Fault-tolerant enhanced bijective soft set with applications

Ke Gong^{a, b}, Panpan Wang^{a, c}, Yi Peng^b

Abstract

As an extension of the soft set, the bijective soft set can be used to mine data from soft set environments, and has been studied and applied in some fields. However, only a small proportion of fault data will cause bijective soft sets losing major recognition ability for mining data. Therefore, this study aims to improve the bijective soft set-based data mining method on tolerate-fault-data ability. First some notions and operations of the bijective soft set at a β -misclassification degree is defined. Moreover, algorithms for finding an optimal β , reductions, cores, decision rules and misclassified data are proposed. This paper uses a real problem in gaining shoreline resources evaluation rules to validate the model. The results show that the proposed model has the fault-tolerant ability, and it improves the tolerate-ability of bijective soft set-based data mining method. Moreover, the proposed method can help decision makers to discover fault data for further analysis.

Keywords: shoreline resources; data mining; soft set; variable precision; reduction

1. Introduction

Scholars have made many efforts to model uncertainties in fields of economic, engineering, medical science. These uncertainties are often deemed as incomplete, vagueness and irregularity. Molotov [1] pointed out that traditional mathematical tools for modeling uncertainties are usually inadequate in parameterization process. He proposed the soft set theory [1], which uses parameters to describe the observed data. The concise definition of soft sets makes it easy to describe data with variable forms of uncertainty, which shows a promising research domain on dealing with multi-form uncertainty-mixed problems. Many scholars focused this theory and developed it very fast in theory and practice. Recently, it has been combined with theories, such as fuzzy set [2-5], rough set [6], algebras [7, 8], group [9, 10], ideal [11], topology [12] and intuitionistic fuzzy soft sets [13]. Moreover, soft set theory has been applied in areas, such as parameters reduction [14-16], decision making [17-20], risk assessment [21], medical diagnosis [22-24], group decision [25].

Bijective soft set (BSS) [26] is a particular extension of the soft set. The operations of the BSS make the soft set theory have the ability to reduce and mine decision rules from data. It developed rapidly in theory recently. Gong and Xiao [26] proposed the notion of BSS and used it in mining decision rules of auto sales. Xiao et al. [27] updated the BSS to the exclusive disjunctive soft set. Miao and Wei [28] proposed a discernible matrix-based algorithm to improve the reduction algorithm proposed by the reference [26]. Kim and Min [29] extended the BSS to the full soft set. Gong et al. [30] extended the BSS to fuzzy environments. The applications of the BSS are rising. It was used in some areas, such as the performance measurement of logistics service supply chain [31], the identification of heart valve disease [32], using gene expression profiles to forecast cancer [33] and the classification of electrocardiogram arrhythmia [34].

However, BSS-based methods mainly rely on parameters' dependency degree, which are

obtained by precise classification. It is inevitable that the decision rules calculated by those methods do not reflect the real situation if those methods are not improved. This situation is due to the precise classification, which have not any fault-tolerant ability. Therefore, dependency degree cannot always reflect the reality. Fault-tolerant ability should be considered in the face of noise data.

The purpose of this paper is to improve the fault-tolerance ability of the BSS in obtaining reductions and decision rules. In this research, first the misclassification degree [35] to describe the intensity of misclassification is introduced and some related notions and operations of the BSS are redefined. Second, an algorithm to find the optimal misclassification degree for the BSS decision system is proposed. Third, an algorithm to find reductions, cores and decision rules is proposed. Lastly, on the basis of the proposed algorithms, an empirical analysis of the assessment of shoreline resources is processed as well. The results indicate that the length of shoreline, the stability of shoreline and the traffic condition are the most important factors that decide the value of shoreline. The obtained rules can be used as knowledge to construct an expert system for evaluation shoreline and making a decision. Moreover, the proposed algorithm recognizes the possible fault data of this experiment. A comparison between the proposed method and the method of the literature [26] is conducted as well. The result reveals that the proposed method can obtain a better consequence.

There are three contributions of this paper. First, the introducing of misclassification degree to the BSS enriches the soft set theory in data mining. Second, the proposed method of this paper makes the BSS-based data mining process have a toleration ability, which is not supported in the literature. It can both deal with mining decision rules and discover possible fault data. Third, this method uses real data from shoreline resource to obtain the evaluation rules, which provides a novel angle of view in solving this problem.

This paper is organized as follows. Section 2 introduces the basic concepts and properties of the soft set and the bijective soft set. Section 3 is the notion and algorithm of the proposed method. Section 4 describes a shoreline resources assessment model based on the proposed method. Section 5 is the discussion.

2. Preliminary

2.1. Soft sets

Assume that U is a common universe and E is a set of parameters throughout this paper.

Definition 2.1. ([1]) A pair (F, E) is called a soft set over U if and only if F is a mapping of E into the power set of U , where F is a mapping given by $F : E \rightarrow P(U)$.

Example 1. Suppose that $U = \{h_1, h_2, h_3, h_4\}$ be a set of houses, $E = \{E_1, E_2, E_3, E_4\}$ be a set of parameters that stand for “beautiful”, “cheap”, “in green surroundings”, and “in good location” respectively. These four houses can be represented by the soft set $(F, E) = \{(e_1, \{x_1, x_4\}), (e_2, \{x_1, x_2\}), (e_3, \{x_1, x_3\}), (e_4, \{x_2, x_4\})\}$.

Definition 2.2. ([36]) (*The union of two soft sets*) The union of two soft sets (F, A) and (G, B) over U is the soft set (H, C) , where $C = A \cup B$ and $\forall \varepsilon \in C$,

$$H(\varepsilon) = \begin{cases} F(\varepsilon), & \text{if } \varepsilon \in A - B \\ G(\varepsilon), & \text{if } \varepsilon \in B - A \\ F(\varepsilon) \cup G(\varepsilon), & \text{if } \varepsilon \in A \cap B \end{cases} \quad (2.1)$$

This is denoted by $(F, A) \tilde{\cup} (G, B) = (H, C)$.

Definition 2.3. ([36]) For two soft set (F, A) and (G, B) over U , (F, A) is called a soft subset of (G, B) if (i) $A \subseteq B$, and (ii) $\forall \varepsilon \in A$, $F(\varepsilon)$ and $G(\varepsilon)$ are identical approximations. It is denoted as $(F, A) \tilde{\subset} (G, B)$.

Example 2. Suppose that two soft sets describing five houses.

$$(F, A) = \{(beautiful, \{x_1, x_4\}), (cheap, \{x_1, x_2\})\},$$

$$(G, B) = \{(beautiful, \{x_1, x_4\}), (cheap, \{x_1, x_2, x_3\}), (in\ green\ surroundings, \{x_4\})\}.$$

(F, A) is a soft subset of (G, B) .

Definition 2.4. ([36]) (*AND operation on Two Soft Sets*) If (F, A) and (G, B) are two soft sets, where $A \cap B = \emptyset$, then “ (F, A) AND (G, B) ” denoted by $(F, A) \wedge (G, B)$ is defined by $(F, A) \wedge (G, B) = (H, A \times B)$, where $H(\alpha, \beta) = F(\alpha) \cap G(\beta), \forall (\alpha, \beta) \in A \times B$.

Example 3. Suppose that two soft sets representing four houses as following.

$$(F, A) = \{(beautiful, \{h_1, h_3\}), (cheap, \{h_2, h_4\})\},$$

$$(G, B) = \{(in\ green\ surroundings, \{h_1, h_2\}), (in\ good\ location, \{h_3, h_4\})\}.$$

We have:

$$(F, A) \wedge (G, B) = \{(beautiful \& in\ green\ surrounding, \{h_1\}), (beautiful \& in\ good\ location, \{h_3\}), (cheap \& in\ green\ surrounding, \{h_2\}), (cheap \& in\ good\ location, \{h_4\})\}.$$

2.2. Bijective soft sets

The bijective soft set is a particular soft set, which describes all the elements in a universe completely and without ambiguity.

Definition 2.5. ([26]) Let (F, B) be a soft set over U , where F is a mapping $F : B \rightarrow P(U)$ and B is non-empty parameter set. We say that (F, B) is a bijective soft set, if (F, B) such that (i) $\bigcup_{e \in B} F(e) = U$ (ii) For any two parameters $e_i, e_j \in B, e_i \neq e_j$, $F(e_i) \cap F(e_j) = \emptyset$.

Example 4. Suppose that $U = \{x_1, x_2, x_3, x_4\}$, (F, E) is a soft set over U , $E = \{e_1, e_2\}$, where $F(e_1) = \{x_1, x_2, x_3\}$ and $F(e_2) = \{x_4\}$. (F, E) is a bijective soft set. All of the elements are mapped to a parameter, and every element has a unique parameter. In **Example 3**, (F, A) and (G, B) are both bijective soft sets in universe $U = \{h_1, h_2, h_3, h_4\}$.

3. Fault-tolerant enhanced bijective soft sets

To keep the consistency, we still use the example 4 in [26] to help understanding.

Example 5. Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$. Suppose that nine stores are characterized by a soft sets (F, E) over U . $E = \bigcup_{i=1}^5 E_i$ is the parameter set, which describes the perceived quality of merchandise, the purchasing cost, the empowerment of sales personnel, the high traffic location and “store profit or loss”, respectively. Each soft set is given as following:

$$\begin{aligned}
(F_1, E_1) &= \{(\text{avg.}, \{x_1, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}), (\text{good}, \{x_2\})\}, \\
(F_2, E_2) &= \{(\text{expensive}, \{x_1, x_2, x_3\}), (\text{cheap}, \{x_4, x_5, x_6, x_7, x_8, x_9\})\}, \\
(F_3, E_3) &= \{(\text{high}, \{x_2, x_8\}), (\text{med.}, \{x_4, x_5, x_6, x_7, x_9\}), (\text{low}, \{x_1, x_3\})\}, \\
(F_4, E_4) &= \{(\text{no}, \{x_1, x_3\}), (\text{yes}, \{x_2, x_4, x_5, x_6, x_7, x_8, x_9\})\}, \\
(F_5, E_5) &= \{(\text{profit}, \{x_1, x_2, x_3\}), (\text{balance}, \{x_4, x_5, x_7, x_9\}), (\text{loss}, \{x_6, x_8\})\}.
\end{aligned}$$

3.1. Some operations on bijective soft sets with misclassification degree

In this subsection, we introduce the misclassification to the BSS to enhance the fault-tolerant ability and redefine all the notions and operations of the BSS. We also study some new properties of the BSS with misclassification degree.

Definition 3.1. (see[35]) (*Misclassification degree*) Let $U = \{x_1, x_2, \dots, x_n\}$, $X, Y \subseteq U, X \neq \emptyset$. The misclassification degree that classify X into Y by mistake ($0 \leq \beta \leq 0.5$) is denoted by $\beta = C(X, Y) = 1 - |X \cap Y| / |X|$.

Example 6. Suppose that $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, $X = \{x_1, x_2, x_3, x_4\}$, $Y = \{x_1, x_2, x_3\}$. We have $\beta = C(X, Y) = 1/4$. It indicates that there are 25% elements of X misclassified into Y . Apparently, if $X \subseteq Y$, then X can be entirely classified into Y .

Definition 3.2. (*Misclassified elements*) Let $U = \{x_1, x_2, \dots, x_n\}$, $X, Y \subseteq U, X \neq \emptyset$. The elements in $X - X \cap Y$ is called the misclassified elements.

Example 7. The set of misclassified elements of the data in **Example 6** is x_4 .

Definition 3.3. (β misclassification degree restricted AND operation on a BSS and a Subset of Universe). Let (F, E) be a BSS over U and $X \subseteq U$. The operation of “ (F, E) restricted AND X at β misclassification degree” is defined by $(F, E) \underset{\beta}{\wedge} X = \bigcup_{e \in E} \{F(e) : C(F(e), X) \leq \beta\}$.

If β equals to 0, the operation degenerates into the classical restricted AND operation of a bijective soft set.

Example 8. Let (G, B) be a BSS over U , $U = \{x_1, x_2, x_3, x_4\}$. Suppose the following: $(G, B) = \{(e_1, \{x_1\}), (e_2, \{x_2, x_3, x_4\})\}$, $X = \{x_2, x_3\}$. It can be written $(G, B) \underset{1/3}{\wedge} X = \{x_2, x_3, x_4\}$.

Definition 3.4. (β misclassification degree relaxed AND operation on a BSS and a Subset of Universe) Let (F, E) be a bijective soft set over U and $X \subseteq U$. The operation of “ (F, E) relaxed AND X at β misclassification degree” is defined by $(F, E) \tilde{\wedge}^\beta X = \bigcup_{e \in E} \{F(e) : C(F(e), X) \leq 1 - \beta\}$.

Like the β restricted AND operation, the relaxed one also degenerates into the classical relaxed AND operation of a BSS.

Example 9. Let (G, B) be a BSS over U , $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. Suppose that $(G, B) = \{(e_1, \{x_1\}), (e_2, \{x_2, x_4\}), (e_3, \{x_3, x_5, x_6\})\}$ and $X = \{x_2, x_3\}$. There is $(G, B) \tilde{\wedge}^{1/3} X = \{x_2, x_3, x_4, x_5, x_6\}$.

Similar to the BSS, $Y = (F, E) \tilde{\wedge}_\beta X - (F, E) \underset{\beta}{\wedge} X$ can be used to represent the boundary region of the BSS (F, E) with respect to X at β misclassification degree.

Obviously, there are the following properties.

Property 3.1. Suppose that (F, E) is a BSS over U , and $X \subseteq U$. There is

$$(F, E) \tilde{\wedge}_{\beta} X \supseteq (F, E) \wedge_{\beta} X.$$

Proof. Straightforward. $\square$

Property 3.2. Let $X \subseteq U$ and (F, E) be a BSS over U . For the two misclassification degree β_1, β_2 , $(0 \leq \beta_1 < \beta_2 \leq 0.5)$, There is $(F, E) \wedge_{\beta_1} X \subseteq (F, E) \wedge_{\beta_2} X$.

Proof. $\because \beta_1 < \beta_2$,

$$(F, E) \wedge_{\beta_1} X = \bigcup_{e \in E} \{F(e) : 1 - \frac{|F(e) \cap X|}{|X|} \leq \beta_1\}, (F, E) \wedge_{\beta_2} X = \bigcup_{e \in E} \{F(e) : 1 - \frac{|F(e) \cap X|}{|X|} \leq \beta_2\}$$

$$\therefore (F, E) \wedge_{\beta_1} X \subseteq (F, E) \wedge_{\beta_2} X \square$$

This property indicates that a larger misclassification degree leads to a bigger size of the result set. Similarly, there is the following property.

Property 3.3. Let $X \subseteq U$ and (F, E) be a BSS over U . For the two misclassification degrees β_1, β_2 , $(0 \leq \beta_1 < \beta_2 \leq 0.5)$, there is $(F, E) \tilde{\wedge}_{\beta_1} X \subseteq (F, E) \tilde{\wedge}_{\beta_2} X$.

Proof. Straightforward. $\square$

The β misclassification degree can be introduced to redefine the dependency between two BSSs.

Definition 3.5. (β dependency between two BSS) Let $(F, E), (D, C)$ be two BSSs over U , where $E \cap C = \emptyset$. $(F, E) \rightrightarrows_{k_{\beta}} (D, C)$ is used to represent the degree that (F, E) depends on (D, C) at β misclassification degree, denoted by k_{β} ($0 \leq k_{\beta} \leq 1$) and defined by

$$k_{\beta} = \gamma((F, E), (D, C), \beta) = \frac{|\bigcup_{e \in C} (F, E) \wedge_{\beta} D(e)|}{|U|} \quad (3.1)$$

, where $|\bullet|$ is the cardinal number of a set.

The following example is provided to illustrate the definition.

Example 10. The data in **Example 5** is still used.

$$(F_3, E_3) = \{\{x_2, x_8\}, \{x_4, x_5, x_6, x_7, x_9\}, \{x_1, x_3\}\},$$

$$F_5(\text{profit}) = \{x_1, x_2, x_3\}, F_5(\text{balance}) = \{x_4, x_5, x_7, x_9\}, F_5(\text{loss}) = \{x_6, x_8\}.$$

$$\text{Suppose } \beta=0.2, \text{ there is } (F_3, E_3) \wedge_{0.2} F_5(\text{profit}) = \{x_1, x_3\},$$

$$(F_3, E_3) \wedge_{0.2} F_5(\text{balance}) = \{x_4, x_5, x_6, x_7, x_9\}, (F_3, E_3) \wedge_{0.2} F_5(\text{loss}) = \emptyset.$$

From **Definition 3.5**, it can be obtained

$$k_{0.2} = \gamma((F_3, E_3), (F_5, E_5), 0.2)$$

$$= \frac{|(F_3, E_3) \wedge_{0.2} F_5(\text{profit}) \cup (F_3, E_3) \wedge_{0.2} F_5(\text{balance}) \cup (F_3, E_3) \wedge_{0.2} F_5(\text{loss})|}{|U|}$$

$$= \frac{7}{9}$$

Property 3.4. Suppose that $(F, E), (D, C)$ are two BSSs over U , where $E \cap C = \emptyset$. For two misclassification degrees β_1, β_2 , $(0 \leq \beta_1 < \beta_2 \leq 0.5)$, (F, E) depends on (D, C) to degrees of k_{β_1}, k_{β_2} ($0 \leq k_{\beta_1}, k_{\beta_2} \leq 1$). There is $k_{\beta_1} < k_{\beta_2}$.

Proof. Denote $P_{\beta_1} = \bigcup_{e \in C} (F, E) \wedge_{\beta_1} D(e)$, $P_{\beta_2} = \bigcup_{e \in C} (F, E) \wedge_{\beta_2} D(e)$, where $0 \leq \beta_1 < \beta_2 \leq 0.5$.

According to **Definition 3.5**, there is $P_{\beta_1} \subseteq P_{\beta_2}$.

$$k_{\beta_1} = \gamma((F, E), (D, C), \beta_1) = |P_{\beta_1}|/|U|$$

$$k_{\beta_2} = \gamma((F, E), (D, C), \beta_2) = |P_{\beta_2}|/|U|$$

Then $k_{\beta_1} < k_{\beta_2}$. $\square$

When misclassification degree gets larger, i.e., the fault-tolerant ability gets better, more information of (D, C) can be contained and explained by (F, E) .

Definition 3.6. (β misclassification degree Bijective soft decision system dependency) Let $((F, E), (G, B), U)$ be a bijective soft decision system [26], where $(F, E) = \tilde{U}_{i=1}^n (F_i, E_i)$ and (F_i, E_i) is a BSS. (F, E) is called a condition soft set. The bijective soft β -dependency between $(F_1, E_1) \wedge (F_2, E_2) \wedge \dots \wedge (F_n, E_n)$ and (G, B) is referred to as the β -dependency of $((F, E), (G, B), U)$, denoted κ_β and defined by

$$\kappa_\beta = \gamma(\bigwedge_{i=1}^n (F_i, E_i), (G, B), \beta) \quad (3.2)$$

Example 11. The data of **Example 5** is used here. Suppose that $(\tilde{U}_{i=1}^3 (F_i, E_i), (F_4, E_4), U)$ is a bijective soft decision system on how to choose profitable stores. Suppose that $(H, C) = (F_1, E_1) \wedge (F_2, E_2) \wedge (F_3, E_3) \wedge (F_4, E_4)$.

It can be obtained that $H(e_1) = \{x_1, x_3\}$, $H(e_2) = \{x_2\}$, $H(e_3) = \{x_4, x_5, x_6, x_7, x_9\}$ and $H(e_4) = \{x_8\}$, where e_1 = avg. and expensive and low and no, e_2 = good and expensive and high and yes, e_3 = avg. and cheap and med. and yes, and e_4 = avg. and cheap and high and yes.

$$\begin{aligned} F_5(\text{profit}) &= \{x_1, x_2, x_3\}, F_5(\text{balance}) = \{x_4, x_5, x_7, x_9\}, F_5(\text{loss}) = \{x_6, x_8\} \\ (H, C) &= \{\{x_1, x_3\}, \{x_2\}, \{x_4, x_5, x_6, x_7, x_9\}, \{x_8\}\} \\ (H, C) \wedge_{0.2} F_5(\text{profit}) &= \{x_1, x_2, x_3\} \\ (H, C) \wedge_{0.2} F_5(\text{balance}) &= \{x_4, x_5, x_6, x_7, x_9\} \\ (H, C) \wedge_{0.2} F_5(\text{loss}) &= \{x_8\} \end{aligned}$$

From **Definition 3.6**, it can be obtained that

$$\begin{aligned} \kappa_{0.2} &= \gamma((H, C), (F_5, E_5), 0.2) \\ &= \frac{|(H, C) \wedge_{0.2} F_5(\text{profit}) \cup (H, C) \wedge_{0.2} F_5(\text{balance}) \cup (H, C) \wedge_{0.2} F_5(\text{loss})|}{|U|} \\ &= 1 \end{aligned}$$

Definition 3.7. (*Reduction of β misclassification degree bijective soft decision system*) Let $((F, E), (G, B), U)$ be a bijective soft decision system and M is the set contains all the BSSs of (F, E) . κ_β is the β -dependency of $((F, E), (G, B), U)$. If $N \subseteq M$ such that $\gamma(\bigwedge_{(F_i, E_i) \in N} (F_i, E_i), (G, B), \beta) = \kappa_\beta$ and $\forall \varepsilon \in M \quad \gamma(\bigwedge_{(F_i, E_i) \in N-\varepsilon} (F_i, E_i), (G, B), \beta) < \kappa_\beta$, M is deemed to be a β -reduction of the bijective soft decision system $((F, E), (G, B), U)$.

The reductions of bijective soft decision system with β misclassification degree represent the minimum BSSs to maintain the dependency of entire decision system with β misclassification. We can use the reductions to describe the whole decision system without the disturbing of redundancy.

Example 12. The data of **Example 5** is used here. Suppose that $(\tilde{U}_{i=1}^4(F_i, E_i), (F_5, E_5), U)$ is a β misclassification degree bijective soft decision system. According to **Equation (3.1)** and **Equation (3.2)**, $\gamma((F_2, E_2) \wedge (F_3, E_3), (F_5, E_5), 0.2) = \kappa_{0.2} = 1$. Thus $(F_2, E_2) \tilde{\cup} (F_3, E_3)$ is a reduction of $(\tilde{U}_{i=1}^4(F_i, E_i), (F_5, E_5), U)$.

Like the BSS, the following notions can be redefined by introducing the misclassification degree β as well.

Definition 3.8. (Core reduction of β misclassification degree bijective soft decision system) Suppose that the BSS (H, C) belongs to every β -reduction of the bijective soft decision system $((F, E), (G, B), U)$. (H, C) is called a β -core BSS of $((F, E), (G, B), U)$.

Example 13. The bijective soft decision system of **Example 12** is employed here. According to **Equation (3.1)**, **Equation (3.2)** and **Definition 3.7**, two reductions can be obtained, $(F_2, E_2) \tilde{\cup} (F_3, E_3)$ and $(F_1, E_1) \tilde{\cup} (F_3, E_3)$ when $\beta=0.2$. Therefore, (F_3, E_3) is the core when $\beta=0.2$.

Definition 3.9. (Significance of β misclassification degree soft sets to decision soft set) Suppose that $(\tilde{U}_{i=1}^n(F_i, E_i), (G, B), U)$ is a β misclassification degree bijective soft decision system. The β -significance of BSS to decision soft set, denoted by $\sigma((F_j, E_j), \tilde{U}_{i=1}^n(F_i, E_i), (G, B), \beta)$, is defined as the following:

$$\sigma((F_j, E_j), \tilde{U}_{i=1}^n(F_i, E_i), (G, B), \beta) = \kappa_\beta - \gamma((H, C), (G, B), \beta) \quad (3.3)$$

, where $(H, C) = \bigwedge_{i=1}^n (F_i, E_i)$ ($i \neq j$).

Example 14. The bijective soft decision system of **Example 12** is also used here. According to **Equation (3.3)**, the significances of all the BSSs to decision soft set are 0, 0, 0.111 and 0 when $\beta=0.2$, respectively. Obviously, the core (F_3, E_3) has a non-zero significance.

Property 3.5. Let $((F, E), (G, B), U)$ be a bijective soft decision system and M is the set contains all the BSSs of (F, E) . If a BSS $m \in M$ and $\sigma(m, \tilde{U}_{i=1}^n(F_i, E_i), (G, B), \beta) = 0$ then m is not belong to a reduction.

Proof. Straightforward. $\square$

Definition 3.10. (Decision rules in β misclassification degree bijective decision system) Let $((F, E), (G, B), U)$ be a bijective soft decision system. $\tilde{U}_{i=1}^m(F_i, E_i) \tilde{\subset} (F, E)$ is the β -reduction of $((F, E), (G, B), U)$. Suppose that $(H, C) = \bigwedge_{i=1}^m (F_i, E_i)$. Define a decision rule induced by $\tilde{U}_{i=1}^m(F_i, E_i)$ as following:

$$\text{if } e_i \text{ then } e_j \quad (|H(e_i)| / |G(e_j)|) \quad (3.4)$$

, where $e_i \in C$, $e_j \in B$, and $|H(e_i)| / |G(e_j)|$ denotes the coverage proportion of rule.

3.2. Algorithm for finding an appropriate tolerance level

The value of β is a vital factor, which decides the intensity of misclassification. However, there is a dilemma: if the fault-tolerance ability is enhanced at a larger β , the accuracy of the model's results will be reduced inevitably. To figure out this dilemma, a new algorithm is developed to generate the optimal β as the following.

Algorithm 1.

Input: a bijective soft decision system $((F, E), (G, B), U)$.

Step 1: $i = 0 : s$;

Step 2: $\beta(0) = 0; \kappa_{\beta(0)} = \gamma(\bigwedge_{i=1}^n (F_i, E_i), (G, B), \beta(0))$;

Step 3: while $\kappa_{\beta(i)} \leq 0.5$; // Here $\kappa_{\beta(i)}$ is supposed to smaller than 0.5

 if $\kappa_{\beta(i)} < \kappa_{\beta(i+1)}$,
 $i = i + 1$;
 $\beta(i + 1) = \beta(i) + 1 / 2s$; // where s is the number of step and $s \in I, s > 1$
 $i++$;
 else
 return " $\beta(i)$ is optimal"
 end.

Output: The optimal β .

This algorithm is designed according to the **Property 3.4**. By using the iteration, the minimum β can be obtained such that the κ is the biggest one. The complexity of the algorithm is $O(s)$.

3.3. Algorithm for finding reductions

Algorithm 2.

Input: The bijective soft decision system $(\tilde{\cup}_{i=1}^{n-1} (F_i, E_i), (F_n, E_n), U)$ and the optimal β .

Step 1: Suppose that M is the set of all BSS in $\tilde{\cup}_{i=1}^{n-1} (F_i, E_i)$.

$M \rightarrow red_{front}, M \rightarrow red_{behind}$ // red_{front} and red_{behind} are reductions that obtained by removing bijective soft from ascending and descending order

Step 2: for each $m \in M$ in ascending order
 if $\sigma(m, \tilde{\cup}_{F_i, E_i \in M} (F_i, E_i), (G, B), \beta) = 0$ then
 $red_{front} \rightarrow red_{front} - m$

 End if
End for

Step 3: for each $m \in M$ in descending order
 if $\sigma(m, \tilde{\cup}_{F_i, E_i \in M} (F_i, E_i), (G, B), \beta) = 0$ then
 $red_{behind} \rightarrow red_{behind} - m$

 End if
End for

Step 4: $core \rightarrow red_{front} \cap red_{behind}$

Step 5: Suppose that $\emptyset \rightarrow Misclassified_elements$

for each $x_i \in U$

For each $r \in red_{front}, red_{behind}$

If $x_i \in Misclassified_elements$ of r

$Misclassified_elements \rightarrow Misclassified_elements \cup \{x_i\}$

End for
End for

Output: The reductions, the cores, the misclassified elements and decision rules.

The complexity of the algorithm is $O(N)$, where N is the number of condition BSSs. Step 3 and Step 4 are methods that remove BSSs to find reductions in two directions. Finding all the reductions is a searching process in the power set, which is an algorithm with $O(2^n)$ complexity. We just want to two reductions at most to get the core, obtain the misclassified elements, and get the decision rules. Therefore, the ascending-descending-order strategy is conducted to simplify the complexity of this algorithm. We will discuss this issue in the discussion.

4. Application of mining assessment rules for shoreline resources

4.1. Background

Shoreline resources are one of the most valuable land resources, which are the basis of coasts, quays, and shipping infrastructure. Popular evaluation methods of shoreline resources are limiting-factor-method [37] and statistical methods [38]. Scholars used statistical methods, such as Principal Component Analysis (PCA)[38], Clustering Analysis (CA)[39, 40] and Grey Forecasting (GF)[41], to evaluate the shoreline resources. These methods only deal with quantitative and precise data. However, in the real world most of the indices used in the assessment of shoreline resources are qualitative which is obtained by linguistic description and logical reasoning. Classical methods mentioned above are difficult to solve this kind of problem.

Other than the methods in the literature, the soft set theory can be used to solve this issue. The parameters of the soft set theory can be established in the form of quantity, linguistic description, predicates and the mapping of the parameter with respect to an object.

In this research, the evaluation indicator system in [37] is mainly followed and some of the economic and environmental aspects of the assessment indices, proposed by Yang et al. [39], are integrated. The evaluation indices are shown in **Table 4.1**.

Table 4.1

Evaluation indicator system and the ranking evidence of shoreline resources

Evaluation Index	Rank	Ranking evidence
The length of shoreline(see [40])	A	The length of shoreline is longer than 3000m
	B	The length of shoreline is between 2000 and 3000m
	C	The length of shoreline is between 1000 and 2000m
	D	The length of shoreline is shorter than 1000m
The maximum berthing capacity(see [40])	A	The maximum berthing capacity is 5000 tonnes
	B	The maximum berthing capacity is 3000 tonnes
The stability of shoreline (see [37])	A	Most stable
	B	A little stable
	C	A little unstable
	D	Most unstable
The depth of terrain(see [37])	A	The depth of terrain is bigger than 1000m
	B	The depth of terrain is between 800 and 1000m
	C	The depth of terrain is between 500 and 800m
	D	The depth of terrain is smaller than 500m
The distance between the port to the administration center(see [37])	A	The distance is smaller than 10m
	B	The distance is between 10 and 20m

[37])	C	The distance is between 20 and 40m
	D	The distance is bigger than 40m
The traffic condition(see [37])	A	There are highway, railway and air transportations
	B	There are highway and railway transportations
	C	There is highway transportation
The percentage of port capacity among overall shipping(see [39])	H	High percentage
	L	Small percentage
The existence of protection zone or places of interest nearby(see [39])	Y	Yes, there exists one or more protection zones or places of interest nearby
	N	No, there is no protection zone or places of interest nearby
If the port development satisfies national or local regulations(see [39])	Y	Yes, it satisfies national or local regulations
	N	No, it does not satisfy national or local regulations

4.2. Experiments and results

The 32 ports' shoreline-resources data of Chongqing from China port statistical yearbook, documents of Port and Waterway Bureau of Chongqing, Chongqing municipal people's government are collected. The value of shoreline is ranked to "high shoreline value" and "low shoreline value", which is set by experts.

To conduct the analysis, **Algorithm 1** and **Algorithm 2** are implemented by using JAVA language. The integrated development environment (IDE) is Eclipse 4.4.0. **Fig 1** is the schematic of this model.

Fig 1. Schematic diagram of the proposed model

First, the bijective soft decision system $(\tilde{U}_{i=1}^9(F_i, E_i), (F_{10}, E_{10}), U)$ is constructed for these algorithms as the input. Let the $U = x_1, x_2, \dots, x_{32}$ be a universal set of 32 ports, and E be the

parameter set, where $E = E_1 \cup E_2 \cup \dots \cup E_{10}$. $E_1, \dots, E_9$ represents the nine evaluation indices in **Table 4.1** respectively, and E_{10} describes the general value of shoreline resources. It is assumed that $E_{10} = \{H, L\}$, where H and L represent “high shoreline value” and “low shoreline value”, separately. **Table 4.2** is the tabular form of the soft set $\tilde{U}_{i=1}^{10}(F_i, E_i)$. In this table, if there is a mapping between an element and a parameter, the corresponding cell will be labeled by 1.

Table 4.2

The tabular form of soft set

	E_1				E_2		E_3				E_4				E_5				E_6			E_7		E_8		E_9		E_{10}	
	A	B	C	D	A	B	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	H	L	Y	N	Y	N	H	L
x_1		1			1			1					1		1				1			1			1	1			1
x_2		1			1			1		1						1			1				1		1	1			1
x_3	1				1		1					1					1				1		1		1	1			1
x_4		1			1					1			1	1						1			1	1			1		1
x_5			1		1		1			1								1		1		1			1	1			1
x_6			1		1			1				1				1			1				1			1	1		
x_7				1	1				1				1			1					1	1		1			1		1
x_8				1	1			1					1		1				1			1			1		1	1	
x_9		1			1			1					1				1			1		1	1			1			1
x_{10}			1		1			1					1		1				1			1		1			1		1
x_{11}			1		1			1					1	1							1	1					1		1
x_{12}	1				1		1					1					1				1		1		1				1
x_{13}				1		1	1						1		1						1		1		1		1		1
x_{14}				1		1	1						1		1					1		1		1	1				1
x_{15}			1			1			1				1					1		1		1		1	1				1
x_{16}	1				1			1				1				1			1			1		1	1			1	
x_{17}			1			1		1				1		1					1		1		1		1	1			1
x_{18}		1				1		1				1		1				1				1		1	1			1	
x_{19}		1				1		1				1		1					1			1	1		1				1
x_{20}		1			1			1		1							1			1		1		1		1			1
x_{21}			1			1		1					1		1					1		1		1	1				1
x_{22}			1			1		1					1		1					1		1		1	1			1	
x_{23}			1			1		1					1		1					1		1	1				1		1
x_{24}	1				1			1		1							1		1			1		1	1			1	
x_{25}				1	1			1					1		1						1		1		1	1			1
x_{26}			1			1		1				1		1					1			1		1	1				1
x_{27}				1		1		1				1		1						1		1		1	1				1
x_{28}			1			1		1				1		1					1			1		1	1				1
x_{29}			1			1		1				1		1					1			1		1	1				1
x_{30}		1				1		1				1		1							1		1	1	1				1
x_{31}				1		1		1				1		1					1			1		1	1			1	
x_{32}			1			1		1				1		1					1			1		1	1				1

Second, the bijective soft decision system is put as an input of **Algorithm 1** to obtain the optimal β . It is calculated that the optimal β is **0.15**.

Third, **Algorithm 2** is used to find the reductions and the decision rules. Two reductions can be obtained, which are $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_6, E_6) \tilde{\cup} (F_7, E_7)$, deemed as the reduction 1, and $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_4, E_4) \tilde{\cup} (F_6, E_6)$, deemed as the reduction 2, and the cores are (F_1, E_1) , (F_3, E_3) and (F_6, E_6) . For convenience, these reduced factors are listed in **Table 4.3**. The result of reductions and cores can have the following managerial implication:

- 1) The reduction 1 or the reduction 2 can substitute all the indicators listed in **Table 4.1** to determine the variations of the value for shoreline. In other words, the other factors, such as the maximum berthing capacity [40] and the distance between the port to the administration

center [37] etc., are redundant because their variations cannot affect the variations of the value for shoreline. Decision makers can leverage the results to focus on the most important factors, the factors in reductions, to answer questions like, “why are these factors vital?”, “How to improve the value of shoreline by adjust the most important factors?”.

- 2) The cores of the result exposes that the length of shoreline [40], the stability of shoreline [37] and the traffic condition [37] are indispensable in any reduction. If any one of these factors is removed from any reduction, the dependency of the value of shoreline to the conditional factors will decrease. Namely, the absence of any factor in the cores will make it difficult to illustrate the variations of the value for shoreline, and decision makers should pay more attention to these cores.
- 3) In this case, we consult with the industry. These experts told us that the shoreline resources in Chongqing municipality are rare. A shoreline resource which has a long shoreline $((F_1, E_1))$ with stable geological condition $((F_3, E_3))$ is suitable for developing to a high efficient port. Besides, a valuable shoreline for developing a large and high efficient port depends tightly on the traffic condition $((F_6, E_6))$, which ensures the prominent loading and unloading capacity. The depth of terrain $((F_4, E_4))$ is another vital factor. If a shoreline resource has large terrain, the port-vicinity economy may be prosperous for the large terrain providing enough land resources, which enhances value to the shoreline resource. Lastly, the percentage of port capacity among overall shipping $((F_7, E_7))$ means the contribution to the overall shipping. Obviously, a port which is a top port contributes to the overall shipping should be valuable as well. In summary, the result calculated by the proposed method consists with domain knowledge for the given data.

Table 4.3

The factors in reductions

Symbol	Evaluation Index	Is core
(F_1, E_1)	The length of shoreline (see [40])	Yes
(F_3, E_3)	The stability of shoreline (see [37])	Yes
(F_4, E_4)	The depth of terrain (see [37])	No
(F_6, E_6)	The traffic condition (see [37])	Yes
(F_7, E_7)	The percentage of port capacity among overall shipping(see [39])	No

Two groups of rules can be calculated by the **Algorithm 2** as well. When β is 0.15 and the reduction is $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_6, E_6) \tilde{\cup} (F_7, E_7)$, there are rules:

- (1) $E_1 = A$ and $E_3 = A$ and $E_6 = B$ and $E_7 = L \rightarrow H$ 1 / 32
- (2) $E_1 = A$ and $E_3 = B$ and $E_6 = A$ and $E_7 = H \rightarrow H$ 1 / 32
- (3) $E_1 = B$ and $E_3 = A$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 1 / 32
- (4) $E_1 = B$ and $E_3 = B$ and $E_6 = A$ and $E_7 = H \rightarrow H$ 2 / 32
- (5) $E_1 = B$ and $E_3 = B$ and $E_6 = A$ and $E_7 = L \rightarrow H$ 2 / 32
- (6) $E_1 = B$ and $E_3 = B$ and $E_6 = B$ and $E_7 = H \rightarrow H$ 1 / 32
- (7) $E_1 = B$ and $E_3 = B$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 2 / 32
- (8) $E_1 = B$ and $E_3 = B$ and $E_6 = C$ and $E_7 = L \rightarrow L$ 1 / 32

- (9) $E_1 = B$ and $E_3 = D$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 1 / 32
- (10) $E_1 = C$ and $E_3 = A$ and $E_6 = B$ and $E_7 = H \rightarrow H$ 1 / 32
- (11) $E_1 = C$ and $E_3 = B$ and $E_6 = A$ and $E_7 = L \rightarrow H$ 1 / 32
- (12) $E_1 = C$ and $E_3 = B$ and $E_6 = B$ and $E_7 = H \rightarrow L$ 1 / 32
- (13) $E_1 = C$ and $E_3 = B$ and $E_6 = B$ and $E_7 = L \rightarrow H$ 1 / 32
- (14) $E_1 = C$ and $E_3 = B$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 7 / 32
- (15) $E_1 = C$ and $E_3 = B$ and $E_6 = C$ and $E_7 = H \rightarrow H$ 1 / 32
- (16) $E_1 = C$ and $E_3 = C$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 1 / 32
- (17) $E_1 = D$ and $E_3 = A$ and $E_6 = B$ and $E_7 = L \rightarrow L$ 1 / 32
- (18) $E_1 = D$ and $E_3 = A$ and $E_6 = C$ and $E_7 = L \rightarrow L$ 1 / 32
- (19) $E_1 = D$ and $E_3 = B$ and $E_6 = A$ and $E_7 = H \rightarrow H$ 1 / 32
- (20) $E_1 = D$ and $E_3 = B$ and $E_6 = B$ and $E_7 = L \rightarrow H$ 1 / 32
- (21) $E_1 = D$ and $E_3 = B$ and $E_6 = C$ and $E_7 = L \rightarrow L$ 2 / 32
- (22) $E_1 = D$ and $E_3 = C$ and $E_6 = C$ and $E_7 = H \rightarrow L$ 1 / 32

When β is 0.15 and the reduction is $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_4, E_4) \tilde{\cup} (F_6, E_6)$, there are rules:

- (1) $E_1 = A$ and $E_3 = A$ and $E_4 = B$ and $E_6 = B \rightarrow H$ 1 / 32
- (2) $E_1 = A$ and $E_3 = B$ and $E_4 = A$ and $E_6 = A \rightarrow H$ 1 / 32
- (3) $E_1 = B$ and $E_3 = A$ and $E_4 = B$ and $E_6 = B \rightarrow L$ 1 / 32
- (4) $E_1 = B$ and $E_3 = B$ and $E_4 = A$ and $E_6 = A \rightarrow H$ 1 / 32
- (5) $E_1 = B$ and $E_3 = B$ and $E_4 = A$ and $E_6 = B \rightarrow H$ 1 / 32
- (6) $E_1 = B$ and $E_3 = B$ and $E_4 = C$ and $E_6 = A \rightarrow H$ 2 / 32
- (7) $E_1 = B$ and $E_3 = B$ and $E_4 = D$ and $E_6 = A \rightarrow H$ 1 / 32
- (8) $E_1 = B$ and $E_3 = B$ and $E_4 = D$ and $E_6 = B \rightarrow L$ 2 / 32
- (9) $E_1 = B$ and $E_3 = B$ and $E_4 = D$ and $E_6 = C \rightarrow L$ 1 / 32
- (10) $E_1 = B$ and $E_3 = D$ and $E_4 = D$ and $E_6 = B \rightarrow L$ 1 / 32
- (11) $E_1 = C$ and $E_3 = A$ and $E_4 = A$ and $E_6 = B \rightarrow H$ 1 / 32
- (12) $E_1 = C$ and $E_3 = B$ and $E_4 = B$ and $E_6 = A \rightarrow H$ 1 / 32
- (13) $E_1 = C$ and $E_3 = B$ and $E_4 = D$ and $E_6 = B \rightarrow H$ 1 / 32
- (14) $E_1 = C$ and $E_3 = B$ and $E_4 = D$ and $E_6 = B \rightarrow L$ 8 / 32
- (15) $E_1 = C$ and $E_3 = B$ and $E_4 = D$ and $E_6 = C \rightarrow H$ 1 / 32
- (16) $E_1 = C$ and $E_3 = C$ and $E_4 = D$ and $E_6 = B \rightarrow L$ 1 / 32
- (17) $E_1 = D$ and $E_3 = A$ and $E_4 = D$ and $E_6 = B \rightarrow L$ 1 / 32
- (18) $E_1 = D$ and $E_3 = A$ and $E_4 = D$ and $E_6 = C \rightarrow L$ 1 / 32
- (19) $E_1 = D$ and $E_3 = B$ and $E_4 = C$ and $E_6 = A \rightarrow H$ 1 / 32
- (20) $E_1 = D$ and $E_3 = B$ and $E_4 = D$ and $E_6 = B \rightarrow H$ 1 / 32

$$(21) E_1 = D \text{ and } E_3 = B \text{ and } E_4 = D \text{ and } E_6 = C \rightarrow L \quad 2 / 32$$

$$(22) E_1 = D \text{ and } E_3 = C \text{ and } E_4 = C \text{ and } E_6 = C \rightarrow L \quad 1 / 32$$

Comparing rules with the largest probabilities in these two groups of rules, i.e., Rule (14) in $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_6, E_6) \tilde{\cup} (F_7, E_7)$ group and Rule (14) in $(F_1, E_1) \tilde{\cup} (F_3, E_3) \tilde{\cup} (F_4, E_4) \tilde{\cup} (F_6, E_6)$ group, we also find that values of E_1, E_3, E_6 happen to be the same—C, B, B. This indicates that more efforts should be made on protecting shoreline resources and increasing the accessibility of vehicles. For the brief of this paper, the managerial implications of the rest rules will not be extended in this paper. There is no doubt that the obtained rules can be used as knowledge to construct an expert system for evaluating shoreline and making a decision.

Finally, the proposed method can find the misclassified element set which is $\{x_{22}\}$. Decision makers can leverage this feature to find those elements which may be fault data, and investigate why these data are not consistent with the major part of similar data or revise these data to expected value.

4.3. Comparison

In the former literature [26], Gong et al. proposed the BSS. Our proposed method is an extension of the literature [26], which considers misclassification degree. Therefore, the bijective soft set based method proposed by Ref. [26] can be deemed as the benchmark for the comparison.

This comparison is conducted by the following steps. First, the algorithm for finding reductions, core and decision rules which is proposed by Ref. [26] is implemented by JAVA language, and the integrated development environment (IDE) is Eclipse 4.4.0 as well. Second, the same data which are mentioned in §4 are utilized for running the program. Finally, the result of the comparison between the proposed method and the method of literature [26] is revealed in **Table 4.4**, which contains κ_β (see **Equation (3.2)** and **Ref. [26] Equation (3.3)**), reductions, cores and the number of rules.

Table 4.4 describes five aspects of differences, which reveals huge differences. It is the fault-tolerant ability, enhanced by introducing the misclassification degree, which makes these huge differences. Due to the existence of misclassified data the method of Ref. [26] loses a lot of dependency, whose κ_β is 71.875%. Moreover, the possible fault data also damage the results of cores and reductions, which incur the number of rules increases vastly. If there were not the misclassified data $\{x_{22}\}$, the results of the method of Ref. [26] should be similar to the results of the proposed method. This situation indicates that the misclassified data can damage the results a lot when there is not any fault data tolerant mechanism. Hence, in contrast, the proposed method possesses a fault data tolerant mechanism, which can retain the original dependency when there is disturbing of fault data. Thus, the proposed method is superior to the method of Ref. [26].

Besides, the complexity of the proposed algorithm and the algorithm in Ref. [26] are $O(N)$ and $O(2^n)$ respectively, which means that the proposed algorithm improves the performance of the algorithm. Furthermore, the proposed method can recognize the misclassified elements while the method in Ref. [26] cannot find due to the absence of the ability that tolerates misclassified data.

Table 4.4

The comparison of two methods

Method	κ_β	Reductions	Cores	The number of rules	Can find misclassified element?

The proposed method	1	$\tilde{U}_{i=1,3,6,7}(F_i, E_i)$ $\tilde{U}_{i=1,3,4,6}(F_i, E_i)$	$(F_1, E_1) (F_3, E_3)$ (F_6, E_6)	44	Yes
The method of Ref. [26]	0.71875	$\tilde{U}_{i=1,5,6,7,9}(F_i, E_i)$ $\tilde{U}_{i=1,2,3,6,8}(F_i, E_i)$ $\tilde{U}_{i=1,3,5,6,7,8}(F_i, E_i)$ $\tilde{U}_{i=1,2,3,4,6,9}(F_i, E_i)$ $\tilde{U}_{i=1,3,4,5,6}(F_i, E_i)$	$(F_1, E_1) (F_6, E_6)$	123	No

5. Discussion

The misclassified elements are harmful to the data mining method without misclassification degree. For the reduction (E_1, E_3, E_4, E_6) , x_{22} has the same parameter mapping as the elements (10, 17, 21, 22, 23, 26, 28, 29 and 32). Nevertheless, the decision value is inconsistent. As a result, the bijective soft decision system dependency calculated by the method of [26] is 0.71875, which means about 28% elements cannot explain the variation of the decision values. Accordingly, x_{22} and the reduction (E_1, E_3, E_6, E_7) has the same situation. The evidence reveals that only one fault data may cause a huge loss in dependency and introducing misclassification degree to tolerate these fault data is necessary.

Our proposed method can be used to find possible fault data and help the decision maker to find the problem of data and result. According to the analysis of the misclassified elements in the former paragraph, the evaluation value of x_{22} may be incorrect. There are many reasons lead to the fault data. For example, the data are collected with mistakes; the discretization method chosen for the data processing is inapplicable; the number of discretization ranks is not set rational.

According to the definition of misclassification degree and the operations of the BSS, the phenomenon of misclassification will mostly occur in three cases. First, the sample is large. Second, the parameters to describe the elements are few. Third, the number of parameters in decision BSS is small. In a word, data with large samples requires considering the tolerate ability, the misclassification degree.

Finding all the reductions is a combinational optimization problem because the number of all the subset of conditional BSSs is 2^n [42]. Since the purpose of the **Algorithm 2** is to find two reductions at most to get the core and obtain the misclassified elements. Therefore, removing conditional BSS in ascending and descending order is acceptable and can reduce the complexity of this algorithm, whose complexity is $O(N)$. If there is a requirement to find all the reductions, an updated algorithm proposed by [28] can help achieve, whose complexity is $O(n^2)$.

6. Conclusion

The BSS is a new tool to mine decision rules from data. However, only a minor part of data that have some small fault will cause the BSS-based data mining method to get an undesirable result. In our experiment, only one fault-data, which is only 3%, will cause a 28% loss of dependency. Therefore, this paper introduces the misclassification degree to the operations of the BSS to improve the fault-tolerance ability of the BSS in obtaining reductions and decision rules. Moreover, the

proposed method is applied to the assessment of shoreline resources of 32 real ports. The results give some suggestion for decision makers. Additionally, the proposed algorithm recognizes the possible fault data of this experiment. A comparison is also conducted, which provides support for the proposed method.

In our future work, the fault-tolerant enhanced BSS to other types of soft set would be extended. Some clustering based on the optimization of the fault-tolerant enhanced BSS would be studied. Furthermore, a data-repair approach based on our proposed method would be proposed. The applications of the proposed algorithm in the field of fault data recognition, such as e-commerce, credit risks, are another area of our further study as well.

Acknowledgements

Our work is supported by National Science Foundation Project of CQ CSTC(Grant No. cstcjj A30007), Humanity and Social Science Youth Foundation of Ministry of Education(Grant No. 12YJC630053), China Postdoctoral Science Foundation (Grant No. 2014M560711, 2015T80974), and the National Science Foundation of China(Grant No. 71301180, 71402011). The authors are very grateful for the anonymous reviewers' valuable suggestions.

References

- [1] D. Molodtsov, Soft set theory - First results, *Computers & Mathematics With Applications*, 37 (1999) 19-31.
- [2] Y. Jiang, Y. Tang, Q. Chen, H. Liu, J. Tang, Interval-valued intuitionistic fuzzy soft sets and their properties, *Computers & Mathematics With Applications*, 60 (2010) 906-918.
- [3] P. Majumdar, S. Samanta, Generalised fuzzy soft sets, *Computers & Mathematics With Applications*, 59 (2010) 1425-1432.
- [4] X. Yang, T.Y. Lin, J. Yang, Y. Li, D. Yu, Combination of interval-valued fuzzy set and soft set, *Computers & Mathematics With Applications*, 58 (2009) 521-527.
- [5] I. Delia , N. Çağman, Intuitionistic fuzzy parameterized soft set theory and its decision making, *Applied Soft Computing*, 28 (2015) 109-113.
- [6] F. Feng, X. Liu, L.-F. Violeta, Y.B. Jun, Soft sets and soft rough sets, *Information Sciences*, 181 (2011) 1125-1137.
- [7] Y.B. Jun, K.J. Lee, C.H. Park, Fuzzy soft set theory applied to BCK/BCI-algebras, *Computers & Mathematics with Applications*, 59 (2010) 3180-3192.
- [8] J. Zhan, Y. Jun, Soft BL-algebras based on fuzzy sets, *Computers & Mathematics With Applications*, 59 (2010) 2037-2046.
- [9] H. Aktas, N. Cagman, Soft sets and soft groups, *Information Sciences*, 177 (2007) 2726-2735.
- [10] A. Aygünöglu, H. Aygün, Introduction to fuzzy soft groups, *Computers & Mathematics with Applications*, 58 (2009) 1279-1286.
- [11] Y. Jun, C. Park, Applications of soft sets in ideal theory of BCK/BCI-algebras, *Information Sciences*, 178 (2008) 2466-2475.
- [12] B. Qin, Topological structures of soft fuzzy rough sets, *Journal Of Computational Analysis And Applications*, 17 (2014) 459-467.
- [13] M. Agarwal, K.K. Biswas, M. Hanmandlu, Generalized intuitionistic fuzzy soft sets with applications in decision-making, *Applied Soft Computing*, 13 (2013) 3552-3566.
- [14] D. Chen, E.C.C. Tsang, D. S. Yeung, X. Wang, The parameterization reduction of soft sets and its applications *Computers & Mathematics with Applications*, 49 (2005) 757-763.
- [15] Z. Kong, W. Jia, G. Zhang, L. Wang, Normal parameter reduction in soft set based on particle swarm optimization algorithm, *Applied Mathematical Modelling*, 39 (2015) 4808-4820.
- [16] M.I. Ali, Another view on reduction of parameters in soft sets, *Applied Soft Computing*, 12 (2012) 1814-1821.
- [17] N. Cagman, S. Enginoglu, Soft set theory and uni-int decision making, *European Journal of Operational Research*, 207 (2010) 848-855.
- [18] F. Feng, Y. Li, V. Leoreanu-Fotea, Application of level soft sets in decision making based on interval-valued fuzzy soft sets, *Computers & Mathematics With Applications*, 60 (2010) 1756-1767.
- [19] Y. Jiang, H. Liu, Y. Tang, Q. Chen, Semantic decision making using ontology-based soft sets, *Mathematical and*

Computer Modelling, 53 (2011) 1140-1149.

- [20] Q. Feng, Y. Zhou, Soft discernibility matrix and its applications in decision making, *Applied Soft Computing*, 24 (2014) 749-756.
- [21] K.-H. Chang, A more general risk assessment methodology using a soft set-based ranking technique, *Soft Computing*, 18 (2014) 169-183.
- [22] S.K. Sujit Das, Group decision making in medical system: An intuitionistic fuzzy softset approach, *Applied Soft Computing*, 24 (2014) 196-211.
- [23] N.K.M. T. Mitra Basu, S.K. Mondal, A balanced solution of a fuzzy soft set based decision making problem in medical science, *Applied Soft Computing*, 12 (2012) 3260-3275.
- [24] T. Basu, N. Mahapatra, S. Mondal, A balanced solution of a fuzzy soft set based decision making problem in medical science, *Applied Soft Computing*, 12 (2012) 3260-3275.
- [25] S. Das, S. Kar, Group decision making in medical system: An intuitionistic fuzzy soft set approach, *Applied Soft Computing*, 24 (2014) 196-211.
- [26] K. Gong, Z. Xiao, X. Zhang, The bijective soft set with its operations, *Computers & Mathematics With Applications*, 60 (2010) 2270-2278.
- [27] Z. Xiao, K. Gong, S. Xia, Y. Zou, Exclusive disjunctive soft sets, *Computers & Mathematics With Applications*, 59 (2010) 2128-2137.
- [28] B. Miao, W. Wei, The Parameter Reduction Algorithm and Its Application in Decision-making Based on the Bijective Soft Set, *Systems Engineering*, 30 (2012) 115-119.
- [29] Y.K. Kim, W.K. Min, Full soft sets and full soft decision systems, *Journal of Intelligent and Fuzzy Systems*, 26 (2014) 925-933.
- [30] K. Gong, P. Wang, Z. Xiao, Bijective soft set decision system based parameters reduction under fuzzy environments, *Applied Mathematical Modelling*, 37 (2013) 4474-4485.
- [31] K. Gong, H. Yan, Performance Measurement of Logistics Service Supply Chain Using Bijective Soft Set, *Journal of Advanced Manufacturing Systems*, 14 (2015) 23-40.
- [32] S.U. Kumar, H.H. Inbarani, A.T. Azar, A.E. Hassanien, Identification of Heart Valve Disease using Bijective Soft Sets Theory, *International Journal of Rough Sets and Data Analysis (IJRSDA)*, 1 (2014) 1-14.
- [33] S.U. Kumar, H.H. Inbarani, S.S. Kumar, Improved Bijective-Soft-Set-Based Classification for Gene Expression Data, in: *Computational Intelligence, Cyber Security and Computational Models*, Springer, 2014, pp. 127-132.
- [34] S.U. Kumar, H.H. Inbarani, Classification of ECG Cardiac Arrhythmias Using Bijective Soft Set, in: *Big Data in Complex Systems*, Springer, 2015, pp. 323-350.
- [35] W. Ziarko, Variable precision rough set model, *Journal of Computer and System Sciences*, 46 (1993) 39-59.
- [36] P.K. Maji, R. Biswas, A.R. Roy, Soft set theory, *Computers & Mathematics with Applications*, 45 (2003) 555-562.
- [37] C. Wang, The water-front resources evaluation along the lower and middle reaches of Yangtze River, in: *Chinese Academy of Sciences*, Nanjing, 2000.
- [38] B. Chen, Y. Chen, Application of PCA in port shoreline level evaluation, *Port & Waterway Engineering*, (2008) 26-30.
- [39] D. Yang, C. Wang, P. Zhang, Port Shoreline Evaluation Model and Its Application, *Port & Waterway Engineering*, (2003) 38-42.
- [40] B. Qiu, The Evaluation and Applicability Research of Port Coast-line Resource, in: *Shanghai Maritime University*, Shanghai Maritime University, 2007.
- [41] B. Chen, Y. Chen, Application of entropy-based grey system theory in shoreline evaluation, *Journal of Waterway and Harbor*, 29 (2008) 377-380.
- [42] Q. Hu, D. Yu, Z. Xie, Neighborhood classifiers, *Expert Systems with Applications*, 34 (2008) 866-876.