A New Criterion for Soft Set Based Decision Making Problems under Incomplete Information

José Carlos R. Alcantud [1,2], Gustavo Santos-García [2]

[1] *BORDA Research Unit*

[2] *Multidisciplinary Institute of Enterprise (IME),*
University of Salamanca, E-37007 Salamanca, Spain

E-mail: $\{jcr, santos\}$*@usal.es*

Abstract

We put forward a completely redesigned approach to soft set based decision making problems under incomplete information. An algorithmic solution is proposed and compared with previous approaches in the literature. The computational performance of our algorithm is critically analyzed by an experimental study.

Keywords: Soft set; Incomplete soft set; Choice value; Decision making.

1. Introduction

In this paper we revisit the soft set based decision making problem under incomplete information as approached by Han *et al.* [16], Qin *et al.* [23], and Zou and Xiao [38].

Many real life problems require the use of imprecise or uncertain data (cf., Kahraman *et al.* [17]). Their analysis must involve the application of mathematical principles capable of capturing these features. Fuzzy set theory meant a paradigmatic change in mathematics which allows partial membership. The publication of Zadeh's seminal article [33] has triggered a vast literature on fuzzy sets and their applications, which includes a number of successful generalizations.

Of these variations we are especially interested in the application of soft sets theory and their extensions to decision making problems (cf., Molodtsov [22] for a definition and arguments about

its applicability to several fields). Further relevant references include Aktaş and Çağman [1], Alcantud [3], Ali [7,8], Ali *et al.* [9], and Maji *et al.* [21]. Ali *et al.* [10] define lattice ordered soft sets for situations where some order [5,6] exists among the elements of the parameter set. Qin *et al.* [25] combine interval sets and soft sets, and Zhang [37] studies interval soft sets. Maji, Biswas and Roy [19] introduce fuzzy soft sets (see also Ali and Shabir [11] for techniques that permit to study logic connectives for soft sets and fuzzy soft sets, and Alcantud [2] for a recent decision-making procedure based on fuzzy soft sets). Relatedly, Shao and Qin [29] define fuzzy soft lattices and discuss their structure. Wang, Li and Chen [31] introduce hesitant fuzzy soft sets. As mentioned above, Han *et al.* [16], Qin *et al.* [23], and Zou and Xiao [38] are concerned with incomplete soft sets. Rodríguez *et al.* [26,27] are recent surveys on hesitant fuzzy sets (cf., Torra [30]). There are also interesting hybrid models in recent literature, e.g., rough soft hemirings (a new rough

set theory which is an extended notion of a rough hemiring and a soft hemiring) in Zhan *et al.* [34,35], or Z-soft rough fuzzy ideals of hemirings (cf., Zhan and Zhu [36]).

Maji, Biswas, and Roy [20] pioneered soft set based decision making. They established the criterion that an object can be selected if it maximizes the choice value of the problem. Relatedly, Zou and Xiao [38] argued that in the process of collecting data there may be unknown, missing or inexistent data. Therefore, standard soft sets under incomplete information must be taken into account, which calls for the inspection of incomplete soft sets. Afterwards Han *et al.* [16] and even Qin *et al.* [23] present other interesting approaches to incomplete soft set based decision making. However, in the frequent cases in which there is perfect uncertainty about the real value of the missing data, it seems appropriate to proceed with a decision making procedure that avoids estimations. This is the purpose of the present contribution.

Our paper contributes to decision making studies in the context of incomplete soft sets from an altogether different perspective. We propose to go through all the filled tables that can be produced from the original incomplete table. All cases are evaluated according to their respective choice values, and finally the alternatives are ranked by the proportion of tables in which they receive the highest choice value.

Our proposal is justified by a classical Laplacian argument from probability theory. Under Laplace's principle of indifference, due to our complete ignorance we are entitled to assume that all tables in which the $*$'s can be replaced with either 0 or 1 are equiprobable. It is sensible to compute the objects which should be selected according to soft-set based decision making in each of these cases in order to make a decision for any object that is selected in the highest proportion of cases.

The use of probabilities in decision making under incomplete information in soft sets already appears in Zou and Xiao's [38] seminal article as a tool that permits to attach weights to the possible choice values of the options. Nevertheless, we provide examples which prove that our solution is indeed different from theirs and other proposed approaches in the existing literature. Consequently, pretty simple situations confirm that making decisions according to our methodology produces decisions that do not necessarily coincide with solutions in the literature. We discuss some issues of efficiency of our Algorithm by means of a computational experiment.

This paper is organized as follows: Section 2 retrieves some terminology and definitions. We define a notion of domination of alternatives that is used in our decision making algorithm. Section 3 contains a review of the solutions proposed by earlier investigations as well as our proposal and an illustrative example. Then we compare our algorithm with previous solutions in order to ensure its novelty. In subsection 3.5 the computational features of our proposal are examined, too. Finally, our conclusions are in Section 4.

2. Definitions: Soft Sets and Incomplete Soft Sets

We adopt the usual description and terminology for soft sets and their extensions: U denotes a universe of objects and E denotes a universal set of parameters.

Definition 1. [Molodtsov [22]] A pair (F,A) is a *soft set* over U when $A \subseteq E$ and $F : A \longrightarrow \mathscr{P}(U)$, where $\mathscr{P}(U)$ denotes the set of all subsets of U.

A soft set over U is regarded as a parameterized family of subsets of the universe U, the set A being the parameters. For each parameter $e \in A$, $F(e)$ is the subset of U approximated by e or the set of e-approximate elements of the soft set. Many scholars have performed formal investigations of this and related concepts. For example, Maji, Bismas and Roy [21] develop this notion and define among other concepts: soft subsets and supersets, soft equalities, intersections and unions of soft sets, *et cetera*. Furthermore, Feng and Li [14] give a systematic study on several types of soft subsets and various soft equal relations induced by them. Concerning (pure) soft set based decision making, we refer the reader to Çağman and Enginoğlu [12], Maji, Biswas and Roy [20], and Feng *et al.* [15]

In order to model increasingly general situations, Definition 2 below has been subsequently proposed and investigated:

Definition 2. [Han *et al.* [16]] A pair (F,A) is an *incomplete soft set* over U when $A \subseteq E$ and $F : A \longrightarrow \{0, 1, *\}^U$, where $\{0, 1, *\}^U$ is the set of all functions from U to $\{0, 1, *\}$.

Obviously, every soft set can be considered an incomplete soft set. The $*$ symbol in Definition 2 permits to capture "lack of information": when $F(e)(u) = *$ we interpret that u belongs to the subset of U approximated by e is unknown. As in the case of soft sets, when $F(e)(u) = 1$ (resp., $F(e)(u) = 0$), we interpret that u belongs (resp., does not belong) to the subset of U approximated by e.

As is well known since the early antecedent in Yao [32], when both U and A are finite (as in the application references mentioned above), soft sets and incomplete soft sets can be represented either by matrices or in tabular form. Rows are attached with objects in U, and columns are attached with parameters in A. In the case of a soft set, these representations are binary (i.e., all cells are either 0 or 1).

3. The Problem: Preliminaries and New Proposals of Solution

In this section we tackle the problem of applying incomplete soft sets in decision making practice.

3.1. Preliminaries and Antecedents

Concerning (pure) soft set based decision making, the fundamental reference is Maji, Biswas and Roy [20]. When a soft set (F,A) is represented in matrix form through the matrix $(t_{ij})_{k \times l}$, where k and l are the cardinals of U and A respectively, then the *choice value* of an object $h_i \in U$ is $c_i = \sum_j t_{ij}$. A suitable choice is made when the selected object h_k verifies $c_k = \max_i c_i$. In other words, objects that maximize the choice value are satisfactory outcomes of this decision making problem.

Concerning incomplete soft set based decision making, the most successful approaches are probably Han *et al.* [16], Qin *et al.* [23], and Zou and Xiao [38]. Let us survey their ideas.

Zou and Xiao [38] initiated the analysis of soft sets and fuzzy soft sets under incomplete information. In the first, standard case they propose to calculate all possible choice values for each object, and then calculate their respective "decision values" d_i by the method of weighted-average. To this purpose the weight of each possible choice value is computed by existing complete information. In particular, they put forward some simple indicators that can eventually be used to prioritize the alternatives, namely, $c_{i(0)}$ (the choice value if all missing data are assumed to be 0), $c_{i(1)}$ (the choice value if all missing data are assumed to be 1) and d_{i-p} (which is easily shown to correspond to $(c_{i(0)} + c_{i(1)})/2$).

Because Zou and Xiao's definition of d_i is complex and difficult to understand, Kong *et al.* [18] have designed a simplified approach to decision making for incomplete soft sets that is equivalent to the weight-average of all possible choice value approach.

Inspired by the data analysis approach in Zou and Xiao [38], Qin *et al.* [23] propose a new way to fill out the missing data in an incomplete soft set. To that purpose, they introduce the relation between parameters. Thus, they prioritize the association between parameters rather than the probability of objects appearing in $F(e_i)$. This way they attach a *completed soft set* with any incomplete soft set. However, the procedure of Qin *et al.* [23] presupposes that there are associations among some of the parameters. Their proposal relies on to Zou and Xiao's method [38]: when an exogenously given threshold is not reached, the data is filled out according to Zou and Xiao's approach. Incidentally, there is no clue as to which threshold is suitable for each problem.

Qin *et al.* hint that their procedure can be used to implement subsequent applications involving incomplete soft sets, but they do not make any explicit statement as to decision making. In fact, these authors criticize the approach for (crisp) soft sets in Zou and Xiao [38] because the missing data are still missing at the end of the process, and "the soft sets cannot be used in other fields but decision making". Nevertheless, it seems appropriate to complement their filling procedure with a prioritization of the objects according to their choice values Q_i, as is stan-

dard in soft-set based decision making. Han *et al.* [16], section 1.3, explain that "this method is good when objects in U are related with each other".

Relatedly, we also mention that Khan *et al.* [28] deal with predicting the performance of handling missing data in incomplete soft sets, but they do not advocate for any decision making mechanism. Therefore Khan *et al.* [28], like Qin *et al.* [23], are not directly concerned with decision making mechanisms.

Finally, Han *et al.* [16] develop and compare several elicitation criterions for decision making of incomplete soft sets which are generated by restricted intersection. Here we do not pursue that avenue.

Table 1 summarizes the definitions of elements that we have explained above.

Index	Description	Source
d_i	Calculate all possible choice values for each object. Then calculate their d_i by method of weighted-average	Zou & Xiao [38] Kong *et al.* [18]
d_{i-p}	$(c_{i(0)} + c_{i(1)})/2$	Zou & Xiao [38]
$c_{i(0)}$	Choice value if missing data are 0	Zou & Xiao [38]
$c_{i(1)}$	Choice value if missing data are 1	Zou & Xiao [38]
Q_i	Choice value of soft set completed by data filling approach	Own ellaboration from Qin *et al.* [23]

Table 1. Indicators that permit to make decisions about soft sets in an environment with incomplete information.

3.2. A Preparatory Step

Here we comment upon some solutions in Zou and Xiao [38]. Let us fix an incomplete soft set.

The researcher can conduct a pre-screening operation before selecting a final option. We calculate the maximum value c_0 of all choice values $c_{j(0)}$ across options u_j. If this value is strictly greater than the choice value $c_{k(1)}$ of an alternative u_k, this alternative can be removed from the initial matrix. The reason is that if all missing data for u_k are assumed to be 1, there is another option i verifying that when all missing data for u_i are assumed to be 0, still option i has a greater choice value than option k. This argument suggests the following novel definition:

Definition 3. Let (F,A) be an incomplete soft set over U. An option i *dominates* an option k when $c_{k(1)} < c_{i(0)}$.

Clearly, if we adhere to any choice value based solution, we can freely discard dominated options. For example, if we use either d_j, $c_{j(0)}$, $c_{j(1)}$ or d_{j-p} as an indicator for any option j, option k cannot maximize the selected indicator when option i dominates it.

This simplification is basically inconsequential in the case of Zou and Xiao's solutions, however we can apply it in other computationally costly algorithms in order to reduce calculations.

3.3. A New Proposal of Solution

We must emphasize that our purpose here is not to fill any missing data but to give advice on which choice should be made when data are missing in the context of soft sets. Due to the criticisms raised above, to that purpose we cannot support the idea that averages, probabilities or any other specific evaluations should be used to fill missing data. Given that generally there is perfect uncertainty about the real value of these absent data, we propose a completely different approach. We evaluate each feasible filled table according to their choice value. Then we order the alternatives by the proportion of tables where they receive the highest choice value.

The intuition for our proposal is as follows. According to Laplace's principle of indifference in probability theory, under complete ignorance we must assume that all tables where the $*$'s are replaced with either 0 or 1 are equiprobable. Hence the best we can do is to compute which objects should be selected according to soft-set based decision making in each of these cases, and then opt for any object that is optimal in the highest proportion of cases with completed information.

In accordance with this idea and our arguments in section 3.2, we endorse the following algorithm for the problems where both U and A are finite:

Algorithm 1. Incomplete Soft Sets Algorithm

---

1. Input the incomplete soft set on k' objects with l parameters in the form of an input table whose cell (i, j) is denoted $t_{ij} \in \{0, 1, *\}$.

2. Calculate $c_0 = \max_{i=1,\ldots,k'} c_{i(0)}$ from the incomplete matrix.

   Remove each row i from the incomplete matrix such that $c_{i(1)} < c_0$. Denote by k the number of remaining objects.

3. In the trimmed $k \times l$ matrix, enumerate the cells with value $*$ as $((i_1, j_1), \ldots, (i_w, j_w))$.

4. For each vector $v = (v_1, \ldots, v_w) \in \{0,1\}^w$, construct a $k \times l$ matrix $C_v = (c_{ij})_{k \times l}$ where

   (a) $c_{ij} = t_{ij}$ if (i,j) is not listed in $((i_1, j_1), \ldots, (i_w, j_w))$.

   (b) $c_{ij} = v_z$ if $(i,j) = (i_z, j_z)$, $z \in \{1, 2, \ldots, w\}$.

5. For each $i = 1, \ldots, k$, let n_i denote the number of vectors $v = (v_1, \ldots, v_w) \in \{0,1\}^w$ for which object i (denoted h_i) maximizes the choice value at C_v. Define $s_i = n_i/2^w$ for each i. For dominated alternatives, we let $s_i = 0$.

6. The decision is any object u_l that maximizes the score computed in Step 5, i.e., any u_l such that $s_l = \max_{i=1,\ldots,k} s_i$.

Observe that because choice values can be repeated at each C_v matrix, one has $\sum_{i=1}^{k} n_i \geqslant 2^w$.

We can reinterpret the Algorithm above as follows. Firstly, we eliminate dominated alternatives at Step 2. Suppose that due to our complete ignorance of the real data, all completed soft sets from the input of the problem are considered equally probable. Suppose that in case of (complete) soft set problems, we opt for any object with the highest choice value. Then a solution for our original problem is any object such that the probability of its being a solution of a randomly selected completed soft set problem (that is to say, $s_i = n_i/2^w$) is maximal.

The following example from real practice illustrates our proposal. Afterwards we use it to explain the fundamentals of our proposal in practical terms.

Example 1. Let $U = \{h_1, h_2, h_3\}$ be a universe of houses. With respect to the set of parameters (or attributes or house characteristics) $E_0 = \{e_1, e_2, e_3, e_4\}$

the following information is known in the form of an incomplete soft set (F_0, E_0):

(a) $h_1 \in F_0(e_1) \cap F_0(e_3)$ and $h_1 \notin F_0(e_4)$, but it is unknown whether $h_1 \in F_0(e_2)$ or not.

(b) $h_2 \in F_0(e_2)$ and $h_2 \notin F_0(e_3) \cup F_0(e_4)$, but it is unknown whether $h_2 \in F_0(e_1)$ or not.

(c) $h_3 \in F_0(e_1) \cap F_0(e_4)$ and $h_3 \notin F_0(e_2) \cup F_0(e_3)$.

(d) $h_4 \notin F_0(e_1) \cup F_0(e_2) \cup F_0(e_4)$, but it is unknown whether $h_4 \in F_0(e_3)$ or not.

Table 2 captures the information defining (F_0, E_0). We observe that houses 1 and 3 dominate house 4, hence h_4 is eliminated and a 3×4 table remains. In such trimmed table we have $w = 2$, and we enumerate the cells with value $*$ as $((1,2),(2,1))$.

	e_1	e_2	e_3	e_4
h_1	1	$*$	1	0
h_2	$*$	1	0	0
h_3	1	0	0	1
h_4	0	0	$*$	0

Table 2. Tabular representation of the incomplete soft set (F_0, E_0) in Example 1.

For each $v \in \{0,1\}^w = \{v_1 = (0,0), v_2 = (0,1), v_3 = (1,0), v_4 = (1,1)\}$ one feasible completed table arises. These four tables are represented in Table 3, together with the choice values of the houses at each table. We observe that h_1 attaches the highest choice value at all these four tables, h_2 attaches the highest choice value at C_{v_2} only, and h_3 attaches the highest choice value exactly at C_{v_1} and C_{v_2}. Therefore we easily obtain the s_1, s_2, s_3 scores as in Step 5.

Table 4 contains this information as well as the indicators by other focal proposals of solution for this problem, as defined in section 3.1. The optimal alternatives for each procedure are represented too. In addition, observe that the fact that houses 1 and 3 dominate house 4 stems from Table 4 by comparing the maximum of column $c_{i(0)}$ —which is attained at 1 and 3— and the values in column $c_{i(1)}$ that are strictly smaller than such maximum.

C_{v_1} matrix						C_{v_2} matrix				
	e_1	e_2	e_3	e_4	c_i		e_1	e_2	e_3	e_4	c_i
h_1	1	0	1	0	2	h_1	1	0	1	0	2
h_2	0	1	0	0	1	h_2	1	1	0	0	2
h_3	1	0	0	1	2	h_3	1	0	0	1	2

C_{v_3} matrix						C_{v_4} matrix				
	e_1	e_2	e_3	e_4	c_i		e_1	e_2	e_3	e_4	c_i
h_1	1	1	1	0	3	h_1	1	1	1	0	3
h_2	0	1	0	0	1	h_2	1	1	0	0	2
h_3	1	0	0	1	2	h_3	1	0	0	1	2

Table 3. The four completed tables for the incomplete soft set (F_0, E_0) according to Step 4 in our Algorithm, with the respective choice values for each alternative.

	s_i	d_i	d_{i-p}	$c_{i(0)}$	$c_{i(1)}$	Q_i
h_1	1.00	2.50	2.50	2	3	3
h_2	0.25	2.00	1.50	1	2	1
h_3	0.50	2.00	2.00	2	2	2
h_4	0	0.33	0.5	0	1	1
Optimal	$\{h_1\}$	$\{h_1\}$	$\{h_1\}$	$\{h_1, h_3\}$	$\{h_1\}$	$\{h_1\}$

Table 4. Solutions for the problem represented by (F_0, E_0) in Example 1 according to various indicators (v., Algorithm 1 and Table 1).

In order to explain how the intuition of our idea applies to this example, we take advantage of the very competent introduction to incomplete soft sets in real practice by Han *et al.* [16] . Maybe the indeterminacy at attribute e_1 is due to the fact that the real estate salesman refuses to show the required certificate for the second house. Similarly, if the second attribute concerns "convenient underground facilities" it may happen that it is not yet known whether a subway will be built nearby the first house; one needs to consult the municipal construction planning. But if all the possibilities are accounted for, ultimately one of the four tables represented in Table 3 contains the complete information that is needed to make the decision. Since we do not know which one will be correct, we must assume that these tables are equiprobable according to Laplace's principle of indifference. It is then sensible to compute which objects should be selected according to soft-set based decision making in each of these cases, and then select any object that is optimal in most cases.

3.4. A Comparison with Existing Solutions in the Literature

The primary objective of this subsection is to prove that our proposal is different from any of the solutions provided by previous literature. To that purpose, we first perform an extensive analysis of the incomplete soft set that Qin *et al.* [23] use to illustrate their data filling approach. Specifically, in Example 2 below we compute various indicators that induce procedures for the prioritization of the objects, including our own proposal.

Example 2. A University department is recruiting researchers, and 8 persons apply for the job. The universe of applicants is $U = \{a_1, \ldots, a_8\}$. $E = \{e_1, \ldots, e_6\}$ is the parameter set, the e_i ($i = 1, \ldots, 6$) standing for the parameters "experienced", "young age", "good command of the language", "the highest academic degree is PhD", "the highest academic degree is Master's" and "studied abroad", respectively. The recruiting committee must make a recommendation on the basis of an incomplete soft set (F, E) that captures the "capabilities of the candidates" according to the information contained in Table 5.

	e_1	e_2	e_3	e_4	e_5	e_6
a_1	1	0	1	0	1	0
a_2	1	0	0	1	0	0
a_3	0	1	0	0	1	0
a_4	0	1	*	1	0	*
a_5	1	0	1	1	0	0
a_6	0	1	0	0	*	0
a_7	1	*	1	0	1	0
a_8	0	0	1	1	0	0

Table 5. Tabular representation of the incomplete soft set (F, E) in Qin *et al.* [23] .

In order to analyze this choice situation, Table 6 collects several indicators suggested by previous procedures as well as s_i provided by our own Algorithm. Recall that the Q_i indicators correspond to the choice values in the table completed by Qin *et al.*'s filling algorithm, which they give at Qin et al. [23] . At the bottom of Table 6 we present the respective recommendations by these methodologies. As is apparent, this example readily proves that our proposal of solution does not coincide with the solutions provided by either the $c_{i(0)}$ or the $c_{i(1)}$ or the

Q_i indicators.

	s_i	d_i	d_{i-p}	$c_{i(0)}$	$c_{i(1)}$	Q_i
a_1	0.375	3	3	3	3	3
a_2	0	2	2	2	2	2
a_3	0	2	2	2	2	2
a_4	0.5	2.57	3	2	4	2
a_5	0.375	3	3	3	3	3
a_6	0	1.42	1.5	1	2	2
a_7	0.875	3.42	3.5	3	4	3
a_8	0	2	2	2	2	2
	$\{a_7\}$	$\{a_7\}$	$\{a_7\}$	$\{a_1,a_5,a_7\}$	$\{a_4,a_7\}$	$\{a_1,a_5,a_7\}$

Table 6. Indicators by several focal proposals and prescribed solutions for the incomplete soft set (F, E) in Table 5.

In order to complete our comparative analysis, Example 3 below proves that our solution is indeed different from the solutions provided by the d_i and the d_{i-p} indicators as well. Of course, the reason is that when Zou and Xiao define the weights of all possible choice values, they acquiesce that there is an inherent relationship between object.

Example 3. In the situation of Example 1, three agents need to decide on the basis of the information contained in the incomplete soft set (F_1, E_0). The first agent follows the recommendations by the d_{i-p} indicator, whereas the second one follows the recommendations by the d_i indicator, and the third one follows the recommendations by our Algorithm.

Now the first agent observes a tie since the evaluations of d_{i-p} at h_1 and h_2 are 2 whereas its value is 1.5 at h_3, thus her decision is $\{h_1, h_2\}$. The second agent's only winning option is h_1 because $d_1 = 2$, $d_2 = 1$, and $d_3 = 1.5$. However, the third agent's only winning option is h_3 because $s_1 = 0.6562$, $s_2 = 0.6875$, and $s_3 = 0.4062$.

Hence, we confirm that neither the solutions by the d_{i-p} indicators nor the solutions by the d_i indicators necessarily coincide with the solutions by our s_i indicators.

	e_1	e_2	e_3	e_4
h_1	1	1	0	0
h_2	1	0	*	*
h_3	*	*	*	0

Table 7. Tabular representation of the incomplete soft set (F_1, E_0) in Example 3.

3.5. Computational Experiment

An operational objection to the Algorithm in subsection 3.3 is that it is computationally inefficient: the number of completed tables in Step 4 is 2^w, thus when the number of missing data is very high we cannot reach a solution in a timely manner. Put more precisely, the Algorithm execution time has order $\mathcal{O}(2^w) + \mathcal{O}(n \times m)$, where $n \times m$ is the size of the matrix and w is the number of unknown data in the matrix associated to the soft incomplete set. It is nevertheless possible to apply our Algorithm to all the examples that we have found in related literature in a very reasonable execution time. We proceed to give experimental grounds for this assurance.

The Algorithm we have developed is written in R2014b Matlab language. We run it on a Mac computer with OSX Yosemite system, processor Intel Core i5 CPU I5-2557M at 1,7 GHz and 4 GB RAM.

# unknown	Matrix Dimensions (n, m)				
values	$(5, 10)$	$(25, 100)$	$(50, 200)$	$(75, 250)$	$(100, 500)$
1	0.0008	0.0009	0.0009	0.0011	0.0018
2	0.0012	0.0014	0.0014	0.0014	0.0029
3	0.0017	0.0024	0.0025	0.0028	0.0050
4	0.0032	0.0038	0.0048	0.0049	0.0075
5	0.0056	0.0068	0.0088	0.0093	0.0126
6	0.0109	0.0127	0.0148	0.0174	0.0271
7	0.0226	0.0244	0.0301	0.0318	0.0460
8	0.0446	0.0490	0.0584	0.0702	0.1004
9	0.0926	0.0959	0.1157	0.1542	0.2041
10	0.1836	0.1926	0.2331	0.2844	0.4128
11	0.3782	0.3879	0.4726	0.5603	0.8252
12	0.7535	0.7700	0.9245	1.0916	1.5072
13	1.5563	1.5493	2.0415	2.1525	2.9577
14	3.0519	3.0954	3.8883	4.2896	5.8508
15	6.0364	6.2123	7.3964	9.0436	12.2864
16	12.1147	13.2758	15.4481	17.3812	26.9188

Table 8. Matlab implementation performance of our Algorithm: running times in seconds.

In order to perform a more practical experimental analysis, in Table 8 we consider the following matrix sizes: (a) $n = 5, m = 10$; (b) $n = 25, m = 100$; (c) $n = 50$, $m = 200$; (d) $n = 75$, $m = 250$; (e) $n = 100, m = 500$. Figure 1 shows only cases (a), (d) and (e) to avoid cumbersome displays. For each matrix size we represent a series. In the abscissa axis, the number of unknown data in our matrix is represented. The ordinate axis represents the average runtime in seconds of our Algorithm for 10 random examples of matrices with the same conditions.
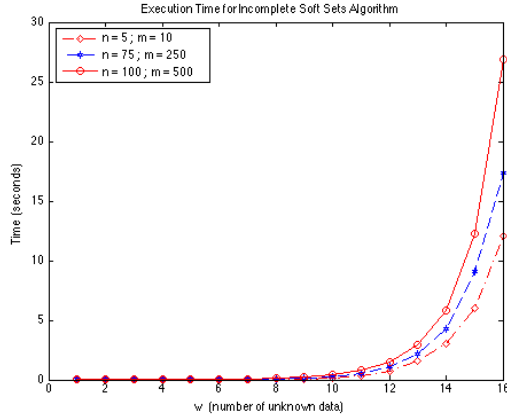
Fig. 1. Matlab implementation performance of our Algorithm. For convenience we partially display the data in Table 8.
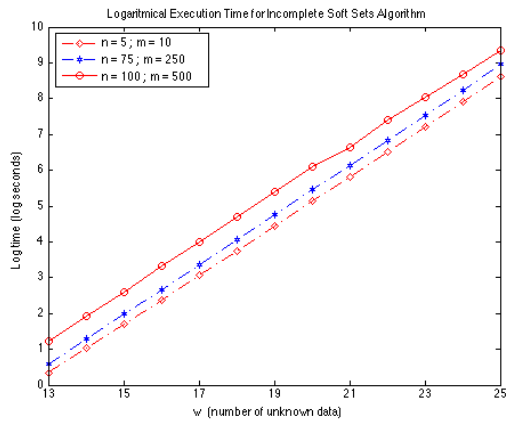

Fig. 2. Matlab implementation performance of our Algorithm. For interpretation of the data, see Section 4.

For larger sets of data, experimental analyses become tediously lengthy. Consequently, Figure 2 is a logarithmic representation analogous to Figure 1 of the same cases (a), (d), and (e) *. However, in this graph instead of using 10 random matrices, we only consider three experiments for the sake of minimizing execution time.

In all series we found that the running time of the Algorithm is exponential to the number of unknown elements in the matrix. For example, we tested it for a matrix with $n = 100$, $m = 500$, and $w = 30$ and the execution time was $2.2848 \cdot 10^5$ seconds ($= 63, 46$ hours $= 2.64$ days) with the same basic computer.

We must acknowledge that when the value of w grows above a certain level, the running time makes the Algorithm impracticable as it stands. But that level does not seem to be surpassed in the literature we have sampled. In this regard it seems convenient to compare our proposal where every missing data are either 0 or 1, with other possible studies where the incomplete infomation is filled with numbers ranging in a larger set. We believe that even if the researcher attempts to estimate the missing data, the majority of the approaches restrict the options to 0 or 1 (cf., DFIS in Qin *et al.* [24], ADFIS in Khan *et al.* [28], and Qin *et al.* [23]). As an exceptional case where larger sets of values can be used, we can only cite the recent Kong *et al.* [18] (see their Tables 4, 7 and 12).

4. Final Comments and Conclusion

Our proposal provides a new criterion for soft set based decision making problems under incomplete information. We have demonstrated that this method is different from the approaches in the existing literature.

The main advantage of our work is that under the conditions in which there is no inherent relationship between objects and between parameters, the decision result will be more convincing than other approaches. In the general framework where no such information on relationships is available, the user must avoid the recourse to possible similarities between objects and/or parameters that are the premise of other approaches. According to Laplace's principle of indifference, in case of complete ignorance we are entitled to assume that all tables in which the ∗'s are replaced with either 0 or 1 are equiprobable. In each of these cases it is reasonable to compute the objects selected according to soft-set based decision making in order to make a decision that is optimal in the highest proportion of cases.

In view of the computational analysis in subsection 3.5, in our future investigations we will focus on related algorithms that permit to preserve the ethos of our approach and take on the cases that require a large execution time.

---

*We use logarithmic representations in order to avoid extreme distortions in the graphical display.

The following two techniques can be used for that purpose.

1. The researcher can select a random sample out of the 2^w possible matrices in such way that with a sufficiently large number of samples, the estimate $\bar{s}_i$ of our true coefficient s_i is deemed sufficiently reliable (cf., Alcantud and Santos-García [4]).

This procedure is efficient since it can be stopped at any time. The verdict as to which options are optimal is adjustable because it depends on the proportion of filled tables that are used. This main feature is common to other approaches (e.g., the use of thresholds in Qin *et al.* [23] in this context; or the adjustable approach by Feng *et al.* [13] in fuzzy soft set based decision making). The choice of the thresholds in the latter proposals is under the control of the users, and the technique in Alcantud and Santos-García [4] leaves the decision of such proportions to the users too.

2. One can produce a modification of our idea in subsection 3.3 that exploits the fact that when there are several missing data for an option, the result of its choice value does not depend on the precise alternatives that are completed with a 1, but rather it is dependent on the exact number of alternatives that are completed with a 1, since the choice value is computed by addition. Hence, our Algorithm can be refined to run in less time by replacing Steps 4 and 5 with a better procedure that goes through all the possibilities and computes how many times each case takes place. We emphasize that a non-trivial combinatorial problem emerges.

Finally, it would be interesting to extend our approach to the theories formulated by Han *et al.* [16] regarding decision making of incomplete soft sets generated by restricted intersection.

Table 4 summarizes the techniques available to the practitioner for the problem we have studied.

Acknowledgment

References

1. H. Aktaş and N. Çağman. Soft sets and soft groups. *Information Sciences*, 177:2726–2735, 2007.
2. J. C. R. Alcantud. A novel algorithm for fuzzy soft set based decision making from multiobserver input parameter data set. *Information Fusion*, 29:142–148, 2016.
3. J. C. R. Alcantud. Some formal relationships among soft sets, fuzzy sets, and their extensions. *International Journal of Approximate Reasoning*, 68:45–53, 2016.
4. J. C. R. Alcantud and G. Santos-García. *Decision Economics, In Commemoration of the Birth Centennial of Herbert A. Simon 1916-2016 (Nobel Prize in Economics 1978): Distributed Computing and Artificial Intelligence, 13th International Conference*, chapter Incomplete Soft Sets: New Solutions for Decision Making Problems, pages 9–17. Springer International Publishing, 2016.
5. JCR Alcantud. Topological properties of spaces ordered by preferences. *International Journal of Mathematics and Mathematical Sciences*, 22:17–27, 1999.
6. JCR Alcantud. Weak utilities from acyclicity. *Theory and decision*, 47(2):185–196, 1999.
7. M.I. Ali. A note on soft sets, rough soft sets and fuzzy soft sets. *Applied Soft Computing*, 11:3329–3332, 2011.
8. M.I. Ali. Another view on reduction of parameters in soft sets. *Applied Soft Computing*, 12(6):1814–1821, 2012.
9. M.I. Ali, F. Feng, X. Liu, W.K. Min, and M. Shabir. On some new operations in soft set theory. *Computers & Mathematics with Applications*, 57(9):1547–1553, 2009.
10. M.I. Ali, T. Mahmood, M.M. Ur Rehman, and M.F. Aslam. On lattice ordered soft sets. *Applied Soft Computing*, 36:499–505, 2015.
11. M.I. Ali and M. Shabir. Logic connectives for soft sets and fuzzy soft sets. *IEEE Transactions on Fuzzy Systems*, 22(6):1431–1442, 2014.
12. N. Çağman and S. Enginoğlu. Soft set theory and uni-int decision making. *European Journal of Operational Research*, 207(2):848–855, 2010.
13. F. Feng, Y.B. Jun, X. Liu, and L. Li. An adjustable approach to fuzzy soft set based decision making. *Journal of Computational and Applied Mathematics*, 234:10–20, 2010.
14. F. Feng and Y. Li. Soft subsets and soft product operations. *Information Sciences*, 232:44–57, 2013.
15. F. Feng, Y. Li, and N. Çağman. Generalized uni-

Method(s) and tool(s)		Indexes	Ref.	Comments
Weight-average of all possible choice values of objects Weights of each choice value given by distribution of the other objects		d_i, d_{i-p}, $c_{i(0)}$, $c_{i(1)}$	38	Original approach
Data filling based on association between parameters Choice from choice values for completed set		Q_i	23	Proposed here. Suitable for cross-related objects Inspired by 38
Elicitation criterions for incomplete soft sets generated by restricted intersections		-	16	Different approach to related problem
Elimination of dominated options Laplacian argument: equal probability to all completed tables Choice mechanism: choice values	*Random sample*	s_i^2	4	Results depend on sample † Sample size under control of user
	Brute force	s_i	Alg. 1	Computationally costly for sizes larger than problems in applied literature †
	Combinatorics	-	-	Future research †

† Suitable when there is no guarantee that objects and parameters are related to each other.

Table 9: Summary table of studies about incomplete soft set based decision making.

int decision making schemes based on choice value soft sets. *European Journal of Operational Research*, 220(1):162–170, 2012.

16. B.-H. Han, Y.-M. Li, J. Liu, S.-L. Geng, and H.-Y. Li. Elicitation criterions for restricted intersection of two incomplete soft sets. *Knowledge-Based Systems*, 59:121–131, 2014.

17. C. Kahraman, S. Cevik Onar, and B. Oztaysi. Fuzzy multicriteria decision-making: A literature review. *International Journal of Computational Intelligence Systems*, 8(4):637–666, 2015.

18. Z. Kong, G. Zhang, L. Wang, Z. Wu, S. Qi, and H. Wang. An efficient decision making approach in incomplete soft set. *Applied Mathematical Modelling*, 38(7–8):2141–2150, 2014.

19. P.K. Maji, R. Biswas, and A.R. Roy. Fuzzy soft sets. *Journal of Fuzzy Mathematics*, 9:589–602, 2001.

20. P.K. Maji, R. Biswas, and A.R. Roy. An application of soft sets in a decision making problem. *Computers and Mathematics with Applications*, 44:1077–1083, 2002.

21. P.K. Maji, R. Biswas, and A.R. Roy. Soft set theory. *Computers and Mathematics with Applications*,

45:555–562, 2003.

22. D. Molodtsov. Soft set theory - first results. *Computers and Mathematics with Applications*, 37:19–31, 1999.

23. H. Qin, X. Ma, T. Herawan, and J. Zain. Data filling approach of soft sets under incomplete information. In NgocThanh Nguyen, Chong-Gun Kim, and Adam Janiak, editors, *Intelligent Information and Database Systems*, volume 6592 of *Lecture Notes in Computer Science*, pages 302–311. Springer Berlin Heidelberg, 2011.

24. H. Qin, X. Ma, T. Herawan, and J. Zain. DFIS: A novel data filling approach for an incomplete soft set. *International Journal of Applied Mathematics and Computer Science*, 22:817–828, 2012.

25. K. Qin, D. Meng, Z. Pei, and Y. Xu. Combination of interval set and soft set. *International Journal of Computational Intelligence Systems*, 6(2):370–380, 2013.

26. R.M. Rodríguez, B. Bedregal, H. Bustince, Y.C. Dong, B. Farhadinia, C. Kahraman, L. Martínez, V. Torra, Y.J. Xu, Z.S. Xu, and F. Herrera. A position and perspective analysis of hesitant fuzzy sets on information fusion in decision making. towards

high quality progress. *Information Fusion*, 29:89–97, 2016.

27. R.M. Rodríguez, L. Martínez, V. Torra, Z.S. Xu, and F. Herrera. Hesitant fuzzy sets: state of the art and future directions. *International Journal of Intelligent Systems*, 29:495–524, 2014.

28. M. Sadiq Khan, M.A. Al-Garadi, A.W.A. Wahab, and T. Herawan. An alternative data filling approach for prediction of missing data in soft sets (adfis). *SpringerPlus*, 5(1):1348, 2016.

29. Y. Shao and K. Qin. Fuzzy soft sets and fuzzy soft lattices. *International Journal of Computational Intelligence Systems*, 5(6):1135–1147, 2012.

30. V. Torra. Hesitant fuzzy sets. *International Journal of Intelligent Systems*, 25(6):529–539, 2010.

31. F. Wang, X. Li, and X. Chen. Hesitant fuzzy soft set and its applications in multicriteria decision making. *Journal of Applied Mathematics*, pages Article ID 643785, 10 pages, 2014.

32. Y.Y. Yao. Relational interpretations of neighbourhood operators and rough set approximation operators. *Information Sciences*, 111:239–259, 1998.

33. L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

34. J. Zhan, Q. Liu, and B. Davvaz. A new rough set theory: rough soft hemirings. *Journal of Intelligent & Fuzzy Systems*, 28(4):1687–1697, 2015.

35. J. Zhan, Qi Liu, and W. Zhu. Another approach to rough soft hemirings and corresponding decision making. *Soft Computing*, pages 1–12, 2016.

36. J. Zhan and K. Zhu. A novel soft rough fuzzy set: Z-soft rough fuzzy ideals of hemirings and corresponding decision making. *Soft Computing*, pages 1–14, 2016.

37. X. Zhang. On interval soft sets with applications. *International Journal of Computational Intelligence Systems*, 7(1):186–196, 2014.

38. Y. Zou and Z. Xiao. Data analysis approaches of soft sets under incomplete information. *Knowledge-Based Systems*, 21(8):941–945, 2008.