A Citation Analysis of Top Research Papers of Computer Science

Dr. Akhtar Hussain
Web Librarian, Civil Engineering Department
King Saud University, Riyadh
Kingdom of Saudi Arabia
Email: ahusain.c@ksu.edu.sa

Dr. Dillip K Swain
Lecturer,
P. G. Department of Library and Information Science
North Orissa University, Baripada, Odisha, India
E-mail: swaindk_69@yahoo.co.in

Abstract

The study intends to evaluate the top papers of Computer Science as reflected in Science Direct. Moreover, it aims to find out authorship pattern, ranking of authors, ranking of country productivity, ranking of journals, and highly cited papers of Computer Science. The citations data have been collected from the quarterly list of hottest 25 research articles in the subject field of Computer Science from Science Direct database. In the present study, 20 issues of the alert service beginning from January/March 2005 to October/December 2010 containing a total number of 495 articles in Computer Science have been taken up for analysis. The study reveals that out of 495 top papers; three-authored articles are little ahead than two authored articles followed by four-authored articles and the country productivity of USA is at the top followed by UK, Taiwan, Chaina, and Canada. Moreover, it finds that European Journal of Operational Research occupies the top position followed by Computers in Human Behavior, and Pattern Recognition.

Keywords: Citation Analysis, Bibliometric Study, Equal Credit Method, Citation Counts, Country Productivity

Introduction

Citation analysis is a bibliometric technique that uses citation patterns in documents to trace the relationships between those documents and the original sources and authors. The relationships found provide a picture of the cultures of those disciplines. According to bibliometric pioneer Eugene Garfield, "citation links provide a quantitative picture of journal utility and relationships that is useful" (Garfield, 1979)[1]. Citation analysis is now commonly used to determine what titles to purchase, to discontinue, or to weed (Smith, 1981)[2]. The main idea in citation analysis is that citations are the real reflection of the impact of published scientific results, and that the majority of important information will be found in the core, highly cited journals. Nevertheless, citation analysis is often used as a tool for evaluating the performance and measuring the impact of scientists, institutions, journals, regions etc. (Matutinovic, 2007)[3]. Although this measure is easy to use, there are many factors to be considered as to why a journal is being cited, such as circulation and acceptance rate and other factors such as the group of researchers with whom the author associates (Buffardi and Nichols, 1981, LaBonte, 2005)[4-5].Moreover, the scientific community takes the support of bibliometric data,

including citation counts of articles, impact factors of journals, etc. in their subject areas for selecting preferred journals for research papers/scholarly communications (Sahu, Goswami, & Chaudhury, 2011)[6]. Therefore, the study attempts to assess, evaluate and analyse best papers, core journals and the most potential authors that have bestowed significant contributions to the field of computer science during 2005 to 2010.

Objectives

The study keeps the following key objectives in its ambit of analysis:

- To ascertain the authorship pattern in top research papers of Computer Science,
- To prepare a rank list of authors;
- To study the country wise share of contributions in the identified top research papers of Computer Science;
- To prepare a rank list of journals that have published high yielding scholarly papers; and
- To find out top ten papers which have received significant impact by their number of citations.

Literature Review

A number of studies have been carried out on citation analysis in different disciplines. In this context, Line and Sandison (1974)[7] in their work stated that citation analysis of documents not only shows the relationships among journals, papers and authors, but also investigates the quality and quantity of research work. In this direction, Salton and Bergmark (1979)[8] evaluated the importance of individual authors, documents, and journals through a clustering study of computer science literature using bibliographic citations as clustering criterion. Goodrum, *et al*. (2001)[9] in their work entitled, "scholarly publishing in the internet age-a citation analysis of computer science literature" analysed two views of information production and use in computer-related research based on citation analysis of publications on the Web using autonomous citation indexing and parallel citation of journal literature indexed by the Institute of Science Information (ISI) in SCISEARCH. Shi, Tseng, and Adamic (2009)[10] examined the impact of cross-community information flows in computer science through their empirical observations of citations of citations of computer science articles, focusing specially on information flow across community boundaries and temporal gaps. Gupta, Kshitij, and Singh (2010)[11] analysed research output of computer science in 11-subfields in India during 1999 to 2008 and found that the average citations per paper registered by India's publications output were 2,10 during 1999-2006 and the cumulative collaborative publications output accounted for 19.92 percent in the cumulative output of India in computer science. Unlike the present study, Maharana, Majhi, and Sethi (2011)[12] conducted a citation analysis of top research papers in chemistry retrieved from Science Direct Database and revealed that India secured 3rd position after USA and China. However, the present study which has yet been unexplored becomes imminent.

Methodology

The present study employs required bibliometric measures for analysis of citations. Keeping the objectives of the study in mind, the authors have collected citations data from the quarterly list of hottest 25 research articles in the subject field of Computer Science from Science Direct Database. The hottest 25 articles is a free quarterly alert service which provides lists of most read articles counted by article

downloads on *Science Direct*. In the present study, 20 issues of the alert service beginning from January/March 2005 to October/December 2010 containing a total number of 495 articles (January/March, 2005 issue contained only top 20 papers) in Computer Science have been taken up for analysis.

Authorship Pattern

It is evident from Table-1(Fig-1) that out of 495 top papers, *three-authored* articles (162 articles) are little ahead than *two authored* articles (161 articles) followed by *four-authored* articles (61 articles), while single-authored articles (59 articles) are at the back foot. Hence, it is inferred that, the trend of collaborative research has taken the major part in the pinnacle of computer Science research during the stated period.

Table-1: Authorship pattern

Authors	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*Total*
Single	14	16	6	9	10	4	59
Two	30	31	30	25	29	16	161
Three	16	16	25	31	32	42	162
Four	8	13	7	13	15	5	61
Five	6	1	3	5	6	9	30
Six	1	4	1	2	2	5	15
Seven	0	1	0	0	3	0	4
> Seven	1	2	0	0	0	0	3
Total	76	84	72	85	97	81	495

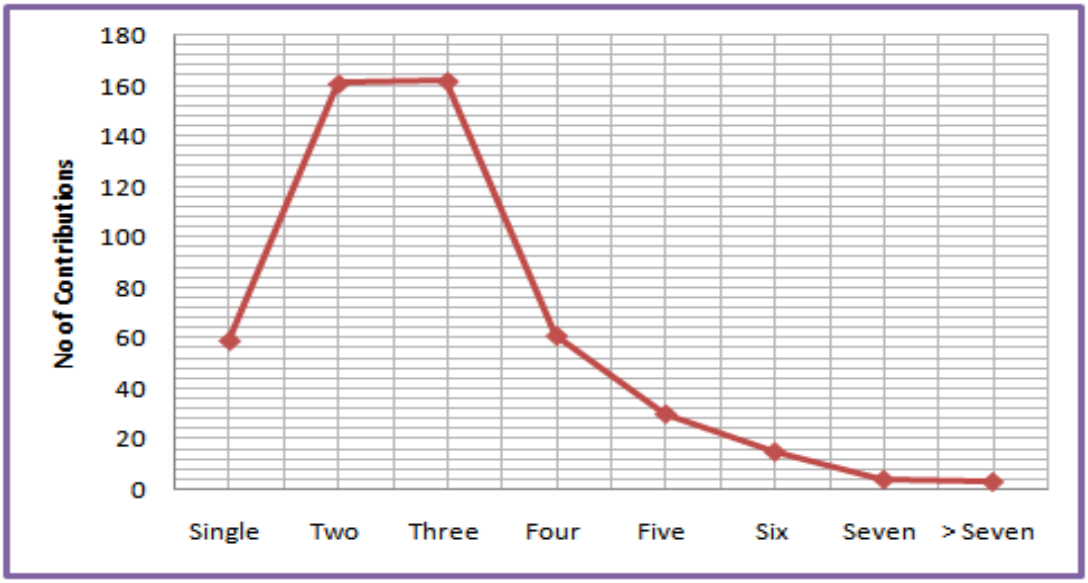

Fig-1 Authorship pattern

Ranking of Authors

Table-2 shows that Akyildiz, I. F. (USA) who has been most prominently featured for a record numbers of *ten* times in the Science Direct's hottest 25 lists tops the table. Concurrently, the 2nd rank is shared by Broberg, J (Australia), Fogel, J (USA), Ngai, E. W.T. and Wang, X (Both from China) with *six* appearances each. It is evident that the works of these authors have received momentous impact in the field of Computer Science during 2005 to 2010.

Table-2: Ranking of Authors

Author	Country	No of Citations in top 25 List	Rank
Akyildiz, I.F.	USA	10	1
Broberg, J.	Australia	6	2
Fogel, J.	USA	6	2
Ngai, E.W.T.	China	6	2
Wang, X.;	China	6	2
Yeo, C.S.	Australia	6	2
Anastasi, G.	Italy	5	3
Arseneault, J.M.	Canada	5	3
Brandic, I.	Austria	5	3
Buyya, R.	Australia	5	3
Ghosal, D.	USA	5	3
Mukherjee, B.	USA	5	3
Nehmad, E.	USA	5	3
Orr, E.S.	Canada	5	3
Ross, C.	Canada	5	3
Simmering, M.G.	Canada	5	3
Sisic, M.	Canada	5	3
Venugopal, S.	Australia	5	3
Yick, J.	USA	5	3
Grasmuck, S.	USA	4	4
Lee, W.Y.	USA	4	4
Martin, J.	USA	4	4
Mohanty, S.	USA	4	4
Orr, R.R.	Canada	4	4
Pierce, T.	USA	4	4
Shu, C.	Singapore	4	4
Vuran, M.C.	USA	4	4
Wang, W.	USA	4	4

Zhao, S.	USA	4	4
16 authors	-	3 each	5
70 authors	-	2 each	6
1178 authors	-	1 each	7

Ranking of country productivity

Ranking of country productivity in top research papers has been determined by using equal credit method in the light of calculations made by Chua, et al. (2002) [13], Lowry, *et al.* (2007)[14], and Serenko, *et al.* (2009)[15]. According to this method, if an article is contributed by n authors, each author will earn a score of 1/n for his county. Table-3 shows that the country productivity of USA is far ahead of all other 57 contributing countries followed by UK, Taiwan, China, and Canada. However, the country productivity of Russia and Korea stays at the bottom with a credit score of just 0.2 each.

Table-3: Country wise share of contributions

Sl No	*Rank*	*Country*	*Score*	*No of authors*
1	1	USA	151.74	422
2	2	UK	34.59	92
3	3	Taiwan	34.25	83
4	4	China	29.85	85
5	5	Canada	27.9	101
6	6	Spain	23.25	62
7	7	Australia	18.58	69
8	8	Germany	16.2	40
9	9	Hong Kong	14.82	42
10	10	Italy	14.46	40
11	11	France	13.59	38
12	12	Republic of Korea	10.98	28
13	13	Netherlands	10.26	30
14	14	Switzerland	10.04	34
15	15	Singapore	8.91	26
16	16	Sweden	8.09	31
17	17	Turkey	7.74	20
18	18	Greece	7.41	26
19	19	India	7.22	21
20	20	Denmark	5.32	15
21	21	Japan	5	18
22	21	Israel	5	12
23	22	Iran	4.16	13
24	23	Finland	4	11
25	24	Brazil	3.98	14

26	25	South Korea	3.25	12
27	25	Malaysia	3.25	9
28	25	Belgium	3.25	11
29	26	UAE	3	10
30	27	Mexico	2.45	9
31	28	Portugal	2.18	10
32	29	Austria	2.11	6
33	30	South Africa	1.5	4
34	30	Slovenia	1.5	3
35	30	Republic of Korea	1.5	5
36	30	Hungary	1.5	4
37	31	Thailand	1.33	2
38	32	Venezuela	1	2
39	32	Slovak Republic	1	3
40	32	Northern Cyprus	1	2
41	32	Jordon	1	2
42	32	Egypt	1	2
43	32	Czech Republic	1	2
44	32	Argentina	1	3
45	33	Belarus	0.66	2
46	34	Ukraine	0.5	2
47	34	Saudi Arabia	0.5	1
48	34	Serbia	0.5	2
49	34	Columbus OH	0.5	1
50	34	Azerbaijan	0.5	2
51	35	Norway	0.33	1
52	35	Nigeria	0.33	1
53	35	Fiji	0.33	1
54	36	Lebanon	0.25	1
55	36	Chile	0.25	1
56	37	Russia	0.2	1
57	37	Korea	0.2	1

Ranking of journals

A total number of 74 journals appeared in the list of hottest 25 articles in Science Direct released in different issues ranging from 2005 to 2010. Table-4 reveals that *European Journal of Operational Research* (78 appearances) occupies the top position followed by *Computers in Human Behavior* (52 appearances), and *Pattern Recognition* (38 appearances). The other two leading journals which occupy 4th and 5th rank respectively were *Expert Systems with Applications* (33 appearances), and *Information and management* (23 appearances). Interestingly, *Medical Image Analysis*, and *Journal of Informatics* which

have received a relatively higher impact factor of 4.248 and 3.111 respectively occupy 16th rank jointly. It may be due to the fact that articles download of these journals might have been less irrespective of their citations.

Table-4: Ranking of journals

Name of Journals	No of appearances	Impact Factor*	Rank
European Journal of Operational Research	78	2.158	1
Computers in Human Behavior	52	1.865	2
Pattern Recognition	38	2.607	3
Expert Systems with Applications	33	1.924	4
Information & management	23	2.627	5
Decision Support Systems	22	2.135	6
Journal of Computational Physics	21	2.345	7
Computer Networks	21	1.690	7
Computers & Geosciences	14	1.416	8
Ad Hoc Networks	12	1.592	9
Design Studies	10	1.354	10
Information sciences	10	2.833	10
Computers & Structures	8	1.719	11
Image and Vision Computing	8	1.525	11
Computer in Human Behavior	7	1.865	12
Future Generation Computer Systems	6	2.365	13
Signal Processing	6	1.351	13
Telecommunication Policy	6	0.963	13
Computer Methods in Applied Mechanics and Engineering	5	2.082	14
Computer Standards & Interfaces	5	0.825	14
Computer Vision and Image Understanding	5	2.404	14
Information and Organization	5	-	14
Information Processing & Management	5	1.673	14
Finite Elements in Analysis and Design	4	1.030	15
Neural Networks	4	1.955	15
Optical Fiber Technology	4	0.841	15
Advances in Engineering Software	3	1.004	16
Computational Biology and Chemistry	3	1.281	16
Computational Statistics & Data Analysis	3	1.089	16
Computer Communications	3	0.815	16
ISPRS Journal of Photogrammetry and Remote Sensing	3	2.158	16

Journal of informetrics	3	3.119	16
Journal of Molecular Structure: THEOCHEM	3	1.288	16
Knowledge-Based Systems	3	1.574	16
Medical Image Analysis	3	4.248	16
Microelectronics Reliability	3	1.066	16
World Patent Information	3	-	16
17 journals	2 times each	-	17
20 Journals	1 time each	-	18

*Retrieved from Thomson Reuters Journal Citation Report, 2011

Distribution of papers by frequency of citations

Table-5 shows that only 29 (5.86%) out of 495 articles are without citations but, they have been featured in the category of top papers may be by their quantum of downloads. However, all other top articles carried citations. It is found that more than one-third of articles were cited in the range of 1 to 10 and more than half of the articles cited a maximum up to 20 times. Undoubtedly, 52 articles which were cited more than 100 times are presumed to be some of outstanding papers carrying high impact that carry high academic interest to the teachers, research scholars, and students of computer science.

Table-5: Distribution of papers by frequency of citations

Range of Citations	No. of Appearances	%	Cumulative %
Nil	29	5.86	5.86
1 to 10	167	33.74	39.60
11 to 20	114	23.03	62.63
21 to 30	53	10.71	73.33
31 to 40	28	5.66	78.99
41 to 50	20	4.04	83.03
51 to 60	14	2.83	85.86
61 to 70	9	1.82	87.68
71 to 80	3	0.61	88.28
81 to 90	6	1.21	89.50
91 to 100	0	0.00	89.50
>100	52	10.51	100.00
Total	495	100.00	-

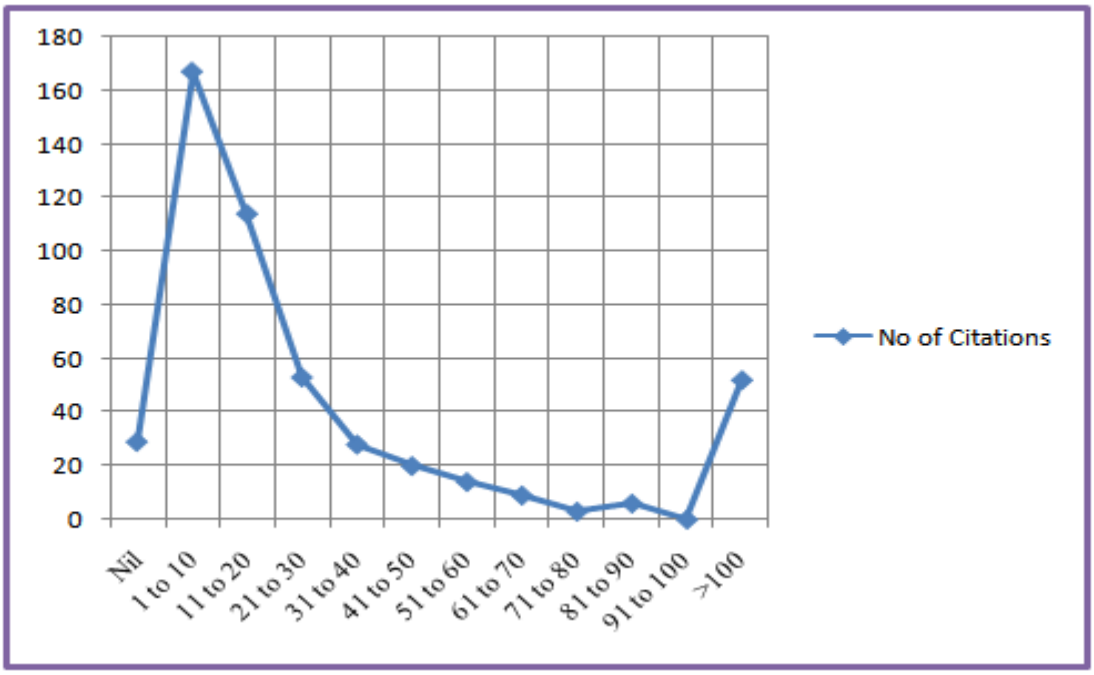

Fig-2 Distribution of papers by frequency of citations

Highly cited papers

Some of the exceptionally brilliant papers which have been so far recorded for more than 300 scopus citations are worthy of examinations. One such paper at the top is *Wireless sensor networks: a survey* contributed by Akyldiz; Su; Sankarasubramaniam; and Cayirci which has been cited for a record number of 3879 times appears to be a masterpiece which would be required to be read by any computer scientist. Similar such works which have been frequently referred to by many authors from various parts of the word are depicted in Table-6 for a view.

Table-6: Highly cited papers (top ten)

Rank	Title	Authors	No of Citations	Name of the Journals
1	Wireless sensor networks: a survey	Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E	3879	Computer Networks
2	Independent component analysis: algorithms and applications	Hyvarinen, A.; Oja, E.	1514	Neural Networks
3	Image registration methods: a survey	Zitova, B.; Flusser, J.	1469	Image and Vision Computing
4	Wireless mesh networks: a survey	Akyildiz, I.F.; Wang,	1099	Computer

		X.; Wang, W.		Networks
5	NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey	Akyildiz, I.F.; Lee, W.Y.; Vuran, M.C.; Mohanty, S.	1025	Computer Networks
6	A survey on routing protocols for wireless sensor networks	Akkaya, K.; Younis, M.	653	Ad Hoc Networks
7	Face Detection: A Survey	Hjelmas, E.; Low, B.K.	438	Computer Vision and Image Understanding
8	A survey of advances in vision-based human motion capture and analysis	Moeslund, T.B.; Hilton, A.; Kruger, V.	433	Computer Vision and Image Understanding
9	A survey on wireless multimedia sensor networks	Akyildiz, I.F.; Melodia, T.; Chowdhury, K.R.	403	Computer Networks
10	Why do people use information technology? A critical review of the technology acceptance model	Legris, P.; Ingham, J.; Collerette, P	399	Information & Management

Summary of the findings

The major findings of the study are summarized below:

- It is found that from among a total of 495 top papers; *three-authored* articles are little ahead than *two authored* articles followed by *four-authored* articles, while single-authored articles (59 articles) are at the back foot. Hence, it is inferred that, the trend of collaborative research has taken the major part in the pinnacle of computer Science research;
- It is evident that Akyildiz, I. F (USA) is ranked first. Concurrently, the 2nd rank is shared by Broberg, J. (Australia), Fogel, J. (USA), Ngai, E. W.T. and Wang, X (Both from China) with *six* appearances each. It is evident that the works of these authors have received momentous impact in the field of Computer Science;
- The analysis depicts that the country productivity of USA is at the top followed by UK, Taiwan, Chaina, and Canada;
- It is found that *European Journal of Operational Research* occupies the top position followed by *Computers in Human Behavior*, and *Pattern Recognition*. The other two leading journals which occupy 4th and 5th rank respectively were *Expert Systems with Applications*, and *Information and management*; and
- The top paper entitled, *Wireless sensor networks: a survey* contributed by Akyldiz; Su; Sankarasubramaniam; and Cayirci which has been cited for a record number of 3879 times that has in

turn carried significant impact in the field of Computer Science research.

Conclusion

21st century has witnessed remarkable developments and swift innovations in computer technologies and applications which have been resulted out of incredible researches in the field of computer science. The study unearths the epitome of outstanding researches in the subject that are of vital significance to the academic arena. All 74 journals cited in this piece of work are indeed worthy of examinations. The findings of the study can immensely help the research scholars which authors they should adopt as their role model and in which journals they would like to be published. Highly cited papers, best journals, country productivity, and certain other parameters reflected in this study may provide a solid platform to the young and energetic researchers of computer science to promote, support, and sustain future research.

References

1. Garfield, E. (1979), "Citation indexing: its theory and in science, technology, and humanities", New York: Wiley.
2. Smith, L. (1981), "Citation analysis", *Library Trends*, Vol.30, pp. 83-106.
3. Matutinovic, S. F. (2007), "Citation analysis for five serbian authors in web of science, scopus and google scholar", INFOTHECA-Journal of Informatics and Librarianship, Vol.8 No, 1/ 2, available at: http://www.unilib.bg.ac.rs/zajednica01/english/izdanja/infoteka/br1_2_2007/INFOTHECA_VIII_1-2_October2007_25a-34a.pdf (accessed 14 October, 2011).
4. Buffardi, L. C.,and Nichols, J. A. (1981), "Citation impact, acceptance rate, and APA journals", *American Psychologist,*Vol. 36,pp. 1453-1456.
5. LaBonte, K. B.(2005), "Citation analysis-a method for collection development for a rapidly developing field", *Issues in Science and Technology librarianship*, available at: http://www.istl.org/05-summer/refereed.html (accessed 10 October, 2011).
6. Sahu, A. K., Goswami, N. G. and Choudhury, B. K.(2011), "Research publications of metallurgical laboratory during the year 2001-2010: a study on citation patterns, *Annals of Library and Information Studies*, Vol.58, June, pp.151-160.
7. Line, M. B. and Sandison, A. (1974), "Obsolescence and changes in the use of literature with time", *Journal of Documentation*, Vol.30No.3, pp. 283-350.
8. Salton, G. and bergmark, D.(1979), "A citation study of computer science literature", available at: http://ecommons.cornell.edu/handle/1813/7481 (accessed 12 October, 2011).
9. Goodrum, A. A., McCain, K. W., Lawrence, S. and Giles, C. L.(2001), "Scholarly publishing in the Internet age- a citation analysis of computer science literature", *Information Processing and Management*, Vol.37, pp.661-675, available at: http://clgiles.ist.psu.edu/papers/IPM-2001-scholarly-publishing.pdf (accessed 10 October, 2011).

10. Shi, X., Tseng, B. and Adamic, L. (2009), "Information diffusion in computer science citation networks", available at: http://www.stanford.edu/~shixl/papers/CitationNetworks.pdf (accessed 12 October, 2011).
11. Gupta, B. M., Kshitij, A. and Singh, Y.(2010), "Indian computer science research output during 1999-2008-qualitative analysis", *DESIDOC Journal of Library & Information Technology*, Vol.30 No.6, pp.39-54.
12. Maharana, B., Majhi, S. and Sethi, B. B. (2011), "Citation analysis of top research papers in chemistry with specific reference to India", *Library Review*, Vol. 60 No.6, pp.501 – 512.
13. Chua, C., Coa, L.,Cousins, K. and Straub, D. W. (2002), "Measuring research-production in information systems", *Journal of the Association for Information Systems*, Vol.3 No.1, pp.23-33.
14. Lowry, P. B., Karuga, G. G. and Richardson, V. J.(2007), "Assuring leading institutions, faculty, and articles in premier information systems research journals", *Communications of the Association for Information Systems*, Vol.20, pp.142-203.
15. Serenko, A. et al.(2010), "A scientometric analysis of knowledge management and intellectual capital academic literature (1994-2008)", *Journal of Knowledge Management*, Vol.4 No.1, pp.3-23.
