Using the Probabilistic Weighted Average in Decision Making with Distance Measures

José M. Merigó, *Member, IAENG*

***Abstract*—We develop a new decision making method based on distance measures that uses the probabilistic weighted averaging (PWA) operator. The PWA operator is an aggregation operator that unifies the weighted average and the probability in the same formulation and considering the degree of importance that each concept has in the aggregation. We introduce the probabilistic weighted averaging distance (PWAD) operator. It is a new aggregation operator that uses probabilities, weighted averages and distance measures. We study some of its main properties and particular cases such as the arithmetic weighted Hamming distance and the arithmetic probabilistic Hamming distance. We also develop an application in a decision making problem concerning the selection of investment strategies.**

***Index Terms*—Probability; Weighted average; Distance Measures; Decision making.**

I. Introduction

In the literature, we find a wide range of methods for decision making [3,7-10,14,17]. A very useful technique for doing so is the Hamming distance [4] and more generally all the distance measures [3-7,11-14]. The main advantage of using distance measures in decision making is that we can compare the alternatives of the problem with some ideal result [3,6]. Thus, by doing this comparison, the alternative with a closest result to the ideal is the optimal choice.

Usually, when using distance measures in decision making, we normalize it by using the arithmetic mean or the weighted average (WA) obtaining the normalized Hamming distance (NHD) and the weighted Hamming distance (WHD), respectively. However, sometimes it would be interesting to consider the possibility of using other types of aggregation operators [1-2]. For example, Merigó and Gil-Lafuente have suggested the use of the OWA operator [7,15-17] obtaining the OWA distance [7,11,14].

Recently, Merigó has suggested a new model that unifies the weighted average with the probability [8]. He called it the probabilistic weighted averaging (PWA) operator. The main advantage of the PWA is that it is able to unify the probability and the weighted average in the same formulation considering the degree of importance that each concept has in the aggregation.

The aim of this paper is to present the probabilistic weighted averaging distance (PWAD) operator. It is a new aggregation operator that uses the WA and the probability in the same formulation and considering the degree of importance that each concept has in the aggregation. Moreover, it also uses distance measures in the aggregation process. Note that in this paper we consider the use of the Hamming distance but it is also possible to consider other distance measures such as the Euclidean and the Minkowski distance. The main advantage of the PWAD is that it is able to deal with probabilities and WAs in the Hamming distance. We analyze several families of PWAD operators such as the probabilistic Hamming distance, the weighted Hamming distance, the arithmetic probabilistic Hamming distance and the arithmetic weighted Hamming distance.

We study the applicability of the PWAD operator in a decision making problem concerning the selection of strategies. We see that depending on the particular type of PWAD operator used, the results may lead to different decisions.

This paper is organized as follows. In Section 2, we briefly review some basic concepts about the Hamming distance, the PWA and the OWAD operator. In Section 3 we introduce the PWAD operator and in Section 4 we develop an application in a decision making problem. In Section 5 we present a numerical example. Section 6 summarizes the main conclusions of the paper.

II. Preliminaries

A. The Hamming Distance

The Hamming distance [4] is a very useful technique for calculating the differences between two elements, two sets, etc. For two sets $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$ it can be defined as follows.

Definition 1. A weighted Hamming distance of dimension n is a mapping d_{WH}: $[0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$WHD\ (A, B) = \left(\sum_{i=1}^{n} w_i \mid a_i - b_i \mid \right) \tag{1}$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Note that it is possible to generalize this definition to all the real numbers by using $R^n \times R^n \rightarrow R$. Note also that if $w_j = 1/n$, we get the normalized Hamming distance. For the formulation used in fuzzy set theory, see for example [6].

Manuscript received March 22, 2010.
J.M. Merigó is with the Department of Business Administration, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain (corresponding author: +34-93-4021962; fax: +34-93-4039882; e-mail: jmerigo@ ub.edu).

B. The Probabilistic Weighted Average

The probabilistic weighted averaging (PWA) operator [8] is an aggregation operator that unifies the probability and the weighted average in the same formulation considering the degree of importance that each concept has in the aggregation. It is defined as follows.

Definition 2. A PWA operator of dimension n is a mapping PWA: $R^n \rightarrow R$ such that:

$$PWA\ (a_1, \ldots, a_n) = \sum_{j=1}^{n} \hat{v}_i a_i \tag{2}$$

where the a_i are the argument variables, each argument a_i has an associated weight (WA) v_i with $\sum_{i=1}^{n} v_i = 1$ and $v_i \in [0, 1]$, and a probabilistic weight p_i with $\sum_{i=1}^{n} p_i = 1$ and $p_i \in [0, 1]$, $\hat{v}_i = \beta p_i + (1-\beta) v_i$ with $\beta \in [0, 1]$ and $\hat{v}_i$ is the weight that unifies probabilities and WAs in the same formulation.

C. The OWAD Operator

The OWAD (or Hamming OWAD) operator [7,11,14] is an extension of the traditional normalized Hamming distance by using OWA operators. The main difference is the reordering of the arguments of the individual distances according to their values. Then, it is possible to calculate the distance between two elements, two sets, two fuzzy sets, etc., modifying the results according to the interests of the decision maker. It can be defined as follows.

Definition 3. An OWAD operator of dimension n is a mapping $OWAD$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector W, with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$ such that:

$$OWAD\ (\langle \mu_1, \mu_1^{(k)} \rangle, \ldots, \langle \mu_n, \mu_n^{(k)} \rangle) = \sum_{j=1}^{n} w_j D_j \tag{3}$$

where D_j represents the jth largest of the individual distances $|\mu_i - \mu_i^{(k)}|$, μ_i and $\mu_i^{(k)} \in [0, 1]$, and $k = 1, 2, \ldots, m$.

Note that this definition can be generalized to all the real numbers R by using OWAD: $R^n \times R^n \rightarrow R$. Note also that it is possible to distinguish between ascending and descending orders. The weights of these operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the descending OWAD (DOWAD) operator and w^*_{n-j+1} the jth weight of the ascending OWAD (AOWAD) operator.

III. The Probabilistic Weighted Averaging Distance Operator

The probabilistic weighted averaging distance (PWAD) operator is a distance measure that uses the WA and the probability in the normalization process of the Hamming distance by using the PWA operator. It can be defined as follows for two sets $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$.

Definition 4. A PWAD operator of dimension n is a mapping $PWAD$: $R^n \times R^n \rightarrow R$ that has an associated weighting vector P such that $\hat{v}_i \in [0, 1]$ and $\sum_{i=1}^{n} \hat{v}_i = 1$, according to the following formula:

$$PWAD\ (\langle x_1, y_1 \rangle, \ldots, \langle x_n, y_n \rangle) = \sum_{i=1}^{n} \hat{v}_i |x_i - y_i| \tag{4}$$

where each argument (individual distance) $|x_i - y_i|$ has an associated weight (WA) v_i with $\sum_{i=1}^{n} v_i = 1$ and $v_i \in [0, 1]$, and a probabilistic weight p_i with $\sum_{i=1}^{n} p_i = 1$ and $p_i \in [0, 1]$, $\hat{v}_i = \beta p_i + (1-\beta) v_i$ with $\beta \in [0, 1]$ and $\hat{v}_i$ is the weight that unifies probabilities and WAs in the same formulation.

Note that it is also possible to formulate the PWAD operator separating the part that strictly affects the probabilistic distance aggregation and the part that affects the weighted Hamming distance. This representation is useful to see both models in the same formulation but it does not seem to be as a unique equation that unifies both models.

Definition 5. A PWAD operator is a mapping $PWAD$: $R^n \times R^n \rightarrow R$ of dimension n, if it has an associated weighting vector P, with $\sum_{i=1}^{n} p_i = 1$ and $p_j \in [0, 1]$ and a weighting vector V that affects the WAD, with $\sum_{i=1}^{n} v_i = 1$ and $v_i \in [0, 1]$, such that:

$$PWAD\ (\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle) = \\ = \beta \sum_{i=1}^{n} p_i |x_i - y_i| + (1-\beta) \sum_{i=1}^{n} v_i |x_i - y_i| \tag{5}$$

where $|x_i - y_i|$ are the individual distances and $\beta \in [0, 1]$.

If D is a vector corresponding to the arguments $|x_i - y_i|$, we shall call this the argument vector, and W^T is the transpose of the weighting vector, then the PWAD operator can be represented as follows:

$$PWAD\ (\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle) = W^T D \tag{6}$$

Note that if the weighting vector is not normalized, i.e., $\hat{V} = \sum_{i=1}^{n} \hat{v}_i \neq 1$, then, the PWAD operator can be expressed as:

$$PWAD\ (\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle) = \frac{1}{\hat{V}} \sum_{i=1}^{n} \hat{v}_i |x_i - y_i| \tag{7}$$

Note that $PWAD$ ($\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle$) = 0 if and only if $x_i = y_i$ for all $i \in [1, n]$. Note also that $PWAD$ ($\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle$) = $PWAD$ ($\langle y_1, x_1 \rangle, \langle y_2, x_2 \rangle, \ldots, \langle y_n, x_n \rangle$).

The PWAD operator is monotonic, bounded and idempotent. It is monotonic because if $|x_i - y_i| \geq |s_i - t_i|$, for all $|x_i - y_i|$, then, $PWAD$ ($\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle$) $\geq$ $PWAD$ ($\langle s_1, t_1 \rangle, \langle s_2, t_2 \rangle, \ldots, \langle s_n, t_n \rangle$). It is bounded because the PWAD aggregation is delimitated by the minimum and the maximum. That is, Min$\{|x_i - y_i|\} \leq PWAD$ ($\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle$) $\leq$ Max$\{|x_i - y_i|\}$. It is idempotent because if $|x_i - y_i| = |x - y|$, for all $|x_i - y_i|$, then, $PWAD$ ($\langle x_1, y_1 \rangle, \langle x_2, y_2 \rangle, \ldots, \langle x_n, y_n \rangle$) = $|x - y|$.

for all $|x_i - y_i|$, then, $PWAD(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle, \ldots, \langle x_n, y_n\rangle) = |x - y|$.

The PWAD operator includes many particular types of distance measures. For example, we can consider the two main cases found by analyzing the coefficient β. Basically, if $\beta = 0$, we get the weighted Hamming distance (WHD) and if $\beta = 1$, the probabilistic Hamming distance (PHD). Note that if $v_i = 1/n$, for all i, then, we get the arithmetic probabilistic Hamming distance (APHD). And if $p_i = 1/n$, for all i, then, we get the arithmetic weighted Hamming distance (AWHD).

IV. Decision making with the PWAD operator

The process to follow in decision making with the PWAD operator is similar to the process developed in [3,7,11-14], with the difference that now we are considering a problem of selection of investment strategies. The 5 steps to follow can be summarized as follows:

Step 1: Analysis and determination of the significant characteristics of the available investment strategies for the company. Theoretically, it will be represented as follows: $C = \{C_1, C_2, \ldots, C_i, \ldots, C_n\}$, where C_i is the ith characteristic of the investment and we suppose a limited number n of required characteristics.

Step 2: Fixation of the ideal levels of each characteristic in order to form the ideal investment strategy.

Table 1: Ideal investment

	C_1	C_2	…	C_i	…	C_n
P	μ_1	μ_2	…	μ_i	…	μ_n

where P is the ideal investment expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i \in [0, 1]$; $i = 1, 2, \ldots, n$, is a number between 0 and 1 for the ith characteristic.

Step 3: Fixation of the real level of each characteristic for all the investments considered.

Table 2: Available alternatives

	C_1	C_2	…	C_i	…	C_n
P_k	$\mu_1^{(k)}$	$\mu_2^{(k)}$	…	$\mu_i^{(k)}$	…	$\mu_n^{(k)}$

with $k = 1, 2, \ldots, m$; where P_k is the kth investment expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i^{(k)} \in [0, 1]$; $i = 1, \ldots, n$, is a number between 0 and 1 for the ith characteristic of the kth investment.

Step 4: Comparison between the ideal investment and the different alternatives considered using the PWAD operator. In this step, the objective is to express numerically the removal between the ideal investment and the different alternatives considered. Note that it is possible to consider a wide range of PWAD operators such as those described in Section 3 and 4.

Step 5: Adoption of decisions according to the results found in the previous steps. Finally, we should make the decision about which investment select. Obviously, our choice will be the investment with the best results according to the particular type of PWAD operator used.

V. Numerical example

In the following, we are going to develop a brief illustrative example of the new approach in a decision making problem concerning the selection of investment strategies. Assume a decision maker wants to invest some money in a market. After analyzing the market he considers five possible alternatives.

- Invest in Europe: A_1.
- Invest in North America: A_2.
- Invest in Asia: A_3.
- Invest in the three regions: A_4.
- Do not develop any investment: A_5.

After careful review of the information, the decision maker establishes the following general information about the investments. He has summarized the information of the investments in five general characteristics $C = \{C_1, C_2, C_3, C_4, C_5\}$.

- C_1: Benefits in the short term.
- C_2: Benefits in the mid term.
- C_3: Benefits in the long term.
- C_4: Risk of the investment.
- C_5: Other factors.

The results are shown in Table 3. Note that the results are valuations (numbers) between 0 and 1.

Table 3: Characteristics of the investment strategies.

	C_1	C_2	C_3	C_4	C_5
A_1	0.7	0.6	0.8	0.6	0.5
A_2	0.6	0.8	0.4	0.5	0.7
A_3	0.6	0.6	0.8	0.4	0.8
A_4	0.4	0.8	0.6	0.7	0.5
A_5	0.8	0.6	0.4	0.8	0.6

According to the objectives of the decision maker, he establishes the following ideal investment. The results are shown in Table 4.

Table 4: Ideal investment strategy.

	C_1	C_2	C_3	C_4	C_5
I	0.8	0.9	1	0.9	0.8

With this information, it is possible to develop different methods based on the PWAD operator for selecting an investment strategy. In this example, we consider the normalized Hamming distance (NHD), the weighted Hamming distance (WHD), the probabilistic Hamming distance (PHD), the arithmetic weighted Hamming distance (A-WHD), the arithmetic probabilistic Hamming distance (A-PHD) and the probabilistic weighted Hamming distance (PWAD). We assume that $\beta = 0.6$, that is, the probability has a degree of importance of 60% while the WA a degree of 40%. We also assume the following weights: $P = (0.1, 0.2, 0.2, 0.2, 0.3)$ and $V = (0.3, 0.3, 0.2, 0.1, 0.1)$. The results are shown in Table 5.

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 5: Aggregated results.

	NHD	WHD	PHD	A-WHD	A-PHD	PWAD
A_1	0.24	0.22	0.26	0.232	0.252	0.244
A_2	0.28	0.26	0.27	0.272	0.274	0.266
A_3	0.22	0.24	0.22	0.228	0.22	0.228
A_4	0.28	0.28	0.27	0.28	0.274	0.274
A_5	0.24	0.24	0.26	0.24	0.248	0.248

As we can see, depending on the particular type of PWAD operator used, the optimal choice is different. Therefore, it is interesting to establish an ordering of the investments for each particular case.

Table 6: Ordering of the investments

	Ordering		Ordering
NHD	$A_3 \succ A_1 = A_5 \succ A_2 = A_4$	A-WHD	$A_3 \succ A_1 \succ A_5 \succ A_2 \succ A_4$
WHD	$A_1 \succ A_3 = A_5 \succ A_2 \succ A_4$	A-PHD	$A_3 \succ A_5 \succ A_1 \succ A_2 = A_4$
PHD	$A_3 \succ A_1 = A_5 \succ A_2 = A_4$	PWAD	$A_3 \succ A_1 \succ A_5 \succ A_2 \succ A_4$

As we can see, depending on the particular type of distance aggregation operator used, the results may be different leading to different decisions.

VI. CONCLUSIONS

We have presented a new decision making model based on the use of the PWAD operator. The PWAD operator is a new aggregation operator that uses probabilities and WAs in the Hamming distance. Its main advantage is that it is able to unify the probability and the WA in the same formulation and considering the degree of importance that each concept may have in the aggregation. We have studied several properties and particular cases including the weighted Hamming distance, the probabilistic Hamming distance, the PWA operator, the arithmetic weighted Hamming distance and the arithmetic probabilistic Hamming distance.

We have also developed an application of the new approach in a decision making problem concerning the selection of investment strategies. We have seen that depending on the particular type of PWAD operator used, the results may lead to different decisions.

In future research, we expect to develop further developments by using generalized and quasi-arithmetic means. We will also develop a more complete formulation by using OWA operators and unified aggregation operators.

ACKNOWLEDGEMENTS

Support from the Spanish Ministry of Science and Innovation under project "JC2009-00189" is gratefully acknowledged.

REFERENCES

[1] G. Beliakov, A. Pradera and T. Calvo, *Aggregation Functions: A Guide for Practitioners*. Berlin: Springer-Verlag, 2007.
[2] T. Calvo, G. Mayor and R. Mesiar, *Aggregation Operators: New Trends and Applications*. New York: Physica-Verlag, 2002.
[3] J. Gil-Aluja, *The interactive management of human resources in uncertainty*. Dordrecht: Kluwer Academic Publishers, 1998.
[4] R.W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Techn. J.*, 29:147-160, 1950.
[5] N. Karayiannis, "Soft learning vector quantization and clustering algorithms based on ordered weighted aggregation operators," *IEEE Trans. Neural Networks*, 11:1093-1105, 2000.
[6] A. Kaufmann, *Introduction to the Theory of Fuzzy Subsets*. New York: Academic Press, 1975.
[7] J.M. Merigó, *New extensions to the OWA operators and its application in decision making* (In Spanish). PhD Thesis, Department of Business Administration, University of Barcelona, 2008.
[8] J.M. Merigó, "The probabilistic weighted averaging operator and its application in decision making", in: *Operations Systems Research & Security of Information*, Tecumseh, Canada, 2009, pp. 55-58.
[9] J.M. Merigó, "Fuzzy decision making using immediate probabilities," *Comp. & Indust. Engineering*, 2010, doi:10.1016/j.cie.2010.01.007.
[10] J.M. Merigó and M. Casanovas, "Induced aggregation operators in decision making with Dempster-Shafer belief structure," *Int. J. Intelligent Syst.*, 24:934-954, 2009.
[11] J.M. Merigó and A.M. Gil-Lafuente, "The ordered weighted averaging distance operator," *Lectures on Modelling and Simulation*, 8: 1-11, 2007.
[12] J.M. Merigó and A.M. Gil-Lafuente, "On the use of the OWA operator in the Euclidean distance," *Int. J. Comp. Sci. Eng.*, 2:170-176, 2008.
[13] J.M. Merigó and A.M. Gil-Lafuente, "Using the OWA operator in the Minkowski distance," *Int. J. Comp. Sci.*, 3:149-157, 2008
[14] J.M. Merigó and A.M. Gil-Lafuente, "New decision-making techniques and their application in the selection of financial products," *Inform. Sci.*, 180:2085-2094, 2010.
[15] R.R. Yager, "On ordered weighted averaging aggregation operators in multi-criteria decision making," *IEEE Transactions on Systems, Man and Cybernetics B*, 18:183-190, 1988.
[16] R.R. Yager, "Families of OWA operators," *Fuzzy Sets and Systems*, 59:125-148, 1993.
[17] R.R. Yager and J. Kacprzyk, *The Ordered Weighted Averaging Operators: Theory and Applications*. Norwell: Kluwer Academic Publishers, 1997.

ISBN: 978-988-17012-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)