RANKING JOURNALS, CONFERENCES AND AUTHORS IN COMPUTER GRAPHICS: A FUZZY REASONING

Sajid Hussain
Department of Systems and Software Engineering
Blekinge Institute of Technology, Sweden
shu@bth.se

Hakan Grahn
Department of Systems and Software Engineering
Blekinge Institute of Technology, Sweden
hgr@bth.se

ABSTRACT

Rankings of different research bodies are of particular interest for academia and bibliometrics is used to measure the quality of these research bodies. Different factors affecting this quality have been proposed. In this paper, we have demonstrated a new approach based on fuzzy models and taken into account different proposed factors to access the overall quality. Our model actually refines the ranking by adding more factors which affects this quality. Thus, improving the information retrieval system based on human reasoning. A dimensionless index called Fuzzy Index (FI) has been proposed and used to shuffle the previously ranked research bodies. We have successfully demonstrated the application of our FI in ranking journals, conferences and authors in the field of computer graphics in particular and computer science in general.

KEYWORDS

Information retrieval, Fuzzy ranking, Bibliometrics, Data mining.

1. INTRODUCTION

The ranking of academic journals, conferences and authors is an important but contentious issue. Bibliometrics is the branch of science which measures and analyse the quality of these research bodies. Within a given field, it describes specific patterns of publications (Cathrine 2005). The influence of researchers, impact of journals and quality of venues are very important factors for decision making and ranking. Today's modern internet technologies, availability of large digital libraries, databases and overflow of information have made it a difficult task to categorise different research bodies within a field of interest. This gives rise to fuzzy sets instead of classical or normal set theory.

Computer graphics in specific and computer science in general with its rapid growth in the last three decades has become one of the major research domains. The fast growing research pace and abruptly changing technologies has given rise the need to deliver findings and results to research community faster and quicker. Peer-review conferences have become a major trend now and conferences with strict acceptance criteria (10% - 30%) play an equally important role as that of established journals in that particular area (Rahm & Thor 2005). Also with the fast growth of computer science, the growth of conferences and journals has also increased over the last decade (DBworld 2008) (Rahm & Thor 2005).

The idea behind is to search for the most valuable and important research bodies (journals, conferences, publications and authors) within the focused area. The problem could be formulated as Z is a set where $z \in Z$ contains multiple parameters or attributes like {impact factor, h-index, programme committee, sponsors, acceptance rate …} and we have to identify a set of research bodies $R \subseteq Z$ such that $r \in R$ satisfies some conditions and constraints. Different important bibliometrics like impact factor (Garfield 1983) and h-index (Hirsch 2005) is associated with the research bodies to quantify their qualities. That is, if a certain research body exhibits bibliometrics above a certain threshold, it is categorised in higher quality category. In this paper, we gather a fuzzy set of all important bibliometrics of a research body to be quantified. Based on these parameters, we classify some top ranks within a particular research body. These top ranks then further analysed with a set of some fuzzy parameters like research body themes, relatedness index, locations, programme committees, sponsors or co-authors (to quantify individual authors) etc. We use human reasoning and formulate a set of rules for individual research body. Fuzzy inference models combined with human reasoning called fuzzy reasoning are used to classify the importance of a particular research body.

The rest of the paper is organised as follows. Section 2 gives some previous related work. Section 3 gives basics of fuzzy inference and fuzzy logic. Section 4 identify top ranked journals in computer graphics and further analyse them in detail to refine the ranking. Section 5 focuses important conferences in the area of interest. Section 6 describes top authors in the area. Finally, section 7 concludes the paper with future directions.

2. PREVIOUS WORK

Bibliometrics deals with the quality of research bodies and it is very important measure for research organisations, universities or individual researchers to identify high quality research bodies to publish their work. Impact factor (Garfield 1983), a very important bibliometric in this area and is defined as "The average number of time, published papers are cited within two years after its publications", is used to rank high quality journals in particular field of interest. Different researchers have proposed different metrics to quantify research bodies like h-index (Hirsch 2005) and g-index (Egghe 2006) are used to measure the impact of a scientist in a particular field. Moreover, many researchers (Newman 2001) (Barabasi et al. 2002) (Elmacioglu & Lee 2005) proposed and analysed different scientific networks and their collaborations.

What we have described in this paper is not an improvement to existing parameters, but a model which analyses the existing parameters and adds the effect of these parameters with some weighting function to the overall effect to rank research bodies. Some work has already been done in this context like in (Zhuang et al. 2007); conference quality has been measured by mining programme committee characteristics and use usage based metrics to detect scholarly impacts.

3. FUZZY INFERENCE AND FUZZY LOGIC

The idea was first introduced by Lotfi Zadeh (Zadeh 1965). Fuzzy logic deals with approximate reasoning. It is the theory of relative importance of precision. Like in real world, human have imprecise reasoning to elaborate or understand some physical process or problem, fuzzy logic theory comes in and it is the ability of this theory to compare human imprecise reasons in some sense to precise ones understandable by computers. Problems like launching satellites, communication between two points with laser beams etc. are required with precise logical reasoning and hence outside of the scope of fuzzy systems so far. But human's daily problems like driving a car, judging a beauty contest; allocating budgets and many more are solved with human imprecise reasoning and hence controlled by fuzzy systems (Ross 2004). These problems are presented by many linguistic labels like unclear, approximate, amorphous, random, unpredictable, conflicting and dissonant etc.

Fuzzy systems work by mapping input space to output space and this is actually the working principle of everything. In contrast, fuzzy process takes into account the human imprecise reasons to produce such mapping. Following are some examples of human reasoning about a tipping problem in a restaurant.

*"If the service is **poor** then tip is **cheap**"*
*"If the service is **good** then tip is **average**"*
*"If the service is **excellent** then tip is **generous**"*

These are just if-then statements in any programming language we write. Although, we can make the above system work with some other means, fuzzy is cheaper and faster (Zadeh 1965). We can combine as many reasons as we like in the input space. For example if we want to add food quality also we can reconstruct the above reasons like

*"If the service is **poor** or food is **rancid** then tip is **cheap**"*
*"If the service is **good** then tip is **average**"*
*"If the service is **excellent** or food is **delicious** the tip is **generous**"*

Here we have combined food quality factor in our first and third reason. These three reasons are the heart of our fuzzy system. The fuzzy set theory is a bit different from classical set theory. In classical set theory, set A is a mapping for the elements of set S to the set $\{0,1\}, i.e, A: S \rightarrow \{0,1\}$ and the characteristic function $\mu_{A(x)}$ is

$$\mu_{A(x)} = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases}. \tag{1}$$

In fuzzy set theory, set F is a mapping for the elements of set S to the interval $\{0,1\}, i.e, F: S \rightarrow [0,1]$ and the characteristic function $\mu_{F(x)}$ would be between the interval $0 \le \mu_{F(x)} \ge 1$, where 1 means full membership, 0 means no membership and any value between them means graded membership. The process of constructing the system (mapping input space to output space) based upon imprecise reasons or fuzzy logic is called fuzzy inference. A very popular fuzzy inference system was proposed by Mamdani (Mamdani & Assilian 1975) based on theory presented by Zadeh (Zadeh 1965). The process is shown in Figure 1 for the tipping problem described above (Fuzzy Logic Toolbox 2008).

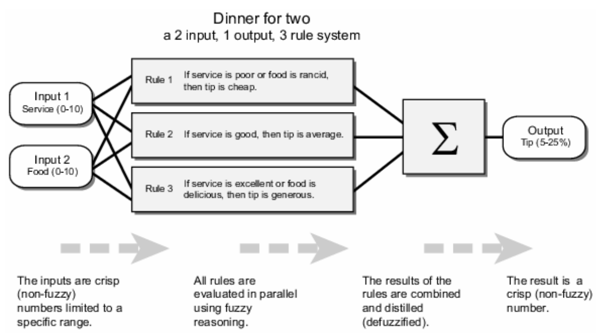

Figure 1. Main steps in fuzzy inference

In Figure 1, there are two inputs to the system, service and food. Both the inputs are given numbers from 0-10 based on their qualities. There is one output called tip which varies from 5% to 25% based upon the quality of two inputs and the three rules who governs the system. The first step in building the system is to fuzzify the inputs according to some member function and calculate to which degree a given input belongs to the fuzzy set. Different member functions are used to fuzzify the inputs which include triangular, trapezoid, Gaussian, sigmoid, bell and their various versions. If there is more than one input, a fuzzy operator is applied to obtain one number that depicts the effects of all the inputs to the system. An implication method is then applied which gives the fuzzy set represented by the membership function and which depicts the consequences of linguistic characters found in the reasoning. We then aggregate the output fuzzy sets of each rule and finally, we defuzzify. Defuzzification takes a fuzzy set as an input and produces a single crisp number that best represents that set. The common method in defuzzification is finding the centroid. More elaborated theory on fuzzy sets can be read from any good book on fuzzy logic like (Ross 2004). Figure 2 describes the whole fuzzy inference process in detail (Fuzzy Logic Toolbox 2008).

4. JOURNAL RANKING

Stated preferences and revealed preferences are two major approaches in ranking journals (Mingers & Harzing 2007). We will analyse journals ranking based upon revealed preferences by analysing publication patterns and impact factors (IF) of journals from ISI web of knowledge journal citation reports (JCR).

Table 1 lists top 10 computer graphics journals ranked according to their impact factor (IF). Relatedness index (RI), immediacy index (II) and fuzzy index (FI) are also shown in Table 1. Relatedness index (RI) is calculated here as the average of maximum relatedness (Rmax) values for a particular journal to other journals in the ranking. Relatedness values are calculated as the number of citations a particular journal has to its related journal and vice versa. We have taken average of these values scaled down within the range of [1 5] and in RI/N field; N is the number of journals the current journal is related with e.g. 4.2/8 in the Table 1 means the journal "ACM T GRA" has RI of 4.2 and is related to 8 other journals in the table. This relationship factor N is also considered in the final calculation of FI. II measures how fast in time the proceedings of a journal are cited and in fact should not be ignored while ranking journals. Table 1 shows the journal in computer graphics with different parameters. The ranges used are IF [0 5], RI [0 5], N [0 5], II [0 0.30] and FI [0 10]. Fuzzy index (FI) is calculated on the bases of four parameters IF, RI, N and II. The following statements are made for fuzzy reasoning and a fuzzy inference model has been implemented in MATLAB®.

*"If IF is **low** and RI is **less** and N is **less** and II is **small** then FI is **low**"*
*"If IF is **high** and RI is **less** and N is **less** and II is **small** then FI is **average**"*
*"If IF is **high** and RI is **average** and N is **average** and II is **medium** then FI is **good**"*
*"If IF is **high** and RI is **good** and N is **good** and II is **large** then FI is **better**"*
*"If IF is **highest** and RI is **good** and N is **good** and II is **large** then FI is **best**"*

Table 1. Journal ranking : computer graphics

Rnk	**JName**	**IF**	**RI/N**	**II**	**FI**
1	ACM T GRA	4.081	4.2/8	0.243	9.03(1)
2	IEEE T VIS C GRA	1.794	2.2/7	0.212	7.65(2)
3	IEEE COMP GRA	1.429	4.4/5	0.275	7.41(3)
4	COMP GRA FORU	1.164	3.5/7	0.095	6.68(4)
5	VISUAL COMP	0.708	3.1/7	0.101	5.87(5)
6	GRAPH MODEL	0.702	2.6/5	0.033	2.04(7)
7	COMP AN V WOR	0.644	2.5/4	0.041	1.96(8)
8	COMP GRAPH UK	0.601	2.1/6	0.132	4.70(6)
9	COMP GRAPH US	0.536	0.0/0	-	1.89(9)
10	J OF COMP GEOM	0.449	0.8/1	0.034	1.78(10)

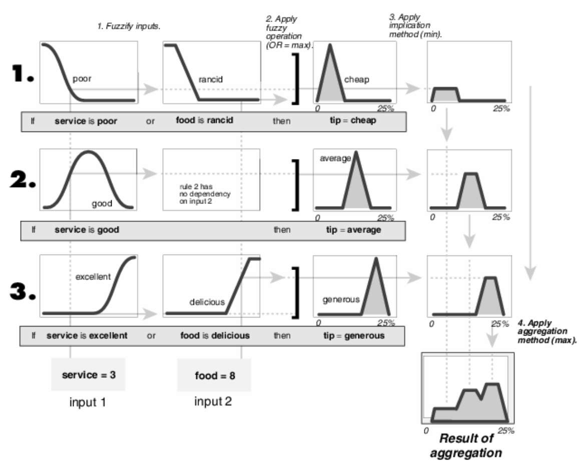

Figure 2. Detailed steps in fuzzy inference

Note that FI ranges from 0 (low) to 10 (best) and according to FI, journal "COMP GRAPH UK" should take number 6 in the final ranking as it has high II and N as compared to "GRAPH MODEL" and "COMP AND V WOR". In this fuzzy inference model, we have used Gaussian rules for input parameters and triangular rule for output parameter. We call this fuzzy index (FI) as we can increase number of rules as large as possible and we can include other parameters as well like programme committee characteristics, number of times a journal is printed each year etc. We have included II here as it is very important factor in this fast growing research environment. The final surface which decides FI is complex; Figure 3 shows a deciding surface (generated by fuzzy inference model) only dependent on two parameters just for visual simplicity.

Similarly Table 2 lists top 10 journals in computer science field (where computer graphics is the sub-field of computer science). The ranking is based on impact factor (IF) from ISI web of knowledge journal citation reports (JCR). Here, we have used RI as the total number of journals the particular journal is related with. The value is further scaled down into the interval [0 35]. Further, there is no N parameter in this case. We have made a fuzzy inference model depending upon the above statements but without N field. The FI column of Table 2 lists the new ranking of the journals according to the fuzzy inference model. As you can see from Table 2 that the journal "INT J COMP VIS" despite its high IF has taken rank 5 and "BIOINFORMATIC" and "IEEE T MED IMA" has taken ranks 1 and 2 respectively. The reason behind is the fast and rapid growing medical research in computer science instead of other fields. So, the two medical related journals have taken higher rank because of their RI and II. The parameter ranges are defined as follows. IF [0 10], RI [0 35], II [0 0.50] and FI [1 10].

Table 2. Journal ranking : computer science

Rnk	JName	IF	RI	II	FI
1	INT J COMP VIS	6.085	5.1	0.483	4.70(5)
2	ACM T INF SYS	5.059	1.6	0.212	2.87(7)
3	BIOINFORMATIC	4.894	33.6	0.712	8.04(1)
4	MIS QUART	4.731	2.8	0.610	3.99(6)
5	IEEE T PATT ANA	4.306	11.1	0.489	5.07(3)
6	ACM COMP SURV	4.130	4.4	0.077	2.72(9)
7	ACM T GRA	4.081	3.3	0.243	2.74(8)
8	J AM M INF ASS	3.979	6.7	0.587	4.80(4)
9	IEEE T EVO COM	3.770	3.1	0.200	2.64(10)
10	IEEE T MED IMA	3.757	12.5	0.532	5.21(2)

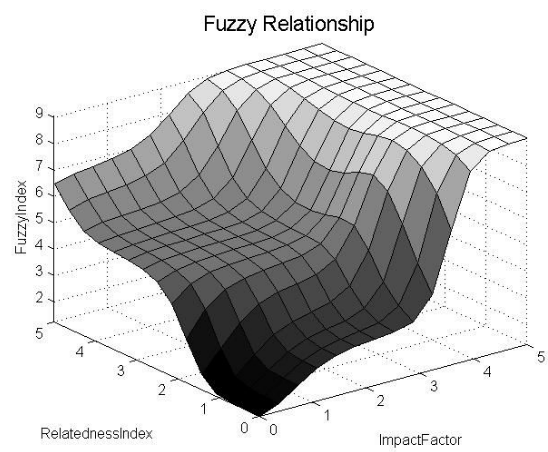

Figure 3. Fuzzy relationship of two parameters: journal ranking

5. CONFERENCE RANKING

Conference ranking is as important as journals ranking. Due to the facts describes in section I, conferences are increasing in number each year as it is a fast way of making your research public and conferences with strict reviewing criteria are considered to be as equally important as journals in that particular area. The main factors affecting the ranking of conferences includes quality of committee members, ratio of acceptance of papers, impact of papers, convenient locations, relationship to industry (sponsors) and many more.

We have used a number of resources to rank conferences in our focused field of computer graphics (Bhowmick 2007) (Zaiana 2008) (Core 2007). Table 3 lists important conferences in the field of computer graphics. Column 3 of Table 3 shows estimated impact of conference (EIC) from (CS 2008). Column 4 shows average acceptance rate for the last 4-5 years taken from (Ulicny 2005). The following statements have been made for fuzzy reasoning.

*"If EIC is **poor** and AR is **relaxed** then FI is **average**"*
*"If EIC is **poor** and AR is **average strict** then FI is **average**"*
*"If EIC is **average** and AR is **average strict** then FI is **good**"*
*"If EIC is **good** and AR is **average strict** then FI is **better**"*
*"If EIC is **good** and AR is **strict** FI is **best**"*

Column 4 of Table 4 shows the fuzzy index (FI) and according to FI, conferences ranking has been changed. Note that "EUROGRAPHICS" has gone to third rank despite its EIC equal to 0.97. Acceptance rate (AR) has affected the results and actually shuffled the rankings. The ranges EIC [0.90 1], AR [0.15 0.40] and FI [0 10] are used in the fuzzy inference model. Table 4 shows top conferences in computer science.

Note that our fuzzy inference model has correctly ranked the conferences according to their acceptance rate. Two conferences, "VLDB" and "ISCA" have acceptance rates 16.8% and 16.7% respectively. Our fuzzy inference model ranks them equal and "SIGCOMM" wins with its acceptance rate of 10.3%. In this case EIC is constant and we have used the ranges AR [0 0.30] and FI [0 10].

Table 3. Conference ranking : computer graphics

Rnk	**JName**	**EIC**	**AR**	**FI**
1	SIGGRAPH	0.97	0.216	8.88(1)
2	IEEE VISU	0.97	0.356	8.58(4)
3	I3DG	0.97	0.302	8.80(2)
4	EUROGRAPHICS	0.96	0.236	8.75(3)
5	CGI	0.96	0.400	5.00(6)
6	GH	0.96	0.390	5.31(5)
7	IEEE SYM ON RT	0.95	0.530	4.22(8)
8	EGSR	0.94	0.368	5.23(7)

Table 4. Conference ranking : computer science

Rnk	**JName**	**EIC**	**AR**	**FI**
1	SIGMOD	0.99	0.160	8.85(2)
2	VLDB	0.99	0.168	8.84(3)
3	AAAI	0.99	0.238	8.72(4)
4	ISCA	0.99	0.167	8.84(3)
5	SIGCOMM	0.99	0.103	8.91(1)
6	INFOCOM	0.99	0.200	8.51(7)
7	FOCS	0.99	0.272	8.66(5)
8	LICS	0.99	0.298	8.60(6)

6. AUTHORS RANKING

Authors are ranked according to different factors like number of papers in a particular field of which papers in top ranked conferences and journals. An important parameter is h-index of an author proposed by J.E.Hirsch (Hirsch 2005). The h-index actually shows the scientific impact and productivity of an author. Various other indexes are also used to measure the impact of a researcher in a particular field. The one other than h-index is g-index suggested by Leo Egghe (Egghe 2006). The g-index is an advanced version of h-index addressing its shortcomings. The main database we have used in ranking authors as harzing database (www.harzing.com). We have taken names of major researchers (section VI) in computer graphics who contributed their work in top ranked conferences and journals. We have also searched top authors in ISI web of knowledge database under the field of computer graphics. We have then analyzed different parameters and indexes related to these researchers using harzing's "publish or parish" software. Table 5 shows the authors ranking in computer graphics. We have used five parameters to rank different authors in this field. Column 3 of Table 5 shows number of articles (NA) of a researcher and in our fuzzy inference model, its range is from [60 240]. Column 4 shows h-index with range [10 40]. Column 5 shows g-index with range [15 70]. Column 6 shows first author (FA) values in range [15 40]. These are the number of articles where the particular researcher was the principal author. Column 7 shows co-author index (CAI) and range is [0 1]. The co-author index has been calculated by dividing total number of articles with total number of co-authors participated in these articles (author instances appearing more than once are counted each time means a co-author participating in 10 articles has been counted 10 times). The CAI is good if its value approaches 1. The following rules have been defined to be used in fuzzy inference model.

*"If NA is **less** and HI is **low** and GI is **low** and FA is **less** and CAI is **low** then FI is **low**"*
*"If NA is **less** and HI is **average** and GI is **average** and FA is **medium** and CAI is **average** then FI is **average**"*
*"If NA is **less** and HI is **high** and GI is **high** and FA is **medium** and CAI is **average** then FI is **above average**"*
*"If NA is **average** and HI is **average** and GI is **average** and FA is **medium** and CAI is **average** then FI is **above average**"*
*"If NA is **average** and HI is **high** and GI is **high** and FA is **medium** and CAI is **average** then FI is **above average**"*
*"If NA is **average** and HI is **high** and GI is **high** and FA is **large** and CAI is **average** then FI is **high**"*
*"If NA is **average** and HI is **high** and GI is **high** and FA is **large** and CAI is **high** then FI is **high**"*
*"If NA is **high** and HI is **high** and GI is **high** and FA is **large** and CAI is **high** then FI is **highest**"*

Note that in Table 5, author "I Wald" has been ranked at position 3 as compared to author "P Slusallek" at position 4 despite his number of articles (NA), HI and GI more than "I Wald" and CAI is also better, 0.51 as compared to 0.47. The reason is because of first author (FA). The author "P Slusallek" has been a first author in his articles 16.3% times and "I Wald" was 43.4%. Similar is the case with "J Dorsey" if compared to "I Wald". The deciding surface in our fuzzy inference model is complex and Figure 4 shows the fuzzy index (FI) dependent only on two variables NA and HI for visual clarity.

Table 5. Authors ranking : computer graphics

Rnk	Author	NA	HI	GI	FA	CAI	FI
1	M Gross	240	32	60	20	0.55	7.8
2	G Drettakis	116	25	42	26	0.61	3.2
3	I Wald	76	17	33	33	0.47	2.8
4	P Slusallek	153	21	37	25	0.51	2.7
5	J Dorsey	76	20	40	20	0.48	1.7
6	Ph Christen	84	12	24	36	0.95	1.5
7	Y Dohashi	71	12	22	22	0.39	0.9

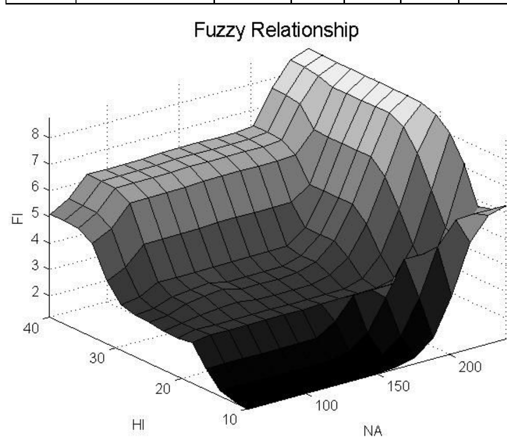

Figure 4. Fuzzy relationship of two parameters: authors ranking

7. CONCLUSION AND FUTURE WORK

We presented in this paper a fuzzy inference model works on human reasoning to refine the ranking of already ranked different research bodies in our focused area (computer graphics in particular and computer science in general). We have added different parameters which were to be considered in ranking different research bodies. We have showed that our model successfully incorporates the effect of these parameters and the more we give importance (weight) to any parameter, the more we get the effect of that parameter in the final results.

Although, the work done in this paper is limited to a certain domain and a few parameters, the proposed model is very flexible and we can add the effect of as many parameters as we like. The final ranking could change by adding or deleting input parameters. This is the main reason behind the index we have calculated and is called Fuzzy Index (FI).

Future work could be in different directions. We plan to apply the model in different domains other than computer science and computer graphics. We also plan to incorporate effects of more indexes to produce the final rankings. These indexes include programme committee characteristics, sponsors, organizers, locations and number of attendees in terms of conferences.

REFERENCES

Bhowmick, S. S. 2007,' Computer science conference ranking', [Online] Available at: http://www.ntu.edu.sg/home/assourav/crank.htm

Barabasi, A. & Jeong, H. & Neda, Z. & Ravasz, E. & Schubert, A. & Vicsek, T. 2002, 'Evolution of the social network of scientific collaborations', *Physica A*, vol. 311, pp. 590-614.

Cathrine, H. 2005, 'Bibliometrics Information Kit', *Council of Australian university librarians (CAULS)*, [Online] Available at: http://www.caul.edu.au/stats/caul20052bibliometrics.doc

Core, 2007, 'Australian ranking of ICT conferences', [Online] Available at: http://www.core.edu.au/rankings/Conference%20Ranking%20Main.html

CS, 2008, 'Computer science conference ranking', [Online] Available at: http://www.cs-conference-ranking.org/conferencerankings/topicsi.html

DBworld 2008 'Discussion email list announcing conferences', *ACM SIGMOD* [Online] Available at: http://www.cs.wisc.edu/dbworld/browse.html

Egghe, L. 2006, 'Theory and practice of g-Index', *Scientometrics*, vol. 69, no. 1, pp. 131-152.

Elmacioglu, E. & Lee, D. 2005, 'On six degrees of separation on DBLP-DB and more', *ACM SIGMOD Record*, vol. 34 no. 2, pp 33-40.

Fuzzy Logic Toolbox, *MathWorks* [Online]. Available at: http://www.mathworks.com/access/helpdesk/help/toolbox/fuzzy/index.html?/access/helpdesk/help/toolbox/fuzzy/fp351dup8.html&http://www.google.com/search?hl=en&q=fuzzy+inference

Garfield, E. 1983, 'Citation indexes for science: a new dimension in documentation through association of ideas', *Information Science Reprint*, vol. 6, pp. 468-471.

Hirsch, J. E. 2005, 'An index to quantify an individual's scientific research output', *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp.16569-16572.

Mamdani, E.H. & S. Assilian, 1975, 'An experiment in linguistic synthesis with a fuzzy logic controller', *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1-13.

Mingers, J. & Harzing, A. 2007, 'Ranking journals in business and management: a statistical analysis of the harzing dataset' *European Journal of Information Systems*, vol. 16, no. 4, pp. 1-44.

Newman, M. 2001, 'Who is the best connected scientist? a study of scientific co-authorship networks', *Physical Review*, vol 64.

Rahm, E. & Thor, A. 2005, 'Citation analysis of database publications', *ACM SIGMOD*, vol. 34, no. 4, pp. 48-53.

Ulicny, B., 2005, 'Acceptance rate statistics for publications in graphics', [Online] Available at: http://vrlab.epfl.ch/~ulicny/statistics/

Zadeh, L. 1965, 'Fuzzy sets', *Information and Control,* vol. 8, pp. 338-353.

Ross, T. J. 2004, 'Fuzzy logic with engineering applications, John Wiley, New York.

Zaiana, O. 2008, 'Computer science conference ranking', [Online] Available at: http://www.cs.ualberta.ca/~zaiane/htmldocs/ConfRanking.html

Zhuang, Z. & Elmacioglu, E. & Lee, D. & Giles, C. L. 2007, 'Measuring conference quality by mining program committee characteristics', *Joint Conference on Digital Libraries (JCDL)*, pp. 225-234.