. This indicates that computer science is still a relatively new field of science, compared to other traditional science fields such as physics, chemistry, and biology. Before 1995 the number of publications was fewer than 100 for each of the five-year periods. After 1995 we saw the biggest increase rate in the number of publications: the numbers in 1996–2000 and

Table 2: The number of publications by each researcher

Researcher ID	Total no. of publications	Conference papers		Journal articles	
		No.	%	No.	%
Researcher 1	115	95	82.61	20	17.39
Researcher 2	105	76	72.38	22	20.95
Researcher 3	47	28	59.57	18	38.30
Researcher 4	56	40	71.43	14	25.00
Researcher 5	5	1	20.00	4	80.00
Researcher 6	149	79	53.02	58	38.93
Researcher 7	20	7	35.00	12	60.00
Researcher 8	63	23	36.51	36	57.14
Researcher 9	48	38	79.17	10	20.83
Researcher 10	45	39	86.67	6	13.33
Researcher 11	159	82	51.57	53	33.33
Researcher 12	58	5	8.62	48	82.76
Researcher 13	110	69	62.73	39	35.45
Researcher 14	65	54	83.08	7	10.77
Researcher 15	136	96	70.59	38	27.94
Researcher 16	54	26	48.15	22	40.74
Researcher 17	44	27	61.36	17	38.64
Researcher 18	151	29	19.21	121	80.13
Researcher 19	192	114	59.38	64	33.33
Researcher 20	41	34	82.93	4	9.76
Researcher 21	40	37	92.50	3	7.50
Researcher 22	16	9	56.25	4	25.00
Researcher 23	69	49	71.01	19	27.54
Researcher 24	21	14	66.67	6	28.57
Researcher 25	111	65	58.56	42	37.84
Researcher 26	26	4	15.38	19	73.08
Researcher 27	44	31	70.45	12	27.27
Researcher 28	102	60	58.82	37	36.27
Researcher 29	27	21	77.78	2	7.41
Researcher 30	283	194	68.55	84	29.68
Total	2,402	1,446	N/A	841	N/A
Mean	80.1	48.2	59.33	28.0	35.49
SD	11.3	7.4	22.05	5.0	21.45

Note: The sum of journal articles and conference papers may not add up to the total number of publications because some publications were of other document types (e.g., editorials, reviews, letters).

2001–2005 each doubled that in the previous period. Between 2006 and 2010, the number of publications still increased significantly but at a somewhat slower rate. Note that the number of publications in the period 2011–2014 is lower than that in the period 2006–2010. This is because 2011–2014 is only a four-year term, and documents published in this period might not yet have been added to the database owing to an indexing delay. Figure 3 illustrates the evolution of the percentages of journal articles and conference papers. As can be seen, the percentage of journal articles was higher than that of conference papers during each of the five-year periods before 2000. However, the trend has changed since the start of the new century, and the percentage of conference papers

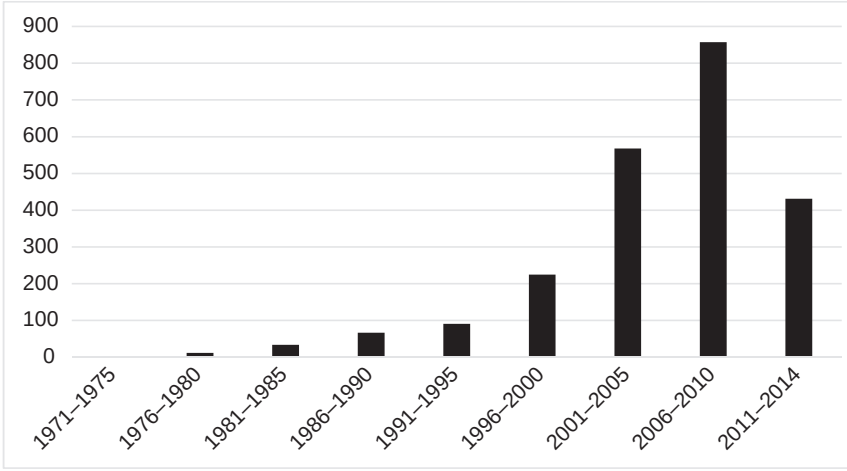

Figure 2. Publication trend of total number of publications (journal articles and conference papers only; other types of documents were excluded)

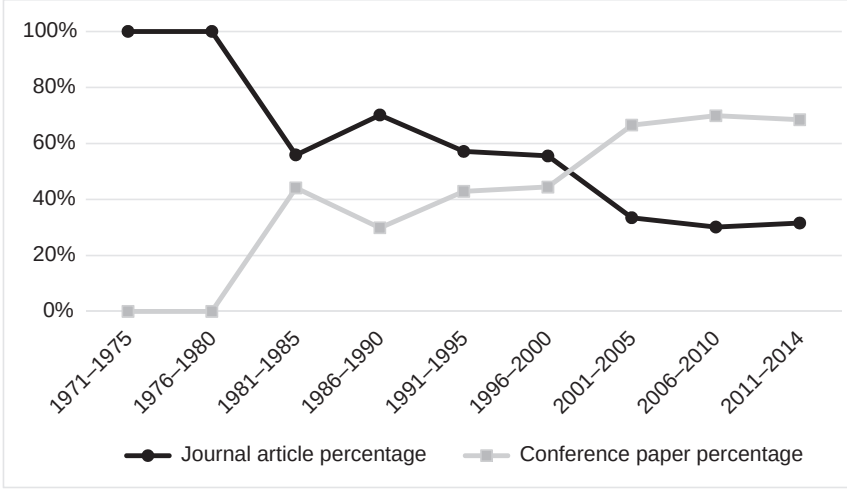

Figure 3. Evolution of the percentages of journal articles and conference papers in total publications

outperformed that of journal articles since then. The percentage of conference papers remains steady at about 65%, while that of journal articles is about 35%.

Publication structure

The publication structure of the documents published by these computer science researchers is shown in Table 3. The average number of authors of each journal article is 3.13, and that number for conference papers is 3.12. A paired

Table 3: Comparison of document structure

	Journal articleaverage (SD)	Conference paperaverage (SD)
No. of authors/doc	3.13 (0.78)	3.12 (0.77)
No. of affiliations/doc	1.90 (0.48)	1.62 (0.43)
No. of pages/doc	15.25 (3.91)	9.21 (3.25)
No. of references/doc	26.24 (8.93)	19.70 (6.57)

Table 4: Comparison of citation counts by document type

	Journal article	Conference paper
Average number of times cited/doc (SD)	14.38 (16.37)	4.66 (5.27)
Median number of times cited/doc	4	1
h-Index of the collection	51	33
No. of docs with times cited ≥ 50 (%)	54 (6.4%)	17 (1.2%)
No. of docs with times cited ≥ 100 (%)	25 (2.9%)	8 (0.6%)
No. of docs with times cited = 0 (%)	175 (20.6%)	619 (42.8%)
Maximum times cited	507	754

t-test shows that the difference is not significant ($p = .95$). The average number of affiliations per document is 1.9 and 1.62, the average page counts are 15.25 and 9.21, and the average number of references used in each document is 26.24 and 19.70 for journal articles and conference papers, respectively. Journal articles outperformed conference papers in all three of these categories, and the differences are significant ($p < .003$ for each variable).

Citation indicators

There is a significant difference between journal articles and conference papers in the average citation counts per document ($p < .001$). On average, each journal article was cited 14.38 times, whereas each conference paper was cited only 4.66 times.

To find whether there are differences in the percentages of highly cited documents for each publication type, we compared the number of documents cited at least 50 times and at least 100 times. Of the journal articles, 54 (6.4%) were cited more than 50 times, as compared to 17 (1.2%) for conference papers; 25 (2.9%) journal articles were cited at least 100 times, but that number was only 8 (0.6%) for conference papers. About 20% of the journal articles were never cited; however, that percentage was significantly higher for conference papers (42%). The higher *h*-index of journal articles (51) also confirmed that journal articles had a higher impact than conference papers, which had an *h*-index of 33. Interestingly, the most cited conference paper received 754 citations, 247 more than the most cited journal article, which received 507 citations. The citation data details can be found in Table 4.

We compared the page counts and number of references of the journal articles and the conference papers that were cited at least 50 times, hoping to

Table 5: Citation data of the sample computer science researchers' conference papers and journal articles

Researcher ID	*h*-Index	% of CPs in *h*-core	% of JAs in *h*-core	Total citations to *h*-core	% of citations to *h*-core from CPs	% of citations to *h*-core from JAs
Researcher 1	8	50.00	50.00	161	36.65	49.69
Researcher 2	11	18.18	72.73	256	53.91	43.75
Researcher 3	8	12.50	87.50	157	44.59	52.87
Researcher 4	8	50.00	50.00	133	57.14	35.34
Researcher 5	3	0.00	100.00	24	45.83	45.83
Researcher 6	33	39.39	60.61	4,177	56.31	39.84
Researcher 7	9	11.11	88.89	431	69.37	25.52
Researcher 8	19	0.00	84.21	833	37.94	60.38
Researcher 9	14	64.29	35.71	1,865	56.94	39.36
Researcher 10	7	85.71	14.29	65	55.38	38.46
Researcher 11	12	16.67	83.33	300	43.67	54.33
Researcher 12	22	0.00	95.45	1,100	9.45	86.73
Researcher 13	22	31.82	63.64	1,049	57.01	40.61
Researcher 14	13	92.31	0.00	334	71.56	25.45
Researcher 15	12	33.33	66.67	695	46.76	46.91
Researcher 16	8	12.50	87.50	165	48.48	43.03
Researcher 17	5	60.00	40.00	56	58.93	39.29
Researcher 18	23	13.04	86.96	1,160	26.47	70.09
Researcher 19	19	63.16	36.84	778	60.41	38.30
Researcher 20	3	66.67	33.33	128	53.91	44.53
Researcher 21	8	87.50	12.50	274	77.01	20.44
Researcher 22	3	33.33	33.33	42	59.52	38.10
Researcher 23	16	62.50	37.50	563	59.50	37.83
Researcher 24	7	14.29	71.43	124	34.68	53.23
Researcher 25	19	21.05	73.68	869	37.74	52.82
Researcher 26	7	14.29	71.43	92	43.48	51.09
Researcher 27	5	40.00	60.00	64	43.75	48.44
Researcher 28	11	45.45	54.55	267	55.43	38.20
Researcher 29	6	83.33	16.67	123	70.73	26.02
Researcher 30	19	42.11	52.63	696	54.31	42.10
Average	12	38.81	57.38	566	50.90	44.29
SD	7.2	27.93	26.96	811	14.05	13.29

Note: CP = conference paper; JA = journal article. The sum of the percentages of the journal articles and conference papers in the *h*-core may be less than 100% because other document types might be present in the *h*-core.

determine whether there were any similarities in the document structure of these highly cited documents. The average page counts of these journal articles were 18.83, compared to 10.70 for the conference papers; the average number of references was 27.48 for these journal articles and 18.53 for these conference papers. Based on *t*-tests, the differences were significant (p = .004, and .05, respectively), meaning that the highly cited journal articles demonstrated a higher degree of completeness than the highly cited conference papers, as indicated by the higher page counts and the larger number of references.

Table 5 displays the *h*-index of each researcher, the percentage of conference papers and journal articles in the *h*-core, total citations to the *h*-core, and the percentage of citations to the *h*-core from conference papers and journal articles. The *h*-index of these researchers ranges from 3 to 33, with an average *h*-index of 12 (SD 7.2). Three researchers' *h*-cores consisted of journal articles only, and one researcher's *h*-core consisted of conference papers only. Out of the documents in the *h*-core, on average 58% were journal articles, and 39% were conference papers ($p = .07$). The results indicate that most of the influential works of these researchers were in the form of journal articles, although conference papers also had a strong presence in the *h*-core.

When considering the citations to the *h*-core, it was found that, on average, 50.9% of the citations to *h*-core were from conference papers, while 44.3% were from journal articles. The difference is not significant ($p = .19$). This means that the most influential works of these computer scientists were cited heavily in both conference papers and journal articles.

Republication

In many disciplines, conference papers are often seen as the first or preliminary publication of research results. Results published in a conference proceeding can later be extended as a journal article. To study the extent of republication in computer science, we tested the republication rate of the sample researchers. Because of resource limitations, we applied a simplified approach that Bar-Ilan (2010, 820) used to identify republications: a journal article is a republication of a conference paper if they have identical or almost identical titles. It was found that out of the total 2,402 publications, only 61 were republications, a rate of 2.5%.

Discussion

This article studied the publication pattern of Canadian computer scientists and compared the impact of conference papers and journal articles published by these researchers. As expected, the results confirm that conference proceedings are the preferred venue for scholarly communication for Canadian computer scientists, as evidenced by the fact that 61% of the total publications were conference papers, whereas 35% were journal articles for these sample researchers. The study results also find that Canadian computer scientists often use conference proceedings as the final venue for disseminating their research results, as indicated by the low republication rate. The republication rate of 2.5% for this sample is lower than that from Bar-Ilan's (2010) study, which detected a republication rate of 8.5% using the same method. However, when applying a more accurate method for detecting republication, Eckmann, Rocha, and Wainer (2012) found that 30% of the papers in the top three computer vision journals were republications of conference papers. The simplified method used in our study considered only the titles of the documents; therefore, it is quite possible that some republications were not identified if the authors used different titles. Thus, the actual republication rate may be higher than the rate found in this

study. But still, even if our sample had a republication rate of 30%, which might be unlikely, conference proceedings continue to be used for disseminating final research results.

Although conference papers dominate the scholarly publications in computer science, their overall quality or impact may not be comparable to that of journal articles. In this study, the quality and impact were analysed from two perspectives: publication structure and citation indicators. The results show that journal articles exhibited a higher number of author affiliations, higher page counts, and higher reference counts, which indirectly indicate that research published in journal articles may be more complex and more complete, and thus extra efforts may be needed to have research published in journals. From the perspective of citation indicators, journal articles also outperformed conference papers in almost all the dimensions, including the average and median number of times a document was cited, the *h*-index, and the number of documents cited 50 and 100 times, and had a significant lower rate of non-citation (Table 4). The study results suggest that computer scientists may want to consider publishing more in journals if they would like to maximize the impact of their research. However, computer science research is roughly composed of two research areas: theory and experimentation (Computing Research Association 1999). Researchers in the two areas may have different publishing traditions. The work of theoreticians is more related to mathematics, and they may already have had a publishing tradition that resembles that of other science disciplines: publishing more in journals. On the other hand, experimentalists focus their research on creating computational artefacts, and they may be more likely to disseminate their research results via conference presentations. Because we did not collect the research areas of the sample, it was not possible to differentiate the publishing pattern of the theoreticians from that of the experimentalists. Therefore, it is unknown whether the suggestion presented above would apply only to a specific group of researchers or to all researchers. Further research needs to be conducted in this regard.

When the composition of the most influential works of a researcher—that is, the *h*-core—was analysed, it was found that both types of documents contributed heavily to the *h*-core although journal articles still made up a higher percentage of the *h*-core than did conference papers. This finding is consistent with Rahm's (2008) finding that while there are many conference papers of comparatively low scientific impact, the impact of the top conference papers of a researcher is comparable to that of top journal articles.

Several reasons might account for the lower citation impact of conference papers. First, it might be that the overall quality of conference papers is indeed lower than that of journal articles. However, a few other factors may also affect the citation data of conference papers. Conference proceedings have different dissemination approaches from traditional journals. Conference organizers are inconsistent in how they publish conference proceedings. Some conference proceedings are published as individual books, in either print or electronic format; some appear as special issues of a journal; and some are available on the Internet

for free. In cases where a conference occurs regularly, the proceedings may be treated as a book series or a journal. Furthermore, some proceedings are distributed only to conference attendees, and not all conference proceedings are available for purchase. Libraries are known as the major information provider for scientific research. The inconsistency in the dissemination of conference proceedings makes it hard for libraries to collect these types of materials, and libraries usually do not collect conference proceedings as systematically as they collect journals. The different dissemination approaches of conference proceedings and journals have put conference proceeding in a disadvantaged position when users are trying to identify and access relevant literature. Another factor is the discoverability of conference proceedings. Even when a library does have access to a conference paper, locating it can be challenging. Many libraries use OpenURL link resolver software such as SFX to streamline their collection and database records (ExLibris 2015), and each database record has such a software link embedded to help users to locate the full text of the desired document. These link resolvers are mainly designed for use with journals, and they do not work well with conference proceedings. When a user finds a relevant conference paper through database searching, the link resolver will usually provide wrong information by indicating that the full text of this document is not available in the library even when the library does have it. However, when the desired document is a journal article, the link resolver will usually give the right information. It is natural that many users will bypass those documents that are not available (even if interlibrary loan is available) and choose those that are currently available to them, in this case, journal articles. Therefore, the discoverability issue will affect the usage of conference papers. The major publishers of conference proceedings, libraries, and link resolver vendors need to work together to improve the discoverability of conference proceedings. Librarians will also need to develop strategies to help users to overcome these challenges in locating conference proceedings.

In the computer science field, evaluating the impact of conference papers is controversial. In 1999 the Computing Research Association issued a report with recommendations on evaluating computer scientists. In this report, it specifically states that conference publications are preferred to journal articles and that "[p]ublication in the prestige conferences is inferior to the prestige journals only in having significant page limitations and little time to polish the paper. In those dimensions that count most, conferences are superior" (Computing Research Association 1999, A). Many universities have adopted these recommendations for the evaluation of computer scientists' research since 1999. Note that the publication date of this recommendation (1999) roughly corresponds to the time when computer science started to grow rapidly, as indicated by the dramatic increase of total publications in this field from 2000 (Figure 2). It is also since then that the prevalence of conference proceedings has outperformed that of journal publications (Figure 3). Fifteen years later, since computer science as a research field has evolved and become more mature, are these practices still applicable?

One of the primary functions of conferences is to bring researchers together and foster collaboration in the community. Therefore, it is not unreasonable to

assume that the number of authors and the number of affiliations would be higher for conference papers than for journal articles. However, our study did not find a significant difference in the average number of authors for the two types of publications. On the contrary, the average number of affiliations for journal articles (1.90) was significantly higher than that for conference papers (1.62). The results indicate that computer scientists are actually more motivated to collaborate when writing journal articles, although it is possible that they had previously established collaboration during a conference meeting. The current evaluation criteria for computer science research focus on conference papers, which has led to a significant increase in the number of conferences since 2000. Owing to funding and time constraints, conferences are usually limited to a group of attendees whose papers have been accepted, and very few conferences draw many attendees beyond the group of authors (Fortnow 2009). Therefore, the opportunities to identify potential collaboration at a conference may be compromised. The results from this study provided indirect evidence on this matter. Moreover, with the wide spread of social networking, other means have emerged for researchers to connect within the scientific community. For example, computer scientists have started to use social media as a forum to discuss the newest research findings and to identify collaboration opportunities (Hadgu and Jaschke 2014).

There are also other concerns about conference publications. The quality of peer review is probably the biggest concern. Peer review of scientific papers is the most accepted method for quality assurance of scientific publications. The reviewing of conference submissions is usually conducted by conference program committees. Because of extremely tight time lines and high workloads, the quality of peer review in conference publications may not rise to the level of journal publications (Birman and Schneider 2009). Second, computer scientists submit their papers when a conference deadline is reached instead of when the research has been properly executed and completed (Fortnow 2009), which may result in the publication of premature research. Third, conference papers are usually shorter than journal articles, as indicated by the research results in this study. This has led some scientists to break up their research results into smaller pieces for publishing in conference proceedings, producing more publications per researcher and per project, yet the combined scientific value of all these papers is likely similar to that of one more complete journal article (Birman and Schneider 2009; Parnas 2007). While this problem is not unique for conference papers, the extent of the problem might be smaller for journal articles because our results indicate that journal articles are significantly longer than conference papers. Fourth, the scholarly communication landscape has changed significantly since the new century, and there are many other channels for fast dissemination of scientific research in addition to conference publications. For example, e-print resources such as CiteSeer^{X} (http://citeseerx.ist.psu.edu) and arXiv (http://arxiv.org) have been widely used by many computer scientists to publish and archive their preliminary and final research immediately. In addition, many sub-fields of computer science are becoming interdisciplinary in nature, and there are strong connections between computer science and other science fields such as biology,

medicine, physics, economics, and education. The different publication culture in computer science may discourage collaboration with other science fields (Fortnow 2009). Because of the limitations of the current systems for disseminating computer science research, it may be time for computer scientists to start to consider using journal systems as the main vehicle to advance computer science research and having fewer but larger-scale conferences so that conferences will fulfil their primary role: bringing the community together and enhancing collaboration (Hermenegildo 2012).

This study has several limitations. This study used bibliometrics to evaluate the impact/quality of publications in computer science. Many computer science research outputs involve the creation of an artefact. Therefore, using citation indicators might not be able to capture the entire impact. The study sample is based on computer scientists in Canada. As stated in the introduction, different reward systems in different countries may directly affect the scholarly communication practices in that country; therefore, these results may not be applicable to the computer science field in other countries. Scopus was used as the data source in this study, and thus the results were limited to what was indexed in Scopus. Changes in Scopus's indexing policy over time may have an effect on the results of this research. For example, if more recent publications had been indexed in Scopus, the number of publications in recent years would increase. Similarly, if Scopus had increased its coverage of recent conference proceedings, the percentage of conference papers would increase too.

Conclusion

This study confirms that conference proceedings are the main venue for the dissemination of computer science research in Canada. However, when the document structure and impact/quality of conference papers and journal articles were compared, it was found that journal articles outperformed conference papers in many of the dimensions. To achieve a higher scientific impact, computer scientists may want to consider publishing more in journals, though this requires more effort than publishing in conference proceedings.

The current evaluation practice of computer science research focuses heavily on conference papers. The scholarly communication landscape has changed significantly in the past 15 years, and computer science as a research field has evolved and become more mature; therefore, using conferences as the main channel to disseminate computer science research may not be able to fully meet the current needs of this field. It may be time for the computer science community to consider changing the current scholarly communication model and adopting a model that has been used by other science fields. In this model, journals will be used as the main channel to publish and archive complete, in-depth, well-written research but with an increased publishing speed; conferences will revert to their primary role of bringing the community together, and their role as a publishing venue will be de-emphasized (Fortnow 2009).

This study also has implications for academic libraries. By raising awareness of the quality and impact of the two types of publications, librarians will have a

role to play in shaping the information behaviours and perhaps the publishing behaviours of computer scientists in the future. As conference papers represent a large body of the literature in computer science but with lower quality compared to journal articles, how to balance the collection of the two types of materials in computer science is a topic that librarians should consider. This is particularly important nowadays when many libraries are facing the challenges of budget constraints but the prices of these materials are still increasing. Identifying and locating conference papers are more challenging than for journal articles for various reasons; therefore, librarians need to incorporate strategies for identifying and locating this type of literature in their information literacy programs for computer science researchers and students. Further, libraries, software vendors, and conference proceeding publishers need to work collaboratively to identify ways to improve the discoverability of conference proceedings. These actions will improve the usage of the conference proceedings and thus perhaps help to improve the overall impact of conference papers.

References

Aguillo, Isidro F. 2012. "Is Google Scholar Useful for Bibliometrics? A Webometric Analysis." *Scientometrics* 91 (2): 343–51. http://dx.doi.org/10.1007/s11192-011-0582-8.

Bar-Ilan, Judith. 2010. "Web of Science with the Conference Proceedings Citation Indexes: The Case of Computer Science." *Scientometrics* 83 (3): 809–24. http://dx.doi.org/10.1007/s11192-009-0145-4.

Birman, Ken, and Fred B. Schneider. 2009. "Viewpoint Program Committee Overload in Systems." *Communications of the ACM* 52 (5): 34–37. http://dx.doi.org/10.1145/1506409.1506421.

Bordons, María, and Isabel Gómez. 2000. "Collaboration Networks in Science." In *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, ed. Blaise Cronin and Helen Barsky Atkins, 197–213. Medford, NJ: Information Today.

Burrell, Quentin L. 2007. "On the *h*-Index, the Size of the Hirsch Core and Jin's *A*-Index." *Journal of Informetrics* 1 (2): 170–77. http://dx.doi.org/10.1016/j.joi.2007.01.003.

Chen, Jilin, and Joseph A. Konstan. 2010. "Conference Paper Selectivity and Impact." *Communications of the ACM* 53 (6): 79–83. http://dx.doi.org/10.1145/1743546.1743569.

Computing Research Association. 1999. "Best Practices Memo Evaluating Computer Scientists and Engineers for Promotion and Tenure." http://cra.org/resources/bp-view/evaluating_computer_scientists_and_engineers_for_promotion_and_tenure/.

Delgado López-Cózar, Emilio, Nicolás Robinson-García, and Daniel Torres-Salinas. 2014. "The Google Scholar Experiment: How to Index False Papers and Manipulate Bibliometric Indicators." *Journal of the Association for Information Science and Technology* 65 (3): 446–54. http://dx.doi.org/10.1002/asi.23056.

Drott, M. Carl. 1995. "Reexamining the Role of Conference Papers in Scholarly Communication." *Journal of the American Society for Information Science* 46 (4): 299–305. http://dx.doi.org/10.1002/(SICI)1097-4571(199505)46:4<299::AID-ASI6>3.0.CO;2-0.

Eckmann, Michael, Anderson Rocha, and Jacques Wainer. 2012. "Relationship between High-Quality Journals and Conferences in Computer Vision." *Scientometrics* 90 (2): 617–30. http://dx.doi.org/10.1007/s11192-011-0527-2.

ExLibris. 2015. "SFX—the OpenURL Link Resolver and Much More." http://www.exlibrisgroup.com/category/SFXOverview.

Fortnow, Lance. 2009. "Viewpoint Time for Computer Science to Grow Up." *Communications of the ACM* 52 (8): 33–35. http://dx.doi.org/10.1145/1536616.1536631.

Franceschet, Massimo. 2010. "The Role of Conference Publications in CS." *Communications of the ACM* 53 (12): 129–32. http://dx.doi.org/10.1145/1859204.1859234.

———. 2011. "The Skewness of Computer Science." *Information Processing & Management* 47 (1): 117–24. http://dx.doi.org/10.1016/j.ipm.2010.03.003.

Freyne, Jill, Lorcan Coyle, Barry Smyth, and Padraig Cunningham. 2010. "Relative Status of Journal and Conference Publications in Computer Science." *Communications of the ACM* 53 (11): 124–32. http://dx.doi.org/10.1145/1839676.1839701.

González-Albo, B., and M. Bordons. 2011. "Articles vs. Proceedings Papers: Do They Differ in Research Relevance and Impact? A Case Study in the Library and Information Science Field." *Journal of Informetrics* 5 (3): 369–81. http://dx.doi.org/10.1016/j.joi.2011.01.011.

Hadgu, Asmelash Teka, and Robert Jaschke. 2014. "Identifying and Analyzing Researchers on Twitter." In *Proceedings of the 2014 ACM Conference on Web Science*, 23–32. Bloomington, IN: Association for Computing Machinery. http://dx.doi.org/10.1145/2615569.2615676.

Hermenegildo, Manuel V. 2012. "Conferences vs. Journals in CS, What to Do? Evolutionary Ways Forward and the ICLP/TPLP Model." http://boemund.dagstuhl.de/mat//Files/12/12452/12452.HermenegildoManuel.Paper.pdf.

Jacsó, Péter. 2006. "Deflated, Inflated and Phantom Citation Counts." *Online Information Review* 30 (3): 297–309. http://dx.doi.org/10.1108/14684520610675816.

———. 2011. "Google Scholar Duped and Deduped: The Aura of 'Robometrics.'" *Online Information Review* 35 (1): 154–60. http://dx.doi.org/10.1108/14684521111113632.

Laplante, Phil, Jon Rockne, Paolo Montuschi, Thomas Baldwin, Mike Hinchey, Linda Shafer, Jeffrey Voas, and Wenping Wang. 2009. "Quality in Conference Publishing." *IEEE Transactions on Professional Communication* 52 (2): 183–96. http://dx.doi.org/10.1109/TPC.2009.2017989.

Maclean's. 2012. "2013 Medical Doctoral University Ranking." http://www.macleans.ca/education/unirankings/2013-medical-doctoral/.

McVeigh, Marie E., and Stephen J. Mann. 2009. "The Journal Impact Factor Denominator: Defining Citable (Counted) Items." *Journal of the American Medical Association* 302 (10): 1107–9. http://dx.doi.org/10.1001/jama.2009.1301.

Parnas, David Lorge. 2007. "Stop the Numbers Game." *Communications of the ACM* 50 (11): 19–21. http://dx.doi.org/10.1145/1297797.1297815.

Rahm, Erhard. 2008. "Comparing the Scientific Impact of Conference and Journal Publications in Computer Science." *Information Services and Use* 28 (2): 127–28.

Seglen, Per O. 1997. "Citations and Journal Impact Factors: Questionable Indicators of Research Quality." *Allergy* 52 (11): 1050–56. http://dx.doi.org/10.1111/j.1398-9995.1997.tb00175.x.

Wainer, Jacques, Siome Goldenstein, and Cleo Billa. 2011. "Invisible Work in Standard Bibliometric Evaluation of Computer Science." *Communications of the ACM* 54 (5): 141–46. http://dx.doi.org/10.1145/1941487.1941517.

Wainer, Jacques, Henrique Przibisczki de Oliveira, and Ricardo Anido. 2011. "Patterns of Bibliographic References in the ACM Published Papers." *Information Processing*

and Management 47 (1): 135–42. http://dx.doi.org/10.1016/j.ipm.2010.07.002.

Weingart, Peter. 2005. "Impact of Bibliometrics upon the Science System: Inadvertent Consequences?" *Scientometrics* 62 (1): 117–31. http://dx.doi.org/10.1007/s11192-005-0007-7.

Wouters, Paul, and Rodrigo Costas. 2012. "Users, Narcissism and Control: Tracking the Impact of Scholarly Publications in the 21st Century." In *Proceedings of 17th International Conference on Science and Technology Indicator*, ed. Éric Archambault, Yves Gingras, and Vincent Larivière, 847–57. Montreal: Science-Metrix and Observatoire des sciences et des technologies. http://www.sticonference.org/Proceedings/vol2/Wouters_Users_847.pdf.

Zhang, Li. 2014. "The Impact of Data Source on the Ranking of Computer Scientists Based on Citation Indicators: A Comparison of Web of Science and Scopus." *Issues in Science and Technology Librarianship* 75 (Winter). http://dx.doi.org/10.5062/F4D798CW.