Proceedings of 10th International Conference on Webometrics, Informetrics and Scientometrics & 15th COLLNET Meeting 2014

September 3-5, 2014
Technische Universität Ilmenau, Germany

Edited by
Bernd Markscheffel • Daniel Fischer •
Daniela Büttner • Hildrun Kretschmer

THOMSON REUTERS

TECHNISCHE UNIVERSITÄT ILMENAU

Elsevier
Research Intelligence

Proceedings
of
10th International Conference on
Webometrics, Informetrics and Scientometrics &
15th COLLNET Meeting 2014

September 3-5, 2014

Technische Universität Ilmenau, Germany

Edited by

Bernd Markscheffel,

Daniel Fischer,

Daniela Büttner and

Hildrun Kretschmer

Bernd Markscheffel,
Daniel Fischer,
Daniela Büttner and
Hildrun Kretschmer

Technische Universität Ilmenau
Fakultät für Wirtschaftswissenschaften und Medien
Institut für Wirtschaftsinformatik
P.O. Box 100565
98684 Ilmenau
Germany

bernd.markscheffel@tu-ilmenau.de
daniel.fischer@tu-ilmenau.de
daniela.buettner@tu-ilmenau.de
kretschmer.h@onlinehome.de

Ilmenau, 2014

Index

Invited Papers

Weight of Webometrics Criteria using Entropy Method

Handaru Jati

Universitas Negeri Yogyakarta Karangmalang Campus Indonesia
handaru@uny.ac.id

Abstract

The aim of this study was to propose some of the basic tools for Decision Making. The purpose of this paper is to show a methodology test for the selection of the weighted method, as aid to decision making in the design stage in the area of webometrics. Selecting the weighted method is one of the problems of Multicriteria Decision Analysis in which decision-makers have had disadvantages in weighting assignment criteria. To resolve this problem arises weighting variables using the entropy method. The model presented in this article is limited to display application in a webometric case. This model can be applied as a way to supplement the technical studies to select the weighted method of a webometrics and it gives the relative importance weights of the various elements, and gives an empirical analysis, explain the role of the entropy weight in webometrics study. Entropy weighted method enables rank all the alternatives in question without decisor bias and calculates the specific weight of criteria.

Introduction

In the world there are thousands of universities, and since 2004 it has been published a Web Ranking whichshows the results in every six months (January and July) and covers about 20,000 Higher Education Institutions worldwide. The composite index (Ranking) is calculated by combining standardized positions instead of values. The visibility is calculated giving an extra inbound links that are not from generic domain importance (.Com, .Org, .Net). Figures for rich files (pdf, doc, ppt, ps, Dox, pptx, eps) are combined and have not been treated individually. The intention with this system of analysis and projection of cybermetric indicators under the parameters set Webometrics is to strengthen and indicate the type of information being generated in each of the institutions and thereby improve certain characteristics that further enrich university of university webometrics ranking has changed the setting of higher education and is likely to continue to influence further development nationally and internationally. This moment is a new era for university, characterized by global competition, in which university ranking systems have assumed an importance factor for surviving. Their emergence *has* also *been* a matter of *controversy*, often controversial and subject to considerable debate, has been met with a lot of scepticism, some enthusiasm and an institutional unease. Academic rankings are here to stay and it is results that count for most of higher education's stakeholders.

Literature Review

Webometrics

Although the subfield of webometrics is considered as one of the most recent quantitative studies within the field of library and information science, there are already several international studies that address this topic. Many authors have directed their focus of study for this new environment, for finding web immense diversified network of information resources, easily accessible and still little explored. In this sense, Cronin and McKim (1996) argue that as the Web is becoming a medium increasingly important to science and academia , it is logical that quantitative studies extend well to this medium. Also Thelwall, Vaughan and Björneborn (2003) consider that being a global network of Web documents, initially developed for academic use and then extended to general users , it is obvious that it is a fertile field of research for

bibliometrics , the scientometrics and informetrics. The Webometrics is a ranking based on measurements of the presence of the universities on the Web. It is prepared by the Laboratory Cybermetrics, a group of research is part of the Superior Council of Scientific Research of Spain, and not for commercial purposes. In contrast to other rankings, Webometrics classifies a large number of universities, more than 20,000 in its latest edition (January 2012). Published twice a year (January and July). The system also allows universities ordered by country and region (Aguillo, Ortega et al. 2008). According to its website, the ranking aims to promote open access to information on the Internet by universities access. Also, as most of the rankings, insist on the superiority of his method: "As other rankings focused only on a few relevant aspects, specially research results, our ranking based on indicators of the presence reflects best the Web overall activity of the institutions, as there are many other tasks performed by teachers and researchers that appear on the Web. However, this method also has its limitations, since it favors large universities or those with large budgets for technology.

Entropy Method

The entropy method was developed as an objective method of allocation weights depending on the decision matrix without affecting the preference of the decision maker (Zeleny 1982), the relative importance of criterion j in a decision situation, wj measure its weight is directly related to the amount of information provided by the intrinsically set of alternatives with respect to that criterion (Barba Romero and Pomerol 1997). How much have greater diversity in the evaluations of the alternatives greater importance should be the criterion. Far this diversity is conceptually based on solid and accepted concept of entropy in an information channel posed by Claude Shannon (Shannon and WEAVER 1949) . The procedure is as follows:

a. The evaluations ij (i = 1, m) (j = 1, n) are taken as normalized as a fraction of the sum i ij Σ to the original assessments of each criterion j.

$$a_{ij} = \frac{k_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} k_{ij}} \quad \text{for m > 1 and i=1, 2, …, m; and j=1, 2, …, n.} \qquad (1)$$

b. Entropy (Ej) is calculated.

$$E_j = [\frac{-1}{\ln(m)}] \sum_{i=1}^{m} [a_{ij} \ln(a_{ij})] \qquad (2)$$

where m = number of alternatives in the matrix standardized assessments and ij = Criteria or standardized attributes.

c. Diversity criterion (Dj) is calculated.

$$D_j = 1 - E_j \qquad (3)$$

d. The normalized weight of each criterion (Wj) is calculated.

$$w_j = \frac{D_j}{\sum D_j} \qquad (4)$$

Research Method

Weighted indicators that take into account are:

- Size: number of pages recovered from 4 search engines: Google, Yahoo, Live Search and Exalead (20%).
- Visibility: The total number of unique external links received (inlinks) by a site that you can den get consistently from Yahoo Search, Live Search and Exalead (50%).
- Rich files: the following file formats were selected after considering their relevance in academic and publication activities and considering the volume of use: Adobe Acrobat (pdf.), Adobe PostScript (ps.), Microsoft Word (. Doc) and Microsoft Powerpoint (.

Ppt). These data are extracted through Google, Yahoo Search, Live Search and Exalead (15%).

- Academic: Google Scholar provides the number of papers and citations for each domain academic. The results obtained from the database of Google Scholar papers, reports and other academic papers (15%).

Results

The four number of criteria that should typically be considered in selecting the best university website are Size(C1), Visibility (C2), Rich Files (C3), and scholar (C4). First of all we form the decision matrix, after that we compute hi ,di and wi base on Shannon method that are shown in Table 1.

Table 1. Data

Universitas	Size	Visibility	Rich Files					Scholar
			.pdf	.ps	.ppt	.doc	Total	
Uni A	9950	177,321	259000	84200	9110	22900	375210	9950
Uni B	8970	307,113	390000	26400	10800	13400	440600	8970
Uni C	33200	4.616,437	317000	22300	18100	19900	377300	33200
Uni D	30100	362,854	268000	10100	8650	22800	309550	30100
Uni E	26700	113,286	269000	12900	20000	20500	322400	26700

We want to obtain a weight for each criterion by using the proposed approach. According to Eq.1, normalized matrix data are presented.

Table 2. Normalized Data

Size	Visibility	Rich Files	Scholar
0,040	0,014	0,501	0,040
0,000	0,043	1,000	0,000
1,000	1,000	0,517	1,000
0,872	0,055	0,000	0,872
0,732	0,000	0,098	0,732

The evaluations of these five alternatives according to the previously stated criteria, i.e., evaluation matrix, are displayed in Table 3.

Table 3. Normalized Data

Size	Visibility	Rich Files	Scholar
0,960	0,986	0,499	0,960
1,000	0,957	0,000	1,000
0,000	0,000	0,483	0,000
0,128	0,945	1,000	0,128
0,268	1,000	0,902	0,268

In our analysis we calculate diversity criteria and the result shows in the table 4.

Table 4. Diversity Criterion

	Size	Visibility	Rich File	Scholar
	-0,665112338	-0,683290598	-0,345859551	-0,665112338
	-0,693147181	-0,663312374	0	-0,693147181
	0	0	-0,334805162	0
	-0,088681645	-0,654732445	-0,693147181	-0,088681645
	-0,185945385	-0,693147181	-0,625181204	-0,185945385
E(C) = ln(2) *total sum	**0,471151465**	**0,777463336**	**0,576787486**	**0,471151465**
d = 1-E(C)	**0,528849**	**0,222537**	**0,423213**	**0,528849**

The final rank of each criterion by using the entropy weighted method can be seen in table 5. The obtained values of criterion Size, visibility, rich files, and scholar are 0,310458; 0,130639; 0,248445; and 0,310458 respectively.

We see that the rank of size and scholar are just better than the rank of rich file and visibility. Therefore, size locates at rank 1. Other criteria can be ranked in the same way. For problems with more complexity, with a small program (for example Excel) we can determine the rank of each criterion. In the last Table 5, the rank of each criterion can be seen.

Table 5. Weight of Criterion

Criteria	Weight (W) = d/total
Size	**0,310458**
Visibility	**0,130639**
Rich File	**0,248445**
Scholar	**0,310458**

Conclusion

There are several methods for obtaining the weights of criteria of an MADM problem, one of which is the entropy method.How to ascertain weights and subjectivity of evaluation model are the main aspects which influence evaluation result in the present quantitative evaluation methods. During ascertaining weights, either subjectivity can't be avoided, or calculation is too complex. On the other hand, subjectivity can't be avoided in some evaluation methods. based on entropy weight can avoid not only subjectivity or complex calculation in ascertaining weights but also subjectivity of evaluation model via the evaluation criteria of weighted relative adjacent degree. Entropy weighted method is a new advancement in quantitative evaluation methods for webometrics.

References

Aguillo, I. F., J. L. Ortega, et al. (2008). "Webometric ranking of world universities: Introduction, methodology, and future developments." *Higher Education in Europe* 33: 233-244.

Barba Romero, S. and J. Pomerol (1997). Decisiones Multicriterio Fundamentos Teóricos y Utilización Práctica. *Universidad de Alcalá de Henares*.

Shannon, C. E. and W. WEAVER (1949). *The Mathematical Theory of Communication*. Chicago, University of Illinois Press.

Zeleny, M. (1982). *Multiple Criteria Decision Making*. New York, Mc Graw Hill.