Annals of **F**uzzy **M**athematics and **I**nformatics
Volume 19, No. 3, (June 2020) pp. 239–274
ISSN: 2093–9310 (print version)
ISSN: 2287–6235 (electronic version)
http://www.afmi.or.kr
https://doi.org/10.30948/afmi.2020.19.3.239

@FMI

On balance of uncertainty in shadowed sets

M. A. Ibrahim, T. O. William-West, A. F. D. Kana, D. Singh

Reprinted from the
Annals of Fuzzy Mathematics and Informatics
Vol. 19, No. 3, June 2020

Annals of Fuzzy Mathematics and Informatics
Volume 19, No. 3, (June 2020) pp. 239–274
ISSN: 2093–9310 (print version)
ISSN: 2287–6235 (electronic version)
http://www.afmi.or.kr
https://doi.org/10.30948/afmi.2020.19.3.239

@FMI

On balance of uncertainty in shadowed sets

M. A. Ibrahim, T. O. William-West, A. F. D. Kana, D. Singh

Received 26 August 2019; Revised 4 November 2019; Accepted 18 January 2020

Abstract. A shadowed set, S, facilitates crisp decision-making with a fuzzy set F. It is constructed with the aid of different optimization-based principles. Among these principles, the requirement of uncertainty balance guarantees preservation of the uncertainty of F in S. In order to gain further insight on uncertainty balance, some essential mathematical properties which characterize uncertainty-balance-based objective function, $J(\alpha)$, are studied. These properties provide theoretical explanation for interpreting and analyzing $J(\alpha)$ and its ensuing optimum partition threshold α. Two senses of uncertainty balance are discussed in this paper. Their combined efficiency in enhancing clustering results is illustrated with the aid of synthetic data set used in shadowed C-means clustering. Finally a need for five-region shadowed sets, S_5, is pointed out. A closed-form formula for determining its optimum thresholds is proposed and exemplified on typical fuzzy set and synthetic dataset.

2010 AMS Classification: 03B52, 03E72, 28E10

Keywords: Fuzzy set, Shadowed set, Three-way approximation.

Corresponding Author: T. O. William-West (westtamuno@gmail.com)

1. Introduction

Shadowed sets, introduced by Pedrycz [20], are direct algorithmic construction of fuzzy sets. By relying on Kleene's three-valued logic [10], they represent fuzzy sets with the aid of three regions (i.e., core, shadow and excluded zones). In fact, shadowed sets stem from the necessity to make crisp decision with fuzzy sets [18, 31].

A key aspect of shadowed sets is determination of the required pair of thresholds which balances the uncertainty distorted as a result of transforming a given fuzzy set into three regions [20, 24]. There are various methods of inducing shadowed sets

[5, 9, 20, 27, 8, 29, 33, 36, 40, 37]. These methods anchor on different optimization-based principles.

The main goal of shadowed sets is to relocate the fuzziness inherently associated with the original fuzzy set into a shadow region. As used in this paper, we determine a pair of symmetric thresholds, $(\alpha, 1-\alpha)$ and are guided by a principle of uncertainty balance. This principle, follows from a general optimization-based principle known as principle of uncertainty and information invariance [11].

Klir [11, 12], in general terms, suggested a principle of uncertainty (and its related information) invariance; which states that when making transformations between different mathematical theories characterized by uncertainty, the amount of uncertainty should be preserved under these transformations. In the framework of shadowed sets, Klir's principle of uncertainty invariance may be viewed as a principle of uncertainty balance. This principle advocates for preservation of the uncertainty (and related information) of the original fuzzy set in the ensuing shadowed set.

It is difficult to retain the total amount of fuzziness of a fuzzy set, F, in the ensuing shadowed set, S. Therefore, studies [20, 27, 29] which tackle this issue have been carried out to improve the initial formulation of uncertainty balance. In order to study the concept of uncertainty balance, some essential mathematical properties for comprehending its notion, as well as, characterizing an optimum solution for trisecting F are needed. This underlines the significance of this paper.

In this study, based on a principle of uncertainty balance, some mathematical properties related to shadowed sets are studied. By exploiting a notion of fuzziness set, φ_F, of F [27], we show that for optimum threshold $(1-\alpha)$-cut, the number of patterns, x with $\mu_{\varphi_F}(x) \geq \varphi(1-\alpha)$ are exactly the overall amount of fuzziness in F. Based on this idea, we give an algorithm to obtain the optimum threshold. Its performance is illustrated on some typical fuzzy sets, as well as synthetic data set. We investigate the usefulness of a principle of uncertainty balance in shadowed C-means clustering. It is observed that uncertainty balance is theoretically meaningful for retention of the fuzziness encountered in a dataset X. However, the related information in F can be adequately preserved by exploiting the two senses of uncertainty balance discussed in this paper. Further, illustrative examples are provided to underline that better approximation of F in terms of minimum approximation error and adequate fulfilment of a principle of uncertainty balance calls for development of five-region shadowed sets.

1.1. **Motivation and contribution of the study.** This paper is motivated by rare provision of mathematical properties of the objective function of uncertainty balance in shadowed sets.

The idea of uncertainty balance in shadowed sets has been discussed in [27] by exploiting a concept of gradual number of fuzziness, specific gradual elements, and partial assignment functions. In order to effectively communicate a notion of balance of uncertainty, Yao [33], modified the method in [27] in a simplified notational system. Without delving into deeper mathematical concepts such as gradual number of fuzziness, specific gradual elements, and partial assignment functions, a concept of uncertainty balance in shadowed sets can be easily expressed and explained in simple terms for the interest of early researchers in the field. Also, for a variety of reasons which we enunciate.

From the formulation in [27, 33], it is difficult to envision the contribution of the elements in the shadow region to the overall amount of fuzziness of F. In this paper, we strive to overcome this difficulty by: (1) establishing simple terms that let us easily communicate about uncertainty balance in shadowed sets and also allow us suggest, precisely, a quick way to determine the optimum thresholds, (2) presenting some new mathematical properties which guides selection of the required thresholds, and (3) using an alternative closed-form formula for induction of shadowed sets.

In fact, uncertainty balance can be interpreted from two perspectives: (a) the relationship between the size of the shadow region of a shadowed set and the overall amount of fuzziness of F and, (b) the relationship between the total amount of fuzziness to be minimized in the core and excluded zones, and to be maximized in the shadow zone. This two interpretations come with two closed-form formula and achieves the same results. However, they give distinct insight on the distribution of uncertainty and how it can be balanced and controlled in practical situations (e.g., data clustering).

The first interpretation is what we pursue. The second interpretation has been discussed in [27]. Both interpretations when put together facilitate decision-making, especially in classification of patterns in shadowed clustering. In fact, the membership of a pattern to a cluster can be evaluated by using $(\alpha, 1-\alpha)$-cuts to define the size of the core, shadow and excluded zones of a cluster. The relationship (i.e., fuzziness-cut) between the total amount of fuzziness to be minimized and maximized in the aforesaid zones can then be exploited to decide which object are to be used to compute a cluster centroid. Therefore, with the aid of the two interpretations, a fuzzy cluster can be approximated from two thresholds: α-cut and fuzziness-cut.

From numerical point of view, five-way approximation of fuzzy sets, when compared to three-way approximations, promises better approximation results, and it is worth developing. So, this paper introduces five-region shadowed sets.

This paper contributes to studies on shadowed set approximation of fuzzy sets in the following aspects. (1) it provides theoretical analysis of a notion of uncertainty balance in shadowed sets; (2) existence, uniqueness and determination of the optimum thresholds which satisfy a principle of uncertainty balance is discussed; (3) two approaches (i.e, direct procedure and closed-form formula), together with an algorithm for determining the the optimum thresholds are provided. The performance of these approaches are evaluated using typical fuzzy sets and synthetic data sets; (4) by considering the fuzziness cut, $\varphi(1-\alpha)$ and the pair of optimum thresholds, $(\alpha, 1-\alpha)$, we propose an idea for inducing shadowed C-means clustering. Essentially, by adopting, $\varphi(1-\alpha)$ and $(\alpha, 1-\alpha)$, an approach to deal with objects which exhibit very high degree of uncertainty to a fuzzy cluster is introduced; (5) we provide a starting point for the development of five-way approximation of fuzzy sets.

Throughout the study, we adhere to the following symbols:

(α, β): a pair of optimum thresholds (not necessarily symmetric),
$(\alpha, 1-\alpha)$: a pair of symmetric optimum thresholds,
S_n: n-region shadowed set
S_α: shadowed set induced by α,
φ: measure of fuzziness,
μ_F: membership function of fuzzy set F,
φ_F: fuzziness set of F,
μ_{φ_F}: membership function of the fuzziness set of F defined from μ_F to $[0,1]$,
f_S: membership function of shadowed set S,
$E(A)$: error function which measures the error incurred in set A,
$J(\alpha)$: optimization function of uncertainty balance,
c_k: kth cluster,
v_k: centroid of the kth cluster,
μ_{ik}: membership grade of the ith object in the kth cluster,
C: number of clusters,
n: number of objects.

2. Theoretical foundations

2.1. **Fuzzy set and Zadeh's three-way approximations.** A fuzzy set, F is characterized by a membership function

$$\mu_F : X \longrightarrow [0,1] \tag{2.1}$$

which establishes the degree of membership, $\mu_F(x) \in [0,1]$, of an element, $x \in F$. It can be cut at a point $\alpha \in [0,1]$ in order to filter irrelevant elements.

Zadeh [34] suggested two cuts α and β such that F is partitioned into the following three regions:

Quantitative approximation:
$Pos(F) = \{x \in X | \mu_F(x) > \beta\}$,
$Bnd(F) = \{x \in X | \alpha \le \mu_F(x) \le \beta\}$,
$Neg(F) = \{x \in X | \mu_F(x) < \alpha\}$.

Qualitative approximation:
$Pos(F) = \{x \in X | \mu_F(x) = T\}$,
$Bnd(F) = \{x \in X | \mu_F(x) = U\}$,
$Neg(F) = \{x \in X | \mu_F(x) = F\}$.
Here the truth values T, U and F denote *true, unknown* and *false*, respectively [10].

2.2. **Fuzziness of fuzzy sets.** A function $\varphi : P(X) \longrightarrow \mathbf{R}$ which characterize the degree of fuzziness of each fuzzy subset A of X is called a measure of fuzziness, where $P(X)$ denotes the set of all fuzzy subsets of X [11]. Therefore, the fuzziness of a (fuzzy) set, F, is a measure that quantifies the underlying vagueness or uncertainty in F. In the literature, various measures of fuzziness have been proposed (see, for example, [4, 12, 14]). In general, any measure of fuzziness should have the following properties:

P1) $\varphi(A) = 0$, for all non-fuzzy sets A,

P2) $\varphi(\mu_F(x)) = 1$, if $\mu_F(x) = \frac{1}{2}$ (i.e., level of maximum fuzziness),

P3) $\varphi(\mu_F(x)) = \varphi(1 - \mu_F(x))$, where $1 - \mu_F(x)$ is the complement of $\mu_F(x)$,

P4) $\varphi(\mu_F(x))$ is monotonically increasing in $[0, \frac{1}{2}]$ as $\mu_F(x) \to \frac{1}{2}$ (whether from left or right) and, monotonically decreases in $[\frac{1}{2}, 1]$ as $\mu_F(x) \to 1$.

In [11, 12, 13], the measure of fuzziness of a given fuzzy set F is defined as:

$$\varphi(F) = \sum_{i=1}^{|F|}[1 - |2\mu_F(x_i) - 1|] \tag{2.2}$$

If $\mu_F(x_i) = \frac{1}{2}$ for all i, then we have

$$\varphi(F) = Card(F), \tag{2.3}$$

where $Card(F)$ or $|F|$ is the cardinality of fuzzy set F.

Fuzziness set of Fuzzy sets: Given a fuzzy set $F = \{(x, \mu_F(x))|x \in X\}$ drawn from a universe

$$X = \{x_i|i = 1, 2, ..., n\},$$

its corresponding fuzziness set, φ_F, is defined in [27] as

$$\varphi_F = \{(x, \mu_{\varphi_F}(x))|x \in X \wedge \mu_{\varphi_F}(x) = \varphi(\mu_F(x))\}.$$

The fuzziness of F is completely described by its fuzziness set.

2.3. **Three-way decisions and shadowed sets.** A three-way decision is conceived as follows.

Three-way decisions: Let $X = \{x_i : i = 1, 2, 3, ..., n\}$ be a finite nonempty set of objects and $C = \{c_j : j = 1, 2, 3, ..., k\}$ be a finite set of conditions. A three-way decision T relies on an evaluation function, $\mu_T : X \longrightarrow \{a, r, u\}$, which classifies $x \in X$ into T according to its fulfilment of $c_j \in C$. Here a, r and u represent accept, reject and non-commitment decision, respectively.

Shadowed sets: Shadowed sets are concrete models of three-way decisions. They approximate a fuzzy set by exploiting Kleene's three-valued valuations.

Let $(0, 1)$ denote uncertainty of an element's membership to a set S. Given a non-empty (finite) universe X, a shadowed set is characterized by the mapping [20]:

$$f_S : X \longrightarrow \{0, (0, 1), 1\} \tag{2.4}$$

An element, $x \in X$, which (completely) belong to S is assigned the membership grade $f_S(x) = 1$ and put in the core region, $Cor(S)$. An element $x \in X$ which is (completely) excluded from S is assigned the membership grade $f_S(x) = 0$, and put in the reduced area, $Red(S)$. Subsequently, an element, $x \in X$, which is doubtful is assigned the membership grade $f_S(x) = (0, 1)$ or $f_S(x) = \frac{1}{2}$, as suggested in [2], and it is put in the shadow region, $Shd(S)$.

When a shadowed set is constructed from a given fuzzy set, F, with the aid of some criteria $(\alpha, 1 - \alpha)$, $(\alpha \in (0, \frac{1}{2}])$, the mapping in Equation (2.4) is defined on the membership grades of F by

$$f_S(\mu_F(x)) = \begin{cases} 1, \text{ if } \mu_(x) > 1 - \alpha \\ \frac{1}{2}, \text{ if } \alpha \leq \mu_F(x) \leq 1 - \alpha \\ 0, \text{ if } \mu_F(x) < \alpha. \end{cases} \tag{2.5}$$

2.4. **Optimization-based principles for shadowed sets.** Principle of uncertainty balance and principle of minimum transformation error are key ingredients in constructing shadowed sets.

A principle of uncertainty balance requires retention of the overall amount of fuzziness of a given fuzzy set and its proper relocation in the ensuing shadowed set [20, 27].

A principle of minimum error demands that the optimized thresholds required to induce a shadowed set should be determined by minimizing the total error incurred in changing elements membership grade from $\mu_F(x)$ to $f_S(\mu_F(x))$.

It is important to underline that the original idea for constructing shadowed sets does not require fulfillment of a principle of minimum error. However, Deng and Yao [5] proposed error-based interpretation of shadowed sets; which suggests that this principle is desirable.

In what follows, we formally explicate principle of uncertainty balance, and some situations under which it is useful.

Balance of uncertainty: A principle of uncertainty balance captures the essence of uncertainty of fuzzy sets, particularly when they are transformed into a different formalism (i.e., three-way approximation). It guides construction of three-way approximation of F based on available amount of fuzziness in F.

The rationale behind a principle of uncertainty balance is to deliver a concise characterization of F such that the produced shadow region is justified in the light of the fuzziness encountered in F. Intuitively, this principle quantifies the extent to which the shadow region accounts for $\varphi(F)$. Some situations under which this principle is useful are:

(1) Data mining; where it is necessary to transform a data set into more interpretable representation in order to support knowledge discovery. Here the aforesaid principle guarantees information preservation.
(2) Granular computing. When information granules are to be constructed from data, a principle of uncertainty balance is deployed in the form of a principle of justifiable granularity [23] in order to assess the extent to which the information granule is supported by data.

2.5. **Induction of shadowed sets from fuzzy sets.** Throughout the rest of the paper, we will use the symbol $|.|$ to denote the absolute value function. Also, we note that there are other ways of inducing shadowed sets, for example, in [36], game-theoretic shadowed sets has been suggested. However, as we do not wish to discuss aspects of game-theory, this work has been omitted. Moreso, in [33] a three-way approximation of F is discussed from three principles (i.e., generalized three-way approximation, principles of minimum distance and minimum risk cost). As our focus is on a special model of three-way approximation viz. shadowed sets, and our objective is not to review decision-threoretic approach to approximation of F, we will not delve into such methods.

Furthermore, we note that the authors in [5, 6, 23, 27] initially used a pair of feasible thresholds, (α, β) in their formulations. However, they chose a single threshold in terms of α (i.e., $\beta = 1-\alpha$) for the purpose of calculational convenience. Therefore, we adopt a pair of symmetric thresholds to report their methods.

2.5.1. *Pedrycz's model.* In [20], the following condition is suggested as a requirement for balance of uncertainty:

$$\varphi(Cor(S)) + \varphi(Red(S)) = \varphi(Shd(S)). \tag{2.6}$$

Equation (2.6) interprets the sum of the elevated area and the reduced area as the amount of decreased uncertainty and, the shadow is the amount of increased uncertainty [33].

For a finite (discrete) fuzzy set F, Pedrycz expanded Equation (2.6) as

$$\sum_{\mu_F(x_i)<\alpha} \mu_F(x) + \sum_{\mu_F(x_i)>1-\alpha} (1-\mu_F(x_i)) = Card(\{x_i \in X | \alpha \leq \mu_F(x) \leq 1-\alpha\}). \tag{2.7}$$

In order to obtain an α-cut that satisfies Equation (2.6) and (2.7), Equation (2.7) is reformulated as an optimization problem:

$$\alpha(opt) = arg(min_\alpha |a+b-c|), \tag{2.8}$$

where

$$a = \sum_{\mu_F(x_i)<\alpha} \mu_F(x_i),$$

$$b = \sum_{\mu_F(x_i)>1-\alpha} (1-\mu_F(x_i)),$$

$$c = Card(\{x_i \in X | \alpha \leq \mu_F(x_i) \leq 1-\alpha\}).$$

2.5.2. *Tahayori, Sadeghian and Pedrycz model.* In line with a principle of uncertainty balance, Yao et al. [33] reformulated Equation (2.7) to explicate Tahayori *et al.* [27] model as follows:

$$\sum_{\mu_F(x_i)<\alpha} \varphi(\mu_F(x_i)) + \sum_{\mu_F(x_i)>1-\alpha} \varphi(\mu_F(x_i)) = \sum_{\alpha \leq \mu_F(x_i) \leq 1-\alpha} (\varphi(\frac{1}{2}) - \varphi(\mu_F(x_i))) \tag{2.9}$$

and the optimized α-cut is computed using

$$\alpha(opt) = arg(min_\alpha |d+e-f|), \tag{2.10}$$

where
$d = \sum_{\mu_F(x_i)<\alpha} \varphi(\mu_F(x_i))$ is the total amount of fuzziness in $Red(S)$,
$e = \sum_{\mu_F(x_i)>1-\alpha} \varphi(\mu_F(x_i))$ is the total amount of fuzziness in $Cor(S)$,
$f = \sum_{\alpha\leq\mu_F(x_i)\leq 1-\alpha} (\varphi(\frac{1}{2}) - \varphi(\mu_F(x_i)))$ is the total amount of fuzziness introduced when the original amount of fuzziness of elements placed in $Shd(S)$ is changed from $\varphi(\mu_F(x))$ to $\varphi(f_S(x))$.

2.5.3. *Deng and Yao model.* As an alternative approach, Deng and Yao [5] rely on a principle of minimum error, and suggest computation of the optimized thresholds as

$$\alpha(opt) = arg(min_\alpha E(\mu_F)), \tag{2.11}$$

where

$$E(\mu_F) = E(Red(S)) + E(Cor(S)) + E(Shd(S)) \tag{2.12}$$

is the summation of errors in the three regions:

$$E(Red(s)) = \sum_{\mu_F(x_i)<\alpha} \mu_F(x_i),$$

$$E(Cor(S)) = \sum_{\mu_F(x_i)>1-\alpha} (1 - \mu_F(x_i))$$

and

$$E(Shd(S)) = \sum_{\alpha\leq\mu_F(x_i)\leq 1-\alpha} |\frac{1}{2} - \mu_F(x_i)|.$$

2.5.4. *Ibrahim and William-West model.* Ibrahim and William-West [9] observed that the threshold, α, which satisfies Equation (2.7), may not satisfy the principle of uncertainty balance. The attitudes of different decision makers towards the level of uncertain they want to retain is taken into consideration. It is observed that if decision makers are optimistic, less of the original amount of uncertainty of F will be retained in S, although the misclassification error may increase. If decision makers are pessimistic, more of the original amount of uncertainty of F will be retained in S. Based on this perspective, a shadowed sets which embraces a compromise between optimistic and pessimistic way of uncertainty preservation are introduced by searching for the average of the balance of uncertainty as follows. Therefore, Equation (2.7) was modified to

$$V(\alpha) = |\sum_{\mu_F(x_i)<\alpha} \mu_F(x_i) + \sum_{\mu_F(x_i)>1-\alpha} (1 - \mu_F(x_i)) - \sum_{\alpha\leq\mu_F(x_i)\leq 1-\alpha} \mu_F(x_i)| \tag{2.13}$$

and Equation (2.8) now becomes

$$\alpha(opt) = arg(avg_\alpha V(\alpha)), \tag{2.14}$$

where

$$avg_\alpha V(\alpha) = \frac{V(\alpha_1) + V(\alpha_2) + ... + V(\alpha_k)}{k}. \tag{2.15}$$

$V(\alpha)$ is the balance of uncertainty, $\alpha_i \in [\mu_{min}, \frac{1}{2}], i = 1, 2, ..., k$, and μ_{min} is the minimum membership grade in F. Here k is the number of feasible thresholds α.

From Equation (2.15), the α-cut producing a value of $V(\alpha_i)$ nearest to the average $avg_\alpha V(\alpha)$, is selected as the optimal threshold.

2.5.5. *William-West, Ibrahim and Kana model.* William-West et al. [29] proposed a method for calculating $\alpha(opt)$ as follows:

$$\alpha(opt) = arg(min_\alpha|\varphi(F) - \varphi(S_{\alpha_i})|), \tag{2.16}$$

where S_{α_i} are candidate shadowed sets determined from feasible thresholds $\alpha_i \in [\mu_{min}, \frac{1}{2}], i = 1, 2, ..., k$.

Also, in line with studies in [5], the membership grade of elements in $Shd(S)$ is taken as

$$\frac{\mu_F(x_1) + \mu_F(x_2) + ... + \mu_F(x_k)}{k},$$

where $\alpha \leq \frac{\mu_F(x_1)+\mu_F(x_2)+...+\mu_F(x_k)}{k} \leq 1-\alpha$ and k is the number of objects embraced by the shadow region.

A comparative study on shadowed sets induced by these aforesaid methods has been carried out in [29]. The findings show that for randomly generated fuzzy sets, different methods may produce different shadowed sets.

In summary, uncertainty balance is generally achieved by formulating an optimization-based objective function whose terms are representations of an amount of fuzziness of objects distributed in various regions of a shadowed set. So, Equations (2.8), (2.11) and (2.14) in their present form are appropriate only for certain situations. Consequently, these formulations may suffer from some limitations in reaching adequate fulfillment of a principle of uncertainty balance.

3. The proposed method

Studies in shadowed sets point out that; to fulfill a principle of uncertainty balance, the main focus should be to achieve the following equation [27, 29]:

$$\varphi(F) = \varphi(Shd(S)). \tag{3.1}$$

We note that the fuzziness of a shadowed set is couched in the shadow region. Equation (3.1) infers that the size of $Shd(S)$, relative to $\varphi(F)$, should determine the optimality of a given threshold value.

Noticeably, if $\mu_F(x) = \frac{1}{2}$, then $\mu_{\varphi_F}(x) = 1$. Since element in $Shd(S)$ have $\mu_S(x) = \frac{1}{2}$. Therefore, we have $Card(Shd(S)) = \varphi(Shd(S))$.

The meaningfulness of the idea proposed in this paper anchors on this notion, and the fact that a shadowed set partitions a fuzzy set F into three disjoint regions, such that the overall amount of fuzziness of F is distributed into the three regions, and should be equivalent to the cardinality of $Shd(S)$. Based on a principle of uncertainty balance, we have the following closed-form formula:

$$\sum_{x \in Cor(S)} \mu_{\varphi_F}(x_i) + \sum_{x \in Red(S)} \mu_{\varphi_F}(x_i) + \sum_{x \in Shd(S)} \mu_{\varphi_F}(x_i) = Card(Shd(S)). \tag{3.2}$$

One can interpret Equation (3.2) as the extent to which the shadow region (i.e., right hand side of Equation (3.2)) accounts for $\varphi(F)$. That is, the optimal threshold to be determined, should induce a shadow region satisfying Equation (3.1).

More explicitly, for any shadowed set S, $Cor(S)$ and $Red(S)$ are crisp sets, whereas $Shd(S)$ is not. Due to the elevation, reduction and fixing actions of S, the uncertainty reduced (UR) and uncertainty introduced (UI) in S are respectively calculated as:

$$UR = \sum_{x \in Cor(S)} \mu_{\varphi_F}(x_i) + \sum_{x \in Red(S)} \mu_{\varphi_F}(x_i)$$

and

$$UI = Card(Shd(S)) - \sum_{x \in Shd(S)} \mu_{\varphi_F}(x_i).$$

Therefore, a principle of uncertainty balance is interpreted as the next equation:

$$UR = UI$$

which simplifies to Equation (3.2).

3.1. **Theoretical analysis of uncertainty balance in shadowed sets.** The following properties (i.e., theorems) facilitate discussion of uncertainty balance.

Throughout this section, $(\alpha_i, 1-\alpha_i), 1 \leq i \leq r$ denote the feasible thresholds for constructing a shadowed set from a given fuzzy set, where $\alpha_i \in (0, \frac{1}{2}]$. Further, we assume that there are at least two of such pairs of feasible thresholds. Here α is selected from the set consisting of all membership grades, $\mu_F(x)$, of an object, $x \in F$, between 0 and $\frac{1}{2}$.

We note that for a given (discrete) fuzzy set $F = \{(x, \mu_F(x)) : x \in X\}$, the choice of α_i is exactly the membership grades $\mu_F(x) \in (0, \frac{1}{2}]$. Also, it is important to underline that selecting thresholds which do not coincide with $\mu_F(x) \in (0, \frac{1}{2}]$ may produce an approximation which does not conform with the structure of data, F.

Theorem 3.1. *Let $(\alpha, 1-\alpha) \in \{(\alpha_i, 1-\alpha_i)\}$ be any pair of feasible threshold values. Then $\mu_{\varphi_F}(x) < \varphi(1-\alpha)$ if and only if $x \notin Shd(S_\alpha), \alpha_i \in (0, \frac{1}{2}]$.*

Proof. Suppose $\forall x \in F, \mu_{\varphi_F}(x) < \varphi(1-\alpha)$. By implication, $\mu_F(x) < \alpha$ or $\mu_F(x) > 1-\alpha$. Then $\mu_F(x) \notin [\alpha, 1-\alpha]$. Thus, $x \notin Shd(S_\alpha)$.

Conversely, suppose $x \notin Shd(S_\alpha)$. Then either $\mu_F(x) < \alpha$ or $\mu_F(x) > 1-\alpha$. Let $\varphi(1-\alpha) = r \in \mathbf{R}$. Since φ is monotonically increasing in $[0, \frac{1}{2}]$ and monotonically decreasing in $[\frac{1}{2}, 1]$, we have $\mu_{\varphi_F}(x) < \varphi(1-\alpha), \forall x \in F$.

This completes the proof. □

Theorem 3.2. *Let B be the set of all patterns, $x \in F$, satisfying $\mu_{\varphi_F}(x) \geq \varphi(1-\alpha)$ with $\alpha \in (0, \frac{1}{2}]$. Then*

(1) *$Red(S_\alpha) \cup Elv(S_\alpha) = B' = \{x \in F : \mu_{\varphi_F}(x) < \varphi(1-\alpha)\}$,*
(2) *$Shd(S_\alpha) = B$.*

Proof. Recall that by the symmetric property of α, $\varphi(\alpha) = \varphi(1-\alpha)$. Let $F = \{(x, \mu_F(x)) : i = 1, 2, ..., n\}$, $B = \{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha)\}$ and

$B' = \{x \in F : \mu_{\varphi_F}(x) < \varphi(1-\alpha)\}$. For any $x \in B'$, since $\varphi(\mu_F(x)) = \mu_{\varphi_F}(x)$, property (P4) of Subsection 2.2 guarantees that one of the inequalities hold:

* $\mu_F(x) < \alpha \leq \frac{1}{2}$, or
* $1-\alpha < \mu_F(x) \leq 1$.

That is, either $\mu_F(x)$ is trailing behind α or, $1-\alpha$ is trailing behind $\mu_F(x)$. This is supported by Theorem 3.1. Then the membership grade of any $x \in B'$ is in one of the intervals $(0, \alpha)$ or $(1-\alpha, 1]$. Thus

$$Red(S_\alpha) \cup Elv(S_\alpha) = B' = \{x \in F : \mu_{\varphi_F}(x) < \varphi(1-\alpha)\}.$$

Further, the pair of thresholds $(\alpha, 1-\alpha)$ partitions F into disjoint regions:

$$Red(S_\alpha), Elv(S_\alpha) \text{ and } Shd(S_\alpha).$$

So $B = F/B'$. Theorem 3.1 guarantees that $Shd(S_\alpha) = B$. In fact, $\forall x \in B$, $x \in Shd(S_\alpha)$. Conversely, $\forall x \in Shd(S_\alpha)$, $x \in B$.
This completes the proof. □

Theorem 3.3. *For any fuzzy set F, there exists a three-way approximation, $S_\alpha(\alpha \in \{\alpha_i : i = 1, 2, ..., r\})$, of F having the nearest amount of fuzziness, $m = Card(Shd(S_\alpha))$, to the total amount of fuzziness, $\varphi(F)$, in F. Here α_i denote any feasible threshold.*

Proof. Let $\alpha_1 < \alpha_2 < ...\alpha_r$ denote the relation among the feasible threshold values and φ_F be the fuzziness set of F. Then $\varphi(1-\alpha_r), \varphi(1-\alpha_{r-1})...$ upto $\varphi(1-\alpha_1)$ cut φ_F from r different parts. Since α_1 corresponds to $min\{\mu_F(x_i)\}$, from Theorem 3.2, we can find numbers

$$m_1 = Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_1)\}) = Card(Shd(S_{\alpha_1}))$$

and

$$m_r = Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_r)\}) = Card(Shd(S_{\alpha_1})),$$

respectively having the maximum and minimum amount of fuzziness associated with their corresponding shadow regions $Shd(S_{\alpha_1})$ and $Shd(S_{\alpha_r})$. Other threshold values will induces shadowed sets whose total amount of fuzziness lie between m_1 and m_r. Sorting m_i in ascending order and, fixing $\varphi(F)$ in a suitable position among the m_i, say m_α, we obtain the nearest amount of fuzziness to $\varphi(F)$. Hence, the result holds. □

Remark 3.4. (1) Theorem 3.2 shows that every pair of feasible thresholds, $(\alpha, 1-\alpha)$, partitions a given fuzzy set in such a way that objects in the shadow region of the tri-partition are always at and/or above the fuzziness cut, $\varphi(1-\alpha)$, of a fuzziness set φ_F. By cutting φ_F at the right position, $\varphi(1-\alpha)$ can be used to induce an S which retains the amount of fuzziness most equivalent to $\varphi(F)$. In fact, the principal idea of three-way decisions, requires that the optimized threshold, $1-\alpha$, facilitate an evaluation of an object for acceptance or rejection whenever its associated uncertainty is below $\varphi(1-\alpha)$, otherwise a non-commitment decision is reached.

(2) Theorem 3.3 discusses the existence of a shadowed set which fulfills a principle of uncertainty balance. In three-region shadowed sets, such m_α (i.e., as observed in the theorem) is the best possible.

In what follows, we discuss more detailed issue on uncertainty balance.

Consider the fuzzy set F depicted in Figure 1. Suppose $(\alpha_i, 1-\alpha_i), 1 \leq i \leq r$, are feasible thresholds for constructing shadowed sets S_{α_i}. Let Figure 2 depict the graph of the fuzziness set of F.

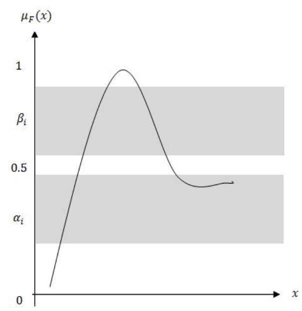

FIGURE 1. A curve representing a fuzzy set.

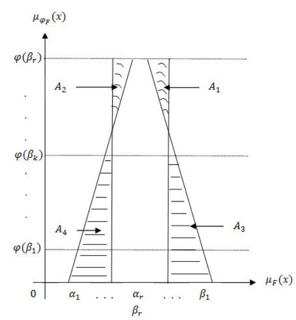

FIGURE 2. A graph of fuzziness set of a fuzzy set.

The pair of optimum thresholds, say, $(\alpha_k, 1-\alpha_k)$ to be found can be comprehended from Equation (3.1) and Theorem 3.2. From Theorem 3.2, $A_3 + A_4$ is the union of $Red(S_{\alpha_k})$ and $Elv(S_{\alpha_k})$. Also $A_1 + A_2$ captures the patterns in $Shd(S_{\alpha_k})$. Hence, the goal of a principle of uncertainty balance is to search for $1-\alpha_k$-cut such that the line $\varphi(1-\alpha_k)$ cuts the graph of Figure 2 to the horizontal in such a way that the total number of objects, $x \in F$, satisfying the inequality $\mu_{\varphi_F}(x) \geq \varphi(1-\alpha)$ (i.e., objects on and/or above the horizontal line drawn to $\varphi(1-\alpha_k)$) is the most equivalent to $\varphi(F)$.

Illustratively, suppose F is depicted as in Figure 3 such that $\varphi(F) = 10.4$, say, and if the total number, $n[\varphi(1-\alpha_k)]$, of patterns, $x \in F$, on and/or above the line $\varphi(1-\alpha_k)$ is 10. Then $(\alpha_k, 1-\alpha_k)$ becomes the desired pair of optimum thresholds,

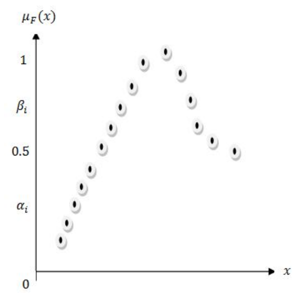

FIGURE 3. A graph depicting a discrete fuzzy set

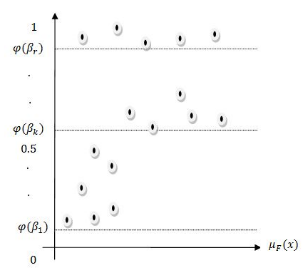

FIGURE 4. The optimum fuzziness cuts, $\varphi(1-\alpha_k)$, of the fuzziness set of the discrete fuzzy set

as it produces the closest number to 10.4. Figure 4 brings home the idea.

In this instance (i.e., Figure 2),

$$Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_k)\}) = Card(Shd(S_{\alpha_k})) = 10.$$

Noticeably,

$$\varphi(F) = A_3 + A_4 + A_1 + A_2. \tag{3.3}$$

By assuming that the integral exist, the overall fuzziness of F can be expressed from Equation (3.1), (3.2) and (3.3) to obtain:

$$\varphi(F) = \int_{x:\mu_{\varphi_F}(x)<\varphi(1-\alpha_k)} \mu_{\varphi_F}(x) + \int_{x:\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_k)} \mu_{\varphi_F}(x). \tag{3.4}$$

Following the ideas obtained from Theorem 3.2 and Equation (3.1), we have the following modification of Equation (3.3):

(3.5)

$$Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_k)\}) = \int_{x:\mu_{\varphi_F}(x)<\varphi(1-\alpha_k)} \mu_{\varphi_F}(x) + \int_{x:\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_k)} \mu_{\varphi_F}(x)$$

or
(3.6)
$$Card(\{x \in F : \alpha_k \leq \mu_F(x) \leq 1-\alpha_k\}) = \int_{x:\mu_{\varphi_F}(x)<\varphi(1-\alpha_k)} \mu_{\varphi_F}(x) + \int_{x:\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_k)} \mu_{\varphi_F}(x).$$

Therefore, the required pair of thresholds which fulfills uncertainty balance is obtained by minimizing the absolute difference:

$$J(\alpha) = |\int_{x:\mu_{\varphi_F}(x)<\varphi(1-\alpha)} \mu_{\varphi_F}(x) + \int_{x:\mu_{\varphi_F}(x)\geq\varphi(1-\alpha)} \mu_{\varphi_F}(x) - z|, \tag{3.7}$$

where $z = Card(\{x \in F : \alpha \leq \mu_F(x) \leq 1-\alpha\}), \alpha \in (0, \frac{1}{2}]$.

In passing, we search for the following optimum threshold:

$$\alpha(opt) = argmin_\alpha J(\alpha). \tag{3.8}$$

For discrete membership values $\mu_F(x_i), (i = 1, 2, ..., n)$, Equation (3.6) is expressed as:

$$\alpha(opt) = argmin_\alpha J'(\alpha), \tag{3.9}$$

where

$$J'(\alpha) = |\sum_{x\in Cor(S)} \mu_{\varphi_F}(x_i) + \sum_{x\in Red(S)} \mu_{\varphi_F}(x_i) + \sum_{x\in Shd(S)} \mu_{\varphi_F}(x_i) - z|.$$

Equivalently, we have

$$\alpha(opt) = arg(min_\alpha |g + h + i - z|), \tag{3.10}$$

where

$$g = \sum_{\mu_F(x_i)<\alpha} \varphi(\mu_F(x_i)),$$
$$h = \sum_{\mu_F(x_i)>1-\alpha} \varphi(\mu_F(x_i)),$$
$$i = \sum_{\alpha\leq\mu_F(x_i)\leq 1-\alpha} \varphi(\mu_F(x_i)),$$
$$z = Card(\{x \in X : \alpha \leq \mu_F(x_i) \leq 1-\alpha\}).$$

Theorem 3.5. *Let F be a fuzzy set and $(\alpha_i, 1-\alpha_i), 1 \leq i \leq r$, be pairs of feasible thresholds. For fixed k, the pair $(\alpha_k, 1-\alpha_k) \in \{(\alpha_i, 1-\alpha_i)\}$ satisfying a principle of uncertainty balance is unique. That is, F has exactly one optimum solution $S_{\alpha_k(opt)}$.*

Proof. Suppose for contradiction that F has two solutions S_{α_k} and $S_{\alpha_{k'}}$. Then we must have $Shd(S_{\alpha_k})$ and $Shd(S_{\alpha_{k'}})$ as the nearest amount of fuzziness to $\varphi(F)$ such that $Card(Shd(S_{\alpha_k})) = Card(Shd(S_{\alpha_{k'}}))$.

Selecting the feasible α-cuts in F and arranging them in ascending order in a set: $A = \{\alpha_1 = min\mu_F(x_i), \alpha_2, ..., \alpha_r\}$, it can be observed that A is well-ordered. Consequently, the set of intervals $\{(\alpha_i, 1-\alpha_i)\}, 1 \leq i \leq r$, is well-ordered. That is, for any two pairs of feasible thresholds $(\alpha, 1-\alpha), (\alpha', 1-\alpha') \in \{(\alpha_i, 1-\alpha_i)\}$, only one of the proper subsethood relation holds:

(1) $(\alpha', 1-\alpha') \subset (\alpha, 1-\alpha)$,
(2) $(\alpha, 1-\alpha) \subset (\alpha', 1-\alpha')$.
This contradicts the initial assumption which infers that

$$Card(Shd(S_{\alpha_k})) = Card(Shd(S_{\alpha_{k'}})).$$

Then each pair $(\alpha, 1-\alpha)$ must induce a shadow region of distinct cardinality. Further, in Equations (3.5) or (3.7), the value of $J(\alpha)$ is influenced by the size of z. That is, different $Shd(S_\alpha)$, $\alpha \in (min\mu_F(x), \frac{1}{2}]$ give rise to different $Card(Shd(S_\alpha))$. Thus there is only one pair of thresholds, $(\alpha_k, 1-\alpha_k)$, generating a number $Card(Shd(S_{\alpha_k}))$ nearest to $\varphi(F)$, and this pair of thresholds is unique. □

Theorem 3.6. *The objective function* $J(\alpha)$ *attains its maximum and minimum values in* $(0, \frac{1}{2}]$, *where*

$$J(\alpha) = |\sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_i)} \mu_{\varphi_F}(x_i) + \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_i)} \mu_{\varphi_F}(x_i) - Card(Shd(S_{\alpha_i}))|.$$

Proof. Let $\alpha_1 < \alpha_2 < ... < \alpha_r, \alpha_i \in (0, \frac{1}{2}]$, be the relation among feasible threshold values. Then the following subsethood relation holds:

$$(\alpha_r, 1-\alpha_r) \subset (\alpha_{r-1}, 1-\alpha_{r-1}) \subset ... \subset (\alpha_1, 1-\alpha_1).$$

Thus we have

$$Card(Shd(S_{\alpha_r})) < Card(Shd(S_{\alpha_{r-1}})) < ... < Card(Shd(S_{\alpha_1})).$$

Further, the relation $\varphi(1-\alpha_1) < \varphi(1-\alpha_2) < ... < \varphi(1-\alpha_r)$ holds. By implication, the following inequalities are obtained:

$$\sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_1)} \mu_{\varphi_F}(x_i) < \sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_2)} \mu_{\varphi_F}(x_i) < ... < \sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_r)} \mu_{\varphi_F}(x_i)$$

and

$$\sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_1)} \mu_{\varphi_F}(x_i) > \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_2)} \mu_{\varphi_F}(x_i) > ... > \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_r)} \mu_{\varphi_F}(x_i).$$

Moreover, for all $p \neq q$, we have the following equation:

$$\sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_p)} \mu_{\varphi_F}(x_i) + \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_p)} \mu_{\varphi_F}(x_i) = \sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha_q)} \mu_{\varphi_F}(x_i) + \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha_q)} \mu_{\varphi_F}(x_i). \tag{3.11}$$

However, $J(\alpha_p) \neq J(\alpha_q), p \neq q$. So Equation (3.9) infers that the value of $J(\alpha)$ is influenced by

$$Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_i)\}) = Card(Shd(S_{\alpha_i})).$$

Since $\varphi(1-\alpha_1) < \varphi(1-\alpha_2) < ... < \varphi(1-\alpha_r)$, the set

$$\{Card(\{x \in F : \mu_{\varphi_F}(x) \geq \varphi(1-\alpha_i)\}) : 1 \leq i \leq r\}$$

is well-ordered. Hence, it must contain two numbers nearest to $\varphi(F)$ and farthest from $\varphi(F)$. These numbers guaranty the existence of the required minimum and maximum.

This completes the proof. □

3.2. **Algorithm to obtain optimum thresholds from the proposed method.** A constructive algorithm to obtain the optimum threshold which fulfills a principle of uncertainty balance can be drawn in the following steps:

Input: $F = \{(x_i, \mu_F(x_i)) : i = 1, 2, ..., n\}$,
Output: Optimum threshold value $1 - \alpha$, Let

$$\alpha_j \in (min\{\mu_F(x_i)\}, \mu_F(x_{k+1}), ..., \mu_F(x_{k+r}) \leq \frac{1}{2})$$

(i.e., α_j, for each j is drawn from the ordered pair

$$(min\{\mu_F(x_i)\}, \mu_F(x_{k+1}), ..., \mu_F(x_{k+r}) \leq \frac{1}{2})),$$

$j, k, r = 1, 2, ...p$,

(1) Determine $\varphi(1 - \alpha_j)$ by $\varphi(1 - \alpha_j) = 1 - |2(1 - \alpha_j) - 1|$, $j = 1, 2, ...p$,
(2) Compute $\varphi(F)$ by Equation (2.2),
(3) Compute $\mu_{\varphi_F}(x_i)$ from $\mu_F(x_i)$ by $\mu_F(x_i) = 1 - |2\mu_F(x_i) - 1|$, $\forall i$,
(4) Initialization:
$Shd(S_{1-\alpha_j}) = \{\}$,
(5) *while* $\mu_{\varphi_F}(x_i) \geq \varphi(1 - \alpha_j)$ (for any i and fixed j):
put x in $Shd(S_{1-\alpha_j})$. i.e., $Shd(S_{1-\alpha_j}).append(x_i)$
do for all $i \leq n$,
(6) Determine $Card(Shd(S_{1-\alpha_j}))$, $\forall j$,
(7) Compute $d_j = |\varphi(F) - Shd(S_{1-\alpha_j})|$, for each j,
(8) Determine $min\{d_j\}$ over j,
(9) Return $1 - \alpha_k$ as optimum threshold value if $d_k = min\{d_j\}$, for fixed index k.

In the algorithm, α_j denote the feasible thresholds drawn from the interval

$$[min\{\mu_F(x)\}, max\{\mu_F(x)\} \leq \frac{1}{2}].$$

(5) and (6) determine the amount of fuzziness in each feasible shadowed set, $S_{1-\alpha_j}$. (7) computes the discrepancies between the fuzziness in F and $S_{1-\alpha_j}$. Once the minimum discrepancy is determined by line 8, the optimum threshold which satisfies a principle of uncertainty balance is returned.

Another approach for obtaining the optimum threshold is to apply the closed-form formula in Equation (3.2). From (5), we may also compute the number, $n(1 - \alpha_k)$, of patterns $x \in F$ satisfying the inequality therein. Line 6 can be replaced with "if $\|floor\varphi(F) \leq n(1 - \alpha_k) \leq \|ceil\varphi(F)$", then return $1 - \alpha_k$ as optimum threshold.

3.3. **Relationship between the proposed idea and Tahayori *et al.* method.** As reported in [27], the optimum threshold (which corresponds to $\varphi(1-\alpha)$ in Figure 2), to be found should be such that

$$A_1 + A_2 = A_3 + A_4. \tag{3.12}$$

That is, for the fuzziness set, φ_F, the threshold to be found should be such that the membership values of elements with $\mu_{\varphi_F}(x) < \varphi(1 - \alpha)$ would compensate the membership values of all elements with $\mu_{\varphi_F}(x) \geq \varphi(1-\alpha)$ to become full members of φ_F. Theoretically, this is captured with the aid of the next equations:

$$\sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha)} \mu_{\varphi_F}(x) = \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha)} (1-\mu_{\varphi_F}(x)) \tag{3.13}$$

or

$$\sum_{\mu_F(x)<\alpha} \mu_{\varphi_F}(x) + \sum_{\mu_F(x)>1-\alpha} \mu_{\varphi_F}(x) = \sum_{\alpha\leq\mu_F(x)\leq1-\alpha} (\varphi(\frac{1}{2}) - \mu_{\varphi_F}(x)). \tag{3.14}$$

The next equation (i.e., Equation (3.14) and (3.15) shows that this formulation agrees with our proposed idea (Equations (3.4)–(3.80). However, Equation (3.12) may come with the following misleading interpretations, which may misinform a user:

(1) one may interpret it as the optimum threshold to be found should be such that the total amount of fuzziness in the reduced and elevated area should be equal to the the total amount of fuzziness in the shadow area. Such threshold value may be too restrictive and may not guaranty $\varphi(F) \equiv \varphi(Shd(S_\alpha))$. Tahayori *et al.*, [27] use Equation (3.12) to denote the fuzziness which is to be minimized and maximized in the $Cor(S), Red(S)$ and $Shd(S)$, respectively.
(2) from semantic standpoint, the notational system of Equation (3.13) may not give a clear idea on the overall amount of fuzziness in $Shd(S_\alpha)$. So, the relationship between the expected size of the shadow region of S and the overall amount of fuzziness in F as distributed over the three regions of S is not easily visualized.

Notice that from right hand side of Equation (3.12), we can set

$$\sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha)} (1-\mu_{\varphi_F}(x)) = Card(\{x \in F : \alpha \leq \mu_F(x) \leq 1-\alpha\}) - \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha)} \mu_{\varphi_F}(x). \tag{3.15}$$

Therefore, Equation (3.12) can now be expressed as:

$$\sum_{\mu_{\varphi_F}(x)<\varphi(1-\alpha)} \mu_{\varphi_F}(x) + \sum_{\mu_{\varphi_F}(x)\geq\varphi(1-\alpha)} \mu_{\varphi_F}(x) = Card(\{x \in F : \alpha \leq \mu_F(x) \leq 1-\alpha\}). \tag{3.16}$$

This is equivalent to Equation (3.5).

3.4. **Two senses of uncertainty balance.** The ideas drawn from the proposed formulation and the work in [27], when combined together gives two interpretations (i.e., thresholds $(\alpha, 1-\alpha)$ for abridging membership grades and, the fuzziness-cut for making crisp and/or non-commitment decisions) of balance of uncertainty. They point out two senses of uncertainty balance:

(1) structure-sensitive approximation. That is, balanced partitioning of F in terms of its related information and uncertainty. This is achieved with the aid of $(\alpha(opt), 1-\alpha(opt))$,
(2) information filtering. That is, separating objects for which there is sufficient information for taking clear decision action (i.e., objects with low enough

fuzziness) from the ones we are not clear about (i.e., objects with very high fuzziness). This is achieved with the aid of the fuzziness-cut, $\varphi(1-\alpha(opt))$.

Practically speaking, an effective decision-making strategy can be drawn from (1) and (2) above. That is, on one hand, one could partition a data set involving uncertainty with the aid an optimum criteria. On the other hand, the fuzziness-cut can facilitate selection of objects that could be used in precisely interpreting the data set of concern.

In order to bring home this idea, we will apply this strategy in shadowed C-means clustering.

4. Examples

4.1. **Synthetic data set.** A synthetic fuzzy set F_1 consists of 30 items. Their membership values is given in Table 1. The optimized thresholds, fuzziness of the

Table 1. Synthetic dataset of typical fuzzy set, F_1

Items	$\mu_{F_1}(x_i)$	Items	$\mu_{F_1}(x_i)$	Items	$\mu_{F_1}(x_i)$
x_1	0.05	x_{11}	0.65	x_{21}	0.15
x_2	0.10	x_{12}	0.65	x_{22}	0.40
x_3	0.20	x_{13}	0.75	x_{23}	0.45
x_4	0.29	x_{14}	0.75	x_{24}	0.55
x_5	0.30	x_{15}	0.75	x_{25}	0.28
x_6	0.30	x_{16}	0.80	x_{26}	0.38
x_7	0.35	x_{17}	0.80	x_{27}	0.15
x_8	0.35	x_{18}	0.85	x_{28}	0.12
x_9	0.50	x_{19}	0.90	x_{29}	0.26
x_{10}	0.60	x_{20}	0.95	x_{30}	0.31

induced shadowed set, discrepancy and error incurred by various methods are reported in Table 2:

Table 2. Optimized thresholds for discrete fuzzy set F_1

Method	Optimized $(\alpha, 1-\alpha)$	$\varphi(S)$	$\varphi(F_1)$	Discrepancy	Error
Pedrycz	$(0.38, 0.62)$	6	15.88	9.88	6.2
Taha *et al*	$(0.26, 0.74)$	16	15.88	**0.12**	4.74
Deng and Yao	$(0.30, 0.70)$	13	15.88	2.88	4.90
Ibra and Will	$(0.25, 0.75)$	19	15.88	3.12	4.74
Willi *et al*	$(0.25, 0.75)$	19	15.88	3.12	4.74
Proposed	$(0.26, 0.74)$	16	15.88	**0.12**	4.74

The methods whose threshold values fulfill a principle of uncertainty balance are in bold font. Their outstanding performance are easily seen (i.e., row 4 and 8 of Table 2) from their minimal discrepancy in retention of overall amount of fuzziness encountered in F_1.

As another example, we consider the commonly used fuzzy set [5, 27] viz. Gaussian fuzzy set.

4.2. **Gaussian fuzzy set.** Consider a Gaussian membership function

$$\mu_F(x_i) = e^{-(\frac{x_i-\bar{x}}{\sigma})^2}$$

describing ages, x_i, of players in Aduvie School baseball team. The mean age is $\bar{x} = 30.68$, the spread, $\sigma = 6.09$, and

$x_i = 21, 21, 22, 23, 23, 24, 24, 25, 25, 28, 29, 29, 31, 32, 33, 33, 34, 35, 36, 36, 36, 36, 38, 38, 40$ represent the ages of the players.

The relationship explaining the *normal* distribution of the ages may be view from the membership values summarized in Table 3:

Table 3. Gaussian fuzzy set, F_2

Items	$\mu_{F_2}(x_i)$	Items	$\mu_{F_2}(x_i)$	Items	$\mu_{F_2}(x_i)$
x_1	0.08	x_9	0.83	x_{17}	0.61
x_2	0.08	x_{10}	0.92	x_{18}	0.47
x_3	0.13	x_{11}	0.92	x_{19}	0.47
x_4	0.20	x_{12}	1.00	x_{20}	0.47
x_5	0.29	x_{13}	0.95	x_{21}	0.47
x_6	0.29	x_{14}	0.87	x_{22}	0.24
x_7	0.42	x_{15}	0.87	x_{23}	0.24
x_8	0.42	x_{16}	0.74	x_{24}	0.10

The optimum thresholds for crisply interpreting ages, as well as, the discrepancy and error induced by various methods for the Gaussian fuzzy set are shown in Table 4:

Table 4. Optimized thresholds for discrete fuzzy set F_2

Method	Optimized $(\alpha, 1-\alpha)$	$\varphi(S)$	$\varphi(F_2)$	Discrepancy	Error
Pedrycz	$(0.47, 0.53)$	4	11.32	7.32	3.82
Taha *et al*	$(0.24, 0.76)$	12	11.32	**0.68**	2.72
Deng and Yao	$(0.24, 0.76)$	12	11.32	**0.68**	2.72
Ibra and Will	$(0.08, 0.92)$	22	11.32	10.68	5.02
Willi *et al*	$(0.20, 0.80)$	13	11.32	1.68	2.72
Proposed	$(0.24, 0.76)$	12	11.32	**0.68**	2.72

Here (i.e., in Table 4) the methods in bold font achieve the same level of discrepancies and have outstanding performance.

Remark 4.1. (1) It can be deduced from column 5 and 6 of Tables 2 and 4 that when constructing a shadowed set, an optimum threshold value which minimizes approximation error, may not balance the underlying uncertainty of F in S. Also, a threshold value which balances the underlying uncertainty may not minimize the error in approximation.

(2) Adequate fulfillment of both uncertainty balance and minimization of approximation error may be achieved from $n(> 3)$-region shadowed sets. This will be demonstrated in Section 5.

4.3. **Application example.** We consider shadowed C-means (SCM) clustering [16]; which relies on shadowed sets, as our application example. In what follows, brief introduction of SCM clustering is provided.

4.3.1. *Shadowed C-means (SCM) clustering.* Shadowed set approximation of fuzzy c-means clustering is performed by exploiting an approach known as shadowed C-mean clustering (SCM) [17]. An SCM clustering approach represents each cluster as a shadowed set. For desired number of clusters C, it computes the means (or centroids) $v_1, v_2, ..., v_C$ of the clusters, associates objects to each centroid according to some computed membership degree, and determines a collection of thresholds

$\alpha_1(opt), \alpha_2(opt), ..., \alpha_C(opt)$ which facilitate the three-way clusters. Accordingly, each fuzzy cluster $\bar{c_k}$ is transformed into c_k and partitioned in three regions with the aid of $\alpha_k(opt)$ ($\alpha \in [\mu_{i_{min}}, \frac{\mu_{ik_{min}}+\mu_{ik_{max}}}{2}]$) determined bu using a shadowed set approximation method. Here $\mu_{ik_{min}}$ and $\mu_{ik_{max}}$ are minimal and maximal membership grades in the kth cluster.

Mitra *et al.* [17] modified Equation (2.7) and obtained:

$$\alpha_k(opt) = min_{\alpha_k} | \sum_{\mu_{ik}<\alpha_k} \mu_{ik} + \sum_{\mu_{ik}>(\mu_{ik_{max}}-\alpha_k)} (\mu_{ik_{max}} - \mu_{ik}) - c'|, \tag{4.1}$$

where $c' = Card(\{i|\alpha_k \leq \mu_{ik} \leq (\mu_{ik_{max}} - \alpha_k)\})$ and μ_{ik} is calculated as

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} (\frac{||x_i^a - v_k||^2}{||x_i^a - v_j||^2})^{\frac{2}{m-1}}} \tag{4.2}$$

similar to fuzzy C-means clustering (FCM). Here we take m as 2.
After completing the computation of $\alpha_k(opt)$, the SCM weighs the objects based on their fulfilment of the critria $1-\alpha(opt)$, $[\alpha(opt), 1-\alpha(opt)]$ and $\alpha(opt)$, inducing regions:

$$Cor(c_k) = \{x \in c_k : \mu_{c_k}(x) > 1 - \alpha_k\},$$

$$Shd(c_k) = \{x \in c_k : \alpha \leq \mu_{c_k}(x) \leq 1 - \alpha_k\}$$

and

$$Red(c_k) = \{x \in c_k : \mu_{c_k}(x) < \alpha_k\}$$

which are then used to update the centroid, v_k, defined by:

$$v_k = \frac{\sum_{x \in Cor(c_k)} x_i^a + \sum_{x \in Shd(c_k)} \mu_{ik}^m x_i^a + \sum_{x \in Red(c_k)} (\mu_{ik})^{m^m} x_i^a}{u' + v' + w'}, \tag{4.3}$$

where

$$u' = Card(\{i|\mu_{ik} > (\mu_{ik_{max}} - \alpha_k)\}),$$

$$v' = \sum_{\alpha_k \leq \mu_{ik} \leq (\mu_{ik_{max}} - \alpha_k)} \mu_{ik}^m$$

and

$$w' = \sum_{\mu_{ik}<\alpha_k} \mu_{ik}^{m^m}.$$

The elements in $Cor(c_k)$ do not have any fuzzy weight factor, while the elements in other regions are treated as in FCM clustering technique. In particular, the elements in $Red(c_k)$ are assigned double exponential fuzzifier. This shows a very low bias for excluded objects, thereby minimizing the effect of noise and outliers [16].

The SCM clustering method repeats computation of μ_{ik} from Equation (3.18) and then the computation of the centroid is repeated until convergence (i.e., $|\mu_{ik}(t) - \mu_{ik}(t-1)| < \epsilon$, at maximum iteration t) is reached.

4.3.2. *Cluster validity indexes.* In the literature [7, 19, 26], some validity indexes have been proposed to evaluate the quality of clustering, namely DB, XB, PBM indexes. This indexes anchor on a principle that minimizes intra-cluster distance (compactness) and maximizes inter-cluster distance (separation). As observed in [25], formulation of these indexes are primarily couched with *minimum* and/or *maximum* distance between centroids and, are therefore sensitive to outliers. In order to overcome this issue, Saitta et al. [25] proposed a bounded index for cluster validity called the *score function* (SF). This index will be used to evaluate the performance of various SCM clustering technique. Here SF is defined as:

$$SF = 1 - \frac{1}{e^{e^{bcd-wcd}}}, \tag{4.4}$$

where
$bcd = \frac{1}{nc}\sum_{j=1}^{C} ||v_j - \bar{x}||^2.n_j$ is the between cluster distance,
$wcd = \frac{1}{C}\sum_{j=1}^{C} \sqrt{\frac{1}{Card(c_j)}\sum_{x \in c_j} ||x - v_j||^2}$ is the within cluster distance,
n and n_j are respectively the number of patterns in dataset X and the numbers of patterns in the j-th cluster. Also, C and $\bar{x}$ denote the number of cluster and the mean of the dataset, respectively.

Accordingly, we define the H-index as follows:

$$H = -log(\frac{wcd}{bcd}). \tag{4.5}$$

The goodness of clustering is determined by minimum values of SF and H.

Example

The aim of the following example is to demonstrate how two-senses perspective of uncertainty balance is fast to discover meaningful structures (i.e., clusters) in datasets and to underline that uncertainty balance,when applied in data clustering, on its own may not effectively learn the underlying structure inherent in the dataset. Noticeably, all shadowed set approximation methods may learn the structure of data at different (higher or lower) number of iterations. We purpose to observe the method that is quick to learn the underlying pattern of the dataset.
As the rate of convergence to the original structure of the dataset can be visualized from the third iteration, we report three iterations which help in explicating the expected centroids determined by various approximation methods.

A two dimensional dataset (i.e., D_{32}) ([17, 38]), is used in the SCM clustering experiments. In order to minimize the number of iterations and accelerate the process and quality of clustering, we adopt the method of obtaining initial centroid described in [1]. D_{32} is presented in Table 5, where a and b denote the attributes of the pattern described in Figure 5. We note that the initial cluster centroids (i.e., patterns with x marks in Figure 5) that are used are calculated as $[6.9, 3.3]$ and $[7.1, 6.0]$. Also, our choice of fuzzifier m [22, 38, 39] is taken as 2.

Table 5. Clustering dataset 1

Pattern	a	b	Pattern	a	b	Pattern	a	b
x_1	1	3	x_{12}	5	3	x_{23}	10	4
x_2	2	3	x_{13}	5	4	x_{24}	10	5
x_3	3	2	x_{14}	6	3	x_{25}	11	2
x_4	3	3	x_{15}	7	3	x_{26}	11	3
x_5	3	4	x_{16}	8	3	x_{27}	11	4
x_6	4	1	x_{17}	9	3	x_{28}	12	3
x_7	4	2	x_{18}	9	3	x_{29}	13	3
x_8	4	3	x_{19}	9	4	x_{30}	4	20
x_9	4	4	x_{20}	10	1	x_{31}	7	20
x_{10}	4	5	x_{21}	10	2	x_{32}	10	20
x_{11}	5	2	x_{22}	10	3	.	.	.

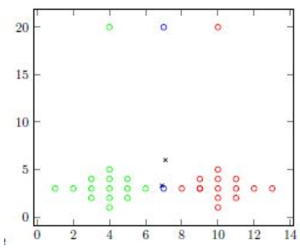

FIGURE 5. Topology of D_{32}

Two main clusters can be easily observed by eyeballing. Patterns x_{30}, x_{31} and x_{32} (see Table 6) are inserted into the dataset as noisy objects. An efficient representation of D_{32} should effectively deal with the aforesaid fuzziness by classifying them into the boundary region of the clusters. One can identify two notable clusters (to the left and right of Figure 5) lying between 0 to 7 and 8 to 14 of the a-axis. The next table shows the thresholds, and the centroids obtained by various methods in the third iteration: The deviation in column 5 of Table 6 is calculated by using the

Table 6. Thresholds and centroids for D_{32}

Methods	Optimized threshold (α, β)	Cluster centre 1	Cluster centre 2	$D(c_i)$	SF	H
Pedrycz	$(0.132, 0.856), (0.044, 0.955)$	$[6.943, 3.021]$	$[7.033, 19.899]$	650.736	0.9998	1.3124
Taha. *et al.*	$(0.120, 0.834), (0.084, 0.915)$	$[6.878, 3.025]$	$[7.060, 19.590]$	642.378	1.0000	1.0524
Deng and Yao	$(0.151, 0.814), (0.064, 0.935)$	$[6.926, 3.011]$	$[7.045, 19.873]$	650.077	0.9998	1.3096
Ibra and Willi	$(0.021, 0.920), (0.079, 0.920)$	$[7.009, 2.948]$	$[8.468, 18.691]$	621.064	1.0000	1.0284
Willi *et al*	$(0.034, 0.930), (0.066, 0.934)$	$[6.934, 3.010]$	$[6.969, 19.487]$	639.561	0.9999	0.5485
Proposed	$(0.120, 0.834), (0.084, 0.915)$	$[6.878, 3.025]$	$[7.060, 19.590]$	642.378	1.0000	1.0524
Two-senses	$(0.145, 0.838), (0.159, 0.840)$	$[3.771, 2.946]$	$[10.003, 3.075]$	331.378	**0.9979**	**0.1460**

following equation:

$$D(c_i) = \sum_{j=1}^{C} \sum_{i=1}^{n} ||x_i^a - v_j||^2. \tag{4.6}$$

In Equation (4.6), x_i^a is the attribute of the i-th object in the dataset. This equation measures the degree to which the computed cluster centroid captures the central position of the patterns in a given cluster. In fact, it shows the appropriateness of the calculated cluster centroid relative to the data structure. Minimum value of

D underlines the efficiency of the clustering technique in recognizing the pattern described by the dataset.

4.4. **Discussion.** The schema of an SCM clustering classifies pattern in a cluster by relying on a criteria $(\alpha, 1-\alpha)$. Computation of cluster centorid utilizes the attributes and membership grades of all patterns in the cluster. However, the quality of the computed centroid is degraded by patterns exhibiting very high degree of fuzziness to a cluster. Consequently, the obtained centroids, no matter how well separated they may be, may not adequately capture the true position of their representatives. In order to determine cluster centroids which maintain the true position of their representatives, two types of cuts are used to (i.e., α-cuts; for upgrading membership values and, fuzziness-cut; for separating objects for which there is sufficient information for making clear decision about their membership to a cluster) partition a fuzzy cluster. We call this approach "two-senses" clustering method.

The procedure for identifying patterns which exhibit very high degree of fuzziness is comprehended from what follows:

First of all, we define the fuzziness set, φ_{c_k}, of c_k as follows:

$$\varphi_{c_k} = \{(x_i, \varphi(x_i)) | \varphi(x_i) = 1 - |2\mu_{ik} - 1|\}.$$

Let $\varphi_{max_{c_k}}$ and $\varphi_{min_{c_k}}$ denote the maximum and minimum degree of fuzziness of the objects in c_k. We search for a pair of optimum fuzziness values,

$$(\varphi^*, \varphi^{**}),$$

where $\varphi^* = \varphi_{max_{c_k}} - \varphi^{**}$ and $\varphi^* \in [\varphi_{min_{c_k}}, \frac{\varphi_{max_{c_k}} - \varphi_{min_{c_k}}}{2}]$.
This optimum threshold is calculated as:

$$\varphi^*(opt) = min_{\varphi^*} |u + v + w - z|, \tag{4.7}$$

where

$$u = \sum_{\varphi_{c_k}(x) < \varphi^*} \varphi(x),$$

$$v = \sum_{\varphi_{c_k}(x) > \varphi^{**}} \varphi(x),$$

$$w = \sum_{\varphi' \leq \varphi_{c_k}(x) \leq \varphi^{**}} \varphi(x),$$

$$z = Card(\{x | \varphi^* \leq \varphi_{c_k}(x) \leq \varphi^{**}\}).$$

With the aid of the aforesaid optimized fuzziness thresholds, the following decision rules are used to refine a three-way cluster:

Accept fuzziness (AF):

An object with lower fuzziness is accepted as belonging to a cluster, if its membership grade to the fuzziness set, φ_{c_k}, is less than φ^*. That is,

$$AF = \{x | \varphi_{c_k}(x) < \varphi^*\}.$$

Undecided fuzziness (UF):

We postpone the decision as to the inclusion or exclusion of an object to a cluster, if its membership grade to the fuzziness set φ_{c_k} is between φ^* and φ^{**}. That is,

$$UF = \{x|\varphi^* \leq \varphi_{c_k}(x) \leq \varphi^{**}\}.$$

Reject fuzziness (RF):

An object with higher fuzziness is rejected as belonging to a cluster, if its membership grade to the fuzziness set φ_{c_k} is greater than φ^{**}. That is,

$$RF = \{x|\varphi_{c_k}(x) > \varphi^{**}\}.$$

In view of the aforesaid, any pattern in RF will not be used during computation of cluster centroids. The thresholds and centroids obtained by using this idea are shown in the last row of Table 6.

Let us discuss the results presented in Table 6. In the following figures, the patterns represented in green, blue and red circles are core, fringe and outlier (excluded) to cluster 1, respectively. Patterns in red are outliers and, are evaluated for possible consideration in cluster 2. The centroids are depicted with purple x marks.
The centroids computed by deploying Equation (2.8) into the schema on D_{32} yield the following plots:

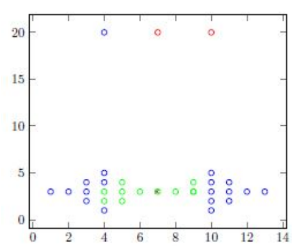

FIGURE 6. SCM clustering of D_{32} with Equation (2.8): Cluster 1

Cluster 2 yields Figure 7:

When the formulation in Equation (2.10) (Tahayori *et al.*, 2013) is deployed to cluster patterns in D_{32}, we obtain the next to figures:
Cluster 2 yields Figure 9:

When the formulation in Equation (2.11) (Deng and Yao, 2014) is deployed in the schema on D_{32}, we obtain the next two figures:

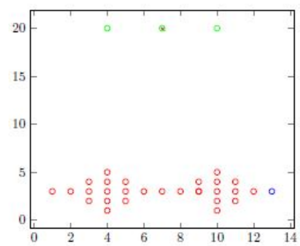

FIGURE 7. SCM clustering of D_{32} with Equation (2.8): Cluster 2

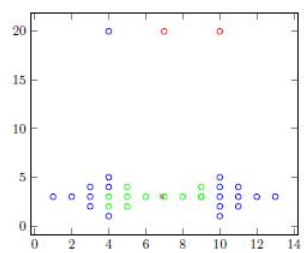

FIGURE 8. SCM clustering of D_{32} with Equation (2.10): Cluster 1

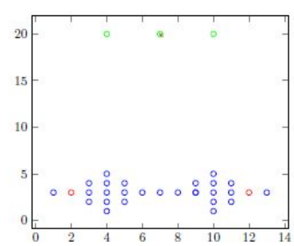

FIGURE 9. SCM clustering of D_{32} with Equation (2.10): Cluster 2

Cluster 2 yields Figure 2.11:
The clustering induced by deploying Equation (2.14) (Ibrahim and William-West, 2019) yields the next two figures:

Cluster 2 yields Figure 13:

The clustering determined by deploying Equation (2.16) (William-West *et al.*, 2019) yields the next two figures:

Cluster 2 yields Figure 15:

The clustering obtained from Two-senses approximation is shown in the last two

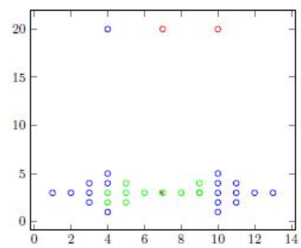

FIGURE 10. SCM clustering of D_{32} with Equation (2.11): Cluster 1

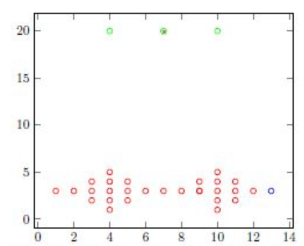

FIGURE 11. SCM clustering of D_{32} with Equation (2.11): Cluster 2

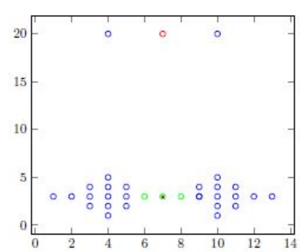

FIGURE 12. SCM clustering of D_{32} with Equation (2.14): Cluster 1

figures:

The clustering of patterns in cluster 2 is shown in the next figure:

Noticeably from Figures 6–17, only the two-senses approach show better performance in terms of fast learning of the structural pattern described by D_{32}. So, its clustering effectively align with the intuitive idea of the topology of D_{32}. The other methods produce two clusters: classifying noisy patterns into one cluster and the patterns directly below them into another cluster. This is merely a structural assumption, and does not reflect our intuition about D_{32}.

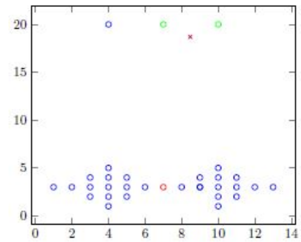

FIGURE 13. SCM clustering of D_{32} with Equation (2.14): Cluster 2

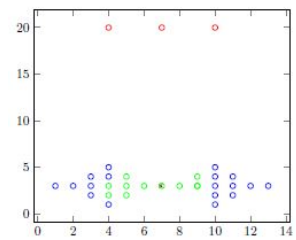

FIGURE 14. SCM clustering of D_{32} with Equation (2.16): Cluster 1

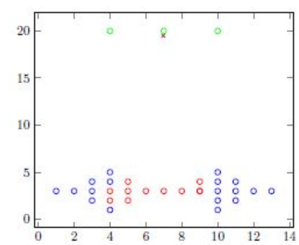

FIGURE 15. SCM clustering of D_{32} with Equation (2.16): Cluster 2

Further, it can be observed from the cluster centroids (i.e., the patterns with purple x marks) that the two-senses approach determines a central position of their representatives, which adequately conform with the intuitive idea, as portrayed by the patterns. Hence, the method gives a more appealing representation of D_{32}. This is supported by the minimum value of D in Table 6.

From the insightful gain obtained from clustering D_{32} with various shadowed set approximation techniques, the following conclusions could be made:

(1) without isolating patterns which exhibit very high degree of uncertainty to a cluster when the centroids are computed, an SCM algorithm may not effectively minimize the influence of uncertain patterns. Also, it may not ensure that we have centroids which do not drift from the true position of

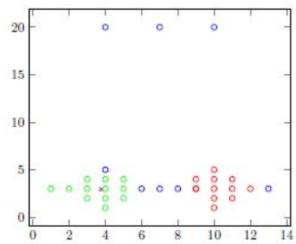

FIGURE 16. SCM clustering of D_{32} with Two-senses approach: Cluster 1

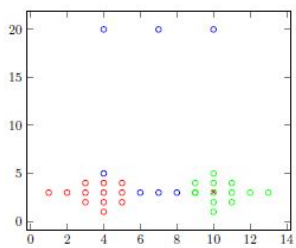

FIGURE 17. SCM clustering of D_{32} with Two-senses approach: Cluster 2

their representatives. Consequently, the underlying structure in the dataset may be perceived differently (i.e., as in D_{32}, uncertain patterns may be grouped together, while every other pattern may be viewed as been *related* to the same group),

(2) shadowed set approximation based on a principle of uncertainty balance plays a key role in data mining. However, being primarily couched in uncertainty and information preservation, on its own, may not effectively deal with pattern recognition task, especially when the structure of the dataset is such that the within-cluster distance is bigger than the between-cluster distance. In fact, it attempts to preserve the original uncertainty in the dataset and thereby, producing the largest (respectively smallest) fringe region according to the magnitude of the uncertainty encountered in each cluster. For the case of having large fringe size, this may lead to a clustering with patterns concentrating at the boundary and, may misguide formation of core region.

(3) the two-senses approach learns the structural configuration of the dataset faster. This is evident from Figures 16 and 17. However, as shown in Figures 16 and 17, a few data objects are apparently partitioned into the wrong regions, namely patterns $x_{14}, x_{16}, x_{30}, x_{31}$ and x_{32},

(4) The ratio between compactness and separation of clusters, numerically expressed as SF, as well as H-index, evaluates the performance of two-senses shadowed C-means clustering method for D_{32} to be outstanding. This

method exhibits the capability to effectively deal with pattern recognition task that cannot be handled by uncertainty balance method.

5. A need for shadowed sets with higher number of approximation regions

Remark 4.1 motivates the introduction of $n(>3)$-way approximation of fuzzy sets via shadowed sets.
In an attempt to increase the number of approximation regions in shadowed sets, the first point of departure is to consider four-region shadowed sets. This may be in line with Lukasiewicz n-valued logic, when $n = 4$. However, two problems may arise:

(1) how to define the complement of the fourth region. Suppose we can assign the following membership values to the four regions: $0, 0.25, 0.75$ and 1. This would lead to a new type of problem:
(2) a shadowed set which lacks well-defined boundary region.

We note that introduction of four-region shadowed sets may lead to introduction of five-region shadowed set. This is supported by the profound findings on truth-value judgment involving vagueness from a study by Zehr [35].

The cognitive basis for five-region shadowed sets, S_5, can be explained from the studies on techniques of chunking (i.e., a method of transforming information into manageable size) in [3, 15].

Formally, by exploiting the four threshold values α, β, γ and ρ, we partition a fuzzy set into five-regions:

(1) $Cor(S_5) = \{x \in F : \mu_F(x) \in [\rho, 1]\}$,
(2) $SCor(S_5) = \{x \in F : \mu_F(x) \in (\gamma, \rho)\}$,
(3) $Shd(S_5) = \{x \in F : \mu_F(x) \in [\beta, \gamma]\}$,
(4) $SRed(S_5) = \{x \in F : \mu_F(x) \in (\alpha, \beta)\}$,
(5) $Red(S_5) = \{x \in F : \mu_F(x) \in [0, \alpha]\}$.

The threshold values are such that $0 < \alpha \leq \beta$, $\alpha + \beta = \frac{1}{2}$, $\alpha + \rho = 1$, $\beta + \gamma = 1$ and $\gamma \leq \rho$. Also, 1- 5 represent the core (i.e., highlighting clearly true instances), semi-core (i.e., highlighting *possibly* true instances), shadow (i.e., highlighting borderline instances), semi-reduced (i.e., highlighting *possibly* false instances) and reduced area (i.e., highlighting clearly false instances). The semi-core and semi-reduced zones consist of objects for which there is no sufficient information as to their full inclusion into the core and exclusion from the set, respectively. Essentially, these zones are introduced to control the effects of taking wrong/quick decision.

Let us discuss key aspect of induction of S_5. Consider the fuzzy set described in Figure 18 and its resulting fuzziness set in Figure 20. Construction of S_5 requires that Figure 18 give rise to Figure 19. That is, the approximation procedure (as summarized in Figure 19) involves elevating elements' membership values to $0.25, 0.5, 0.75$ and 1 if they are greater than α, β, γ and ρ, respectively. Similarly, we reduce the membership values to $0, 0.25, 0.5$ and 0.75 if they are less than α, β, γ and ρ.

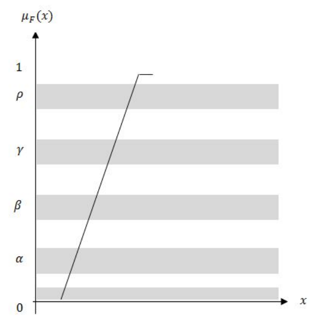

FIGURE 18. Graph of fuzzy set.

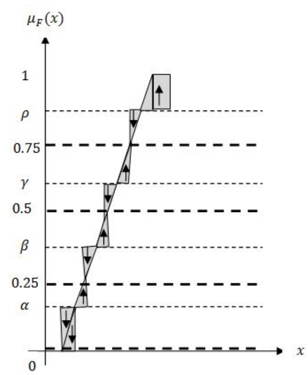

FIGURE 19. Five-region approximation of fuzzy set

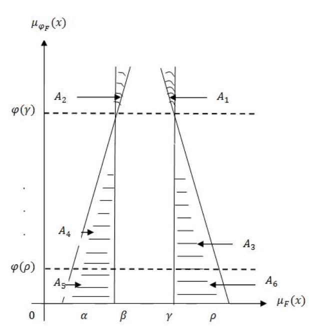

FIGURE 20. Fuzziness graph.

The following property facilitate discussion of uncertainty balance in five-region shadowed sets. Determination of the optimum thresholds is deduced from the same property.

Theorem 5.1. *Let (α, ρ) and (β, γ) be any pair of feasible thresholds for constructing S_5 from F. Then $\forall x \in F$, we have*

(1) *$x \in (Cor(S_5) \cup Red(S_5))$ if and only if $\mu_{\varphi_F}(x) \leq \varphi(\rho)$,*
(2) *$x \in (SCor(S_5) \cup SRed(S_5))$ if and only if $\varphi(\rho) < \mu_{\varphi_F}(x) < \varphi(\gamma)$,*
(3) *$x \in Shd(S_5)$ if and only if $\mu_{\varphi_F}(x) \geq \varphi(\gamma)$*

Proof. Recall that $\varphi(\rho) = \varphi(\alpha)$ and $\varphi(\beta) = \varphi(\gamma)$.

(1) Suppose $x \in (Cor(S_5) \cup Red(S_5))$. Then $\mu_F(x) \in (0, \alpha]$ or $\mu_F(x) \in [\rho, 1]$. Since φ is monotonic increasing in $(0, \alpha]$ and monotonic decreasing in $[\rho, 1]$ and, attains its maximum value for both interval at α and ρ,n $\mu_{\varphi_F}(x) \leq \varphi(\rho)$.

Conversely, suppose $\mu_{\varphi_F}(x) \leq \varphi(\rho)$. Then $\mu_F(x) \in (0, \alpha]$ or $\mu_F(x) \in [\rho, 1]$. Thus x must belong to either $Cor(S_5)$ or $Red(S_5)$. So $x \in (Cor(S_5) \cup Red(S_5))$.

(2) Suppose $x \in (SCor(S_5) \cup SRed(S_5))$. Then $\mu_F(x) \in (\alpha, \beta)$ or $\mu_F(x) \in (\gamma, \rho)$. From the monotonic property of φ, we can find a maximum value in (α, β) and (γ, ρ) at β and ρ. Since $\varphi(\beta) = \varphi(\gamma)$ and $\varphi(\rho) < \varphi(\gamma)$, $\varphi(\rho) < \mu_{\varphi_F}(x) < \varphi(\gamma)$. We note that φ attains its minimum value in (α, β) and (γ, ρ) at α and ρ.

Conversely, suppose $\varphi(\rho) < \mu_{\varphi_F}(x) < \varphi(\gamma)$. Then $\mu_F(x) \in (\alpha, \beta)$ or $\mu_F(x) \in (\gamma, \rho)$. Thus x must belong to either $SCor(S_5)$ or $SRed(S_5)$. So $x \in (SCor(S_5) \cup SRed(S_5))$.

(3) Suppose $x \in Shd(S_5)$. Then $\mu_F(x) \in [\beta, \gamma]$. Now φ attains its maximum value at $\frac{1}{2} \in [\beta, \gamma]$. The minimum value of φ in $[\beta, \gamma]$ is at β and γ. Thus $\forall x \in [\beta, \gamma]$, $\mu_{\varphi_F}(x) \geq \varphi(\gamma)$.

Conversely, suppose $\mu_{\varphi_F}(x) \geq \varphi(\gamma)$. Then $x \in [\beta, \gamma]$. Thus we must have that x belongs to $Shd(S_5)$.

This completes the proof. □

Remark 5.2. Theorem 5.1 infers that in five-region shadowed sets, the optimum threshold values to be found depends on two fuzziness-cuts $\varphi(\rho)$ and $\varphi(\gamma)$ which cuts the entire fuzziness set of F into five regions.

Note that S_5 redistributes the fuzziness in F into $Shd(S_5)$, $SCor(S_5)$ and $SRed(S_5)$.

The optimum pairs of thresholds, (α, ρ) and (β, γ), to found can be comprehended from the following theoretical analysis:

$$\varphi(F) = \varphi(Shd(S_5)) + \varphi(SCor(S_5)) + \varphi(SRed(S_5)). \tag{5.1}$$

From Theorem 5.1, we have the following equation:

$$A_1 + A_2 + A_3 + A_4 + A_5 + A_6 = \varphi(F). \tag{5.2}$$

That is,

$$\sum_{\mu_{\varphi_F}(x) \leq \varphi(\rho)} \mu_{\varphi_F}(x) + \sum_{\varphi(\rho) < \mu_{\varphi_F}(x) < \varphi(\gamma)} \mu_{\varphi_F}(x) + \sum_{\mu_{\varphi_F}(x) \geq \varphi(\gamma)} \mu_{\varphi_F}(x) = \varphi(F). \tag{5.3}$$

This expands to the next equation:

$$\sum_{x \in Cor(S_5)} \mu_{\varphi_F}(x) + \sum_{x \in Red(S_5)} \mu_{\varphi_F}(x) + \sum_{x \in SCor(S_5)} \mu_{\varphi_F}(x) + \sum_{x \in SRed(S_5)} \mu_{\varphi_F}(x) + \sum_{x \in Shd(S_5)} \mu_{\varphi_F}(x) = \varphi(F). \tag{5.4}$$

By combining Equations (4.2) and (4.5), we have the next equation:

$$w_1 + w_2 + w_3 + w_4 + w_5 = v_1 + v_2 + v_3, \tag{5.5}$$

where

$w_1 = \sum_{x\in Cor(S_5)} \mu_{\varphi_F}(x)$,
$w_2 = \sum_{x\in Red(S_5)} \mu_{\varphi_F}(x)$,
$w_3 = \sum_{x\in SCor(S_5)} \mu_{\varphi_F}(x)$,
$w_4 = \sum_{x\in SRed(S_5)} \mu_{\varphi_F}(x)$,
$w_5 = \sum_{x\in Shd(S_5)} \mu_{\varphi_F}(x)$,
$v_1 = \varphi(Shd(S_5))$,
$v_2 = \varphi(SCor(S_5))$,
$v_3 = \varphi(SRed(S_5))$.

Recall that whenever $\mu_F(x) = 0.75$ or $\mu_F(x) = 0.25$, we have $\mu_{\varphi_F}(x) = \mu_{\varphi_F}(x) = \frac{1}{2}$. Also, when $\mu_F(x) = 0.5$, we have $\mu_{\varphi_F}(x) = 1$. Hence, we have the following relationship:

$$\varphi(Shd(S_5)) = Card(Shd(S_5)), \tag{5.6}$$

$$\varphi(SCor(S_5)) = \frac{1}{2}Card(SCor(S_5)), \tag{5.7}$$

$$\varphi(SRed(S_5)) = \frac{1}{2}Card(SRed(S_5)). \tag{5.8}$$

Putting Equations (5.6)–(5.8) in the right hand side of Equation (5.5), the optimum threshold is computed by minimizing the next equation:

$$Q(\alpha) = |w_1 + w_2 + w_3 + w_4 + w_5 - (v_1' + v_2' + v_3')|, \tag{5.9}$$

where

$v_1' = Card(Shd(S_5))$,
$v_2' = \frac{1}{2}Card(SCor(S_5))$,
$v_3' = \frac{1}{2}Card(SRed(S_5))$.

Remark 5.3. Five-region shadowed sets provide additional decision alternatives for approximation of F. From the examples considered on F_1 and F_2 of Tables 1 and 3, their efficiency in fulfilling a principle of uncertainty balance and error minimization are shown in the next table.

Table 7. Optimized thresholds for discrete fuzzy set F_2

Fuzzy set	Optimized threshold $(\alpha,\rho),(\beta,\gamma)$	$\varphi(S_5)$	$\varphi(F_i)$	Discrepancy	Error
F_1	$(0.15, 0.85), (0.35, 0.65)$	16	15.88	0.12	1.93
F_2	$(0.10, 0.90), (0.40, 0.60)$	11.5	11.32	0.18	1.49

A comparison of the discrepancy and error in three-region and five-region shadowed set approximation of fuzzy sets (see column 5 and 6 of Tables 2, 4 and 5) underlines the advantage of five-region shadowed sets over three-region shadowed sets.

It is observed that the higher the (odd) number of approximation regions of a shadowed set the smaller the discrepancy and approximation error. We conclude the analysis with the following conjecture, which has been checked from several randomly constructed fuzzy sets.

Conjecture 1

For two shadowed sets, S_3 and S_5, approximation of fuzzy set F, the following inequalities hold:
(1) $|\varphi(F) - \varphi(S_5)| \leq |\varphi(F) - \varphi(S_3)|$,
(2) $E(S_5) \leq E(S_3)$,
where S_i denotes an i-region shadowed set and $E(A)$ is the error in approximating F into A.

5.1. **Five-way approximation of sets as a model for thinking in fives.** Granular computing(GrC) [24, 21] is a theoretical and computational platform for modeling human-data-interaction. It facilitates construction of information granules arising from abstraction, generalization, approximation, aggregation or other forms of derivation of knowledge from data or information [28]. As fuzzy sets and their associated n-way approximations are computational tools in GrC [30], five-way approximation of sets, as well, facilitate construction of information granules. It is a form of information granulation that allow us differentiate *necessary and sufficient*, *sufficient*, and uncertain from *seemingly unnecessary* and *unnecessary* detail in a given data system.

From a broad perspective, five-way approximation introduces the notion of five-way decision. It can be used to interpret five types of decision actions obtained from the five regions of the ensuing approximation of fuzzy sets. Immediately, one can easily link its methodology to a concept of *thinking in fives* (an extended form of thinking in threes [32].

More explicitly, when handling data sets, human cognitive methods of processing deploys a concept in GrC known as *zoom-out* (mapping information granules from finer level to coarser level) and *zoom-in* (mapping information granules from coarser level to finer level) operations. For example, a day may be zoom-in as morning, afternoon and night. Alternatively, one can zoom-out a day as morning, mid-day, afternoon, evening and night. In this context, these two operations define thinking in threes and fives, respectively.

Concluding, it is well-known that given the same knowledge source, distinct views may induce distinct granular knowledge structures. So, a fusion of three-way and five-way approximation models in GrC methodologies may provide practical strategies on how to use granular knowledge structures. Moreover, this notion may help us select and switch between levels and views in a given granular space. Therefore, both models should rather be harmonized than compete against each other. Hence, one can take either five-way approximation as a literal sense of *five*, as being taken in the present study, or as a figurative sense of *five* as some zoom-out operation in three-way decision models.

6. Conclusions

This paper explicates a concept of uncertainty balance in shadowed sets and provides some of its related properties. It identified two senses of uncertainty balance and underlines their usefulness when combined in approximating fuzzy clusters. The study revealed that uncertainty balance is theoretically meaningful for retention of the fuzziness encountered in a dataset F. However, the related information in F can be adequately preserved by exploiting the two-senses of uncertainty balance. This is exemplified with pattern recognition task involving shadowed C-means clustering.

Further, five-region shadowed sets has been introduced in this paper. An imperative for its introduction is to effectively deal with the problem of uncertainty balance, and in the same vein, minimize the approximation error associated with three-region shadowed sets. A blind spot of five-region shadowed set is that its objective function $Q(\alpha)$ is nonconvex. Hence, more than one optimal solution for $Q(\alpha)$ exists. Therefore, it is important to investigate the conditions that can be imposed on $Q(\alpha)$ for its solution to be unique.

Acknowledgements. The authors wish to thank the Editor-in-Chief and the anonymous reviewers for their technical comments towards the improvement of this paper.

References

[1] H. A. Arnaldo and B. C. R. Bedregal, A new way to obtain the initial centroids in fuzzy C-means algorithm 2013 2nd Workshop-School on Theoretical Computer Science (2013) 139–144.
[2] G. Cattaneo G and D. Ciucci, Shadowed sets and related algebraic structures, Fundamenta Informaticae 55 (2003) 255–284.
[3] N. Cowan, The magic number 4 in short-term memory: a reconsideration of mental storage capacity, Behav Brain Sci. 24 (2000) 87–185.
[4] A. De Luca and S. Termini, A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, Information and Control 20 (1972) 301–312.
[5] X. F. Deng and Y. Y. Yao, Mean-value-based decision-theoretic shadowed sets, Proceedings of 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS) (2013) 1382–1387.
[6] X. F. Dengand Y. Y. Yao, Decision-theoretic three-way approximations of fuzzy sets, Inform. Sci. 279 (2014) 702–715.
[7] D. L. Dubes and D. W. Boulding, Cluster separation measure, IEEE Transaction on Pattern Analysis and Machine Intelligence 1 (1979) 224–227.
[8] P. Grzegorzewski, Fuzzy number approximation via shadowed sets, Inform. Sci. 225 (2013) 35–46.
[9] A. M. Ibrahim. and T. O. William-West, Induction of shadowed sets from fuzzy sets, Granular Computing 4 (2019) 27–38.
[10] S. C. Kleene, Introduction to Meta-mathematics, D. van Nostrand, NewYork 1952.
[11] G.J. Klir, Where do we stand on measures of uncertainty,ambiguity , fuzziness and the like , Fuzzy Sets and Systems 24 (1987) 141–160.
[12] G. J. Klir, A principle of uncertainty and information invariance, International Journal of General Systems 17 (2) (1990) 249–275.

[13] G. J. Klir, U. H. St. Clair and B. Yuan, Fuzzy Set Theory: Foundations and Applications, Englewood Cliffs, NJ: Prentice Hall 1997.

[14] S. Knopfmacher, On measures of fuzziness, Journal of Mathematical Analysis and Applications 49 (1975) 529–534.

[15] G. A. Miller, The magical number seven, plus or minus two: some limits on our capacity for processing information, Pyschol Re. 63 (1956) 81–97.

[16] S. Mitra and P. P. Kundu, Satelite image segmentation with shadowed C-means, Pattern Recognition 181 (2011) 3601–3613.

[17] S. Mitra, W. Pedrycz and B. Barman, Shadowed C-Means: Interpreting Fuzzy and Rough Clustering, Pattern Recognition 43 (2010) 1282–1291.

[18] H. T. Nguyen, W. Pedrycz and V. Kreinovich, On approximation of fuzzy sets by crisp sets: from continuous control-oriented defuzzification to discrete decision making, In Proceedings of international Conference on Intelligent Technologies (2000) 254–260.

[19] K. Pakhira, S. Bandiyopadhyay and U. Maulik, Validity index for crisp and fuzzy clusters, Pattern Recognition 43 (2010) 1281–1291.

[20] W. Pedrycz, Shadowed sets: Representing and processing fuzzy sets, IEEE Transactions on System, Man and Cybernetics 28 (1998) 103–109.

[21] W. Pedrycz, Granular computing with shadowed sets, in: D. Slezak, G.Y. Wang, M. Szczuka, I. Dntsch, Y.Y. Yao (Eds.), RSFDGrC 2005: LNCS(LNAI) Springer, Heidelberg 3641 (2005) 23–32.

[22] W. Pedrycz, Interpretation of Clusters in the Framework of Shadowed sets, Pattern Recognition Letters 26 (2005) 2439–2449.

[23] W. Pedrycz, Granular Computing for Data Analytics: A Manifesto of Human-Centric Computing, IEE/CAA Journal of Automatica Sinica 5 (6) (2018) 1025–1034.

[24] W. Pedrycz and G. Vukovich, Granular computing with shadowed sets, International Journal of Intelligent Systems 17 (2002) 173–197.

[25] S. Saitta, B. Raphaely and I. F. C. Smith, A comprehensive validity index for clustering, Intelligent Data Analysis 12 (2008) 529–548.

[26] X. Xie and G. Beni, Validity measure for fuzzy clustering, IEEE Transaction on Pattern Analysis and Machine Intelligence 13 (1991) 841–847.

[27] H. Tahayori, A. Sadeghian and W. Pedrycz, Induction of shadowed sets based on the gradual grade of fuzziness, IEEE Transactions on Fuzzy Systems 21 (2013) 937–949.

[28] G. Wilke and E. Portmann, Granular computing as a basis of human data interaction: a cognitive cities use case, Granular Computing, DOI 10.1007/s41066-016-0015-4.

[29] T. O. William-West, A. M. Ibrahim and A. F. D. Kana, Shadowed sets approximation of fuzzy sets based on nearest quota of fuzziness, Ann. Fuzzy Math. Inform. 17 (2) (2019) 133–145.

[30] W. Xu, J. Pang and S. Luo, International Journal of Approximate Reasoning, 55 (2014) 853–866.

[31] Y. Y. Yao, An outline of a theory of three-way decisions, In: Yao, J., Yang, Y., Slowinski, R., Greco, S., Li, H., Mitra, S., Polkowski, L. (eds.) RSCTC 2012. LNCS (LNAI) 7413 (2012) 1–17. Springer, Heidelberg.

[32] Y.Y. Yao, Three-way decision and granular computing, Int. J. Approx. Reas 107 (2018) 107–123.

[33] Y. Yao, S. Wang and X. Deng, Constructing Shadowed Sets and Three-way Approximations of Fuzzy Sets, Inform. Sci. 413 (2017) 132–153.

[34] A. Zadeh. Fuzzy Sets, Information and Control 8 (1965) 338–353.

[35] J. Zehr, ST5: A 5-Valued Logic for Truth-Value Judgments Involving Vagueness and Presuppositions, M. Colinet et al. (Eds.): ESSLLI 2012/2013, LNCS 8607 (2014) 247–265.

[36] Y. Zhang and J. T. Yao, Game theoretic approach to shadowed sets: A three-way tradeoff perspective, Information Sciences (2018), doi:10.1016/j.ins.2018.07.058.

[37] J. Zhou, D. Miao, C. Gao, Z. Lai and X. Yue, Constrained three-way approximations of fuzzy sets: From the perspective of minimal distance, Inform. Sci. 502 (2019) 247–267.

[38] J. Zhou, W. Pedrycz and D. Miao, Shadowed sets in the characterization of rough-fuzzy clustering, Pattern Recognition 44 (2011) 1738–1749.

[39] J. Zhou, Z. Lai, C. Gao and W. Wong, Rough-fuzzy clustering based on two-stage three approximations, IEEE Access 6 (2018) 27541–27554.

[40] J. Zhou1, C. Gao, W. Pedrycz and Z. Lai, Constrained shadowed sets and fast optimization algorithm, Int. J. Intell. Syst. DOI: 10.1002/int.22170.

M. A. Ibrahim (amibrahim@abu.edu.ng)
Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria

T. O. William-West (westtamuno@gmail.com)
Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria

A. F. D. Kana (donfackkana@gmail.com)
Department of Computer Science, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria

D. Singh (mathdss@yahoo.com)
Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria