

Impact of Double Blind Reviewing on SIGMOD Publication: A More Detail Analysis

Anthony K. H. Tung
National University of Singapore
atung@comp.nus.edu.sg

1. Introduction

In [1], a set of statistic had been provided with the conclusion that double blind reviewing make no impact on SIGMOD publication. Our studies here will show results contrary to that finding.

2. Use of Median instead of Mean

Our first study will use the median instead of the mean for our analysis. The use of median is more robust to the existence of outliers which can skew the mean drastically [2]. For example, the mean of

papers/famous person from 2001-2005 for SIGMOD is 0.912 while the median is 0.81. Also, 4 out of the 5 values used in the computation of the mean are in fact smaller than 0.912. Table 1 depict the data obtain from [1] with four additional rows at the end. Of these four additional rows, the first three are the column median for the period 1994-2005, 1994-2000 and 2001-2005 respectively. The last row computes the gain of the column median from the period 2001-2005 over the period 1994-2000. SIGMOD adopt double blind reviewing from the year 2001 onwards.

Table 1: Publication Statistics by Year Per Conference from [1].

Year	Papers/Famous Person		Fraction Famous Papers		Total Papers	
	SIGMOD	VLDB	SIGMOD	VLDB	SIGMOD	VLDB
1994	0.81	0.73	0.48	0.28	42	65
1995	0.54	0.73	0.37	0.31	36	59
1996	0.88	1.07	0.47	0.55	47	49
1997	0.92	0.85	0.55	0.38	42	55
1998	0.73	0.69	0.43	0.33	42	52
1999	0.88	0.81	0.53	0.35	42	58
2000	1	0.88	0.52	0.38	48	58
2001	0.77	0.81	0.44	0.31	44	66
2002	0.81	1.15	0.4	0.32	50	91
2003	0.85	1.15	0.4	0.34	53	84
2004	<u>*1.34</u>	1.53	0.49	0.38	69	102
2005	0.81	0.92	0.31	0.22	66	103
Median(1994-2005)	0.83	0.865	0.455	0.335	45.5	62
Median(1994-2000) A	0.88	0.81	0.48	0.35	42	58
Median(2001-2005) B	0.81	1.15	0.4	0.32	53	91
(B-A)/A	-7.95%	+41.98%	-16.67%	-8.57%	+26.19%	+56.90%

Legends: **Red Bold**: Higher than Median(1994-2005) i.e. "good year"
* : Outlier

Based on Table 1, we have the following observations

Observation 1 (OB1): In term of total number of papers accepted, both SIGMOD and VLDB see an increase of 26.19% and 56.90% respectively for the period 2001-2005 when compared to the period 1994-2000.

Observation 2 (OB2): In term of papers/famous person, VLDB approximately follow the trend in (OB1) with a growth of 41.98%. However, this is not the case for SIGMOD which see a drop of 7.95%, a rather unexpected and surprising result considering (OB1) where the number of SIGMOD paper accepted increase by 26.19%.

Observation 3 (OB3): In term of fraction famous paper, both SIGMOD and VLDB see a drop indicating that the increase in accepted papers from (OB1) are not coming from the famous people. However, the drop for SIGMOD is 16.67% which is two times that of VLDB with a drop of only 8.57%.

Based on observations 1-3, we believe that there are indications that double blind reviewing does have an impact in term of papers accepted for famous people in SIGMOD.

3. Probability of “Good Year”

We first define a “good year” for famous person to be a year in which the value is higher than Median(1994-2005). For example, in term of fraction/famous person, the good years for famous person in SIGMOD are 1996, 1997, 1999, 2000, 2003 and 2004 since their corresponding values are all higher than 0.83. This is similar in spirit to the definition of “famous person” with a more compelling reason of ensuring that there are enough samples for both classes i.e. “good year” and “bad year”.

Based on this definition, we will expect famous people to have 50% chance of having “good year” and “bad year” irregardless of DB or non-DB if they are in fact independent of each other. We have the following observations:

Observation 4 (OB4): In term of papers/famous person, we see that there is a 57% (i.e. 4/7) chance of having “good year” for SIGMOD from 1994-2000 while there is only a probability of 40%(i.e. 2/5) for having “good year” for SIGMOD from 2001-2005. This again is surprising given (OB1).

For VLDB, this trend is reversed. The chance of having “good year” for VLDB from 1994-2000 is 28.57%(2/7) while the chance of doing so from 2001-2005 is 80%(4/5). Comparing SIGMOD(2001-2005) to VLDB(2001-2005), the chance of having “good year” for SIGMOD is only half (40% vs 80%) that of VLDB which indicate that double blind reviewing halved the chance of “good year” for famous people.

Observation 5 (OB5): In term of fraction famous paper, both SIGMOD and VLDB see a drop in the probability of “good year” for 2001-2005. The explanation is similar to that of (OB3). However, the chance of “good year” for SIGMOD(2001-2005) is still only half that of VLDB(2001-2005) i.e. 20% vs 40%.

From observations 4-5, we can see that double blind review in fact reduce the probability of a “good year” for famous people by half, indicating its strong impact.

4. Final Conclusion

In this report, we make two studies which indicate that double blind review in SIGMOD do have its impact on the performance of “famous person” compared to VLDB. As mentioned in [1], there are probably a lot of other factors that must be taken into consideration before the database community makes a final choice on whether to continue with double blind review. We hope that our studies here can provide a little more information towards making the final choice.

Acknowledgement:

It is unfortunate that we have to disagree politely with the finding of Samuel and David in [1]. We nonetheless like to thank them for their effort in collecting the data, without which the study here will not be possible.

Reference:

1. Impact of Double-Blind Reviewing on SIGMOD Publication Rates. Samuel Madden and David Dewitt. SIGMOD Record, Volume 32, No. 2, June 2006.
2. <http://www.ltcconline.net/greenl/courses/201/descstat/mean.htm>