AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(12): 20420–20440.
DOI: 10.3934/math.20221119
Received: 06 December 2021
Revised: 19 January 2022
Accepted: 23 January 2022
Published: 19 September 2022

*Research article*

# A soft set based approach for the decision-making problem with heterogeneous information

**Sisi Xia[1,*], Lin Chen[2] and Haoran Yang[1]**

[1] School of Economics, Southwest University of Political Science and Law, Chongqing 401120, China

[2] Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China

* **Correspondence:** Email: xiasisi@swupl.edu.cn; Tel: +8615123300009.

**Abstract:** This paper proposes the concept of a neighborhood soft set and its corresponding decision system, named neighborhood soft decision system to solve decision-making (DM) problems with heterogeneous information. Firstly, we present the definition of a neighborhood soft set by combining the concepts of a soft set and neighborhood space. In addition, some operations on neighborhood soft sets such as "restricted/relaxed AND" operations and the degree of dependency between two neighborhood soft sets are defined. Furthermore, the neighborhood soft decision system and its parameter reduction, core attribute are also defined. According to the core attribute, we can get decision rules and make the optimal decision. Finally, the algorithm of DM with heterogeneous information based on the neighborhood soft set is presented and applied in the medical diagnosis, and the comparison analysis with other DM methods is made.

**Keywords:** soft set; neighborhood soft set; decision-making problems; heterogeneous information
**Mathematics Subject Classification:** 03E75, 91B06

## 1. Introduction

Various mathematical theories, such as fuzzy set theory [6, 11], rough set theory [38, 39], vague set theory [2, 31], etc., have been proposed by researchers to deal with vagueness and uncertainty in practical problems in engineering, economics, social science, medical science and so on. However, Molodtsov [1] pointed out that the parameterization tools of the above aforementioned theories were inadequate due to inherent limitation. Then he instead proposed soft set theory, which is a new mathematical tool to deal with uncertain problems. It is free from the inherent limitations, and its parameterization tools are adequate to process uncertainties. Objects can be described based on soft set theory through establishing the mapping from parameters to universe. The result from the mapping in a soft set is not a set, but a parameterized family of subsets of universe, so parameters of soft sets can be any form such as numerical values, words, logical language, sentences and so on.

On the basis of Molodtsov's research [1], many researchers have enriched the theory of soft set through discussing some classical set theoretic laws and algebraic properties in soft set theory, including the equality of two soft sets, subset and super set of soft set, complement of a soft set, null soft set, absolute soft set, the binary operations between two soft sets, such as AND, OR, union, intersection, De Morgan's law and so on [10, 12, 15, 16, 18, 23, 28]. With the establishment and development of soft set theory, it was applied to various research fields such as data analysis and recognition [37], combined forecasting [40], information system [24], decision-making [4, 5, 17, 26, 36], evaluation [14], economic [25], medical diagnosis [42], fixed point theory [27], and feature selection [13, 30, 34]. The application of soft sets in decision-making (DM) problems, among many others, is of great interest to many researchers [8].

Vagueness and uncertainty of information widely exist in practical DM problems and prevent an appropriate decision-making. Expressing the information accurately by mathematical languages to modeling DM problems is an effective method. Therefore, the first step of an efficient decision-making is to transform the vague and uncertain information into numerical variables through mathematical languages, such as fuzzy language, vague language, rough language, interval language, etc. In addition, by exploring the advantage of soft sets in providing unlimited parameterization tools, the above mathematical languages and soft set theory were combined to study practical DM problems. For example, Yang [32] combined soft sets with interval-valued fuzzy set to deal with the DM problems with interval-valued fuzzy information. Guan [33] introduced a new order relation on fuzzy soft sets, called soft information order, by combining soft sets with fuzzy sets to analyze DM problems. Likewise, many other researches focused on proposing new DM methods by constructing different kinds of hybrid soft sets, including N-soft set [7], trapezoidal interval type-2 fuzzy soft sets [41], Pythagorean fuzzy soft sets [29] and so on. Apparently, these methods facilitated DM under different types of information environment.

However, by each hybrid soft set mentioned above, only one type of information in practical DM problems can be processed. In order to describe objects exactly and make an optimal decision, several attributes with different types of data should be used to represent alternatives comprehensively and accurately. As a consequence, the information in practical DM problems may be heterogeneous. That is to say, there may be coexistence of continuous data, discrete data, dual/Multilevel semantic data and other data types in one DM model, which cannot be processed effectively by any kinds of hybrid soft sets at hand. Therefore, the focus of this paper is to develop a new method to process heterogeneous information in DM problems in order to deliver better decisions in practice.

At present, the main method for processing the heterogeneous information in DM problems is either by limiting the description of objects under an information circumstance with only one data type, or by transforming different types of data into the same [35]. Either way, the information can be used is restricted or may be lost in the process of transformation, which may result in incorrect decisions.

Neighborhood space, introduced by Fix and Hodges [3], is a popular learning and classification technique and a more general topological space than equivalence space. Some new types neighborhoods system in classical rough set theory have been recently presented [21, 22], and it has been applied to many fields such as attribute reduction, feature selection, classification, information recognition [9, 19, 20] and so on. By using some gathering rules in neighborhood space, objects can be classified into several groups according to their similarity and some neighborhood granules can be generated. More importantly, these gathering rules and neighborhood granules can be generated from a heterogeneous information environment. That means there is no limitation on the types of data in description on objects, or no need to transfer different types of data into the same in the process of DM. On the contrary, neighborhood space can process the heterogeneous information directly. Considering soft sets have the advantage of providing adequate parameterization tools and neighborhood space is capable of classifying objects with heterogeneous information, this paper proposes a new method to deal with DM problems which contain heterogeneous information by combining soft sets with neighborhood space.

The rest of this paper is organized as follows: Section 2 introduces the basic definitions of soft sets and neighborhood space. The concept of a neighborhood soft set which is a combination of the soft set and neighborhood space is presented in section 3. Besides, the operation rules on the neighborhood soft set are discussed in this section. In section 4, the definition of a neighborhood soft decision system is proposed and the method of DM under heterogeneous information environment based on the neighborhood soft decision system is proposed, followed by an illustrative example and a comparison analysis. Section 5 applies the new method in the medical diagnosis, and the last section discusses our main conclusions.

## 2. Preliminaries

### 2.1. Soft sets

**Definition 1.** *[1] Suppose that $U$ is an initial universe set and $A$ is a set of parameters. Let $P(U)$ denote the set of all subsets of $U$, a pair $(F, A)$ is called a soft set over $U$, where $F$ is a mapping given by*

$$F : A \to P(U). \tag{2.1}$$

Clearly, a soft set is a mapping from parameters to $P(U)$, and it is not a set, but a parameterized family of subsets of $U$. For $e \in A$, $F(e)$ can be considered as the set of $e$-approximate elements of the soft set $(F, A)$.

**Example 1.** *Suppose the following:*
*i. $U$ is a set of houses under consideration;*
*ii. $A$ is a set of parameters, each parameter is a word or a sentence.*
*For example,*

$$U = \{h_1, h_2, h_3, h_4, h_5, h_6\}$$

*and*

$$A = \{e_1, e_2, e_3, e_4, e_5\} = \{cheap, beautiful, bigarea, goodlocation, inthegreensurroundings\}.$$

*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.*

*Thus, we can view the soft set $(F, A)$ as a collection of approximations as below:*

$$\begin{aligned}(F,A) = \{&cheap\ houses = \{h_2, h_4\},\\ &beautiful\ houses = \{h_1, h_3\},\\ &big\ houses = \{h_3, h_4, h_5\},\\ &good\ location\ houses = \{h_1, h_3, h_5\},\\ &in\ the\ green\ surroundings = \{h_1\}\}.\end{aligned}$$

The soft set of 'attractiveness of the houses' in Example 2.1 can also be tabulated as in Table 1.

**Table 1.** The soft set $(F, A)$.

| $U$ | 'cheap' | 'beautiful' | 'big area' | 'good location' | 'in the green surroundings' |
|---|---|---|---|---|---|
| $h_1$ | 0 | 1 | 0 | 1 | 1 |
| $h_2$ | 1 | 0 | 0 | 0 | 0 |
| $h_3$ | 0 | 1 | 1 | 1 | 0 |
| $h_4$ | 1 | 0 | 1 | 0 | 0 |
| $h_5$ | 0 | 0 | 1 | 1 | 0 |
| $h_6$ | 0 | 0 | 0 | 0 | 0 |

*2.2. Neighborhood space*

**Definition 2.** *[3] Let $U = \{x_1, x_2, \cdots, x_n\}$ be an initial universe set, and $A = \{e_1, e_2, \cdots, e_m\}$ be a set of parameters. $\forall x_i \in U$ and $B \subseteq A$, the neighborhood of $x_i$ in the subspace $B$ denoted by $\delta_B(x_i)$ is defined as:*

$$\delta_B(x_i) = \{x_j | x_j \in U, \Delta_B^P(x_i, x_j) \leq \delta\} \tag{2.2}$$

*where $\delta$ is an arbitrary small nonnegative number and $\Delta$ is a metric function which satisfies:*
*i. $\Delta\{x_i, x_j\} \geq 0$;*
*ii. $\Delta\{x_i, x_j\} = 0$, if and only if $x_i$ and $x_j$ are the same;*
*iii. $\Delta\{x_i, x_j\} = \Delta\{x_j, x_i\}$;*
*iv. $\Delta\{x_i, x_k\} \leq \Delta\{x_i, x_j\} + \Delta\{x_j, x_k\}$ and*

$$\Delta_B^P(x_i, x_j) = \left(\sum |f(x_i, e_k) - f(x_j, e_k)|^P\right)^{1/P} \ (P = 1, 2, \cdots, \infty) \tag{2.3}$$

*where $\Delta_B^P(x_i, x_j)$ is a Minkowsky distance [19], $x_i, x_j \in U$, $e_k \in B \subseteq A$, $f(x, e_k)$ is the value function of object $x$ in the $k$th dimension.*

From Definition 2, it is obvious that the family of neighborhood granules $\{\delta_B(x_i) | x_i \in U\}$ forms an elemental granule system, which gathers similar objects from the universe set, rather than partitions off it into several mutual exclusive subsets.

## 3. Neighborhood soft sets

### *3.1. The concept of a neighborhood soft set*

Based on Definition 1 and Definition 2, we can combine the concepts of soft sets and neighborhood space to get the following definition of a neighborhood soft set.

**Definition 3.** *Let $U = \{x_1, x_2, \cdots, x_n\}$ be an initial universe set and $A = \{e_1, e_2, \cdots, e_m\}$ be a set of parameters. A pair $(F_\delta, A)$ is called a neighborhood soft set over U, where $F_\delta$ is a mapping given by:*

$$F_\delta : A \to P_\delta(U), \tag{3.1}$$

where $P_\delta(U)$ denotes the set of all neighborhoods of each object in $U$. For any $x_i \in U(i = 1, 2, \cdots, n)$ and $e_k \in A(k = 1, 2, \cdots, m)$, let $\delta_{e_k}(x_i)$ denotes the neighborhood of $x_i$ on $e_k$ and

$$\delta_{e_k}(x_i) = \{x_j | x_j \in U, \Delta_{e_k}(x_i, x_j) \leq \delta\}, \tag{3.2}$$

where $j = 1, 2, \cdots, n$ and $\delta$ is a threshold parameter which defines the range of the neighborhood $\delta_{e_k}(x_i)$. $\delta$ can be defined on any type of data but should be small enough. We can calculate the range according to the Eq (2.3), and for the purpose of simplicity, let $P = 1$, then

$$\Delta_{e_k}(x_i, x_j) = |f(x_i, e_k) - f(x_j, e_k)|, \tag{3.3}$$

where $f(x_i, e_k)$ is the value function of the object $x_i$ on the parameter $e_k$.

Obviously, a neighborhood soft set is also a special case of a soft set, because it is still a mapping from parameters to the universe. For $e \in A$, $F_\delta(e)$ can be considered as the set of $e$-approximate elements of the neighborhood soft set $(F_\delta, A)$. Unlike the other subtypes of soft sets, such as fuzzy soft sets, interval-valued fuzzy soft sets and so on, which express information by a uniform type of data (the discrete data 0 or 1, or the interval number between 0 and 1), neighborhood soft sets store information by neighborhood granules, which are determined by the threshold parameter $\delta$ and are the sets of approximate objects of $x_i$ on parameter $e_k$.

According to the Definition 3, there is no restriction on the data types of the threshold parameter $\delta$. It can be any form if it can weigh the distance between two objects. We can define it according to the actual situation. For example, for numerical data, $\delta$ can be an arbitrarily small nonnegative number, which denotes the maximum acceptable difference between two objects; for semantic data, $\delta$ can be a word/sentence, which can identify the difference of objects precisely. Therefore, neighborhood soft sets can process the heterogeneous information in DM problems. The following example can make this easier to understand.

**Example 2.** *Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ be an initial universe set which represents a set of six houses, and $A = \{e_1, e_2, e_3, e_4\}$ be a parameter set which describes the status of the houses. Specifically, $e_1$ denotes the area of house; $e_2$ denotes the appearance of house; $e_3$ describes the public transportation of house; and $e_4$ represents the price of house. The data are listed in Table 2:*

**Table 2.** Status of the six alternative houses.

| | $e_1(m^2)$ | $e_2$ | $e_3$(number of buses/trains) | $e_4$(thousand $yuan/m^2$) |
|---|---|---|---|---|
| $x_1$ | 122 | beautiful | 4 | 13 |
| $x_2$ | 83 | ordinary | 7 | 11 |
| $x_3$ | 75 | beautiful | 8 | 12.5 |
| $x_4$ | 69 | ordinary | 3 | 9 |
| $x_5$ | 91 | beautiful | 7 | 12 |
| $x_6$ | 155 | beautiful | 0 | 15 |

From Table 2 we can see that the types of data which describe the six houses are different: $e_1$ and $e_4$ are continuous variables, $e_2$ is a semantic variable, $e_3$ is a discrete variable. For the four variables we specify the following threshold parameters $\delta$ respectively:

For $e_1$, let $\delta_1 = 10m^2$. According to the definition of neighborhood soft sets, it means if the difference between the area of two houses $x_i$ and $x_j$ is not bigger than 10 $m^2$, then they can be regarded as similar area, and $x_j$ should be in the neighborhood of $x_i$. Analogously, we can define that $\delta_2$= "the same", $\delta_3$= 2 bus/train routes, and $\delta_4$=1.5 thousand $yuan/m^2$.

Then according to Definition 3, a neighborhood soft set $(F_\delta, A)$ can be used to describe the six houses on the shortlist, the mapping from parameters to the universe in $(F_\delta, A)$ is given as follows:

$$(F_\delta, A) = \{F_\delta(e_1), F_\delta(e_2), F_\delta(e_3), F_\delta(e_4)\},$$

where:

$$F_\delta(e_1) = \left\{\frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\},$$

$$F_\delta(e_2) = \left\{\frac{\{x_1, x_3, x_5, x_6\}}{x_1}, \frac{\{x_2, x_4\}}{x_2}, \frac{\{x_1, x_3, x_5, x_6\}}{x_3}, \frac{\{x_2, x_4\}}{x_4}, \frac{\{x_1, x_3, x_5, x_6\}}{x_5}, \frac{\{x_1, x_3, x_5, x_6\}}{x_6}\right\},$$

$$F_\delta(e_3) = \left\{\frac{\{x_1, x_4\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_5\}}{x_3}, \frac{\{x_1, x_4\}}{x_4}, \frac{\{x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\},$$

$$F_\delta(e_4) = \left\{\frac{\{x_1, x_3, x_5\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_1, x_2, x_3, x_5\}}{x_3}, \frac{\{x_4\}}{x_4}, \frac{\{x_1, x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\}.$$

In Example 2, the raw data which describes the houses can be processed directly by the neighborhood soft set defined by this paper. It is obvious that through setting the value of $\delta$ carefully, the new mapping method under the framework of neighborhood soft sets is capable of processing various types of data, and there is no need to transform the different types of data into the same. Therefore, neighborhood soft sets can provide a holistic approach to process heterogeneous information directly and precisely. It can classify the universe $U$ into several categories through finding the neighborhood of each object.

Figure 1 illustrates an example of classification based on two parameters in the neighborhood soft set $(F_\delta, A)$: $e_1$ and $e_3$, which is a simplified version of Example 2. Consider $x_3$, its neighborhood on parameter $e_1$ with $\delta_1$ can be defined by the space between parallel dashed lines $R_1$ and $R_2$, so $x_2$, $x_4$, and of course including $x_3$ itself, are in the neighborhood of $x_3$ based on $e_1$. That means the houses $x_2$,

$x_3$ and $x_4$ in $U$ can be treated as similar to $x_3$ in terms of the area of house; similarly, $x_2$, $x_5$ and also $x_3$ itself are in the neighborhood of $x_3$ based on $e_3$ with $\delta_3$. If both $e_1$ and $e_3$ are considered, then only $x_2$ and $x_3$ are in the neighborhood of $x_3$.

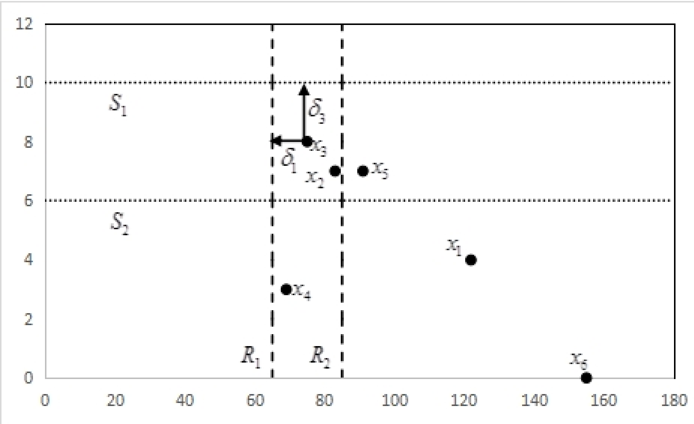


**Figure 1.** An example of classification with two parameters in the neighborhood soft set.

Based on the definition of neighborhood soft sets, the following properties of neighborhood soft sets can be derived:

**i. Reflexive**: $x_i \in \dfrac{F_\delta(e)}{x_i}$.

*Proof.* $\Delta_e(x_i, x_i) = 0 \le \delta$, then according to Definition 3, $x_i \in \delta_e(x_i)$, i.e. $x_i \in \dfrac{F_\delta(e)}{x_i}$. □

**ii. Symmetric**: $x_i \in \dfrac{F_\delta(e)}{x_j} \Leftrightarrow x_j \in \dfrac{F_\delta(e)}{x_i}$.

*Proof.* $x_i \in \dfrac{F_\delta(e)}{x_j}$, i.e. $x_i \in \delta_e(x_j)$, then $\Delta_e(x_i, x_j) \le \delta$, that is to say $x_j \in \delta_e(x_i)$. Thus, $x_j \in \dfrac{F_\delta(e)}{x_i}$. □

**iii. Nontransitive:** $x_j \in \dfrac{F_\delta(e)}{x_i}, x_k \in \dfrac{F_\delta(e)}{x_i} \nRightarrow x_k \in \dfrac{F_\delta(e)}{x_j}$ or $x_j \in \dfrac{F_\delta(e)}{x_k}$.

*Proof.* $x_j \in \dfrac{F_\delta(e)}{x_i}, x_k \in \dfrac{F_\delta(e)}{x_i}$, then $\Delta_e(x_j, x_i) \le \delta$ and $\Delta_e(x_k, x_i) \le \delta$, which implies $\Delta_e(x_k, x_j) \le 2\delta$, not $\Delta_e(x_k, x_j) \le \delta$. Therefore, the neighborhood soft set is nontransitive. A counterexample in Example 2 can be given to prove it: $x_2 \in \dfrac{F_\delta(e_1)}{x_3}, x_4 \in \dfrac{F_\delta(e_1)}{x_3}$, but $x_4 \notin \dfrac{F_\delta(e_1)}{x_2}$. □

**Definition 4.** *Let $U$ be an initial universe set and $E$ be a set of parameters. Suppose that $A, B \subseteq E$, $(F_\delta, A)$, and $(G_\delta, B)$ are two neighborhood soft sets, we say that $(F_\delta, A)$ is a neighborhood soft subset of $(G_\delta, B)$, denoted by $(F_\delta, A) \subseteq^\delta (G_\delta, B)$ if and only if:*
*i. $A \subseteq B$, and;*
*ii. $\forall e \in A$, $F_\delta(e)/x_i \subseteq G_\delta(e)/x_i$.*

$(F_\delta, A)$ is a neighborhood soft super set of $(G_\delta, B)$, denoted by $(F_\delta, A) \supseteq^\delta (G_\delta, B)$, if $(G_\delta, B)$ is a neighborhood soft subset of $(F_\delta, A)$.

**Example 3.** *Given two neighborhood soft sets $(F_\delta, A)$ and $(G_\delta, B)$, $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. Here $U$ is the set of houses on the list.*

*Let $A = \{e_1, e_2\} = \{$cheap, beautiful$\}$, $B = \{e_1, e_2, e_3\} = \{$cheap, beautiful, big area$\}$, and*

$$F_\delta(e_1) = \left\{\frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3\}}{x_2}, \frac{\{x_2, x_3\}}{x_3}, \frac{\{x_4\}}{x_4}, \frac{\{x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\},$$

$$F_\delta(e_2) = \left\{\frac{\{x_1, x_3, x_5, x_6\}}{x_1}, \frac{\{x_2, x_4\}}{x_2}, \frac{\{x_1, x_3, x_5, x_6\}}{x_3}, \frac{\{x_2, x_4\}}{x_4}, \frac{\{x_1, x_3, x_5, x_6\}}{x_5}, \frac{\{x_1, x_3, x_5, x_6\}}{x_6}\right\},$$

$$G_\delta(e_1) = \left\{\frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\},$$

$$G_\delta(e_2) = \left\{\frac{\{x_1, x_3, x_5, x_6\}}{x_1}, \frac{\{x_2, x_4\}}{x_2}, \frac{\{x_1, x_3, x_5, x_6\}}{x_3}, \frac{\{x_2, x_4\}}{x_4}, \frac{\{x_1, x_3, x_5, x_6\}}{x_5}, \frac{\{x_1, x_3, x_5, x_6\}}{x_6}\right\},$$

$$G_\delta(e_3) = \left\{\frac{\{x_1, x_4\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_5\}}{x_3}, \frac{\{x_1, x_4\}}{x_4}, \frac{\{x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6}\right\}.$$

*Therefore, we have $(F_\delta, A) \subseteq^\delta (G_\delta, B)$.*

**Definition 5.** *For two neighborhood soft sets $(F_\delta, A)$ and $(G_\delta, B)$, we say $(F_\delta, A)$ and $(G_\delta, B)$ are neighborhood soft equal, denoted by $(F_\delta, A) =^\delta (G_\delta, B)$, if and only if:*
*i. $(F_\delta, A) \subseteq^\delta (G_\delta, B)$,*
*ii. $(F_\delta, A) \supseteq^\delta (G_\delta, B)$.*

**Definition 6.** *Let $U$ be an initial universe set, $E$ be a set of parameters, and $A, B \subseteq E$. $(F_\delta, A)$ is called a null neighborhood soft set (with respect to the parameter set $A$), denoted by $\varnothing_\delta^A$, if $F_\delta(e) = \varnothing$ for all $e \in A$; $(G_\delta, B)$ is called a whole neighborhood soft set (with respect to the parameter set $B$), denoted by $U_\delta^B$, if $G_\delta(\varepsilon) = U$ for all $\varepsilon \in B$.*

However, because of the reflexive in neighborhood soft sets as proved above, i.e. $x_i \in \dfrac{F_\delta(e)}{x_i}$. So $F_\delta(e) \neq \varnothing$, the null neighborhood soft set does not exist.

### *3.2. Operations on neighborhood soft sets*

#### 3.2.1. Restricted/relaxed AND on a neighborhood soft set and a subset

**Definition 7.** *Let $U = \{x_1, x_2, \cdots, x_n\}$ be an initial universe set, $(F_\delta, A)$ be a neighborhood soft set defined on $U$ and $X$ be a subset of $U$. The operation of "$(F_\delta, A)$ restricted AND $X$" denoted by $(F_\delta, A) \wedge_\delta X$ is given by:*

$$(F_\delta, A) \wedge_\delta X = \bigcup_{e\in A} \{F_\delta(e) \wedge_\delta X\} = \bigcup_{e\in A} \left\{x_i | \frac{F_\delta(e)}{x_i} \subseteq X, x_i \in U\right\}. \tag{3.4}$$

**Example 4.** *Reconsider Example 2. Let $X = \{x_1, x_3, x_5, x_6\}$ be the set of houses which are preferred by most consumers and are the bestselling houses on the market. Then*

$$\begin{aligned}(F_\delta, A) \wedge_\delta X &= \bigcup_{e\in A} \{F_\delta(e) \wedge_\delta X\} \\ &= \{x_1, x_6\} \cup \{x_1, x_3, x_5, x_6\} \cup \{x_6\} \cup \{x_1, x_6\} \\ &= \{x_1, x_3, x_5, x_6\}.\end{aligned}$$

From the above definition, we have:

**i.** $(F_\delta, A) \wedge_\delta U = U$**.**

*Proof.* For $\forall e \in A$, and $x_i \in U$, $F_\delta(e)/x_i \subseteq U$, i.e. $(F_\delta, A) \wedge_\delta U = U$. □

**ii.** $(F_\delta, A) \wedge_\delta \varnothing = \varnothing$**.**

*Proof.* For $\forall e \in A$, and $x_i \in U$, according to the reflexive property: $x_i \in F_\delta(e)/x_i$, $F_\delta(e)/x_i \neq \varnothing$, i.e. $F_\delta(e)/x_i \nsubseteq \varnothing$. Therefore, $(F_\delta, A) \wedge_\delta \varnothing = \varnothing$. □

**iii.** $(F_\delta, A) \wedge_\delta X \subseteq X$**.**

*Proof.* For $\forall e \in A$, $X \subseteq U$ and $x_i, x_j \in U$.

For $x_i \in (F_\delta, A) \wedge_\delta X$, that means $F_\delta(e)/x_i \subseteq X$, according to the reflexive property: $x_i \in F_\delta(e)/x_i$, then $x_i \in X$.

For $x_j \in X$, it is not conclusive that $F_\delta(e)/x_j \subseteq X$, then $x_j$ may be not in $(F_\delta, A) \wedge_\delta X$.

Therefore, $(F_\delta, A) \wedge_\delta X \subseteq X$. □

**Definition 8.** *Let $U = \{x_1, x_2, \cdots, x_n\}$ be an initial universe set, $(F_\delta, A)$ be a neighborhood soft set defined on $U$ and $X$ be a subset of $U$. The operation of "$(F_\delta, A)$ relaxed AND $X$" denoted by $(F_\delta, A) \wedge^\delta X$ is given by:*

$$(F_\delta, A) \wedge^\delta X = \bigcup_{e \in A} \left\{F_\delta(e) \wedge^\delta X\right\} = \bigcup_{e \in A} \left\{x_i | \frac{F_\delta(e)}{x_i} \cap X \neq \varnothing, x_i \in U\right\}. \tag{3.5}$$

**Example 5.** *Continue Example 4:*

$$\begin{aligned}(F_\delta, A) \wedge^\delta X &= \bigcup_{e \in A} \left\{F_\delta(e) \wedge^\delta X\right\} \\ &= \{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_1, x_3, x_5, x_6\} \cup \{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_1, x_2, x_3, x_5, x_6\} \\ &= \{x_1, x_2, x_3, x_4, x_5, x_6\}.\end{aligned}$$

Similarly, from the above definition, we have:

**i.** $(F_\delta, A) \wedge^\delta U = U$**;**

**ii.** $(F_\delta, A) \wedge^\delta \varnothing = \varnothing$**;**

**iii.** $X \subseteq (F_\delta, A) \wedge^\delta X$**.**

*Proof.* Straightforward. □

3.2.2. The degree of dependency between two neighborhood soft sets

To explore the ability of classification of neighborhood soft sets, we give the definition of the degree of dependency between two neighborhood soft sets.

**Definition 9.** *Suppose that $(F_\delta, A)$ and $(G_\delta, B)$ are two neighborhood soft sets over $U$, where $A \cap B = \varnothing$. We say $(F_\delta, A)$ has a 'k degree of dependency' on $(G_\delta, B)$ denoted by $k((F_\delta, A), (G_\delta, B))$ and*

$$k((F_\delta, A), (G_\delta, B)) = \frac{\left|\bigcup_{\varepsilon_j \in B} \bigcup_{x_i \in U} \left\{(F_\delta, A) \wedge_\delta G_\delta(\varepsilon_j)/x_i\right\}\right|}{|U|}, \tag{3.6}$$

where $|*|$ denotes the number of elements in a set.

Apparently, $k((F_\delta, A), (G_\delta, B))$ is the ratio of the number of elements in two sets, one is the result of $(F_\delta, A)$ restricted AND $G_\delta(\varepsilon_j)/x_i$, the other is $U$, and:

**i**. $k(0 \leq k \leq 1)$, we say $(F_\delta, A)$ is partially depended on $(G_\delta, B)$, which measures the degree of approximation in classification between two neighborhood soft sets.

**ii**. If $k = 1$ we say $(F_\delta, A)$ is completely depended on $(G_\delta, B)$, which means the results of classification by the two neighborhood soft sets are exactly the same.

**iii**. If $k = 0$ we say $(F_\delta, A)$ is not depended on $(G_\delta, B)$, which means the results of classification by the two neighborhood soft sets are completely different.

Through the definition of the degree of dependency between two neighborhood soft sets, we can compare the similarity of classification results between two neighborhood soft sets.

**Example 6.** *Let $(F_\delta, A)$ and $(G_\delta, B)$ be two neighborhood soft sets, and*

$$F_\delta(e_1) = \left\{ \frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$F_\delta(e_2) = \left\{ \frac{\{x_1, x_3, x_5, x_6\}}{x_1}, \frac{\{x_2, x_4\}}{x_2}, \frac{\{x_1, x_3, x_5, x_6\}}{x_3}, \frac{\{x_2, x_4\}}{x_4}, \frac{\{x_1, x_3, x_5, x_6\}}{x_5}, \frac{\{x_1, x_3, x_5, x_6\}}{x_6} \right\},$$

$$F_\delta(e_3) = \left\{ \frac{\{x_1, x_4\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_5\}}{x_3}, \frac{\{x_1, x_4\}}{x_4}, \frac{\{x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$F_\delta(e_4) = \left\{ \frac{\{x_1, x_3, x_5\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_1, x_2, x_3, x_5\}}{x_3}, \frac{\{x_4\}}{x_4}, \frac{\{x_1, x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$G_\delta(\varepsilon_1) = \left\{ \frac{\{x_1, x_2, x_5, x_6\}}{x_1}, \frac{\{x_1, x_2\}}{x_2}, \frac{\{x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_1, x_5, x_6\}}{x_5}, \frac{\{x_1, x_5, x_6\}}{x_6} \right\}.$$

$$k\left((F_\delta, A), (G_\delta, B)\right) = \frac{\left| \bigcup_{\varepsilon_j \in B} \bigcup_{x_i \in U} \left\{ (F_\delta, A) \wedge_\delta (G_\delta, \varepsilon_j)/x_i \right\} \right|}{|U|} = \frac{|\{x_1, x_4, x_5, x_6\}|}{|U|} = \frac{4}{6} = \frac{2}{3}.$$

### *3.3. The neighborhood soft decision system*

#### 3.3.1. The concept of a neighborhood soft decision system

To store the heterogeneous information in DM problems and develop a DM method under a heterogeneous information environment based on neighborhood soft sets, it is necessary to introduce neighborhood soft decision systems.

**Definition 10.** *Suppose that $(F_\delta, A)$ and $(G_\delta, B)$ are two neighborhood soft sets over a common universe $U$, where $A \cap B = \emptyset$. Then the information system $(U, (F_\delta, A), (G_\delta, B))$ is called a neighborhood soft decision system over the common universe $U$, where $A$ is a condition attributes set, $B$ is a decision attributes set, and $(F_\delta, A)$ is a condition neighborhood soft set, $(G_\delta, B)$ is a decision neighborhood soft set.*

In a neighborhood soft decision system $(U, (F_\delta, A), (G_\delta, B))$, the condition attributes set $A$ is used for describing characteristics of each alternative object. The decision attributes set $B$ is used for describing the rating of each alternative object.

Then, the degree of dependency of $(U, (F_\delta, A), (G_\delta, B))$ is defined as the degree of dependency between $(F_\delta, A)$ and $(G_\delta, B)$, and is denoted by $k(U, (F_\delta, A), (G_\delta, B))(0 \leq k \leq 1)$. It provides a measure of the similarity of classification results between the condition neighborhood soft set $(F_\delta, A)$ and the decision neighborhood soft set $(G_\delta, B)$.

**Example 7.** *Reconsider Example 2. Let* $A = \{e_1, e_2, e_3, e_4\}$ *be a condition attribute set,* $(F_\delta, A)$ *is the corresponding condition neighborhood soft set.*

*Table 3 listed the rating of sold houses.*

**Table 3.** The rating of sold houses.

| | $\varepsilon_1(\%)$ |
|---|---|
| $x_1$ | 90 |
| $x_2$ | 86 |
| $x_3$ | 72 |
| $x_4$ | 76 |
| $x_5$ | 91 |
| $x_6$ | 92 |

*According to Definition 3, let* $\delta = 4\%$*, which means the two houses are in the same class if their difference of rating is no more than 4%, we get another neighborhood soft set* $(G_\delta, B)$*, where* $B = \{\varepsilon_1\}$ *is the decision attribute set. The two neighborhood soft sets* $(F_\delta, A)$ *and* $(G_\delta, B)$ *are given as follows:*

$$F_\delta(e_1) = \left\{ \frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$F_\delta(e_2) = \left\{ \frac{\{x_1, x_3, x_5, x_6\}}{x_1}, \frac{\{x_2, x_4\}}{x_2}, \frac{\{x_1, x_3, x_5, x_6\}}{x_3}, \frac{\{x_2, x_4\}}{x_4}, \frac{\{x_1, x_3, x_5, x_6\}}{x_5}, \frac{\{x_1, x_3, x_5, x_6\}}{x_6} \right\},$$

$$F_\delta(e_3) = \left\{ \frac{\{x_1, x_4\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_5\}}{x_3}, \frac{\{x_1, x_4\}}{x_4}, \frac{\{x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$F_\delta(e_4) = \left\{ \frac{\{x_1, x_3, x_5\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_1, x_2, x_3, x_5\}}{x_3}, \frac{\{x_4\}}{x_4}, \frac{\{x_1, x_2, x_3, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$G_\delta(\varepsilon_1) = \left\{ \frac{\{x_1, x_2, x_5, x_6\}}{x_1}, \frac{\{x_1, x_2\}}{x_2}, \frac{\{x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_1, x_5, x_6\}}{x_5}, \frac{\{x_1, x_5, x_6\}}{x_6} \right\}.$$

*Apparently according to Definition 10,* $(U, (F_\delta, A), (G_\delta, B))$ *can be called a neighborhood soft decision system, and its degree of dependency is:*

$$k\left(U, (F_\delta, A), (G_\delta, B)\right) = \frac{\left|\bigcup_{\varepsilon_1 \in B} \bigcup_{x_i \in U} \{(F_\delta, A) \wedge_\delta (G_\delta, \varepsilon_1)/x_i\}\right|}{|U|} = \frac{|\{x_1, x_4, x_5, x_6\}|}{|U|} = \frac{4}{6} = \frac{2}{3}.$$

### 3.3.2. The reduction of a neighborhood soft decision system

To explore which variables in the condition neighborhood soft set are decisive for DM, we give the definition of the reduction of a neighborhood soft decision system based on the degree of dependency of a neighborhood soft decision system.

**Definition 11.** *Let $(U, (F_\delta, A), (G_\delta, B))$ be a neighborhood soft decision system, $C \subseteq A$, we say attribute subset $C$ is a reduction of $(U, (F_\delta, A), (G_\delta, B))$ if:*
*i. $k\,(U, (F_\delta, C), (G_\delta, B)) = k\,(U, (F_\delta, A), (G_\delta, B))$;*
*ii. $\forall e \in C$, $k\,(U, (F_\delta, C), (G_\delta, B)) > k\,(U, (F_\delta, (C - e)), (G_\delta, B))$.*

Obviously, according to Definition 4, $(F_\delta, C)$ is a neighborhood soft subset of $(F_\delta, A)$ since $C \subseteq A$. Through Definition 11, we can compare the degree of dependency between each neighborhood soft subset of $(F_\delta, A)$ and $(G_\delta, B)$, and identify the subset $C$ which not only contains the minimum number of variables (the key attributes), but also produces the same degree of dependency between $(F_\delta, C)$ and $(G_\delta, B)$ as that between $(F_\delta, A)$ and $(G_\delta, B)$.

**Example 8.** *Consider the neighborhood soft decision system in Example 7. Our aim is to figure out which attributes in the condition attribute set $A$ are the major factors in deciding which house is the most favorite house for Mr. X. The burden to make a good choice can be reduced by getting rid of unnecessary information or attributes. Therefore, the reduction of $(U, (F_\delta, A), (G_\delta, B))$ is indispensable.*

$$\bigcup_{x \in U} \{(F_\delta, e_1) \wedge_\delta G_\delta(\varepsilon_1)/x_i\} = \{x_1, x_4, x_5, x_6\},$$

$$\bigcup_{x \in U} \{(F_\delta, e_2) \wedge_\delta G_\delta(\varepsilon_1)/x_i\} = \varnothing,$$

$$\bigcup_{x \in U} \{(F_\delta, e_3) \wedge_\delta G_\delta(\varepsilon_1)/x_i\} = \{x_6\},$$

$$\bigcup_{x \in U} \{(F_\delta, e_4) \wedge_\delta G_\delta(\varepsilon_1)/x_i\} = \{x_4, x_6\}.$$

*Then*

$$k((F_\delta, e_1), (G_\delta, B)) = k((F_\delta, A), (G_\delta, B)) = 2/3.$$

*Moreover, subset $C = \{e_1\}$ can not be reduced anymore, so it is a reduction of $(U, (F_\delta, A), (G_\delta, B))$. That is to say, the area of house is the major factor in determining which house is the best for Mr. X.*

*It should be noted that $C = \{e_1\}$ is the sole reduction of $(U, (F_\delta, A), (G_\delta, B))$, because the $k$ between the other minimal neighborhood soft subsets of $(F_\delta, A)$ and $(G_\delta, B)$ is not equal to $k((F_\delta, A), (G_\delta, B))$.*

Based on the definition of the reduction of a neighborhood soft decision system, we can define the core attribute of a neighborhood soft decision system and the core attribute set of a neighborhood soft decision system as follows:

**Definition 12.** *The attribute $e(e \in A)$ is a core of a neighborhood soft decision system if it belongs to every reduction of a neighborhood soft decision system.*

**Definition 13.** *The attribute set $C(C \subseteq A)$ is a core attribute set of a neighborhood soft decision system, if all of the elements in $C$ are the core attributes.*

**Example 9.** *In Example 8, the attribute $e_1$ belongs to the unique (every) reduction of $k((F_\delta, A), (G_\delta, B))$, so $e_1$ is a core of it.*

*In addition, in the attribute set $C = \{e_1\}$, all of its elements ($e_1$) is a core of $k((F_\delta, A), (G_\delta, B))$, so $C$ is a core attribute set of it.*

### 3.3.3. Classification rules of a neighborhood soft decision system

To make a decision based on the neighborhood soft decision system, we need to classify objects in the first place. The follows are classification rules based on the neighborhood soft decision system:

In a neighborhood soft decision system, if $F_\delta(e)/x_i \cap F_\delta(e)/x_j = \varnothing$, then $x_i$ and $x_j$ are in different groups, otherwise they are in the same group.

**Example 10.** *Consider $(U, (F_\delta, A), (G_\delta, B))$ in Example 7.*

*For the condition attribute $e_1$, we have:*

$$F_\delta(e_1) = \left\{ \frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\},$$

$$F_\delta(e_1)/x_1 = \{x_1\},$$

$$F_\delta(e_1)/x_2 = \{x_2, x_3, x_5\},$$

$$F_\delta(e_1)/x_1 \cap F_\delta(e_1)/x_2 = \varnothing.$$

*Therefore, for the condition attribute $e_1$, the two objects $x_1$ and $x_2$ are not the same. From the raw data on $e_1$, we can regard $x_1$ as a medium area house, and $x_2$ as a small area house.*

*For the decision attribute $\varepsilon_1$, we have:*

$$G_\delta(\varepsilon_1) = \left\{ \frac{\{x_1, x_2, x_5, x_6\}}{x_1}, \frac{\{x_1, x_2\}}{x_2}, \frac{\{x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_1, x_5, x_6\}}{x_5}, \frac{\{x_1, x_5, x_6\}}{x_6} \right\},$$

$$G_\delta(\varepsilon_1)/x_1 = \{x_1, x_2, x_5, x_6\},$$

$$G_\delta(\varepsilon_1)/x_2 = \{x_1, x_2\},$$

$$G_\delta(\varepsilon_1)/x_1 \cap G_\delta(\varepsilon_1)/x_2 \neq \varnothing.$$

*Therefore, for the decision attribute $\varepsilon_1$, the two objects $x_1$ and $x_2$ are the same. From the raw data on $\varepsilon_1$, we can regard $x_1$ and $x_2$ as two good-selling houses.*

### 3.3.4. Decision rules of a neighborhood soft decision system

Based on the above classification, we can get the decision rules of the neighborhood soft decision system. Through classification rules, the alternatives in a neighborhood soft decision system can be classified into several categories, each category can be seen as a decision rule, that means we can make a decision according to the relationship between the condition attributes and its decision attributes of each category. Therefore, a neighborhood soft decision system is a set of rules actually.

**Example 11.** *In Example 10, there are two categories for condition attribute $e_1$ in $(U,(F_\delta,A),(G_\delta,B))$, so we can get two decision rules:*

(1) *If the area of house is like $x_1$ (medium area house), then its rating is like $x_1$ and $x_2$ (good-selling house);*

(2) *If the area of house is like $x_2$ (small area house), then its rating is like $x_1$ and $x_2$ (good-selling house).*

That is to say, if a house is a medium area house or a small area house, then it is a good-selling house. However, we cannot decide whether a big area house is a good-selling house or a bad-selling house based on the above decision rules, and the final decision rules may be different when all objects are considered.

## 4. The method of DM under heterogeneous information environment

### *4.1. The algorithm of DM based on neighborhood soft sets*

This section attempts to demonstrate how to apply the newly developed neighborhood soft sets to DM problems with heterogeneous information. Based on the above definitions about the neighborhood soft set and the neighborhood soft decision system, the algorithm of DM under heterogeneous information environment is given as follows.

Figure 2 shows the flow of the new algorithm to get the classification rules and decision rules, and then get the optimal decision.

Step 1. Construct a neighborhood soft decision system $(U,(F_\delta,A),(G_\delta,B))$ using Eqs (3.1)–(3.3) and Definition 10;

Step 2. Calculate the degree of dependency of $(U,(F_\delta,A),(G_\delta,B))$ using Eq (3.5);

Step 3. Identify the reduction attributes of $(U,(F_\delta,A),(G_\delta,B))$ according to Definition 11;

Step 4. Derive the classification rules of $(U,(F_\delta,A),(G_\delta,B))$;

Step 5. Get the decision rules and the optimal decision.

Then we can complete Example 2.

In step 2, according to the Eq (3.5), the degree of dependency of $(U,(F_\delta,A),(G_\delta,B))$ is given by

$$k = k((F_\delta,A),(G_\delta,B)) = \frac{\left|\bigcup_{e\in B}\bigcup_{x\in U}(F_\delta,A)\wedge_\delta (G_\delta,e)/x_i\right|}{|U|} = \frac{2}{3}.$$

In step 3, the degree of dependency between the neighborhood soft subset $(F_\delta,e_i)$ and $(G_\delta,B)$ in Example 7 is given by:

$$k_1 = k((F_\delta,e_1),(G_\delta,B)) = 2/3,$$

$$k_2 = k((F_\delta,e_2),(G_\delta,B)) = 0,$$

$$k_3 = k((F_\delta,e_3),(G_\delta,B)) = 1/6,$$

$$k_4 = k((F_\delta,e_4),(G_\delta,B)) = 1/3.$$

Then according to Definition 11, we can conclude that attribute $e_1$ is the reduction of $(U,(F_\delta,A),(G_\delta,B))$, because $k_1 = k((F_\delta,e_1),(G_\delta,B)) = k((F_\delta,A),(G_\delta,B)) = 2/3$. That means the rating of sold houses is mainly determined by $e_1$, i.e. the area of a house.

In step 4, for the reduction attribute $e_1$:

$$F_\delta(e_1) = \left\{ \frac{\{x_1\}}{x_1}, \frac{\{x_2, x_3, x_5\}}{x_2}, \frac{\{x_2, x_3, x_4\}}{x_3}, \frac{\{x_3, x_4\}}{x_4}, \frac{\{x_2, x_5\}}{x_5}, \frac{\{x_6\}}{x_6} \right\}.$$

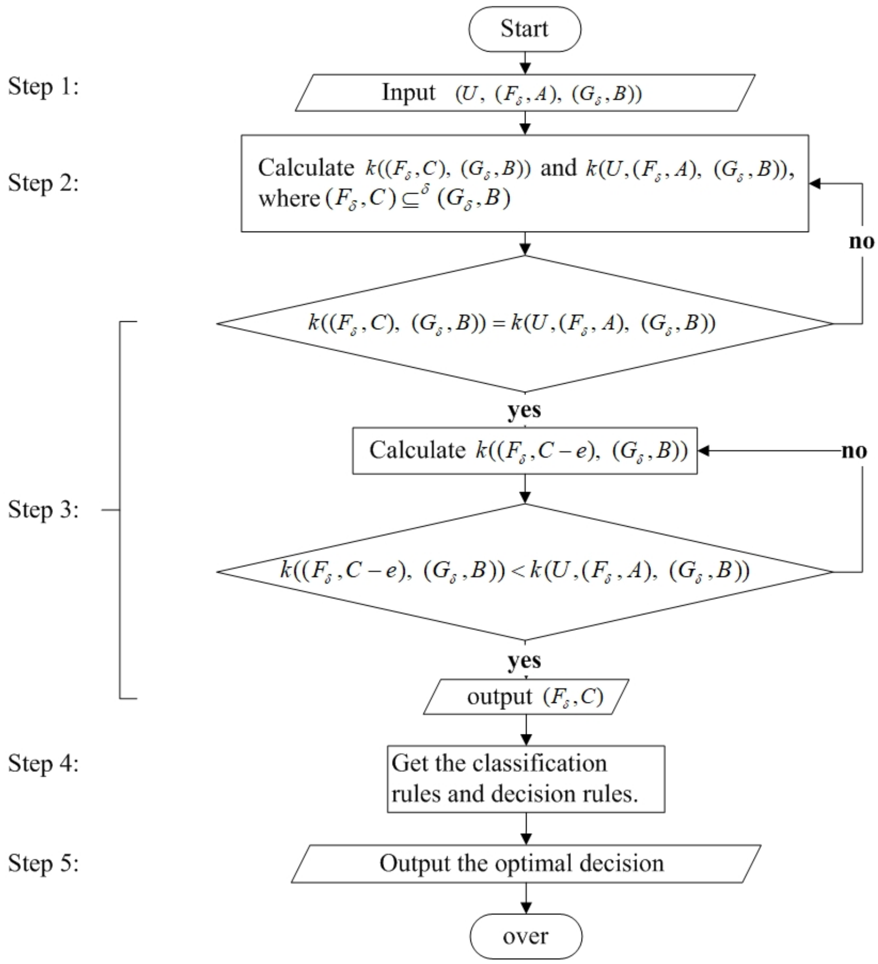


**Figure 2.** The algorithm of DM under heterogeneous information environment based on the heterogeneous soft set.

We can classify objects according to the reduction attribute $e_1$, as demonstrated in Table 4:

**Table 4.** Classification according to $e_1$.

| $F_\delta(e_1)/x_i \cap F_\delta(e_1)/x_j$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|
| $x_1$ | | $\varnothing$ | $\varnothing$ | $\varnothing$ | $\varnothing$ | $\varnothing$ |
| $x_2$ | $\varnothing$ | | $\{x_2, x_3\}$ | $\{x_3\}$ | $\{x_2, x_5\}$ | $\varnothing$ |
| $x_3$ | $\varnothing$ | $\{x_2, x_3\}$ | | $\{x_3, x_4\}$ | $\{x_2\}$ | $\varnothing$ |
| $x_4$ | $\varnothing$ | $\{x_3\}$ | $\{x_3, x_4\}$ | | $\varnothing$ | $\varnothing$ |
| $x_5$ | $\varnothing$ | $\{x_2, x_5\}$ | $\{x_2\}$ | $\varnothing$ | | $\varnothing$ |
| $x_6$ | $\varnothing$ | $\varnothing$ | $\varnothing$ | $\varnothing$ | $\varnothing$ | |

Then we get the classification result: $\{\{x_1\},\{x_2,x_3,x_4,x_5\},\{x_6\}\}$, which means the objects were categorized into three classes by condition attribute $e_1$: medium area house $x_1$, small area houses $x_2,x_3,x_4,x_5$, and large area house $x_6$.

For the decision attribute $\varepsilon_1$:

$$G_\delta(\varepsilon_1)=\left\{\frac{\{x_1,x_2,x_5,x_6\}}{x_1},\frac{\{x_1,x_2\}}{x_2},\frac{\{x_3,x_4\}}{x_3},\frac{\{x_3,x_4\}}{x_4},\frac{\{x_1,x_5,x_6\}}{x_5},\frac{\{x_1,x_5,x_6\}}{x_6}\right\}.$$

We can classify objects according to the decision attribute $\varepsilon_1$, as shown in Table 5:

**Table 5.** Classification according to $\varepsilon_1$.

| $G_\delta(\varepsilon_1)/x_i\cap G_\delta(\varepsilon_1)/x_j$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|
| $x_1$ | | $\{x_1,x_2\}$ | $\varnothing$ | $\varnothing$ | $\{x_1,x_5,x_6\}$ | $\{x_1,x_5,x_6\}$ |
| $x_2$ | $\{x_1,x_2\}$ | | $\varnothing$ | $\varnothing$ | $\{x_1\}$ | $\{x_1\}$ |
| $x_3$ | $\varnothing$ | $\varnothing$ | | $\{x_3,x_4\}$ | $\varnothing$ | $\varnothing$ |
| $x_4$ | $\varnothing$ | $\varnothing$ | $\{x_3,x_4\}$ | | $\varnothing$ | $\varnothing$ |
| $x_5$ | $\{x_1,x_5,x_6\}$ | $\{x_1\}$ | $\varnothing$ | $\varnothing$ | | $\{x_1,x_5,x_6\}$ |
| $x_6$ | $\{x_1,x_5,x_6\}$ | $\{x_1\}$ | $\varnothing$ | $\varnothing$ | $\{x_1,x_5,x_6\}$ | |

Then we get the classification result: $\{\{x_1,x_2,x_5,x_6\},\{x_3,x_4\}\}$, which means decision attribute $\varepsilon_1$ classifies the objects into two groups: Good-selling houses $x_1,x_2,x_5,x_6$, and poor-selling houses $x_3,x_4$.

And the final decision rules can be gotten as follows:

1) If the area of a house is medium, then it is a best-selling house;

2) If the area of a house is small, then it may be a best-selling house or may not;

3) If the area of a house is large, then it is a best-selling house.

In step 5, according to the above rules, the optimal choice(s) for Mr. X is to buy a medium area house or a large area house, depends on his budgets, because both of them are good-selling houses.

### *4.2. Comparison analysis*

DM methods based on soft sets have been analyzed by many researchers [4,5,17,36]. The common feature of these methods, when dealing with heterogeneous information, is to homogenize the variables or attributes. That means different types of data have to be transformed into the same one before applying other DM methods based on soft sets. Take the research of Maji [17] for example, one of the classical works in this field, Maji [17] presented the rough mathematics soft sets (RMSS) method of decision-making by combining soft sets with rough mathematics of Pawlak [38]. This section compares RMSS with the neighborhood soft sets (NSS) we proposed in this paper to the problem of choosing the best house in Example 2.

#### 4.2.1. Results from RMSS

The first step of RMSS is to construct a soft set $(F,A)$ according to the raw data set. However, the raw data in Example 2 contains various types of variables which can not be described by soft set theory. In order to implement RMSS, all of the variables should be transformed into binary variables with values of “0” and “1”, where “1” indicates “cheap”, “beautiful”, “big”, and “good location”

respectively, and "0" indicates otherwise as in table 6. Then the soft set $(F, A)$ is given as follows according to the mapping $F$ from $A$ to $U$, which consists of the sets of "cheap houses", "beautiful houses", "big houses", and "good location houses":
$(F, A) = \{F(e_1) = \{\text{cheap houses}\} = \{x_1, x_6\},$
$F(e_2) = \{\text{beautiful houses}\} = \{x_1, x_3, x_5, x_6\},$
$F(e_3) = \{\text{big houses}\} = \{x_2, x_3, x_5\},$
$F(e_4) = \{\text{good location houses}\} = \{x_1, x_3, x_5, x_6\}\}.$

In $(F, A)$, $F(e_1) = \{\text{cheap houses}\} = \{x_1, x_6\}$, for example, is a subset in $U$, in which the houses $x_1$ and $x_6$ are cheap in price, while the houses $x_2$, $x_3$, $x_4$ and $x_5$ are not cheap in price. $F(e_2)$, $F(e_3)$ and $F(e_4)$ are defined in the same manner.

The choice value $c_i$ of each house is given by:

$$c_i = \sum_{j}^{m} h_{ij} \tag{4.1}$$

where $i = 1, \cdots, n$, $j = 1, \cdots, m$, $h_{ij}$ is the value of the house $x_i$ on the attribute $e_j$.

Table 6 is the tabulation of $(F, A)$ with the choice values presented in the last column. The optimal choices are the houses with the largest choice value, i.e. $x_1$, $x_3$, $x_5$ and $x_6$.

**Table 6.** Tabular representation of $(F, A)$.

| | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $c_i$ |
|---|---|---|---|---|---|
| $x_1$ | 1 | 1 | 0 | 1 | 3 |
| $x_2$ | 0 | 0 | 1 | 0 | 1 |
| $x_3$ | 0 | 1 | 1 | 1 | 3 |
| $x_4$ | 0 | 0 | 0 | 0 | 0 |
| $x_5$ | 0 | 1 | 1 | 1 | 3 |
| $x_6$ | 1 | 1 | 0 | 1 | 3 |

4.2.2. Comparison

From the results of RMSS and NSS, the differences between them can be summarized as follows: Firstly, decision-making with the NSS method is much simpler than with the RMSS method. The former has the capacity to reduce redundant parameters, so decisions can be made based on less but essential parameters and the burden of decision-making is lessened. In RMSS, however, the final decision had to be made based on all attributes. In the house buying example, four houses were determined as the best houses in both methods. With limited budgets, consumers may have to choose one from the four best alternatives. The NSS method simplified the core condition attributes to only one: The area of the houses. Therefore, the final decision can be derived relying on this core attribute and her/his budget. But in the RMSS method, consumers still need to go through all the attributes of the best alternatives to determine her/his final choice. Therefore, the difficulties of making decision with RMSS method were not fundamentally reduced. Secondly, in RMSS method, serious loss of information was obvious. For example, the houses were classified into big and non-big roughly by the area of them. However, the difference of area between houses within one class could be significant for decision-making. Therefore, the RMSS method based on the binary variables may drop a lot of valuable information during transferring. In the NSS method, heterogeneous information in decision-making problems can be integrated straightway, and information losing or distortion can be prevented. Thirdly, the RMSS method did not make full use of the decision values in the process of decision-making, and there was no connection between the conditional attributes and the decision attributes. A single decision rule was generated by ranking the objectives according to their choice values, the objectives with the maximum choice value were selected as the optimum. On the contrary, the NSS method can generate decision rules by making a connection between the conditional parameters and decision parameters, and the decision can be made based on multiple decision rules, which is more suitable for the actual decision-making environment because of its variety and convenience.

## 5. Application in the medical diagnosis

In this section, we use the newly developed method to facilitate the diagnosis of heart disease. We got a dataset of 297 patients from UCI machine learning repository, which contains 14 parameters for their heart disease diagnosis. Among all of the parameters, the classes of heart disease was regarded as the decision attribute, and the other 13 variables were regarded as the condition attributes for describing the patients. This dataset was chosen because it contains various data types, including binary variables, categorical variables, discrete variables and continuous numerical variables. So the medical diagnosis of heart disease is a practical decision-making problem under a heterogeneous information environment, which can be solved by the NSS method proposed by this paper.

We randomly selected 267 observations for training the NSS algorithm, and used the rest 30 observations for testing the predictive accuracy of the trained decision rules. We also applied the RMSS method in the testing dataset to predict the patients' type of heart disease. The prediction errors of the two methods are shown in Table 7, and it is obvious that our method outperforms RMSS. Moreover, in the process of diagnosis, even though there were more steps in our algorithm compared with Maji's, decision-making based on our algorithm relies on fewer parameters (derived from parameter reduction) and more precise decision rules (obtained from the connection between the conditional attributes and the decision attributes), which can make the decision-making simpler and more efficient. But in Maji's method, there is no parameter reduction involved, and no decision rules specifically defined. As a result, decision-making and prediction using Maji's method is not efficient and less accurate.

**Table 7.** The prediction errors of NNS and RMSS.

| | The prediction error (%) |
|---|---|
| NSS | 3.3 |
| RMSS | 30.0 |

## 6. Conclusions

Based on the research of Molodtsov [1], this paper proposed the concepts of neighborhood soft sets and the neighborhood soft decision system. After that, a new decision-making method under heterogeneous information environment was presented. This method not only can provide adequate parameterization tools inherited from the properties of soft sets, but also can integrate heterogeneous information directly with the help of neighborhood space. More accurate decision can be derived based on the new method, because information losing or distortion caused by transformation of information can be avoided. An example of choosing the best houses was used to demonstrate the operation of the newly developed decision-making method. With the same example, we also compared our method with the decision-making method of Maji [17]. The results showed not only the advantage of our method to process heterogeneous information, but also its capacity to develop concise and effective decision rules. Moreover, the decision-making method proposed by this paper can be applied to a wide range of areas such as feature selection, evaluation and forecasting problems with heterogeneous information.

## Acknowledgments

This work has been supported by the National Natural Science Foundation of China (grants No. 61902370), the Natural Science Foundation of Chongqing (grants No. cstc2020jcyj-msxmX0945), and the Chongqing Municipal Education Commission (grants No. KLQN202100314).

## Conflict of interest

The authors declare no conflicts of interest.

## References

1. D. Molodtsov, Soft set theory–First results, *Comput. Math. Appl.*, **37** (1999), 19–31. https://doi.org/10.1016/S0898-1221(99)00056-5
2. D. H. Hong, C. H. Choi, Multicriteria fuzzy decision-making problems based on vague set theory, *Fuzzy Sets Syst.*, **114** (2000), 103–113. https://doi.org/10.1016/S0165-0114(98)00271-1
3. E. Fix, J. L. Hodges, Discriminatory analysis–Nonparametric discrimination: Consistency properties, *Int. Stat. Rev.*, **57** (1989), 238–247. https://doi.org/10.2307/1403797
4. F. Ye, An extended TOPSIS method with interval-valued intuitionistic fuzzy numbers for virtual enterprise partner selection, *Expert Syst. Appl.*, **37** (2010), 7050–7055. https://doi.org/10.1016/j.eswa.2010.03.013
5. H. Garg, R. Arora, TOPSIS method based on correlation coefficient for solving decision-making problems with intuitionistic fuzzy soft set information, *AIMS Math.*, **5** (2020), 2944–2966. http://dx.doi.org/10.3934/math.2020190
6. J. A. Goguen, L-fuzzy sets, *J. Math. Anal. Appl.*, **18** (1967), 145–174. https://doi.org/10.1016/0022-247X(67)90189-8
7. J. C. R. Alcantud, F. Feng, R. R. Yager, An $N$-soft set approach to rough sets, *IEEE Trans. Fuzzy Syst.*, **28** (2020), 2996–3007. https://doi.org/10.1109/TFUZZ.2019.2946526
8. J. L. Yang, Y. Y. Yao, Semantics of soft sets and three-way decision with soft sets, *Knowledge-Based Syst.*, **194** (2020), 105538. https://doi.org/10.1016/j.knosys.2020.105538
9. J. S. Mi, Y. Leung, W. Z. Wu, Dependence-space-based attribute reduction in consistent decision tables, *Soft Comput.*, **15** (2011), 261–268. https://doi.org/10.1007/s00500-010-0656-1

10. W. A. Khan, A. Rehman, A. Taouti, Soft near-semirings, *AIMS Math.*, **5** (2020), 6464–6478. https://doi.org/10.3934/math.2020417
11. L. A. Zadeh, Fuzzy sets, *Inf. Control*, **8** (1965), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X
12. M. Abu Qamar, A. G. Ahmad, N. Hassan, An approach to Q-neutrosophic soft rings, *AIMS Math.*, **4** (2019), 1291–1306. https://doi.org/10.3934/math.2019.4.1291
13. M. Mohamad, A. Selamat, O. Krejcar, H. Fujita, T. Wu, An analysis on new hybrid parameter selection model performance over big data set, *Knowledge-Based Syst.*, **192** (2020), 105441. https://doi.org/10.1016/j.knosys.2019.105441
14. M. Y. Li, Z. P. Fan, T. H. You, Screening alternatives considering different evaluation index sets: A method based on soft set theory, *Appl. Soft Comput.*, **64** (2018), 614–626. https://doi.org/10.1016/j.asoc.2017.12.037
15. M. I. Ali, F. Feng, X. Y. Liu, W. K. Min, M. Shabir, On some new operations in soft set theory, *Comput. Math. Appl.*, **57** (2009), 1547–1553. https://doi.org/10.1016/j.camwa.2008.11.009
16. P. Yiarayong, On interval-valued fuzzy soft set theory applied to semigroups, *Soft Comput.*, **24** (2020), 3113–3123. https://doi.org/10.1007/s00500-019-04655-3
17. P. K. Maji, A. R. Roy, R. Biswas, An application of soft sets in a decision making problem, *Comput. Math. Appl.*, **44** (2002), 1077–1083. https://doi.org/10.1016/S0898-1221(02)00216-X
18. P. K. Maji, R. Biswas, A. R. Roy, Soft set theory, *Comput. Math. Appl.*, **45** (2003), 555–562. https://doi.org/10.1016/S0898-1221(03)00016-6
19. Q. H. Hu, D. R. Yu, Z. X. Xie, Neighborhood classifiers, *Expert Syst. Appl.*, **34** (2008), 886–876. https://doi.org/10.1016/j.eswa.2006.10.043
20. S. B. Tan, Neighbor-weighted k-nearest neighbor for unbalanced text corpus, *Expert Syst. Appl.*, **28** (2005), 667–671. https://doi.org/10.1016/j.eswa.2004.12.023
21. T. M. Al-shami, An improvement of rough sets' accuracy measure using containment neighborhoods with a medical application, *Inform. Sci.*, **569** (2021), 110–123. https://doi.org/10.1016/j.ins.2021.04.016
22. T. M. Al-shami, Improvement of the approximations and accuracy measure of a rough set using somewhere dense sets, *Soft Comput.*, **25** (2021), 14449–14460. https://doi.org/10.1007/s00500-021-06358-0
23. T. M. Al-shami, M. E. El-Shafei, $T$-soft equality relation, *Turk. J. Math.*, **44** (2020), 1427–1441. https://doi.org/10.3906/mat-2005-117
24. T. M. Al-shami, Compactness on soft topological ordered spaces and its application on the information system, *J. Math.*, **2021** (2021), 6699092. https://doi.org/10.1155/2021/6699092
25. T. M. Al-shami, On soft separation axioms and their applications on decision-making problem, *Math. Probl. Eng.*, **2021** (2021), 8876978. https://doi.org/10.1155/2021/8876978
26. T. M. Al-shami, M. E. El-Shafei, Partial belong relation on soft separation axioms and decision-making problem, two birds with one stone, *Soft Comput.*, **24** (2020), 5377–5387. https://doi.org/10.1007/s00500-019-04295-7
27. T. M. Al-shami, Infra soft compact spaces and application to fixed point theorem, *J. Funct. Space*, **2021** (2021), 3417096. https://doi.org/10.1155/2021/3417096

28. T. M. Al-shami, Investigation and corrigendum to some results related to g-soft equality and gf-soft equality relations, *Filomat*, **33** (2019), 3375–3383. https://doi.org/10.2298/FIL1911375A
29. T. M. Athira, S. J. John, H. Garg, A novel entropy measure of Pythagorean fuzzy soft sets, *AIMS Math.*, **5** (2020), 1050–1061. https://doi.org/10.3934/math.2020073
30. W. Xu, Z. Xiao, X. Dang, D. L. Yang, X. L. Yang, Financial ratio selection for business failure prediction using soft set theory, *Knowledge-Based Syst.*, **63** (2014), 59–67. https://doi.org/10.1016/j.knosys.2014.03.007
31. W. L. Gau, D. J. Buehrer, Vague sets, *IEEE Trans. Syst. Man Cybern.*, **23** (1993), 610–614. https://doi.org/10.1109/21.229476
32. X. B. Yang, T. Y. Lin, J. Y. Yang, Y. Li, D. J. Yu, Combination of interval-valued fuzzy set and soft set, *Comput. Math. Appl.*, **58** (2009), 521–527. https://doi.org/10.1016/j.camwa.2009.04.019
33. X. C. Guan, Y. M. Li, F. Feng, A new order relation on fuzzy soft sets and its application, *Soft Comput.*, **17** (2013), 63–70. https://doi.org/10.1007/s00500-012-0903-8
34. X. Q. Ma, H. W. Qin, N. Sulaiman, T. Herawan, J. H. Abawajy, The parameter reduction of the interval-valued fuzzy soft sets and its related algorithms, *IEEE Trans. Fuzzy Syst.*, **22** (2014), 57–71. https://doi.org/10.1109/TFUZZ.2013.2246571
35. X. D. Peng, Y. Yang, Algorithms for interval-valued fuzzy soft sets in stochastic multi-criteria decision making based on regret theory and prospect theory with combined weight, *Appl. Soft Comput.*, **54** (2017), 415–430. https://doi.org/10.1016/j.asoc.2016.06.036
36. Y. Yang, X. Tan, C. C. Meng, The multi-fuzzy soft set and its application in decision making, *Appl. Math. Model.*, **37** (2013), 4915–4923. https://doi.org/10.1016/j.apm.2012.10.015
37. Y. Zou, Z. Xiao, Data analysis approaches of soft sets under incomplete information, *Knowledge-Based Syst.*, **21** (2008), 941–945. https://doi.org/10.1016/j.knosys.2008.04.004
38. Z. Pawlak, Rough classification, *Int. J. Man-Mach. Stud.*, **20** (1984), 469–483. https://doi.org/10.1016/S0020-7373(84)80022-X
39. Z. Pawlak, Rough sets and decision tables, In: A. Skowron, *Computation theory*, Lecture Notes in Computer Science, Symposium on Computation Theory 1984, Berlin, Germany, **208** (1985), 187–196. https://doi.org/10.1007/3-540-16066-3_18
40. Z. Xiao, K. Gong, Y. Zou, A combined forecasting approach based on fuzzy soft sets, *J. Comput. Appl. Math.*, **228** (2009), 326–333. https://doi.org/10.1016/j.cam.2008.09.033
41. Z. M. Zhang, S. H. Zhang, A novel approach to multi attribute group decision making based on trapezoidal interval type-2 fuzzy soft sets, *Appl. Math. Model.*, **37** (2013), 4948–4971. https://doi.org/10.1016/j.apm.2012.10.006
42. Z. W. Li, G. Q. Wen, N. X. Xie, 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, *Artif. Intell. Med.*, **64** (2015), 161–171. https://doi.org/10.1016/j.artmed.2015.05.002