

Measurement Methods in the field of benchmarking

Istvan Szuts, Dr.,

Budapest Polytechnic Institute for Entrepreneurship Management
6 Doberdo street, Budapest, 1034
szuts@bmf.hu

Abstract: In benchmarking¹ we often come across with parameters being difficult to measure while executing comparisons or analyzing performance, yet they have to be compared and measured so as to be able to choose the best practices. The situation is similar in the case of complex, multidimensional evaluation as well, when the relative importance and order of different dimensions, parameters to be evaluated have to be determined or when the range of similar performance indicators have to be decreased with regard to simpler comparisons. In such cases we can use the ordinal or interval scales of measurement elaborated by S.S. Stevens.

1. Ordinal scale

In case of ordinal scale the entities can be compared by the desired features, by means of it we can achieve relations like for example: better than ..., more useful than ..., bigger than ..., or their opposites: worse, less useful, smaller, etc. When measuring with ordinal scale entities must be comparable and transitive by one common criterion.

To indicate comparison the so called preference relation is used. Preference relation shows the rank order of the entities. A prefers to B ($A \succ B$) means that A can be regarded as better with regard to the actual criterion. For indication the natural numbers are used in an increasing or decreasing order. By means of it the above transitive criterion can be formulated as follows.

If $A \succ B$ and $B \succ C$, entity A is better than B and B is better than C, then A is also better than C.

¹ Competitive benchmarking involves analyzing the performance and practices of best-in-class companies. Their performance becomes a benchmark to which a firm can compare its own performance and their practices are used to improve that firm's practices.

Assignment of rank numbers can be carried out in an increasing sequence in the following way. We allocate 3 to A, 2 to B, 1 to C.

A→3
B→2
C→1

Only the ranking is important while allocating. We can not state that entity A is twice as good as B or C is three times better. The allocation of numbers is totally arbitrary with the exception of sticking to the sequence, it is just a matter of decision that the first three natural numbers have been used. Differences and ratios between the values of the scale tell nothing about real differences and proportions, they only establish the sequence. The scale will remain untouched by any sequence preserving transformation.

Measuring by ordinal scale can be carried out in two ways in accordance with practical requirements. In the example presented so far no criterion has been permitted, the preference between the two entities had to be decided. In case of such rules we deal with a so called strong preference. In case we permit criterion between the entities, the sequence can be regarded as a so called weak preference. A sequence can significantly limit the applicable evaluating statistic methods. On the ordinal scale natural numbers are usually used. The entities on the scale are not at identical intervals, they are not of the same magnitude. In this case only those operations can be accomplished which do not presume the identical magnitude of intervals. With regard to ordinal measuring we have to speak about the problems of multidimensional comparisons as it is important from the point of view of performance evaluation. As mentioned, comparisons can be accomplished only if entities have at least one common criterion. If comparison is executed only by one well defined criterion we have to do with a one dimensional comparison. However, we are well aware that a product, a process or a company may have theoretically infinite number of criterians. With regard to benchmarking [5] analysis not all the criterians can be taken into consideration, so we choose only some important ones. The selected set of criterians serve also as a base for the evaluation of products, processes and companies, so the criterians can be termed as evaluation factors as well. If comparison is accomplished with several criterians and evaluation factors we are faced with a multidimensional comparison. In this case specific problems arise that are not easy to solve. Some evaluation factors can be measured by ordinal, others by interval scales. The problem is how to compare the different dimensions.

In product comparison investigations in the so called “criterion range” provide solutions in case of ratio scales, but their practical applications are significantly limited by extremely difficult mathematical operations. Some criterians to be evaluated can not be measured on ratio scales. Besides solving the dimensional and measuring problems of comparisons, priorities between evaluation factors

have also to be established. For this we have to be familiar with the importance of evaluation factors with regard to each other.

2. Interval scale

The interval scale possesses the characteristics of the ordinal scale and the range between two numbers is known and has a definite value. This scale can be regarded as a measuring scale in its traditional sense. The differences in a numerical sense show equal differences in reality as well. Measuring units, the 0 point can be determined arbitrarily. So linear transformation ($x' = ax + b$) can be permitted. Adding up proportions and quantities make no sense as both change according to the position of 0 point. But if the differences of the entities are considered, the entities have additive properties and as such they are suitable for comparisons. In this case the “b” constant of the relation “ax+b” is eliminated and as a consequence the 0 point is also eliminated so the different entities can be regarded as an absolute quantity.

So the proportion of any two intervals is independent from measurement units and 0 point on the interval scale. Several criterians can be measured on the interval scale in benchmarking, but this is much more difficult than on the ordinal scale. The elaboration of a proper interval scale for measuring a criterion of a phenomenon often means a complicated scale designing technique. Data gained from ordinal scales can be transformed into interval scales by using specific measuring methods.

The axioms of ecriterion (1., 2., 3.) and the axioms of rank ordering are valid (4., 5.,) on ordinal scales. Newer axioms can not be provided for the interval scales in the above described axiom system, but it must be noted that the axioms of additivity (6., 7., 8., 9.,) are valid for the differences of values on the scales. [1]

- | | | | |
|----|---|---|------------------|
| 1. | Or $a=b$ or $a \neq b$ | } | |
| 2. | If $a = b$, then $b = a$ | } | ecriterion |
| 3. | If $a = b$ and $b = c$, then $a = c$ | } | |
| 4. | If $a > b$, then $b \triangleright a$ | } | |
| 5. | If $a > b$ and $b > a$, then $a > c$ | } | (ordinal axioms) |
| 6. | If $a = p$ and $b > 0$, then $a + b > p$ | } | additivity |
| 7. | $a + b = b + a$ | | |
| 8. | If $a = p$ and $b = q$, then $a + b = p + q$ | | |
| 9. | $(a + b) + c = a + (b + c)$ | | |

3. Pairwise Comparison

The comparative method of pairs means comparing the criterians of the entities in pairs and establishing the preference between two things (or dimensions of multidimensional comparison). The comparison, as stated earlier, may be one or multidimensional. Of course usually not two but several entities are compared and the limits of our ability to process the information may cause problems as the number of comparisons increases.

If we meet the requirements of transitivity (see ordinal scale) a consistent opinion or priority order is formulated. In the opposite case inconsistent result is obtained that can be described as follows: If $A \succ B$ and $B \succ C$, that $C \succ A$.

An inconsistent decision impairs or makes impossible the transformation from the ordinal scale into the interval scale.

As a great number of pairs must be compared in the course of analyzing performance in benchmarking, the control of consistency, the determination of non parametric factors for evaluation is an essential condition. Indicators for consistency can be formulated by establishing the inconsistent decision:

$$C = 1 - \frac{24d}{n^2 - n}, \quad \text{if } n \text{ is odd number}$$

$$C = 1 - \frac{24d}{n^3 - 4n}, \quad \text{if } n \text{ is even number;}$$

Where:

n = the number of compared entities

d = the number of inconsistent triads.

The inconsistent triads is shown by Figure.



Figure 1.: Inconsistent triads

Results obtained by pairwise comparisons are usually demonstrated in a preference matrix, where lines and bars mean all compared entities. Such a preference matrix is shown in Table 1. with regard to seven compared entities with large evaluation factor. Number 1 in row A and bar B means that is entity. A was found preferable with regard to entity B ($A \succ B$). Table 1. contains the result

of comparing in pairs with the final result of the following:

$$A \succ B \succ D \succ E \succ F \succ G \succ C$$

pairwise comparison provide possibility only for creating ordinal scales. The preference numbers obtained on the basis of preference matrix containing the results of the comparisons supply only ordinal scales without determining the magnitudes of intervals on the scales.

	A	B	C	D	E	F	G	Preference numbers
A	-	1	1	1	1	1	1	6
B	0	-	1	1	1	1	1	5
C	0	0	-	0	0	0	0	0
D	0	0	1	-	1	1	1	4
E	0	0	1	0	-	1	1	3
F	0	0	1	0	0	-	1	2
G	0	0	1	0	0	0	-	1

Table 1.: contains the result of comparing in pairs with the final result

If ordinal scales are obtained by putting them in priority order end not by comparing in pairs, then numbers can be allocated directly to the entities in an increasing or decreasing order. Of course in this way no information is given about the differences of scale values in reality.

In case of consistent decisions the preference matrix contains consistent partial square matrixes at all diagonals. For example: if 3,4,5, etc. square matrixes are chosen along the diagonal of the matrix in case of a ten-element matrix, the partial matrixes will be consistent. Let us see that in the case of a consistent matrix how many relations are needed between two elements to get a matrix. For example in a matrix where $N=6$, we have fifteen possible pairwise comparisons. This is the maximal number of comparisons but supposing the matrix is being consistent 5 pairwise comparisons will determine the remaining 10 preferences and the matrix can be filled in. In general it can be stated that a consistent matrix containing N elements can be determined by $N-1$ definite, not arbitrary comparisons.

An important problem of pairwise comparisons from the point of view of our investigation can be regarded as quantitative. If the number (N) of element to be compared exceed 8-10 the following problems may arise. The consistency of the decision makers will impair, matrixes will contain inconsistent decisions leading to ecriterians on interval scales and it will decrease the efficiency of the analyses. pairwise comparisons to the performed will drastically increase with the proliferation of entities.

While in case of $N=4$ the number of pairwise comparisons is 6, then in case of 10 elements it is 45, in case of 20 elements is as many as 190 [2]. The priority statistic and graphic methods provide possibility to determine the priority order of evaluation factors the basis of limited comparisons or to choose the ones belonging to the same set of criterians, this way the multidimensional evaluation can be simplified. This method can be very useful when establishing consumer preferences or when analyzing performance with benchmarking.

Let us consider the following example of purchasing a car, it is easily understandable for everyone, namely what comparisons occur most frequently.

Among the criterians mentioned we can find both expert and consumer preferences. By means of pairwise comparisons the following preference priority order of the pairs were obtained.

Preferences are marked by: \succ

Final Speed \succ Comfort	Function \succ Appearance
Braking Distance \succ Acceleration	Acceleration \succ Comfort
Design \succ Image	Safety \succ Consumption
Consumption \succ Final Speed	Comfort \succ Appearance
Final speed \succ Function	Slowing Down \succ Final Speed
Comfort \succ Image	Appearance \succ Image
Function \succ Design	Safety \succ Acceleration
Consumption \succ Braking Distance	Maximal performance \succ Function
Safety \succ Slowing Down	Slowing Down \succ Max. performance
Comfort \succ Design	

In graphic theory the problem can be formulated as follows. Let us consider criterians, features as a set where certain criterians end features are more important than others. We would like to set up a priority order so as to meet all preferences.

We would like to note that in case of 12 factors we ought to perform 66 pairwise comparisons.

T1 \succ T5,	T1 \succ T8,	T2 \succ T9,	T2 \succ T10,	T2 \succ T11
T4 \succ T11,	T5 \succ T3,	T5 \succ T7,	T5 \succ T12,	T6 \succ T8
T7 \succ T3,	T8 \succ T7,	T8 \succ T12,	T9 \succ T1,	T9 \succ T6
T10 \succ T1,	T10 \succ T4,	T11 \succ T5,	T12 \succ T3	

These 12 criterians and their 19 relevant requirements can be represented by a graph where peaks substitute criterians and (TI,TJ) curve is shown in the graph if $T_i \prec T_j$ [3].

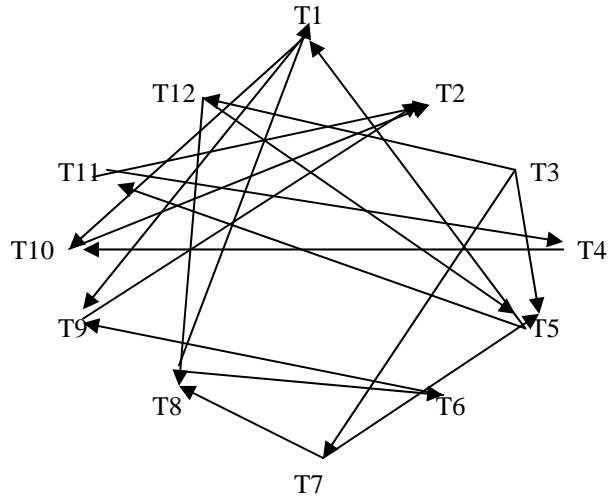


Figure 2.: 12 criterians and their 19 relevant requirements

It must be noted that this graph can not contain a cycle as in that case one event would precede itself and this can not happen because of the nature of the investigated matter.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12		V0	V1	V2	V3	V4	V5	V6
T1					1			1						2	2	2	0	x	x	x
T2									1	1	1			3	3	3	3	2	1	0
T3														0	x	x	x	x	x	x
T4											1			1	1	1	1	0	x	x
T5			1				1					1		3	2	0	x	x	x	x
T6								1						1	1	1	0	x	x	x
T7			1											1	0	x	x	x	x	x
T8							1					1		2	2	0	x	x	x	x
T9	1					1								2	2	2	2	0	x	x
T10	1			1										2	2	2	2	1	0	x
T11					1									1	1	1	0	x	x	x
T12			1											1	0	x	x	x	x	x
														T3	T7 T12	T5 T8	T1 T6 T11	T4 T9	T10	T2
														0	1	2	3	4	5	6

Table 2.: Matrix

Let us take now the matrix representing the earlier graph. This matrix has been enlarged with a certain number of bars, will speak about it later. For better transparency deliberately no 0-s were written in the matrix where according to the definition we ought to have written them. Let us mark the column vectors of the matrix by $V_{T1}, V_{T2}, \dots, V_{T12}$.

First we can calculate vector

$$V_0 = V_{T1} + V_{T2} + \dots + V_{T12}$$

and the result is written in bar V_0 . In this vector 0 can be found in the place corresponding to row T_3 indicating that this peak (i.e. criterion) is not followed by any other.

So we can state that the level of T_3 is 0.

Now let us calculate vector

$$V_1 = V_0 - V_{T3}$$

and let's write x in the T_3 row of vector V_1 . We find a new 0 in rows of T_7 and T_{12} of bar V_1 ; so if T_3 is omitted T_7 and T_{12} are not followed by any peaks. Therefore, we state that the level of T_7 and T_{12} is one. Now we calculate vector

$$V_2 = V_1 - V_{T3} - V_{T12}$$

and we write x in vector V_2 in all places where 0 was to be found in the earlier vector, etc.

At least the 12 peaks were divided into 7 levels: $N_0, N_1, N_2, \dots, N_6$. These levels define what we call: graph priority function free of circular triads [4] [6].

The next figure (Figure 3.) shows the representation of this level determination: numbering was started at T_2 , we also could have made it in the reverse direction starting from T_3 . The figure shows the priority order of criterians and characteristics: they are shown not only in relation to each other but with regard to the whole structure.

It is obvious that operation T_3 is the last, it is preceded by T_7 and T_{12} (their priority between each other is not important), then they are preceded by T_5 and T_8 (again priority can be neglected) and so on.

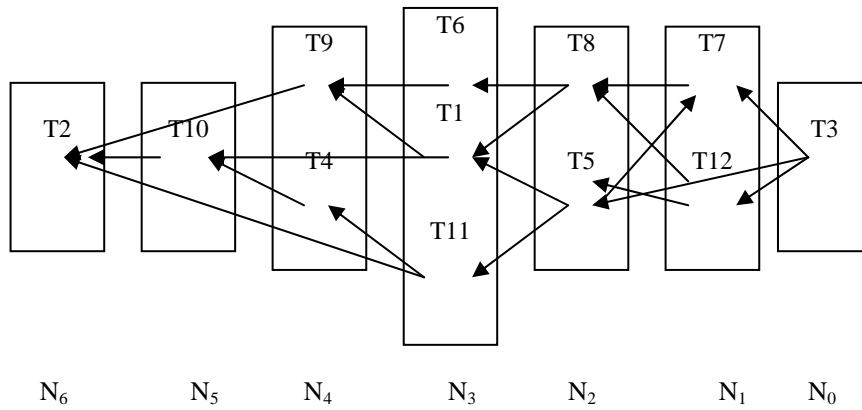


Figure 3.: Representation of the whole structure

Let us be aware that an other priority function would have been obtained if we had worked with the inverse relation “TI follows TJ”. In this case we would have dealt with the row vectors of the matrix and we would have gained a priority function different from the previous one. However, all peaks have an arrangement that is compatible with all priority functions. On the basis of above said we can formulate the preference groups (consumer or benchmarking expert) determining the choice or comparison.

The previously obtained 19 preferences are the following (Figure 4.):

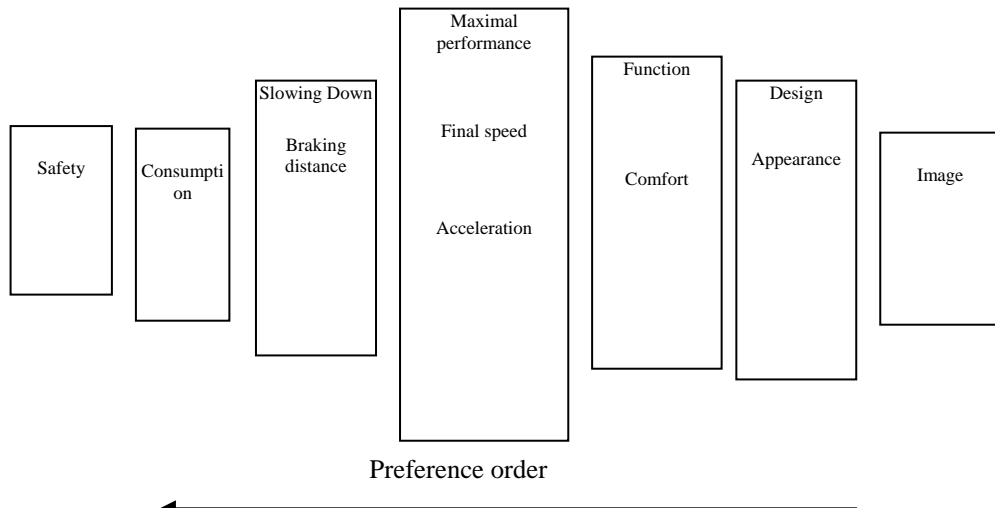


Figure 4.: The previously obtained 19 preferences

References:

- [1.] Stevens, S. S.: Handbook of Experimental Psychology
Wiley, New York 1951.
- [2.] Dr. Szűts István: Methods for comparative analysis of company efficiency
KJK Budapest 1983 (In hungarian)
- [3.] Istvan Szuts dr.: An operation research model for the credit system
In-Tech-Ed '96, Budapest, 1996.
- [4.] Istvan Szuts dr.: Potential algorithms for the Credit System Analysis
KMF Budapest, 1999.
- [5.] Robert C. Camp: Business Process Benchmarking
ASQC Criterion Press, Milwaukee 1996.
- [6.] Dr. Bakó András, Dr. Szűts István: Gráfelméleti algoritmusok az oktatás
korszerűsítésére. KMF Tudományos Közlemények, 1999. 19-29. o.