A Replicable Approach for Rankings in Computer Fields

Abstract: *The goal of this work is to aid funding agencies, universities, and researchers to maximize their research impact in an emerging, multidisciplinary and fast-paced field by developing a replicable methodology for bibliometric evaluation of the relative impact of publication venues in that field of security science.*

1. Introduction

Publishing is a criteria used by research organizations, universities, and funders, for evaluating the scientific merit of researchers for fellowships, funding awards, academic tenure and promotion. Researchers need to be able to identify and assess quality work in their respective fields in order to grow in their profession. Friedman & Schneider (2015) argue that computer science can only grow and mature if scholars are judged based on the quality, rather than quantity, of their work. Researchers also need some basis for deciding on the venues that are most appropriate for their work so as to maximize its impact. Likewise, universities need a basis by which to evaluate faculty and their work for promotion and tenure, and funding agencies need a means for evaluating research in order to inform funding decisions and for evaluating outcomes of funded work.

The goal of this work is to aid funding agencies, universities, and researchers to maximize their research impact in an emerging, multidisciplinary and fast-paced field by developing a replicable methodology for bibliometric evaluation of the relative impact of publication venues in that field of security science.

To address our goal, we develop a scalable and replicable methodology for bibliometric evaluation at the publication venue level that would address limitations of current ranking systems for publications in computer-related fields, and that could be used across a large number of publications. Our bibliometric method is developed to aid organizations in evaluating the relative impact of publication venues in emerging, multidisciplinary, and fast-paced fields, but the methodology could be applied across a variety of disciplines.

Specifically, our work addresses two research questions:
RQ 1 (assessing): How can funding agencies and universities assess the quality of research works?
RQ 2 (identifying): How can researchers identify appropriate venues for their work?

To address the research questions, we (1) develop a list of publication venues relevant to the science of security, (2) identify potential experts to rate publication venues, (3) collect expert ratings of publication venues, (4) analyze the experts' ranking data to develop a metric, 5) utilize quantitative statistics about the publication venues to validate our metric and create a comparison baseline.

We select science of security as the study unit for our work, given that:
(1) science of security is emerging as a multi-disciplinary field involving computer and information science, engineering, and social sciences, there is no existing ranking system that captures relevant publication venues across all of those disciplines;
(2) nor is it possible to use existing ranking systems for evaluating science of security research, as there is no common baseline citation rate.

2. Review of Existing Ranking Systems

Our ranking methodology addresses several limitations of current ranking systems. Most ranking systems rely solely or partially on citation-based metrics which are subject to bias in their judgement of quality and productivity of publication venues. Traditional approaches for evaluating research impact have relied heavily on citation-based metrics such as impact factors, h-factors, and citation counts. For example, Web of Science and Google Scholar rank publication venues in the field of security science using a totally citation-based system. [Describe h-5 calculation index, relationship to impact factor?] Their impact factors are based on the average number of citations to articles in a publication venue over a given period of time. This is problematic in fast-paced fields, as citations are slow to accrue and they are primarily used for academic purposes (Maggio and Leroux et al., 2018). Moreover, citation-based metrics rely on a two year average citation rate (Journal Citation Reports database, add link?). Yet, in fields where the science and technology is evolving very rapidly, research that is two years old may be considered out of date.

A second limitation to citation-based ranking systems is that the quantity of papers published in a journal is an indicator of the venue's productivity, which does not consider the quality of the publication venue. A citation in a well-known publication is worth more than a citation in an obscure publication (Mingers & Leydesdorff, 2015); and publishing in one high-impact journal is arguably worth more than a few publications in low-impact journals.

Third, there is a lack of standardization for citation density across different research fields. For example, Iglesias & Pecharroman (2007) found that citation rates of molecular biology were ten times greater than that of computer science, as the physical sciences typically have greater citation density than other scientific fields. This is a problem for evaluating multidisciplinary research or research in emerging fields of study because there may be no common baseline citation rate to compare against

Most of the ranking systems we reviewed used a combination of citation-based metrics and other methods to assess quality of venues. Two ranking systems were developed by university professors who are experts in computer security. Developed for informal use only, these ranking systems are based on quantitative metrics and expert personal opinion. Although, personal opinions are based on an assessment of multiple more objective variables (e.g. acceptance rates, impact factors, etc.). About half of the ranking systems we identified only included conferences in their lists and half included both journals and conferences. Of the sources that considered both venue types, conferences were typically ranked on a more regular basis than journals and with a more well-defined ranking methodology.

Table 1 provides a brief description of the ranking systems we identified, some of which are explained in more detail in the Methodology section.

Table 1. Ranking System Methods

Ranking System	**Topical area**	**Ranking method**	**Venues**
Google Scholar	Security and cryptography	Citation-based, h-index	Conferences and journals
CORE Rankings Portal	Computing disciplines	Conferences are partially citation-based, as well as mix of indicators including acceptance rates, visibility and track record of key conference hosts; journals have not been ranked since 2010	Conferences and journals
Jianying Zhao, university professor	Security and cryptography	Conference impact factor (CIF) formula based on acceptance rates and impact factors	Conferences
Guofei Gu, university professor	Academic and scientific security	Opinion-based (e.g. acceptance ratio, paper quality and impact, committee member quality, etc.)	Conferences
University of California, Santa Barbara (UCSB)	Systems and networking	Tracks number of submitted and accepted papers and other factual conference-level information (e.g. number of attendees)	Conferences
National University of Singapore	Computer science	Methodology not described, and rankings not updated since 2011	Conferences and journals

3. Methodology

In order to develop a customized bibliometric for the science of security, we needed to undertake the following steps: 1) develop a list of publication venues relevant to the science of security, 2) identify potential experts to rate publication venues, 3) collect expert ratings of publication venues, 4) analyze the experts' ranking data to develop a metric, 5) utilize quantitative statistics about the publication venues to validate our metric and create a comparison baseline.

To carry out this work, we used a Science of Security research group as a case study and compiled an initial list of publication venues in which the group's papers were published and presented. [[We also included publication venues nominated by group's researchers, nominations from the NSA, and those nominated by our experts]] To provide a more comprehensive view of the type of venues that these papers were published and presented, we developed a metric to assess the quality of both journals and conferences. Next, we used a variety of structured methods based on ratings from 21 computer science experts and quantitative non-citation based metrics for the venues selectivity and scale. The process of data collection and analysis is described in the next section.

Data Collection

We compiled a list of publication venues and number of papers published in each venue by Lablet researchers from 2014-2016. Next, we collected information on acceptance rates for each of these venues. A first attempt to collect the bibliometric information was done by accessing the venue website, conference proceedings, or an aggregator site. An aggregator site in this particular context brings together disparate sources of information (e.g. publication venue rankings, acceptance rates, number of submissions, etc.) to form a cohesive list that can be accessed in one place. We used the following two aggregator sites (listed directly below) and referenced a third site as needed to ensure no relevant statistical data was overlooked:

1. A site administered by the University of California, Santa Barbara (UCSB) that tracks submission numbers and acceptance rates for Systems and Networking conferences: http://www.cs.ucsb.edu/~almeroth/conf/stats/
2. A list maintained by Jianying Zhou at Singapore University of Technology and Design to track acceptance rates and impact factors once per year for the top Crypto and Security conferences (http://jianying.space/conference-ranking.html). This list is intended for informal use only. Security conferences are ranked by their Conference Impact Factor (CIF) which is calculated with the following formula: CIF=1/AR+PR+CR where AR= number of accepted papers, PR = Security conferences are ranked using a Conference Impact Factor (CIF) system [CIF= 1/AR+PR+CR, where AR=number of accepted papers/number of submissions, PR=number of accepted papers/number of registered participants, and CR=number accepted papers/number of citations]. CR data for years through 2016 are from Microsoft Academic Search, and 2017 data are pulled from Google Scholar. This list does not rank all security conferences and is intended for informal use only. and Cand conferences with less than 20 accepted papers and less than 50 participants are excluded.
3. A more extensive list of academic and scientific security conferences compiled by Guofei Gu at Texas A&M to track the venue's acceptance rates, accepted papers, and number of submissions. (http://faculty.cse.tamu.edu/guofei/sec_conf_stat.htm). This list was developed for informal use only. Conferences are ranked based on consideration of the following variables: acceptance ratio, paper quality and impact, committee member quality, attendee/number of paper ratio, conference location and history, and industry connection. Rankings reflect the quality of the venue as determined by personal assessment of the criteria. Based on the acknowledged bias in the ranking methodology, we only consulted this list for the acceptance rate data. http://faculty.cse.tamu.edu/guofei/sec_conf_stat.htm

In addition to the above-mentioned sites, acceptance rate data were gathered through a variety of online sources including conference and journal websites, publishing venues, and secondary data sources. If online information was unavailable, we contacted the conference chair, or journal editor, or listed contact person via email to request the data. If we did not get a response after 1 initial contact and 1 follow up request, we attempted to contact the society which hosted the conference or the publisher of the journal. If no reply was received, we followed up with a second email a week after the first. A final communication attempt was made by phone if the second email yielded no response, and calls were only made to phone numbers within the United States.

For some venues, only partial data were available through online searches or email communication with publication venue editors. The most common examples of partial data were average acceptance rates for multiple years instead of rates for each year, and data that only reported the number of accepted papers or acceptance rates but not both. We present average acceptance rates where specific data could not be identified. We documented these instances to ensure that the results are interpreted appropriately. When possible, the third method to account for missing data was to calculate the missing data points using partial data (e.g. number of submissions and acceptance rates, or number of accepted papers and submission rates, or number of accepted papers and acceptance rates). For conferences that do not occur on an annual basis, data were still considered complete if all possible data were collected.

Inclusion Criteria

The list of SoS publication venues was developed from a number of sources. First, quarterly reports for the Lablet were used to develop a database of NCSU Lablet publications. We used that database to identify publication venues to be included in the list. Second, we invited NCSU Lablet participants and participants in the other Lablets to nominate publication venues to nominate relevant publication venues. Finally we consulted with NSA program officers and a SoS experts. All publication venues identified were included in the initial list. The list of publication venues was validated by an online search of each venue. Altogether, we compiled a list of X publications and X publication venues prior to applying exclusion criteria.

Exclusion Criteria

Publication venues were excluded from consideration in the analysis if they met any of the exclusion criteria. Venues were excluded if there were less than 10 papers submitted to the venue, if the venue was a one-time only event or a local meeting, if the conference was a sub-conference that was already included in the analysis, and if no relevant information about the venue could be found online, or from the venue's editor, publisher, or coordinator. After applying the exclusion criteria below, X venues were removed from the list, resulting in a total of X publication venues. Overall, there were 169 publications across these 84 venues, resulting in a response rate of X/X=X.57/79 = 72%

Expert Ratings

Twenty-one experts were nominated by the National Security Agency (NSA) program director nominees and the SoS Lablet. Many of the experts were identified based on their involvement on review boards for major computer science conferences. Experts were provided a set of directions for rating the publication venues, and asked to rate the venues in terms of their potential impact on the science of security. Impact Experts provided two separate ratings for each venue; a rating for quality of the work published in the venue and a rating for relevance to science of security. Experts were asked to consider six criteria when making their ratings determination, and they assigned a rating between 0 and +1. The ratings are described below.

[Why did we develop the criteria, one line and what criteria, bulleted list]

[Create a small table to display ratings descriptions?]

- +1: for premier publication venues in the science of security
- 1: for excellent publication venues
- 2: for good publication venues
- 3: for other publication venues
- 0: for publication venues for which you are unable to provide a rating

Data analysis

Expert ratings were compiled and based on all expert ratings assigned to each venue, a mode score was determined in Excel for quality and relevance combined, and as separate metrics.

[DESCRIBE ADDITIONAL ANALYSES]

4. Results

Of 170 publication venues, the experts most frequently assigned a top or middle tier score to the majority of venues (72%) in the list. The top tier rating (3) was the most frequent score assigned to 43% of the venues, and the middle tier rating (2) was the most frequent score given to 29% of the publication venues. Results were similar when examining quality and relevance scores separately, though experts were slightly more likely to rate the venue higher for its relevance to the field than for the quality of the publication venue. Experts most frequently assigned a top tier quality score to 42% of venues and top tier relevance score to 43% of venues, indicating that publication venues were viewed as nearly equal in quality and relevance to the field. Table X includes the number of publication venues in each rating category.

Table X Publication Venues by Rating Category

Category	Number of venues in each rating category (combined score)	Number of venues in each rating category (quality score)	Number of venues in each rating category (relevance score)
Top tier (3)	**73**	**70**	**71**
Middle tier (2)	49	44	52
Bottom tier (1)	44	47	43
Not assigned rating (0)	4	9	4

We also determined the percentage of publication venues included in our analysis that also appear in the list of venues rated by other ranking systems (Table X). Most ranking systems only rated conference venues, therefore we determined the percentage of conference venues from our analysis that appeared on lists compiled by other sources.

Table X [Add title]

Ranking System	Venues ranked	% of our venues ranked in other lists
Google Scholar	Conferences and journals	6.74
Computer Science Rankings	Computer science schools	
CORE Rankings Portal	Conferences and journals	
Jianying Zhao, university professor	Conferences	
Guofei Gu, university professor	Conferences	
University of California, Santa Barbara (UCSB)	Conferences	
National University of Singapore	Conferences and journals	

Conclusion

Our bibliometric analysis was specifically designed to be relevant to the Science of Security. We used a variety of structured methods based on ratings from a pool of several subject matter experts and quantitative non-citation based metrics for the venues. Our comprehensive approach aims to reduce bias and inaccuracy that can result from relying entirely on a citation-based system as a quality measure. It is a scalable method at the publication venue level that can be used to analyze a large number of publications as well as measure potential impact for papers published in these venues.

Although our bibliometric analysis has several notable strengths, a few limitations should be acknowledged. First, we were unable to get complete data to determine acceptance rates for several journals. Journal metrics were much more limited in availability through online sources than conference data, and in some instances we could only obtain aggregated acceptance rates over multiple years. Although we made many attempts to locate data for publication venues where it was not readily available, there were instances when it was difficult to connect with the appropriate contact for obtaining data. Also, we acknowledge the possibility that our experts may have personal biases toward certain publication venues, which is why we chose to engage several different experts in the field. The consistency of venue ratings across this pool of experts strengthens the credibility of the results.

[add more content, limitations of expert ratings, etc.]

References

[cite the ranking systems described in review]

Friedman & Schneider (2015)
Iglesias, Pecharroman. (2007).

Web of Science. (2019). Journal Citation Reports. Available at: https://ucsd.libguides.com/ResearchImpact/JCR

Maggio, L. A., Leroux, T. C., Meyer, H. S., & Artino, A. R., Jr (2018). #MedEd: exploring the relationship between altmetrics and traditional measures of dissemination in health professions education. *Perspectives on medical education*, *7*(4), 239–247. doi:10.1007/s40037-018-0438-5

Mingers, J. and Leydesdorff, L. (2015). A review of theory and practice in Scientometrics. European Journal of Operational Research, 246 (1), 1-19.

Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. Journal of the
American Society for Info Science and Tech, 57(8), 1060-1072.

Glänzel, W., Schubert, A., & Czerwon, H. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. Scientometrics, 44(3), 427-439.

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. D., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. Nature, 520, 429–431.

Freyne, J., Coyle, L., Smyth, B., & Cunningham, P. (2010). Relative status of journal and conference publications in computer science. Communications of the ACM, 53(11), 124-132.

Patterson, D., Snyder, L. and J. Ullman, J. (1999) Evaluating computer scientists and engineers for promotion and tenure, CRA Best Practice Memos.

Friedman, B. and Schneider, F. (2015). Incentivizing quality and impact: Evaluating scholarship in hiring, tenure, and promotion, CRA Best Practice Memos.

Meyer, B., Choppy, C., Staunstrup, J., & van Leeuwen, J. (2009). Viewpoint research evaluation for computer science. Communications of the ACM, 52(4), 31-34.

Wainer, J., Goldenstein, S., & Billa, C. (2011). Invisible work in standard bibliometric evaluation of computer science. Communications of the ACM, 54(5), 141-146.