

Errata: The DCI Index

The paper by K. Järvelin & O. Persson published in JASIST 59(9):

Järvelin, K. & Persson, O. (2008). The DCI-Index: Discounted Cumulated Impact based Research Evaluation. Journal of the American Society for Information Science and Technology (JASIST) 59(9): 1433-1440.

contains an unfortunate error in one of its formulas, Formula 3.

The incorrect formulation in the original paper was:

$$dDCI[i] = \sum_{j=1 \dots i} DCI_i[j] \quad (3)$$

while the correct one is:

$$dDCI[i] = DCI_i[i] \quad (3)$$

In the original one, the summation is in excess because the calculation of the DCI already contains cumulating. The example below (also used in the original paper but slightly extended here) clarifies the calculation of *dDCI*. Despite the error in the formula, the calculations in the empirical case presented in the original paper are correct.

An Example

Assume that some scholar has received the following weighted citation loads over the past 15 years, from year 1 (past) to year 15 (current), represented as a vector of values:

$$F' = \langle 8, 0, 8, 0, 0, 3, 5, 0, 0, 0, 0, 0, 0, 0, 0 \rangle$$

The cumulated impact at a given interval i is computed by summing from position 1 to i . Formally, let us denote the i th component in the impact vector F by $F[i]$. Now the cumulated impact vector CI is defined recursively as the vector CI where:

$$CI[i] = \begin{cases} F[1], & \text{if } i=1 \\ CI[i-1] + F[i], & \text{otherwise} \end{cases} \quad (1)$$

For example, from F' we obtain $CI' = \langle 8, 8, 16, 16, 16, 19, 24, 24, 24, 24, 24, 24, 24, 24, 24 \rangle$.

The discounted cumulated impact DCI does an analogical accumulation but discounts each year's value by a discount factor based on the log of the number of years passed since that year until the current year of analysis:

$$DCI_t[i] = \begin{cases} F[1]/\max(\{1, \log_b(t-1)\}), & \text{if } i=1 \\ DCI_t[i-1] + F[i]/\max(\{1, \log_b(t-i)\}), & \text{otherwise} \end{cases} \quad (2)$$

For example, let $b = 2$ be the base of the discount logarithm and $t=15$ the current year of analysis, modeling a rapidly developing (or decaying) area. From F' given above we obtain $DCI_{15}' =$

$$\langle 2.10, 2.10, 4.33, 4.33, 4.33, 5.28, 6.95, \dots \rangle$$

The DCI_t vector is static, representing the chosen interval t of analysis. To analyze and visualize dynamic time we need a dynamic DCI vector, which accumulates and discounts past performance at each interval of evaluation. For each interval $i, 1 \leq i \leq t$, the dynamic discounted cumulated impact $dDCI$ is:

$$dDCI[i] = DCI_i[i] \quad (3)$$

For example, let $b = 2$, modeling a rapidly developing area. From F' given in the preceding section we obtain $dDCI' =$

$$\langle 8, 8, 16, 13.05, 12, 11.5, 15.09, \dots \rangle.$$

The following Table 1 exemplifies the calculation of dynamic DCI. The $dDCI$ values can be read from the diagonal of the table (the shaded cells).

Acknowledgement

We are thankful to Mr. Tarun Sharma (visiting University of Tampere, Finland, 2008), for pointing out the error.

i	DCI _i , t= 1... 15														
	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10	t=11	t=12	t=13	t=14	t=15
1	8,00	8,00	8,00	5,05	4,00	3,45	3,09	2,85	2,67	2,52	2,41	2,31	2,23	2,16	2,10
2		8,00	8,00	5,05	4,00	3,45	3,09	2,85	2,67	2,52	2,41	2,31	2,23	2,16	2,10
3			16,00	13,05	12,00	8,49	7,09	6,30	5,76	5,37	5,07	4,84	4,64	4,47	4,33
4				13,05	12,00	8,49	7,09	6,30	5,76	5,37	5,07	4,84	4,64	4,47	4,33
5					12,00	8,49	7,09	6,30	5,76	5,37	5,07	4,84	4,64	4,47	4,33
6						11,49	10,09	9,30	7,65	6,87	6,37	6,00	5,71	5,47	5,28
7							15,09	14,30	12,65	10,03	8,87	8,15	7,64	7,26	6,95
8								14,30	12,65	10,03	8,87	8,15	7,64	7,26	6,95
9									12,65	10,03	8,87	8,15	7,64	7,26	6,95
10										10,03	8,87	8,15	7,64	7,26	6,95
11											8,87	8,15	7,64	7,26	6,95
12												8,15	7,64	7,26	6,95
13													7,64	7,26	6,95
14														7,26	6,95
15															6,95

Table 1. An example of calculating the dynamic DCI index for the sample citation data F'.