Engineering Letters, 14:2, EL_14_2_7 (Advance online publication: 16 May 2007)

Fuzzy Measures Acquisition Methods

Amol Wagholikar[1], Peter Deer[2]

School of Information and Communication Technology[1,2],
Faculty of Engineering and Information Technology
Griffith University, PMB 50, Gold Coast Mail Centre, Queensland 9726,Australia

Abstract - **Fuzzy measures are monotonic set functions used for modelling strength of coalition of criteria in multicriteria decision-making. However, the practical application of fuzzy measures is difficult. The problem of fuzzy measures acquisition can be thought as an engineering problem. In our study, we have investigated the issue of fuzzy measures acquisition from an engineering perspective. Many methods have been suggested for resolving this issue. We are presenting our theoretical study on these methods. We extend the state-of-art by proposing similarity-based methods for fuzzy measure acquisition.**

Index Terms - Fuzzy measures, multicriteria decision-making, similarity

I. INTRODUCTION

Fuzzy measure is an important and fundamental concept used in many theories. It has emerged from the concept of capacities [1]. It is the extension of the classical measure theory [2]. Fuzzy measures are mainly used to model either uncertainty (as in the Dempster-Shafer theory of evidence or in possibility theory) or the strength of coalition of elements (as in game theory or multicriteria decision making) [3]. Dempster-Shafer theory of evidence deals with means of manipulating degrees of belief. It is based on the belief functions, which is used to combine separate pieces of information to calculate the probability of an event. Fuzzy measures model the belief functions in the Dempter-Shafer theory.

Fuzzy measures are considered conceptually important in these theories. However, their pragmatic determination can be problematic. This is mainly due to 2^n-1 coefficients.

We studied some data driven methods for fuzzy measure determination. We propose some new methods. One is using the concept of semantic similarity and the second one is using difference of weights between the attributes. We also propose an approach based on case-based reasoning.

II. LITERATURE REVIEW

We are addressing the issue of fuzzy measure determination in the context of multicriteria decision-making (MCDM). MCDM is a systematic and formal decision model. It consists of evaluation of given set of alternatives against a given set of criteria. The chosen criteria can be dependent or independent. They can be positively or negatively related with each other.

This phenomenon of interaction is well explained in the literature [4][5].

The interaction between the criteria should be considered while aggregating the criteria evaluations. Normally, additive measures such as weighted average are used for aggregation. Due to the additive nature of these measures, the importance of the sub-set of criteria is equal to algebraic summation of their individual importance. However, this is not true in the presence of interactive criteria [6]. Also while determining the global evaluation of the objects, along with the individual importance of criteria, the importance of sub-sets of criteria must be considered. Hence additive measures are inadequate to model the phenomenon of interaction.

A. *Non-additive measures*

Non-additive measures such as fuzzy measures and fuzzy integrals have the ability to model the issue of dependence of criteria.

B. Fuzzy measures

Fuzzy measures model the combined importance of a sub-set of criteria. These are monotone set functions defined over a universal set C and a non-empty family ρ of sub-sets (i.e. power set) of C. It is defined as $\mu : \rho \rightarrow [0,1]$, where

$\mu(\phi) = 0$ & $\mu(X) = 1$

The fuzzy measures possess the property of super additively and sub additively, which models the positive and negative type of interaction respectively. For more detailed description of these properties, readers are kindly referred to the literature [2].

C. *Fuzzy integrals*

Fuzzy integrals are the incremental summation of the product of the criteria evaluation and its fuzzy measures. There are two types of commonly used fuzzy integrals viz Choquet integral [1] and Sugeno integral [7]. Choquet integral is used for a quantitative setting whereas Sugeno integral is proposed for a qualitative setting. Since we have assumed a quantitative setting, we would use Choquet integral for this work. Let C = $\{c_1, c_2, \ldots, c_n\}$ be the set of elements. μ be the fuzzy measure on C. The elements of C can be criteria in MCDM, players in cooperative games or set of beliefs in

Demester-shafer theory. The Choquet integral of the function f *is* $f: C \rightarrow [0,1]$ with respect to μ is given by.

$$Ch_{\mu} = \sum_{i=1}^{n} \{f(c_{(i)}) - f(c_{(i-1)})\}\ \mu(A(i)) \qquad (1)$$

Where $._{(i)}$ indicates that the indices have been permuted such that –

$0 \le f(c_{(1)}) \le f(c_{(2)}) \le \ldots\ldots \le f(c_{(n)}) \le 1$ and $A(i) = \{c_{(i)}, \ldots\ldots, c_{(n)}\}$.

We can observe that for aggregating "n" elements, we need to determine 2^n-1 fuzzy measure coefficients. Hence their practical determination is the main issue especially for higher values of "n".

III. FUZZY MEASURE DETERMINATION METHODS

In this section, we present some of the methods suggested to solve the problem of fuzzy measure determination.

A.. Constructing the fuzzy measures by using Choquet or Sugeno integrals

In this method by using a given monotone function μ and some properties that new fuzzy measure ν is required to possess, an acceptable fuzzy measure can be constructed by determining an appropriate measurable function f on C. f can be determined directly from the decision maker or experts [8]. The new set function ν can be obtained by using the Sugeno or Choquet integral. Then as per this approach, the new fuzzy measure ν on the given data set would be ν = cho (f). μ

For this method, the main requirement is the definition of μ and required characteristics of v. In this method, we believe that there is no precise way of determining μ .

B. Constructing Fuzzy Measures by Transformations

Given a monotone function μ , the new monotone function can be found out by using suitable transformations. In this approach, the fuzzy measure can be determined as

$\nu(C) = \theta.\mu(C)$

Where θ is the transformation operator. Various types of transformations are suggested. These transformations are done with the help of conditional parameters [8]. Transforms are applicable where there is some initial value of the fuzzy measure. Based on that value a new fuzzy measure can be computed.

C. Data-driven construction methods

This is one class of methods where fuzzy measure determination is governed by data about the global evaluations obtained through experts or certain experiments. The general description of this method is as follows –

Let a set $C = \{c_1, c_2, \ldots\ldots\ldots, c_n\}$ where each element is a criterion for given decision-making problem. Let $A = \{a_1, a2, \ldots\ldots\ldots a_m\}$ be the set of objects, which are to be evaluated against C. Assume the criteria evaluations for these alternatives are x_{ij}. Assume that the global evaluations for these alternatives be ξ_i obtained by some way.

Let μ be the fuzzy measure on the above set C. Then the fuzzy measure on C can be determined by using the least square method, which is

$$\text{Minimize } e = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\xi_i - \xi_{i'})}^{\,2}$$

$$\text{Subject to } \xi_m = \sum_{i=1}^{n}(x_i - x_{i-1}).\mu(A(i))$$

In all of the above methods mentioned in previous section, we note that a new fuzzy measure is determined by a given fuzzy measure or using the global evaluation of the objects Unlike previous methods, the following methods are based on the similarity-based reasoning.

IV. SUGGESTED NEW METHODS

In this section, we would propose an approach, which is not based on pre-defined fuzzy measure. The main factor behind suggesting the new methods is the ease of application of fuzzy measure for a given practical problem. It is based on the similarity between the two data sets. The generalised description of our approach is as follows –

Let $C = \{c_1, c_2, \ldots\ldots, c_n\}$ be the factor space for an object space, $A = \{a_1, a_2, \ldots\ldots\ldots, a_n\}$.

Let $W = \{w_1, w_2, \ldots\ldots, w_n\}$ be the weights assigned to the elements of C. The elements of the factor space could be criteria in MCDM, players in game theory or beliefs in Dempster-shafer theory. The objects are to be evaluated by combining the individual evaluations of the elements of C. Consider the objects as a set of alternatives in MCDM. Suppose there is 'n' such factor & object spaces viz. C' & A', which are closely related with the above data. Since this data is related with the new factor space, there exists a similarity between the elements of the new factor space with the previous one. While computing such similarity we construct the factor space in such a way that the elements of the factor space are the elements of power set of C. The new fuzzy measure can be determined as the importance of the most similar element from the factor space. This similarity computation involves computing the similarity between a single element and sub set of elements for which the following methods are suggested [11].

A. Determining Fuzzy measures Using Weights

The underlying idea in this approach is that the similarity between two criteria can be considered as the function of the importance or weights assigned to them.

$$sim(C'_k, c_j) = f(w)$$

$$f(w) = c'_{kw} - c_{wj}$$

Intuitively speaking, the above expression indicates that two criteria can be considered similar if they are equally important in a given decision domain. Based on this concept, the similarity between elements of C' and sub-sets of C would be

$$sim(c'_k, c_{1\ldots n}) = \sum_{i=1}^{n} c'_{kw} - c_{jw} \qquad (2)$$

The above computation would result in determining the closest element from the factor space C' whose weight would give the required fuzzy measure μ.

B. Determining Fuzzy Measures Using Taxonomical Data

Other approach to resolve this problem would be constructing a data set of semantics of attributes. By using this approach the similarity between the attributes can be determined by constructing the data set of possible suitable senses of each attributes considered in the decision problem. In this approach, the semantic similarity between the attributes is governed by its information content [9]. The information content can also be termed as entropy. Thus Eqn (2) could be evaluated as

$$sim(s_k, c_j) = \max_{c \,\varepsilon\, (s_k, c_j)} [-\log p(c)]$$

$$p(c) = \frac{freq(c)}{N}$$

p(c) is the individual probability of the frequency encountering a suitable sense of s_k and c_j out of the possible N senses of the words in the given taxonomy. We have used taxonomy of Wordnet 2.0 for this computation [10]. The above approach is illustrated by choosing a random data set [11] in the context of MCDM. The results show such an approach can be a useful one for practical implementation of the fuzzy measures.

C. Fuzzy Measures Acquisition using CBR

Case-based reasoning (CBR) is a methodology that solves a new problem on the basis of previous problems and their solutions [12] [13]. Case-based reasoning can also be termed as similarity-based reasoning. Similarity-based reasoning is a sub-set of reasoning by analogy where a problem is solved using the analogies between the same domain. The problem definition is based on the context or a particular model chosen to represent the problem [14]. In a particular context, it is possible that the system has to rely on incomplete, imprecise or vague information. Fuzzy logic is proven to be a good tool for modelling the imprecise information. Combination of fuzzy logic with CBR results in Fuzzy logic-based CBR approaches [15][16]. Especially in our problem, the decision maker may not be able to express the criteria evaluations precisely. Hence we have chosen to adopt a fuzzy CBR approach for our problem. Fuzzy logic can be used in case representation [17] [18]. We are using fuzzy sets for representing criteria evaluations in a given decision problem.

We now propose an approach in a purely quantitative setting. We propose an interactive model of fuzzy measure determination using a set of past decisions in the form of a case-base. We chose to use absolute values of the attributes. The model is presented in the following text.

1.Let $C = \{C_1, C_2, \ldots\ldots, C_n\}$ be the set of criteria against which a set A of alternatives $\{A_1, A_2, \ldots\ldots\ldots, A_n\}$ is to be evaluated. The individual weights, W_i on the criteria is elicited from the decision maker such that $\sum_{i=1}^{n} w_i = 1$.

For the above decision problem, we have case-base in the following form.

i. A set of attributes same as the new decision problem or target problem

ii. Absolute values of each attribute.

iii. A set of individual attribute weights defined by the decision maker at the previous instance of making a decision.

iv. A set S of fuzzy measure coefficients defined at the previous instance of decision-making. We assume that these coefficients are constructed using expert's feedback.

v. Global evaluation of the object obtained by using the Choquet integral and the fuzzy measure coefficients obtained from the experts.

2. After constructing the above data set, the similarity of the cases w.r.t to the new decision problem is determined using City-block distance metric or Manhattan distance metric. There are various similarity measures in the literature [19]. However, since the case library consists of absolute values of the attributes, the Manhattan distance metric is a suitable measure for this approach. The Manhattan distance metric is defined as the absolute difference between the values of pair of attributes or criteria I.e.

$$d_{ij} = | x_{ik} - x_{jk} | \qquad (3)$$

where x_i & x_j are the values of the attribute for attribute "k". For computing the similarity, the distance values should be normalized in the range [0,1] by the maximum distance $d_{\max}$ between the attribute value in the target case and corresponding values in the case base. The most similar case then can be computed by computing the global similarity of a given case from the case base with given alternative in the new decision problem. The global similarity will be calculated by simple weighted average of the local similarities of the given attributes.

Thus we have,

$$d(Case\_t, Case\_u) = \frac{\sum_{k=1}^{n} w_k d_{ij}}{\sum_{k=1}^{n} w_k} \qquad (4)$$

The weight of the criteria "k" is denoted by w_k. Based on this distance, the similarity is calculated as

$$Sim(Case\_t, Case\_u) = 1 - d(Case\_t, Case\_u) \qquad (5)$$

where $Case\_t$ = Target case i.e. the new decision problem & $Case\_u$ = Case from the case library i.e. past decision problems.

1. After such similarity computation the case with maximum similarity would be retrieved and the values of the fuzzy measures would be considered for solving the new decision problem. However, the decision maker is likely to have different importance values. Hence using a direct rating method [20], we propose an importance scale designed to obtain the new fuzzy measure coefficients from the decision maker. The importance scale is designed to ask the decision maker about the importance of sub-set of criteria. The decision maker can directly rate the sub-set of criteria in a range between 0 and 1. The chosen range is divided into different linguistic variables. It is assumed that the decision maker would normally rate the importance of subset of criteria using the linguistic variables. The designed scale is shown in the tabular form below.

Table 1: Importance Scale

Sub-set is Less important	Extremely	0
	Highly	0.1
	Very	0.2
	Strongly	0.3
	Moderately	0.4
	Equally	0.5
Sub-set is More important	Moderately	0.6
	Strongly	0.7
	Very	0.8
	Highly	0.9
	Extremely	1

Based on this scale, the new fuzzy measures can be elicited from the decision maker. The agreement of the decision maker with the values elicited from this experiment and the values given by the similarity computation indicates the utility of the proposed system.

D. ILLUSTRATION

The sample data set should involve absolute values of the criteria. There should be a mix of numerical and ordinal values. Our search for a pre-defined and standard data set shows that it is difficult to obtain such data set for the illustration. Hence we have constructed a customized data set. The following is a classic car selection problem where 4 cars are to be evaluated against a set of 5 criteria.

Table 2. Sample Data set for car selection problem

	Price	Power	Fuel Economy	Comfort
Car A	$30,000	2000	10	Good
Car B	$25,000	1000	12	OK
Car C	$20,000	1500	12	OK
Car D	$35,000	2500	8	Very Good

Suppose for the above data set, we have case base of past fuzzy measure values in the form of a vector of dimension j x k where j is the number of cases and k is 2^n-1 fuzzy measure coefficients. The case base for this data set would be represented as below.

$$C = [Case_i, \{k_1, k_2, \ldots\ldots, k_{2n-1}\}].$$

This sample data set is built using Microsoft Access DBMS. We have chosen this tool due to its simplicity and also smaller scale of the application. After performing the similarity computations, we get the following

Car A is most similar to Case 2

Hence for computing the worth of Car A, the default values of the fuzzy measures applied in case 2 could be used for evaluating worth of the car A. Similarly,

Car B is most similar to Case 4
Car C is most similar to Case 4
Car D is most similar to Case 1

And hence based on the above result, the worth of cars is shown in Table 4 below.

Table 3. Car Evaluations

Alternatives	Global Evaluations
Car A	0.78
Car C	0.77
Car D	0.68
Car B	0.675

By visual observation of the data in Table 2 and the results in table 4, the decision maker may not agree with this ranking as one can argue about the higher ranking of B. For testing the above result, the fuzzy measures are elicited from the decision makers. We performed hypothetical experimentation with 10 decision makers.

Table 4. Evaluations based on decision maker's input

Alternatives	Global Evaluation based on the Importance Scale
Car A	0.78
Car C	0.77
Car D	0.77
Car B	0.68

From the above results, we can observe that the decision maker had ranked only car C & D equally. Other evaluations are same as the system's evaluations using the default case-base. Hence the system results are in an approximate agreement with the decision maker and it provides an optimal solution.

V. CONCLUSION AND FUTURE DIRECTION

In this work, we have studied various methods for determining monotonic set functions i.e. fuzzy measures. We mainly studied the pragmatic issues associated with the fuzzy measures and some approaches to resolve those issues. Data driven methods suggest an optimal approximate solution.

However, in those methods fuzzy measure is determined based on a pre-defined value of the fuzzy measure. We have proposed new approaches based on the similarity between the two data sets. These approaches give offer a significantly different perspective towards fuzzy measure acquisition problems. Hence these approaches make a valuable contribution in the field of fuzzy measures theory.

ACKNOWLEDGMENT

We would like to thank School of Information and Communication Technology, Griffith University for giving all kind of support to undertake this research.

REFERENCES

[1] G. Choquet, Theory of capacities, *Ann.Inst.Fourier 5*, 1953, 131-195.

[2] Z. Wang, G.J. Klir, Fuzzy Measure Theory, Plenum Press, New York, 1992.

[3] Kojadinovic, I., Modelling interaction phenomena using fuzzy measures: on the notions of interaction and independence, Fuzzy sets and Systems 135 (2003) 317-340.

[4] M.Grabisch, *Fuzzy integral in multi-criteria decision making"*, Fuzzy Sets and Systems, 69, 1995, 279-298.

[5] J. Marichal, *Aggregation Operators for Multicriteria decision aid*, Ph.D. thesis, University of Li`ege, Belgium, 1999.

[6] A. Wagholikar, P. Deer, Non-additive Measures In Multicriteria Decision Making, *International Conference on Computational Intelligence for Modelling*, Control & Automation, Gold Coast, Australia, July 2004.

[7] M. Sugeno, *Theory of fuzzy integrals and its applications*, PhD Dissertation, Tokyo Institute of Technology, Japan, 1974.

[8] Klir G, J., Constructing fuzzy measures in expert systems, fuzzy sets and systems 92 (1997), 251-264.

[9] D. Lin, An Information-theoretic definition of similarity, *Proceedings of the Fifteenth International Conference on Machine Learning*, ISBN: 1-55860-556-8, 1998, Pages: 296 – 304.

[10] G. Miller, WordNet: An on-line lexical database, *International Journal of Lexicography*, 3(4), 1990.

[11] A. Wagholikar, P. Deer, Determination of fuzzy measures using similar data set, the 4th Asia Pacific International Symposium on Information Technology, Gold Coast, Australia, January 2005, 319-322.

[12] Leake, B., D., Case-Based Reasoning: Experiences, Lessons, & Future Directions, AAAI Press, Menlo Park, California (1996).

[13] Hansen, B. K.: Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory, Master of Computer Science Thesis, Technical University of Nova Scotia, Halifax, Nova Scotia, Canada (2000).

[14] Jaczynski, M., Trousse, B.: Fuzzy logic for the retrieval step of a case based reasoner, Proceedings of the Second European Workshop on Case-Based Reasoning, Chantilly, France (1994), 313—321.

[15] Aggour, K.S., Pavese, M., Bonissone, P.P., &. Cheetham, W.E. (2003). SOFT-CBR: A self-optimizing fuzzy tool for case-based reasoning. Proceedings of the Fifth International Conference on Case-Based Reasoning (pp. 5-19). Trondheim, Norway: Springer.

[16] Cheetham, W., Graf,J.: Case-based reasoning in color matching, Proceedings of the Second International Conference on Case-Based Reasoning,1-12, Providence, RI: Springer (1997).

[17] Yager, R.R.: Soft aggregation methods in case based reasoning. Applied Intelligence, 21(3),(2004) 277-288.

[18] Pal, S.K., Dillon, T.S., & Yeung, D.S. (Eds.): Soft computing in case based reasoning. Springer Verlag, London (2001).

[19] Kardi, T.: Similarity Measurement., Web Resource Accessed March (2005).

[20] Bilgic, T., Turksen, I.: Measurement of Membership Functions: Theoretical and Empirical Work, Handbook of Fuzzy Sets and Systems, Vol.1, foundations, (1997).