Research Article

A Threat Intelligence Analysis Method Based on Feature Weighting and BERT-BiGRU for Industrial Internet of Things

Jingchen Yan , Zhe Du, Jifang Li, Shiduo Yang, Jinghao Li, and Jianbin Li

North China Electric Power University, School of Control and Computer Engineering, Beijing, 102206, China

Correspondence should be addressed to Jianbin Li; lijb87@ncepu.edu.cn

Received 9 April 2021; Revised 2 December 2021; Accepted 10 January 2022; Published 25 February 2022

Academic Editor: Xun Yi

Copyright © 2022 Jingchen Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The combination of 5G technology and the industrial Internet of things (IIoT) makes it possible to realize the interconnection of all things. Still, it also increases the risk of attacks such as large-scale DDoS attacks and IP spoofing attacks. Threat intelligence is a collection of information causing potential and nonpotential harm to the industrial Internet. Extracting network security entities and their relationships from threat intelligence text and constructing structured threat intelligence information are particularly important for IIoT security protection. However, threat intelligence is mostly text reports, which means the value information needs to be extracted manually by security analysts, and it is highly dependent on personnel experience. Therefore, this study proposes an IIoT threat intelligence analysis method based on feature weighting and BERT-BiGRU. In this method, BERT-BiGRU is used to classify attack behavior and attack strategy. Then, the attack behavior is weighted to make the classified result more accurate according to the relationship between attack strategy and attack behavior in ATT&CK for ICS knowledge. Finally, the possibility of attack and the harm degree of attack are calculated to form the threat value of the attack. The security analysts can judge the emergency response sequence by the threat value to improve the accuracy and efficiency of emergency response. The results indicate that the proposed method in this study is more accurate than the other standard methods and is more suitable for the unstructured threat intelligence analysis of IIoT.

1. Introduction

The developing application of 5G technology [1] improved the communication quality and made it possible to enable the perception and interconnection of infrastructure, personnel, and their environment. However, the interconnection between the network and external devices brought new threats in various angles and forms. Not only does the “threat surface” of external attacks become more extensive, but also the probability of equipment failures, software defects, and user errors increases, all of these will have a tremendous negative impact on the operation of the system. The blackout caused by the BlackEnergy Malware in the Ukrainian power grid [2] and the large-scale blackout in Venezuela [3] are two notable examples. According to the US Securities and Exchange Commission report and European financial report, the loss caused by an industrial infection in 2017 is as high as 1 billion dollars.

Facing the increasingly severe security situation, a new network security defense mechanism driven by threat intelligence has emerged. In 2013, Gartner proposed the concept of threat intelligence (TI) [4], which includes scenarios, tools, indicators, inferences, and feasible suggestions. It is evidence-based knowledge about the existing or emerging threats faced by assets, providing a decision-making basis for threat response [5]. Based on what was mentioned above, the threat intelligence contains detailed information about current or upcoming network security threats, which can help enterprises mine and analyze the attack behavior [6] and implement active network defense against network security threats.

At present, most of the threat intelligence in the industrial field provided by security companies [7, 8] is in an unstructured format, so it is difficult for security analysts and organizations to obtain standardized and structured threat information. Moreover, the extracted attack information

does not include a threat value, so it is difficult to provide accurate and effective emergency response countermeasures for IIoT security situational awareness systems or other defense mechanisms. Therefore, effective extraction of the valuable information of threat intelligence and converting it into a standardized and structured form are essential and significant for practical applications and fundamental research on the security of IIoT.

To solve the problems above, this study proposes a threat intelligence analysis method based on feature weighting and BERT-BiGRU for IIoT. Firstly, according to the matrix knowledge of ATT&CK for ICS, a large amount of threat intelligence was collected and standardized [9]. Secondly, a multilabel classification model was built based on the BERT-BiGRU model to identify the attack behavior of threat intelligence. Then, all the attack behaviors were weighted based on the dependence between strategy labels and behavior labels so that more accurate results of attack behavior identification were obtained. Finally, the attack behavior risk index was measured to form the attack behavior threat value.

Through this method, the attack behavior extraction of the threat intelligence of the IIoT is realized. By sorting the threat degree of the attack behavior, it provides a reference for emergency response and disposal, improving the security of the IIoT.

This study is organized as follows: in Section 2, the relevant research works are introduced. In Section 3, we describe the proposed method based on feature weighting and BERT-BiGRU in detail. In Section 4, the experiments are conducted, and the results are analyzed. Finally, we conclude the work and give the prospects for future research.

2. Related Work

2.1. Threat Intelligence Analysis. Threat intelligence analysis extracts unstructured data such as security warning notification, vulnerability notification, and threat notification from threat intelligence using natural language processing technology and helps the attacked analyze the behavior and vulnerabilities exploited by the attackers so that the attacked can make emergency defense decisions promptly.

Gao and Fan [10] used a graph database to analyze threat intelligence, indicated their properties and association relationship of industrial Internet security vulnerability data effectively and intuitively, and realized in-depth analysis and evaluation of vulnerability data. Wu et al. [11] proposed group tracer to automatically extract the TTP curve, to dig out behind the complex attack and potential attackers through the combination of network attack behavior threat intelligence knowledge. Liu et al. [12] analyzed the attack behavior events through threat intelligence and correlated the similar behavior according to the direction of the attack events to investigate the attack stage and protect it. Zhang et al. [13] proposed the EX-Action framework for extracting threat behavior from CTI reports. Ex-Action could detect threat actions using natural language processing (NLP) technology and identify threat actions using a multimodal learning algorithm. Zhang et al. [14] proposed a prediction method SIoT account malicious behavior based on threat

intelligence. It used SVMs to obtain the threat intelligence related to the target account's negative behavior. It analyzed the contextual data in the threat intelligence to predict the behavior of the malicious version. Preuveneers et al. [15] proposed the security enhancement framework of TATIS to timely respond to new vulnerabilities and attack forms in network attacks via threat intelligence analysis. Hinne [16] established a joint analysis model in network attack events and threat intelligence to analyze the attacker's motive and exploit the vulnerability, steps, and specific actions. In the process of event response, the attacker status is updated in real time, and decision analysis is provided.

However, the above methods do not provide comprehensive guidance to security analysts. In the face of an attack, it is difficult for security analysts to determine the emergency response sequence, resulting in heavy losses.

2.2. Multilabel Classification Analysis. A noticeable problem exists to automatically extract threat behavior from cyber threat intelligence reports in threat intelligence analysis. The threat intelligence contains various categories of data, such as threat behavior and attack stage. Thus, the threat behavior extraction problem can be abstracted into a multilabel classification problem.

Multilabel classification refers to separately analyzing the task text data with multiple labels. The calculation of multilabel classification tasks is more complicated than that of traditional classification tasks. It is mainly reflected in that the text features of a sample need to be associated with multiple labels, which require more advanced feature extraction and correct mapping to the corresponding labels. However, due to the complexity of data expression and the exponentiality of the label output space, the research on multilabel classification is still limited.

The current research mainly focuses on problem conversion (considering the labels are independent, converting the problem into two (multiple)) and algorithm adaptation (adapting the learning model to cope with the multilabel classification task) of these two aspects. Bernhard et al. [17] proposed a chain binary classification model to model the high-order association between labels. Yen et al. [18] proposed PDSparse to learn a separate linear classifier for each label. In the training process, all the positive labels and a small amount of active negative labels of each training sample can be distinguished by classifier via optimizing the label distribution. Yang et al. [19] proposed a labeled implicit Dirichlet model based on subdividing the data to reduce the time complexity of the multilabel classification algorithm. Tan and Liu [20] used the K-nearest neighbor graph to segment the relationship before the text label as a weakly supervised method. Then, the maximized posterior probability of the label value was utilized to construct a multilabel classification model, resulting in predicting the new label. Prabhu and Varma [21] optimized the nDCG algorithm to learn the structure pattern of the tree in the feature space dimension, then trained a binary classifier for each internal node, and, finally, predicted the label distribution of a given instance. Literature [22] used CNN and RNN to capture the

inner relationship of local and global semantic feature modeling labels. Xiao et al. [23] designed the LSAN model to determine the semantic connection between labels and documents using label semantic information. The self-attention mechanism is used to capture label data, and a label-specific document feature representation is constructed. Wehrmann et al. [24] proposed a multilayer output neural network model for multilabel classification; this structure has an output layer at each hierarchical level and provides a global output layer for the entire network to track the label dependency in the hierarchy as a whole by optimizing the sum of the global and each level of a loss function.

3. Threat Intelligence Analysis Method for IIoT Based on Feature

3.1. Weighting and BERT-BiGRU

3.1.1. Basic Notions

(1) *Threat Intelligence*. Threat intelligence [4] is a collection of information that can cause potential and non-potential harm to an enterprise. Threat intelligence describes attack events and attack behavior.

(2) *Attack Event*. IIoT attack event refers to a security threat event causing potential harm to the system or damage to system assets through various technical means. Usually, attackers use configuration defects, protocol defects, program defects, or violent attacks to attack the IIoT.

(3) *Attack Behavior*. Attack behavior refers to an action performed by an attacker to achieve a goal or gain some resources. Any attack event on the IIoT is composed of a series of attack behavior, based on which the whole process of an attack event can be depicted entirely. For example, the attacker can obtain the target system's TCP/IP subnet mask information by "network connection enumeration," and the "network connection enumeration" is recognized as an attack behavior.

(4) *ATT&CK for ICS*. ATT&CK for ICS [9] is a model and knowledge base reflecting the attack behavior of the industrial control system in each attack life cycle. It consists of three parts: strategy, technology, and process. The design represents what the attacker tries to achieve. Technology and process represent the behavior performed by the attacker to achieve the goal. At present, ATT&CK for ICS covers 11 attack strategies and 81 attack behaviors.

3.1.2. Overview of the Method. This study proposed a threat intelligence analysis method of IIoT based on feature weighting and BERT-BiGRU. The overview of the method is shown in Figure 1. Firstly, the threat intelligence data of IIoT on the open-source threat intelligence platform are collected, and the data preprocessing operations such as cleaning and denoising are completed; secondly, word segmentation and BERT sentence vector acquisition are

carried out on the preprocessed data, and a multilabel classification model based on BERT-BiGRU is constructed. The attack strategy and attack behavior of the threat intelligence are classified and identified. Based on the recognition result, all the behavior labels were weighted based on the dependence of the strategy label and its internal behavior labels to obtain more accurate attack behavior recognition results. Finally, the attack risk indicators were measured to obtain the attack behavior threat value. The threat value of attack behavior represents the harm degree of attack behavior, providing a reference for emergency response and disposal.

3.1.3. Threat Intelligence Analysis Method

(1) *Threat Intelligence Data Preprocessing*. There are usually multiple data sources during the threat intelligence data collection of IIoT, including homogeneous or heterogeneous databases, file systems, and service interfaces. Different data sources generally have complementarity and difference in data integrity, accuracy, and representation format. Different data sources are generally complementary and different in data integrity, accuracy, and presentation format and are vulnerable to noise data, missing data values, data conflicts, etc. Therefore, the collected data sets need to be preprocessed to ensure the accuracy, consistency, and high quality of the data analysis results.

It can be found from Figure 2 that the threat intelligence data preprocessing in this study is divided into three stages: standardization, cleaning, and reduction in threat intelligence data:

Standardization: the obtained data may have multiple structures and types. By the threat intelligence standard, we standardize the presentation of threat intelligence data from different sources. The concrete operation includes word root processing and morpheme processing. This process contributes to transforming these complex data into a single or manageable structure to achieve the goal of rapid analysis and processing.

Cleaning: not all data in attack events are valuable. There are some negligible data or even some data that are completely wrong distractions. Therefore, it is necessary to use various verification methods to remove inaccurate data (word abbreviations, unusual spacing, nonword characters, and any non-computer-related terms) that hinder classification. This study uses the filtering method to extract valuable data and label the information with confidence.

Reduction: this process is to merge the threat intelligence data. Feature reduction technology can reduce and simplify the size of the data set without compromising the accuracy of analysis results, which contributes to increasing the value density of the threat intelligence data. The feature reduction formulas are shown in

where α and β are, respectively, the set of measured values of two different types of features. n_1 and n_2 are the corresponding sample numbers. $SE(\alpha - \beta)$ is the variance of the

FIGURE 1: Overview of the method.

feature. The conflict of the feature is used to normalize the mean of the feature. The TEST function is built for comparison. As the deviation increases severely, the importance of this feature enhances. Otherwise, the importance of this feature decreases.

$$SE(\alpha - \beta) = \sqrt{\frac{\text{var}(\alpha)}{n_1} + \frac{\text{var}(\beta)}{n_2}}, \quad (1)$$

$$\text{TEST: } \frac{|\text{mean}(\alpha) - \text{mean}(\beta)|}{SE(\alpha - \beta)} > \text{threshold_value}, \quad (2)$$

3.1.4. Attack Behavior Identification and Classification

(1) *Identification of Attack Behavior and Attack Strategy.* This study designed a multilabel classification model based on feature weighting and BERT-BiGRU for attack recognition, as shown in Figure 3. The classification model consists of the BERT model and the BiGRU model. The BERT model is only used to extract a sentence representation, whereas the BiGRU model is used to classify attack behavior and attack strategy in threat intelligence. Firstly, the preprocessed threat intelligence content is input into the BERT model, and the vector representation is performed after two pretraining tasks of the model. Subsequently, the vector representation fused with the full-text semantic information is output. Then, the output of the BERT model is input into the BiGRU model. The BiGRU model extracts the abstract features of threat intelligence through the fully connected (FC) layer by word vector mapping. It facilitates feature extraction by adding an attention mechanism before the FC layer to give a higher weight to essential attributes. To complete the multilabel classification task of attack behavior and attack strategy in threat intelligence, an FC layer and softmax need to be connected to the model to classify the deep semantic features of the threat intelligence text.

(2) *Feature Weighting of Attack Behavior.* According to the ATT&CK for ICS knowledge, an attack strategy connects many different attack behaviors, and there exists a dependency between attack strategy and attack behavior. For

example, when the probability of an attack strategy increases, the probability of an attack behavior within the strategy will rise accordingly. The current attack threat can be dealt with more accurately by analyzing and extracting the relationship between attack strategy and attack behavior. As a result, based on the relationship between attack strategy and attack behavior, the attack behavior feature weighting method is designed in this study, and the critical steps of this method are shown in the following formulas:

$$\partial = \left(\frac{1}{\text{EXP}} \left((\text{Labeled} - \text{Tact}_{\phi}) * (-1) \right) * \left(\frac{1}{10} \right) \right), \quad (3)$$

$$\text{Labeled} - \text{Tech}_{\Delta} = \text{Labeled} - \text{Tech}_{\Delta} * (1 + \partial). \quad (4)$$

When a specific attack behavior occurs, its associated attack strategy must also exist. Analyzing attack strategy is more accessible than analyzing attack behavior, and the attack strategy analysis result is more accurate. The result of attack strategy identification is processed exponentially to optimize the analysis effect of attack behavior. ∂ is the value of attack strategy identification result $\text{Labeled} - \text{Tact}$ processed in exponential form. $\text{Labeled} - \text{Tech}$ is the feature weighting attack behavior identification result.

Based on what is mentioned above, the in-depth analysis of the threat intelligence data is successfully realized, which can output the structured attack behavior labels and the corresponding probability value with high accuracy and readability.

3.1.5. *Generation of Attack Behavior Threat Value.* Based on the method above, the identification of the attack behavior in the threat intelligence is realized. However, it is difficult to judge the threat degree of the attack behavior, resulting in difficulty in priority warning to the attack with a higher harm degree. Therefore, it is necessary to consider the threat of attack behavior and quantify such important risk indicators, including the possibility of attack behavior and the harmful degree of attack behavior. Based on the data of Common Attack Pattern Enumeration and Classification (CAPEC) [25] published by MITRE, the possibility and harm degree of the attack behavior were calculated in this study, and the threat value of attack behavior was formed.

FIGURE 2: Threat intelligence data preprocessing.

CAPEC is an enumeration and classification data set of attack types established by the United States Department of Homeland Security in 2007, a widely accepted and recognized public standard for attack modes, as shown in Table 1. CAPEC comprises two indicators of “likelihood of attack” and “typical severity.” Both were classified into five levels, including “very low,” “low,” “medium,” “high,” and “very high.” The “likelihood of attack” represents the probability of successful attack behavior. It considers relevant factors, including attack prerequisites, required attacker resources, and possible countermeasures. The “typical severity” aims to reveal the severity degree of the consequences of successful attack behavior.

The attack behavior labels in CAPEC and ATT&CK for ICS are the same. When we calculate the threat value of attack behavior, we firstly map the attack behavior in ATT&CK for ICS to the CAPEC, as shown in formula (5). Secondly, we quantify the unstructured CAPEC level labels to 1–5, as shown in formulas (6) and (7). Since the “typical severity” indicator is more valuable for attack defense, it is given a higher weight, as shown in formula (8). Then, the CAPEC indicator score and the attack behavior label classification result are combined to form the threat score of each attack behavior, as shown in formula (9).

$$\text{Tech}_i, \text{Tact}_i \leftarrow \text{analysisCAPECtext}, \quad (5)$$

$$\{x | x = 1, 2, 3, 4, 5\} \leftarrow \text{LikelihoodScore}, \quad (6)$$

$$\{y | y = 1, 2, 3, 4, 5\} \leftarrow \text{SeverityScore}, \quad (7)$$

$$\text{Sc}(\text{CAPEC}) = \text{LikelihoodScore} + \text{EXP}(\text{SeverityScore}), \quad (8)$$

$$\text{Sc}(\text{Tech}_\Delta) = \text{Labeled} - \text{Tech}_\Delta * \text{Sc}(\text{CAPEC})_\Delta. \quad (9)$$

The threat value of each attack behavior can be obtained based on the above steps. Security personnel can determine the corresponding emergency response sequence according to the threat value of the attack behavior.

4. Experiments and Analysis

4.1. Experiment 1: Demonstration of the Results of Proposed Method. The analysis of the Industroyer attack [26] is taken as an example. It is shown in Figure 4. Industroyer is a malicious program that could destroy the critical assets of the industrial control system. It invaded and attacked the Ukrainian power grid in 2016, causing a significant impact. It has been one of the biggest threats to the industrial control system since the “earthquake network virus” appeared, bringing about a large-scale power outage and property loss. Figure 4 shows an Industroyer attack event text on the IBM security platform, and Figure 5 shows the identification results obtained using the proposed method in this study. To enhance the readability of the identification results, the truth map was constructed using the Neo4j technology, as shown in Figure 6. After matching with the detailed list of Industroyer attacks provided by the MITRE platform and widely identified by security experts [27], the accuracy and recall values of the identification results using the present method are as high as 89.87% and 87.1%, which are much higher than those obtained using other methods.

4.2. Experiment 2: Comparative Experiment with Other Methods. In this section, we conduct three groups of comparative experiments, which are, respectively, as follows: 1. comparison between the present method and the BERT-BiGRU method without feature weighting; 2. comparison between the present method and the KNN and random


```

C:\WINDOWS\system32\cmd.exe
C:\Users\msduz\Desktop\ICS\ICS>python TI_cmd.py -p -i example_ctr/Industroyer.txt
The result of the given report are :
Tactics :
(1)The 'Initial Access' is: 0.8756234684519051
(2)The 'Impact' is: 0.8075148792700383
(3)The 'Inhibit Response Function' is: 0.7747942905623486
(4)The 'Impair Process Control' is: 0.7389815956263774
(5)The 'Collection' is: 0.6252294290204957
(6)The 'Execution' is: 0.5745100389107707
(7)The 'Discovery' is: 0.5628070034703921
(8)The 'Evasion' is: 0.5431484373380369
(9)The 'Lateral Movement' is: 0.47246689263774555
(10)The 'Command and Control' is: 0.3603938858768858
(11)The 'Persistence' is: 0.2827586085827272

Techniques :
(1)The 'Denial of Service' is: 0.7586039217099104
(2)The 'Loss of Control' is: 0.6851403913585424
(3)The 'Loss of Productivity and Revenue' is: 0.6621759368015425
(4)The 'Spearphishing Attachment' is: 0.6364714729199587
(5)The 'Denial of Control' is: 0.6211624928825441
(6)The 'External Remote Services' is: 0.572684936262641
(7)The 'Automated Collection' is: 0.5663564165641558
(8)The 'Role Identification' is: 0.5227586085824928
(9)The 'Device Restart/Shutdown' is: 0.5116752962116461
(10)The 'Damage to Property' is: 0.4745792567632514
(11)The 'Unauthorized Command Message' is: 0.4638579152750431
(12)The 'Loss of Safety' is: 0.43919810438112556
(13)The 'Loss of View' is: 0.42967043812296753
(14)The 'Data Historian Compromise' is: 0.4251528906592567
(15)The 'Modify Control Logic' is: 0.41926161565751090
(16)The 'Remote System Discovery' is: 0.4113321311090233
(17)The 'Block Reporting Message' is: 0.40431211337337273
(18)The 'Block Command Message' is: 0.4031212311231031
(19)The 'Control Device Identification' is: 0.39970234709954995
(20)The 'Manipulation of Control' is: 0.38619678817492865
(21)The 'Service Stop' is: 0.37027874925620737
(22)The 'Denial of View' is: 0.36639123411357796
(23)The 'Data Destruction' is: 0.3657156579779771
(24)The 'Manipulation of View' is: 0.3538611111501616
(25)The 'Masquerading' is: 0.34752470253393912
(26)The 'Network Connection Enumeration' is: 0.3393795950713506
(27)The 'Modify Parameter' is: 0.32662181332790811
(28)The 'Serial Connection Enumeration' is: 0.31713590501761838
(29)The 'Brute Force I/O' is: 0.306966572666892
(30)The 'Command-Line Interface' is: 0.30196111711351013
(31)The 'Activate Firmware Update Mode' is: 0.29668027809184928
(32)The 'Block Serial COM' is: 0.2858751699126088
(33)The 'Theft of Operational Information' is: 0.27086641135460255
(34)The 'Supply Chain Compromise' is: 0.2643857852965705

```

FIGURE 5: Identification results by the proposed method.

forest; and 3. comparison between the present method and the SyntaxNet method. The results are analyzed and discussed from two aspects of accuracy and recall.

The Python platform was used to train the model. We collected more than 4000 threat intelligence of industrial control systems from the open-source threat platform,

FIGURE 6: Logic map of attack behavior.

TABLE 2: Accuracy and recall of the two methods.

	No feature weighting added	Feature weighting added (the present method) (%)
Accuracy rate	86.69%	89.87
Recall rate	83.6%	87.1

FIGURE 7: Accuracy of the proposed method, KNN, random forest, and bagging.

FIGURE 8: Recall rate of the proposed method, KNN, random forest, and bagging.

TABLE 3: Accuracy and recall of each method.

	The present method (%)	SVM (%)	KNN (%)	Bagging (%)	Random forest (%)
Accuracy rate	89.87	71.67	61.4	67.81	64.8
Recall rate	87.1	62.14	50.78	56.12	58.4

forming the training set. In addition, we randomly selected 100 threat intelligence of more than ten mainstream attacks faced by the current industrial control system, including Industroyer, WannaCry, and Stuxnet, forming the test data set of the present experiment.

To verify the value of “feature weighting operation” in improving attack analysis capability, we compared the methods with feature weighting and without feature weighting. The results are shown in Table 2. The accuracy and recall of the method without feature weighting are 86.69% and 83.6%, respectively, whereas they are significantly improved with feature weighing, which can be as high as 89.87% and 87.1%. The performance of the method with feature weighing appears much better.

Many researchers have carried out text analysis in recent years and achieved good results using SVM [28], bagging algorithm, KNN classification algorithm, decision tree algorithm, random forest algorithm, neural network model, and other methods. Therefore, we selected four typical methods of SVM, bagging algorithm, K-nearest neighbor classification algorithm, and random forest algorithm to compare and analyze the attack behavior identification results with the method proposed in this study to verify the effectiveness of the current method. It can be seen from Figures 7 and 8 that the accuracy and recall rate can be significantly improved with the iterative training of the amount of a large sample. As shown in Table 3, the accuracy rate of the present method, SVM, K-nearest neighbor classification algorithm (KNN),

bagging algorithm, and random forest algorithm is 89.87%, 71.67%, 61.4%, 67.81%, and 64.8%; the recall rate of the present method is 87.1%, 62.14%, 50.78%, 56.12%, and 58.4% [29].

The results indicate that the proposed method in this study is ideal with higher accuracy and a better recall. The result of the SVM model is slightly better than other methods, but the accuracy is much lower than that of this method. KNN classification method needs no training, and it is time-saving, but it possesses the disadvantages of common computing capability. The accuracy of the bagging method is high. Still, all the predicted variables are considered during training, and the more robust predicted variables are placed at the top split point of the method. Hence, the reliability of this method is relatively low. The random forest method uses the decision tree as the primary classifier, improving the overall recall rate. However, due to the high number of iterations, it is time-wasting and accessible to overfitting.

4.3. Experiment 3: Usefulness Verification of Attack Behavior Threat Value. We hired four safety experts to carry out the experiments with us. By comparing the experimental results of the proposed method in this study with the evaluation results of security experts, the rationality of the attack threat value generation method proposed in this study is verified.

In the experiment, “man-in-the-middle attack,” “flooding,” “spear phishing,” and “code inclusion” are selected to simulate

TABLE 4: Evaluation scores.

Attack behavior	Expert evaluation Score_1	Expert evaluation Score_2	Expert evaluation Score_3	Expert evaluation Score_4	Our approach score
Man-in-the-middle attack	0.76	0.73	0.79	0.75	0.75
Flooding	0.89	0.86	0.91	0.88	0.92
Spear phishing	0.59	0.57	0.58	0.60	0.55
Code inclusion	0.12	0.10	0.15	0.13	0.14

the attack on the industrial control system. The present method and security experts in this study analyze and evaluate the threat degree of different attacks, respectively. Among them, the evaluation score is in the range of 0–1, and the threat degree of attack behavior is directly proportional to the score. The evaluation results are shown in Table 4.

It can be seen from Table 4 that the evaluation results of the method proposed in this study are consistent with the evaluation opinions given by security experts. The threat of attack behavior is flooding > man-in-the-middle attack > spear phishing > code inclusion. Experiments show that the attack behavior threat value generation method proposed in this study can effectively analyze the threat degree of attack behavior and early warning response. According to the threat score of attack behavior, the information security analysts of IIoT can take corresponding precautions to ensure the safe operation of IIoT.

By comparing the experiments and the current classical text analysis and attack behavior evaluation methods, the IIoT threat intelligence analysis method based on feature weighting and BERT-BiGRU proposed in this study possess advantages in accuracy and recall. In addition, it is more effective in evaluating threat behavior score, which is much closer to the score assessed by experts, resulting in more practical.

5. Conclusions

This study presents a threat intelligence analysis method of IIoT based on feature weighting and BERT-BiGRU. This method can automatically identify and classify the attacks in threat intelligence and calculate the threat value of each attack behavior. The threat value can provide a reference for the judgment of emergency response sequence and improve the accuracy and efficiency of emergency response, resulting in adequate security protection for 5G-oriented IIoT. The experiments show that the proposed method is more accurate than the other common methods and is more suitable for the unstructured threat intelligence analysis of IIoT.

In the future, we will complete an affair map based on threat intelligence to improve our emergency response capabilities to attack further.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] D. Minoli and B. Occhiogrosso, "Practical aspects for the integration of 5G networks and IoT applications in smart cities environments," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 5710834, 2019.
- [2] B. Antiy, "Comprehensive analysis of attacks on Ukraine's power system," *Information Security Research*, vol. 2, no. 6, p. 243, 2016.
- [3] L. Fang, L. Huang, Q. Zhao, and A. Pan, "Power grid security in super large cities from the perspective of blackout in Venezuela," *Power and Energy*, vol. vol. 4, no. 6, pp. 674–677, 2019.
- [4] M. Bromiley, "Threat intelligence: what it is, and how to use it effectively," *SANS Institute InfoSec Reading Room*, vol. 15, p. 172, 2016.
- [5] M. Broda, M. Hervieux, and H. Habib, "Cyber threat intelligence threat and vulnerability assessment of service supplier chain," *U.S. Patent Application*, 2020.
- [6] X. Rong and D. Song, "Research on cyber attack defense system based on big data and threat intelligence," *Journal of Information Security Research*, vol. 5, pp. 383–387, 2019.
- [7] A. AlienVault, "Threat intel resources," <https://start.me/p/rxRbpo/ti>.
- [8] IBM, "Threat Intel Resources," <https://exchange.xforce.ibmcloud.com/>.
- [9] B. Strom, A. Applebaum, D. Miller, N. Nickels, A. Pennington, and C. Thomas, "Mitre ATT&CKTM: Design and Philosophy," MITRE, McLean, VA, USA, 2018, <https://www.mitre.org/sites/default/files/publications/pr-18-0944-11-mitre-attack-design-and-philosophy.pdf>.
- [10] J. Gao and L. Fan, "Kernel-based weighted discriminant analysis with QR decomposition and its application to face recognition," *WSEAS Transactions on Mathematics archive*, vol. 10, no. 10, pp. 358–367, 2011.
- [11] Y. Wu, C. Huang, X. Zhang, and H. Zhou, "GroupTracer: automatic attacker TTP profile extraction and group cluster in Internet of things," security and communication networks," vol. 2020, Article ID 8842539, 14 pages, 2020.
- [12] L. Liu, J. Lin, Q. Wang, and X. Xu, "Research on network malicious Code detection and provenance tracking in future network," in *Proceedings of the 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, Lisbon, Portugal, July 2018.
- [13] H. Zhang, G. Shen, C. Guo, Y. Cui, and C. Jiang, "EX-action: automatically extracting threat actions from cyber threat intelligence report based on multimodal learning," *Security and Communication Networks*, vol. 2021, Article ID 5586335, 12 pages, 2021.

- [14] H. Zhang, Y. Yi, J. Wang, N. Cao, and Q. Duan, "Network attack prediction method based on threat intelligence for IoT," *Multimedia Tools and Applications*, vol. 78, no. 21, pp. 30257–30270, 2019.
- [15] D. Preuveneers, W. Joosen, J. B. Bernabe, and A. Skarmeta, "Distributed security framework for reliable threat intelligence sharing," *Security and Communication Networks*, vol. 2018, Article ID 8833765, 15 pages, 2018.
- [16] H. Hinne, "Rationality constraints in cyber defense: incident handling, attribution and cyber threat intelligence," *Computers & Security*, vol. 109, 2021.
- [17] P. Bernhard, H. Geoff, and F. Eibe, "Classifier chains for multi-label classification," *Machine Learning*, vol. 85, no. 3, 2011.
- [18] I. Yen, X. Huang, W. Dai, P. Ravikumar, I. Dhillon, and E. Xing, "Ppdsparse: a parallel primal-dual sparse method for extreme classification," in *Proceedings of the 2017 Conference on ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 545–553, ACM, Halifax NS, Canada, August 2017.
- [19] J. Yang, Y. Liu, and J. Luo, "Multi-label extreme classification based on subset topic model," *Computer Engineering and Design*, vol. 12, pp. 140–145, 2020.
- [20] H. Tan and Z. Liu, "Multi-label K nearest neighbor algorithm by exploiting label correlation," *Journal of Computer Applications*, vol. 35, no. 10, pp. 2761–2765, 2015.
- [21] Y. Prabhu and M. Varma, "Fastxml: a fast, accurate and stable tree-classifier for extreme multi-label learning," in *proceedings of the 2014 Conference on ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 263–272, August 2014.
- [22] C. Guibin, Y. Deheng, X. Zhenchang, C. Jieshan, and C. Erik, "Ensemble application of convolutional and recurrent neural networks for multi-label text categorization," in *Proceedings of the 2017 International Joint Conference on Neural Networks-Anchorage*, pp. 2377–2383, AK, USA, May 2017.
- [23] L. Xiao, X. Huang, B. Chen, and L. Jing, "Label-specific document representation for multi-label text classification," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pp. 466–475, Hong kong, China, January 2019.
- [24] J. Wehrmann, R. Cerri, and R. Barros, "Hierarchical multi-label classification networks," in *Proceedings of the 35th International Conference on Machine Learning ICML*, pp. 8321–8330, Stockholm SM, Sweden, Stockholm SWEDEN, February 2018.
- [25] S. Barnum, "Common attack pattern enumeration and classification (CAPEC) schema," *Department of Homeland Security*, 2008.
- [26] A. Cherepanov and R. Lipovsky, "Industroyer: biggest threat to industrial control systems since stuxnet," *WeLiveSecurity, ESET*, vol. 12, 2017.
- [27] A. Cherepanov, *Win32/Industroyer: A New Threat for Industrial Control Systems*, ESET, 2017, https://www.welivesecurity.com/wpcontent/uploads/2017/06/Win32_Industroyer.pdf.
- [28] M. Awad and R. Khanna, "Support vector machines for classification," in *Efficient Learning Machines*, pp. 39–66, 2015.
- [29] F. Pedregosa, G. Varoquaux, A. Gramfort, and V. Michel: API Reference, <https://scikit-learn.org/stable/modules/classes.html>.