Research on Optimization of 4G-LTE Wireless Network Cells Anomaly Diagnosis Algorithm based on Multidimensional Time Series Data

Bing Qian[1], Chong Ma[1][a] and Tong Zhang[2]

[1]*Beijing Research Institute, China Telecom Corporation Limited, Beijing, China*

[2]*Intel Corporation, Santa Clara, California, U.S.A.*

Keywords: Multidimensional Time Series Data, Anomaly Detection, Unsupervised Learning.

Abstract: With the continuous increase of network terminal equipment, the operation scenarios of 4G-LTE wireless networks are becoming more and more complex. The traditional manual method of analysis and screening of network cell equipment can no longer meet the needs of production. Therefore, an efficient wireless network cell abnormality diagnosis algorithm is needed to screen abnormalities of equipment to improve operation and maintenance efficiency. In view of the fact that the existing single-dimensional anomaly diagnosis algorithm cannot achieve fully automated detection and the existing multidimensional anomaly diagnosis algorithm has low detection efficiency on multidimensional time series data, there are a large number of errors and omissions. This paper proposes a multidimensional time series data based on 4G-LTE wireless network cell anomaly diagnosis optimization algorithm uses small-sample supervised algorithms to assist the training of massive-sample unsupervised algorithms, thereby improving the detection performance of unsupervised learning algorithms. This paper verifies the effectiveness of the optimization algorithm through experiments, and has a great improvement in the four commonly used unsupervised algorithms, which can well improve the anomaly detection capabilities of the existing algorithms.

1 INTRODUCTION

With the continuous development of communication technology, the layout of wireless networks has become more complex, and the operation and maintenance of network equipment has become more and more challenging. The number of existing 4G-LTE base stations is huge and there are many problems. However, the limited maintenance resources, the shortage of personnel, and the lack of support methods and platforms make it difficult to achieve in-depth and detailed maintenance. How to reduce the impact of faults on the business and improve user experience under the existing circumstances is the top priority of maintenance work. At present, the traditional operation and maintenance method of wireless base stations is to monitor equipment alarms and network indicators by engineers, identify abnormal points, and manually analyze, screen, locate, and process. The efficiency of manual screening is low, and the skill level of maintenance personnel is uneven, resulting in an inability to effectively improve maintenance efficiency. Therefore, in order to realize fault detection automation and reduce manual participation, it is necessary to develop a detection algorithm for wireless network cell abnormality.

The anomalies of wireless network cell can be classified into three categories: anomalous outliers, anomalous cycles, and anomalous collections (Chandola et al., 2009). As shown in Figure 1, in aperiodic data, if a single data point can be considered anomalous relative to other data, the data is called an outlier. In a periodic sequence, if the data is abnormal in a certain period but normal in other periods, the data is called abnormal period data. In time series collections, if the collection where the data is located is inconsistent with other sibling collections, the collection is an abnormal collection. This paper performs anomaly detection on wireless network cell devices. The above three anomalies need to be included. For a 4G-LTE wireless network cell, the device reports monitoring data every hour. The monitoring data contains multiple indicators, including PDCP (Packet Data Convergence Protocol) layer data flow, RRC (Radio Resource Control) connection times, CQI (Channel Quality Indicator) excellent

[a] https://orcid.org/0000-0001-8602-4676

Qian, B., Ma, C. and Zhang, T.
Research on Optimization of 4G-LTE Wireless Network Cells Anomaly Diagnosis Algorithm based on Multidimensional Time Series Data.
DOI: 10.5220/0010434000480057
In *Proceedings of the 6th International Conference on Internet of Things, Big Data and Security (IoTBDS 2021)*, pages 48-57
ISBN: 978-989-758-504-3

rate, and so on. Within a week's time sequence window, the point abnormality and periodic abnormality of each indicator at a certain moment will affect the failure judgment of a single network cell. At the same time, different sets of network cells need to be compared to detect anomalies that are different from other sibling network cell collections.

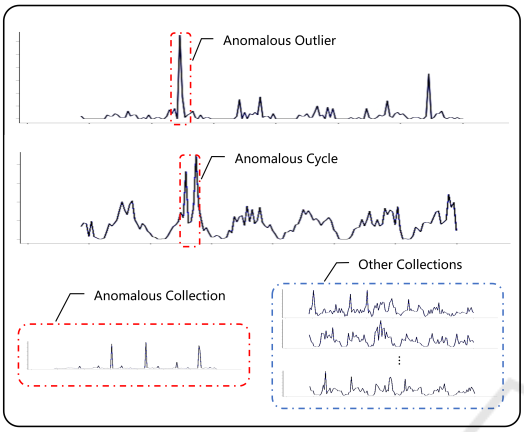

Figure 1: Three kinds of 4G-LTE wireless network cell anomalies: anomalous outliers, anomalous cycles, and anomalous collections.

2 RELATED WORK

In wireless network cell anomaly detection, the existing single-dimensional anomaly diagnosis algorithm, whether it is traditional machine learning such logistic regression (Kleinbaum et al., 2002) or deep learning algorithms such TCN (Bai et al., 2018), these algorithms firstly predict the index value at the future moment, then set the threshold of the difference between the predicted data and the real data to decide whether it is abnormal. This method has some limitations. On the one hand, it can only judge the abnormal value of a single indicator. To determine whether the network cell is abnormal according to the single indicator, it also needs to rely on the voting between the indicators or other manually formulated combination rules. On the other hand, this method can only detect point anomalies and partial periodic anomalies, and cannot compare the wireless network cell data set with other sibling sets. Therefore, the single indicator anomaly detection algorithm is not suitable for the scenario in this paper. This paper needs to be modeled by combining statistical feature extraction and multidimensional anomaly diagnosis algorithm. Statistical feature extraction mainly includes the construction of time series features and set features. Multidimensional anomaly diagnosis algorithms include supervised algorithms with labeled data, such as SVM(George and Vidyapeetham, 2012), ANN(Pradhan et al., 2012), and unsupervised algorithms with unlabeled data, such as k-Means(Wazid and Das, 2016). Generally, the results of supervised algorithms are more reliable and accurate than unsupervised algorithms. However, due to the amount of abnormal data is much less than normal data, a larger amount of data is required to train an effective supervision model, which means that it will cost a lot to label the data. Therefore, supervised anomaly detection algorithms are actually not suitable for large-scale multi-dimensional anomaly detection Scenes. Although unsupervised anomaly detection algorithms do not require labeling data and are more suitable for massive data scenarios, multidimensional unsupervised algorithms cannot select useful features, these mixed useless features will reduce the accuracy of unsupervised models. This paper designs a method of coupling supervised and unsupervised algorithms for training. We have obtained a small number of 4G-LTE wireless network cell annotation data. These data come from multiple operation and maintenance engineers, but we found that different operation and maintenance engineers have different understandings of the same data. They rely on their own operation and maintenance experience, and it is difficult to unify their opinions. Therefore, we believe that these annotation data not only contain reliable abnormal labels, but may also contain noisy normal data (False alarms), which is a low-quality annotation data. If a model with high accuracy is obtained through supervised algorithm training with this data, then its generalization performance on a large number of samples is not excellent. We first analyze these low-quality annotation data to find useful features, and then use these useful features to train unsupervised algorithms. The anomaly detection ability of the unsupervised model is improved through the coupling training of the unsupervised algorithm and the supervised algorithm.

General anomaly diagnosis algorithms such as anomaly detection based on measure density and KNN (Angiulli and Pizzuti, 2002), Auto Encoder based on neural network (Aggarwal, 2015), anomaly detection based on projected distance and PCA (Shyu et al., 2003), Isolation Forest (Aryal et al., 2014), One Class SVM (Wang et al., 2004), KDE (Kim and Scott, 2012), etc., cannot simultaneously find abnormal outliers, abnormal cycles, and abnormal collections. After comparing various algorithms, we selected the four algorithms with the best effects for analysis and subsequent experiments. As shown in Figure 2, it can be seen that KNN and

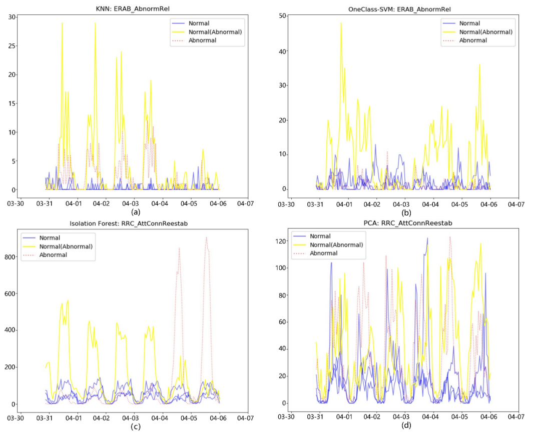

Figure 2: Abnormal state that could not be fully detected. Each curve represents the change in the value of a single indicator of the network cell within a week. Figure (a) represents the abnormal detection of the E-RAB_Abnormal indicator of the network cell using the anomaly detection algorithm based on measurement density and KNN. The abscissa is the time point, the ordinate is the indicator value. The red legend represents the detected abnormal curve. Blue represents the detected normal curve. Yellow represents a curve that the algorithm detects as normal but is actually abnormal. Figure (b) represents the detection result of the E-RAB_Abnormal indicator by the One Class SVM algorithm. Figure (c) represents the detection result of the RRC_AttConnReestab indicator by the Isolation Forest algorithm. Figure (d) represents the detection result of the PCA algorithm on the RRC_AttConnReestab indicator.

One Class SVM cannot perfectly detect wireless network cells different from other collections, such as E-RAB_AbnormRel (Evolved Radio Access Bearer Abnormally Released) anomaly. Isolation Forest and PCA also have the problem of missed detection of RRC_AttConnReestab (Radio Resource Control Attach Connection Reestablish) anomaly. These unsupervised algorithms are often unable to find out the anomalies in this scenario comprehensively. Therefore, based on the existing small number of expert system annotated samples and massive non-annotated samples, this paper designs a training method that combines supervised and unsupervised algorithms, which can improve the detection performance of unsupervised algorithms.

3 METHODOLOGY

We first defined 3 anomaly types for the time series data of 4G-LTE wireless network cells, and then we proposed a method to train an unsupervised anomaly diagnosis algorithm assisted by a supervised model.

3.1 Problem Definition

In this paper, all the time series window data is shown in Figure 3, which can be regarded as the set X, the single network cell time series window is the set X_n, the relationship between the two can be expressed as $X = X_1, X_2, \ldots, X_n$, n represents the number of network cells included in set X. The multidimensional data at a single moment in the time series window is S_t, $X_i = S_1, S_2, \ldots, S_t$, t is the time series length, and the multidimensional data $S_t = s_t^1, s_t^2, \ldots, s_t^k$, k represents the indicator dimension.

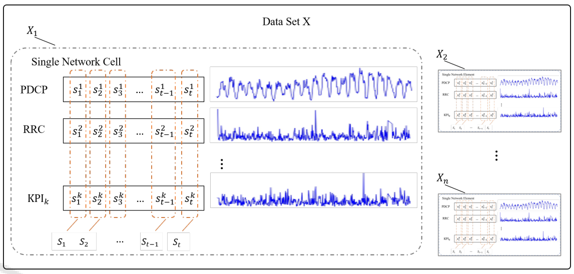

Figure 3: Time series window data.

The problem to be solved in this paper is that in the data set X containing many network cells, an abnormal network cell X_i is detected by a multidimensional unsupervised algorithm. The basis for judging the abnormality of the network cell X_i is that an indicator sequence $S_1^l, S_2^l, \ldots, S_n^l$ $, l \in (1, \ldots, k)$ in X_i has an anomalous outlier $S_{abnormal}^l$ (anomalous outliers) or an abnormal sub-sequence $S_a^l, S_(a+1)^l, \ldots, S_(a+t)^l, a \in (1, \ldots c_(n-t))$ (anomalous cycles), or the sequence (anomalous collection) is inconsistent with the indicator sequence changes of other network cells. Synthesize abnormal outlier, abnormal cycle detection and abnormal detection of network cell collections to determine abnormal network cell.

3.2 Our Method

This paper mainly assists the unsupervised algorithm to select important features through a supervised algorithm, and improves the performance of the unsupervised anomaly detection algorithm. As shown in Figure 4, in the 4G-LTE wireless network cell anomaly detection scenario, first, the features of the original data are constructed based on the statistical method to form the original feature data set, and the data is preprocessed. And then divided into annotated set and non-annotated set according to whether it has been labeled. Then use supervised algorithms such as XGBoost (Chen and Guestrin, 2016) to train the annotated set and calculate the feature importance, select the important feature set by sorting the important features (Chen et al., 2019), and filter the non-annotated data features, and finally use KNN, PCA, Isolation Forest, One Class SVM and other unsupervised algorithms are trained on non-annotated sets to obtain classification results.

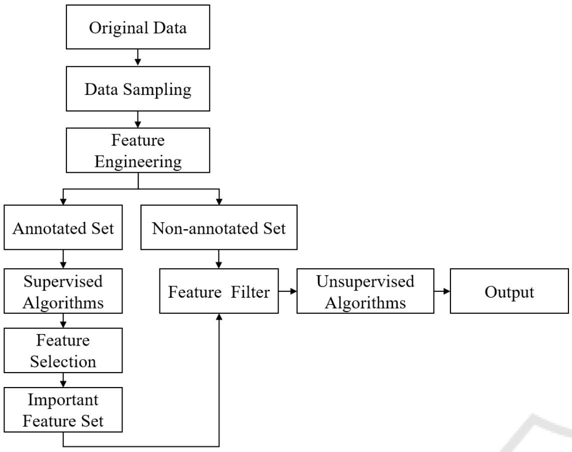

Figure 4: Our anomaly detection algorithm process.

Because the non-annotated data cannot be verified, it is only used in the real reasoning stage. As shown in Figure 5, in order to verify the effectiveness of the algorithm in this paper, we conduct experiments on the annotated data. First, 4 unsupervised algorithms (KNN, PCA, Isolation Forest, One Class SVM) are used to calculate the anomaly labels, and then vote together with the labels marked by experts. The rule is that if 3 of the 5 tags are marked as abnormal, the data is counted as an abnormal point, otherwise it is a normal point. Then construct features on the voting data through feature engineering, divide the training data and test data, and then train the XGBoost model, sort the importance of the features constructed by the feature engineering according to the XGBoost algorithm, and intercept the first 100 features as important feature sets. Then filter the features of the original annotation data, respectively train 4 unsupervised algorithms, and calculate the evaluation indicator according to the predicted label and ground truth. Finally, the effectiveness of the algorithm is verified by comparing the evaluation indicator of four unsupervised algorithms before and after feature selection.

4 DATA PREPROCESSING

In this paper, the original data is first screened, some data with more missing time series are removed. Then, some of the original indicators with higher correlation coefficients are deleted, because indicators with higher correlations have lower discrimination and will affect the training of the linear model. Next, construct statistical features and time series features of the remaining indicators through feature engineering. Finally, since what we obtained is a kind of low-quality and unreliable annotation data, in order to enhance the credibility of the annotation data, we use the unsupervised anomaly detection algorithm and the expert mark to perform majority voting to determine the anomaly label.

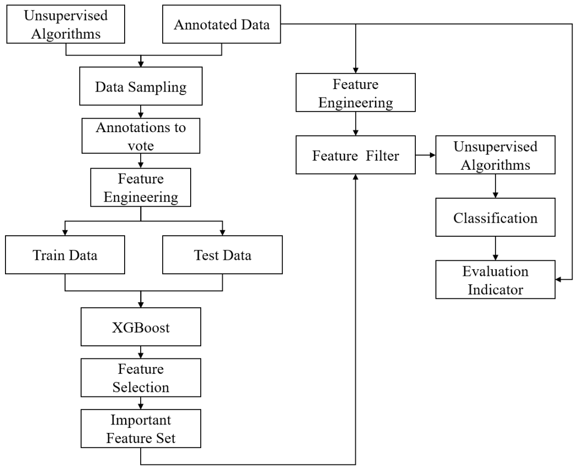

Figure 5: The process of validating the algorithm.

4.1 Data Sampling

Data Scenes Screening. The original data contains a total of 6 scenes of data, including high-speed rail, colleges, residential buildings, subways, etc. This paper selects the wireless network cell data of the residential scene. Because the data of the residential scene has a high proportion, and the data of the residential scene has a certain periodicity in time, it is convenient for experiment and analysis.

Data Cleaning. The data set of each wireless network cell should contain 7×24 hours of time series data, but in the actual data collection process, there are some data reports that are repeated or lost. This paper first removes the data with the same wireless network cell id and the same timestamp, then, the collection with less than 3% of missing cells is screened, and finally the number of wireless network cell collections is 4188, and the hourly granularity data is 688747.

4.2 Feature Engineering

Original Indicators. The original indicators are shown in Table 1, which contains 24 kinds of indicators.

Table 1: Original indicators.

Meaning	Name	Meaning	Name
PDCP traffic	pdcp	Same frequency switching success rate	HO_SuccOutIntraFreq_Rate
RRC connection times	rrc	Number of failed same frequency switching	HO_FailOutIntraFreq
Radio initial connection success rate	Radio_InitSuccConn_Rate	Inter-frequency switching success rate	HO_SuccOutInterFreq_Rate
S1 signaling connection establishment failure times	S1Sig_FailConnEstab	Number of failed inter-frequency switching	HO_FailOutInterFreq
RRC connection establishment failure times	RRC_FailConnEstab	CQI excellent rate	cqi_rate
E-RAB establishment failure times	ERAB_FailEstab	PRB average interference noise	phy_rrurxrssimean_chan1
Number of abnormal releases of UE context	UECNTX_AbnormRel	Packet loss number of uplink user interface of air port	PDCP_SduLossPktUl
UE context drop rate	UECNTX_Drop_Rate	Packet loss rate of uplink user interface of air port	PDCP_SduLossPktUl_Rate
E-RAB abnormal release times	ERAB_AbnormRel	Packet loss number of downlink user interface of air port	PDCP_SduLossPktDl
E-RAB drop rate	ERAB_Drop_Rate	Packet loss rate of downlink user interface of air port	PDCP_SduLossPktDl_Rate
RRC connection reestablish rate	RRC_ConnReestab_Rate	Packet discard number of downlink user interface of air port	PDCP_SduDiscardPktDl
RRC reconstruction request times	RRC_AttConnReestab	Packet discard rate of downlink user interface of air port	PDCP_SduDiscardPktDl_Rate

Correlation Analysis. Calculate the Pearson correlation (Lee Rodgers and Nicewander, 1988) between the original indicators two by two. The results are shown in Figure 6. The original indicators with correlation coefficient > 0.7 are selected and deleted. The deleted indicators are shown in Table 2.

Generate Features. Construct features from the 21 original indicators retained through feature engineering. This paper constructs 3 feature sets, namely statistical feature set, time feature set, and time series feature set. The statistical feature set calculates the maximum, minimum, mean, standard deviation, and median on the time series for a single indicator of each wireless network cell; the time feature set includes the hour corresponding to the time stamp and the day of the week, whether it is a weekend, whether it is a holiday; time series feature set include the maximum, minimum, mean, standard deviation, and median of a single indicator at the same hour in a week, and the value of a single indicator in the previous hour. The generated feature set is shown in Table 3.

Generate Labels. After the data is constructed through feature engineering, 4 unsupervised algorithms KNN, PCA, Isolation Forest, and One Class SVM are trained separately, and the prior anomaly ratios of the four algorithms are set to 1%, calculate the abnormal label through unsupervised algorithm, and

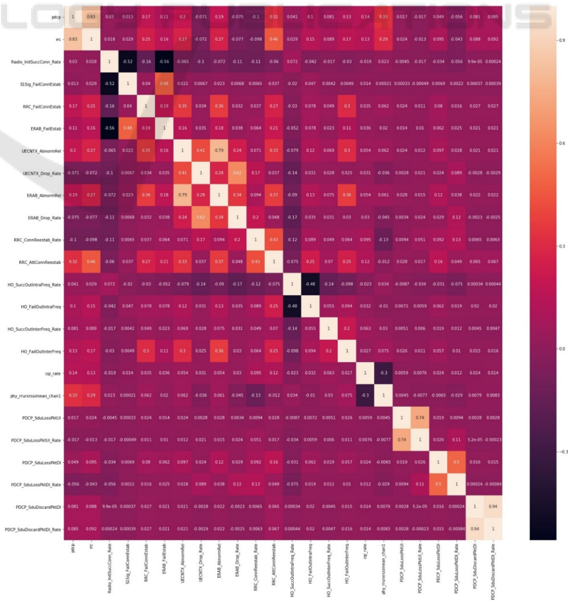

Figure 6: Correlation between original indicators.

Table 2: Delete the original indicator.

Indicator 1	Status	Indicator 2	Status	Correlation
pdcp	Keep	rrc	Delete	0.83
PDCP_SduDiscardPktDl	Keep	PDCP_SduDiscardPktDl_Rate	Delete	0.74
PDCP_SduLossPktUl	Keep	PDCP_SduLossPktUl_Rate	Delete	0.94

Table 3: Constructing features based on original indicators.

Feature set	Meaning	Name	Input (dim)	Output (dim)
Statistical Features	The maximum value of a single indicator in time series	kpi_max	21	105
	The minimum value of a single indicator in time series	kpi_min		
	The mean value of a single indicator in time series	kpi_mean		
	The standard deviation of a single indicator in time series	kpi_std		
	The median of a single indicator in time series	kpi_med		
Time Features	Current hour	hours	1	4
	Current day of the week	day_of_the_week		
	Whether it is weekend	is_week_day		
	Whether it is a holiday	is_vacation		
Time Series Features	The maximum value of a single indicator at the same time within a week	kpi_samehour_max	21	105
	The minimum value of a single indicator at the same time within a week	kpi_samehour_min		
	The average value of a single indicator at the same time within a week	kpi_samehour_mean		
	The standard deviation of a single indicator at the same time within a week	kpi_samehour_std		
	The median value of a single indicator at the same time within a week	kpi_samehour_med		
	The value of a single indicator at the previous moment within a week of the wireless network cell	kpi_last_hour	21	21

then vote with the label marked by the expert. If 3 of the 5 types of tags are marked as abnormal, the data is counted as an abnormal point, otherwise it is a normal point. As shown in Table 4, there were 684765 normal samples and 3982 abnormal samples.

Table 4: Data distribution.

Normal	Abnormal
684765	3982

5 FEATURE SELECTION

In this paper, the data processed by feature engineering is trained by supervised algorithms, and the important indicators are found through supervised algorithm. The purpose of this is to try to improve the detection performance of unsupervised algorithms through these important indicators.

5.1 Training Supervision Model

Data Set Division. After data processing, there are a total of 688747 pieces of training data, and each piece of data corresponds to 256 features. After the data is shuffled, the training data and the validation data are divided according to the ratio of 7:3. The ratio of the divided data set is shown in Table 5.

Table 5: Data set division.

Data Set	Normal	Abnormal
Train	418014	2545
Validation	179215	1025

Training Data Augmentation. In the training data in Table 5, the ratio of the positive sample to the negative sample is 1:164, and the data skew is serious, so the generalization ability of the model obtained by directly using the original data for training will be poor. Considering that the down-sampling data will cause too few samples and the model is easily overfitted, this paper uses the up-sampling algorithm to generate more abnormal samples. This paper uses the SMOTE (Chawla et al., 2002) algorithm to augment 2545 metadata, and finally the ratio of positive and negative samples approaches 1:1.

Hyper Parameter Optimization. This paper uses random search to adjust the hyperparameters, and the evaluation indicator is AUC (Walter, 2005). The parameter search range and optimal parameters are shown in Table 6.

Table 6: XGBoost parameter settings.

Parameter	Range	Optimal Value
Subsample ratio of columns when constructing each tree	[0.6, 0.7, 0.8, 0.9, 1.0]	0.7
Boosting learning rate	[0.1, 0.4, 0.45, 0.5, 0.55, 0.6]	0.55
Maximum tree depth for base learners	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	8
Minimum sum of instance weight needed in a child	[0.001, 0.003, 0.01]	0.001
Number of trees to fit	[1, 2, 3, 4, 5, ... , 18, 19, 20]	16

Evaluation. This paper selects Precision, Recall (Buckland and Gey, 1994), F1-Score (Sokolova et al., 2006), and AUC as evaluation indicators. The results are shown in Table 7. The experimental results show that the Recall and AUC indicators of the model are above 95%, which can distinguishes positive and negative samples well.

Table 7: XGBoost evaluation indicators.

Evaluation Indicator	Value(%)
Precision	83.9
Recall	95.8
F1-Score	89.6
AUC	97.9

5.2 Important Features Selection

This paper selects three ways to calculate the importance of features to XGBoost model, which are Frequency, Average Gain and Average Cover (Hastie et al., 2009). For each calculation method, the first 48 important features are calculated and the important feature set $f_i, i \in 1,2,3$ is formed. The final important feature set F $(F = \bigcup_{i=1}^{3} f_i)$ is obtained by combining the three sets. The capacity of the final feature set F is 100, as shown in Table 8, including 20 basic indicator fields and 2 information indicator fields.

6 RESULT

After the important feature selection in the previous chapter, we finally retained 100 important features for unsupervised model training. We respectively calculated the prediction results of the four unsupervised algorithms, KNN, PCA, Isolation Forest, and One Class SVM under all features (256 columns) and only important features (100 columns), then compared them with the expert annotation labels to obtain the Accuracy, Recall, F1-Score, and AUC of the prediction labels. As shown in Table 9, we find that the evaluation indicators of the four algorithms have been improved after the important feature selection, especially the Recall and F1-Score have improved significantly. Therefore, it can be proved that the detection performance of the unsupervised algorithm can be improved by screening important features with a small sample of supervised algorithms. Finally, we fused the prediction results of the four algorithms. The abnormal scores predicted by the four algorithms were weighted and fused according to the coefficient of 0.4: 0.3: 0.2: 0.1. The final Recall was 31.1% and F1-Score was 17.7%. Compared with the Recall of the four algorithms, the fusion result can cover more abnormal situations, and the F1-Score is not much lower, and the false detection of normal samples is also maintained at a reasonable level.

As shown in Figure 7, this paper shows the detection results of PDCP under the four algorithms. It can be seen from the figure that the algorithm results after feature selection are more accurate than the previous results. It can effectively detect indicator sets with large fluctuation ranges and less obvious fluctuations (compared to other stable sets), as well as some subsequences that are quite different from the normal period.

7 CONCLUSION

Combining Table 9 and Figure 7, we can find that the four unsupervised algorithms Isolation Forest, One Class SVM, KNN, and PCA are better than the results under the original features after the extraction of important features, and the Recall of each algorithm has Significantly improved, especially Isolation Forest and PCA, increased by 16.5% and 12.1% respectively, it shows that more abnormal samples have been detected. Combining normal samples and abnormal samples, the F1-Score of the four algorithms have also been greatly improved. The Isolation Forest and PCA have improved significantly, with 12.3% and 9.1% respectively. This shows that when more abnormal samples are detected, a large number of normal samples are not mistakenly detected as abnormal samples, which reduces the occurrence of false detections while reducing missed detections. Finally, with the support of computing power, the four algorithms

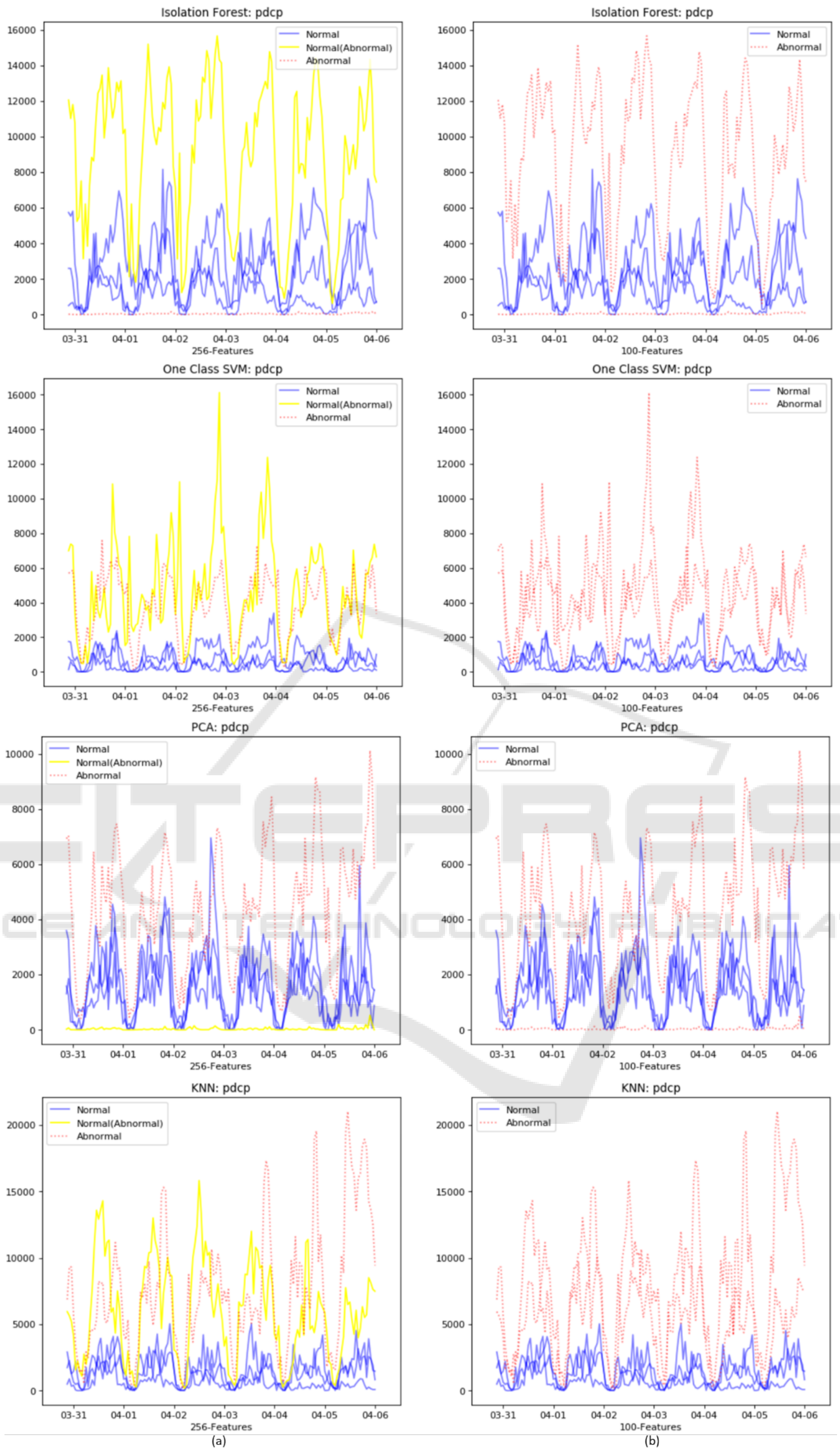

Figure 7: Comparison of detection results before and after feature selection. Figure (a) is the detection result obtained by training an unsupervised model based on 256-dimensional original features, and Figure (b) is an abnormal wireless network cell detected using 100-dimensional important features. From top to bottom, the detection results of Isolation Forest, One Class SVM, PCA, and KNN are selected. Each figure selected the PDCP indicators of multiple wireless network cells for display. The abscissa represents the time series point, and the ordinate represents the indicator value. The red legend represents the detected abnormal wireless network cell, the blue is the normal, and the yellow represents the abnormal data but the algorithm does not detect the situation (the algorithm judges the normal wireless network cell).

Table 8: Important features set.

Indicator Fields	Statistical Features
pdcp	kpi, last_hour, max, mean, med, min, samehour_max, samehour_mean, samehour_med, samehour_min, samehour_std, std
Radio_InitSuccConn_Rate	kpi, last_hour, min, samehour_max, samehour_mean
S1Sig_FailConnEstab	kpi, mean, samehour_med, samehour_min
RRC_FailConnEstab	last_hour, std
ERAB_FailEstab	kpi, mean, samehour_mean, std
UECNTX_AbnormRel	kpi, last_hour, mean, med, std
UECNTX_Drop_Rate	kpi, med, samehour_mean, samehour_med
ERAB_AbnormRel	kpi, last_hour, mean, samehour_mean
ERAB_Drop_Rate	med
RRC_ConnReestab_Rate	kpi, last_hour, mean, med, samehour_max, samehour_min
RRC_AttConnReestab	max, mean, samehour_mean, samehour_med
HO_SuccOutIntraFreq_Rate	kpi, last_hour, min, samehour_min, samehour_std, std
HO_FailOutIntraFreq	kpi, last_hour, samehour_med
HO_FailOutInterFreq	kpi, med, samehour_med, samehour_min
cqi_rate	kpi, last_hour, samehour_max, samehour_min, samehour_std
phy_rrurxrssimean_chan1	kpi, last_hour, min, samehour_max, samehour_mean, samehour_med, samehour_std, std
PDCP_SduLossPktUl_Rate	kpi, samehour_max, samehour_mean
PDCP_SduLossPktDl	kpi, last_hour, samehour_max, samehour_std
PDCP_SduLossPktDl_Rate	kpi, samehour_min, samehour_std
PDCP_SduDiscardPktDl_Rate	kpi, last_hour, max, mean, med, samehour_max, samehour_mean, samehour_med, samehour_min, samehour_std, std
hours	hours
day_of_the_week	day_of_the_week

Table 9: Comparison of evaluation indicators before and after feature selection.

Algorithm	Eval	256-D (%)	100-D (%)	Inc (%)
Isolation Forest	Accuracy	98.5	98.7	+0.2
	Recall	11.0	27.5	+16.5
	F1-Score	**8.2**	**20.5**	**+12.3**
	AUC	55.0	63.3	+8.3
OneClass-SVM	Accuracy	96.7	97.5	+0.8
	Recall	23.4	30.3	+6.9
	F1-Score	**8.2**	**12.5**	**+4.3**
	AUC	60.3	64.1	+3.8
PCA	Accuracy	98.5	98.6	+0.1
	Recall	8.4	20.5	+12.1
	F1-Score	**6.2**	**15.3**	**+9.1**
	AUC	53.7	59.8	+6.1
KNN	Accuracy	98.6	98.7	+0.1
	Recall	11.1	16.2	+5.1
	F1-Score	**8.5**	**12.6**	**+4.1**
	AUC	55.1	57.7	+2.6
Ensemble algorithms	Accuracy	-	98.3	-
	Recall	-	31.1	-
	F1-Score	**-**	**17.7**	**-**
	AUC	-	64.9	-

can be weighted and fused, and the Recall index after fusion is increased by 3.6%, and more abnormal wireless network cells can be detected after the fusion. In summary, the effect of constructing a feature set on the original data and performing anomaly detection through an unsupervised algorithm is relatively poor, while the detection effect of the same algorithm on the feature set after feature screening has been greatly improved. The meaning of this paper mainly includes two aspects. On the one hand, in the massive unlabeled data, building important feature sets through small samples of labeled data and supervised algorithms can assist the training of unsupervised algorithms, thereby improving the detection performance of unsupervised algorithms. On the other hand, through the optimization training of unsupervised algorithm, a large amount of data can be pre-annotated to provide an auxiliary decision-making role for the follow-up annotation work of experts.

In the future work, we will try to unify the opinions of different operation and maintenance engineers as much as possible to obtain higher quality annotation results. Although the evaluation indicators of this article have been improved, the current inconsistencies in the annotations have caused the final recall to be unsatisfactory, and this experiment only selected 4G-LTE wireless network cell data in a few regions. In the future, we will use data from more provinces for optimization and verification to better improve the current wireless network base station operation and maintenance methods.

REFERENCES

Aggarwal, C. C. (2015). Outlier analysis. In *Data mining*, pages 237–263. Springer.

Angiulli, F. and Pizzuti, C. (2002). Fast outlier detection in high dimensional spaces. In *European conference on principles of data mining and knowledge discovery*, pages 15–27. Springer.

Aryal, S., Ting, K. M., Wells, J. R., and Washio, T. (2014). Improving iforest with relative mass. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 510–521. Springer.

Bai, S., Kolter, J. Z., and Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Buckland, M. and Gey, F. (1994). The relationship between recall and precision. *Journal of the American society for information science*, 45(1):12–19.

Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority oversampling technique. *Journal of artificial intelligence research*, 16:321–357.

Chen, M., Liu, Q., Chen, S., Liu, Y., Zhang, C.-H., and Liu, R. (2019). Xgboost-based algorithm interpretation and application on post-fault transient stability status prediction of power system. *IEEE Access*, 7:13149–13158.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

George, A. and Vidyapeetham, A. (2012). Anomaly detection based on machine learning dimensionality reduction using pca and classification using svm. *International Journal of Computer Applications*, 47(21):5–8.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Kim, J. and Scott, C. D. (2012). Robust kernel density estimation. *The Journal of Machine Learning Research*, 13(1):2529–2565.

Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., and Klein, M. (2002). *Logistic regression*. Springer.

Lee Rodgers, J. and Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1):59–66.

Pradhan, M., Pradhan, S. K., and Sahu, S. K. (2012). Anomaly detection using artificial neural network. *International Journal of Engineering Sciences & Emerging Technologies*, 2(1):29–36.

Shyu, M.-L., Chen, S.-C., Sarinnapakorn, K., and Chang, L. (2003). A novel anomaly detection scheme based on principal component classifier. Technical report, Miami Univ Coral Gables FL Dept of Electrical and Computer Engineering.

Sokolova, M., Japkowicz, N., and Szpakowicz, S. (2006). Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation. In *Australasian joint conference on artificial intelligence*, pages 1015–1021. Springer.

Walter, S. D. (2005). The partial area under the summary roc curve. *Statistics in medicine*, 24(13):2025–2040.

Wang, Y., Wong, J., and Miner, A. (2004). Anomaly intrusion detection using one class svm. In *Proceedings from the Fifth Annual IEEE SMC Information Assurance Workshop, 2004.*, pages 358–364. IEEE.

Wazid, M. and Das, A. K. (2016). An efficient hybrid anomaly detection scheme using k-means clustering for wireless sensor networks. *Wireless Personal Communications*, 90(4):1971–2000.