



Article Simultaneous-Fault Diagnosis of Satellite Power System Based on Fuzzy Neighborhood ζ-Decision-Theoretic Rough Set

Laifa Tao ^{1,2,3}, Chao Wang ^{1,2,3}, Yuan Jia ⁴, Ruzhi Zhou ⁵, Tong Zhang ⁶, Yiling Chen ^{1,2,3}, Chen Lu ^{1,2,3} and Mingliang Suo ^{1,2,3,*}

- ¹ Institute of Reliability Engineering, Beihang University, Beijing 100191, China
- ² Science & Technology on Reliability & Environmental Engineering Laboratory, Beijing 100191, China
- ³ School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
- ⁴ Beijing Institute of Radio Metrology and Measurement, China Aerospace Science and Industry Corporation Limited, Beijing 100039, China
- ⁵ Shanghai Institute of Satellite Engineering, China Aerospace Science and Technology Corporation, Shanghai 201109, China
- ⁶ Marine Design and Research Institute of China, China State Shipbuilding Corporation Limited, Shanghai 200011, China
- Correspondence: suozi@buaa.edu.cn

Abstract: Due to the increasing complexity of the entire satellite system and the deteriorating orbital environment, multiple independent single faults may occur simultaneously in the satellite power system. However, two stumbling blocks hinder the effective diagnosis of simultaneous-fault, namely, the difficulty of obtaining the simultaneous-fault data and the extremely complicated mapping of the simultaneous-fault modes to the sensor data. To tackle the challenges, a fault diagnosis strategy based on a novel rough set model is proposed. Specifically, a novel rough set model named FN ζ DTRS by introducing a concise loss function matrix and fuzzy neighborhood relationship is proposed to accurately mine and characterize the relationship between fault and data. Furthermore, an attribute rule-based fault matching strategy is designed without using simultaneous-fault data as training samples. The numerical experiments demonstrate the effectiveness of the FN ζ DTRS model, and the diagnosis experiments performed on a satellite power system illustrate the superiority of the proposed approach.

Keywords: simultaneous-fault diagnosis; rough set; attribute reduction; satellite power system

MSC: 94C12

1. Introduction

The power system is regarded as the heart of a satellite, whose health management is critical to the on-orbit operation of the entire satellite. The satellite power system is mainly composed of a solar array and battery pack. The solar array exposes in the outer space environment for a long time, and is very vulnerable to external environment intrusion. The battery pack is in a frequent and long-term working state with the periodic operation of the satellite. Therefore, with the increasing probability of space junk collisions, intense radiation of space particles, and striking temperature differences in space, the satellite power system may have multiple independent single faults occurring at the same time, which is called simultaneous-fault [1]. Accurate fault diagnosis is the basis for the health management of a satellite. At present, the research on the diagnosis of single-fault has achieved great success [2,3]. However, as satellites become more complex in their functional composition and longer in their mission time, the mode of simultaneous-fault has become the key factor affecting the normal on-orbit operation of satellites, and the risk and influence caused by such a fault mode cannot be ignored, because the development speed and destructive



Citation: Tao, L.; Wang, C.; Jia, Y.; Zhou, R.; Zhang, T.; Chen, Y.; Lu, C.; Suo, M. Simultaneous-Fault Diagnosis of Satellite Power System Based on Fuzzy Neighborhood ζ-Decision-Theoretic Rough Set. *Mathematics* **2022**, *10*, 3414. https://doi.org/ 10.3390/math10193414

Academic Editor: Denis N. Sidorov

Received: 12 August 2022 Accepted: 7 September 2022 Published: 20 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). power of such a fault mode are far more than that of a single-fault mode. Therefore, it is necessary to diagnose the simultaneous-fault precisely to make sound decisions to enable satellites to perform their missions smoothly and safely. This is the core motivation of our work, namely, we try to solve the diagnosis problem of simultaneous-fault, which is more complex and more harmful than that of single-fault.

With respect to the diagnosis of simultaneous-fault, there are two major challenges that can be listed as follows: (1) The historical simultaneous-fault data are scarce, which greatly limits the effectiveness of data-driven models; (2) The simultaneous-faults would involve multiple sensors, and the mapping between sensor data and fault modes is complicated, which leads to the uncertainty in the diagnosis process. Therefore, new cognitive methods and further research are needed for simultaneous-fault cognition and diagnosis. These can be considered as the technical motivation of our research.

Regarding the first challenge issue, the absence of historical simultaneous-fault data is a thorny problem that needs to be solved urgently. Unlike the traditional fault diagnosis studies that require all kinds of samples during the training phase [4–6], some literature has shown that multi-label classification is expected to achieve simultaneous-fault diagnosis without historical simultaneous-fault data [1,7–10]. The multi-label classification task focuses on the problem where each training simple is represented by a single instance with a single label, and the task is to yield a model that can predict the proper label sets for unseen instances [11]. Multi-label classification methods can be divided into two categories, one of which is the problem transformation methods, including Binary Relevance [12], Classifier Chains [13], Calibrated Label Ranking [14], and other classical methods; the other includes the algorithm adaptation methods, including multi-label K-nearest neighbor (ML-KNN) [15], multi-label decision tree (ML-DT) [16], etc. However, in the face of complex problems, the above methods cannot effectively deal with the problem of insufficient data, and there is still a need for long-term and in-depth research.

For the second challenge issue, some data mining methods are good solutions. In terms of the cognition of things, rough set theory provides a perspective of knowledge and data fusion. This is the main reason why this paper chooses the rough set model as the basic model. The setting of condition attribute and decision attribute can provide multiple information for the characterization of fault, which is conducive to extracting the mapping information between sensor data and fault modes. Rough set theory initiated by Pawlak [17] provides an authoritative mathematical framework for analyzing and handling ambiguous and uncertain data, which can be used to attribute reduction [18–22], rule extraction [23–26], and uncertainty reasoning [22,27–29]. Among kinds of rough set models, the decision-theoretic rough set (DTRS) model has been proved to be a generalized model of many other rough set models [30,31]. At present, there have been related studies on various decision-theoretic rough set models for fault diagnosis, which have proved that the models can effectively select the fault attributes when the pair of the threshold parameters is set appropriately [30,31]. Nevertheless, how to determine the appropriate threshold parameters is the biggest difficulty in the research and application of DTRS. In our previous work [32], we have presented a single-parameter decision-theoretic rough set (SPDTRS) model by setting only one parameter named compensation coefficient rather than two or six, which facilitates the convenient application of the DTRS model. However, the setting of the compensation coefficient in this model is still not clear enough, and the setting of the loss function matrix is defective. In addition, this model lacks the consideration of uncertain information in data description, which makes it unable to deal with continuous data directly. Therefore, in order to make the rough set model (i.e., SPDTRS) more effective in dealing with the simultaneous-fault problem, we need to carry out more targeted improvement work. The details are as follows.

Motivated by the analyses mentioned above, in this work, we propose a fault matching strategy for simultaneous-fault diagnosis based on a revised DTRS named fuzzy neighborhood ζ -decision-theoretic rough set model (FN ζ DTRS). Since there is a coupling relationship of fault characteristics between a single-fault and its associated simultaneous-fault, this

paper proposes the fault matching strategy based on this principle. The main idea of the proposed strategy is that when an unknown simultaneous-fault occurs, its fault attributes are first selected by the FN ζ DTRS and then classified according to the correlation between the obtained fault attributes and the fault attributes of each single-fault selected by the FN ζ DTRS model beforehand. Therefore, the main novelties and contributions of this study can be listed as follows.

- (1) A novel and concise data-driven loss function matrix is designed for DTRS.
- (2) A fuzzy neighborhood ζ-decision-theoretic rough set model is proposed with the help of the fuzzy neighborhood relationship and the proposed loss function matrix, which can deal with hybrid data common in engineering.
- (3) The proposed FNζDTRS model, used for attribute reduction, has a significant advantage in classification accuracy compared with other existing rough sets. This proves that it is more suitable for real fault diagnosis.
- (4) A diagnosis strategy of simultaneous-fault is put forward based on a coupling mapping relationship between single-fault and its associated simultaneous-fault. This ensures that our strategy can handle both single-fault and simultaneous-fault.
- (5) The proposed strategy is successfully applied to the simultaneous-fault diagnosis of the satellite power system and only requires single-fault samples in the training phase, which is highly feasible for practical applications.

The remainder of this paper starts with some preliminaries and related work, then puts forward the presentation of the FN ζ DTRS model in Section 2 and presents the basic framework of simultaneous-fault diagnosis in Section 3. The effectiveness and superiority of the FN ζ DTRS model is verified through some numerical experiments in Section 4, and further demonstrated by a comparative analysis with several baseline algorithms for simultaneous-fault diagnosis in Section 5. The paper closes with main conclusions in Section 6.

2. Preliminaries and Related Work

This subsection will review some notions about rough sets that are relevant to the development of our theory.

Definition 1. (Decision system) A binary group: $DS = (U, C \cup D)$ can describe a decision system. Among them, $U = \{x_1, x_2, \dots, x_m\}$ is called the universe, which is a finite and nonempty set. D is the set of decision attributes which is a nonempty set. C is the collection of conditional attributes, $C \cap D = \emptyset$, $D \neq \emptyset$ 31. Therefore, the relationship between each element in a decision system can be represented as shown in Figure 1. To better understand the above definition, we describe the above-mentioned elements in combination with the fault diagnosis problem. C represents the parameters output by the sensor or the extracted feature attributes, D represents the category of the failure mode, and U denotes the collected data.



Figure 1. The illustration of a decision system.

The decision-theoretic rough set (DTRS) presented by Yao et al. [33]. provides a concise semantic interpretation through a loss function matrix. The loss function matrix is described in Table 1.

Table 1. The detailed information of a loss function matrix.

	Q	Q^c
a_P a_B	$\lambda_{PP} \ \lambda_{BP}$	$\lambda_{PN} \ \lambda_{BN}$
	λ_{NP}	λ_{NN}

Consider that *X* is the subset of samples with the same label d_k . The state *Q* suggests that a related sample defined as *x* is in *X*, and the state Q^c suggests that *x* is not in *X*. The set of actions a_P , a_B , and a_N indicate the classification of *x* into three regions, which are $x \in POS(X)$, $x \in BND(X)$, $x \in NEG(X)$. POS(X) denotes the acceptance of the event $x \in X$. BND(X) denotes the deferment of the event $x \in X$, also considers BND(X) denotes the non-commitment of the event $x \in X$. NEG(X) denotes the rejection of $x \in X$. Furthermore, $\lambda_{\bullet P}$ denotes the loss caused by taking actions (a_P, a_B, a_N) while $x \in X$. $\lambda_{\bullet N}$ is the loss caused by taking actions (a_P, a_B, a_N) while $x \notin X$.

Consider this scenario: the risk of delaying the execution of the correct action is increased compared to that of the correct action, and both are less than the loss of taking the wrong action, the DTRS model therefore made a reasonable assumption: $0 \le \lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$ and $0 \le \lambda_{NN} \le \lambda_{BN} < \lambda_{PN}$, which is the basis for generating this rough set model.

Thanks to the above assumption, a pair of threshold parameters is used to define the positive region POS(X), the boundary region BND(X) and the negative region NEG(X) to construct the DTRS model, which is guided by the Bayesian risk minimization principle and the three-way decision theory. Thus, we have the form of a DTRS model as follows 31:

$$POS_{(\alpha,\beta)}(X) = \{ x \in U | P(X|[x]) \ge \alpha \},\tag{1}$$

$$BND_{(\alpha,\beta)}(X) = \{ x \in U | \beta < P(X[x]) < \alpha \},$$
(2)

$$NEG_{(\alpha,\beta)}(X) = \{ x \in U | P(X|[x]) \le \beta \}.$$
(3)

The following equations represent the relationship between the two threshold parameters and the six loss functions:

$$\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})},\tag{4}$$

$$\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}.$$
(5)

The key parts of the DTRS model are the loss function or threshold parameter (α , β). To study and employ the DTRS model, an important issue is how to determine these parameters. Inspired by the idea of being data-driven, our previous work proposed a single-parameter decision-theoretic rough set (SPDTRS) model [32] that simplifies the traditional DTRS model. Specifically, the model requires only one parameter to be preset rather than the pair of (α , β) or the six parameters in the loss function matrix. However, the solution of employing two truncation functions utilized in the model calculation makes the model relatively complex. Moreover, the interpretability of the loss function matrix in the SPDTRS model is slightly insufficient. The above two disadvantages are the focus of this paper in proposing a new rough set model.

3. Fuzzy Neighborhood ζ-Decision-Theoretic Rough Set

3.1. Granular Computing Based on Fuzzy Neighborhood Relationship

In order to be more applicable to practical problems, rough set models need to be able to handle a hybrid dataset, including continuous and discrete data. Fuzzy relationship and neighborhood relationship are two effective means to deal with the spatial relationship of samples. Their combined form is used by a variety of models [34]. The fuzzy neighborhood relationship can analyze the relationship between the entities in the decision system more precisely. Therefore, to overcome the inability of the SPDTRS model to handle the hybrid dataset, we introduce this fuzzy neighborhood relationship.

Definition 2. (Fuzzy neighborhood relationship) Given a decision system $DS = (U, C \cup D)$, for an arbitrary sample $x \in U$, the fuzzy neighborhood subset of x is defined as:

$$[x]^{\delta} = \{ y \in U | r(x, y) \ge \delta \}, \tag{6}$$

where δ is fuzzy neighborhood radius. The range of δ is $0 \le \delta \le 1$. If x and y are continuous data, we have $r(x, y) = 1 - \frac{1}{n} \sqrt{\left(\sum_{i=1}^{n} (x_i - y_i)^2\right)}$. While the two elements x and y are discrete data, we have

$$r(x,y) = \begin{cases} 1, \text{ if } x_i = y_i \\ 0, \text{ if } x_i \neq y_i' \end{cases}$$
(7)

Thus, the fuzzy neighborhood subset is also called equivalence class. On this basis, the fuzzy conditional probability of *x* could be described as:

$$\widetilde{P}\left(X|[x]^{\delta}\right) = \frac{\sum\left\{r(x,y)|y\in\left(X\cap[x]^{\delta}\right)\right\}}{\sum\left\{r(x,z)|z\in[x]^{\delta}\right\}},\tag{8}$$

where *X* is the subset of samples with the same label d_k . Under the assumption of the fuzzy neighborhood subset $[x]^{\delta} \cap X \neq \emptyset$, we have $0 < \widetilde{P}(X|[x]^{\delta}) \leq 1$, while $\widetilde{P}(X|[x]^{\delta}) = 1$ if and only if $[x]^{\delta} \subseteq X$. Σ {} represents the sum of all elements in its set.

3.2. Determination of the Two Threshold Parameters

Considering the disadvantage of the SPDTRS model, a new loss function matrix is proposed, which is under fuzzy neighborhood relationship by a concise loss function relationship to avoid introducing the truncation functions. The novel SPDTRS model avoids the discussion of multiple situations and reduces the computational complexity of the SPDTRS model.

In the new loss function matrix, the data-driven loss functions under fuzzy neighborhood relationship is shown in Table 2. Besides, $\widetilde{P}(X|[x]^{\delta})$ is the fuzzy neighborhood conditional probability, which can be calculated by Equation (8). The compensation coefficient is ζ with $0 \leq \zeta < 1$. $\widetilde{S}(X|[x]^{\delta})$ and $\widetilde{S}^{c}(X|[x]^{\delta})$ are the significance coefficients, which can be described as follows:

$$\widetilde{S}\left(X|[x]^{\delta}\right) = \frac{\sum\left\{\widetilde{P}\left(X|[y]^{\delta}\right)|y\in\left(X\cap[x]^{\delta}\right)\right\}}{\sum\left\{\widetilde{P}\left(X|[z]^{\delta}\right)|z\in X\right\}},\tag{9}$$

$$\widetilde{S}^{c}\left(X|[x]^{\delta}\right) = \frac{\sum\left\{\widetilde{P}\left(X^{c}|[y]^{\delta}\right)|y\in\left(X^{c}\cap[x]^{\delta}\right)\right\}}{\sum\left\{\widetilde{P}\left(X|[z]^{\delta}\right)|z\in X\right\}}.$$
(10)

where Σ {} represents the sum of all elements in its set.

Under the assumption of the equivalence class $[x]^{\delta} \cap X \neq \emptyset$, the relationships $\widetilde{S}(X|[x]^{\delta}) > 0$ and $\widetilde{S}^{c}(X|[x]^{\delta}) \ge 0$ hold, and $\widetilde{S}^{c}(X|[x]^{\delta}) = 0$ if and only if $[x]^{\delta} \subseteq X$.

	Q	Q ^c
a _p	$\lambda_{PP}=0$	$\lambda_{PN} = \widetilde{S}^{c} \left(X [x]^{\delta} ight)$
a _B	$\lambda_{BP} = \widetilde{S}\Big(X [x]^{\delta}\Big)\widetilde{P}\Big(X [x]^{\delta}\Big)\zeta$	$\lambda_{BN} = \widetilde{S}^{c} \Big(X [x]^{\delta} \Big) \Big(1 - \widetilde{P} \Big(X [x]^{\delta} \Big) \Big) \zeta$
a_N	$\lambda_{NP} = \widetilde{S}\Big(X [x]^{\delta}\Big)$	$\lambda_{NN}=0$

Table 2. The fuzzy neighborhood data-driven loss function matrix.

Subsequently, we can conclude the pair of threshold parameters according to the fuzzy neighborhood data-driven loss function matrix, which can be represented as follows. In addition, we rewrite $\tilde{S}(X|[x]^{\delta}) = S, \tilde{S}^{c}(X|[x]^{\delta}) = S^{c}, \tilde{P}(X|[x]^{\delta}) = P$ for the sake of convenience.

$$\alpha^{fn} = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}$$

$$= \frac{(S^C - S^C (1 - P)\zeta)}{(S^C - S^C (1 - P)\zeta) + (SP\zeta - 0)} , \qquad (11)$$

$$= \frac{S^C (1 - \zeta + P\zeta)}{S^C (1 - \zeta + P\zeta) + SP\zeta}$$

$$\beta^{fn} = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}$$

$$= \frac{(S^C (1 - P)\zeta - 0)}{(S^C (1 - P)\zeta - 0) + (S - SP\zeta)} . \qquad (12)$$

$$= \frac{S^C (1 - P)\zeta}{S^C (1 - P)\zeta + S(1 - P\zeta)}$$

Subsequently, we can set up three-way decision rules as follows:

Rule (P): Decide $x \in POS(X)$ while $\widetilde{P}(X|[x]^{\delta}) \ge \alpha^{fn}$; Rule (B): Decide $x \in BND(X)$ while $\beta^{fn} < \widetilde{P}(X|[x]^{\delta}) < \alpha^{fn}$; Rule (N): Decide $x \in NEG(X)$ while $\widetilde{P}(X|[x]^{\delta}) \le \beta^{fn}$.

According to the decision rules, the following roots about α^{fn} and β^{fn} can be described in the following two cases.

Case 1: $0 < \zeta < 1$

From the rule (P), we can obtain

$$P \leq \frac{(2\zeta S^{C} - S^{C}) - \sqrt{S^{C}(-4\zeta^{2}S + 4\zeta S + S^{C})}}{2\zeta(S + S^{C})},$$
$$P \geq \frac{(2\zeta S^{C} - S^{C}) + \sqrt{S^{C}(-4\zeta^{2}S + 4\zeta S + S^{C})}}{2\zeta(S + S^{C})}.$$
(13)

From the rule (N), we can obtain

$$P \leq \frac{(2\zeta S^{C} + S) - \sqrt{S(-4\zeta^{2}S^{C} + 4\zeta S^{C} + S)}}{2\zeta(S + S^{C})},$$
$$P \geq \frac{(2\zeta S^{C} + S) + \sqrt{S(-4\zeta^{2}S^{C} + 4\zeta S^{C} + S)}}{2\zeta(S + S^{C})}.$$
(14)

From these results, we can only accept two roots because of the relationship $0 < P \le 1$. Thus, we can rewrite these two roots:

$$\alpha_1^{fn} = \frac{(2\zeta S^C - S^C) + \sqrt{S^C(-4\zeta^2 S + 4\zeta S + S^C)}}{2\zeta(S + S^C)},\tag{15}$$

$$\beta_1^{fn} = \frac{(2\zeta S^C + S) - \sqrt{S(-4\zeta^2 S^C + 4\zeta S^C + S)}}{2\zeta(S + S^C)}.$$
(16)

Case 2: $\zeta = 0$

The values of these two parameters are as follows:

$$\alpha_2^{fn} = 1, \tag{17}$$

$$\beta_2^{fn} = 0. \tag{18}$$

The model arising from this case corresponds to the Pawlak model. Both boundary loss functions are equal to 0. Therefore the model clearly exhibits a two-way decision-making characteristic. Thus, the Pawlak model is one of the specific examples of our model, and $\zeta = 0$ is a necessary non-sufficient condition for it.

Theorem 1. For the fuzzy neighborhood data-driven loss function matrix, assuming the equivalence class $[x]^{\delta} \cap X \neq \emptyset$, when $[x]^{\delta} \nsubseteq X$ holds, namely, the concerned equivalence class is not a consistent class, then we have:

 $\begin{array}{l} (a_1) \ \lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}, \\ (a_2) \ \lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}. \\ When \ [x]^{\delta} \subseteq X \ holds, \ i.e., \ the \ concerned \ equivalence \ class \ is \ a \ consistent \ class, \ then \ we \ have: \\ (b_1) \ \lambda_{PP} \leq \lambda_{BP} < \lambda_{NP} \\ (b_2) \ \lambda_{NN} = \lambda_{BN} = \lambda_{PN} = 0 \end{array}$

Proof.

(a₁) If $[x]^{\delta} \notin X$, then $\widetilde{S}(X|[x]^{\delta}) > 0$, $0 < \widetilde{P}(X|[x]^{\delta}) < 1$. Since $0 \le \zeta < 1$, $\lambda_{BP} = \widetilde{S}(X|[x]^{\delta})\widetilde{P}(X|[x]^{\delta})\zeta$, then $0 \le \lambda_{BP} < \widetilde{S}(X|[x]^{\delta})$. Due to $\lambda_{PP} = 0$ and $\lambda_{NP} = \widetilde{S}(X|[x]^{\delta})$, hence $\lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$.

(a₂) If $[x]^{\delta} \not\subseteq X$, then $\widetilde{S}^{c}(X|[x]^{\delta}) > 0$, $0 < \widetilde{P}(X|[x]^{\delta}) < 1$. Since $0 \leq \zeta < 1$, $\lambda_{BN} = \widetilde{S}^{c}(X|[x]^{\delta})(1 - \widetilde{P}(X|[x]^{\delta}))\zeta$, then $0 \leq \lambda_{BN} < \widetilde{S}^{c}(X|[x]^{\delta})$. Due to $\lambda_{NN} = 0$ and $\lambda_{PN} = \widetilde{S}^{c}(X|[x]^{\delta})$, hence $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$.

(b₁) If $[x]^{\delta} \subseteq X$, then $\widetilde{S}(X|[x]^{\delta}) > 0$, $\widetilde{P}(X|[x]^{\delta}) = 1$. Since $0 \leq \zeta < 1$, $\lambda_{BP} = \widetilde{S}(X|[x]^{\delta})\widetilde{P}(X|[x]^{\delta})\zeta$, then $0 \leq \lambda_{BP} < \widetilde{S}(X|[x]^{\delta})$. Due to $\lambda_{PP} = 0$ and $\lambda_{NP} = \widetilde{S}(X|[x]^{\delta})$, hence $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$.

(b₂) If $[x]^{\delta} \subseteq X$, then $\tilde{S}^{c}(X|[x]^{\delta}) = 0$, and $\tilde{P}(X|[x]^{\delta}) = 1$. Since $0 \leq \zeta < 1$, $\lambda_{BN} = \tilde{S}^{c}(X|[x]^{\delta})(1 - \tilde{P}(X|[x]^{\delta}))\zeta$, then $\lambda_{BN} = 0$. Due to $\lambda_{NN} = 0$ and $\lambda_{PN} = \tilde{S}^{c}(X|[x]^{\delta})$, hence $\lambda_{NN} = \lambda_{BN} = \lambda_{PN} = 0$. QED. \Box

3.3. Establishment of FNZDTRS

Reasoning by Section 3.2, the expressions of these two threshold functions lead to the following results: $\alpha^{fn} = f(S, S^c, \zeta)$ and $\beta^{fn} = f(S, S^c, \zeta)$ under the above two conditions, where $\alpha^{fn} = \left\{\alpha_1^{fn}, \alpha_2^{fn}\right\}$ and $\beta^{fn} = \left\{\beta_1^{fn}, \beta_2^{fn}\right\}$. Based on Equations (9) and (10), it is easily to obtain *S* and *S^c* by confirming the parameter δ and analyzing the distribution information

of the original data. In summary, we can change these two threshold parameters to $\alpha^{fn} = f(\delta, \zeta)$, $\beta^{fn} = f(\delta, \zeta)$. The rough set model below, which is including parameters δ and ζ is defined as fuzzy neighborhood ζ -decision-theoretic rough set (FN ζ DTRS):

$$\widetilde{POS} = \left\{ x \in U | \widetilde{P} \left(X | [x]^{\delta} \right) \ge \alpha^{fn} \right\},$$
(19)

$$\widetilde{BND} = \left\{ x \in U | \beta^{fn} < \widetilde{P} \left(X | [x]^{\delta} \right) < \alpha^{fn} \right\},$$
(20)

$$\widetilde{NEG} = \left\{ x \in U | \widetilde{P} \left(X | [x]^{\delta} \right) \le \beta^{fn} \right\},\tag{21}$$

where both threshold parameters have different descriptions under the following two conditions:

Case 1: $0 < \zeta < 1$, then

$$\alpha_1^{fn} = \frac{(2\zeta S^C - S^C) + \sqrt{S^C (-4\zeta^2 S + 4\zeta S + S^C)}}{2\zeta (S + S^C)},\tag{22}$$

$$\beta_1^{fn} = \frac{(2\zeta S^C + S) - \sqrt{S(-4\zeta^2 S^C + 4\zeta S^C + S)}}{2\zeta(S + S^C)}.$$
(23)

Case 2: $\zeta = 0$, then

$$x_2^{fn} = 1,$$
 (24)

$$\beta_2^{fn} = 0.$$
 (25)

In the above $FN\zeta DTRS$ model, there are only two cases to discuss, in contrast to the SPDTRS model that requires four cases to discuss, which greatly reduces the computational complexity of the SPDTRS model due to the concise setting of loss functions.

Theorem 2. In the FN ζ DTRS model, given two compensation coefficients ζ_1 and ζ_2 with $0 \leq \zeta_1 < 1$ and $0 \leq \zeta_2 < 1$, and the parameter δ is fixed, if there exists $\zeta_1 \geq \zeta_2$, then $\alpha^{fn}(\zeta_1) \leq \alpha^{fn}(\zeta_2)$ and $\beta^{fn}(\zeta_1) \geq \beta^{fn}(\zeta_2)$ hold.

Proof. If the equivalence class $[x]^{\delta} \subseteq X$, then $S^{C} = 0$, according to Equations (22)–(25), when $0 < \zeta < 1$, $\alpha^{fn}(\zeta_{1}) = \alpha^{fn}(\zeta_{2}) = 0$, $\beta^{fn}(\zeta_{1}) = \beta^{fn}(\zeta_{2}) = 0$, when $\zeta = 0$, $\alpha^{fn}(\zeta_{1}) = \alpha^{fn}(\zeta_{2}) = 1$, $\beta^{fn}(\zeta_{1}) = \beta^{fn}(\zeta_{2}) = 0$. If the equivalence class $[x]^{\delta} \nsubseteq X$, then its monotonicity relations will be proved by the following derivations.

Part I: For $\zeta_1 \ge \zeta_2 \Rightarrow \alpha^{fn}(\zeta_1) \le \alpha^{fn}(\zeta_2)$, two cases need to be considered. **Case 1:** $0 < \zeta < 1$

Since $\alpha_1^{fn} = \frac{(2\zeta S^C - S^C) + \sqrt{S^C(-4\zeta^2 S + 4\zeta S + S^C)}}{2\zeta(S + S^C)}$, we set $\eta = 1/\zeta$, due to $0 < \zeta < 1$, then $\eta > 1$ and $\alpha_1^{fn} = \frac{(2S^C - \eta S^C) + \sqrt{\eta^2(S^C)^2 + 4\eta SS^C - 4SS^C}}{2(S + S^C)}$. Due to S > 0, $S^C > 0$, we can set $\overline{\alpha}_1^{fn} = (2S^C - \eta S^C) + \sqrt{\eta^2(S^C)^2 + 4\eta SS^C - 4SS^C}$ for simplicity of derivation. We can find the partial derivative of it, denoted as f_{α}^1 , which is

$$f_{\alpha}^{1} = \frac{\partial \overline{\alpha}_{1}^{fn}}{\partial \eta} = -S^{C} + \frac{\eta \left(S^{C}\right)^{2} + 2SS^{C}}{\sqrt{\eta^{2} \left(S^{C}\right)^{2} + 4\eta SS^{C} - 4SS^{C}}}$$

Its second-order partial derivative is denoted as f_{α}^2 :

$$f_{\alpha}^{2} = \frac{\partial^{2}\overline{\alpha}_{1}^{fn}}{\partial^{2}\eta} = \frac{-4S(S^{C})^{2}(S+S^{C})}{\left(\eta^{2}(S^{C})^{2} + 4\eta SS^{C} - 4SS^{C}\right)^{3/2}} < 0.$$

Because f_{α}^2 is less than 0, f_{α}^1 is monotonically decreasing. Since $\eta > 1$, $f_{\alpha}^1 \to 2S > 0(1 \leftarrow \eta)$, $f_{\alpha}^1 \to 0(\eta \to +\infty)$, then $f_{\alpha}^1 > 0$. Therefore, α_1^{fn} grows monotonically with respect to η . Hence, α_1^{fn} decreases monotonically with respect to ζ , that is $\zeta_1 \ge \zeta_2 \Rightarrow \alpha_1^{fn}(\zeta_1) \le \alpha_1^{fn}(\zeta_2)$. Case 2: $\zeta = 0$

In this case, we have $\alpha_2^{fn} = 1$. Thus, for $\zeta_1 \ge \zeta_2$, we have $\alpha_2^{fn}(\zeta_1) = \alpha_2^{fn}(\zeta_2)$. **Part II:** For $\zeta_1 \ge \zeta_2 \Rightarrow \beta^{fn}(\zeta_1) \ge \beta^{fn}(\zeta_2)$, two cases need to be considered as well.

Part II: For $\zeta_1 \ge \zeta_2 \Rightarrow \beta^{jn}(\zeta_1) \ge \beta^{jn}(\zeta_2)$, two cases need to be considered as well. **Case 1:** $0 < \zeta < 1$

Since $\beta_1^{fn} = \frac{(2\zeta S^C + S) - \sqrt{S(-4\zeta^2 S^C + 4\zeta S^C + S)}}{2\zeta(S+S^C)}$, we also set $\eta = 1/\zeta$, due to $0 < \zeta < 1$, then $\eta > 1$ and $\beta_1^{fn} = \frac{(2S^C + \eta S) - \sqrt{\eta^2 S^2 + 4\eta S S^C - 4S S^C}}{2(S+S^C)}$. Due to S > 0, $S^C > 0$, we have the simple form $\beta_1^{fn} = (2S^C + \eta S) - \sqrt{\eta^2 S^2 + 4\eta S S^C - 4S S^C}$. We describe partial derivatives of β_1^{fn} like:

$$\begin{split} f_{\beta}^1 &= \frac{\partial \overline{\beta}_1^{f_1^n}}{\partial \eta} = S - \frac{\eta S^2 + 2SS^C}{\sqrt{\eta^2 S^2 + 4\eta SS^C - 4SS^C}}, \\ f_{\beta}^2 &= \frac{\partial^2 \overline{\beta}_1^{f_n}}{\partial^2 \eta} = \frac{4S^2 S^C (S + S^C)}{\left(\eta^2 S^2 + 4\eta SS^C - 4SS^C\right)^{3/2}} > 0. \end{split}$$

From $f_{\beta}^2 > 0$, we know that f_{β}^1 increases monotonously along with η . Since $\eta > 1$, $f_{\beta}^1 \to -2S^C < 0(1 \leftarrow \eta)$, $f_{\beta}^1 \to 0(\eta \to +\infty)$, then $f_{\alpha}^1 < 0$, and β_1^{fn} is monotonously decreasing with regard to η . Therefore, β_1^{fn} is monotonously increasing with ζ , that is $\zeta_1 \ge \zeta_2 \Rightarrow \beta_1^{fn}(\zeta_1) \ge \beta_1^{fn}(\zeta_2)$. QED. \Box

Theorem 3. In the FN ζ DTRS model, the relationship $0 \leq \beta^{fn} \leq \alpha^{fn} \leq 1$ holds.

Proof. If the equivalence class $[x]^{\delta} \subseteq X$, then $S^{C} = 0$, according to Equations (22)–(25), when $0 < \zeta < 1$, $\alpha^{fn} = \beta^{fn} = 0$, when $\zeta = 0$, $\alpha^{fn} = 0$, $\beta^{fn} = 1$, satisfying the inequality $0 \leq \beta^{fn} \leq \alpha^{fn} \leq 1$. If the equivalence class $[x]^{\delta} \notin X$, then in following three parts we intend to prove the inequality.

Part I: Proof of $0 \le \beta^{fn}$ under two cases.

Case 1: 0 < *ζ* < 1

According to Equation (23), we could get $\beta_1^{fn} = \frac{(2\zeta S^C + S) - \psi}{2\zeta(S+S^C)}$, $\psi = \sqrt{(S + 2\zeta S^C)^2 - 4\zeta^2 S^C(S+S^C)}$, and then we only need to prove $4\zeta^2 S^C(S+S^C) > 0$. Since S > 0, $S^C > 0$, then $4\zeta^2 S^C(S+S^C) > 0$, so $\beta_1^{fn} > 0$.

Case 2: $\zeta = 0$

In this case, we have $\beta_2^{fn} = 0$. **Part II:** Proof of the inequality $\beta^{fn} \leq \alpha^{fn}$ in two cases.

Case 1: $0 < \zeta < 1$

According to Equations (22) and (23), we could get

$$\alpha_1^{fn} - \beta_1^{fn} = \frac{-S^C - S + \sqrt{S^C (-4\zeta^2 S + 4\zeta S + S^C)} + \sqrt{S(-4\zeta^2 S^C + 4\zeta S^C + S)}}{2\zeta(S + S^C)}$$

Since S > 0, $S^C \ge 0$, then

$$\sqrt{S^{C}(-4\zeta^{2}S + 4\zeta S + S^{C})} = \sqrt{4\zeta SS^{C}(1-\zeta) + (S^{C})^{2}} > \sqrt{(S^{C})^{2}} = S^{C},$$

$$\sqrt{S(-4\zeta^{2}S^{C} + 4\zeta S^{C} + S)} = \sqrt{4\zeta SS^{C}(1-\zeta) + S^{2}} > \sqrt{S^{2}} = S,$$

so $\alpha_1^{fn} - \beta_1^{fn} > 0$, that is $\beta_1^{fn} < \alpha_1^{fn}$.

Case 2: $\zeta = 0$

In this case, we have $\alpha_2^{fn} = 1$, $\beta_2^{fn} = 0$, so $\beta_2^{fn} < \alpha_2^{fn}$. **Part III:** Proof of the inequality $\alpha^{fn} \le 1$ in two cases.

Case 1: 0 < *ζ* < 1

According to Equation (22), if we want to prove $\alpha_1^{fn} \leq 1$, then we need to prove $\sqrt{S^C(-4\zeta^2S + 4\zeta S + S^C)} - (2\zeta S + S^C) < 0$, which means we need to prove $S^C(-4\zeta^2S + 4\zeta S + S^C) - (2\zeta S + S^C)^2 < 0$. Since S > 0, $S^C \geq 0$, then

$$S^{C}\left(-4\zeta^{2}S+4\zeta S+S^{C}\right)-\left(2\zeta S+S^{C}\right)^{2}=-4\zeta^{2}SS^{C}-4\zeta^{2}S^{2}=-4\zeta^{2}S\left(S^{C}+S\right)<0.$$

Therefore, we have $\alpha_1^{fn} < 1$.

Case 2: $\zeta = 0$

In this case, we have $\alpha_2^{fn} = 1$. QED. \Box

Theorem 4. For a decision system, which is described as $DS = (U, C \cup D)$ with a fixed parameter δ , and two parameters ζ_1 and ζ_2 with $0 \le \zeta_1 \le \zeta_2 < 1$, we have $\widetilde{POS}_1 \subseteq \widetilde{POS}_2$, $\widetilde{NEG}_1 \subseteq \widetilde{NEG}_2$, $\widetilde{BND}_1 \supseteq \widetilde{BND}_2$.

Proof. At the very beginning of the proof, we assume an arbitrary sample *y* to facilitate the proof of the theorem. While $y \in \widetilde{POS}_1$, we have $P(X|[y]) \ge \alpha^{fn}(\zeta_1)$. Since $0 \le \zeta_1 \le \zeta_2 < 1$, the relation $\alpha^{fn}(\zeta_1) \ge \alpha^{fn}(\zeta_2)$ holds according to Theorem 2. Thus, $P(X|[y]) \ge \alpha^{fn}(\zeta_2)$ and $y \in \widetilde{POS}_2$ hold. Hence, we conclude that $\widetilde{POS}_1 \subseteq \widetilde{POS}_2$.

Likewise, to conclude that $\widetilde{NEG}_1 \subseteq \widetilde{NEG}_2$ and $\widetilde{BND}_1 \supseteq \widetilde{BND}_2$ is easy via Theorem 2. From the above, we can find that ζ is inversely correlated with the range of neutrality and positively correlated with the uncertainty of the decision. QED. \Box

3.4. FNZDTRS-Based Attribute Reduction Algorithm

Jia et al. [35] presented a reduction principle in response to the problem of attribute reduction by using the DTRS model. The core idea of it is minimizing the risk of the reduction subset. On this principle, our previous work has designed an attribute reduction algorithm based on the SPDTRS model [32,34]. It is built on minimizing the risk of overall decisions, which can be utilized for the attribute reduction in our proposed FNζDTRS model. Therefore, the detailed attribute reduction algorithm is not repeated in this paper. For details, the readers could refer to our previous work [32,34].

4. Strategy of Simultaneous-Fault Diagnosis

Under the assumption that the data in all single-fault modes are fully available, when an unknown fault occurs, we use the FN ζ DTRS model to mine the fault attributes of the unknown fault. If the fault attributes of the unknown fault are different from the fault attributes of the existing single-fault, then the unknown fault can be considered as a simultaneous-fault. Furthermore, we can use the attribute reduction results obtained from the FN ζ DTRS model to analyze and identify the corresponding fault modes of the simultaneous-fault. Finally, a strategy of simultaneous-fault diagnosis called fault matching strategy is formed, as shown in Figure 2.

The proposed fault matching strategy consists of two main parts, prior knowledge acquisition and rule matching. In the first part, the single-fault data with abnormal labels and normal data with normal labels are sent into the FN ζ DTRS model as the training data set, and the optimal fault attribute subsets of each single-fault are obtained by attribute reduction, used as the prior knowledge for subsequent diagnosis. In the second part, the simultaneous-fault data with abnormal labels and normal data with normal labels are combined to form the data to be diagnosed, and then the data are fed into the FN ζ DTRS model to obtain the optimal fault attribute subset. The optimal fault attribute subset is then obtained here and the optimal fault attribute subsets obtained in the first part are measured by using the Jaccard similarity coeffective. It is worth noting that there may be some single faults with the same fault attributes, therefore we set some rules based on the differences between attribute data to subdivide the faults and complete the fault matching, which can be seen in the subsequent experiments based on the satellite power system.



Figure 2. The procedure of fault diagnosis strategy for simultaneous-fault.

Based on the above description, we can write the core pseudo code in the above process, as shown in Table 3.

Input:	Raw data of each single fault and normal state
Output:	Fault mode
Part I	Prior Knowledge Acquisition
	For each DS regarding to single fault or state
	Initialized: $red = \emptyset$, $Cl = C$, $Rred = H$, //H is a large positive number.
	While $Cl \neq \emptyset$
	For $c \in Cl$
	$a = c \cup red$, / / a is a temporary set.
	Compute the risk generated by <i>a</i> .
	End For
	Find such a subset $a = c \cup red$ with the minimum risk, i.e., <i>Ra</i> .
	If $Ra < Rred$
	The subset <i>a</i> is the selected set
	End If
	End While
	End For, return the reduction <i>red</i> set of each state.
Part II	Rule Matching
	Utilize the above code to obtain the reduction set r of the given fault data to be diagnosed.
	For each <i>red</i>
	Compute the similarity between <i>red</i> and <i>r</i> .
	End For
	Find such a <i>red</i> with the maximum similarity, which could be considered as the similar fault mode <i>f</i> .
	Return f

Table 3. The core pseudo code of the diagnosis strategy.

5. Numerical Experiment of Attribute Reduction

The effectiveness and advantage of the proposed FNζDTRS model is verified on several hybrid decision systems from the UCI (http://archive.ics.uci.edu/ml/index.php, accessed on 21 July 2021) and KEEL (https://sci2s.ugr.es/keel/datasets.php, accessed on 21 July 2021) datasets. As shown in Table 4, the test datasets include both discrete and continuous data. Specific comparative experiments regarding parameters test and attribute reduction are conducted on the same hard and soft platforms. Ten baseline classifiers, including NaiveBayes, REPTree, LogitBoost, SMO, Filtered, Bagging, PART, IBk, J48 and JRip, are employed with a 10-fold cross-validation in Weka (https://waikato.github.io/weka-wiki/downloading_weka/, accessed on 21 July 2021) software to demonstrate the accuracy of attribute selection. The input data are normalized into the range of [0, 1] during preprocessing.

Table 4. The information of the employed datasets.

ID	Full Name	Name	Samples	Attribute	Discrete	Continuous	Class	Source
1	Mutagenesis-Atoms	Atoms	1618	10	8	2	2	KEEL
2	Australian Credit Approval	Australian	690	14	8	6	2	UCI
3	Breast Cancer	Breast	277	9	6	3	2	UCI
4	Heart Disease Cleveland	Cleve	296	13	7	6	2	UCI
5	Statlog Heart	Heart	270	13	6	7	2	UCI
6	Iris	Iris	150	4	0	4	3	UCI
7	Website Phishing	Phishing	1353	10	10	0	3	UCI
8	South African Hearth	Saheart	462	9	1	8	2	UCI
9	Seismic-Bumps	Seismic	2584	18	12	6	2	UCI
10	Congressional Voting Records	Vote	435	16	16	0	2	UCI

5.1. Parameters Test for FNζDTRS

For the FN ζ DTRS model, two parameters ζ and δ need to be set in advance. As described in Section 2, the theoretic value field of ζ is [0, 1), and δ is [0, 1]. Therefore, ζ is sampled with an interval of 0.05, and end at 0.99. δ is also sampled with an interval of 0.05, but end at 1. Figure 3 shows the experimental results.



(a) Atom: $0.25 \le \zeta \le 0.99$, $0.75 \le \delta \le 1$



(c) Breast: $0.3 \le \zeta \le 0.99$, $0.7 \le \delta \le 0.95$



(e) Heart: $0.3 \le \zeta \le 0.99$, $0.8 \le \delta \le 1$

Figure 3. Cont.



(**b**) Australian: $0.05 \le \zeta \le 0.99$, $0.55 \le \delta \le 1$



(d) Cleve: $0.1 \le \zeta \le 0.99$, $0.8 \le \delta \le 1$



(f) Iris: $0.05 \le \zeta \le 0.99$, $0.6 \le \delta \le 1$



(g) Phishing: $0.15 \le \zeta \le 0.99$, $0.8 \le \delta \le 1$

(h) Saheart: $0.35 \le \zeta \le 0.99$, $0.55 \le \delta \le 0.95$



(i) Seismic: $0 \le \zeta \le 0.99$, $0 \le \delta \le 0.95$ (j) Vote: $0.05 \le \zeta \le 0.99$, $0 \le \delta \le 1$

Figure 3. The accuracy results of ζ and δ with respect to FN ζ DTRS.

The results show that the appropriate settings of ζ and δ range in [0.3, 0.99] and [0.85, 0.95], respectively. It could be explained by the fact that a smaller ζ will result in a larger boundary region, marking more samples as uncertain state. It means setting a smaller ζ for the FN ζ DTRS decision system will lead to greater uncertainty. When applied in real world, ζ should be adjusted appropriately according to the risk of wrong decision. When the danger of making a bad decision is low, ζ can be set larger, and vice versa. On the other hand, with the fuzzy neighborhood threshold δ closer to 1, the fuzzy neighborhood granules will be finer, allowing for the samples to be classified accurately into the appropriate regions. The above two parameters are the core parameters of the model proposed in this paper, and their setting values directly affect the test results. Therefore, when setting the above parameters, the values of the two parameters need to be adjusted according to the actual needs with the above principles.

5.2. Comparison Experiments on Attribute Reduction

In this part, seven related models, DTRS-EF [36], DTRS-SMDNS [37], SPDTRS-EF [32], SPDTRS-SMDNS [32], NDTRS [38], FDTRS [39] and FN3WD [34], are introduced into a contrastive analysis on the attribute reduction to demonstrate the superiority of the proposed FN ζ DTRS. The settings of the relevant parameters in these comparison models are the same as those of the corresponding models. The number of reduction attributes and the classification accuracy are the common evaluation indicators of the comparison experiment of attribute reduction [40,41]. The standard deviations of ten trials are also calculated, and the results are shown in Tables 5 and 6.

ID	Name	DTRS-EF	DTRS- SMDNS	SPDTRS- EF	SPDTRS- SMDNS	NDTRS	FDTRS	FN3WD	FNζDTRS
1	Atoms	69.65 ± 2.38	71.08 ± 1.66	70.94 ± 1.20	71.27 ± 1.14	70.42 ± 1.61	70.87 ± 2.07	71.71 ± 1.00	$\textbf{72.08} \pm \textbf{1.13}$
2	Australian	81.34 ± 5.09	82.39 ± 5.64	82.27 ± 2.87	83.46 ± 0.85	83.31 ± 2.76	$\textbf{72.06} \pm \textbf{12.31}$	84.56 ± 0.39	84.97 ± 0.42
3	Breast	72.31 ± 1.06	$\textbf{72.53} \pm \textbf{0.82}$	72.67 ± 0.70	72.65 ± 0.72	$\textbf{72.72} \pm 0.67$	70.38 ± 0.75	73.21 ± 0.29	$\textbf{74.01} \pm \textbf{0.81}$
4	Cleve	79.13 ± 1.09	78.34 ± 5.02	78.99 ± 1.01	79.37 ± 0.61	77.69 ± 6.68	66.64 ± 11.29	80.13 ± 0.78	81.34 ± 0.37
5	Heart	78.34 ± 3.38	78.72 ± 5.20	79.99 ± 1.87	$\textbf{79.95} \pm \textbf{0.88}$	76.31 ± 7.16	68.37 ± 10.42	80.27 ± 2.25	$\textbf{80.99} \pm \textbf{1.23}$
6	Iris	94.85 ± 0.50	94.85 ± 0.62	94.93 ± 0.47	94.87 ± 0.55	94.95 ± 0.41	62.96 ± 17.98	94.82 ± 0.40	95.27 ± 0.33
7	Phishing	84.23 ± 9.62	86.05 ± 6.11	87.08 ± 1.10	87.08 ± 1.10	86.40 ± 5.08	69.93 ± 14.06	$\textbf{87.15} \pm \textbf{1.06}$	87.14 ± 1.07
8	Saheart	69.23 ± 0.99	69.01 ± 1.29	69.49 ± 0.38	69.51 ± 0.43	69.35 ± 0.40	67.93 ± 1.99	69.63 ± 1.61	$\textbf{70.35} \pm \textbf{1.06}$
9	Seismic	92.64 ± 0.80	92.33 ± 0.90	92.53 ± 0.46	91.94 ± 0.69	91.91 ± 0.70	92.01 ± 0.76	92.56 ± 0.48	93.21 ± 0.15
10	Vote	94.66 ± 0.38	94.67 ± 0.38	94.63 ± 0.36	94.63 ± 0.36	94.65 ± 0.38	83.86 ± 15.36	94.63 ± 0.36	94.65 ± 0.31
	Average	81.64 ± 2.53	82.00 ± 2.76	82.35 ± 1.04	82.47 ± 0.73	81.77 ± 2.59	72.50 ± 8.70	82.87 ± 0.86	83.40 ± 0.69

Table 5. The classification accuracy of the reduction subset.

* Bolded indicates that the model achieves the best performance on this dataset.

Table 6. The number of reduction attributes.

ID	Name	DTRS-EF	DTRS- SMDNS	SPDTRS- EF	SPDTRS- SMDNS	NDTRS	FDTRS	FN3WD	FNζDTRS
1	Atoms	5.5 ± 1.8	6.4 ± 1.8	6.8 ± 0.4	7.8 ± 0.4	5.2 ± 1.2	4.5 ± 1.3	1.2 ± 0.4	2.0 ± 0.0
2	Australian	11.2 ± 1.9	12.2 ± 3.0	11.3 ± 0.6	13.0 ± 0.1	12.7 ± 1.4	4.7 ± 2.7	10.8 ± 0.6	8.4 ± 0.5
3	Breast	7.7 ± 2.7	8.5 ± 2.0	9.0 ± 0.0	9.0 ± 0.1	9.0 ± 0.2	3.6 ± 1.0	8.0 ± 0.0	6.9 ± 0.6
4	Cleve	9.8 ± 0.5	12.5 ± 2.4	9.9 ± 0.5	13.0 ± 0.2	10.2 ± 2.6	4.1 ± 2.7	7.1 ± 1.6	3.0 ± 0.0
5	Heart	7.1 ± 2.0	12.1 ± 2.6	7.9 ± 0.6	12.7 ± 0.4	10.0 ± 3.0	4.8 ± 3.1	6.8 ± 1.3	$\textbf{3.0} \pm \textbf{0.0}$
6	Iris	3.8 ± 0.7	3.6 ± 0.8	3.6 ± 0.5	4.0 ± 0.1	4.0 ± 0.0	1.7 ± 1.3	2.6 ± 0.5	$\textbf{1.0} \pm \textbf{0.0}$
7	Phishing	8.3 ± 2.3	8.8 ± 1.4	9.0 ± 0.0	9.0 ± 0.0	8.8 ± 1.4	1.1 ± 0.2	9.0 ± 0.0	9.0 ± 0.0
8	Saheart	8.7 ± 1.6	8.5 ± 1.9	9.0 ± 0.0	9.0 ± 0.0	9.0 ± 0.0	6.4 ± 2.8	4.8 ± 0.4	$\textbf{2.6} \pm \textbf{0.3}$
9	Seismic	6.6 ± 4.8	10.9 ± 6.3	4.0 ± 0.2	13.4 ± 0.9	13.8 ± 0.5	8.3 ± 0.7	2.0 ± 0.0	$\textbf{1.0} \pm \textbf{0.0}$
10	Vote	8.5 ± 0.6	8.5 ± 0.6	8.5 ± 0.5	8.5 ± 0.5	8.5 ± 0.8	1.1 ± 0.3	8.5 ± 0.5	8.1 ± 0.3
	Average	7.7 ± 1.9	9.2 ± 2.3	7.9 ± 0.3	9.9 ± 0.3	9.1 ± 1.1	4.0 ± 1.6	6.1 ± 0.5	4.5 ± 0.2

* Bolded indicates that the model achieves the best performance on this dataset.

According to the results in Tables 4 and 5, the following analysis can be obtained:

- (a) The analysis based on the classification accuracy indicates that the FNζDTRS model is superior to other rough set models. The main reason may lie in the different methods to describe spatial granules. Discretization methods such as EF and SMDNS are commonly introduced to process continuous data in the traditional DTRS models, which results in the destruction of the spatial structure of granules. Using special measures (such as fuzzy relationship, neighborhood relationship, etc.) can avoid the distortion of the discretization method, but it also has some disadvantages, such as simple measurement, insufficient description ability, etc. The proposed FNζDTRS model utilizes fuzzy neighborhood relationships to overcome the above shortcomings. Compared with other DTRS models, the description of spatial granules is more precise in our model and results in the higher classification accuracy.
- (b) With respect to the number of reduction attributes, the FDTRS model has the least number of reduction attributes, but it fails to achieve a desired classification accuracy, whereas the FNζDTRS model can maintain high classification accuracy while keeping the number of reduction attributes small. The results show that the classification ability can be maintained or improved only when the reduction attributes are accurately selected. The above conclusion also conforms to the basic principle of attribute

reduction, that is, in the operation of reduction, we want to get a relatively concise set, which can ensure that the original classification accuracy is not reduced, and the purpose is to improve the operation efficiency.

(c) The standard deviation is used to measure the robustness of models. It is obvious that the standard deviation of the FN ζ DTRS model is the smallest regardless of the classification accuracy or the number of reduced attributes, which directly proves that the robustness of the FN ζ DTRS model is the highest compared to other models. The above robustness characteristics also show that we have a large selection range when setting our two parameters, which is conducive to the wide application of the model in practical projects.

6. Simultaneous-Fault Diagnosis of Satellite Power System

In-orbit faults of the power system should be avoided to the maximum extent for satellites. Therefore, simulation is the best platform to mine fault diagnosis knowledge. In this section, the effectiveness of the proposed simultaneous-fault diagnosis scheme is verified with the simulation model of a geosynchronous (GEO) satellite power system [3]. As shown in Figure 4, the power system works in a direct energy transfer mode during the simulation, and ten telemetry parameters can be measured in the marked position. The information of the telemetry parameters is shown in Table 7.



Figure 4. The schematic diagram of the power system.

Table 7. Information	of the	telemetry	parameters.
----------------------	--------	-----------	-------------

ID	Attribute	Rate Range	Data Type	Unit
a ₁	Duty cycle	0–1	Continuous	-
a ₂	Bus current	13.5–17.3	Continuous	А
a ₃	Shunt current	5.3-12.4	Continuous	А
a ₄	Battery current	3.6-19.4	Continuous	А
a_5	Output power	1070-1090	Continuous	W
a ₆	Battery pressure	2.0 - 5.4	Continuous	MPa
a ₇	Battery quantity	54.3-71.2	Continuous	Ah
a ₈	Status word	-101	Discrete	-
a9	Bus voltage	40.5-43.1	Continuous	V
a ₁₀	Battery voltage	33.0-40.5	Continuous	V

The raw data used for simultaneous-fault diagnosis is composed of the above-mentioned ten kinds of attributes, and all the data are selected in the stationary period. The dataset is stored in a time-series format, with each subset representing one of the scenarios shown in Table 8. There are a total of 12 scenarios, where scenario 0 represents the system without

any fault. F1 represents open-circuit failure in solar array. F2 represents the short-circuit failure in the battery. F3 represents shunt regulator failure without shunt. F4 represents shunt regulator failure with constant shunt. The remaining 7 scenarios are concurrent failures composed of the above 4 single failures occurring at the same time, where F3 and F4 cannot occur simultaneously.

Scenario	Fault Name	Scenario	Fault Name
0		6	F1-F3
1	F1	7	F1-F4
2	F2	8	F2-F3
3	F3	9	F2-F4
4	F4	10	F1-F2-F3
5	F1-F2	11	F1-F2-F4

 Table 8. Different scenarios for faults in satellite power system.

It can be seen that this means can effectively solve the problem of insufficient data in the simultaneous-fault diagnosis, which also responds to one the difficulties in the simultaneous-fault diagnosis introduced at the beginning of this paper.

6.1. Simultaneous-Fault Diagnosis Based on the Fault Matching Strategy

The two main parts of the simultaneous-fault diagnosis strategy, namely prior knowledge acquisition and rule matching, are equivalent to the training and testing process. We choose 4 kinds of single-fault data as the training set and the remaining 7 kinds of simultaneous-fault data as the testing set. The results obtained through the first step of prior knowledge acquisition are shown in Table 9. It can be found that the results of the output attribute subset of F3 and F4 are the same. Therefore, in order to distinguish F3 and F4, further information needs to be excavated. For attribute a₃, its corresponding shunt current data can directly distinguish F3 from F4. The shunt current value of F3 fluctuates between 6.42–8.67 and that of F4 is between 12.45–14.77. Therefore, F3 and F4 can be distinguished by setting the threshold value of the shunt current average.

Table 9. The results of the training process.

Fault Name	The Output Attribute Subset	Average Value of the Data for Attribute a ₃	
F1	a ₅	-	
F2	a ₇	-	
F3	a ₂ , a ₃ , a ₉	7.46	
F4	a ₂ , a ₃ , a ₉	13.47	

The results obtained through the second step of rule matching are shown in Table 10. For a simultaneous-fault, the Jaccard similarity coefficient between its output attribute subset and the output attribute subset of each single-fault obtained in the training process can be calculated in turn. If the Jaccard similarity coefficient is 0, the corresponding single fault can be eliminated preliminarily. Since F3 and F4 cannot be distinguished by the Jaccard similarity coefficient, it is necessary to further distinguish F3 and F4 through the set shunt current threshold and to finally obtain the matching result. Through the final matching result, it can be found that the diagnostic accuracy of the fault matching strategy is 100%. The above fault matching process comprehensively utilizes the similarity of attributes and expert knowledge, which can ensure that the obtained diagnosis results are more accurate.

Fault Name	The Output	Ja	ccard Simila	rity Coefficie	Average Value of	Matching	
	Attribute Subset	F1	F2	F3	F4	Attribute a ₃	Result
F1-F2	a ₅ , a ₇	0.50	0.50	0	0	-	F1-F2
F1-F3	a ₂ , a ₃ , a ₅ , a ₉	0.25	0	0.75	0.75	6.63	F1-F3
F1-F4	a ₂ , a ₃ , a ₅ , a ₉	0.25	0	0.75	0.75	12.62	F1-F4
F2-F3	a ₂ , a ₃ , a ₇ , a ₉	0	0.25	0.75	0.75	7.47	F2-F3
F2-F4	a ₂ , a ₃ , a ₇ , a ₉	0	0.25	0.75	0.75	13.46	F2-F4
F1-F2-F3	a ₂ , a ₃ , a ₅ , a ₇ , a ₉	0.20	0.20	0.60	0.60	6.63	F1-F2-F3
F1-F2-F4	a ₂ , a ₃ , a ₅ , a ₆ , a ₇ , a ₉	0.17	0.17	0.50	0.50	12.63	F1-F2-F4

Table 10. The results of the testing process.

6.2. Comparison Experiment on Simultaneous-Fault Diagnosis

6.2.1. Experimental Setup

The superiority of the proposed fault matching strategy is demonstrated through comparison experiments of simultaneous-fault diagnosis with several multi-label classification algorithms (Binary Relevance, Classifier Chain, Calibrated Label Ranking, ML-KNN, and ML-DT). In these comparison algorithms, the classifiers of the first three algorithms are all set to Random Forest, which is the best classifier after the pretest, and the value of k for ML-KNN is 12. Subset accuracy, hamming loss, precision, recall, and F1 are introduced as the metrics of the multi-label classification performance [11].

Subset Accuracy =
$$\frac{1}{n} \sum_{i=1}^{n} I(y_i = \hat{y}_i)$$
., (26)

Hamming Loss =
$$\frac{1}{nL} \sum_{i=1}^{n} \sum_{j=1}^{L} I\left(y_i^j \neq \hat{y}_i^j\right)$$
, (27)

Precision =
$$\frac{1}{n} \sum_{i=1}^{n} \frac{\left| y_i^j = 1 \cap \hat{y}_i^j = 1 \right|}{\left| y_i^j = 1 \right|}$$
, (28)

$$\text{Recall} = \frac{1}{n} \sum_{i=1}^{n} \frac{\left| y_i^j = 1 \cap \hat{y}_i^j = 1 \right|}{\left| \hat{y}_i^j = 1 \right|},$$
(29)

$$F_{1} = \frac{1}{n} \sum_{i=1}^{n} \frac{2\left|y_{i}^{j} = 1 \cap \hat{y}_{i}^{j} = 1\right|}{\left|y_{i}^{j} = 1\right| + \left|\hat{y}_{i}^{j} = 1\right|},$$
(30)

where y_i represents the ground-truth label vector of the *i* sample, \hat{y}_i represents the predicted label vector, y_i^j represents the ground-truth label of *j* position in *i* sample, \hat{y}_i^j represents the predicted label of *j* position in *i* sample. Intuitively, subset accuracy, precision, recall, F1 perform as the multi-label counterparts of traditional metrics. Hamming lose performs as a special metric of multi-label.

The training set of all algorithms uses single-fault data, while the test set uses simultaneousfault data. Ten-fold cross validation is also employed in this part.

6.2.2. Experimental Results and Analysis

The results are shown in Table 11. It is obvious that the proposed method outperforms other methods in the simultaneous-faults diagnosis of a satellite power system, which benefits from both the FN ζ DTRS model and the fault matching strategy (FN ζ DTRS-FMS). Fundamentally, on the one hand, the attribute reduction results obtained from the FN ζ DTRS

Algorithm	Accuracy/Subset Accuracy	Hamming Loss	Precision	Recall	F1
FNZDTRS-FMS	100.0 ± 0.0	-	100.0 ± 0.0	100.0 ± 0.0	100.0 ± 0.0
Binary Relevance	82.34 ± 5.39	4.43 ± 1.35	100.0 ± 0.0	93.88 ± 1.87	96.28 ± 1.14
Classifier Chain	65.78 ± 8.93	9.30 ± 2.63	100.0 ± 0.0	86.42 ± 4.45	91.31 ± 3.07
Calibrated Label Ranking	0.0 ± 0.0	32.14 ± 7.11	100.0 ± 0.0	45.24 ± 0.0	61.90 ± 0.0
ML-KNN	0.0 ± 0.0	32.15 ± 0.0	99.98 ± 0.0	45.23 ± 7.11	61.89 ± 0.0
ML-DT	0.0 ± 0.0	32.35 ± 0.31	99.59 ± 0.62	45.03 ± 0.31	61.63 ± 0.42

model can accurately represent the fault information, and on the other hand, the proposed matching rules are reasonable and reliable.

Table 11. The results of the simultaneous-fault diagnosis.

* Bolded indicates that the model achieves the best performance on this metric.

In addition to the results of our method, as for multi-label classification methods, the false alarm rate performs satisfactorily, while the missed diagnosis rate performs poorly. It can be seen that the recall rates have been lower than 50%, which is unacceptable in engineering applications. At the same time, the results of the F1 index generated by the multi-label classification methods are relatively low, only slightly higher than 60%. Similarly, the results obtained by the Calibrated Label Ranking method are similar to those of ML-KNN and ML-DT. The results obtained by the Classifier Chain method are in the middle level. In addition, Binary Relevance has the best performance among the multi-label classification methods, which means that there are no relevant dependencies between the single faults.

7. Conclusions

In this work, a novel DTRS model called FN ζ DTRS is proposed and the fault match strategy (FMS) based on the FN ζ DTRS model is designed to overcome three fundamental hurdles faced by simultaneous-fault diagnosis. The effectiveness and superiority of our methodology is demonstrated by both numerical experiments conducted on several standard datasets and comparison analysis of simultaneous-fault diagnosis performed on a simulation model of a satellite power system. Consequently, two main conclusions can be drawn, as follows.

- (1) The proposed FNζDTRS model performs attribute reduction more effectively compared with other models, and it has strong generalization ability. This benefits from the concise loss functions and the introduction of the fuzzy neighborhood relationships. The advantages of our model can greatly promote the smooth implementation of the model in simultaneous-fault diagnosis, which reflects the effectiveness and superiority of our selection of this model.
- (2) The proposed FNζDTRS–FMS does not require simultaneous-fault samples to accomplish training and performs excellently in simultaneous-fault diagnosis compared to classic multi-label classification algorithms. This is completely consistent with the real situation, that is, the existing data cannot completely cover all the imagined failure modes. Therefore, the diagnostic strategy proposed in this paper has stronger application value.

Although the model we proposed has the above advantages, it still has the problem of low computational efficiency compared with the classical rough set because of the use of fuzzy neighborhood computing. This is one of our future research directions. Furthermore, our future work will also focus on fusing rough set models and multi-label learning algorithms to make the simultaneous-fault diagnosis framework more general. **Author Contributions:** Data curation, R.Z.; Formal analysis, Y.J. and T.Z.; Funding acquisition, C.L. and M.S.; Methodology, L.T. and C.W.; Writing—original draft, Y.C., Y.J. and M.S.; Validation, Y.J.; Investigation, T.Z.; Supervision, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (Grant Nos. 61903015 and 61973011), the Fundamental Research Funds for the Central Universities (Grant No. KG21003001), National key Laboratory of Science and Technology on Reliability and Environmental Engineering (Grant No. WDZC2019601A304), as well as the Capital Science & Technology Leading Talent Program (Grant No. Z191100006119029).

Data Availability Statement: The public data can be accessed through the following two websits: the UCI (http://archive.ics.uci.edu/ml/index.php, accessed on 21 July 2021) and KEEL (https://sci2s.ugr.es/keel/datasets.php, accessed on 21 July 2021). However, the data set of the satellite power system is not convenient for public display. If necessary, researchers can make further consultation via email.

Acknowledgments: Thank the reviewers for their comments on improving the quality of our papers, and thank the journal editors for their enthusiastic work.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

$U = \{x_1, x_2, \cdots, x_m\}$	the universe, which is a finite and nonempty set.
D	the set of decision attributes that is a nonempty set.
С	the collection of conditional attributes.
X	the subset of samples with the same label d_k .
a_P , a_B , and a_N	the classification of x into three regions, which are $x \in POS(X)$,
	$x \in BND(X), x \in NEG(X).$
POS(X)	the acceptance of the event $x \in X$.
BND(X)	the non-commitment of the event $x \in X$, denotes the deferment
	of the event $x \in X$.
NEG(X)	the rejection of $x \in X$.
$\lambda_{\bullet P}$	the loss caused by taking actions (a_P, a_B, a_N) while $x \in X$.
$\lambda_{\bullet N}$	the loss caused by taking actions (a_P, a_B, a_N) while $x \notin X$.
α, β	the threshold parameters of the DTRS model.
δ	fuzzy neighborhood radius.
ζ	the compensation coefficient.

References

- 1. Li, S.; Cao, H.; Yang, Y. Data-driven simultaneous fault diagnosis for solid oxide fuel cell system using multi-label pattern identification. *J. Power Sources* 2018, 378, 646–659. [CrossRef]
- Suo, M.; Tao, L.; Zhu, B.; Chen, Y.; Lu, C.; Ding, Y. Soft decision-making based on decision-theoretic rough set and Takagi-Sugeno fuzzy model with application to the autonomous fault diagnosis of satellite power system. *Aerosp Sci. Technol.* 2020, 106, 106108. [CrossRef]
- Suo, M.; Zhu, B.; An, R.; Sun, H.; Xu, S.; Yu, Z. Data-driven fault diagnosis of satellite power system using fuzzy Bayes risk and SVM. Aerosp Sci. Technol. 2019, 84, 1092–1105. [CrossRef]
- Asgari, S.; Gupta, R.; Puri, I.K.; Zheng, R. A data-driven approach to simultaneous fault detection and diagnosis in data centers. *Appl. Soft Comput.* 2021, 110, 107638. [CrossRef]
- Liang, P.; Deng, C.; Wu, J.; Yang, Z.; Zhu, J.; Zhang, Z. Single and simultaneous fault diagnosis of gearbox via a semi-supervised and high-accuracy adversarial learning framework. *Knowl.-Based Syst.* 2020, 198, 105895. [CrossRef]
- Zhang, Z.; Li, S.; Xiao, Y.; Yang, Y. Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning. *Appl. Energ.* 2019, 233, 930–942. [CrossRef]
- 7. Pooyan, N.; Shahbazian, M.; Salahshoor, K.; Hadian, M. Simultaneous Fault Diagnosis using multi class support vector machine in a Dew Point process. *J. Nat. Gas. Sci. Eng.* **2015**, *23*, 373–379. [CrossRef]
- 8. Vong, C.; Wong, P.; Ip, W. A New Framework of Simultaneous-Fault Diagnosis Using Pairwise Probabilistic Multi-Label Classification for Time-Dependent Patterns. *Ieee T Ind. Electron.* **2013**, *60*, 3372–3385. [CrossRef]
- 9. Wong, P.K.; Zhong, J.; Yang, Z.; Vong, C.M. Sparse Bayesian extreme learning committee machine for engine simultaneous fault diagnosis. *Neurocomputing* **2016**, 174, 331–343. [CrossRef]

- 10. Wu, B.; Cai, W.; Chen, H.; Zhang, X. A hybrid data-driven simultaneous fault diagnosis model for air handling units. *Energy Build.* **2021**, 245, 111069. [CrossRef]
- 11. Zhang, M.; Zhou, Z. A Review on Multi-Label Learning Algorithms. IEEE Trans. Knowl. Data Eng. 2014, 26, 1819–1837. [CrossRef]
- 12. Boutell, M.R.; Luo, J.B.; Shen, X.P.; Brown, C.M. Learning multi-label scene classification. *Pattern Recognit.* 2004, *37*, 1757–1771. [CrossRef]
- Read, J.; Pfahringer, B.; Holmes, G.; Frank, E. Classifier chains for multi-label classification. *Mach. Learn.* 2011, 85, 333–359. [CrossRef]
- Fuernkranz, J.; Huellermeier, E.; Mencia, E.L.; Brinker, K. Multilabel classification via calibrated label ranking. *Mach. Learn.* 2008, 73, 133–153. [CrossRef]
- 15. Zhang, M.; Zhou, Z. ML-KNN: A lazy learning approach to multi-label learning. Pattern Recognit. 2007, 40, 2038–2048. [CrossRef]
- 16. Clare, A.; King, R.D. Knowledge Discovery in Multi-label Phenotype Data. In *European Conference on Principles of Data Mining and Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 42–53.
- 17. Pawlak, Z. Rough sets. Int. J. Comput. Inf. Sci. 1982, 11, 341–356. [CrossRef]
- 18. Dong, L.; Chen, D.; Wang, N.; Lu, Z. Key energy-consumption feature selection of thermal power systems based on robust attribute reduction with rough sets. *Inf. Sci.* 2020, 532, 61–71. [CrossRef]
- Su, L.; Yu, F. Matrix approach to spanning matroids of rough sets and its application to attribute reduction. *Theor. Comput. Sci.* 2021, 893, 105–116. [CrossRef]
- 20. Sahu, R.; Dash, S.R.; Das, S. Career selection of students using hybridized distance measure based on picture fuzzy set and rough set theory. *Decis. Mak. Appl. Manag. Eng.* 2021, *4*, 104–126. [CrossRef]
- Dash, S.R.; Dehuri, S.; Sahoo, U.K. Interactions and Applications of Fuzzy, Rough, and Soft Set in Data Mining. Int. J. Fuzzy Syst. Appl. 2015, 3, 37–50. [CrossRef]
- Zhang, P.; Li, T.; Wang, G.; Luo, C.; Chen, H.; Zhang, J.; Wang, D.; Yu, Z. Multi-source information fusion based on rough set theory: A review. *Inf. Fusion* 2021, 68, 85–117. [CrossRef]
- Bai, H.; Ge, Y.; Wang, J.; Li, D.; Liao, Y.; Zheng, X. A method for extracting rules from spatial data based on rough fuzzy sets. *Knowl.-Based Syst.* 2014, 57, 28–40. [CrossRef]
- Landowski, M.; Landowska, A. Usage of the rough set theory for generating decision rules of number of traffic vehicles. *Transp. Res. Procedia* 2019, 39, 260–269. [CrossRef]
- Sharma, H.K.; Kumari, K.; Kar, S. A rough set theory application in forecasting models. *Decis. Mak. Appl. Manag. Eng.* 2020, 3, 1–21. [CrossRef]
- 26. Guo, Y.; Tsang, E.C.C.; Xu, W.; Chen, D. Adaptive weighted generalized multi-granulation interval-valued decision-theoretic rough sets. *Knowl.-Based Syst.* 2020, *187*, 104804. [CrossRef]
- Wang, T.; Liu, W.; Zhao, J.; Guo, X.; Terzija, V. A rough set-based bio-inspired fault diagnosis method for electrical substations. *Int. J. Electr. Power Energy Syst.* 2020, 119, 105961. [CrossRef]
- 28. Sang, B.; Yang, L.; Chen, H.; Xu, W.; Guo, Y.; Yuan, Z. Generalized multi-granulation double-quantitative decision-theoretic rough set of multi-source information system. *Int. J. Approx. Reason.* **2019**, *115*, 157–179. [CrossRef]
- 29. Zhang, P.F.; Li, T.R.; Yuan, Z.; Luo, C.; Liu, K.Y.; Yang, X.L. Heterogeneous Feature Selection Based on Neighborhood Combination Entropy. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, 1–14. [CrossRef]
- Wang, L.; Shen, J.; Mei, X. Cost Sensitive Multi-Class Fuzzy Decision-theoretic Rough Set Based Fault Diagnosis. In Proceedings of the 2017 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 6957–6961.
- 31. Yu, J.; Ding, B.; He, Y. Rolling bearing fault diagnosis based on mean multigranulation decision-theoretic rough set and non-naive Bayesian classifier. *J. Mech. Sci. Technol.* **2018**, *32*, 5201–5211. [CrossRef]
- Suo, M.; Tao, L.; Zhu, B.; Miao, X.; Liang, Z.; Ding, Y.; Zhang, X.; Zhang, T. Single-parameter decision-theoretic rough set. *Inf. Sci.* 2020, 539, 49–80. [CrossRef]
- Yao, Y.Y.; Wong, S.K.M. A decision theoretic framework for approximating concepts. *Int. J. Man Mach. Stud.* 1992, 37, 793–809. [CrossRef]
- Suo, M.; Cheng, Y.; Zhuang, C.; Ding, Y.; Lu, C.; Tao, L. Extension of labeled multiple attribute decision making based on fuzzy neighborhood three-way decision. *Neural Comput. Appl.* 2020, *32*, 17731–17758. [CrossRef]
- 35. Jia, X.; Liao, W.; Tang, Z.; Shang, L. Minimum cost attribute reduction in decision-theoretic rough set models. *Inf. Sci.* 2013, 219, 151–167. [CrossRef]
- 36. Yao, Y. Three-way decisions with probabilistic rough sets. Inf. Sci. 2010, 180, 341–353. [CrossRef]
- 37. Jiang, F.; Sui, Y. A novel approach for discretization of continuous attributes in rough set theory. *Knowl.-Based Syst.* 2015, 73, 324–334. [CrossRef]
- Li, W.; Huang, Z.; Jia, X.; Cai, X. Neighborhood based decision-theoretic rough set models. Int. J. Approx. Reason. 2016, 69, 1–17. [CrossRef]
- Song, J.; Tsang, E.C.C.; Chen, D.; Yang, X. Minimal decision cost reduct in fuzzy decision-theoretic rough set model. *Knowl.-Based* Syst. 2017, 126, 104–112. [CrossRef]

- 40. Wang, C.; Qi, Y.; Shao, M.; Hu, Q.; Chen, D.; Qian, Y.; Lin, Y. A Fitting Model for Feature Selection with Fuzzy Rough Sets. *IEEE Trans. Fuzzy Syst.* **2017**, *25*, 741–753. [CrossRef]
- 41. Wang, C.; Shao, M.; He, Q.; Qian, Y.; Qi, Y. Feature subset selection based on fuzzy neighborhood rough sets. *Knowl.-Based Syst.* **2016**, *111*, 173–179. [CrossRef]