

on a 5-star scale, with half-star increments (0.5 stars - 5.0 stars). Users are represented by an ID without other information. Here we give a simple example with a small number of movies and users. We randomly choose six famous movies and six users who had seen more than 500 movies (including all the six movies). The movies are *Forrest Gump* (1994), *Star Wars: Episode IV - A New Hope* (1977), *Toy Story* (1995), *Dead Men Don't Wear Plaid* (1982), *The Fugitive*, (1993), *Seven* (1995), and the userID are 15, 23, 30, 73, 212, 213. We obtain the following score matrix

$$X = \begin{bmatrix} 2 & 3 & 4 & 5 & 3 & 3 \\ 5 & 4.5 & 4 & 5 & 3.5 & 2.5 \\ 1 & 4.5 & 5 & 5 & 4 & 2 \\ 3 & 3.5 & 5 & 4 & 5 & 2.5 \\ 5 & 3.5 & 5 & 4 & 3 & 4 \\ 5 & 4.5 & 4 & 4.5 & 4 & 5 \end{bmatrix}.$$

We run our algorithm on this score matrix, the final scores of these six movies are $\mathbf{v}_1 = \{3.40, 4.12, 3.71, 3.88, 4.07, 4.44\}^1$, while the final scores of the average method is $\mathbf{v}_2 = \{3.33, 4.08, 3.58, 3.83, 4.08, 4.50\}$. The signal-to-noise ratio SNR of our result is 21.7, which is greater than 18.8 for the average method. This is a computational example on a small number of samples to show the computational steps of our algorithm and compare the results. Next, we give an example on a big data set.

The next example is about ranking universities according to the publications. Our data is from the web site of CWTS Leiden Ranking (<http://www.leidenranking.com/>). The CWTS Leiden Ranking 2016 offers key insights into the scientific performance over 800 major universities worldwide. A sophisticated set of bibliometric indicators provide statistics on the scientific impact of universities and on universities involvement in scientific collaboration. We focus on the indicator called ‘top 10% P’, which is the number of a university’s publications belong to the top 10% most frequently cited in the year. CWTS Leiden Ranking 2016 collected the data from the following five main fields of science: Biomedical and health sciences (Bio), Life and earth sciences (Life), Mathematics and computer science (M&C), Physical sciences and engineering (P&E), and Social sciences and humanities (Soc). The default way is to use the sum of the top 10% cited publications in all the five fields as the indicator of the university, i.e., the sum-up method. We use our method to obtain a result. The first 15 universities under our method are listed in Table 1.

In this table, the column ‘ours’ is the results of our method, comparing the column ‘total’ of the results of the sum-up method. In these two columns, the number in the bracket is the final score obtained by the corresponding method and the number out of the bracket is the position ranked by using these scores. The next five columns denote the numbers of top 10% cited publications of these universities on the five main fields.

From the table we can see that some universities with a smaller number of total top 10% cited publications are ranked

¹In fact, the scores were divided by $\sqrt{n} = \sqrt{6}$ to make them of the same stretching factor as the average method.

Table 1: The top 15 universities of our method.

Univ	Ours	Total	Bio	Life	M&C	P&E	Soc
Harvard Univ	1(4949)	1 (7060)	4947	467	136	834	676
Univ Toronto	2(1926)	3 (2956)	1853	201	150	462	289
Johns Hop Univ	3(1917)	7 (2539)	1976	108	51	261	144
Stanford Univ	4(1744)	2 (3223)	1429	300	215	936	342
UC-SF	5(1641)	17(1990)	1775	75	10	46	85
Univ Michigan	6(1602)	4 (2756)	1417	188	169	627	355
Univ Penn	7(1493)	11(2253)	1456	106	80	309	301
UC-Los Angeles	8(1391)	8 (2424)	1231	225	157	516	296
Univ Coll Lon.	9(1381)	13(2162)	1328	118	91	296	329
UW-Seattle	10(1374)	10(2321)	1267	399	89	340	226
Columbia Univ	11(1290)	14(2129)	1217	181	124	286	321
Univ Oxford	12(1286)	9 (2414)	1084	306	130	527	367
UC-SD	13(1270)	15(2126)	1164	253	153	390	166
Yale Univ	14(1199)	16(2001)	1117	239	60	274	312
MIT	15(1184)	6 (2606)	748	254	246	1164	194

higher in our method, for examples, Johns Hopkins Univ. and UC-San Francisco. These universities have a large number of top 10% cited publications in some fields say Bio. and a small top cited publication number in other fields. Our method rank these universities higher because the major fields of these universities are strong, although they may not be strong in every field. We think this is reasonable and even can be regarded as an advantage of our method. We should encourage universities (agents) to develop some strong fields, instead of all fields with a middle level.

We also note that MIT is not ranked very high in our method, although it has the largest numbers of top 10% cited publications in two fields M&C and P&E. The reason is that the total numbers of top 10% cited publications in these two fields are not large. If we normalize all fields to reduce the differences among them, the situation will be better. Now we apply our method directly without any normalization, which indicates that we assume all fields are of the same commensurability. Under this assumption, small total numbers of top 10% cited publications in the two fields M&C and P&E will indicate small influences of these two fields. It is reasonable.

8 Conclusion

This paper suggests a new method for score aggregation, which uses spectral analysis. Our method assumes that all the agents are of the same commensurability and finds an ‘optimal’ aggregate scoring minimizing the ‘noise’ part. The new method may open up a new dimension along which scoring rules could be defined and evaluated. It also gives rise to many interesting problems in social choice for further study.

For the method itself, there are also some important issues worthy of further study. For example, this paper only considers projecting scores on vectors which begin with the origin. Could we get better results if we consider other vectors as well? It is possible. However, we may get strange results for some extreme cases. Here is an example of two candidates with scores (1, 3) and (3, 1) from two agents, the global optimal vector of which will pass through these two points. For this case, it is hard to explain the meaning of the projection. It will also be interesting to study more properties and types of vectors to be projected on.