Int. J. Appl. Math. Comput. Sci., 2021, Vol. 31, No. 3, 417–430
DOI: 10.34768/amcs-2021-0028

AN INCOMPLETE SOFT SET AND ITS APPLICATION IN MCDM PROBLEMS WITH REDUNDANT AND INCOMPLETE INFORMATION

SISI XIA [a,*], HAORAN YANG [a], LIN CHEN [b]

[a] School of Economics
Southwest University of Political Science and Law
No. 301, Baosheng Ave., Yubei District, 401120 Chongqing, China
e-mail: {xiasisi,yanghaoran}@swupl.edu.cn

[b] Chongqing Institute of Green and Intelligent Technology
Chinese Academy of Sciences
No. 266, Fangzheng Ave., Shuitu Town, Beibei District, 400714 Chongqing, China
e-mail: chenlin@cigit.ac.cn

Multiple criteria decision making (MCDM) problems in practice may simultaneously contain both redundant and incomplete information and are difficult to solve. This paper proposes a new decision-making approach based on soft set theory to solve MCDM problems with redundant and incomplete information. Firstly, we give an incomplete soft set a precise definition. After that, the binary relationships of objects in an incomplete soft set are analyzed and some operations on it are provided. Next, some definitions regarding the incomplete soft decision system are also given. Based on that, an algorithm to solve MCDM problems with redundant and incomplete information based on an incomplete soft set is presented and illustrated with a numerical example. The results show that our newly developed method can be directly used on the original redundant and incomplete data set. There is no need to transform an incomplete information system into a complete one, which may lead to bad decision-making due to information loss or some unreliable assumptions about the data generating mechanism. To demonstrate its practical applications, the proposed method is applied to a problem of regional food safety evaluation in Chongqing, China.

Keywords: soft set, incomplete soft set, incomplete information, redundant information, multiple criteria decision making.

1. Introduction

Many practical problems in engineering, economics, social science, medical science, etc. are difficult to solve because they involve vagueness and uncertainty. Various mathematical theories have been proposed by researchers, such as fuzzy set theory (Zadeh, 1965; Goguen, 1967), rough set theory (Pawlak, 1984; 1985), and vague set theory (Gau and Buehrer, 1993; Hong and Choi, 2000). These approaches have been successfully applied to many practical problems to eliminate vagueness and uncertainty in the modeling. It should be pointed out, however, that each of the above theories has its inherent limitation in providing adequate parametrization tools (Molodtsov, 1999).

Soft set theory is a new mathematical tool to deal with vagueness and uncertainty in practical problems that was initiated by Molodtsov (1999). In a soft set, the value domain of the mapping function is a set of all subsets of objects in the initial finite universe, and it can be a classical set, fuzzy set, or intuitive fuzzy set, among others. Therefore, a soft set is free from the limitation of the inadequacy of the parameterization tools. It has been generalized into fuzzy soft sets (Roy and Maji, 2007; Majumdar and Samanta, 2010), rough soft sets (Meng *et al.*, 2011), vague soft sets (Xu *et al.*, 2019), interval-valued fuzzy soft sets (Alkhazaleh and Salleh, 2012; Yang and Yao, 2020), interval-valued intuitionistic fuzzy soft sets (Jiang *et al.*, 2010), bijective soft sets (Gong *et al.* 2010; 2017), hybrid rough-bijective soft set (Inbarani *et al.*, 2018), and other hybrid soft sets in combination with various theories.

Potential applications of soft set theory include

*Corresponding author

smoothing, game theory, operational analysis, integration, probability theory, measurement theory (Molodtsov, 1999), forecasting (Xu *et al.*, 2019), decision-making (Petchimuthu *et al.*, 2020; Sun *et al.*, 2020; Yang and Yao, 2020), evaluation (Li *et al.*, 2018), medical diagnosis (Akram *et al.*, 2020), and concept selection (Tiwari *et al.*, 2017; 2019), to name a few. The application of soft set in decision-making problems is of great interest to many researchers (Zhang and Zhang, 2013; Li *et al.*, 2015; Peng and Yang, 2017; Das *et al.*, 2018; Garg and Arora, 2018).

An MCDM problem is usually characterized by the ratings of several alternatives with respect to certain criteria and weights assigned to each criterion. In a classical MCDM problem, all criteria are indispensable and independent of each other, and they are attached with different weights to describe their relative importance. However, in real-life MCDM problems, it is usually impossible to find indispensable and independent criteria. Searching for the solution to an MCDM problem is hampered by redundant and interrelated information, which is a typical feature of information in the era of information explosion. Moreover, practical decision-making problems are also characterized by incomplete information for some reasons, such as limited approaches in data collection, data missing in storage, and so on. Therefore, the combination of redundant and incomplete information in practical MCDM problems makes decision-making difficult.

For redundant information in practical MCDM problems, some effective parameter reduction methods of soft sets have been proposed to solve complex MCDM problems with complete information (Maji and Roy, 2002; Chen *et al.*, 2005; Kong *et al.*, 2008; Khan and Zhu, 2020), which not only simplifies MCDM problems, but also improves the decision-making efficiency. However, these methods cannot be applied to MCDM problems with incomplete information.

For incomplete information in practical MCDM problems, there are two main solution methods, that is, deletion and data filling (Kryszkiewicz, 1999). Both are aimed at transforming incomplete information into the complete kind. In deletion, objects with missing values are deleted directly. This method creates a complete information system at the cost of discarding valuable information in deletion, which may lead to a suboptimal solution to the MCDM problem because the deleted objects with missing values may happen to be the optimal objects. In data filling, no object is deleted. Instead, the missing values are filled with some estimated values, such as through experts' opinion, sample mean, or machine learning algorithms.

Even though the combination of a soft set with a data filling method can be used effectively to process incomplete information and all objects are retained (Zou and Xiao, 2008; Qin *et al.*, 2012; Deng and Wang, 2013; Liu *et al.*, 2017), each data filling method has its limitations. For example, data filling by experts is relatively subjective; data filling by sample mean is built on the assumption of a uniform distribution for the variable with a missing value, which may be too restrictive in some circumstances; and data filling by Bayesian inference (one of the machine learning methods) requires an arbitrarily chosen prior distribution for the parameters. Moreover, based on these data filling methods, MCDM problems are solved by choosing optimal objects with the maximal choice values calculated by all parameters.

As mentioned above, the information we collect may be interrelated and redundant. That means that not all the parameters are essential for decision-making, including these with missing values. Filling all the missing values cannot render decision-making more efficient, but it may generate excessive parameters and unnecessarily complicate MCDM problems. Therefore, for an MCDM problem with redundant and incomplete information, the main task is to remove redundant information by parameter reduction to find the core parameters, instead of filling the missing data based on some subjective and arbitrary methods to obtain a complete information system. The existing approaches of incomplete data analysis in a soft set do not provide an effective method of parameter reduction, which makes these approaches inappropriate to addressing MCDM problems with redundant and incomplete information.

In this paper, we propose a new MCDM method based on soft set theory to deal with redundant and incomplete information in the decision-making process. This new method can reduce redundant information appropriately to make decisions more effectively, and, in the meantime, it can process incomplete information axiomatically to avoid the subjectiveness and arbitrariness in the previous methods.

The main contributions of this work can be summarized as follows. Firstly, to describe the incomplete information in MCDM problems, we give a specific definition of a soft set with incomplete information (incomplete soft set) according to its own data structure. Based on that, its characteristics and operations are analyzed. Secondly, to process the redundant information in MCDM problems, this paper provides a parameter reduction approach of incomplete soft sets, which can be used in incomplete information circumstances, while the other parameter reduction approaches of soft sets can only be applied to MCDM problems with complete information. Thirdly, instead of using data filling or deleting methods to transform an incomplete data set into a complete one as in previous studies, our work develops a new method to process incomplete information in MCDM problems by comparing each pair of objects directly (including objects with incomplete information) to classify them and generate a series of decision rules.

The rest of this paper is organized as follows. After introducing the basic definition of soft sets in Section 2, the concepts of incomplete soft sets are defined in Section 3. In Section 4, the incomplete soft decision system is analyzed. Section 5 proposes a decision-making approach of MCDM problems with redundant and incomplete information based on an incomplete soft set. Section 6 presents some results of food safety evaluation from Chongqing, China, by using the newly developed methods. Section 7 summarizes and concludes this paper.

2. Preliminaries

For subsequent discussion, we briefly recall the concept of a soft set in this section.

Suppose that $U = \{h_1, h_2, \ldots, h_n\}$ is a common universe set and $A = \{e_1, e_2, \ldots, e_m\}$ is a set of parameters.

Definition 1. (*Soft set* (*Molodtsov, 1999*)) Let $P(U)$ be the set of all subsets of U. We call a pair (F, A) a soft set over U, where F is a mapping given by $F : A \to P(U)$.

In other words, the soft set is a parameterized family of subsets of the set U. Every set $F(e)$, where $e \in A$, from this family can be considered a set of e-approximate elements of the soft set (F, A), and it is a subset of U.

Example 1. (*Soft set*) Suppose that $U = \{h_1, h_2, \ldots, h_6\}$ is a set of houses under consideration; $A = \{e_1, e_2, \ldots, e_5\}$ is a set of parameters which stand for cheap, beautiful, size, location, and in the green surroundings, respectively; and each parameter is a word or a sentence. Then

$$U = \{h_1, h_2, \ldots, h_6\},$$

and

$$\begin{aligned} A &= \{e_1, e_2, \ldots, e_5\} \\ &= \{\text{cheap, beautiful, size, location,} \\ &\qquad \text{in the green surroundings}\}. \end{aligned}$$

In this case, to define a soft set means to point out cheap houses, beautiful houses, and so on. The soft set (F, A) describes the "attractiveness of the houses" which Mr. X is going to buy and consists of five subsets of U as follows:

$$\begin{aligned} (F, A) = \Big\{ & F(e_1) = \{\text{cheap houses}\} = \{h_3, h_5\}, \\ & F(e_2) = \{\text{beautiful houses}\} \\ & = \{h_1, h_2, h_4, h_6\}, \\ & F(e_3) = \{\text{big houses}\} = \{h_4, h_6\}, \\ & F(e_4) = \{\text{good location houses}\} \\ & = \{h_4, h_5, h_6\}, \\ & F(e_5) = \{\text{in the green suroundings houses}\} \\ & = \{h_1, h_4, h_6\} \Big\}. \end{aligned}$$

To store a soft set on a computer and facilitate calculation, we can represent the soft set of the attractiveness of houses in the form of Table 1. ♦

3. Incomplete soft set

3.1. Concept of incomplete soft sets. In many practical MCDM problems, incomplete information in soft sets is inevitable for some objective reasons. In this case, the soft set (F, A) contains missing values. This leads to a situation that the results of mapping F from A to $P(U)$ in the soft set (F, A), that is, $F(e)$ $(\forall e \in A)$, may contain uncertain or unknown objects. Thus, mapping F in (F, A) might be inaccurate, and the soft sets defined from the mapping F are inappropriate as well.

The data structure of a soft set that contains incomplete information is different from that of a soft set with complete information, and the mapping roles on the soft set with complete information cannot be applied directly on the soft set with incomplete information. Currently, there is no clear definition of a soft set with incomplete information and its mapping in terms of its data structure. Therefore, it is necessary to present an accurate definition of an incomplete soft set and discuss its unique data structure characteristics.

Definition 2. (*Incomplete soft set*) A soft set (F, A) is called a complete soft set if and only if $F(e)$ does not contain uncertain or unknown objects; otherwise, it is called an incomplete soft set and denoted by (F', A).

In an incomplete soft set (F', A), each set $F'(e)$ can be regarded as e-approximate elements of (F', A). However, unlike $F(e)$ in a complete soft set, $F'(e)$ contains two parts, that is, the deterministic and the nondeterministic part. Therefore, the incomplete soft set (F', A) can be denoted by

$$(F', A) = \{F'(e) = \{h_i\} \cup \{h_j\}\}, \tag{1}$$

Table 1. Tabular representation of the soft set (F, A).

U	e_1	e_2	e_3	e_4	e_5
h_1	0	1	0	0	1
h_2	0	1	0	0	0
h_3	1	0	0	0	0
h_4	0	1	1	1	1
h_5	1	0	0	1	0
h_6	0	1	1	1	1

where $e \in A$, $h_i, h_j \in U$, $\{h_i\}$ is the set of objects with known information on attribute e, which belong to $F'(e)$ explicitly; and $\{h_j\}$ denotes the set of objects with incomplete information on attribute e, which may or may not belong to $F'(e)$.

Example 2. (*Incomplete soft set*) For better understanding, reconsider Example 1. Suppose that in the soft set (F, A) the unknown values are the information of object h_2 on attributes e_2, object h_3 on attributes e_1 and e_4, object h_5 on attribute e_4, and object h_6 on attribute e_3. Then, the incomplete soft set (F', A) can be denoted by

$$\begin{aligned}(F', A) = \Big\{ & F'(e_1) = \{\text{cheap houses}\} = \{h_5\} \cup \{h_3\},\\ & F'(e_2) = \{\text{beautiful houses}\}\\ & = \{h_1, h_4, h_6\} \cup \{h_2\},\\ & F'(e_3) = \{\text{big houses}\} = \{h_4\} \cup \{h_6\},\\ & F'(e_4) = \{\text{good location houses}\}\\ & = \{h_4, h_6\} \cup \{h_3, h_5\},\\ & F'(e_5) = \{\text{in the green suroundings houses}\}\\ & = \{h_1, h_4, h_6\} \cup \emptyset\Big\}.\end{aligned}$$

In (F', A), $F'(e_1)$, for example, denotes the set of houses that may be cheap. Among them, h_5 is absolutely a cheap house and h_3 is possibly a cheap house. Its tabular representation is shown in Table 2, in which the special symbol '*' is used to indicate incomplete information on attributes. ♦

It should be noted that, in an incomplete soft set, an unknown value does not mean it is useless; instead, it has increased the uncertainty and difficulty in data analysis.

Definition 3. (*Incomplete soft subset*) Let (F', A) and (G', B) be two incomplete soft sets. (F', A) is said to be an incomplete soft subset of (G', B), which is denoted by $(F', A) \subseteq (G', B)$, if and only if $A \subseteq B$, and $\forall e \in A, F'(e) \subseteq G'(e)$.

Correspondingly, (G', B) is said to be an incomplete soft superset of (F', A) if (F', A) is an incomplete soft subset of (G', B). We denote it by $(G', B) \supseteq (F', A)$.

Table 2. Tabular representation of the soft set (F', A).

U	e_1	e_2	e_3	e_4	e_5
h_1	0	1	0	0	1
h_2	0	*	0	0	0
h_3	*	0	0	*	0
h_4	0	1	1	1	1
h_5	1	0	0	*	0
h_6	0	1	*	1	1

[*] Note that * means incomplete information

Example 3. (*Incomplete soft subset*) Given two incomplete soft sets (F', A) and (G', B), suppose that

$$\begin{aligned}U &= \{h_1, h_2, h_3, h_4, h_5, h_6\} \text{ is a set of houses,}\\ A &= \{e_1, e_2\} = \{\text{cheap, beautiful}\},\\ B &= \{e_1, e_2, e_3\} = \{\text{cheap, beautiful, size}\}.\end{aligned}$$

Here, A and B are two sets of parameters, and an incomplete soft set (F', A) and (G', B) can be defined by

$$\begin{aligned}(F', A) = \{ & F'(e_1) = \{h_5\} \cup \{h_3\},\\ & F'(e_2) = \{h_1, h_4, h_6\} \cup \{h_2\}\},\\ (G', B) = \{ & G'(e_1) = \{h_4, h_5\} \cup \{h_3\},\\ & G'(e_2) = \{h_1, h_4, h_6\} \cup \{h_2, h_5\},\\ & G'(e_3) = \{h_1, h_4, h_6\} \cup \emptyset\}.\end{aligned}$$

Therefore, we have $(F', A) \subseteq (G', B)$. ♦

Definition 4. (*Incomplete soft equal*) Let (F', A) and (G', B) be two incomplete soft sets. If, and only if, $(F', A) \subseteq (G', B)$ and $(F', A) \supseteq (G', B)$, then (F', A) and (G', B) are called incomplete soft equal, which is denoted by $(F', A) = (G', B)$.

3.2. Operations on incomplete soft sets.

3.2.1. Binary relationships. Accurate classification of all alternatives is a precondition for decision making. This section analyzes the binary relationships of each pair of objects in an incomplete soft set to judge whether they have the same properties and can be categorized into one class. Before discussing the binary relationships of objects in incomplete soft sets, the binary relationships of objects in complete soft sets are first presented.

Definition 5. (*Indiscernibility relation*) Let (F, A) be a complete soft set over a common universe U, and each subset of attributes $B \subseteq A$ determines a binary indiscernibility relation $IND(B)$ on U as follows:

$$\begin{aligned}\text{IND}(B) = \{(h_i, h_j) \in U \times U \mid \forall e \in B,\\ h_i, h_j \in F(e) \text{ or } h_i, h_j \in \overline{F(e)}\},\end{aligned} \quad (2)$$

where h_i and h_j denote two objects on U; $F(e)$ is the set of e-approximate elements of (F, A), which means the set of objects whose values on attribute e are 1; $\overline{F(e)}$, which is the complement set of $F(e)$, denotes the set of objects whose values on attribute e are 0.

Since the information in the soft set (F, A) is complete, it is certain whether or not each object belongs to the set $F(e)$. That is to say, if one object belongs to $F(e)$, then it would not belong to $\overline{F(e)}$, that is, $F'(e) \cap \overline{F(e)} = \emptyset$.

According to Definition 5, it is clear that, if a pair of objects (h_i, h_j) from $U \times U$ belongs to $F(e)$ or $\overline{F(e)}$ in the soft set (F, A), then they are in $\mathrm{IND}(B)$ and can be perceived as indiscernible. In other words, they may have the same properties with respect to B in reality. The indiscernibility relation $\mathrm{IND}(B)$ is an equivalence one and generates a partition of U. Each partition can be called an indiscernibility class $I_B(h_i)$ and is defined as

$$I_B(h_i) = \{h_j \in U \mid (h_i, h_j) \in \mathrm{IND}(B)\}. \tag{3}$$

Here, $I_B(h_i)$ describes the objects that are indiscernible from h_i in terms of B.

Example 4. (*Indiscernibility relation*) In Example 1,

$$\begin{aligned}
&(F, A) \\
&= \Big\{F(e_1) = \{\text{cheap houses}\} = \{h_3, h_5\}, \\
&\quad F(e_2) = \{\text{beautiful houses}\} \\
&\quad = \{h_1, h_2, h_4, h_6\}, \\
&\quad F(e_3) = \{\text{big houses}\} = \{h_4, h_6\}, \\
&\quad F(e_4) = \{\text{good location houses}\} \\
&\quad = \{h_4, h_5, h_6\}, \\
&\quad F(e_5) = \{\text{in the green suroundings houses}\} \\
&\quad = \{h_1, h_4, h_6\}\Big\}.
\end{aligned}$$

Complementary to the soft set (F, A) we have

$$\begin{aligned}
\overline{F(e_1)} &= \{h_1, h_2, h_4, h_6\}, \\
\overline{F(e_2)} &= \{h_3, h_5\}, \\
\overline{F(e_3)} &= \{h_1, h_2, h_3, h_5\}, \\
\overline{F(e_4)} &= \{h_1, h_2, h_3\}, \\
\overline{F(e_5)} &= \{h_2, h_3, h_5\}.
\end{aligned}$$

Then, according to Definition 5,

$$\begin{aligned}
\mathrm{IND}(A) = \{&(h_1, h_1), (h_2, h_2), (h_3, h_3), (h_4, h_4), \\
&(h_5, h_5), (h_6, h_6), (h_4, h_6), (h_6, h_4)\}, \\
I_A(h_1) &= \{h_1\}, \quad I_A(h_2) = \{h_2\}, \\
I_A(h_3) &= \{h_3\}, \quad I_A(h_4) = \{h_4, h_6\}, \\
I_A(h_5) &= \{h_5\}, \quad I_A(h_6) = \{h_4, h_6\}.
\end{aligned}$$

♦

However, in incomplete soft sets, the situation becomes somewhat different and much more complicated. For example, in an incomplete soft set (F', A), each set $F'(e)(e \in A)$ contains two parts, that is, the subset $\{h_i\}$ that contains alternatives with complete information and the subset $\{h_j\}$ that contains alternatives with incomplete information on attribute e. Thus, $F'(e)$ means the set of objects whose values on attribute e may be 1, and $\overline{F'(e)}$, which is a complementary set of $F'(e)$, means the set of objects whose values on attribute e may be 0. If the information on attribute e for some objects is unknown, then these objects can be simultaneously classified as $F'(e)$ and $\overline{F'(e)}$. As a result, $F'(e) \cap \overline{F'(e)} \neq \emptyset$, but it equals the set of objects with unknown information on attribute e.

For example, in the incomplete soft set (F', A) of Example 2,

$$\begin{aligned}
F'(e_1) &= \{h_5\} \cup \{h_3\} = \{h_3, h_5\}, \\
\overline{F'(e_1)} &= \{h_1, h_2, h_4, h_6\} \cup \{h_3\} \\
&= \{h_1, h_2, h_3, h_4, h_6\},
\end{aligned}$$

thus $F'(e_1) \cap \overline{F'(e_1)} = \{h_3\} \neq \emptyset$.

Considering that $F'(e_1) \cap \overline{F'(e_1)} \neq \emptyset$, an object with incomplete information on attribute e can belong to $F'(e)$ and $\overline{F'(e)}$ at the same time. Therefore, the relationship between objects with incomplete information cannot be defined through $\mathrm{IND}(B)$. Therefore, we propose to use binary similarity relations in Definition 6 instead to describe the incomplete soft sets.

Definition 6. (*Similarity relation*) Let (F', A) be an incomplete soft set over a common universe U, $B \subseteq A$, and a similarity relation $\mathrm{SIM}(B)$ on U be defined as

$$\begin{aligned}
\mathrm{SIM}(B) = \{&(h_i, h_j) \in U \times U \mid \forall e \in B, \\
&h_i, h_j \in F'(e) \text{ or } h_i, h_j \in \overline{F'(e)}\}.
\end{aligned} \tag{4}$$

where h_i and h_j denote the objects on U; $F'(e)$ is the set of e-approximate elements of (F', A), which denotes the set of objects whose values on attribute e may be 1; and $\overline{F'(e)}$ denotes the set of objects whose values on attribute e may be 0.

Correspondingly, we can define the similarity class $S_B(h_i)$ as the objects that may be similar to h_i in terms of the attributes in B in the incomplete soft sets as

$$S_B(h_i) = \{h_j \in U \mid (h_i, h_j) \in \mathrm{SIM}(B)\}. \tag{5}$$

Example 5. (*Similarity relation*) In Example 2, for the incomplete soft set (F', A) we have

$$\begin{aligned}
&(F', A) \\
&\quad = \Big\{F'(e_1) = \{h_5\} \cup \{h_3\} = \{h_3, h_5\}, \\
&\qquad F'(e_2) = \{h_1, h_4, h_6\} \cup \{h_2\} \\
&\quad = \{h_1, h_2, h_4, h_6\}, \\
&\qquad F'(e_3) = \{h_4\} \cup \{h_6\} = \{h_4, h_6\},
\end{aligned}$$

$$F'(e_4) = \{h_4, h_6\} \cup \{h_3, h_5\} = \{h_3, h_4, h_5, h_6\},$$
$$F'(e_5) = \{h_1, h_4, h_6\} \cup \emptyset = \{h_1, h_4, h_6\}\Big\}.$$

Complementary to the incomplete soft set (F', A) we have

$$\overline{F'(e_1)} = \{h_1, h_2, h_3, h_4, h_6\},$$
$$\overline{F'(e_2)} = \{h_2, h_3, h_5\},$$
$$\overline{F'(e_3)} = \{h_1, h_2, h_3, h_5, h_6\},$$
$$\overline{F'(e_4)} = \{h_1, h_2, h_3, h_5\},$$
$$\overline{F'(e_5)} = \{h_2, h_3, h_5\}.$$

Then, according to Definition 6,

$$\begin{aligned}\mathrm{SIM}(A) = \{&(h_1, h_1), (h_2, h_2), (h_3, h_3), (h_4, h_4),\\ &(h_5, h_5), (h_6, h_6), (h_2, h_3), (h_3, h_2),\\ &(h_3, h_5), (h_5, h_3), (h_4, h_6), (h_6, h_4)\}.\end{aligned}$$

In SIM(A), for example, the pair (h_3, h_5) belongs to $F'(e_1)$, $\overline{F'(e_2)}$, $\overline{F'(e_3)}$, $F'(e_4)$, $\overline{F'(e_4)}$ and $\overline{F'(e_5)}$, and thus it is in SIM(A), in which $A = \{e_1, e_2, e_3, e_4, e_5\}$. In contrast, the relationship between h_3 and h_5 cannot be defined through IND(A) because of incomplete information.

Moreover, we can have

$$\begin{aligned}S_A(h_1) &= \{h_1\}, & S_A(h_2) &= \{h_2, h_3\},\\ S_A(h_3) &= \{h_2, h_3, h_5\}, & S_A(h_4) &= \{h_4, h_6\},\\ S_A(h_5) &= \{h_3, h_5\}, & S_A(h_6) &= \{h_4, h_6\}.\end{aligned}$$

♦

3.2.2. Restricted/relaxed AND operation. The operations on incomplete soft sets can be discussed based on the definition of incomplete soft set and its binary relationships.

Definition 7. (*Restricted AND operation*) Given that (F', A) is an incomplete soft set and $X \subseteq U$, the operation of "(F', A) restricted AND X" denoted by $(F', A)\underline{\wedge}X$ is defined as

$$(F', A)\underline{\wedge}X = \{h_i \in U \mid S_A(h_i) \subseteq X\}. \tag{6}$$

Definition 8. (*Relaxed AND operation*) Given that (F', A) is an incomplete soft set and $X \subseteq U$, the operation of "(F', A) relaxed AND X " denoted by $(F', A)\bar{\wedge}X$ is defined as

$$(F', A)\bar{\wedge}X = \{h_i \in U \mid S_A(h_i) \cap X \neq \emptyset\}. \tag{7}$$

Example 6. (*Restricted/relaxed AND operation*) Consider Example 2 and suppose that $X = \{h_3, h_5\}$ is a subset of the universe U; then, according to Definitions 7 and 8,

$$(F', A)\underline{\wedge}X = \{h_5\},$$
$$(F', A)\bar{\wedge}X = \{h_2, h_3, h_5\}.$$

♦

It can be concluded that the result of the restricted AND operation is a set of objects whose similarity class belongs to X with certainty, while the result of relaxed AND operation is a set of objects whose similarity class possibly belongs to X.

Theorem 1. *Let U be a common universe set and (F', A) be an incomplete soft set. Then (F', B_1) and (F', B_2) are two incomplete soft subsets of (F', A), and $(F', B_1) \subseteq (F', B_2) \subseteq (F', A)$. For $X \subseteq U$, we have*

$$(F', B_1)\underline{\wedge}X \subseteq (F', B_2)\underline{\wedge}X \tag{8}$$

$$(F', B_1)\bar{\wedge}X \supseteq (F', B_2)\bar{\wedge}X. \tag{9}$$

Proof. With $\forall h_i \in U$, if $(F', B_1) \subseteq (F', B_2)$, then $S_{B_1}(h_i) \supseteq S_{B_2}(h_i)$. Assume that $S_{B_1}(h_i) \subseteq X$, where $X \subseteq U$, and then $S_{B_2}(h_i) \subseteq X$. At the same time, there may be $h_j \in U$, $S_{B_1}(h_j) \nsubseteq X$ and $S_{B_2}(h_j) \subseteq X$. Therefore, $(F', B_1)\underline{\wedge}X \subseteq (F', B_2)\underline{\wedge}X$.

Similarly, $\forall h_i \in U$, and we assume that $S_{B_2}(h_i) \cap X \neq \emptyset$. Then $S_{B_1}(h_i) \cap X \neq \emptyset$. At the same time, there may be $h_j \in U$, $S_{B_2}(h_j) \cap X = \emptyset$ and $S_{B_1}(h_j) \cap X \neq \emptyset$. Therefore, $(F', B_1)\bar{\wedge}X \supseteq (F', B_2)\bar{\wedge}X$. ■

4. Incomplete soft decision system

To develop an MCDM method based on incomplete soft sets, it is necessary to introduce incomplete soft decision systems based on incomplete soft sets.

4.1. Concept of an incomplete soft decision system.

In this section, the definitions of the soft decision system and incomplete soft decision system are first proposed.

Definition 9. (*Soft decision system*) Suppose that (F, A) and (G, B) are two soft sets over a common universe U and $A \cap B = \emptyset$. Then, the triple $((F, A), (G, B), U)$ is defined as a soft decision system over the common universe U, where (F, A) is the condition soft set and (G, B) is the decision soft set.

Example 7. (*Soft decision system*) Let the soft set (F, A) in Example 1 be the condition soft set, and soft set (G, B) be the decision soft set, where

$$\begin{aligned}(G, B) = \{&G(\varepsilon_1) = \{h_3, h_5\},\\ &G(\varepsilon_2) = \{h_1, h_2, h_4, h_6\}\}.\end{aligned}$$

Here, ε_1 and ε_2 are two attributes in the attribute set B and denote "unattractive houses" and "attractive houses", respectively.

Then, the triple $((F, A), (G, B), U)$ is a soft decision system. ♦

Definition 10. (*Incomplete soft decision system*) In the soft decision system $((F, A), (G, B), U)$, if there is any $* \in F(e)$ for some $e \in A$ and $* \notin G(\varepsilon), \forall \varepsilon \in B$, then it is called an incomplete soft decision system and denoted by $((F', A), (G, B), U)$.

It should be noted that, in an incomplete soft decision system $((F', A), (G, B), U)$, the incomplete information only exists in the condition soft set (F', A) but not in the decision soft set (G, B).

Example 8. (*Incomplete soft decision system*) Let the incomplete soft set (F', A) in Example 2 be the condition soft set and the soft set (G, B) be the decision soft set.

Then, the triple $((F', A), (G, B), U)$ is an incomplete soft decision system. ♦

4.2. Significance of an attribute subset. This section presents the definition of significance of an attribute subset in an incomplete soft set, which is an important indicator for parameter reduction.

Definition 11. (*Significance of an attribute set*) Let $((F', A), (G, B), U)$ be an incomplete soft decision system, (F', A) an incomplete soft set, and $C \subseteq A$ an attribute subset. The significance of the attribute subset C is denoted by $\mathrm{SIG}(C)$ and defined as

$$\mathrm{SIG}(C) = \Big| \bigcup_{\varepsilon_i \in B} (F', C) \underline{\triangle} G(\varepsilon_i) \Big|, \tag{10}$$

where $|\cdot|$ means the cardinal number of a set.

From Definition 11, the quantity of valid information that an attribute subset in an incomplete soft set contains can be measured by comparing the condition soft set and the decision soft set in an incomplete soft information system based on the restricted AND operation.

Example 9. (*Significance of an attribute set*) In Example 8, according to Definition 11, the significance of the attribute set A in (F', A) can be computed by

$$\begin{aligned} \mathrm{SIG}(C) &= \Big| \bigcup_{\varepsilon_i \in B} (F', A) \underline{\triangle} G(\varepsilon_i) \Big| \\ &= | (F', A) \underline{\triangle} G(\varepsilon_1) \cup (F', A) \underline{\triangle} G(\varepsilon_2) | \\ &= | \{h_5\} \cup \{h_1, h_4, h_6\} | = 4. \end{aligned}$$

Theorem 2. *Let $((F', A), (G, B), U)$ be an incomplete soft decision system and $C_1 \subseteq C_2 \subseteq A$; then, we have*

$$\mathit{SIG}(C_1) \le \mathit{SIG}(C_2). \tag{11}$$

Proof. We have

$$\mathrm{SIG}(C_1) = \Big| \bigcup_{\varepsilon_i \in B} (F', C_1) \underline{\triangle} G(\varepsilon_i) \Big|,$$

$$\mathrm{SIG}(C_2) = \Big| \bigcup_{\varepsilon_i \in B} (F', C_2) \underline{\triangle} G(\varepsilon_i) \Big|.$$

From Theorem 1, for each $\varepsilon_i \in B$,

$$(F', C_1) \underline{\triangle} G(\varepsilon_i) \subseteq (F', C_2) \underline{\triangle} G(\varepsilon_i),$$

so that

$$\Big| \bigcup_{\varepsilon_i \in B} (F', C_1) \underline{\triangle} G(\varepsilon_i) \Big| \le \Big| \bigcup_{\varepsilon_i \in B} (F', C_2) \underline{\triangle} G(\varepsilon_i) \Big|,$$

i.e., $\mathrm{SIG}(C_1) \le \mathrm{SIG}(C_2)$. ■

Theorem 2 shows that the significance of an attribute subset monotonically increases with the number of attributes, which means that adding a new attribute in an attribute subset at least does not decrease its significance. This property is very important for parameter reduction.

4.3. Parameter reduction. Based on the definition of significance of attribute subset, the reduct attribute set, optimal reduct attribute set, core attribute, and core attribute set can be defined.

Definition 12. (*Reduct attribute set*) If $((F', A), (G, B), U)$ is an incomplete soft decision system and $C \subseteq A$, then C is a reduct attribute set of $((F', A), (G, B), U)$ if

$$\mathrm{SIG}(C) = \mathrm{SIG}(A). \tag{12}$$

Definition 13. (*Optimal reduct attribute set*) If, for any subset of C, $D \subset C \subseteq A$, $\mathrm{SIG}(D) < \mathrm{SIG}(C) = \mathrm{SIG}(A)$, then C is an optimal reduct attribute set of soft decision system $((F', A), (G, B), U)$.

Example 7 can be reconsidered and the significance of attribute subset $A_1 = \{e_2, e_3, e_4, e_5\}$ can be computed as follows:

$$(F', A_1) = \{F'(e_2), F'(e_3), F'(e_4), F'(e_5)\},$$

when

$$\begin{aligned} F'(e_2) &= \{h_1, h_4, h_6\} \cup \{h_2\} \\ &= \{h_1, h_2, h_4, h_6\}, \end{aligned}$$

$$F'(e_3) = \{h_4\} \cup \{h_6\} = \{h_4, h_6\},$$
$$F'(e_4) = \{h_4, h_6\} \cup \{h_3, h_5\} = \{h_3, h_4, h_5, h_6\},$$
$$F'(e_5) = \{h_1, h_4, h_6\} \cup \emptyset = \{h_1, h_4, h_6\}.$$

Then

$$\begin{aligned}\mathrm{SIM}(A_1) = \{&(h_1, h_1), (h_2, h_2), (h_3, h_3), (h_4, h_4),\\ &(h_5, h_5), (h_6, h_6), (h_2, h_3), (h_3, h_2),\\ &(h_2, h_5), (h_5, h_2), (h_3, h_5), (h_5, h_3),\\ &(h_4, h_6), (h_6, h_4)\},\end{aligned}$$

$$\begin{aligned}&S_{A_1}(h_1) = \{h_1\}, && S_{A_1}(h_2) = \{h_2, h_3, h_5\},\\ &S_{A_1}(h_3) = \{h_2, h_3, h_5\}, && S_{A_1}(h_4) = \{h_4, h_6\},\\ &S_{A_1}(h_5) = \{h_2, h_3, h_5\}, && S_{A_1}(h_6) = \{h_4, h_6\}.\end{aligned}$$

$$\begin{aligned}\mathrm{SIG}(A_1) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_1) \underline{\triangle} G(\varepsilon_i)\Big|\\ &= |(F', A_1)\underline{\triangle} G(\varepsilon_1) \cup (F', A_1)\underline{\triangle} G(\varepsilon_2)|\\ &= |\emptyset \cup \{h_1, h_4, h_6\}| = 3 < \mathrm{SIG}(A).\end{aligned}$$

Thus, A_1 is NOT a reduct attribute subset of $((F', A), (G, B), U)$. Reconsider the attribute subset $A_2 = \{e_1, e_3, e_4, e_5\}$, and compute the significance of A_2 according to Definition 11:

$$(F', A_2) = \{F'(e_1), F'(e_3), F'(e_4), F'(e_5)\},$$

where

$$F'(e_1) = \{h_5\} \cup \{h_3\} = \{h_3, h_5\},$$
$$F'(e_3) = \{h_4\} \cup \{h_6\} = \{h_4, h_6\},$$
$$F'(e_4) = \{h_4, h_6\} \cup \{h_3, h_5\} = \{h_3, h_4, h_5, h_6\},$$
$$F'(e_5) = \{h_1, h_4, h_6\} \cup \emptyset = \{h_1, h_4, h_6\}\}.$$

Then

$$\begin{aligned}\mathrm{SIM}(A_2) = \{&(h_1, h_1), (h_2, h_2), (h_3, h_3), (h_4, h_4),\\ &(h_5, h_5), (h_6, h_6), (h_2, h_3), (h_3, h_2), (h_3, h_5),\\ &(h_5, h_3), (h_4, h_6), (h_6, h_4)\},\end{aligned}$$

$$\begin{aligned}&S_{A_2}(h_1) = \{h_1\}, && S_{A_2}(h_2) = \{h_2, h_3\},\\ &S_{A_2}(h_3) = \{h_2, h_3, h_5\}, && S_{A_2}(h_4) = \{h_4, h_6\},\\ &S_{A_2}(h_5) = \{h_3, h_5\}, && S_{A_2}(h_6) = \{h_4, h_6\}.\end{aligned}$$

$$\begin{aligned}\mathrm{SIG}(A_2) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_2) \underline{\triangle} G(\varepsilon_i)\Big|\\ &= |(F', A_2)\underline{\triangle} G(\varepsilon_1) \cup (F', A_2)\underline{\triangle} G(\varepsilon_2)|\\ &= |\{h_5\} \cup \{h_1, h_4, h_6\}| = 4 = \mathrm{SIG}(A).\end{aligned}$$

Thus, A_2 is a reduct attribute subset of $((F', A), (G, B), U)$, but it may be not the optimal reduct attribute set, because there may be a subset $A_{2*} \subset A_2$ and $\mathrm{SIG}(A_{2*}) = 4$.

Based on the definition of the reduction of an incomplete soft decision system, we give the definitions of the core attribute and the core attribute set of an incomplete soft decision system as follows.

Definition 14. (*Core attribute*) An attribute e is a core attribute of an incomplete soft decision system if it belongs to every reduction of the incomplete soft decision system $((F', A), (G, B), U)$.

Definition 15. (*Core attribute set*) An attribute set $C\,(C \subseteq A)$ is a core attribute set of incomplete soft decision system $((F', A), (G, B), U)$ if all of the elements in C are core attributes of incomplete soft decision system $((F', A), (G, B), U)$.

4.4. Decision rules. Decision rules of the soft decision system $((F, A), (G, B), U)$ can be established according to the attributes set A as follows:

$$\wedge\, (e_i, v) \to \vee(\varepsilon_i, w), \tag{13}$$

where $e_i \in A, \varepsilon_i \in B, v = F(e_i)/h_i, w = G(\varepsilon_i)/h_i$, $\wedge$ means "and", $\vee$ means "or", $\wedge(e_i, v)$ is the condition part of the rule, $\vee(\varepsilon_i, w)$ is the decision part of the rule.

Example 10. (*Decision rules from soft decision systems*) In Example 7, the decision rules of $((F, A), (G, B), U)$ from the attributes set A are

$r_1 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 0) \wedge (e_4, 0) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house),

$r_2 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 0) \wedge (e_4, 0) \wedge (e_5, 0) \to (\varepsilon_2, 1)$ (attractive house),

$r_3 : (e_1, 1) \wedge (e_2, 0) \wedge (e_3, 0) \wedge (e_4, 0) \wedge (e_5, 0) \to (\varepsilon_1, 1)$ (unattractive house),

$r_4 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 1) \wedge (e_4, 1) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house),

$r_5 : (e_1, 1) \wedge (e_2, 0) \wedge (e_3, 0) \wedge (e_4, 1) \wedge (e_5, 0) \to (\varepsilon_1, 1)$ (unattractive house),

$r_6 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 1) \wedge (e_4, 1) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house).

For an incomplete soft decision system, the decision rules can be obtained from the attributes set using the same method.

Example 11. (*Decision rules from incomplete soft decision systems*) In Example 8, the decision rules of $((F', A), (G, B), U)$ from the attributes set A are

$r_1 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 0) \wedge (e_4, 0) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house),

$r_2 : (e_1, 0) \wedge (e_2, *) \wedge (e_3, 0) \wedge (e_4, 0) \wedge (e_5, 0) \to (\varepsilon_2, 1)$ (attractive house),

$r_3 : (e_1, *) \wedge (e_2, 0) \wedge (e_3, 0) \wedge (e_4, *) \wedge (e_5, 0) \to (\varepsilon_1, 1)$ (unattractive house),

$r_4 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, 1) \wedge (e_4, 1) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house),

$r_5 : (e_1, 1) \wedge (e_2, 0) \wedge (e_3, 0) \wedge (e_4, *) \wedge (e_5, 0) \to (\varepsilon_1, 1)$ (unattractive house),

$r_6 : (e_1, 0) \wedge (e_2, 1) \wedge (e_3, *) \wedge (e_4, 1) \wedge (e_5, 1) \to (\varepsilon_2, 1)$ (attractive house).

where '*' denotes incomplete information in $((F', A), (G, B), U)$.

The decision rules generated above used all the information in attribute set A. When there is too much or redundant information contained in A, such a decision rules generating method cannot be efficient in solving a practical MCDM problem. On the contrary, based on the optimal reduct attribute set of a soft information system we propose, the optimal decision rules that not only cast off redundant information but also improve decision making efficiency can be derived. ♦

5. Method based on incomplete soft sets

5.1. Algorithm and a numerical example. In this section, a new decision-making method of MCDM problems that contain redundant and incomplete information based on incomplete soft sets is proposed and illustrated with a numerical example. The new approach can not only tackle the problem of the inadequacy of parameterization in practical MCDM problems, but also provide a direct way to process redundant and incomplete information simultaneously in MCDM.

For the incomplete information in MCDM, it is not crucial or even necessary to fill each missing datum with an arbitrarily determined data filling mechanism in the process of decision making. The optimal objects can be deduced by a series of optimal decision rules proposed in this paper without data filling.

For the redundant information in MCDM, parameter reduction is the key step for decision-making. The parameter reduction rules developed in this paper are based on the significance of each attribute subset in the incomplete soft set defined above. These rules can be used for both complete and incomplete soft sets.

The algorithm to solve MCDM problems with both redundant and incomplete information based on an incomplete soft set is given below. This new algorithm can be applied to Example 2.

Algorithm 1. Algorithm to solve MCDM problems with redundant and incomplete information based on incomplete soft set.

Step 1. Construct an incomplete soft decision system $((F', A), (G, B), U)$.

Step 2. Calculate $\mathrm{SIG}(A)$ according to Definition 11.

Step 3. Calculate $\mathrm{SIG}(A_i)$, where A_i is an attribute subset in $((F', A), (G, B), U)$.

Step 4. Find the optimal reduct attribute set according to Definition 13.

Step 5. Derive the optimal decision rules and make a decision.

In Step 1, an incomplete soft decision system named $((F', A), (G, B), U)$ is established through the initial data set.

In Step 2, the significance of the attribute set A in $((F', A), (G, B), U)$ is given by

$$\begin{aligned}\mathrm{SIG}(A) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A) \underline{\triangle} G(\varepsilon_i)\Big| \\ &= |(F', A) \underline{\triangle} G(\varepsilon_1) \cup (F', A) \underline{\triangle} G(\varepsilon_2)| \\ &= |\{h_5\} \cup \{h_1, h_4, h_6\}| = 4.\end{aligned}$$

In Step 3, the significance of each attribute subset in $((F', A), (G, B), U)$ can be calculated in the same way:

$$\begin{aligned}A_{2345} &= \{e_1\}, \\ \mathrm{SIG}(A_{2345}) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_{2345}) \underline{\triangle} G(\varepsilon_i)\Big| \\ &= |\{h_5\} \cup \emptyset| = 1 \neq 4; \\ A_{1345} &= \{e_2\}, \\ \mathrm{SIG}(A_{1345}) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_{1345}) \underline{\triangle} G(\varepsilon_i)\Big| \\ &= |\emptyset \cup \{h_1, h_4, h_6\}| = 3 \neq 4; \\ A_{1245} &= \{e_3\}, \\ \mathrm{SIG}(A_{1245}) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_{1245}) \underline{\triangle} G(\varepsilon_i)\Big| \\ &= |\emptyset \cup \{h_4\}| = 1 \neq 4; \\ A_{1235} &= \{e_4\}, \\ \mathrm{SIG}(A_{1235}) &= \Big|\bigcup_{\varepsilon_i \in B} (F', A_{1235}) \underline{\triangle} G(\varepsilon_i)\Big| \\ &= |\emptyset \cup \emptyset| = 0 \neq 4; \\ A_{1234} &= \{e_5\},\end{aligned}$$

$$\begin{aligned}\mathrm{SIG}(A_{1234}) &= \left|\bigcup_{\varepsilon_i\in B}(F',A_{1234})\underline{\triangle}G(\varepsilon_i)\right|\\ &= |\emptyset\cup\{h_1,h_4,h_6\}| = 3\neq 4;\\ A_{234} &= \{e_1,e_5\},\\ \mathrm{SIG}(A_{234}) &= \left|\bigcup_{\varepsilon_i\in B}(F',A_{234})\underline{\triangle}G(\varepsilon_i)\right|\\ &= |\{h_5\}\cup\{h_1,h_4,h_6\}| = 4.\end{aligned}$$

In Step 4, the optimal reduct attribute set of $((F',A),(G,B),U)$ can be found. According to Definitions 12 and 13, it can be concluded that attribute subset $A_{234}=\{e_1,e_5\}$ is an optimal reduct attribute set of $((F',A),(G,B),U)$.

In Step 5, the optimal decision rules can be derived as follows:

$r_1 : (e_1,0)\wedge(e_5,1)\to(\varepsilon_2,1)$ (attractive house),

$r_2 : (e_1,0)\wedge(e_5,0)\to(\varepsilon_2,1)$ (attractive house),

$r_3 : (e_1,*)\wedge(e_5,0)\to(\varepsilon_1,1)$ (unattractive house),

$r_4 : (e_1,0)\wedge(e_5,1)\to(\varepsilon_2,1)$ (attractive house),

$r_5 : (e_1,1)\wedge(e_5,0)\to(\varepsilon_1,1)$ (unattractive house),

$r_6 : (e_1,0)\wedge(e_5,1)\to(\varepsilon_2,1)$ (attractive house).

That is to say, if a house is not cheap or $e_1 = 0$, then it is attractive. If a house is cheap or its price is unknown, which means $e_1 = 1$ or $e_1 = *$, then it is unattractive. Therefore, it can be inferred that h_1, h_2, h_4, and h_6 in Example 2 are attractive houses, while h_3 and h_5 are unattractive ones. The rules can also be applied to other alternative houses.

5.2. Comparative analysis. The advantages of this approach to solve MCDM problems with redundant and incomplete information can be demonstrated by comparing it with other similar methods. One of them is the data analysis approach of soft sets under incomplete information (DASI) proposed by Zou and Xiao (2008), and another is the data filling approach for incomplete soft sets (DFIS) developed by Qin *et al.* (2012). To deal with incomplete information in MCDM, both of these two methods rely on data filling to transform the incomplete information system into a complete one.

5.2.1. Results from DASI. Here, $\forall e\in A$, and let $p_e = n_1/(n_1+n_0)$ stand for the probability that an object belongs to $F(e)$, where n_1 and n_0 are the number of objects that do and do not belong to $F'(e)$, respectively.

Zou and Xiao (2008) proposed to replace the missing value $h_{ij} = *$ by p_e. Then, all the alternatives are ranked by their decision values which are the summation of all values under each attribute, and the alternatives with the highest decision value are selected as the optimal.

Example 2 can be reconsidered, and its tabular representation is given in Table 2. By using the DASI method, for object h_1, there is no incomplete data, and thus its decision values $d_1 = 2$; for object h_2, information about attribute e_2 is missing. We have $p_e = 3/5 = 0.6$; thus, $h_{21} = 0.6$ and $d_2 = 0.6$. Following the same method, the decision values for other objects can be calculated, and all the decision values are listed in Table 3. Then, it can be concluded that h_4 is Mr. X's favorite house because it gets the highest decision value.

5.2.2. Results from DFIS. Qin *et al.* (2012) pointed out that each choice value of objects is independent of those of other objects. Consequently, it is inappropriate to use the distribution of other available objects to decide the weight of each possible choice value. To overcome these problems, Qin *et al.* (2012) proposed the novel DFIS, which is based on the maximum association degree of the parameter.

In an incomplete soft set (F',A), $F(e_i)(x), e_i\in A, x\in U$ denotes the information of objects x on attribute e_i and supposes it is missing, that is, $F(e_i)(x) = *$. Then, according to the DFIS method proposed by Qin *et al.* (2012), $F(e_i)(x)$ can be filled based on the maximum association degree of e_i as follows: $\forall e_i\in A$ and $\forall e_j\in A, i\neq j$. The set U_{ij} that yields the objects with complete information on e_i and e_j is given by

$$U_{ij} = \{x|F(e_i)(x)\neq *\wedge F(e_j)(x)\neq *, x\in U\}. \quad (14)$$

Let CN_{ij} and IN_{ij} denote the consistent association number and the inconsistent association number between e_i and e_j, respectively, where

$$\mathrm{CN}_{ij} = |\{x|F(e_i)(x) = F(e_j)(x), x\in U_{ij}\}|, \quad (15)$$

$$\mathrm{IN}_{ij} = |\{x|F(e_i)(x)\neq F(e_j)(x), x\in U_{ij}\}|, \quad (16)$$

and $|\cdot|$ signifies the number of elements in a set.

Then the consistent association degree CN_{ij} and inconsistent association degree ID_{ij} between e_i and e_j can be obtained as follows:

$$\mathrm{CN}_{ij} = \frac{\mathrm{CN}_{ij}}{|U_{ij}|}, \quad (17)$$

$$\mathrm{ID}_{ij} = \frac{\mathrm{IN}_{ij}}{|U_{ij}|}. \quad (18)$$

Table 3. Decision value of each object.

U	h_1	h_2	h_3	h_4	h_5	h_6
d	2	0.6	0.7	4	1.5	3.2

The association degree D_{ij} between e_i and e_j, is defined as

$$D_{ij} = \max\{\mathrm{CD}_{ij}, \mathrm{ID}_{ij}\}. \tag{19}$$

Then the maximal association degree of parameter e_i is

$$D_i = \max\{D_{ij}\}. \tag{20}$$

A threshold λ can be set to judge whether the association degree between e_i and other parameters is big enough. If $D_i \geq \lambda$, then parameter e_j, which has the maximal association degree with e_i, can be used to fill the missing value of e_i. If there is a consistent association between e_i and e_j, then $F(e_i)(x) = F(e_j)(x)$; otherwise, there is an inconsistent association between e_i and e_j, and $F(e_i)(x) = 1 - F(e_j)(x)$. If $D_i<\lambda$, then the missing value $F(e_i)(x)$ is calculated by DASI.

The DFIS method can be applied to the data set of Example 2, and the association degrees for the incomplete soft set (F', E) are listed in Table 4. According to DFIS, for example, the missing value of h_{31} can be filled by the maximal association degrees of e_1. Apparently, $D_{12} = 1$ is the maximal association degree. Therefore, h_{31} can be filled according to the value of h_{32}. Because $h_{32} = 0$ and there is an inconsistent association between parameters e_1 and e_2, we assign 1 to h_{31}. Similarly, the other missing value can be filled, and then the decision values of each house can be obtained.

Although DFIS is an improvement of DASI, one important drawback is that there might be more than one maximum association degree, which leads to inconsistent filling of the missing values. For example, the value of h_{22} is missing. From Table 4 we can see that $D_{21} = D_{25} = 1$ are both the maximal association degree. Because $h_{21} = 0$ and there is an inconsistent association between e_2 and e_1, we assign 1 to h_{22}. However, $h_{25} = 0$, and there is a consistent association between e_2 and e_5, so we should assign 0 to h_{22}. Therefore, the value that should be assigned to h_{22} is undetermined. Under this circumstance, the chosen value of each object in Example 2 cannot be calculated and the optimal object cannot be found through DFIS.

5.2.3. Comparison. This section compares the three decision-making methods based on incomplete soft sets. Firstly, the approaches to handle incomplete information are different. Both DASI (Zou and Xiao, 2008) and DFIS (Qin *et al.*, 2012) have their own data filling method to compensate for the missing values of each object. However, DASI is based on the binomial distribution, and the unique maximal association degree is required for the implementation of DFIS. In the process of incomplete data analysis developed in this paper, objects with incomplete information can be directly compared by using the similarity relationship, there is no need to fill the missing values, and the unreliable assumptions associated with a data filling method can be avoided. Secondly, DASI and DFIS do not make full use of the decision values in the process of decision-making, and there is no straightforward connection between the condition attributes and the decision attributes. The decision-making method proposed in this paper can generate decision rules by making a straightforward connection between the condition parameters and decision attributes. Thirdly, in DASI and DFIS, the optimal objects can be selected by ranking their decision values but alternatives cannot be classified accurately. In the proposed method, a series of precise decision rules can be derived, and alternatives can be classified accurately as required.

Table 4. Association degrees for (F', E).

	e_1	e_2	e_3	e_4	e_5
e_1		1.00I	0.50I	0.50I	0.80I
e_2	1.00I		0.75C	0.67C	1.00C
e_3	0.50I	0.75C		1.00C	0.80C
e_4	0.50I	0.67C	1.00C		0.75C
e_5	0.80I	1.00C	0.80C	0.75C	

6. Application in the evaluation of regional food safety

This section describes the application of the proposed decision-making method to evaluate the regional food safety situation of Chongqing, China. We obtained the inspection results of 40 districts of Chongqing on 12 attributes regarding food safety from 2018, including Pesticides and Veterinary Drugs Residual (PVR), Non-Edible Substances (NES), Food Additives Abuse (FAA), and so on. According to the Satisfaction Level (SL) of local people about food safety issues, the 40 districts were divided into a safe food group (SL$\geq$0.8) and an unsafe food group (SL$<$0.8). Our research question was how to predict the regional SL regarding food safety when we had data about food safety inspection.

To make the data suitable for the construction of an incomplete soft set defined in this paper, the raw data were processed as follows. Firstly, if the inspection passing rate was not 100%, then it was recognized as below the standard, and the value 0 was assigned to the attribute; otherwise, the attribute had a value of 1. Secondly, because there was no missing value in the raw data, we randomly dropped some values to create an incomplete data set. Thirdly, we randomly selected 30 districts to train the model and used the remaining 10 districts for prediction. With the corresponding satisfaction of food safety, we built an incomplete soft decision system. Then, by using the MCDM algorithm proposed in this paper, the significance of each attribute subset and the optimal parameter reduction could be obtained.

The algorithm was programmed and implemented in PYTHON 3.6. Results show that there were five parameters in the optimal reduct attribute set, including Non-Edible Substance (NES), Heavy Metals and Other Elemental Pollutants (HM&O), Microbial Pollution (MP), Quality Index (QI), and Other Contaminants (OC). The decision rules were as follows:

(i) If the NES is below standard, then the district is classified as unsafe.

(ii) If the NES reaches the standard level but others are below standard, then the district is classified as unsafe.

(iii) If the NES and OC reach the standard level, then the district cannot be classified accurately.

(iv) If the NES and HM&O reach the standard level, then the district is classified as safe.

(v) If the MP is below standard but NES and QI reach the standard level, then the district is classified as safe.

These decision rules were utilized to forecast the food safety SL of the remaining ten districts. The forecasting results are given in Table 5. We also used DASI to analyze the same problem. However, because of the non-uniqueness of the maximal association degree, DFIS could not be used to evaluate the regional food safety based on the data we collected; therefore, we only compare the results of DASI with the proposed method. It is obvious that our method outperformed DASI.

7. Conclusion

Based on the research of Molodtsov (1999), this paper proposed a new method to solve MCDM problems with incomplete and redundant information. The new method is based on the concepts of the incomplete soft set and the incomplete soft decision system developed in this paper. Based on the basic definitions about the incomplete soft set, the binary relationship (binary similarity relation) of objects in an incomplete soft set was discussed, and some operations, such as the restricted/relaxed AND operation on an incomplete soft set and a subset of the universe, were defined. After that, the definition of the significance of an attribute subset in an incomplete soft decision system was proposed. Following this definition, we obtained the definitions of a reduct attribute set, an optimal reduct attribute set and core attributes of an incomplete soft decision system. According to the optimal reduct attribute set, the optimal decision rules could be derived. Finally, the incomplete soft set-based decision making algorithm to deal with MCDM that contains incomplete and redundant information was proposed and illustrated with an example.

Compared with other methods, the results demonstrated the capability of the incomplete soft set to explore data effectively and avoid information loss or distortion caused by data deleting or filling. We also applied this new method to evaluate food safety in Chongqing, China. However, it should be noted that the application of the method proposed by this paper could also be extended to a wide range of areas, such as feature selection and forecasting problems.

Table 5. Comparative results.

Method	Forecasting accuracy (%)
DASI	0.67
Our method	0.9

Acknowledgment

This work has been supported by the National Natural Science Foundation of China (grant no. 61902370), the Natural Science Foundation of Chongqing (grant no. cstc2020jcyj-msxmX0945) and the Chongqing Municipal Education Commission (grant no. KJQN20200305).

References

Akram, M., Shumaiza and Arshad, M. (2020). Bipolar fuzzy TOPSIS and bipolar fuzzy ELECTRE-I methods to diagnosis, *Computational and Applied Mathematics* **39**(1), Article no. 7.

Alkhazaleh, S. and Salleh, A.R. (2012). Generalised interval-valued fuzzy soft set, *Journal of Applied Mathematics* **2012**, Article no. 870504.

Chen, D.G., Tsang, E.C.C., Yeung, D.S. and Wang, X.Z. (2005). The parameterization reduction of soft sets and its applications, *Computers and Mathematics with Applications* **49**(5–6): 757–763.

Das, S., Kar, M.B., Kar, S. and Pal, T. (2018). An approach for decision making using intuitionistic trapezoidal fuzzy soft set, *Annals of Fuzzy Mathematics and Informatics* **16**(1): 99–116.

Deng, T. and Wang, X. (2013). An object-parameter approach to predicting unknown data in incomplete fuzzy soft sets, *Applied Mathematical Modelling* **37**(6): 4139–4146.

Garg, H. and Arora, R. (2018). Bonferroni mean aggregation operators under intuitionistic fuzzy soft set environment and their applications to decision-making, *Journal of the Operational Research Society* **69**(11): 1711–1724.

Gau, W.L. and Buehrer, D.J. (1993). Vague sets, *IEEE Transactions on Systems Man and Cybernetics* **23**(2): 610–614.

Goguen, J.A. (1967). L-fuzzy sets, *Journal of Mathematical Analysis and Applications* **18**(1): 145–174.

Gong, K., Wang, P. and Peng, Y. (2017). Fault-tolerant enhanced bijective soft set with applications, *Applied Soft Computing* **54**: 431–439.

Gong, K., Xiao, Z. and Zhang, X. (2010). The bijective soft set with its operations, *Computers and Mathematics with Applications* **60**(8): 2270–2278.

Hong, D.H. and Choi, C.H. (2000). Multicriteria fuzzy decision-making problems based on vague set theory, *Fuzzy Sets and Systems* **114**(1): 103–113.

Inbarani, H.H., Kumar, S.U., Azar, A.T. and Hassanien, A.E. (2018). Hybrid rough-bijective soft set classification system, *Neural Computing and Applications* **29**(8): 67–78.

Jiang, Y., Tang, Y., Chen, Q., Liu, H. and Tang, J. (2010). Interval-valued intuitionistic fuzzy soft sets and their properties, *Computers and Mathematics with Applications* **60**(3): 906–918.

Khan, A. and Zhu, Y. (2020). New algorithms for parameter reduction of intuitionistic fuzzy soft sets, *Computational and Applied Mathematics* **39**(3), Aricle no. 232.

Kong, Z., Gao, L.Q., Wang, L.F. and Li, S. (2008). The normal parameter reduction of soft sets and its algorithm, *Computers and Mathematics with Applications* **56**(12): 3029–3037.

Kryszkiewicz, M. (1999). Rules in incomplete information systems, *Information Sciences* **113**(3–4): 271–292.

Li, M.-Y., Fan, Z.-P. and You, T.-H. (2018). Screening alternatives considering different evaluation index sets: A method based on soft set theory, *Applied Soft Computing* **64**: 614–626.

Li, Z., Wen, G. and Xie, N. (2015). An approach to fuzzy soft sets in decision making based on grey relational analysis and Dempster–Shafer theory of evidence: An application in medical diagnosis, *Artificial Intelligence in Medicine* **64**(3): 161–71.

Liu, Y., Qin, K., Rao, C. and Mahamadu Alhaji, M. (2017). Object-parameter approaches to predicting unknown data in an incomplete fuzzy soft set, *International Journal of Applied Mathematics and Computer Science* **27**(1): 157–167, DOI: 10.1515/amcs-2017-0011.

Maji, P.K. and Roy, A.R. (2002). An application of soft sets in a decision making problem, *Computers and Mathematics with Applications* **44**(8-9): 1077–1083.

Majumdar, P. and Samanta, S.K. (2010). Generalised fuzzy soft sets, *Computers and Mathematics with Applications* **59**(4): 1425–1432.

Meng, D., Zhang, X. and Qin, K. (2011). Soft rough fuzzy sets and soft fuzzy rough sets, *Computers and Mathematics with Applications* **62**(12): 4635–4645.

Molodtsov, D. (1999). Soft set theory—first results, *Computers and Mathematics with Applications* **37**(4–5): 19–31.

Pawlak, Z. (1984). Rough classification, *International Journal of Man-Machine Studies* **20**(5): 469–483.

Pawlak, Z. (1985). Rough sets and decision tables, *in* A. Skowron (Ed.), *Computation Theory. SCT 1984*, Lecture Notes in Computer Science, Vol. 208, Springer, Berlin, pp. 187–196.

Peng, X. and Yang, Y. (2017). Algorithms for interval-valued fuzzy soft sets in stochastic multi-criteria decision making based on regret theory and prospect theory with combined weight, *Applied Soft Computing* **54**: 415–430.

Petchimuthu, S., Garg, H., Kamaci, H. and Atagun, A.O. (2020). The mean operators and generalized products of fuzzy soft matrices and their applications in MCGDM, *Computational and Applied Mathematics* **39**(2), Article no. 68.

Qin, H., Ma, X., Herawan, T. and Zain, J.M. (2012). DFIS: A novel data filling approach for an incomplete soft set, *International Journal of Applied Mathematics and Computer Science* **22**(4): 817–828, DOI: 10.2478/v10006-012-0060-3.

Roy, A.R. and Maji, P.K. (2007). A fuzzy soft set theoretic approach to decision making problems, *Journal of Computational and Applied Mathematics* **203**(2): 412–418.

Sun, B., Zhang, M., Wang, T. and Zhang, X. (2020). Diversified multiple attribute group decision-making based on multigranulation soft fuzzy rough set and TODIM method, *Computational and Applied Mathematics* **39**(3), Article no. 186.

Tiwari, V., Jain, P.K. and Tandon, P. (2017). A bijective soft set theoretic approach for concept selection in design process, *Journal of Engineering Design* **28**(2): 100–117.

Tiwari, V., Jain, P.K. and Tandon, P. (2019). An integrated Shannon entropy and TOPSIS for product design concept evaluation based on bijective soft set, *Journal of Intelligent Manufacturing* **30**(4): 1645–1658.

Xu, W., Pan, Y., Chen, W. and Fu, H. (2019). Forecasting corporate failure in the Chinese energy sector: A novel integrated model of deep learning and support vector machine, *Energies* **12**(12), Article no. 2251.

Yang, J. and Yao, Y. (2020). Semantics of soft sets and three-way decision with soft sets, *Knowledge-Based Systems* **194**, Article no. 105538.

Zadeh, L.A. (1965). Fuzzy sets, *Information and Control* **8**(3): 338–353.

Zhang, Z.M. and Zhang, S.H. (2013). A novel approach to multi attribute group decision making based on trapezoidal interval type-2 fuzzy soft sets, *Applied Mathematical Modelling* **37**(7): 4948–4971.

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

Sisi Xia received her PhD in economics from Chongqing University in 2013. She is currently a lecturer with the School of Economics, Southwest University of Political Science and Law. She has published several articles on soft sets in various international journals. Her research areas are fuzzy sets, soft sets, data mining and decision making.

Haoran Yang received his PhD in agricultural science from Bonn University in 2015. He is currently a lecturer with the School of Economics, Southwest University of Political Science and Law. His research interests include Bayesian statistics and optimization.

Lin Chen received his MD and PhD degrees in computer science and technology from Chongqing University (CQU), China, in 2010 and 2016, respectively. Currently, he is an assistant researcher with the Chongqing Institute of Green and Intelligent Technology (CIGIT), Chinese Academy of Sciences (CAS). His research interests include machine learning and data mining. He has published over 15 papers in domestic and foreign academic journals or conference proceedings.

Received: 23 February 2021
Revised: 21 May 2021
Accepted: 7 June 2021