Normalization Techniques for Multi-Criteria Decision Making: Analytical Hierarchy Process Case Study

Nazanin Vafaei[(✉)], Rita A. Ribeiro, and Luis M. Camarinha-Matos

CTS/UNINOVA and Faculty of Sciences and Technology,
NOVA University of Lisbon, 2829-516 Caparica, Portugal
nv@ca3-uninova.org, {rar,cam}@uninova.pt

Abstract. Multi-Criteria Decision Making (MCDM) methods use normalization techniques to allow aggregation of criteria with numerical and comparable data. With the advent of Cyber Physical Systems, where big data is collected from heterogeneous sensors and other data sources, finding a suitable normalization technique is also a challenge to enable data fusion (integration). Therefore, data fusion and aggregation of criteria are similar processes of combining values either from criteria or from sensors to obtain a common score. In this study, our aim is to discuss metrics for assessing which are the most appropriate normalization techniques in decision problems, specifically for the Analytical Hierarchy Process (AHP) multi-criteria method. AHP uses a pairwise approach to evaluate the alternatives regarding a set of criteria and then fuses (aggregation) the evaluations to determine the final ratings (scores).

Keywords: Normalization · AHP · MCDM · Rank reversal · Cyber-physical systems (CPSs)

1 Introduction

Everybody makes decisions in their daily lives, as for example: "Should I take an umbrella today"? "Where should I go for lunch"? To make decisions we need access to information (or data) and to reach a decision we need to combine the data to obtain a final score for each candidate decision alternative (e.g. combining food prices and service of restaurants to recommend). The aim of Multi-Criteria Decision Making (MCDM) methods is to rate and prioritize a set of alternatives that best satisfy a given set of criteria [1]. Criteria are a set of requirements or independent attributes that have to be satisfied by several alternatives. Each criterion may be measured in different units, for example, degrees, kilograms or meters; but they all have to be normalized to obtain dimensionless classifications, i.e. a common numeric range/scale, to allow aggregation into a final score. Hence, data normalization is an essential part of any decision making process because it transforms the input data into numerical and comparable data, allowing using MCDM methods to rate and rank alternatives [2, 3].

In this work, the main research question that we address is: *Which normalization technique is more suitable for usage with the AHP method?*

The motivation for carrying out this work includes four interconnected issues: (a) the importance of data normalization for decision problems where we need to fuse or aggregate data to obtain a final score per alternative; (b) the reduced number of research studies available in this topic; (c) continuation of previous work on suitability of normalization techniques for well-known MCDM methods (e.g. TOPSIS) [4]; (d) contributing to advances in Cyber Physical Systems [5] research, where huge amounts of available data from heterogeneous sensors need to be fused (aggregated) to determine a combined view. Specifically, in this study we focus on the well-known AHP method because it is a well-known and widely used MCDM method [6–13] but we plan to perform the same study for data fusion problems as well as other MCDM methods in the future, to determine which technique is more suitable for any decision problem that requires combining (fusing) data.

The Analytic Hierarchy Process (AHP) was introduced by Saaty [6, 7] to solve unstructured problems in economics, social sciences, and management [8]. AHP has been used in a vast range of problems from simple ones (e.g. selecting a school) to harder ones (e.g. in allocating budgets and energy domains) [8]. When applying the AHP method, the decision maker is able to structure the decision problem and break it down into a hierarchical top-down process. Then, he/she performs a pairwise matrix comparison of criteria using a [1–9] scale (corresponding to semantic interpretations such has "A is much more important than B" regarding a criterion). After normalization, the priorities are determined using either Eigen vectors or a simplified version with weighted sum (SAW) [9, 10].

AHP involves five main steps [13]: Step 1: Decompose the problem into a hierarchical structure; Step 2: Employ pairwise comparisons. A pairwise comparison is the process of comparing the relative importance, preference, or likelihood of two elements (objectives) with respect to another element (the goal). Pairwise comparisons are carried out to establish priorities. Decision elements at each hierarchy level are compared pairwisely and then the reciprocal matrix is completed; Step 3: Determine the logical consistency and if >10 % revise the pairwise classifications until the consistency index is below 10 %. In the implementation of AHP, we may face with inconsistent judgment of input data that it may cause some bad effects on decision process. For example, A1 may be preferred to A2 and A2 to A3, but A3 may be preferred to A1. So, Saaty [7] defined a measure of deviation from consistency that is called a consistency index, as: $C.I. = (\lambda max - N)/(N - 1)$, where N is the dimension of the matrix and λ is the largest eigenvalue of the matrix A. Then, Saaty calculated a consistency ratio (C.R.) as the ratio of the C.I. to a random index (R.I.) which is the average C.I. of sets of judgments (from a 1 to 9 scale) for randomly generated reciprocal matrices [7]. Step 4: Estimate the relative weights by combining the individual subjective judgments. We can use the eigenvalue method to estimate the relative weights of the decision elements. In order to estimate the relative weight of the decision elements in a matrix, we can use $A.W = \lambda max.W$ where W is the weight of criterion [13]. Step 5: Determine the priority of alternatives by doing aggregation on relative weights which is obtained by combining the criterion priorities and priorities of each decision alternatives relative to each criterion. Since in our work we discuss the suitability of normalization techniques for the AHP method, we focus on step 4 and 5.

In this work we propose an assessment approach for evaluating five common normalization techniques (see Table 1), using an illustrative example solved with AHP method [1, 2]. We choose AHP because it is a well-known and widely used MCDM method [6–13] but we plan to perform the same study for other MCDM methods in the future. Our novel assessment approach calculating Pearson correlation for global weight of alternatives and Spearman correlation for rank of alternatives which are borrowed from [14] to determine mean values in order to ensure a more robust evaluation and selection of the best normalization technique in AHP. The novelty of this study is making adaptation between assessment process and AHP in order to find best normalization technique for AHP method. The next section presents the experimental study performed.

Table 1. Normalization techniques.

Normalization technique	Condition of use	Formula
Linear: Max (N1) [14]	Benefit criteria	$n_{ij} = \frac{r_{ij}}{r_{max}}$
	Cost criteria	$n_{ij} = 1 - \frac{r_{ij}}{r_{max}}$
Linear: Max-Min (N2) [14]	Benefit criteria	$n_{ij} = \frac{r_{ij} - r_{min}}{r_{max} - r_{min}}$
	Cost criteria	$n_{ij} = \frac{r_{max} - r_{ij}}{r_{max} - r_{min}}$
Linear: sum (N3) [14]	Benefit criteria	$n_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}$
	Cost criteria	$n_{ij} = \frac{1/r_{ij}}{\sum_{i=1}^{m} 1/r_{ij}}$
Vector normalization (N4) [2]	Benefit criteria	$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^{m} r_{ij}^2}}$
	Cost criteria	$n_{ij} = 1 - \frac{r_{ij}}{\sqrt{\sum_{i=1}^{m} r_{ij}^2}}$
Logarithmic normalization (N5) [2]	Benefit criteria	$n_{ij} = \frac{\ln(r_{ij})}{\ln(\prod_{i=1}^{m} r_{ij})}$
	Cost criteria	$n_{ij} = \frac{1 - \frac{\ln(r_{ij})}{\ln(\prod_{i=1}^{m} r_{ij})}}{m-1}$

2 Relationship to Cyber-Physical Systems

Cyber-physical systems (CPS) involve merging computation and physical processes, often denoted as embedded systems [15]. In most CPS, the physical inputs and outputs are typically designed as a network of interacting elements. This conceptual model is tied to the notion of robotics and sensor networks and their usage has been increasing day by day [16]. But CPS also inherits ideas from the areas of embedded and real-time systems. CPS have a broad scope of potential application in areas such as reactive interventions (e.g., collision avoidance); precision operations (e.g., robotic surgery and nano-level manufacturing); operation in dangerous or inaccessible environments (e.g., search and rescue, firefighting, and deep-sea exploration); complex systems coordination (e.g., air traffic control, war fighting); efficiency (e.g., zero-net energy buildings);

and augmentation of human capabilities (e.g., healthcare monitoring and service delivery) [16], to name a few.

There are some discussions on the relationship between Cyber-Physical Systems and Internet of Things [17–19]. Camarinha and Afsarmanesh [5] mention that "there is a growing convergence between the two areas since CPSs are becoming more Internet-based". For example, in smart car parking, data from the parking space is transferred to the car drivers with the help of CPS and IoT technologies. Data is collected from sensors, which are installed in the parking lot, and transferred to the data center to be processed with MCDM methods, to determine the ranking of alternatives (best parking spaces). The best parking spaces are provided to the car drivers to support them making more informed decisions. In the illustrative example section, we will compare several normalization techniques for usage with the AHP method to rank alternatives and support car drivers. The smart car parking example shows a robust relationship between cyber physical system (CPS), Internet of Thing (IoT) and multi-criteria decision making (MCDM) concepts.

3   Normalization

There are several definitions for data normalization, depending on the study domain. For example, in Databases, data normalization is viewed as a process where data attributes, within a data model, are organized in tables to increase the cohesion and efficiency of managing data. In statistics and its applications, the most common definition is the process of adjusting values measured on different scales to a common scale, often prior to aggregating or averaging them [19]. Many other definitions exist, depending on the context or study domain (see for example [20]). Here we focus on normalization techniques for MCDM. In general, normalization in MCDM is a transformation process to obtain numerical and comparable input data by using a common scale [4]. After collecting input data, we must do some pre-processing to ensure comparability of criteria, thus making it useful for decision modeling. Furthermore, in MCDM, normalization techniques usually map attributes (criteria) with different measurement units to a common scale in the interval [0-1] [21, 22]. Several studies on the effects of normalization techniques on the ranking of alternatives in MCDM problems have shown that certain techniques are more suitable for specific decision methods than others [14, 23–28].

Chakraborty and Yeh [23] analyzed four normalization techniques (vector, linear max-min, linear max and linear sum) in the MCDM simple additive weight (SAW) method. They used a ranking consistency index (RCI) and calculated the average deviation for each normalization technique and concluded that the best normalization technique for SAW is the vector normalization. Further, the same authors analyzed the effects of those normalizations for order preference by similarity to ideal solution method (TOPSIS) by calculating ranking consistency and weight sensitivity of each normalization and proved that vector normalization technique is the best for implementing in TOPSIS method [24]. The authors [24] defined weight sensitivity as a method to analyze sensitivity level of different normalization procedures under different

problem settings. They assumed same weights for attributes and then they increased their weights to find the sensitivity of the alternatives (normalization techniques) [24].

Also, the result was further validated by Vafaei et al. [4], who used Pearson and Spearman correlation coefficients to also conclude that the best normalization technique for TOPSIS method is the vector normalization.

In this work, we selected five (shown in Table 1) of the most promising normalization techniques [2, 14] and analyzed their effect on the AHP method. In Table 1, each normalization method is divided in two formulas, one for benefit and another for cost criteria, to ensure that the final decision objective (rating) is logically correct, i.e. when it is a benefit criterion for high values it will correspond to high normalized values (maximization - benefit) and when it is a cost criterion high values will correspond to low normalized values (minimization - cost).

Summarizing, the aim of this study is to identify which normalization technique is best suited for the AHP method.

4 Experimental Study with a Smart Car Parking Example

Here, we discuss the suitability of five normalization techniques for AHP with an illustrative example for smart car parking. This illustrative case consists of 3 criteria (C1, C2, C3), which correspond to time to park, distance, and size of the parking space, and 7 alternatives (A1, A2, …, A7), which correspond to candidate location sites for parking. Finding the best place for parking the car is the goal; C1 and C2 are cost criteria, where low values are better, and C3 is a benefit criterion, where high values are desirable. Following the AHP method we defined three pairwise comparison matrices for each criterion (example in Table 2) and then one pairwise comparison matrix between criteria. To these four matrices we applied the five normalization techniques, separately, to determine the ranking of alternatives and compare results. The pairwise decision matrix for criteria "time to park", after steps 1, 2, 3 of AHP, is shown in Table 2.

Table 2. Pairwise comparison matrix with respect to the time.

	A1	A2	A3	A4	A5	A6	A7
A1	1	1/3	1/2	3	1/3	2	1
A2	3	1	1	4	1	3	1
A3	2	1	1	2	1/2	3	2
A4	1/3	1/4	1/2	1	1/4	1	1/3
A5	3	1	2	4	1	3	1
A6	1/2	1/3	1/3	1	1/3	1	3
A7	1	1	1/2	3	1	1/3	1

We started by testing the sum-based normalization (N3 in Table 1), the usual normalization technique for AHP [7], because it ensures column sum per alternative is equal to one that is defined by Saaty [7]. The other normalization techniques do not

include this characteristic and the sum of the normalized values can be bigger than 1; hence, for comparison purposes we opted for re-normalizing the other four using N3. For illustrating the alternatives rating procedure we show the calculation for vector normalization of alternative A1 and the final results for all alternatives are shown in the Tables 3 and 4:

$$P_{11} = \frac{x_{11}}{\sqrt{\sum_{j=1}^{7} x_{1j}}} = \frac{1}{\sqrt{(1^2) + (3^2) + (2^2) + \left(\frac{1^2}{3}\right) + (3^2) + \left(\frac{1^2}{2}\right) + (1)^2}} = 0.7974$$

$$AverageP1 = \frac{0.7974 + 0.8390 + 0.8091 + 0.5991 + 0.8227 + 0.6524 + 0.7583}{7} = 0.7540$$

$$A_{11} = \frac{AverageP1}{Sum} = \frac{0.7974}{4.8050} = 0.1659$$

$$AverageA1 = \frac{0.1659 + 0.1814 + 0.1659 + 0.1304 + 0.1769 + 0.1393 + 0.1598}{7} = 0.1605$$

Table 3. Normalization results for vector normalization technique for cost criteria.

	P1	P2	P3	P4	P5	P6	P7	Average
P1	0.7974	0.8390	0.8091	0.5991	0.8227	0.6524	0.7583	0.7540
P2	0.3922	0.5169	0.6182	0.4655	0.4681	0.4786	0.7583	0.5283
P3	0.5948	0.5169	0.6182	0.7327	0.7341	0.4786	0.5165	0.5988
P4	0.9325	0.8792	0.8091	0.8664	0.8670	0.8262	0.9194	0.8714
P5	0.3922	0.5169	0.2365	0.4655	0.4681	0.4786	0.7583	0.4737
P6	0.8987	0.8390	0.8727	0.8664	0.8227	0.8262	0.2748	0.7715
P7	0.7974	0.5169	0.8091	0.5991	0.4681	0.9421	0.7583	0.6987
Sum	4.8051	4.6247	4.7730	4.5946	4.6508	4.6829	4.7437	4.6964

Table 4. Re-normalization results for vector normalization technique for cost criteria.

	A1	A2	A3	A4	A5	A6	A7	Average
A1	0.1659	0.1814	0.1695	0.1304	0.1769	0.1393	0.1598	0.1605
A2	0.0816	0.1118	0.1295	0.1013	0.1007	0.1022	0.1598	0.1124
A3	0.1238	0.1118	0.1295	0.1595	0.1578	0.1022	0.1089	0.1276
A4	0.1941	0.1901	0.1695	0.1886	0.1864	0.1764	0.1938	0.1856
A5	0.0816	0.1118	0.0495	0.1013	0.1007	0.1022	0.1598	0.1010
A6	0.1870	0.1814	0.1828	0.1886	0.1769	0.1764	0.0579	0.1644
A7	0.1659	0.1118	0.1695	0.1304	0.1007	0.2012	0.1598	0.1485
Sum	1	1	1	1	1	1	1	1

The global weights of alternatives and ranking results for the four tested normalization techniques are shown in Table 5. We discarded the logarithmic normalization

technique from our results because we obtained negative and infinite data (due to the characteristics of pairwise matrices), hence it is not usable (appropriate) for the AHP method. As it can be seen in Table 5, there is consensus on which normalization techniques is better for alternatives A2, A3, A4 and A5 (i.e. they all have the same ranking), but for the other alternatives there was no consensus. Since, it is not possible to distinguish which is the best normalization technique just by looking at the results, we used the evaluation approach proposed in [4] to make the assessment. Hence, we calculated Pearson correlation and mean r_s values [4] with the global weights of alternatives and Spearman correlation with the ranks of alternatives to assess the suitability of the four tested normalization techniques for the AHP method. Table 6 displays that there exists complete consensus between Pearson and Spearman correlation's results and it is clear that the best normalization technique is N1 (linear: max) because it has the highest mean r_s value (P = 0.9606 & S = 0.9524) and the worst one is N3 (linear sum) with the lowest mean r_s value (P = 0.9029 & S = 0.8413).

Table 5. Global weight (G) and Ranking (R) of alternatives for the smart parking example.

	N1		N2		N3		N4	
	G	R	G	R	G	R	G	R
A1	0.1972	2	0.1925	2	0.1505	4	0.1693	2
A2	0.0681	6	0.0634	6	0.0762	6	0.1165	6
A3	0.1143	5	0.1161	5	0.0993	5	0.1297	5
A4	0.2469	1	0.2658	1	0.2876	1	0.1755	1
A5	0.0460	7	0.0291	7	0.0749	7	0.1101	7
A6	0.1765	3	0.1869	3	0.1598	2	0.1450	4
A7	0.1509	4	0.1462	4	0.1517	3	0.1538	3

Table 6. Pearson correlation between global weights and Spearman correlation between ranks of alternatives for each normalization technique.

	N1		N2		N3		N4		Mean r_s		Rank	
	P	S	P	S	P	S	p	S	P	S	P	S
N1			0.9961	1	0.9171	0.8571	0.9687	1	0.9606	0.9524	1	1
N2	0.9961	1			0.9273	0.8571	0.9458	0.9524	0.9564	0.9365	2	2
N3	0.9171	0.8571	0.9273	0.8571			0.8643	0.8095	0.9029	0.8413	4	4
N4	0.9687	1	0.9458	0.9524	0.8643	0.8095			0.9263	0.9206	3	3

* P = Pearson
**S = Spearman

From the example we can conclude that linear max (N1) is the best normalization technique for the AHP method and linear sum (N3) is the worst one. It is interesting to note that the single normalization used in AHP (linear sum- N3) is the worst one from this comparison study. Although N1 is elected as the most suitable normalization

technique it required a re-normalization with N3 because the sum of the normalized values has to be 1. Therefore, we may conclude that a combination of max-normalization (N1) with linear-sum (N3) seems the most appropriate for AHP.

5   Conclusion

Normalization is the first step of any decision making process to transform data in different units into a common scale and comparable units. In this study we tested five common normalization techniques to assess their suitability for the AHP MCDM method. The tests showed that the logarithmic normalization technique (N5) is not usable in the AHP method because it can result in zero or infinite values in the normalized data, which is not acceptable to use in the method. Further, since AHP requires the columns of the pairwise matrices to sum up 1, the techniques: linear max, linear max-min and vector normalization techniques had to be re-normalized with linear sum (N3) before being compared. To assess the suitability of the normalization techniques for AHP we used Pearson and Spearman correlation and mean r_s values; the results showed that the best normalization technique is N1 (linear: max) combined with N3 (linear-sum) to ensure the sum is 1, while the worst one is N3 alone.

In a previous work we did the same assessment study for TOPSIS and in the future we plan to extend it to other well-known MCDM methods, with the aim to support decision makers by recommending the most suitable normalization techniques for usage with each MCDM method.

References

1. Triantaphyllou, E.: Multi-criteria decision making methods. In: Multi-criteria Decision Making Methods: A Comparative Study, vol. 44, pp. 5–21. Springer, New York (2000)
2. Jahan, A., Edwards, K.L.: A state-of-the-art survey on the influence of normalization techniques in ranking: improving the materials selection process in engineering design. Mater. Des. **65**(2015), 335–342 (2014)
3. Nayak, S.C., Misra, B.B., Behera, H.S.: Impact of data normalization on stock index forecasting. Int. J. Comput. Inf. Syst. Ind. Manage. Appl. **6**(2014), 257–269 (2014)
4. Vafaei, N., Ribeiro, R.A., Camarinha-Matos, L.M.: Data normalization techniques in decision making: case study with TOPSIS method. Int. J. Inf. Decis. Sci. (2016, in press)
5. Camarinha-Matos, L.M., Afsarmanesh, H.: Collaborative systems for smart environments: trends and challenges. Collab. Syst. Smart Networked Environ. **434**(2014), 3–14 (2014)
6. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**(3), 234–281 (1977)
7. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
8. Cheng, C.-H., Yang, K.-L., Hwang, C.-L.: Evaluating attack helicopters by AHP based on linguistic variable weight. Eur. J. Oper. Res. **116**(1999), 423–435 (1999)

9. Gaudenzi, B., Borghesi, A.: Managing risks in the supply chain using the AHP method. Int. J. Logist. Manage. **17**(1), 114–136 (2006)
10. Zahedi, F.: The analytic hierarchy process – a survey of the method and its applications. Interfaces (Providence) **16**(4), 96–108 (1986)
11. Tuzkaya, G., Onut, S., Tuzkaya, U.R., Gulsun, B.: An analytic network process approach for locating undesirable facilities: an example from Istanbul, Turkey. J. Environ. Manage. **88** (2008), 970–983 (2007)
12. Saaty, T.L., Vargas, L.G.: Models, Methods, Concepts & Applications of the Analytic Hierarchy Process. Institute for Operations Research and the Management Sciences, Maryland (2006)
13. Tzeng, G.-H., Huang, J.-J.: Multiple Attribute Decision Making: Methods and Applications. Taylor & Francis Group, Boca Raton (2011)
14. Celen, A.: Comparative analysis of normalization procedures in TOPSIS method: with an application to Turkish deposit banking market. INFORMATICA **25**(2), 185–208 (2014)
15. Shi, J., Wan, J., Yan, H., Suo, H.: A survey of cyber-physical systems. In: 2011 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6 (2011)
16. Wiki1: Cyber-physical system. https://en.wikipedia.org/wiki/Cyber-physical_system. Accessed 02 November 2015
17. Camarinha-Matos, L.M., Tomic, S., Graça, P. (eds.) Technological Innovation for the Internet of Things, vol. 394. Springer, Heidelberg (2013)
18. Camarinha-Matos, L.M., Tomic, S., Graça, P. (eds.): DoCEIS 2013. IFIP AICT, vol. 394. Springer, Heidelberg (2013)
19. Jeschke, S.: Everything 4.0? In: Drivers and Challenges of Cyber Physical Systems (2013)
20. Wiki2: Normalization (statistics). https://en.wikipedia.org/wiki/Normalization_%28statistics%29. Accessed 15 October 2015
21. Wiki3: Normalization. https://en.wikipedia.org/wiki/Normalization. Accessed 15 October 2015
22. Pavlicic, D.M.: Normalization affects the results of MADM methods. Yugosl. J. Oper. Res. **11**(2011), 251–265 (2011)
23. Etzkorn, B.: Data normalization and standardization. http://www.benetzkorn.com/2011/11/data-normalization-and-standardization/. Accessed 28 April 2015
24. Chakraborty, S., Yeh, C.-H.: A simulation based comparative study of normalization procedures in multiattribute decision making. In: International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, pp. 102–109 (2007)
25. Chakraborty, S., Yeh, C.-H.: A simulation comparison of normalization procedures for TOPSIS. In: Computers Industrial Engineering, pp. 1815–1820 (2009)
26. Chakraborty, S., Yeh, C.-H.: Rank similarity based MADM method selection. In: International Conference on Statistics in Science, Business and Engineering (ICSSBE 2012) (2012)
27. Milani, A.S., Shanian, A., Madoliat, R., Nemes, J.A.: The effect of normalization norms in multiple attribute decision making models: a case study in gear material selection. Struct. Multidiscip. Optim. **29**(4), 312–318 (2004)
28. Wang, Y.-M., Luo, Y.: Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making. Math. Comput. Model. **51**(2010), 1–12 (2010)