A Regularized Cross-layer Ladder Network for Intrusion Detection in Industrial Internet-of-Things

***Abstract*—As part of BigData trends, the ubiquitous use of the Internet-of-Things (IoT) in the industrial environment has generated a significant amount of network traffic. In this type of IoT industrial network where there is a large equipment heterogeneity, security is a fundamental issue, thus it is very important to detect likely intrusion behaviors. Furthermore, since the proportion of labeled data records is small in IoT environment, it is challenging to detect various attacks and intrusions accurately. This investigation builds a semi-supervised ladder network model for intrusion detection in IIoT. This model considers the manifold distribution of high-dimensional data and incorporated a manifold regularization constraint in the decoder of the ladder network. Meanwhile, the feature propagation between layers is strengthened by adding more cross-layer connections in this model. On this basis, a random attention-based data fusion approach to generate global features for intrusion detection. The experiments on CIC-IDS2018 show that the proposed approach can recognize the intrusion with less false alarm rate, whilst model training is time-efficient.**

***Index Terms*—industrial internet-of-things, heterogeneity, ladder network, network intrusion detection, manifold regularization**

I. Introduction

With the advancement of industrial informatization and the internet-of-things(IoT) development, the external network can openly access the traditional enclosed industrial control system. Various heterogeneous devices are interconnected via the sensing techniques, forming an industrial IoT(IIoT)[1]. The gigantic volume of industrial data has to be transferred over the network through various applications[2]. The significant innovation in industry and business makes the industrial data face more new cybersecurity threats due to the security deficiencies of the massive sensing devices. IIoT is a large-scale network with three layers, including the application layer, the network layer, and the perception layer, as shown in Fig.1. It makes the sensing devices interconnected with the industrial control network, significantly improving manufacturing efficiency.

The traditional control system to be openly accessed through IIoT network, which poses severe threats to its security[3], [4]. In 2014, more than 30% intelligent electric meters of the top-three power supply providers in Spain exhibited serious security holes[5]. Attackers can use these deficiencies for electric charge fraud or closing power supply systems. To further illustrate the importance of privacy and security given the current IoT trends, in 2015, the BlackEnergy virus caused an extensive blackout in Ukraine [6]. Furthermore, the Mirai botnet infected more than 2 million intelligent devices for launching large-scale denial-of-service (DoS) attacks in 2016 [7]. While in 2018, more than 17 risk holes, including default password and bypass identity authentication, were detected from smart cities[8], and these holes can be utilized to control the alarm systems, tamper sensor data, and easily control the urban transport, causing severe results.

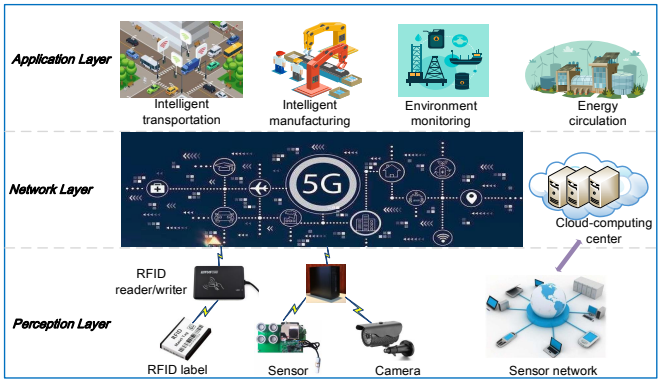


Fig. 1: The architecture of an IIoT

These intrusion events illustrate that the IIoT faces severe threats of complex and diverse malicious attacks. Although intrusion detection has been researched for many years, there are still many security issues to be addressed[9], [10], [11]. Previous researches provide some support for intrusion detection in IIoT. However, it has plenty of difficulties in dealing with the vast amounts of unstructured, noisy, high-dimensional, and unlabeled data [12].

Deep learning can provide a robust solution due to its powerful ability in high-efficient automatic feature extraction from a massive dataset[13]. Many deep learning-based intrusion detection techniques depend on supervised learning on a large amount of labeled training data which is costly and becomes infeasible for the massive unlabeled data generated by IIoT. Therefore, it is critical to use the vast amounts of unlabeled data and build an unsupervised or semi-supervised deep learning-based model for practical applications, such as intrusion detection.

Semi-supervised learning can use the small amounts of available labeled data and huge amounts of unlabeled data for model training, which achieves the equivalent performance to that of the supervised learning. Driven by the practical requirements, this work aims to utilize the high-dimensional and unlabeled data in IIoT and builds a semi-supervised deep learning model for intrusion detection. This research advances state of the art in IoT network intrusion detection mechanisms via the following contributions:

- A semi-supervised ladder network is built in the heterogeneous domains for training with large amounts of unlabeled data in IIoT. This model improves the ladder network by adding dense cross-layer connections and manifold regularization. Therefore, more preserved

features can be used in model training to avoid model degradation issues.

- An improved random attention mechanism estimates the importance score concerning the important feature information and makes it independent of the training data. On this basis, a data fusion technique is proposed to extract heterogeneous features, which are integrated with the features trained by the ladder network. It helps to reduce the dependency on the training data and accelerate the speed of model training.
- The excellent performance of the proposed technique in IIoT intrusion detection is demonstrated. It can satisfy the demands of network intrusion detection in heterogeneous industrial IoT environments.

The remaining paper is organized as follows. Section II analyzes the state-of-art in intrusion detection. Section III introduces the proposed model, and Section IV describes the intrusion detection algorithm. Section IV evaluates the performance of the proposed algorithm, and finally, concluding remarks and future directions are given in Section V.

II. Related Work

With the consideration of the proportion of the labeled data, the deep learning-based intrusion detection techniques can be classified into three types, supervised learning[14][15], unsupervised learning[16][17] and semi-supervised learning [18][19]. At present, research has focused on unsupervised and semi-supervised learning. The latter can fully utilize the small number of labels and achieve better performance. Notably, considering the gigantic volume of unlabeled network data in IIoT, it is exciting to use the gigantic volume of unlabeled data for training, which increases the generalization ability of the models to a new domain. In this section, previous researches in intrusion detection of IIoT by using semi-supervised learning models are investigated as follows.

The hackers may gain authority to attack the IIoT in a long interval, which is difficult to be detected by common machine learning schemes due to the requirements of expert knowledge. By considering this, Li et al.[20] proposed a bidirectional long and short-term memory network with a multi-feature layer. This model can learn the corresponding attack interval from historical data, which greatly enhanced the efficiency to detect attacks in different intervals. Besides, Li et al. [21] focused on semi-supervised learning and designed an intrusion detection technique by applying a disagreement-based semi-supervised learning algorithm for collaborative intrusion detection. Dutta et al. [22] utilized a two-step process for the detection of network anomalies. In the first stage, a Deep Sparse AutoEncoder (DSAE) is employed for data pre-processing, thus solving the feature engineering problem. In the second phase, a stacking ensemble learning approach is utilized for classification.

Several researches concentrate on the time series. Kong et al. [23] proposed an integrated deep generative model by combining the generative adversarial networks based on bi-directional LSTM and attention mechanism, which can capture the time series dependence. The reconstruction loss and generation loss test the input of sample training space and random latent space. Yin et al.[24] utilized the two-stage sliding window in data pre-processing to learn better representations of time series. It is helpful to extract high-level features in the integrated model. The spatial and temporal features are then extracted in CNN and recurrent autoencoder for the classification in fully connected networks. Abdel-Basset et al. [25] proposed a multi-scale residual temporal convolutional module by learning the spatiotemporal representations and introduced an improved attention mechanism for the extraction of the global feature. Finally, a semi-supervised model was implemented for intrusion detection in IoT. Cheng et al. [26] proposed a semi-supervised hierarchical stacking temporal convolutional network for intrusion detection in IoT. It can train unlabeled data based on a small number of labeled data. This algorithm's accuracy is directly related to the effectiveness threshold, and the detection efficiency could potentially be further optimized. Shailendra et al. [27] proposed a semi-supervised model by combining the extreme learning machine and fuzzy C means method for intrusion detection in IoT. It has solved the vulnerability issue of a centrally deployed intrusion detection system, but the model faces the overfitting issue. Furthermore, the model performance achieved with this technique depends on setting a good confidence value. For realizing a distributed intrusion detection framework, Chiu et al. [28] proposed an edge learning system based on semi-supervised learning and federated learning techniques. Besides, a federated swapping operation is proposed to replace the partial federated learning operation based on a few shared data during federated training.

As the unsupervised learning reserves the original information as more as possible and the supervised learning mainly concerns the information related to supervised task. In majority of above semi-supervised schemes, both stages of the unsupervised learning and supervised learning are separate while actually both requirements should be satisfied simultaneously in semi-supervised task.

As a semi-supervised model, ladder network integrates the supervised learning and unsupervised learning in a framework. It was firstly proposed by Valpola[29], which proved the effectiveness of using lateral shortcut connections to aid the deep unsupervised learning. The idea of a ladder network was further extended by Rasmus et al. [30] to support the supervised learning. Their researches added the classification and regression to the unsupervised reconstruction of inputs via a de-noising autoencoder. Finally, Pezeshki et al.[31] investigated the various components that affected the ladder network. They noticed that the lateral connections between encoder and decoder and the addition of noise at each layer of the network could significantly enhance the performance. Many researchers proposed new ladder network architectures for various applications, such as the laplacian ladder network[32].

In this work, we consider the gigantic volume of unlabeled IIoT data with high-dimensional features and attempt to use the ladder network framework in the proposed model. The deep learning models such as Autoencoders and restricted Boltzmann machines may ignore the manifold information of high-dimensional data and generate some unmeaning features[32]. These features are useless for model training in practical applications. Besides, the strengthened feature

propagation from preceding layers is positive to improve model performance[33]. For these considerations, we proposed a ladder network based deep learning model for intrusion detection. It is a semi-supervised learning model that considers the manifold distribution of high-dimensional IIoT data and maximizes the propagation of the detailed features between layers. Besides, a random attention-based data fusion approach is proposed to generate global features for intrusion detection.

III. Proposed Framework and Model

A. Motivation

As stated in Section II, the labels of network traffic in IIoT are limited. This study aims to increase the generalization ability of previous intrusion detection models of IIoT with unsupervised tasks and unlabeled data from several heterogeneous domains. The motivations aim at solving the unsupervised learning problems, which aid the primary intrusion recognition task. The available labeled data records should be fully used, which are expensive to annotate. Therefore, we establish a learning framework where the features of all heterogeneous domains are jointly learned due to the dependencies among multiple attributes of different domains. The ladder network is proposed to leverage unlabeled data effectively. Collectively, we create a semi-supervised intrusion detection model that effectively generalizes to new domains.

B. Regularized Cross-layer Ladder Network

The ladder network usually assumes that the data is distributed in a high-dimensional Euclidean space, which can be mapped to a low-dimensional manifold space. The case that the original data may distribute in a low-dimensional manifold space is not considered. In this case, the ladder network structure easily ignores the local feature of low-dimensional data, which will affect the accuracy of intrusion detection. This work aims to address this issue by proposing a regularized cross-layer ladder network to accurately classify network traffic from different domains and enhance the generalization ability, as described further.

The regularized cross-layer ladder network is established on the basic structure of the traditional ladder network, including a noisy encoder, a noisy decoder, and a clean encoder. This model introduces the concept of DenseNets proposed in [33] and manifold regularization[34] as depicted in Fig.2. In the regularized cross-layer ladder network, more lateral and vertical cross-layer connections are added to preserve more features and maximize the information transmission among layers. In addition, the manifold regularization constraints are included in each decoding layer to preserve the same low-dimensional structure with the input data in the decoding.

1) Encoder: There are two encoders, one for noisy inputs and another for cleaning inputs. The encoder consists of a multilayer perceptron. Gaussian noise with variance σ^2 is added to each encoder layer. Let the Gaussian noise parameter be denoted by ϵ. We have $\epsilon = N(0, \sigma^2)$. Therefore, the input data x of the noisy encoder becomes $\widetilde{x} = x + \epsilon$. After adding the noise, the hidden layer $\widetilde{h}^{(i)}$ can be expressed by Eq.(1).

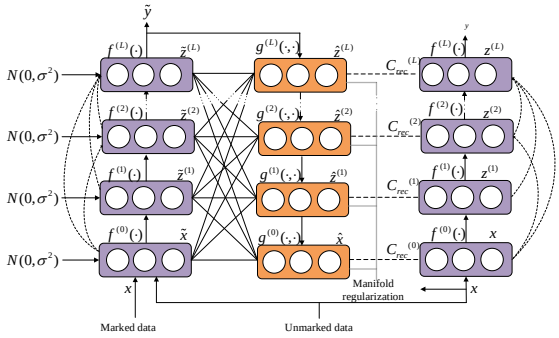


Fig. 2: Regularized cross-layer ladder network

$$\widetilde{h}^{(i)} = f(\mathcal{W}_e^{(i)}\widetilde{x}^{(i)} + \alpha^{(i)}), 1 \le i \le L. \quad (1)$$

f is the encoding mapping function. $\mathcal{W}_e^{(i)}$ and $\alpha^{(i)}$ are respectively the weight matrix and the bias value of the i-th encoding layer. For the first layer of the encoder, we have $\widetilde{h}^{(0)} = \widetilde{x}$.

This model utilizes the thought of cross-layer connection. The vertical connections are added among the encoding layers. Besides, each layer of the noisy encoder is connected to each decoder layer. For this reason, the input of each encoding layer comes from all of the layers ahead. Accordingly, the output of each encoding layer will be connected to the input of the layers behind it. After encoding, the data will be further normalized, and then Gaussian noise will be added as Eq.(2).

$$\widetilde{z}^{(i)} = \mathcal{G}(\mathcal{W}_e^{(i)}[\widetilde{h}^{(0)}, \widetilde{h}^{(1)}, \ldots, \widetilde{h}^{(i-1)}]) + \epsilon, 1 \le i \le L. \quad (2)$$

Here, $\mathcal{G}$ is the normalized function. The output of the hidden layer could be expressed as Eq.(3).

$$\widetilde{h}^{(i)} = f(\mu^{(i)}(\widetilde{z}^{(i)} + \eta^{(i)}), 1 \le i \le L. \quad (3)$$

μ and η are standardized coefficients of the i-th encoding layer. At the L-layer, the classification function can predict the class of network behavior and output the probability of the network behavior belonging to a specific class. The process of the clean encoder is similar to that of the noisy encoder.

2) Decoder: The decoder g can be utilized to reconfigure the encoding value and generate the original input. The expression of the decoder is as Eq.(4).

$$\hat{h}^{(i)} = g(\mathcal{W}_d^{(i)}\hat{h}^{(i)} + \beta^{(i)}), 1 \le i \le L. \quad (4)$$

$\mathcal{W}_d^{(i)}$ and $\beta^{(i)}$ are respectively the weight matrix and the bias value of the i-th decoding layer. The reconfigured $\hat{h}^{(i)}$ has the same shape to $\widetilde{x}$. Each encoding layer will be connected to all of the decoding layers. Thus, it maximally preserves and transmits the features between the encoder and the decoder layers. In this case, each decoding layer should consider the noise from all the encoding layers, and the reconfigured information can be expressed as follows:

$$\hat{z}^{(i)} = \begin{cases} g_k(\mathcal{W}_d^{(i)}\widetilde{\boldsymbol{z}}_{\boldsymbol{k}}, \mathcal{G}(\widetilde{y})), i = L, \\ g_k(\mathcal{W}_d^{(i)}\widetilde{\boldsymbol{z}}_{\boldsymbol{k}}, \mathcal{G}(W_d^{(i+1)}\hat{z}_k^{(i+1)}), 0 \le i \le L-1. \end{cases} \quad (5)$$

Herein, $\widetilde{\boldsymbol{z}}_{\boldsymbol{k}} = [\widetilde{z}_k^{(0)}, \widetilde{z}_k^{(1)}, \ldots, \widetilde{z}_k^{(L)}]$ is the combination of the noise from all encoding layers. $\hat{z}_k^{(i+1)}$ is the output of k-th neuron of $(i+1)$-th decoding layer. $\mathcal{W}_d^{(i)}$ is the weight matrix of the i-th decoding layer. Finally, we have $\hat{x} = \hat{z}^{(0)}$.

The network behavior of data in an industrial IoT environment is high-dimensional. However, the traditional ladder network usually ignores the low-dimensional manifold structure information of high-dimensional data, which will affect the accuracy of intrusion detection; thus, to enhance the discernment ability of feature extraction and generalization ability of the model, the manifold regularization constraints are generated and added into each layer.

The overall loss for the ladder network is given by

$$C_{ladder} = \lambda C_{sup} + C_{rec} + C_{reg}, \tag{6}$$

where C_{sup}, C_{rec} and C_{reg} denote the supervised loss, the reconfiguration error and the manifold regularization constraints respectively. $\lambda = \sqrt{\frac{M_m + M_u}{M_m}}$ is the dynamical weight coefficient of supervised cost function. M_m and M_u are respectively the number of labeled samples and unlabeled samples.

If there are n heterogeneous domains, the addition of the random attention makes the overall loss be given by

$$C_{ladder+att} = \lambda C_{sup} + C_{rec} + C_{reg} + C_{att}, \tag{7}$$

where $C_{att} = \theta_1 C_1 + \theta_2 C_2 + \cdots + \theta_n C_n$ represents learning for the features in n heterogeneous domains. The elements in $\theta_i, i \in [1, 2, \cdots, n]$ are hyper-parameters, where $\theta \in [0, 1]$ and $\theta_1 + \theta_2 + \cdots + \theta_n$=1.

3) Random Attention: The random attention mechanism can make the model concentrate on the important information. In a heterogeneous dataset, given the input $x \in M_{l \times d}$, a feature matrix satisfies $A \in M_{n \times n}$. Herein, l represents the length of the input sequence. d is the number of feature dimensionalities. n is the model dimension. First, the d-dimensional input will be mapped to a n-dimensional B via a parameterized matrix T, namely $B = MT$. Two matrixes $M_1, M_2 \in M_{n \times n}$ are randomly initialized. A matrix M is generated by calculating $M = M_1 \times M_2$, which can be used for calculating the attention score $s_i = \frac{e^{m_1}}{\sum_{j=1}^{n} e^{r_j}}$. Given a matrix M, the score matrix G will be generated via the Softmax function F. Finally, the attention matrix $A_{att} = Softmax(M)B$ is generated.

4) Feature Fusion: The feature fusion is to integrate the local feature and the global feature. In IIoT intrusion detection, a specific attack may have several features. It is critical to adjust the limitation of local feature and global feature fusion. The feature extraction can generate two feature matrixes, which can be integrated by setting a learning parameter.

$$H = [1 - \phi(\alpha)]A_{att} \oplus \phi(\alpha)A_{ln} \tag{8}$$

where $\oplus$ is the connection between two elements. α is a learning parameter which can be set as 0.5 initially. $\phi(\cdot)$ is a function to make the updated α fall into the range of [0,1].

$$\phi(x) = \begin{cases} 1, x > 1 \\ 0, x < 0 \\ x, others \end{cases} \tag{9}$$

IV. Algorithm Description

A semi-supervised intrusion detection algorithm is proposed based on the regularized cross-layer ladder network (RCLN). Since the original network dataset is high-dimensional and redundant, it should be firstly pre-processed. In the RCLN algorithm, the training dataset is then analyzed, and the nonlinear manifold regularization and random attention mechanism are utilized for feature extraction. The redundant features which affect the detection accuracy are then removed. The network intrusion detection issue is subsequently transformed into a training dataset combined with a regularized cross-layer ladder network. Finally, the behavior of network data is then analyzed for accurate recognition and classification of intrusion. The details of the algorithm are described next.

1) Data preprocessing: Each present feature's evaluation metric and data dimension in a network behavior dataset are different. For such data, the direct calculation may cause an inaccurate analysis result. The original data usually requires data preprocessing to eliminate the influence of data dimension and unit on classification results, such as data cleaning, normalization. Data cleaning is to enhance the quality of dataset. As for normalization, the data are mapped into the range between zero and one, and all-dimensional data are transformed into non-dimensional data. Following formulation expresses data normalization as:

$$x' = \frac{x - min(x)}{max(x) - min(x)}, \tag{10}$$

where x' corresponds to the normalized characteristic data, x is the original data sample, $min(x)$ represents the minimum value in the original characteristic data sample, and $max(x)$ is the maximum value in the original characteristic data sample.

2) Data encoding and classification: The normalized data samples are divided into labeled and unlabeled network data. The former can train the standard feedforward neural network, which is regarded as the encoder of the cross-layer ladder network, whilst the latter will be inputted into both the noisy and clean encoders for training.

At each layer, the input data should realize clean encoding and noisy encoding, as the clean encoding can be used to calculate the unsupervised cost. For noisy encoding, the random Gaussian noise is added to the data at each layer for better generalization ability of the training model. Thus, clean and noisy encoding is generated for the input data at each layer after several iterations.

3) Data decoding and de-noising: The construction of decoder aims to support the unsupervised learning of the unlabeled network data. Subsequently, the trained result of supervised learning can be analyzed. By using the designed de-noising function g , $\hat{h}^{(i)} = g(\widetilde{h}^{(k)}, \hat{h}^{(i+1)})$, the optimal estimation of clean data at the hidden layers can be generated.

4) Calculation of loss function: The total loss function can be used to measure the difference between the predicted result and the actual result. It is calculated by considering the supervised loss function, reconfiguration error, and regularization constraint in this section. The optimization goal is to minimize the total loss. After this step, it generates a function $P(y|x)$ to classify the network behavior of the sample data.

5) Error backpropagation and parameter update: Next, the partial derivative of the loss function is calculated and updated using the gradient descent method, and finally, the error is propagated back to optimize the model parameter.

The pseudo-code of the proposed intrusion detection algorithm is described as follows.

Algorithm RCLN Algorithm.

Input:
Dataset x_i with M_m labeled and M_u unlabeled;
Output:
classification result CR ;

$x \leftarrow x + N(0, \sigma^2)$;
$\widetilde{h}^{(0)} \leftarrow \widetilde{z}^{(0)} \leftarrow x$;
for $i = 1$ to L **do**
  $\widetilde{z}_p^{(i)} \leftarrow \mathcal{W}_e^{(i)}[h^{(0)}, h^{(1)}, \ldots, h^{(i-1)}]$;
  $\widetilde{z}^{(i)} \leftarrow \mathcal{G}(\widetilde{z}_p^{(i)}) + N(0, \sigma^2)$;
  $\widetilde{h}^{(i)} \leftarrow f(\mu^{(i)}(\widetilde{z}^{(i)}) + \eta^{(i)})$;
end for
$P(\widetilde{y}|x) \leftarrow \widehat{h}^{(L)}$;
for $i = L$ to 0 **do**
  if $i = L$ **then**
    $\hat{z}^{(i)} \leftarrow g_k(\mathcal{W}_d^{(i)}[\widetilde{z}_k^{(0)}, \widetilde{z}_k^{(1)}, \ldots, \widetilde{z}_k^{(L)}], \mathcal{G}(\widetilde{y}))$;
  else
    $\hat{z}^{(i)} \leftarrow g_K(\mathcal{W}_d^{(i)}[\widetilde{z}_k^{(0)}, \widetilde{z}_k^{(1)}, \ldots, \widetilde{z}_k^{(L)}], \mathcal{G}(\mathcal{W}_d^{(i+1)}\hat{z}_k^{(i+1)}))$;
  end if
end for
$h^{(0)} \leftarrow z^{(0)} \leftarrow x$;
for $i = 1$ to L **do**
  $z_p^{(i)} \leftarrow \mathcal{W}_e^{(i)}[h^{(0)}, h^{(1)}, \ldots, h^{(i-1)}]$;
  $z^{(i)} \leftarrow \mathcal{G}(z_p^{(i)})$;
  $h^{(i)} \leftarrow f(\mu^{(k)}(\widetilde{z}^{(i)}) + \eta^{(i)})$;
end for
$P(y|x) \leftarrow h(^{(L)_1}, h(^{(L)_2}, \cdots, h(^{(L)_n}$;
train with the total cost function $C_{ladder+att} = \sqrt{\frac{M_m+M_u}{M_m}}(C_{ne} + C_{ce}) + C_{rec} + C_{reg} + C_{att}$;
perform error backpropagation;
update the weight of the encoder with gradient descent;
generate CR with P;
return CR.

V. Performance Evaluation

A. Dataset Preparation

In this section, the experiments are conducted on the CIC-IDS2018 dataset[35]. It is a widely used dataset in the intrusion detection field. The dataset will be firstly pre-processed before training. In this work, we use 10% percentage of data records in CIC-IDS2018 for evaluation, which is divided into three groups, including 60% training data, 20% validation data, and 20% testing data. The model is evaluated on the pre-processed dataset for performance comparison.

B. Performance Criteria

Detection accuracy is the ratio of correctly recognized data and the total number of records in a dataset, denoted by $accuracy$, as it directly reflects the classification effect of the proposed model.

Precision measures the number of correct classifications of normal behavior and intrusion outliers penalized by the number of incorrect classifications, which is denoted by $precision$.

The Recall is the ratio of the number of correct classifications to the number of missed entries, as denoted by $recall$.

F1-core ($f1$) is to measure the weighted average of the precision and the recall. It falls in the range of [0,1].

The True Positive Rate (TPR) measures the proportion of intrusion outliers that are correctly identified.

The False Positive Rate (FPR) is the ratio of the number of mislabeled outliers to the number of normal behaviors. This metric can evaluate the reliability of the classification model. So, it is broadly utilized in performance evaluation.

Experiments mainly involve performance in terms of classification accuracy, detection reliability, and detection efficiency. The classification accuracy is measured by $accuracy$. The detection reliability is evaluated by TPR and FPR. The detection efficiency is measured by the detection time and overhead.

C. Comparison Baselines

The proposed model is compared to several ladder network models and two typical semi-supervised intrusion detection models for an overall evaluation of the performance.

The following ladder network models are chosen as the baselines. TypicalLN is the original ladder network[30]. LaplacianLN model is Laplacian Ladder Network[32] by adding Laplacian constraints. DenseLLN model improves the LLN by adding dense connections[33].

Two intrusion detection models are utilized for comparison. The semi-supervised SMLC model [36] realized a multi-clustering model for detection and prevention of intrusion, and the HS-TCN model[26] implemented a temporal convolutional network for intrusion detection.

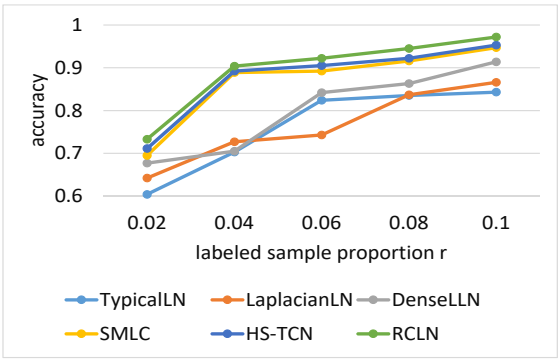
D. Evaluation Results

As a semi-supervised model, the growth of labeled sample proportion will affect intrusion detection results. By analyzing some previous researches, such as [36], it performs a good performance when the labeled sample proportion is 10%, which is equivalent to that when the labeled sample proportion is 90%. However, the performance with the labeled sample proportion less than 10% is not shown. As for a large-scale dataset, 10% labeled samples will be a big dataset. In this case, this section considers the small labeled proportion that is less than 10% for evaluation. The proportion is set as 2%, 4%, 6%, 8% and 10%. For better comparison, results for the labeled sample proportion 10% are available on Table.I. In this table, DenseLLN shows better accuracy than the other two ladder networks, and RCLN has improved the accuracy by 6.35%. Compared to the semi-supervised models SMLC and HS-TCN, the accuracy increase is also encouraging. For the precision, recall, and f1 score metrics, the proposed RCLN is also superior to the baseline models.

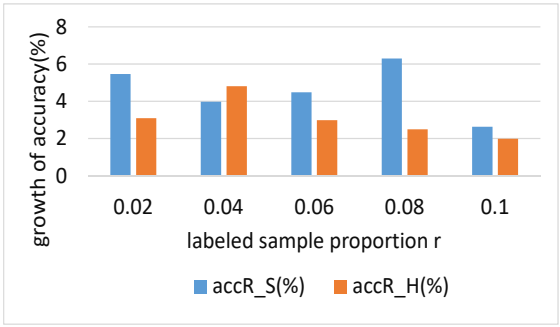
TABLE I: Evaluation results when the labeled sample proportion is 10%

Metric	*accuracy*	*precision*	*recall*	*f1*
TypicalLN	0.843	0.822	0.857	0.832
LaplacianLN	0.866	0.853	0.864	0.859
DenseLLN	0.914	0.908	0.923	0.911
SMLC	0.947	0.955	0.953	0.951
HS-TCN	0.953	0.946	0.961	0.949
RCLN	0.972	0.965	0.968	0.968

1) Accuracy Performance: Accuracy reflects whether the classification is accurate. Here, we consider the evaluation results with various labeled sample proportions and show the comparison curves with SMLC and HS-TCN in Fig.3.



(a)



(b)

Fig. 3: Performance curves with various labeled sample proportions. (a)accuracy. (b)accuracy growth.

In the Fig.3, the classification accuracy is measured. In Fig.3(a), the performance of RCLN outperforms the TypicalLN, LaplacianLN, and DenseLLN, which demonstrates the effective improvements. Both the SMLC and HS-TCN achieve high accuracy in their datasets. Compared to the SMLC and HS-TCN, the RCLN model can rapidly achieve high performance with a labeled sample proportion of more than 4%. The accuracy achieves 0.972 when the labeled sample proportion is 10%. In Fig.3(b), the histograms for the growth of accuracy by comparing SMLC and HS-TCN are described. By analysis, the proposed RCLN model improved an effective semi-supervised model and utilized the random attention mechanism in feature fusion.

The error reduction percentage is utilized to evaluate the significance of the proposed RCLN. Assume the accuracy values of SMLC, HS-TCN and RCLN are respectively acc_S, acc_H, and acc_R when the labeled sample proportion is 10%. the detection error $e = 1 - acc$. For example, $e_R = 1 - acc_R$. Therefore, the error reduction $er_{R\_S}(\%)$ and $er_{R\_H}(\%)$ respectively for SMLC and HS-TCN can be calculated as follows:

$$er_{R\_S}(\%) = \frac{e_S - e_R}{e_S} = \frac{acc_R - acc_S}{1 - acc_S} \tag{11}$$

$$er_{R\_H}(\%) = \frac{e_H - e_R}{e_H} = \frac{acc_R - acc_H}{1 - acc_H} \tag{12}$$

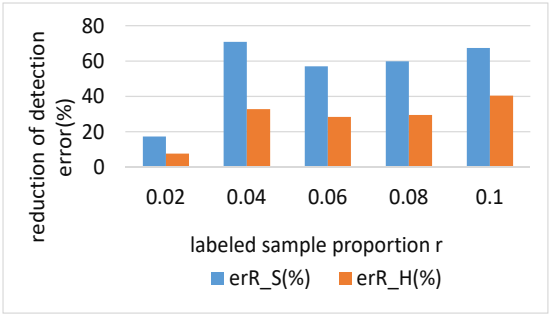
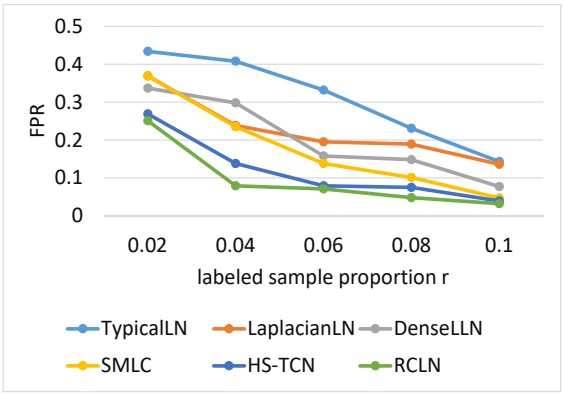


Fig. 4: Error reduction diagram with various labeled sample proportions.

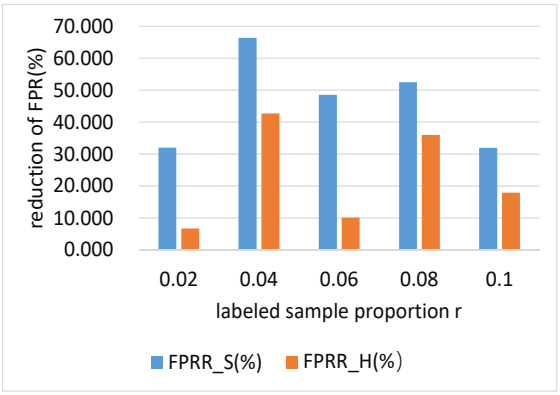
The $er_{R\_S}(\%)$ and $er_{R\_H}(\%)$ are also compared for various labeled sample proportions. The results are shown in Fig.4. In this Figure, RCLN achieves an error reduction of 70.84% based on SMLC using the labeled sample proportion value of 4%. Compared to HS-TCN, RCLN has reduced the detection error by 40.42% when the labeled sample proportion is 10%. Therefore, it demonstrates that the proposed model can significantly reduce the detection error.

2) Detection Reliability: A group of experiments is conducted by comparing the results under different labeled samples, whereas FPR and TPR can measure the detection reliability. By setting the labeled sample proportion r by 2%, 4%, 6%, 8%, and 10%, the RCLN is compared to the baselines. The results are depicted in Fig. 5 and Fig.6.

These models are evaluated for various values of r, directly reflecting the effects of the improved model. In Fig.5 and Fig.6, we evaluate the FPR and TPR with the Gaussian variance $\sigma^2 = 1$. The results show that the RCLN achieves the best performance of FPR and TPR among the ladder network models. It proves that the improvement of the training model is effective. From this figure, we can also see that the change rate of FPR and TPR is less with the growth of the labeled sample proportion. By analyzing the reason, the network model is semi-supervised. It can realize the model training and extract the feature of the dataset with a small labeled sample proportion r. Our model achieves the highest TPR. At this time, the training is equivalent to the supervised learning model. With the growth of r, the classification accuracy performance of the baseline models becomes higher. Since these models are semi-supervised, the r value will directly affect the detection accuracy. By analyzing the change of TPR, we can see that

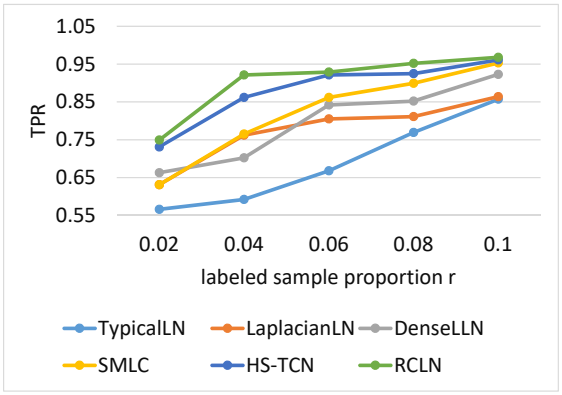


(a)

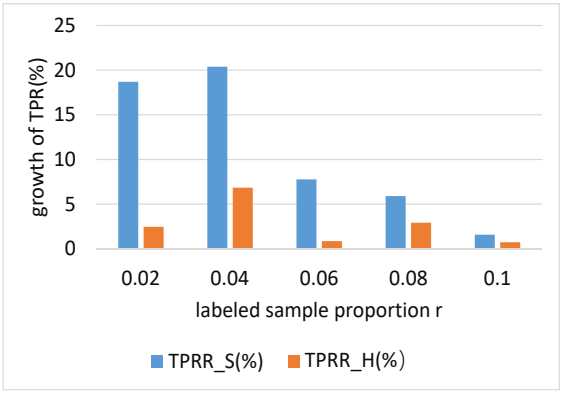


(b)

Fig. 5: FPR curves with various labeled sample proportions. (a)FPR. (b)FPR growth.



(a)



(b)

Fig. 6: TPR curves with various labeled sample proportions. (a)TPR. (b)TPR growth.

the TPR change rate of TypicalLN is the lowest, and that of the RCLN model is the largest. It demonstrates that the improvement of TypicalLN is effective.

3) Detection Efficiency: Due to the resource-constrained feature of the IIoT environment and the time-sensitivity of intrusion detection, the detection efficiency and computational cost are evaluated to attain the best performance. In this work, we have conducted the experiments on a server composed of one Intel Xeon E5-2678 CPU with 62G memory, and the detection time and computation overhead are evaluated in this section. The detection time of three semi-supervised models is evaluated, and the reduction of detection time is compared. The reduction results are shown in Fig.7. In this figure, $dt_{R\_S}(\%)$ and $dt_{R\_H}(\%)$ represent the reduction of detection time by comparing RCLN to SMLC and HS-TCN respectively. With the growth of the labeled sample proportion, the RCLN has a reduction exceeding 10% at least. The results show the reduction of $dt_{R\_S}(\%)$ achieves 21.86% when r=0.1 and $dt_{R\_H}(\%)$ achieves 14.35% when r=0.08. It demonstrates that the RCLN has improved its efficiency by 21.86% and 14.35% compared to SMLC and HS-TCN.

Besides, we also evaluate the overhead in model training, and the results are shown in Fig.8. This figure shows the overhead growth within 30 seconds. The average growth of CPU overhead and memory overhead is respectively 16.92% and 13.58%. In the beginning, the CPU overhead and memory overhead grow rapidly and gradually become stable.

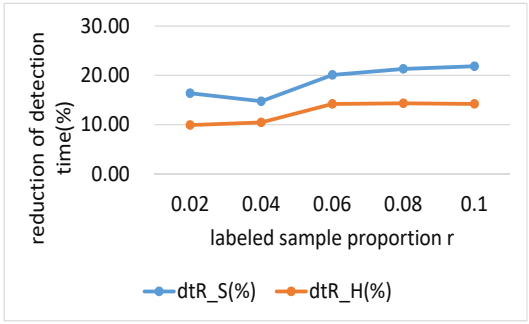


Fig. 7: Error reduction curves with various labeled sample proportions.

Conclusions and Future Work

This work considered the features of heterogeneity and variety in industrial IoT network data and then proposed a regularized cross-layer ladder network for intrusion detection. It is an improved semi-supervised intrusion detection model suitable for dealing with massive unlabeled data. Besides, we proposed a random attention-based data fusion approach to generate the global features on the heterogeneous dataset, as more features are used in model training. The regularization constraints are added to the decoder in the ladder network to enhance the model's generalization ability. The proposed

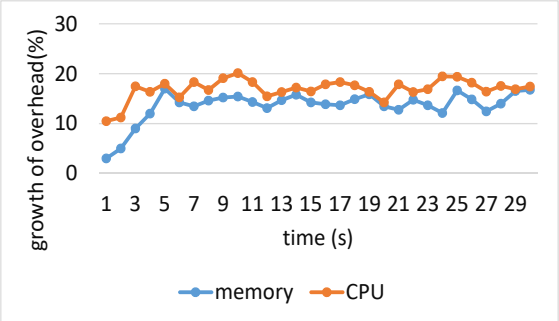


Fig. 8: Growth of CPU and memory overhead.

model is evaluated on the CIC-IDS2018 dataset. The results show it can recognize the intrusion with less false alarm rate, and the model training is time-efficient. As network traffic in IIoT may be collected from different heterogeneous domains. The intrusion detection model should be multi-task learning on the traffic collected from several domains rather than single-task learning on the traffic from a specific domain. So, the next work will consider the intrusion detection based on multi-task learning and research the multi-modal data fusion mechanism.

REFERENCES

[1] Y. Zhang *et al.*, "Survey of internet of things security," *Journal of Computer Research and Development*, vol. 54, no. 10, p. 2130, 2017.

[2] A. K. Mishra, D. Puthal, and A. K. Tripathy, "Graphcrypto: Next generation data security approach towards sustainable smart city building," *Sustainable Cities and Society*, vol. 72, no. 12, p. 103056, 2021.

[3] M. Z. A. Bhuiyan and J. Wu, "Collusion attack detection in networked systems," in *TrustCom/BigDataSE 2018*, 2018, pp. 286–293.

[4] A. K. Mishra, A. K. Tripathy, D. Puthal, and L. T. Yang, "Analytical model for sybil attack phases in internet of things," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 379–387, 2019.

[5] T. Li *et al.*, "Attacks and cyber security defense in cyber-physical power system," *Automation of Electric Power Systems*, no. 22, pp. 162–167, 2017.

[6] Y. Wang *et al.*, "Analysis and defense of the blackenergy malware in the ukrainian electric power system," *Chinese Journal of Network and Information Security*, vol. 3, no. 1, pp. 46–53, 2017.

[7] C. Seaman. (2020) Threat advisory: Mirai botnet. [Online]. Available: https://www.akamai.com/cn/zh/multimdia/documents/state-of-the-internet/akamai-mirai-botnet-threat-advisory.pdf.

[8] E. Ismagilova, L. Hughes, N. P. Rana, and Y. K. Dwivedi, "Security, privacy and risks within smart cities: Literature review and development of a smart city interaction framework," *Information Systems Frontiers*, no. 2, 2020.

[9] M. Abdel-Basset *et al.*, "Semi-supervised spatio-temporal deep learning for intrusions detection in iot networks," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2021.

[10] J. Zhang, M. Z. A. Bhuiyan, X. Yang, T. Wang, X. Xu, T. Hayajneh, and F. Khan, "Anticoncealer: Reliable detection of adversary concealed behaviors in edgeai assisted iot," *IEEE Internet of Things Journal*, 2022.

[11] S. M. Kasongo, "An advanced intrusion detection system for iiot based on ga and tree based algorithms," *IEEE Access*, vol. 9, pp. 113 199 – 113 212, 2021.

[12] W. Liang *et al.*, "Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2021.

[13] W. Liang, S. Xie, J. Cai, J. Xu, and M. Qiu, "Deep neural network security collaborative filtering scheme for service recommendation in intelligent cyber-physical systems," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2021.

[14] L. Lu, X. Zhu, X. Zhang, J. Liu, M. Z. A. Bhuiyan, and G. Cui, "One intrusion detection method based on uniformed conditional dynamic mutual information," in *TrustCom/BigDataSE 2018*, 2018.

[15] J. B. Awotunde, C. Chakraborty, and A. E. Adeniyi, "Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection," *Wireless Communications and Mobile Computing*, pp. 1–17, 2021.

[16] W. Liang *et al.*, "An industrial network intrusion detection algorithm based on multifeature data clustering optimization model," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2063–2071, 2020.

[17] X. Deng *et al.*, "An intelligent outlier detection method with one class support tucker machine and genetic algorithm toward big sensor data in internet of things," *IEEE Transactions on Industrial Electronics*, 2018.

[18] R. A. R. Ashfaq *et al.*, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484 – 497, 2017.

[19] N. Ravi *et al.*, "Semi-supervised learning based security to detect and mitigate intrusions in iot network," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2020.

[20] X. Li *et al.*, "Detection of low-frequency and multi-stage attacks in industrial internet of things," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8820–8831, 2020.

[21] W. Li *et al.*, "Enhancing collaborative intrusion detection via disagreement-based semi-supervised learning in iot environments," 2020.

[22] V. Dutta *et al.*, "A deep learning ensemble for network anomaly and cyber-attack detection," *Sensors*, vol. 20, no. 16, p. 4583, 2020.

[23] F. Kong *et al.*, "Integrated generative model for industrial anomaly detection via bi-directional lstm and attention mechanism," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2021.

[24] C. Yin *et al.*, "Anomaly detection based on convolutional recurrent autoencoder for iot time series," *IEEE Transactions on Systems, Man, and Cybernetics Systems*, vol. 52, no. 1, pp. 112–122, 2022.

[25] M. Abdel-Basset *et al.*, "Semi-supervised spatio-temporal deep learning for intrusions detection in iot networks," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2021.

[26] Y. Cheng *et al.*, "Leveraging semi-supervised hierarchical stacking temporal convolutional network for anomaly detection in iot communication," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2020.

[27] R. Shailendra *et al.*, "Semi-supervised learning based distributed attack detection framework for iot," *Applied Soft Computing*, vol. 72, pp. S1 568 494 618 303 508–, 2018.

[28] T. C. Chiu *et al.*, "Semi-supervised distributed learning with non-iid data for aiot service platform," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 1–1, 2020.

[29] H. Valpola, "From neural pca to deep unsupervised learning," *Eprint Arxiv*, 2015.

[30] A. Rasmus *et al.*, "Semi-supervised learning with ladder networks," *Advances in Neural Information Processing Systems*, 2015.

[31] M. Pezeshki *et al.*, "Deconstructing the ladder network architecture," 2015.

[32] C. Hu *et al.*, "Laplacian ladder networks," *Journal of Software*, vol. Vol.31Issue, no. 5, pp. 1525–1535, 2020.

[33] G. Huang *et al.*, "Densely connected convolutional networks," *IEEE Computer Society*, 2016.

[34] M. T. Kejani *et al.*, "Graph convolution networks with manifold regularization for semi-supervised learning," *Neural Networks*, vol. 127, no. 1, 2020.

[35] (2020) Ids2018. [Online]. Available: IDS2018https://www.unb.ca/cic/datasets/ids-2018.html

[36] O. Y. Al-Jarrah *et al.*, "Semi-supervised multi-layered clustering model for intrusion detection," *Digital Communications and Networks*, 2018.