International Research Journal of Multidisciplinary Science & Technology
www.irjmrs.com ISSN : 2455-930X

Pattern of Collaborative Research on IEEE Transaction on Software Engineering (2000-2005): A Bibliometric Study

Dr.R.Lakshmi Sankari[1], Dr.P.Rani[2]

[1]Librarian, Thiruvalluvar University Constituent College of Arts and Science ,Tittagudi
[2]Dr.P.Rani, Assistant Librarian, Regional Office Anna University Tirunelveli

---

1.INTRODUCTION

Bibliometrics is statistical analysis of write publications, such as books and articles. The most frequently used bibliometric methods are citation analysis and content analysis. Citation links or references give the particulars of earlier work on the topic being discussed. The topics which are cited more shall have greater impact on the area of research. The impact is calculated based on the number of citations received by articles to the number of articles published by the journal in two years. The term bibliometrics was coined in 1969 by Pritchard . Before 1969 it was either called statistical bibliography or it had no title at all. The terms bibliometrics, informetrics, scientometrics, librametrics, webometrics are derived from the term 'metrics'. These are the measuring techniques of knowledge in library and information science. Bibliometrics and scientometrics are two closely related approaches to measuring scientific publications and science in general, respectively. These two methods are widely used for studying growth pattern of publications, citation analysis, to calculate the impact factor of a journal title, development of scope and spectrum of any area of research, and ranking of scholarly output of researchers and institutions and identifying the centres of excellance in academia.

There are famous laws of bibliometric, i.e., (Lokta, 1926) of scientific productivity, Broadford's law (Broadford, 1934) of scattering and Zips law (Zips, 1949) on frequency of words. However, the bibliometric study is unique and common to all the subjects after sixties.

Research has become a collective effort in all fields. It is seen that there is consistent increase towards collaboration in research in various disciplines. Scientists realise the necessity of collaboration in research in present era of Information explosion.

2.SCOPE OF THE JOURNAL

The IEEE Transactions on Software Engineering is published monthly. We are interested in well defined theoretical results and empirical studies that have potential impact on the construction, analysis, or management of software. The scope of Transactions ranges from the mechanisms through the development of the principles to the application of those principles to specific environments. Specific topic areas include;

a. Development and maintenance methods and models, e.g., Techniques and principles for the specification, design and implementation of software systems, including notation and process models;
b. Assessment methods, e.g., Software Test and validation reliability models, test and diagnosis procedures, software redundancy and design for error control, and the measurement and valuation of various aspects of the process and the product;
c. Software Project management, e.g., Productivity factors, cost models, schedule and organizational issues, standards;
d. Tools and environment, e.g., Specific tools, integrated tool environment including the associated architecture, databases and parallel and distributed processing issues;
e. System issues, e.g., Hardware-Software trade-off; and
f. State-of-the-art surveys that provide a synthesis and comprehensive review of the historical development of one particular area of interest.

3.RELATED LITERATURE

Alka Bansal extended the bibliometric analysis to 2012 and concluded that the maximum number of articles was published in 2012 and the maximum number of contributions has the length of 6-10 pages, and majority of the authors preferred journals as the source of their citation.

Garg & Anjana examined authorship pattern, average length of articles in terms of pages, geographical distribution of articles, most prolific institutions and authors, and citations earned by the articles of *Journal of Intellectual Property Rights*.

Kumar & Moorthy carried out bibliometric analysis of *DESIDOC Journal of Library and Information Technology* during 2001-2010 and concluded that the number of papers published increased after 2006, and the maximum number of papers were published in 2008 and 2009. The maximum number of single-authored papers was 37.6 % followed by two-authored papers with 36.9 %.

Rao *et al*., analysed bibliometric analysis of the *Journal of Propulsion and Power* during the period 1985-2013. A total of 4047 articles were published and it is observed that USA produced 61.24 %, Japan 6.92 %, Republic of China' 3.8 %, Germany 3.25 %, UK 3.03 % and India 2.66 % of the total articles.

Rao *et al*., analysed bibliometric analysis of the *Journal of Propulsion and Power* during the period 1985-2013. A total of 4047 articles were published and it is observed that USA produced 61.24 %, Japan 6.92 %, Republic of China' 3.8 %, Germany 3.25 %, UK 3.03 % and India 2.66 % of the total articles.

Thanuskodi studied the *Journal of Social Sciences.* It has published 273 articles during the period of study. The highest number of articles appeared in the area of economics. Also, he examined the authorship pattern, subject-wise distribution of articles, average number of references per articles, forms of documents cited, year wise distribution of cited journals, etc.

Vellaichamy & Jeyshankar investigated that a total of 158 papers were published in the *Journal of Webology* from 2004-2013. The degree of collaboration ranges from 0.182 to 0.693 and its mean value is found to be 0.44. Web analysis (24.68 %) and social media (15.82 %) papers are the top most publications in subject-wise analysis. India has contributed more number of articles compared to any other countries, such as Iran, UK, USA and Australia.

4.METHODOLOGY

The data required for the study was collected from the print version of the journal for the period from 2000 to 2005. The each articles were carefully scanned and tabulated. The following pages discuss and analyse the data collected with the help of bibliometric techniques and present the results in the form of tables and interpretations as per the objectives of the study.

5.OBJECTIVES OF THE STUDY

The present study has been undertaken:

1. To study the yearly distribution of the articles
2. To study the authorship pattern, Degree of Collaboration, Author Index in the field of study
3. To examine the validity of Lotka's law using productivity of authorship and to undertake K-S statistics for the conformity of the results obtained by these methods.
4. To study the Geographical distributions of contributions

6.DATA ANALYSIS AND INTERPRETATION

Table 1-Year-wise Distribution of Contributions

Years	Vol.No	No.of Issues	No. of contributions	%
2000	26	12	65	15.55
2001	27	12	60	14.35
2002	28	12	75	17.94
2003	29	12	87	20.81
2004	30	12	64	15.31
2005	31	12	67	16.03
			418	100.00

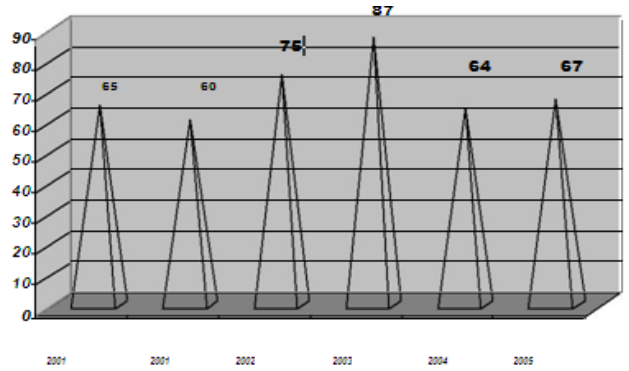

Fig.1 Year wise Distribution of Contributions

7.Year wise Distribution of Contributions

Table 1 provides the information about the Yearly distribution of research articles published in IEEE transaction on Software Engineering, during the period of 2000 to 2005. It is observed from the table that highest number of articles are published in the year 2003, i.e.87 (20.81%) followed by the year 2002 with 75(17.94%) and the lowest number of articles in the year 2004 with 64 (15.31%) articles. A notable attribute of the study is that the year 2005 shows the maximum number of contributions.

8.Authorship pattern

Bibliometric studies have shown that research in the sciences is predominantly carried out by group of researchers rather that by a single researchers, Through collaboration, researchers share and exchange knowledge and techniques, that results in the derivation of positive scientific through and decrease cost (Katz and Martin 1997).

Table-2 Distribution of Authorship pattern

Authors	Year						Total	%
	2000	2001	2002	2003	2004	2005		
Single Author	9	6	8	9	3	7	42	10.05
Two Author	36	17	29	37	15	19	153	36.60
Three Author	11	25	22	23	24	24	129	30.86
More than Four	9	12	16	18	22	17	94	22.49
	65	60	75	87	64	67	418	100.00

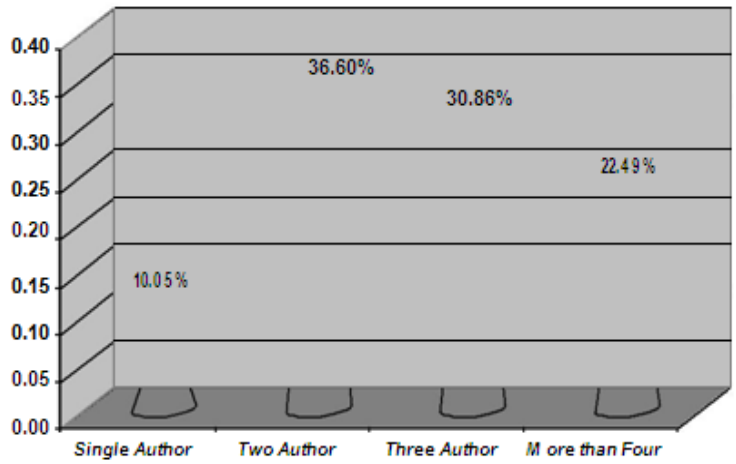

Fig.2 Distribution of Authorship pattern

Distribution of articles according to number of authors is given in Table 2 , reveals that out of 481 Research Articles, 42 (10.05 %) contributed by Single author. 153 (36.60 %) research articles were contributed by two authors.129 (30.86 %) research articles were contributed by Multi authors and 94 (22.49 %) research articles were contributed by more than four authors. It is very clear from the above table that the majority of articles are contributed by two authors.

9.Collaborative Measures

Collaborative coefficient (CC) suggested by Ajiferuke and used by Karki and Garg has been used to measure the extent and strength of collaboration among the IEEE Transaction on Software Engineering. It can be expressed mathematically as:

$$CC = 1 - \sum_{j=1}^{j=k} (1/J)\, f_j/N$$

Where

- f_j is the number of *J* authored papers published in a discipline during a certain period of time.
- *N* is the total number of research papers published in a discipline during a certain period of time and
- *k* is the greatest number of authors per paper in a discipline

10.Degree of Collaboration (DC)

The determine the degree of collaboration in biotechnology, the following formula given by Subramanyam (1983) has been used;

$$DC = \frac{NM}{NM+NS}$$

Where DC = Degree of Collaboration

NM = Number of Multi authored papers

NS = Number of single authored papers.

11.Collaborative Index

Collaborative Index can be obtained by total number of authors divided by total number of published articles.

$$CI = \frac{\text{Total number of Authors}}{\text{Total number of Articles}}$$

Where,

CI = It is the number of authors per paper.

Table - 3 Authorship pattern and collaborative Measures

Year	TP	TA	CI	CC	DC
2000	65	163	2.51	0.50	0.86
2001	60	177	2.95	0.58	0.90
2002	75	212	2.83	0.56	0.89
2003	87	229	2.63	0.55	0.89
2004	64	212	3.31	0.64	0.95
2005	67	196	2.93	0.58	0.90
Total	418	1189	2.84	0.56	0.90

TP=Total Papers; TA=Total Authors; CI=Collaborative Index; CC=Collaborative Co-efficient and DC=Degree of Collaboration.

Table 3 shows that Authorship pattern and collaborative measures. The collaborative Index for the year 2000-2005 was 2.84 which show that collaborative research pattern than solo research; the value of Collaborative Co-efficient (CC) have shown higher (CC) value with more than 0.50, which show greater probability of multiple-authorship and the degree of collaboration (DC) for the year 2000-2005 was 0.90.

12.AUTHOR PRODUCTIVITY AND LOTKA'S LAW

12.1.Author productivity

The term author productivity, scientific productivity, publication productivity and trends of publications are used synonymously. Regarding the author productivity one can say that, author productivity means authors productiveness or author's efficiency in publication. In other words author productivity can be explained as the effectiveness of productive efforts to produce fruitful publication.

Table 4 Author productivity

No. of articles (x)	No. of authors observed (y)	% of authors
1	817	83.79
2	129	13.23
3	13	1.33
4	11	1.13
5	3	0.31
7	1	0.10
9	1	0.10
	975	100.00

The numbers of authors contributing one, two, or more articles each were counted manually. The author data shows 975 authors produced 481 articles with an average of 2.02 authors per article. Table 4 presents author productivity data for Lotka's law. Of the 975 unique author names, 817 (83.79 %) produced one article, 129 (13.23 %) produced two articles and so forth. The number of authors who produced more than 7 and 9 articles is quite small (only 0.10 %).

12.2.Lotka's Law

Lotka's law is a classical method used to test the regularity in the publication activity of authors of scientific literature. It describes the frequency of publication by authors in a given field. It states that the number of authors making *n* contributions is about $1/n^2$ of those making one; and the proportion of all contributors that make a single contribution is in the region of 60 per cent. This means that out of all the authors in a given field, 60 per cent will have just one publication; 15 percent will have two publications ($1/2^2$ times 60); 7 per cent will have three publications ($1/3^2$ times 60), and so on.

This law can be expressed as:

$$y = C \times x^{-n} \quad (1)$$

Where *x* is the number of publications of interest (1, 2, etc.); *n* is an exponent that is constant for a given set of data; *y* is the expected percentage of authors with frequency *x* of publications, and *C* is a constant. The productivity corresponds not to the number of articles published by an author but to its logarithm; it seems that a multiplicative, rather than simply additive, model provides a better fit to this measure or counting method.

The exponent *n* is often fixed at 2, in which case the law is known as the *inverse square law of scientific productivity.* However, given that the exponent *n* predicts the relative number of authors at each productivity level it would seem useful to calculate it. In the present study, least square method has been used. It can be expressed as:

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} \quad (2)$$

Where N is the number of data pairs considered

X is the logarithm of x (x=number of articles) and

Y is the logarithm of y (y=number of authors)

The constant C is calculated using formula:

$$C = \frac{1}{\sum 1/x^n} \quad (3)$$

To verify that the observed distribution of author productivity fits the estimated distribution, Pao (1985) suggested applying the non-parametric Kolmolgorov-Smirnov (K-S) goodness-of fit test. To this end, the maximum difference between the real and estimated accumulated frequencies was calculated and these values were then compared with the critical values (c.v) obtained from the following equation:

$$c.v = \frac{1.63}{\left(\sum yx + \left(\sum yx/10\right)\right)^{1/2}} \quad (4)$$

Table 5 Author productivity Using Lotka's law

x	y	X	Y	x^2	xy	Yx/∑yx	∑(yx/∑yx)	$1/x^n$	fe=C($1/x^n$)	∑fe	D
1	817	0.000	2.912	0.000	0.000	0.838	0.838	1.000	0.871	0.871	0.033
2	129	0.301	2.111	0.091	0.635	0.132	0.970	0.103	0.090	0.961	0.010
3	13	0.477	1.114	0.228	0.531	0.013	0.984	0.027	0.024	0.984	0.001
4	11	0.602	1.041	0.362	0.627	0.011	0.995	0.011	0.009	0.994	0.001
5	3	0.699	0.477	0.489	0.333	0.003	0.998	0.005	0.000	0.994	0.004
7	1	0.845	0.000	0.714	0.000	0.001	0.999	0.002	0.001	0.995	0.004
9	1	0.954	0.000	0.911	0.000	0.001	1.000	0.001	0.001	0.996	0.004
	975	3.879	7.655	2.794	2.127	1.000	6.784	1.148			

n= 3.28, c=0.871, D = 0.033

To validate the Lotka's law a calculation was done using Eqns (1-4), (Table-5), to know the values of n and C to test whether application of Lotka's law fits or not. Thus, based on the data presented in Table 5, the calculated values of n and C are 3.28 and 0.871, respectively. In order to verify that the observed distribution of the productivity of the authors fits the theoretical distribution, we subjected the data to the non-parametric Kolmolgorov-Smirnov test. To this end, we used the data in the last column of Table 5 (Dmax), obtained as the absolute value of the difference between columns 8 and 11 of the same table. The greatest value of this column (Dmax) will be taken as reference for comparison with the "critical value" (c.v), obtained by the asymptotic formula (5)

The critical value is 0.052 and the value of Maximum Difference (D) between the real and estimated accumulated frequencies is 0.033, which is less than the critical value 0.052. This resulted fitting the application of Lotka's law to the data of IEEE Transaction on Software Engineering.

Table 6 – Geographical Distribution of the First Author

Countries	2000	2001	2002	2003	2004	2005	Total	%
America	18	22	31	30	22	29	152	36.36
Canada	8	5	8	8	11	6	46	11.00
Italy	3	7	9	7	5	4	35	8.37
Taiwan	4	8	2	6	5	-	25	5.98
England	5	4	3	6	1	4	23	5.50
Germany	-	1	2	4	5	2	14	3.35
Norway	1	-	-	2	5	5	13	3.11
Australia	3	1	1	4	-	1	10	2.39
China	1	-	-	3	4	2	10	2.39
France	2	2	2	1	1	-	8	1.91
Brazil	3	1	2	1	1	-	8	1.91
Belgium	2	-	2	-	1	3	8	1.91
Spain	1	1	-	4	1	-	7	1.67
India	1	2	2	1	-	1	7	1.67
Japan	2	-	1	1	-	2	6	1.44
Korea	3	2	-	-		-	5	1.20
Switzerland	1	-	-	1	1	1	4	0.96
Netherlands	-	-	1	-	-	3	4	0.96
Maryland	2	1	-	-	-	-	3	0.72
Ireland	-	1	-	2	-	-	3	0.72
Sweden	-	-	1	2	-	-	3	0.72
Argentina	-	-	1	1	-	1	3	0.72
Greece				2	-	1	3	0.72
Rome	2	-	-	-	-	-	2	0.48
Denmark	1	1	-	-	-	-	2	0.48
Singapore	-	-	1	1	-	-	2	0.48
Israel	1	-	-	-	1	-	2	0.48
Newzeland	-	-	1	-	-	1	2	0.48
Hungary	-	-	1	-	-	1	2	0.48
Michigan	1	-	-	-	-	-	1	0.24
Tunisia	-	1	-	-	-	-	1	0.24
Yugoslavia	-	-	1	-	-	-	1	0.24
Finland	-	-	1	-	-	-	1	0.24
Prague	-	-	1	-	-	-	1	0.24
Virgiria	-	-	1	-	-	-	1	0.24
Total	65	60	75	87	64	67	418	100.00

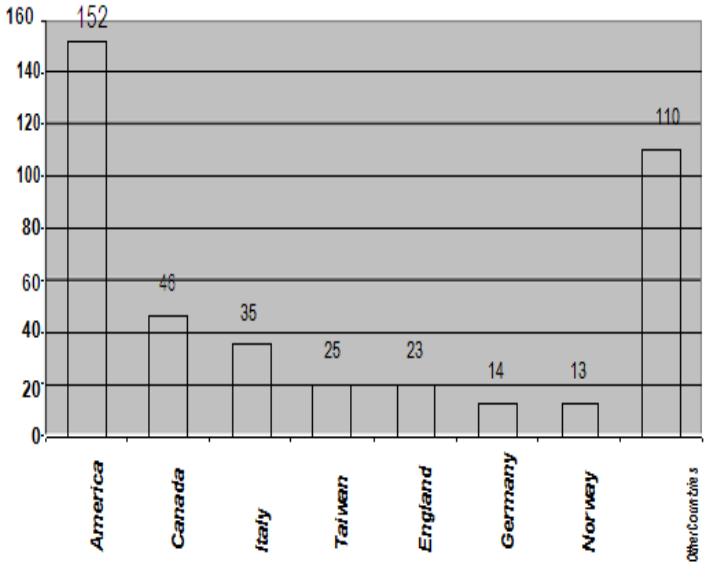

Table: 6 Provides the information about the Geographical distribution of the first Author. Totally 481 Authors have been identified during the study. 152 (36.56 %) First authors contributed were from America, 46 (11 %) from Canada; 35 (8.37 %) from Italy; 25 (5.98 %) from Taiwan, 23 (5.50 %) from England, 14 (3.35 5) from Germany 13 (3.11 %) from Norway and 110 (26.31 %) first authors contributed from other countries. However, it is inferred that out of the above mentioned countries, America gives priority for research when compared to other countries.

13.FINDINGS

The following are the major findings of the study:

1. The maximum number of papers (87) were published in 2003 and minimum of 64 in 2004 (Table 1).
2. Authorship pattern of IEEE Transaction on Software Engineering for the year 2000-2005, the majority of articles are contributed by two authors.153 (36.60 %) (Table 2).
3. The collaborative Index for the year 2000-2003 was 2.84 which show that collaborative research pattern than solo research; the value of Collaborative Co-efficient (CC)) value with more than 0.50, which show greater probability of multiple- authorship and the degree of collaboration (DC) was 0.90. (Table 3)
4. Author Productivity was calculated and that 817 (83.79%) authors contributed one article, and 9 (0.10%) authors contributed more than 9 articles. The productivity of the authors does fit a lotka distribution (Dmax=0.033 and critical value = 0.052).
5. Geographically more number of first authors contributed from America 152(36.36 %) (Table 6).

14.CONCLUSION

A Bibliometric Analysis of "IEEE Transactions on software Engineering" published during 2000-2005. The study reveals that the average number of contribution per volume has come around more than 60.America has contributed more articles for publication. The study highlights that the majority of the articles are contributed by two-authors. Indian authors may be encouraged to write more articles for publication. Joint authors may come to forward to present articles from India. In fact Indian authors are capable of writing the journal articles, they should be given such a opportunity by management as motivation. Indian authors may be permitted to visit abroad and participate in the international conferences and present the articles in the forum. Then Indian will also find the place top in the ranking analysis. Bibliometric analysis is considered as the tool for the measurement in the field of Library and Information science and it has been increasingly used to evaluate the performance and growth of various disciplines.

REFERENCES

[1] Ajiferuke, Isola; Burell, Q. & Tauge, Jean. Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, 1988, 14(5-6), 421-33.
[2] Bansal, Alka. *DESIDOC Journal of Library and Information Technology*: A bibliometric analysis. *DESIDOC J. of Lib. & Inf. Tech.*, 2013, 33(5), 412-17.
[3] Bibliometric study of India's scientific publication outputs during 2001-2010. Dept. of Science and Technology, 2012, 186p.
[4] Garg, K.C. & Anjana, A.K. *Journal of Intellectual Property Rights*: A bibliometric study. *DESIDOC J. of Lib. & Inf. Tech.*, 2014, 34(1), 66-73.
[5] Kumar, Manoj & Moorthy, A.L. Bibliometric analysis of *DESIDOC Journal of Library and Information Technology* during 2001-2010. *DESIDOC J. of Lib. & Inf. Tech.*, 2011, 31(3), 203-08.
[6] Lotka, A. J. The frequency distribution of scientific productivity. Journal of the Washington Academy of Science, 1926, 16, 317-23.
[7] Nageswara Rao, K; Sarma, Rajeev Kumar; Girija Devi, S. & Muralidhar, S. Bibliometric analysis of the *Journal of Propulsion and Power* (1985- 2013). *DESIDOC J. of Lib. & Inf. Tech.*, 2014, 34(3), 271-76.

[8] Salini, C.P.; Nishy, P.; Vishnumaya, R.S. & Mini, S. A bibliometric evaluation of organic chemistry research in India. *Annals of Lib. & Inf. Stud.*, 2014, 61(4), 332-42.
[9] Sengupta, I.N. Bibliometrics, informetrics, scientometrics and librametrics: An overview. *Libri,* 1992, 42(2), 75-98.
[10] Subramanyam, K. Bibliometric studies of research collaboration: A review. *J. of Inf. Sci.*, 1983, 6(1), 33-8.
[11] Thanuskodi, S. *Journal of Social Sciences*: A bibliometric study. *J. of Soc. Sci.*, 2010, 24(2), 77-80.
[12] Vellaichamy, A. & Jeyshankar, R. Bibliometric analysis of the Journal *Webology* from 2004-2013. *J. of Adv. in Lib. and Inf. Sci.*, 2015, 4(1), 7-13.