Feature-weighted Random Forest with Boruta for Fault Diagnosis of Satellite Attitude Control Systems

Shaozhi Chen[1], Xiaopeng Xi[2], Maiying Zhong[3], and Marcos E. Orchard[4]

[1,3] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, Shandong, 266590, China*
szchen@sdust.edu.cn
myzhong@sdust.edu.cn (corresponding author)

[2] *Advanced Center for Electrical and Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile*
xi.xiaopeng@usm.cl

[4] *Department of Electrical Engineering, Faculty of Mathematical and Physical Sciences, Universidad de Chile, Santiago, 8370451, Chile*
morchard@ing.uchile.cl

Abstract

The performance of random forest (RF) based satellite attitude control system (ACS) fault diagnosis methods is limited by uninformative features in high-dimensional data. To solve this problem, we proposed a feature-weighted random forest with Boruta (FWRFB) based fault diagnosis method is proposed for fault diagnosis of ACSs. Firstly, a Boruta feature selection algorithm is used to obtain a feature set and determine significant feature weights. Subsequently, a novel feature-weighted random forest (FWRF) algorithm is designed, which utilizes feature-weighted random sampling instead of simple random sampling to generate feature subsets in the RF. The FWRFB effectively utilizes the feature information while mitigating noise interference. Finally, a FWRFB-based diagnostic module is developed for online fault diagnosis of ACSs. The effectiveness of the proposed method is verified by the ACS data from a semi-physical simulation platform.

1. Introduction

The satellite attitude control system (ACS) is crucial to guarantee the normal operation of onboard loads and even the integrity of the entire satellite (Yuan, Song, Pan, Song, & Ma, 2021). As the subsystem of satellites has a high fault rate, sensor faults and actuator faults may occur in the ACS due to the complex and changeable space environment. Faults without immediate handling can lead to satellite performance degradation or even cause on-orbit mission failure (Ji, Zhang, & Liu, 2024). Therefore, the fault diagnosis of ACS is crucial for improving the safety and stability of satellites (Pourtakdoust, Mehrjardi, & Hajkarim, 2022).

A large amount of test data and telemetry data of ACSs can be easily obtained. Thus, data-driven fault diagnosis methods are more adequate and feasible to implement than model-based approaches; particularly if we cannot rely on prior knowledge or accurate model descriptions (Xiao & Yin, 2021). A variety of intelligent methods have been applied to ACS fault diagnosis (Yang & Zhong, 2022). The random forest (RF) based method is an ensemble method that combines decision trees (DTs) to form a strong classifier, achieving high robustness and accuracy in handling large-scale data (Wu, Chen, Qiu, & Zhou, 2024). It has been demonstrated that the RF algorithm is a suitable method for the fault diagnosis of ACSs (Huang et al., 2021). The effectiveness of RF-based methods usually depends on the classification power of extracted features from original data (Papakonstantinou, Daramouskas, Lappas, Moulianitis, & Kostopoulos, 2022). However, the extracted features affect the fault diagnosis performance differently, and some of them may encumber the fault diagnosis performance improvement.

Feature subsets are constructed by the way of simple random sampling in conventional RF algorithm (S. Chen, Yang, Zhong, Xi, & Liu, 2023). Features enter the feature subset of DT with same probability, which may restrain the function of features with high feature importance (Sun et al., 2011).

In addition, the use of uninformative features may generate noise in the RF-based fault diagnosis model, and affect the prediction accuracy of that model. Feature weighting simplifies mapping data to categories and helps improve classifier performance, making same-category data more compact spatially while allowing different-category data to be looser in structure (Cao, Xu, Liang, Chen, & Li, 2010). The enriched random forest (ERF) applies weighted random sampling, assigning lower weight to less informative features by the feature importance (Ghosh & Cabrera, 2022). With the weighted sampling strategy, the variable importance-weighted random forest (VIRF) can focus on the most informative features (Liu & Zhao, 2017). However, the accuracy of traditional feature importance is limited by the coupling of variables and the randomness of feature selection. Thus, these methods can only partially overcome the influence of uninformative features.

The performance of weighted random forest algorithms may be limited by weights transformed from the traditional feature importance. Moreover, uninformative features cannot be blocked from entering the feature subset due to positive nonzero weights. Feature selection methods can usually be used to reduce uninformative features, which are processes that identify and remove as many irrelevant features as possible. In (Eroglu & Akcan, 2024), a feature importance based feature selection method is designed to address difficulties in high computational demands and extracting valuable insights. A compound fault feature selection method based on the causal feature weighted network effectively reduces the number of features in optimal feature subsets (Yu, Li, Wu, Gao, & Wang, 2024). The accuracy of feature evaluation is crucial in the process of feature selection. However, the fluctuation problem of feature evaluation based on traditional feature importance is usually ignored in the existing literature.

This paper addresses the performance reduction issue caused by uninformative features and information differences, a fault diagnosis method based on feature-weighted random forest with Boruta (FWRFB) is proposed for the ACS fault diagnosis. First, the Boruta algorithm is developed to enhance the feature evaluation and eliminate irrelevant features. Then, the significant feature weights are utilized to improve the construction processes of feature subsets, forming a feature-weighted RF (FWRF) algorithm that enhances the utilization rate of critical features and leverages feature information differences. Finally, a FWRFB-based fault diagnosis method is trained and tested on the enhanced feature dataset, which is applied to the online fault diagnosis of ACSs. Based on the above analysis, the contributions of this paper can be summarized as follows:

1. An FWRFB based fault diagnosis method is proposed for ACSs to improve the efficiency of fault diagnosis.
2. A feature evaluation method based on Boruta algorithm is designed to obtain accurate feature weights and eliminate irrelevant features.
3. An FWRF algorithm is designed by a feature-weighted random sampling method to improve the strength of fault classifiers.
4. The proposed method is helpful because it can overcome uninformative features and enhance the role of features according to the importance.

The rest of this paper is organized as follows. The system and problem formulation is provided in Section 2. In Section 3, the Boruta algorithm and the FWRF algorithm are introduced and explained in detail. In Section 4, the simulation results on a satellite attitude control system are provided and discussed to illustrate the effectiveness of the proposed method. The conclusion and future works are presented in Section 5.

2. SYSTEM DESCRIPTION AND PROBLEM FORMULATION

2.1. System Description

The ACS is mainly composed of an attitude controller, attitude actuator, and sensor, which plays a crucial role in the completion of on-orbit tasks, with the structure shown in Fig. 1. The efficient attitude stabilization and maneuvering of satellites can be realized by the above components. The satellite's orientation information is measured by attitude sensors, and then the orientation is determined by obtained information in the attitude estimation unit. According to the measured data and the expected attitude, the satellite attitude can be determined by the attitude controller that can give a control law. Finally, a required control torque is generated by the actuator to realize the attitude control (Zhong, Liu, Zhou, Li, & Xue, 2019). Multiple signals from components of ACS can be collected for the fault diagnosis subsystem to monitor the system state. The fault diagnosis subsystem, as an essential part of the attitude controller, is critical to the safe and stable operation of the satellite.

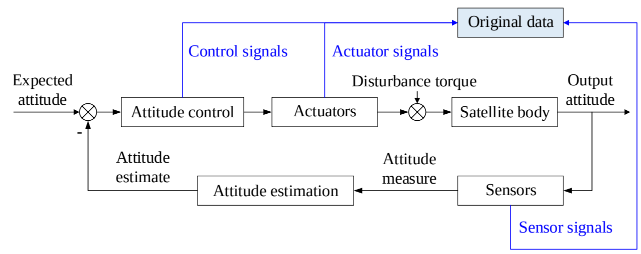

Figure 1. Schematic diagram of ACS.

2.2. Problem Formulation

When the ACS exhibits abnormal behavior, fault states can be directly inferred from variables of telemetry data such as voltage, current, temperature, attitude angle, and angular velocity. In literature (Suo, Zhu, & Yu, 2019), fuzzy Bayes risk and support vector machine methods are employed for diagnosing faults of satellite power systems using the telemetry data. However, the method performance is reduced by uninformative features generated during feature extraction. Irrelevant features degrade the performance of RF by affecting the strength of the decision tree. These irrelevant features may result in problems like overfitting and invalid calculations that degrade the performance of RF-based fault diagnosis algorithms while increasing computational and storage space requirements. In addition, the interpretability of fault diagnosis results is compromised by these irrelevant features.

For an RF classifier, an upper bound of its generalization error can be derived in terms of the strength of individual trees and the correlation between them:

$$PE \leq \frac{\bar{\rho}(1-s^2)}{s^2} \tag{1}$$

where PE is the generalization error, $\bar{\rho}$ is the average pairwise correlation between trees and s is the average single tree strength (Leo, 2001). Therefore, improving the strength of DT is effective way to enhance the RF model performance. The more fault information the used features contain, the better the performance of each DT. However, each feature extracted form telemetry data contains a different amount of fault information, and even some features do not contain fault information.

Due to the limitation of conventional feature importance, it is difficult to design an accurate feature importance based feature evaluation and selection method to leverage the different fault information of features. Although the feature set can be optimized by manual feature selection, relying on expertise in the satellite field, the process is time-consuming. The features can be evaluated by the traditional feature importance of RF, and then an enhanced feature set can be obtained by some feature selection methods. However, these methods are often plagued by inaccurate feature evaluation. It is difficult to get a suitable feature set. Therefore, this study focuses on addressing the performance reduction caused by irrelevant features in RF-based fault diagnosis algorithms and rationally using differences of feature information to improve the strength of DT.

3. MAIN RESULTS

The FWRFB based fault diagnosis approach consists of three steps, as shown in Fig. 2. First, many kind of discriminative features are determined to be extracted from original signals (described in Section 3.1). Second, the Boruta feature selection method (H. Chen, Hu, Han, & Miao, 2024) is designed to obtain a feature set with weights for the subsequent classification algorithm (described in Section 3.2). Third, the feature-weighted random forest algorithm is constructed based on the feature dataset with significant feature weights (described in Section 3.3). Finally, the established fault diagnosis approach is developed for on-line fault diagnosis of ACSs.

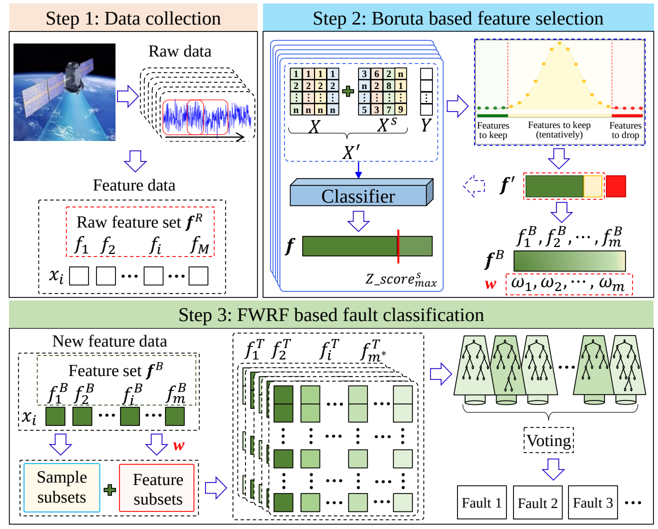

Figure 2. Structure of the proposed method

3.1. Feature Extraction

In order to effectively capture the fault characteristics of the raw signals for accurate fault diagnosis of ACSs, signal feature extraction is a critical step to obtain effective features. The feature dataset of fault diagnosis is formed by consolidating features extracted from original signals collected from the controller, actuators and sensors. In this paper, 16 time-domain features reported in (Guo et al., 2021) are extracted from the original data, as shown in Table 1.

In this paper, the signal data from ACSs is collected under multiple system sates, and consolidated into a dataset D_o. The feature variables are extracted from each segment of original signals. Each feature variable reflects a certain physical meaning. $c_1 \sim c_6$ reflect the amplitude properties, $c_7 \sim c_8$ reflect the energy properties, and $c_9 \sim c_{16}$ reflect the properties of time distribution. An initial feature sample matrix X with feature set $\boldsymbol{f}^R$ can be obtained by consolidating features extracted from the original data D_o. Each sample is labeled by the fault type to form label matrix Y. In matrix X, some columns are uninformative features generated by irrelevant signals and noise. Thus, these features can not provide effective information for fault diagnosis.

Table 1. Feature variable

Variable	Expression	Variable	Expression
c_1	$\frac{1}{N}\sum_{t=1}^{N} s(t)$	c_9	$\sqrt{\frac{\sum_{t=1}^{N}(s(t)-c_1)^2}{N-1}}$
c_2	$\frac{1}{N}\sum_{t=1}^{N} \lvert s(t)\rvert$	c_{10}	$\frac{\sum_{t=1}^{N}(s(t)-c_1)^3}{(N-1)c_9^3}$
c_3	$\max(s(t))$	c_{11}	$\frac{\sum_{t=1}^{N}(s(t)-c_1)^4}{(N-1)c_9^4}$
c_4	$\min(s(t))$	c_{12}	c_6/c_8
c_5	c_3-c_4	c_{13}	c_6/c_7
c_6	$\max(\lvert s(t)\rvert)$	c_{14}	$\frac{Nc_8}{\sum_{t=1}^{N}\lvert s(t)\rvert}$
c_7	$(\frac{\sum_{t=1}^{N}\sqrt{\lvert s(t)\rvert}}{N})^2$	c_{15}	$\frac{Nc_6}{\sum_{t=1}^{N}\lvert s(t)\rvert}$
c_8	$\sqrt{\frac{\sum_{t=1}^{N}(s(t))^2}{N}}$	c_{16}	$\frac{Nc_6}{\sum_{t=1}^{N}(s(t))^2}$

where $s(t)$ $(t = 1,2,\cdots,N)$ represents a time-domain signal.

3.2. Boruta based Feature Selection

During feature extraction, irrelevant features may be generated from irrelevant variables and noise, limiting the performance of subsequent fault classification algorithms. Thus, the feature evaluation method based on Boruta algorithm is designed to obtain accurate feature weights and eliminate irrelevant features. The running flow of the Boruta feature selection algorithm is shown in step 2 of Fig. 2. The Boruta algorithm is built around the RF algorithm, which has the advantages of shadow features and feature importance statistical analysis with the dynamic threshold. The detailed steps of Boruta algorithm are presented as follows.

1) Shadow Feature Formation

To remove correlations between features and label variables and among features, the values of each column of X are shuffled, and then all columns are permutated to obtain the shadow feature matrix X^s. The obtained randomized feature sample matrix X^s is added to the original feature sample matrix X, constructing the extended feature sample matrix X^e.

2) Feature Importance Analysis

The labeled dataset (X^e, Y) is fitted to RF for feature importance. The higher the feature importance, the more important the feature. The Z_score value of each feature is evaluated for comparing features' importance. An original feature is determined as important based on the fact that the feature has a higher Z_score than the maximum Z_score of shadow features $Z\_score^s_{max}$ (MZSF). Otherwise, it is determined to be unimportant. A hit is recorded if an original feature's Z_score is higher than the MZSF.

3) Two-sided Equality Test

A statistical test is performed for all features after running several RFs. The null hypothesis is that the feature's Z_score is equal to the MZSF. A two-sided equality test can reject the hypothesis when feature importance is significantly lower or significantly higher than MZSF. It is straightforward to compute limits for accepting and rejecting features for any number of RFs for a desired confidence level.

4) New Feature Set Construction

The features deemed as unimportant are removed from the original feature set, resulting in a new feature set that ends an iteration. The procedure is performed according to a predefined number of iterations or until all features are either rejected or conclusively deemed important, whichever comes first. In the former case, there are features left that are neither accepted nor rejected and are further referred to as undetermined features.

Finally, important and undetermined features are integrated as $\boldsymbol{f}^B = \{f_1^B, f_2^B, \cdots, f_m^B\}$, and the corresponding weights of features are $\boldsymbol{w} = \{\omega_1, \omega_2, \cdots, \omega_m\}$. To summarize, the pseudocode of the Boruta feature selection method is shown in Algorithm 1.

Algorithm 1 Boruta based feature selection

Input: feature sample data $X \in \mathbb{R}^{n\times M}$, sample label $Y \in \mathbb{R}^{n\times 1}$, feature set $\boldsymbol{f}^R$, number of iteration N_i, number of RF in each iteration N_R and confidence level α.
Output: feature set $\boldsymbol{f}^B$ and feature weights $\boldsymbol{w}$.

for i=1 to N_i **do**
    for j=1 to N_R **do**
        Features of X are shuffled to obtain X^s;
        X^s is added to X by row to obtain X^e;
        Run the RF on (X^e, Y);
        Obtain MZSF based on all features' Z_score;
        Assign the hit by Z_score greater than MZSF;
    end for
    Determine acceptance and rejection thresholds by α;
    Deem important features by acceptance threshold;
    Deem unimportant features by rejection threshold;
    Remove unimportant features from original feature set;
    Remove all shadow features;
    if all features have been deemed **then**
        break
end for
Obtain $\boldsymbol{f}^B$ and $\boldsymbol{w}$ based on retained m features;
return feature set $\boldsymbol{f}^B$ and feature weights $\boldsymbol{w}$.

3.3. FWRF based Fault Classification

From feature weights, it can be seen that the fault information carried by each feature has a different effect on the fault classification. To increase the performance of RF-based fault diagnosis algorithms, it is necessary to improve the feature subset formation way and enhance the contribution of features with high importance. Thus, a feature-weighted random forest algorithm is constructed by using feature-weighted random sampling instead of simple random sampling to form feature sets. The flow of the FWRF algorithm is shown in step 3 of Fig. 2. For building the FWRF-based fault diagnosis model, the feature data set $X^B \in \mathbb{R}^{n\times m}$ with label matrix $Y \in \mathbb{R}^{n\times 1}$ is generated by $\boldsymbol{f}^B$ with feature weights $\boldsymbol{w}$.

The FWRF is an improved RF-based algorithm that is a tree-structured ensemble learning method (Leo, 2001). Thus, the FWRF also consists of k DTs expressed as:

$$H = \{h_1(\mathscr{T}_1), h_2(\mathscr{T}_2), \cdots, h_k(\mathscr{T}_k)\} \tag{2}$$

where $\mathscr{T}_\theta$ $(\theta = 1,2,\cdots,k)$ is the input feature set of each DT, k is the number of DTs and $h_\theta(\mathscr{T}_\theta)$ $(\theta = 1,2,\cdots,k)$ represents the θ-th classification and regression tree (CART). The detailed steps of the FWRF are as follows.

1) Bootstrap Sample Subset

For each decision tree, n samples are sampled form X^B with replacement to form a bootstrap sample subset $X^{\mathscr{T}} = [x_1^{\mathscr{T}}, x_2^{\mathscr{T}}, \cdots, x_n^{\mathscr{T}}]^T$. Out of n times, each sample's probability of not being sampled is $(1-\frac{1}{n})^n$. As n approaches positive infinity, the limit of $(1-\frac{1}{n})^n$ is $\frac{1}{e} \approx \frac{1}{3}$. Thus, approximately one-third of unselected samples become out-of-bag (*oob*) samples, constituting the *oob* dataset for model testing.

2) Weighted Random Sampling based Feature Subset

The strength of DT can be improved by the eligible feature subset including important features. Important features are sampled in high probabilities by the weighted random sampling instead of simple random sampling. The sampling weights $\boldsymbol{w}$ is obtained in the above subsection. Thus, a feature subset $\boldsymbol{f}_i^T$ $(i = 1,2,\cdots,k)$ is generated by sampling $m^* = \lfloor\sqrt{m}\rfloor$ features from $\boldsymbol{f}^B$ in the way of weighted random sampling.

3) Decision Tree Establishment

The process of growing a decision tree requires constant node splitting. The Gini index is used to represent the classification impurity of a sample set (Zhu & Peng, 2022). Samples are divided into C classes under a feature t_i , and the Gini value of t_i can be expressed as follows:

$$Gini(t_i) = \sum_1^C P_j(1-P_j) = 1 - \sum_1^C P_j^2 \tag{3}$$

where P_j is the probability that the samples belongs to the j_{th} category. The optimal feature with a split point is selected from the feature subset by minimizing the Gini index shown below (Leo, 2001):

$$\min_a Gini(t_i, a) = \frac{n_1}{n} Gini(D_{t_{i_1}}) + \frac{n_2}{n} Gini(D_{t_{i_2}}) \tag{4}$$

where n is the total sample number with feature t_i, $D_{t_{i_1}}$ and $D_{t_{i_2}}$ are two sample subsets with feature t_i, and n_1 and n_2 are the sample numbers of $D_{t_{i_1}}$ and $D_{t_{i_2}}$ respectively. The decision tree cannot grow indefinitely and stops growing when each leaf node contains one class of samples.

4) Decision Tree Integration

The FWRF-based fault diagnosis model is obtained from the ensemble of established DTs. All DTs in the FWRF-based fault diagnosis model run in parallel. The fault type result for the sample can be obtained by each DT. Finally, the fault diagnosis result of the FWRF can be obtained by majority voting on the results of all DTs.

The training of the FWRF model can be stopped if the model testing error has met the effective classification. Model structures and parameters are recorded to complete the establishment of the FWRF model. Otherwise, the FWRF model is iteratively trained by repeatedly acquiring newly extracted sample subsets and feature subsets. To summarize, the pseudocode of the FWRF algorithm is shown in Algorithm 2.

Algorithm 2 Feature weighted random forest

Input: feature sample data $X^B \in \mathbb{R}^{n\times m}$, sample label $Y \in \mathbb{R}^{n\times 1}$, feature set $\boldsymbol{f}^B \in \mathbb{R}^{1\times m}$.

Output: fault diagnosis result.

1: Generate sample subset by the Bootstrap sampling method with replacement, simultaneously obtain *oob* dataset;
2: Generate feature subset by weighted random sampling method with weights $\boldsymbol{w}$;
3: Each DT is constructed by a feature sample data determined by a pair of a sample subset and a feature subset;
4: All results of decision trees are integrated by majority voting to obtain the result of FWRF;
5: The fault diagnosis result is determined by the classification result of FWRF;
6: **return** fault diagnosis result.

4. SIMULATION RESULTS

To verify the effectiveness of the proposed method, experiments are designed and performed on the data of ACSs. The original data consists of raw signals from sensors, actuators, and controllers. Finally, the proposed method is verified and compared with some typical machine learning (ML) methods on the feature dataset, and the experimental results are analyzed.

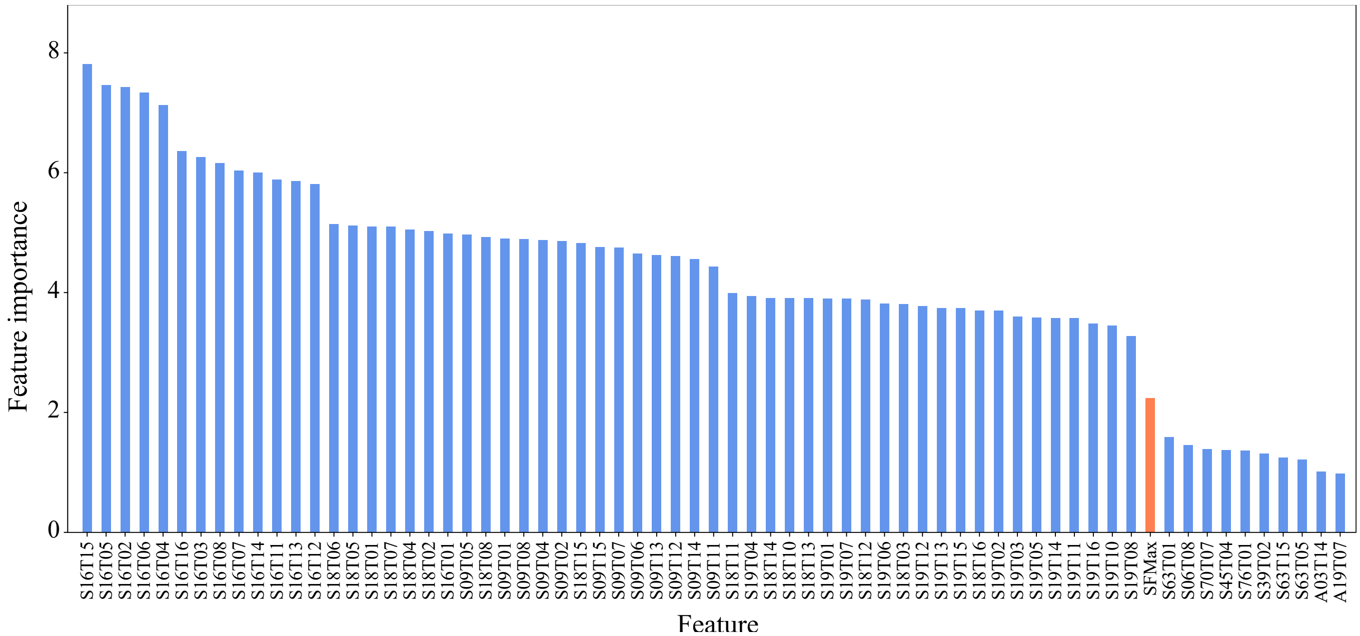

Figure 3. Feature importance of important features

4.1. Experiment Setup

In this paper, experimental data were collected under five typical states of ACSs: normal, constant deviation fault of control moment gyro (CMG), noise increase fault of gyroscope, saturation fault of gyroscope, and constant deviation fault of the gyroscope. A feature sample is generated by integrating features extracted from multiple segments at the same moment. After extracting features from the original signal data, the number of feature samples in each state of the ACS is 185. According to five states, samples are labeled as 1, 2, 3, 4, and 5, respectively. The number of samples and the label of each state are shown in Table 2. The original data contain 107 variables, and 16 domain features are extracted from a signal segment consisting of 8 sampling points for each variable. The feature vector is built by simultaneously integrating 1712 features extracted in a time window. Features are numbered according to their source and characteristics of the time domain in the form of $U-n-T-m$. U represents the source of the feature, consisting of the sensor S, the actuator A, and the controller C, and the corresponding n is numbered $1-76$, $1-27$, and $1-4$. T denotes the time domain feature, and m is the number among the 16 time domain features. For example, $S05T15$ represents the 15th time domain feature extracted from the fifth sensor variable.

4.2. Results and Discussion

The fault diagnosis method proposed in this paper has been validated using the above ACS data. The specific feature selection and fault diagnosis results are analyzed and discussed in the following paragraphs.

Table 2. The samples collected in experiment

Fault type	Fault location	Number	Labels
Normal	None	185	1
Constant deviation	Second CMG	185	2
Noise increase	First gyroscope	185	3
Saturation	Third gyroscope	185	4
Constant deviation	Fourth gyroscope	185	5

4.2.1. Feature selection

The Boruta algorithm is employed to eliminate irrelevant features and obtain the feature set along with feature weights. After dealing with the original feature set $\boldsymbol{f}^R$ containing 1712 features, 56 features are identified as important and 10 features are considered tentative, resulting in a feature set $\boldsymbol{f}^B$ comprising 66 features. A comparison between $\boldsymbol{f}^R$ and $\boldsymbol{f}^B$ reveals a large number of irrelevant features in $\boldsymbol{f}^R$. By averaging the feature importance obtained at each iteration of Boruta, the feature importance of 66 selected features is derived, as illustrated in Fig. 3. Finally, by normalizing feature importance scores, their respective weights in $\boldsymbol{w}$ are determined based on their proportionate contribution to the total sum of importance across all selected features.

4.2.2. FOFWRF based fault classification

The FWRF-based fault diagnosis model was constructed on the feature dataset $X^B \in \mathbb{R}^{925\times 66}$ optimized by the Boruta algorithm. To illustrate the advantages of the proposed method in the fault diagnosis of ACSs, the proposed method is compared with other ML algorithms, support vector machine (SVM), artificial neural network (ANN), K nearest neighbor (KNN), and naive Bayes (NB). Each algorithm is run on the feature set $\boldsymbol{f}^B$, and the performance of each algorithm is compared in terms of fault diagnosis accuracy and running time metrics. The comparison results are shown in Table 3. It can be seen that the metrics of the proposed method are better than other ML methods. These comparison results illustrate the superiority of the proposed method in terms of the ACS fault diagnosis.

Table 3. Performance comparison of different methods

Method	Accuracy (%)	Running time (s)
SVM	49.95	1.52
ANN	70.05	5.51
KNN	93.84	0.37
NB	93.95	0.22
FWRFB	**94.05**	**0.21**

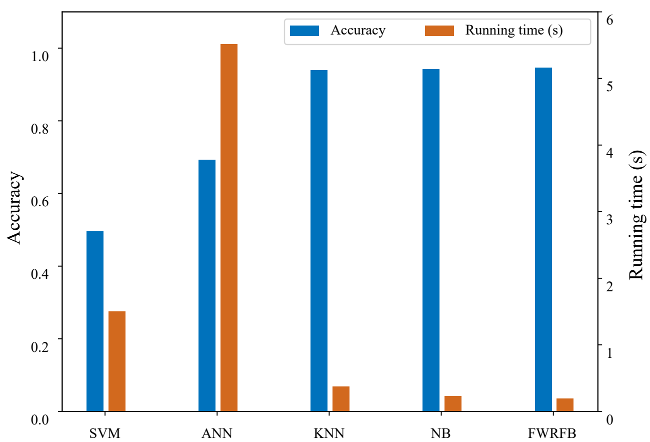

Figure 4. Performance comparison of different methods

5. CONCLUSIONS

This paper proposes a fault diagnosis method based on FWRFB for ACSs. First, the Boruta feature selection algorithm is employed to eliminate irrelevant features and obtain a feature set along with accurate feature weights. Subsequently, a feature-weighted RF-based fault diagnosis model is constructed by using the dataset under the enhanced feature set. Finally, the telemetry data from ACSs are used to validate the effectiveness of the proposed method. Experimental results demonstrate that the proposed method exhibits superior fault diagnosis accuracy and operational efficiency compared to other typical machine learning methods. Due to certain randomness of the feature selection, it is still tricky for the feature set obtained by the Boruta to completely overcome the effect of redundant features. In future work, the transition pattern from feature importance to weighted sampling weights will be optimized to enhance the classification performance of the feature-weighted RF algorithm.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China under Grants (62233012), and the Research Fund for the Taishan Scholar Project of Shandong Province of China. Marcos Orchard would like to thank FONDECYT Chile Grant Nr. 1210031 and the Advanced Center for Electricaland Electronic Engineering, AC3E, Basal Project FB0008, ANID.

REFERENCES

Cao, D., Xu, Q., Liang, Y., Chen, X., & Li, H. (2010). Automatic feature subset selection for decision tree-based ensemble methods in the prediction of bioactivity. *Chemometrics and Intelligent Laboratory Systems*, *103*(2), 129-136.

Chen, H., Hu, C., Han, B., & Miao, K. (2024). A method of diagnosing analog circuit soft faults using boruta features and lightgbm. *Electronics*, *13*(6), 1123.

Chen, S., Yang, R., Zhong, M., Xi, X., & Liu, C. (2023). A random forest and model-based hybrid method of fault diagnosis for satellite attitude control systems. *IEEE Transactions on Instrumentation and Measurement*, *72*, 3518413.

Eroglu, D. Y., & Akcan, U. (2024). An adapted ant colony optimization for feature selection. *Applied Artificial Intelligence*, *38*(1), 2335098.

Ghosh, D., & Cabrera, J. (2022). Enriched random forest for high dimensional genomic data. *IEEE/ACM Transactions On Computational Biology And Bioinformatics*, *19*(5), 2817-2828.

Guo, L., Cao, W., Bai, L., Zhang, J., Xing, L., Xiang, E., & Zhou, L. (2021). Fault diagnosis based on multi-scale texture features of cable terminal on emu of high-speed railway. *IEEE Transactions on Instrumentation and Measurement*, *70*, 3502612.

Huang, J., Liu, Y., Zhong, M., Yang, R., Li, W., & Liu, C. (2021). Fault diagnosis of satellite attitude control systems using random forest algorithm. *Journal of Astronautics (in Chinese)*, *42*(4), 513-521.

Ji, N., Zhang, Q., & Liu, J. (2024). Event-triggered adaptive vibration control for a flexible satellite with time-varying actuator faults. *Neurocomputing*, *584*, 127578.

Leo, B. (2001). Random forest. *Machine Learning*, *45*(1), 5-32.

Liu, Y., & Zhao, H. (2017). Variable importance-weighted random forests. *Quantitative Biology*, *5*(4), 338-351.

Papakonstantinou, C., Daramouskas, I., Lappas, V., Moulianitis, V. C., & Kostopoulos, V. (2022). Determination of threshold failure levels of semiconductor diodes and transistors due to pulse voltages. *Aerospace*, *9*(9), 164.

Pourtakdoust, S. H., Mehrjardi, M. F., & Hajkarim, M. H. (2022). Attitude estimation and control based on modified unscented kalman filter for gyro-less satellite with faulty sensors. *Acta Astronautica*, *191*, 134-147.

Sun, Z., Wang, G., Li, P., Wang, H., Zhang, M., & Liang, X. (2011). An improved random forest based on the classification accuracy and correlation measurement of decision trees. *Expert Systems with Applications*, *237*(Part B), 121549.

Suo, M., Zhu, B., & Yu, Z. (2019). Data-driven fault diagnosis of satellite power system using fuzzy bayes risk and svm. *Aerospace Science and Technology*, *84*, 1092-1105.

Wu, F., Chen, K., Qiu, G., & Zhou, W. (2024). Determination of threshold failure levels of semiconductor diodes and transistors due to pulse voltages. *IEEE Transactions on Industrial Electronics*, 1-11.

Xiao, B., & Yin, S. (2021). A deep learning-based data-driven thruster fault diagnosis approach for satellite attitude control system. *IEEE Transactions on Industrial Electronics*, *68*(10), 10162-10170.

Yang, R., & Zhong, M. (Eds.). (2022). *Machine learning-based fault diagnosis for industrial engineering systems*. CRC Press.

Yu, C., Li, M., Wu, Z., Gao, K., & Wang, F. (2024). Feature selection and interpretability analysis of compound faults in rolling bearings based on the causal feature weighted network. *Measurement Science and Technology*, *35*(8), 086201.

Yuan, Z., Song, N., Pan, X., Song, J., & Ma, F. (2021). Fault detection, isolation, and reconstruction for satellite attitude sensors using an adaptive hybrid method. *IEEE Transactions on Instrumentation and Measurement*, *70*, 3522112.

Zhong, M., Liu, C., Zhou, D., Li, W., & Xue, T. (2019). Probability analysis of fault diagnosis performance for satellite attitude control systems. *IEEE Transactions on Industrial Informatics*, *15*(11), 5867-5876.

Zhu, Y., & Peng, H. (2022). Multiple random forests based intelligent location of single-phase grounding fault in power lines of dfig-based wind farm. *Journal of Modern Power Systems and Clean Energy*, *10*(5), 1152-1163.

BIOGRAPHIES

Shaozhi Chen received his B.E. and M.E. degrees from Shandong University of Technology, Zibo, China, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, China.

His current research interests include fault diagnosis and remaining useful life prediction for satellite attitude control systems.

Xiaopeng Xi is a Post-doctoral Fellow with the Advanced Center for Electrical and Electronic Engineering (AC3E) at Universidad Técnica Federico Santa María, Valparaíso, Chile. His research interests include fault diagnosis, reliability analysis, degradation modeling, and remaining useful life prediction. He received the B.Eng. degree in detection guidance and control technology from the Beijing Institute of Technology, Beijing, China, in 2015, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2020.

Maiying Zhong received the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 1999.

From 2000 to 2001, she was a Visiting Scholar with the University of Applied Sciences Lausitz, Germany. From 2002 to 2008, she was a Professor with the School of Control Science and Engineering, Shandong University, Jinan, China. From 2006 to 2007, she was a Post-Doctoral Researcher Fellow with Central Queensland University, Rockhampton, QLD, Australia. From 2009 to 2016, she was a Professor with the School of Instrument Science and Opto-Electronics Engineering, Beihang University, Beijing, China. In 2016, she joined the Shandong University of Science and Technology, Qingdao, China, where she is currently a Professor with the College of Electrical Engineering and Automation. Her current research interests include model-based fault diagnosis, fault-tolerant systems, and their applications.

Marcos E. Orchard is Professor with the Department of Electrical Engineering at Universidad de Chile, Principal Investigator at the AC3E Center (Universidad Técnica Federico Santa María) and was part of the Intelligent Control Systems Laboratory at The Georgia Institute of Technology. His current research interest is the design, implementation and testing of real-time frameworks for fault diagnosis and failure prognosis, with applications to battery management systems, mining industry, and finance. His fields of expertise include statistical process monitoring, parametric/nonparametric modeling, and system identification. His research work at The Georgia Institute of Technology was the foundation of novel real-time fault diagnosis and failure prognosis approaches based on particle filtering algorithms. He received his M.S. and Ph.D. degrees from The Georgia Institute of Technology, Atlanta, GA, in 2005 and 2007, respectively. He received his B.S. degree (1999) and a Civil Industrial Engineering degree with Electrical Major (2001) from Catholic University of Chile. Dr. Orchard has published more than 100 papers in his areas of expertise.