Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.039733
Article

MEM-TET: Improved Triplet Network for Intrusion Detection System

Weifei Wang[1], Jinguo Li[1,*], Na Zhao[2] and Min Liu[1]

[1]College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, 201306, China
[2]National Pilot School of Software, Yunnan University, Kunming, 650504, China
*Corresponding Author: Jinguo Li. Email: lijg@shiep.edu.cn
Received: 13 February 2023; Accepted: 17 April 2023; Published: 09 June 2023

Abstract: With the advancement of network communication technology, network traffic shows explosive growth. Consequently, network attacks occur frequently. Network intrusion detection systems are still the primary means of detecting attacks. However, two challenges continue to stymie the development of a viable network intrusion detection system: imbalanced training data and new undiscovered attacks. Therefore, this study proposes a unique deep learning-based intrusion detection method. We use two independent in-memory autoencoders trained on regular network traffic and attacks to capture the dynamic relationship between traffic features in the presence of unbalanced training data. Then the original data is fed into the triplet network by forming a triplet with the data reconstructed from the two encoders to train. Finally, the distance relationship between the triples determines whether the traffic is an attack. In addition, to improve the accuracy of detecting unknown attacks, this research proposes an improved triplet loss function that is used to pull the distances of the same class closer while pushing the distances belonging to different classes farther in the learned feature space. The proposed approach's effectiveness, stability, and significance are evaluated against advanced models on the Android Adware and General Malware Dataset (AAGM17), Knowledge Discovery and Data Mining Cup 1999 (KDDCUP99), Canadian Institute for Cybersecurity Group's Intrusion Detection Evaluation Dataset (CICIDS2017), UNSW-NB15, Network Security Lab-Knowledge Discovery and Data Mining (NSL-KDD) datasets. The achieved results confirmed the superiority of the proposed method for the task of network intrusion detection.

Keywords: Intrusion detection; memory-augmented autoencoder; deep metric learning; imbalance data

1 Introduction

As a result of the fast advancement of technology and the broad adoption of the Internet, the network now permeates every aspect of human civilization, and network security is receiving increasing

attention. Cisco's Visual Networking Index (VNI) [1] predicts that by 2022, Internet-connected devices will reach 28.5 billion, dramatically increasing the likelihood of network attacks.

Intrusion Detection Systems (IDS) are one of the most effective approaches for network devices to prevent network intrusion. IDS's primary purpose is to monitor the traffic of network devices and identify anomalous or suspicious data transmission patterns or user actions. Once this behavior is detected during a network or system scan, the IDS can automatically generate protection alarms.

Deep Neural Networks (DNNs) have the potential to extract more potent characteristics from massive, high-dimensional data automatically. Consequently, the development of IDS based on deep learning has been a significant new research focus. Feature extraction utilizes several DNN architectures, such as Recurrent Neural Networks (RNN) and stacked autoencoders. In addition, the deep feed-forward neural network is used to categorize incursion types [2].

Nonetheless, the present-day DNN-based processes have the following limitations: First, unbalanced training is regularly current in the information utilized for intrusion detection, which may result in an excessive false positives (FP) detection charge for a few assaults. Even though several oversampling or records enhancement techniques [3,4] have been developed to stabilize a range of data lessons, they regularly replica or synthesize the found data, which no longer makes the variety of the following data wider. Second, due to the improvement of intrusion techniques, surprising sorts of intrusions may additionally show up on community gadgets in real-world community environments, making contemporary deep learning-based IDS fashions unable to become aware of them.

To solve the IDS data imbalance and increase the ability to detect unknown attacks. This paper proposed innovative memory-augmented autoencoders and triplet networks (MEM-TET) model. Specifically, for the imbalance of the training data, this essay uses memory-augmented autoencoders (MemAE) to capture the most profound features of the data. The strategy studies two independent memory-augmented autoencoders on the regular flow and the attack relatively. Instead of typical sampling, these memory-augmented autoencoders produce positive and negative information for the triplet structure.

Especially each triad construction considers positive and negative pseudo-samples that are exclusive restructuring by recovering the two memory-augmented autoencoders for the anchors. Thus, these pseudo-samples derived from the memory-augmented autoencoders replace the actual training samples randomly selected by the traditional Triplet network from the regular and aggressive parts of the training data. In addition, this essay proposes an effective loss term called re-soft-margin (re-soft). The improved triplet network measures attack based on the Euclidean distance, significantly improving the detection rate of unknown attacks.

In summary, our contribution to this work is as follows:

(1) This paper proposes an effective intrusion detection methodology called MEM-TET. We employ two independent memory-augmented autoencoders to construct the triplet. The memory block can better preserve the prototype features of the data, so the autoencoder can reconstruct the data similar to the training data well and increase the reconstruction error of different data classes.
(2) This paper proposes a new triplet loss that is used to pull the distances of the same class closer while pushing the distances belonging to different classes farther in the learned feature space. It allows better separation of normal and attack traffic.
(3) The higher detection performance of the claimed MEM-TET technique is demonstrated by comparing it with numerous state-of-the-art methodologies using several benchmark datasets.

This article's remaining sections are structured as follows. In Section 2, we examine efforts on intrusion detection that are comparable. Section 3 discusses the proposed MEM-TET technique with complete derivations. Experimental outcomes and further analyses are reported in Sections 4, 5, 6, and 7, 8. Conclusions and future strategies are presented in Section 9.

2 Literature Review

Anderson [5] originally suggested the idea of intrusion detection systems in 1980 to spot aberrant network activities. This section contains work on studying the class imbalance problem of intrusion detection methods, including deep learning methods and Deep metric learning.

2.1 Deep Learning (DL)

Recent research on network intrusion detection has examined numerous methods for addressing the issue of data imbalance.

Nevertheless, Zhang et al. [6] presented a flow-based IDS that used Gaussian Mixture Model (GMM) and Synthetic Minority Oversampling Technique (SMOTE) to handle network data class imbalance. Using a one-dimensional Convolutional Neural Network (CNN), the authors evaluated their model on UNSW-NB15 and CICIDS2017. Xu et al. [7] introduced the LCVAE intrusion detection model, which inherits the capacity of the CVAE. Experiments indicate that utilizing the log-cosh constraint to balance the creation and reconstruction techniques is more successful for generating different intrusion data for unbalanced classes. In [8], a novel approach for sampling an unbalanced dataset was developed. In their study, the authors optimized hyperparameter tuning by combining the SMOTE algorithm with the grid search approach.

Furthermore, the authors of [9] address the unbalanced class issue by adding an unbalanced data filter and neural layers to the conventional Generative Adversarial Network (GAN), producing additional representative samples for the minority classes.

However, these authors disregard that the loss of information is an inherent consequence of down-sampling, whereas the generation of new samples may result in overfitting, noise, or class overlap.

2.2 Deep Metric Learning (DML)

Deep Metric learning (DML) is a paradigm for machines gaining knowledge that depends on distance measures. It seeks to quantify pattern similarity employing reducing the distance between comparable samples and growing the distance between specific examples. While DL's overall performance degrades dramatically when studying unbalanced data, DML methods can also be well-suited to cope with the classification imbalance trouble [10].

DML consists of the main Siamese networks and Triplet networks. DML has started to be utilized in cybersecurity. The Siamese community is a similarity measure approach that goes from the information to analyze a similarity measure. Specifically, in [11,12], using reading Siamese networks, type imbalance in networked intrusion detection structures was once addressed. The effects indicated that the multi-classification overall performance underneath the imbalance was once improved. Li et al. [13] cautioned against a novel methodology for discovering unknown traffic. In particular, the mannequin accepts the cautiously chosen packet traits as input, adopts the Siamese community architecture, and directs the coaching procedure via contrastive loss. Zhou et al. [14] developed a few-shot getting-to-know mannequin using a Siamese CNN shape to minimize overfitting and enhance industrial Cyber Physical System (CPS) anomaly detection.

The triplet impact is additionally beginning to be explored in cyber security. Zhou et al. [15] proposed a particular strategy based on the Triplet community that was once introduced for detecting anomalous Controller Area Network (CAN) bus data. The experimental findings disclose that the proposed device can reply in real-time to anomalies and assaults on the CAN bus, substantially improving the detection ratio. In [16], the community facts are modeled as a format shape to effectively mine the functional elements between information samples. The triplet community shape is used to realize anomalies by evaluating the distance similarity.

However, only some of the above studies concentrated on the sampling strategy of the Triplet network. They used the traditional random sampling strategy to construct the triplet, which led to poor model convergence and thus would lead to the problem of low overall detection efficiency of the model.

3 Methodologies

This section, we introduce MEM-TET, a DML-based network intrusion detection approach that combines Memory-augmented deep Autoencoder with Triplet networks in a unique way. The objective is to discover strong intrusion detection models to identify novel indicators of hostile behavior in network data. This essay combines MemAE to improve the sampling strategy of the Triplet network.

3.1 Triplet Networks

Triplet networks [17] are one of the standard DML techniques. Specifically, a ternary network uses a triad of samples. As shown in Fig. 1, each triad usually consists of training samples selected as anchors, training samples labeled with the same category as the anchors (positive samples), and training samples labeled with the opposite category (negative samples). A ternary network takes the ternary as input and learns an embedding space where the distance between samples labeled with opposite categories is more remarkable than between samples labeled with identical categories. However, existing methods based on ternary networks commonly need better convergence. The random sample selection of the triadic group construction in the training set mainly causes the convergence problem. Traditional ternary networks randomly select positive and negative samples to construct ternary groups, which consumes much time.

3.2 Memory-Augmented Autoencoder (MemAE)

The memory-augmented autoencoder model [18] consists of three components: encoder, memory-augmented module, and decoder. In contrast to the conventional self-encoder model, this model adds the memory-enhanced module after the encoder to learn and record a limited number of prototype patterns of the input data. The stored memory term aggregation results are used for decoding instead of encoding results. The training data can be reconstructed well. In this study, we refer to the encoder and memory-augmented module in MemAE as encoders for convenience. Thus, we used $g \cdot f$ to represent the MemAE model. The f stand for encoder and g stands for decoder.

3.3 The Proposed Method (MEM-TET)

Figs. 2 and 3 show the training and testing phases of MEM-TET, respectively. These Memory-augmented autoencoders are trained independently on routine and attack flows throughout the training phase. Injecting them into a Triplet network enables the construction of robust triplets and the discovery of a new embedding input space that separates regular and attack flows more efficiently. MEM-TET employs the trained MemAE to effectively predict the class of new network streams

following the learned embedding. Section 3.3 thoroughly discusses the training and predictive phases, respectively.

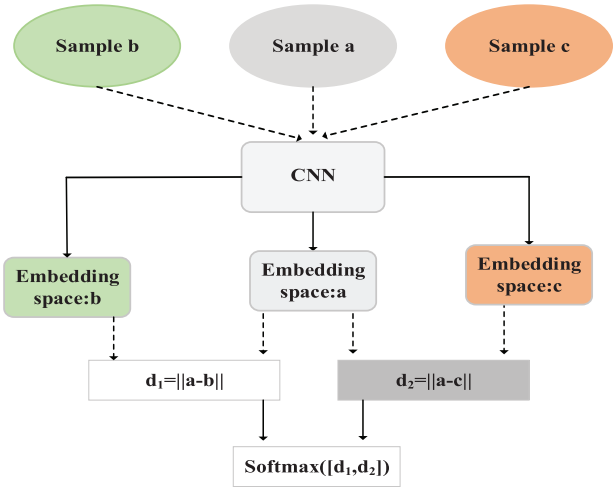

Figure 1: Traditional triplet networks. (1) Need to construct triple (***a***, ***b***, ***c***) as input. (2) Samples ***a*** and ***b*** belong to the same category and samples ***a*** and ***c*** belong to different categories (3) The Euclidean distance between sample ***a*** and sample ***b*** is d_1, and the Euclidean distance between sample ***a*** and sample ***c*** is d_2 (4) The training objective of the Triplet network is to minimize d_1 and maximize d_2

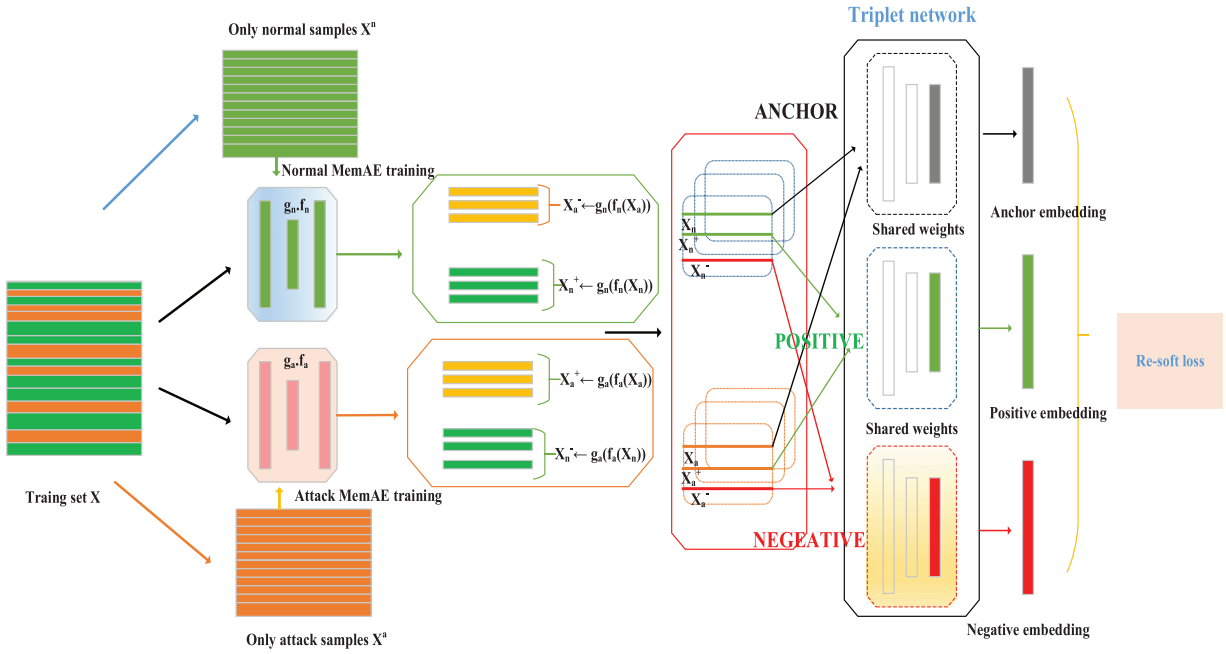

Figure 2: Training phase of MEM-TET model. (1) The data input **x** is divided into normal data $\mathbf{x}^n$ and attack data $\mathbf{x}^a$. (2) The normal MemAE $g_n \cdot f_n$ is trained on. (3) The attack MemAE $g_a \cdot f_a$ is trained on $\mathbf{x}^a$. (4) The two MemAE-reconstructed data train triplet network by forming triples with the original data

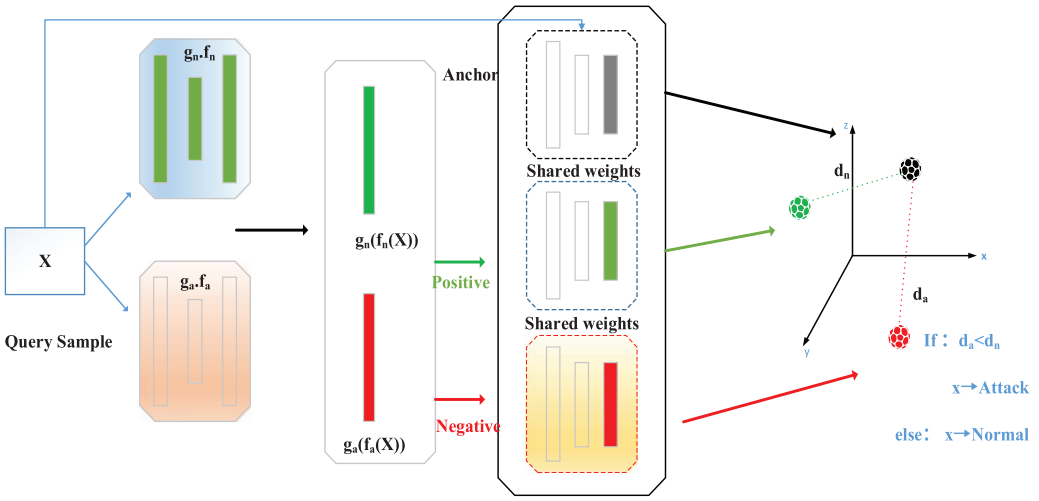

Figure 3: The MEM-TET stage of prediction. MemAE $g_n \cdot f_n$ and $g_a \cdot f_a$ are used to rebuild the query sample **x**. Calculate the distances between **x** and $g_n(f_n(\mathbf{x}))$ and between **x** and $g_a(f_a(\mathbf{x}))$, respectively. Through this, the class of **x** may be predicted

3.3.1 Training Model

As shown in Fig. 2, we explain the training phase of MEM-TET.

Specifically, the training phase of MEM-TET includes two stages:

1. This essay train two separate MemAEs to provide a novel description of the training data that can distinguish between regular and aggressive data.
2. This essay creates the unique positive and negative counterparts of each anchor sample using MemAE and then utilizes these to construct the triplet representation of the training data. The model is then trained using the triples constructed in the previous phase.

3.3.2 Triplet Construction

Triplet Network need to construct triplet comprises ($\mathbf{x}$, $\mathbf{x}^+$, $\mathbf{x}^-$) (1) **x**—**x** as anchor point. (2) $\mathbf{x}^+$— belongs to the same category as **x**. (3) $\mathbf{x}^-$— belongs to the opposite category to **x**.

This study's unique idea is that if **x** is standard data, the trained normal MemAE $g_n \cdot f_n$ can be well reconstructed from the usual traffic data by memory blocks that can retain the prototypical features of the training data. The weight information retained by memory blocks for abnormal data cannot be well reconstructed and can increase the error with the original data. Therefore, regular traffic can be considered positive samples from $g_n \cdot f_n$, while abnormal data can be considered harmful. By analogy, the trained abnormal MemAE $g_a \cdot f_a$ can be well reconstructed from the abnormal data, and the error with the original data can be increased for the standard data. From the perspective of $g_a \cdot f_a$, regular traffic is a negative sample, while attack data is a positive sample. By this idea, we use MemAE to construct positive and negative samples for each anchor point, thus forming a triad.

3.3.3 Triplet Loss

Finally, for a sample **x**, we construct a triplet by memory encoder, which is used for training the triplet network.

Additionally, as shown in Eq. (1) we suggest a novel loss function called re-soft margin which, different from [19], not only considers inter-class distance but also imposes a penalty on intra-class distance. This makes the intra-class distances more aggregated. Where $\|\mathbf{a}-\mathbf{b}\|$ stands for the Euclidean distance calculated among vectors **a** and **b**. Where β represents the hyperparameter, and in this study the value of β maked 0.534. τ_2 represents the new inter-class constraint we impose on ($\mathbf{x}$, $\mathbf{x}^+$).

$$L_{re-soft} = \sum_{\mathbf{x}\in\mathbf{X}} ln\left(1+\exp\left(\| \phi(\mathbf{x}) - \phi\left(\mathbf{x}^+\right) \|^2 - \| \phi(\mathbf{x}) - \phi\left(\mathbf{x}^-\right) \|^2 + \beta max\left\{\| \phi(\mathbf{x}) - \phi\left(\mathbf{x}^+\right) \|^2 - \tau_2, 0.0\right\}\right)\right) \quad (1)$$

3.3.4 Predictive Phase

Fig. 2 illustrates the prediction phase for test data **x**. Initially, average MemAE $g_n \cdot f_n$ and abnormal MemAE $g_a \cdot f_a$ are employed to reconstruct the data as $g_n(f_n(\mathbf{x}))$ and $g_a(f_a(\mathbf{x}))$, respectively. Subsequently, the trained Triplet network ϕ calculates the Euclidean distance between **x** and the reconstructed data. This is shown in the following equation:

$$d_a^{\phi}(\mathbf{x}) = \| \phi(\mathbf{x}) - \phi(g_a(f_a(\mathbf{x}))) \|^2 \quad (2)$$

$$d_n^{\phi}(\mathbf{x}) = \| \phi(\mathbf{x}) - \phi(g_n(f_n(\mathbf{x}))) \|^2 \quad (3)$$

Lastly, if $d_n^{\phi}(\mathbf{x}) > d_a^{\phi}(\mathbf{x})$, then **x** is classified as abnormal data. Otherwise, **x** is considered as a trustworthy data stream.

4 Experiments

4.1 Experimental Details

The proposed MEM-TET method is implemented in the Tensorflow-GPU 2.3 deep learning toolkit, and the experiment is conducted on a desktop machine with Intel i7-6700 K CPU with 24-GB RAM and an NVIDIA 2080Ti GPU card. We randomly use 20% of the data for each training dataset as the validation set and select the best model using automatic parameter optimization.

Each MemAE possesses 5 FC layers of 32–16–8–16–32 neurons. Each buried layer's activation function is the conventional rectified linear unit, whereas the final layer uses Linear. N = 100 memory chunks. A triplet is three feedforward networks with standard weights. Each base network has three intermediate layers, a 512-neuron embedding layer, and two dropout layers.

4.2 Dataset Description

This article examines the CIC-AAGM2017, CICIDS2017, KDDCUP99, and UNSW-NB15 benchmark intrusion detection datasets.

AAGM17 datasets: was collected by the Canadian Institute for Cybersecurity in 2017. In this investigation, we employ the subset built from [19].

CICIDS2017 is the largest dataset of its kind accessible online. In our experimental investigation, we constructed one training set including 100,0000 samples and one testing set containing 900,000 samples. The stratified random sampling method is used to randomly choose training and testing samples from the whole 5-day log to select 80% of normal flows and 20% of attacks, as in the original log.

KDDCup'99 is the most well-known and commonly used dataset for experimentation on anomaly detection in computer networks. In this study, 10% KDDCUP99Train is used for the learning stage, while the whole testing set, designated KDDCUP99Test, is utilized for the assessment stage.

The UNSW-NB15 dataset, in contrast to the typical dataset, this one incorporates various recent synthesis attacks, such as worms, fuzzes, generics, and reconnaissance. The experiment used the UNSW-NB15_training-set (containing 82332 instances) as the training set and the UNSW-NB15_testing-set as the training set (containing 175341 instances). Unlike previous studies, the test set contains multiple unknown attacks.

NSL-KDD is a network intrusion detection dataset based on the KDD-99 dataset, mainly used to evaluate the performance and robustness of intrusion detection algorithms. It is an improved version of the KDD-99 dataset. Compared with the KDD-99 dataset, the NSL-KDD dataset contains more types of network attacks and more normal network traffic, which aligns more with the actual network environment.

Table 1 provides an overview of the properties of the data presented before. We recognize that the traffic distribution is unbalanced in all datasets. The number of regular network traffic is considerably more than the number of assaults in both AAGM17 and CICIDS2017. In KDDCUP99 and UNSW-NB15, however, the number of assaults exceeds the number of regular flows. Scaling the numeric input features using the Min-Max scale is a component of the preprocessing procedure. This procedure is done on features with comparable value ranges. The data preprocessing method used in the experiments converts character features into digital features by unique thermal coding. Accuracy and F1-score (F1) were the primary assessment measures in this study.

Table 1: Dataset description

	Training set			Testing set		
	Total	Normal	Attack	Total	Normal	Attack
AAGM17	100000	80000 (80%)	20000 (20%)	100000	80000 (80%)	20000 (20%)
CICIDS2017	1000000	800000 (80%)	200000 (20%)	900000	720000 (80%)	180000 (20%)
KDDCUP99	494021	97278 (20%)	3967423 (80%)	311029	60593 (19.5%)	250436 (80.5%)
UNSW-NB15	82332	37000 (45%)	45332 (55%)	175341	56000 (32%)	119341 (68%)

5 Ablation Study

To prove the performance as well as the stability of the proposed re-soft loss in this paper, some additional experiments were considered. In this section, we analyze the performance of classical Triplet loss, Soft-margin loss, and the proposed re-soft loss, on the benchmark dataset.

According to [20], the initial triplet loss requires a set of ($\mathbf{x}$, $\mathbf{x}^-$) distances to be one predetermined magnitude larger than a set of ($\mathbf{x}$,$\mathbf{x}^+$) distances, and this requirement is enforced using the following equation:

$$\sum_{\mathbf{x}\in\mathbf{X}} \max\left(\| \phi(\mathbf{x}) - \phi\left(\mathbf{x}^+\right) \|^2 - \| \phi(\mathbf{x}) - \phi\left(\mathbf{x}^-\right) \|^2 + \alpha, 0\right) \tag{4}$$

As shown in Eq. (5) Soft-margin triplet loss was originally presented by [21] also applied in the latest research. Since this loss function does not specify how close the pair ($\mathbf{x}$,$\mathbf{x}^+$) should be, examples of belonging to the identical class may constitute a larger cluster with a relatively large internal meaning.

Learning the class distances in the feature space. Obviously, this is not the expected result and will inevitably harm the overall performance.

$$\sum_{\mathbf{x}\in\mathbf{X}} \ln\left(1+\exp\left(\| \phi(\mathbf{x})-\phi\left(\mathbf{x}^{+}\right)\|^{2}-\| \phi(\mathbf{x})-\phi\left(\mathbf{x}^{-}\right)\|^{2}\right)\right) \tag{5}$$

According to a recent study [22], we put a new constraint in the Soft-margin triplet loss to further require that distance of the pair $(\mathbf{x}, \mathbf{x}^{+})$ be less than a second margin τ_2, Translating this statement into equation, we have:

$$\| \phi(\mathbf{x})-\phi\left(\mathbf{x}^{+}\right)\|^{2}< \tau_2 \tag{6}$$

where τ_2 is a hyperparameter, and in this paper, we choose the value of τ_2 as 0.0265.

The enhanced loss function re-soft (Eq. (1)) seeks to bring the similarity of examples in the same category closer while pushing instances of another category further apart in the learned feature space. This is more compatible with the underlying principle employed by the majority of data clustering and discriminant analysis techniques.

To demonstrate the effectiveness of our proposed loss function, ablation experiments are performed on a benchmark dataset. For the conventional triplet loss we take a random value in the middle of 0–1 for γ.

The performance of our proposed loss function is investigated through the ablation experiments described as follows.

The impact on the model performance compared with the traditional triplet loss with the soft-margin loss applied in Fig. 4. The experimental findings indicate that our suggested loss function is better applicable to the subject of network security. The precision of the conventional margin triplet loss might vary significantly, (the highest F1 with $\gamma = 0.9$ in AAGM17, $\gamma = 1$ in CICIDS2017, $\gamma = 0.1$ in KDDCUP99 and $\gamma = 0.1$ in UNSW-NB15). The soft-margin loss achieved the second highest result (exception on UNSW-NB15). In any event, computing the re-soft loss in the training period enables us to attain the maximum F1 in the prediction phase across all datasets. Therefore, the ablation experiments in this section can demonstrate the effectiveness of the improved loss function.

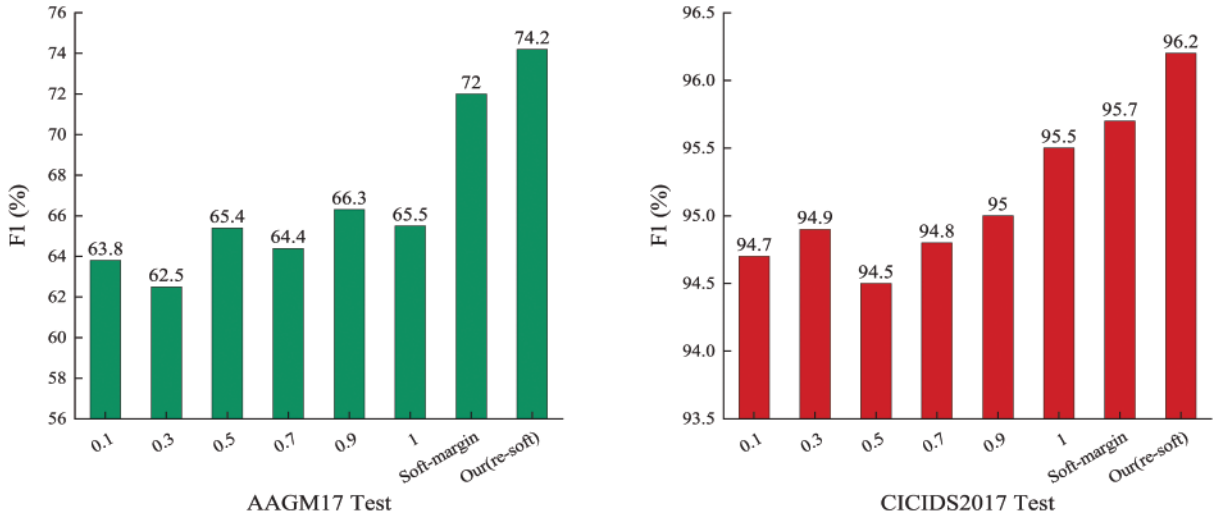

Figure 4: (Continued)

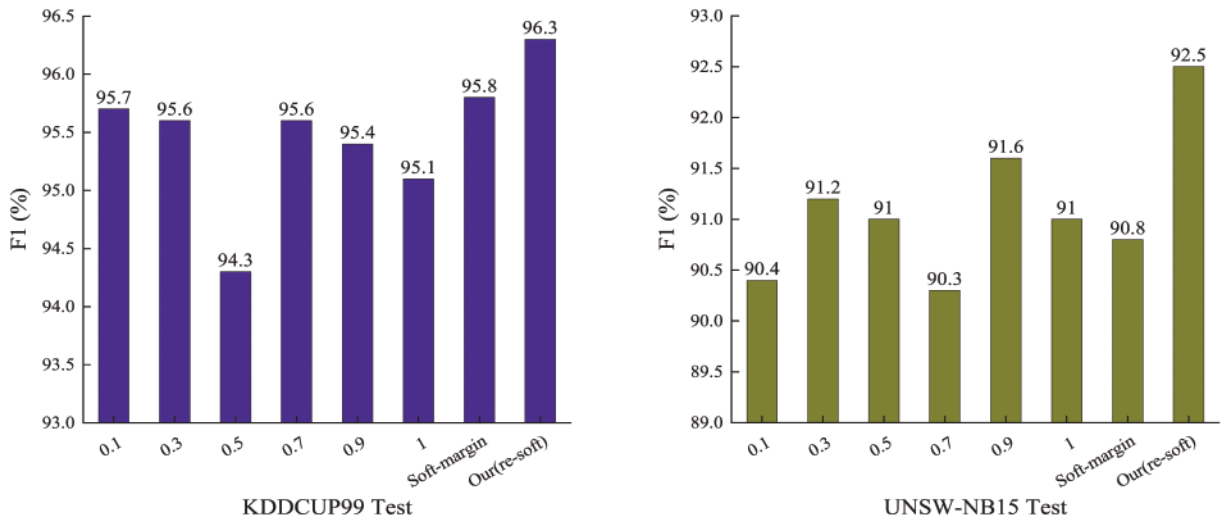

Figure 4: F1-score of traditional triplet loss by altering the margin between 0.1, 0.3, 0.5, 0.7, 0.9, and 1, soft-margin triplet loss, and our suggested re-soft

6 Competitor Analysis

Finally, we compare our findings to the most recent network intrusion detection literature. Specifically, DL architectures based on various:

A competitor with Long short-term memory (LSTM): Variational LSTM [23];

CNN-based competitors: Man [24];

GAN Network-based competitors: Efficient GAN [25], MAGNETO-GAN [26] and ALAD [27], MAD-GAN [28];

Competitors with autoencoder (AE): AIDA [29], DAGMM [30], THEODORA [31], and MINDFUL [32];

Recently DML studies: RENOIR [19];

We note that RENOIR-based, GAN-based, and AE-based rivals are the most analogous to MEM-TET. RENOIR [19] is a recent DML study that uses AE in combination with Triplet, but AE undergoes training overgeneralization, making the overall performance degraded. The memory module in MemAE can save the prototype patterns of training data and increase the reconstruction error for non-training data, which can better solve the problem of overgeneralization in traditional AE. GAN-based competitors [26–28] mostly use generative adversarial learning for unusual traffic detection and usually employ data augmentation techniques to address the class imbalance problem. In contrast, AE-based adversaries exploit encoder info. Specifically, THEODORA [31] uses a multichannel CNN to obtain autoencoder information and employs label reassignment to handle anomalous samples.

The VLSTM [23] model has recently been proposed to cope effectively with imbalances and high-dimensional problems. Although there are many unexpected strikes in the UNSW-NB15, the results demonstrate that our proposed model (MEM-TET) achieves good results, thanks to the combination of MemAE and Triplet network, which can learn the deep features of the data, making

the reconstruction error of non-training data larger, and the classification based on Euclidean distance shows good performance on unbalanced data. It can enhance the rate of unknown assault detection.

We collected the Accuracy and F1 for each approach in this comparison analysis since these metrics are often supplied in reference papers. MEM-TET outperforms all competing models, including the GAN-based model (excluding CICIDS2017) and THEODORA [31] (other than UNSW-NB15), as well as the most recent DML RENOIR [19], as shown in Table 2 for all the datasets (tested on all datasets).

Table 2: Comparative analysis

DataSet	Model	Instruction	Accuracy (%)	F1 (%)
AAGM17	**MEM-TET**	**MemAE + DML**	**91.61**	**74.24**
	RENOIR [19]	AE + DML	89.63	71.90
	MAGNETO-GAN [26]	CNN + GAN	88.03	66.79
	MINDFUL [32]	AE + CNN	86.15	51.62
	THEODORA [31]	AE + CNN	87.62	65.92
	AIDA [27]	AE + MLP	86.06	57.78
CICIDS2017	**MEM-TET**	**MemAE + DML**	**99.45**	**96.23**
	RENOIR [19]	AE + DML	98.24	95.70
	MINDFUL [32]	AE + CNN	97.90	94.93
	AIDA [27]	AE + MLP	94.50	85.80
	THEODORA [31]	AE + CNN	98.03	95.25
KDD CUP99	**MEM-TET**	**MemAE + DML**	**95.27**	**96.34**
	RENOIR [19]	AE + DML	93.50	95.80
	MINDFUL [32]	AE + CNN	92.49	95.13
	THEODORA [31]	AE + CNN	92.97	95.46
	MAGNETO-GAN [26]	CNN + GAN	93.29	95.66
	AIDA [27]	AE + MLP	92.36	95.04
	Efficient GAN [25]	GAN	-	93.72
	ALAD [27]	GAN	-	95.01
	MAD-GAN [28]	GAN	-	90.00
	DAGMM [30]	AE + GMM	-	93.80
UNSW-NB15	**MEM-TET**	**MemAE + DML**	**91.36**	**92.52**
	RENOIR [19]	AE + DML	88.42	90.82
	Variational LSTM [23]	LSTM	88.30	90.70
	MAGNETO-GAN [26]	CNN + GAN	89.73	91.97
	Man [24]	RLF-CNN	88.70	-

Note: (Accuracy and F1 were gauged on test data from AAGM17, CICIDS2017, KDDCUP99, and UNSW-NB15. From the reference papers, the outcomes of rivals are extracted. The top results are highlighted. "–" indicates that no value is reported in the cited work.)

Apart from that, the latest DML-based study [RENOIR [19], Vec2im-SIAM [33]] was also analyzed by us for comparison.

In [33], the authors validated the accuracy of Vec2im-SIAM using the NSL-KDD dataset as a test set. This dataset is an updated version of KDDCUP99, generated by deleting repeat items from the original data. While the code for Vec2im-SIAM is confidential, we repeated the equivalent experimental setup described in [33], which utilized KDDTrain+ and KDDDTest+ as the training and testing sets, respectively. Fig. 5 illustrates the precision of MEM-TET RENOIR [19] and Vec2im-SIAM [33]. We observe that the accuracy of MEM-TET of 99.82% outperforms RENOIR [19] and Vec2im-SIAM [33] in the DML literature, achieving competitive performance. To explore the performance of mem in multiple classification scenarios, we perform more comparisons in Section 8, including DML-based as well as deep neural networks.

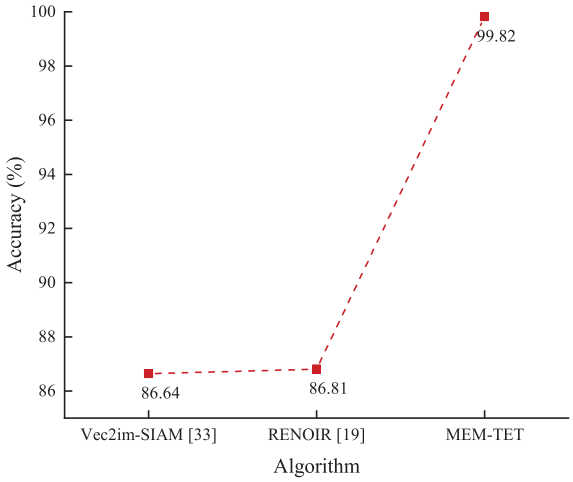

Figure 5: Accuracy of Vec2im-SIAM, RENOIR, and MEM-TET on NSL-KDD. Data were collected from the literature

7 Detection of Unknown Attacks

To demonstrate the performance of the proposed model MEM-TET to detect unknown attacks, we conducted additional experiments using AAGM as the training set and the CICIDS dataset as the test set. Since the AAGM attributes have 80 dimensions and the CICIDS dataset features only have 79 dimensions, we fill the missing column with 0 for the CICIDS dataset. In this current experiment, for the training set, the attack types in the test set are unknown, which can effectively simulate the unknown attacks that occur in natural network environments. The results are shown in Fig. 6. The MEM-TET model scores 47.82% in the CICIDS test dataset F1. In comparison, RENOIR [19] scored only 38.45%, demonstrating that the proposed model also has excellent advantages in unknown attack detection.

8 Multiclassification Analysis

To illustrate the efficacy of our suggested strategy in a scenario involving numerous classifications. To this purpose, we continue to evaluate the NSL-KDD dataset. Table 3 shows the sample size of each category in addition to the data set. It can be observed that U2R and R2L are both uncommon assaults because multiple algorithms have recently validated this dataset. To differentiate between expected flows and attack flows, we ultimately devised a two-stage model for multi-classification, the binary

classification of the MEM-TET. Finally, we use XGBoost to perform multiclassification experiments on the test data.

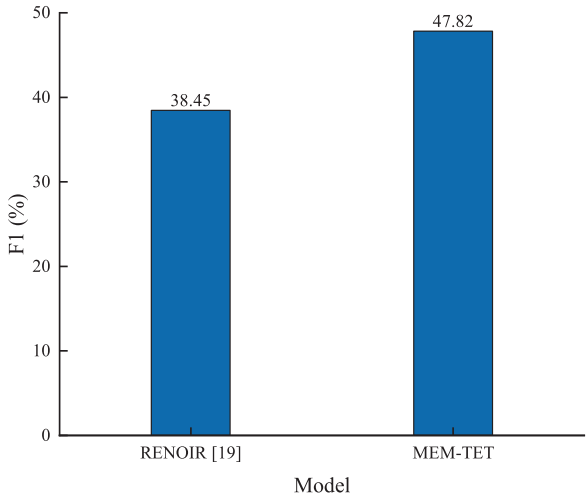

Figure 6: Unknown attack detection, using AAGM as the training set and CICIDS2017 as the test set

Table 3: Description of the NSL-KDD dataset

DataSet	Total	Dos	Probe	U2R	R2L	Normal
KDDTrain+	25192	9234	2289	11	209	13449
KDDTest+	22544	7458	2421	200	2754	9711

We note that the multi-classifier in [12] is also using XGBoost.

MEM-TET + XGBoost differs from [12] in that it uses Triplet rather than a combination of Siamese network, DNN, and XGBoost binary classifier to distinguish regular traffic and intrusive behavior.

For the comparison research, we evaluate the performance of several contemporary rivals that approach the same multi-class problem by incorporating various strategies to counter uncommon assaults.

Primarily, we evaluate the following DML Multi-Category rivals: SIAM-IDS [11], I-SiamIDS [12], and RENOIR + XGBoost [19].

DL for data enhancement: Poulmanogo [34], IGAN-IDS [9], and DGM-RELU [35].

Intensive Learning: AESMOTE [36] and LIO-IDS [37].

Table 4 shows the comparison results. These aggregate metrics confirm that our proposed method is effective. We note that MEM-TET outperforms SIAM-IDS [11], RENOIR [19], and I-SiamIDS [12] in the DML-based results. In addition, we noticed that LIO-IDS [37] achieved second place in the Micro-Average F1 indicator because LIO-IDS [37] also used a two-layer model. By the LSTM classifier, Layer 1 of LIO-IDS [37] recognizes intrusions from a regular data stream. Layer 2 employs ensemble methods to categorize observed incursions into several attack categories. On the other hand,

the IGAN-IDS [9] employs a deep GAN procedure to produce novel minority class data. However, this approach requires more complex operations to make the training dataset increase.

Table 4: Multiclassification results on the NSL-KDD dataset

Method	Macro-average F1(%)	Micro-average(OA)(%)	Weighted F1(%)
our proposed	**72.81**	**85.29**	**85.52**
RENOIR + XGBoost [19]	66.58	83.91	83.05
I-SiamIDS [12]	66.54	79.90	78.49
SIAM-IDS [11]	56.14	76.96	75.28
AESMOTE [36]	-	82.09	82.43
IGAN-IDS [9]	-	84.45	84.17
DGM-RELU(SVM) [35]	-	-	73.00
LIO-IDS [37]	70.20	-	-
Poulmanogo [34]	-	84.25	-

Note: (Compared to existing studies, our technique achieves the best results. Favorable results are bold. Reference articles collect rivals' results. "-" indicates no source value.)

Overall, it is evidenced, based on the experimental results, that the research proposed in this paper has achieved the expected results. In addition, it also lays the foundation for further research on DML in intrusion detection.

9 Conclusions

In the realm of intrusion detection, training is unbalanced most of the time due to the presence of a few categories and presence of a substantial quantity of unknown attacks. It degrades the detection performance of classical machine learning models. This paper introduces MEM-TET, a unique DML methodology that specifies an innovative Triplet network strategy by exploiting memory-augmented autoencoder information to increase the detection rate of unknown attacks. Apart from this, a new loss function called re-soft is proposed, which can better pull in the intra-class distance and push out the inter-class distance. The generic benchmark dataset shows that our implemented model exceeds existing state-of-the-art models for binary and multivariate classifications. In the future, we will further investigate the utilization of DML in network intrusion detection.

Acknowledgement: We declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere.

Funding Statement: This research is a basic research project carried out with the support of National Natural Science Foundation of China (U1936213), Yunnan Provincial Natural Science Foundation, "Robustness analysis method and coupling mechanism of complex coupled network system" (202101AT070167), Yunnan Provincial Major Science and Technology Program, "Construction and application demonstration of intelligent diagnosis and treatment system for childhood diseases based on intelligent medical platform" (202102AA100021).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Weifei Wang, Jinguo Li; analysis and interpretation of results: Weifei Wang, Jinguo Li and

Na Zhao; Manuscript proofing: Min Liu. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The publicly available data set can be found at: https://www.unb.ca/cic/datasets/android-adware.html, https://www.unb.ca/cic/datasets/ids-2017.html, https://www.tensorflow.org/datasets/catalog/kddcup99, https://research.unsw.edu.au/projects/unsw-nb15-dataset, https://www.unb.ca/cic/datasets/nsl.html.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] U. Cisco, "Cisco annual internet report (2018–2023) white paper," *Cisco: San Jose, CA, USA*, vol. 10, no. 1, pp. 1–35, 2020.
[2] C. Li, J. Wang and X. Ye, "Using a recurrent neural network and restricted boltzmann machines for malicious traffic detection," *NeuroQuantology*, vol. 16, no. 5, pp. 823–831, 2018.
[3] H. He, Y. Bai, E. A. Garcia and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE Int. Joint Conf. on Neural Networks (IEEE World Congress on Computational Intelligence)*, IEEE, Hong Kong, China, pp. 1322–1328, 2008.
[4] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
[5] J. P. Anderson, "Computer security threat monitoring and surveil-lance," *Technical Report, James P. Anderson Company*, 1980.
[6] H. Zhang, L. Huang, C. Q. Wu and Z. Li, "An effective convolutional neural network based on smote and Gaussian mixture model for intrusion detection in imbalanced dataset," *Computer Networks*, vol. 177, pp. 107315, 2020.
[7] X. Xu, J. Li, Y. Yang and F. Shen, "Toward effective intrusion detection using log-cosh conditional variational autoencoder," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6187–6196, 2020.
[8] D. Gonzalez-Cuautle, A. Hernandez-Suarez, G. Sanchez-Perez, L. K. Toscano-Medina, J. Portillo-Portillo *et al.,* "Synthetic minority over-sampling technique for optimizing classification tasks in botnet and intrusion-detection-system datasets," *Applied Sciences*, vol. 10, no. 3, pp. 794, 2020.
[9] S. Huang and K. Lei, "Igan-ids: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *AdHoc Networks*, vol. 105, pp. 102177, 2020.
[10] L. Gautheron, A. Habrard, E. Morvant and M. Sebban, "Metric learning from imbalanced data with generalization guarantees," *Pattern Recognition Letters*, vol. 133, pp. 298–304, 2020.
[11] P. Bedi, N. Gupta and V. Jindal, "Siam-ids: Handling class imbalance problem in intrusion detection systems using siamese neural network," *Procedia Computer Science*, vol. 171, pp. 780–789, 2020.
[12] P. Bedi, N. Gupta and D. V. Jindal, "I-Siamids: An improved siam-ids for handling class imbalance in network-based intrusion detection systems," *Applied Intelligence*, vol. 51, pp. 1133–1151, 2021.
[13] J. Li, C. Gu, F. Wei, X. Zhang, X. Hu *et al.,* "Light-seen: Real-time unknown traffic discovery via lightweight Siamese networks," *Security and Communication Networks*, vol. 2021, pp. 1–12, 2021.
[14] X. Zhou, W. Liang, S. Shimizu, J. Ma and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5790–5798, 2020.
[15] A. Zhou, Z. Li and Y. Shen, "Anomaly detection of can bus messages using a deep neural network for autonomous vehicles," *Applied Sciences*, vol. 9, no. 15, pp. 3174, 2019.
[16] Y. Wang, Y. Jiang and J. Lan, "Intrusion detection using few-shot learning based on triplet graph convolutional network," *Journal of Web Engineering*, vol. 20, no. 5, pp. 1527–1552, 2021.

[17] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *Similarity-Based Pattern Recognition: Third Int. Workshop, SIMBAD 2015, Copenhagen, Denmark, October 12–14, 2015. Proc. 3*, Springer, Copenhagen, Denmark, pp. 84–92, 2015.
[18] D. Gong, L. Liu, V. Le, B. Saha, M. R. Mansour *et al.,* "Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, Seoul, Korea (South), pp. 1705–171, 2019.
[19] G. Andresini, A. Appice and D. Malerba, "Autoencoder-based deep metric learning for network intrusion detection," *Information Sciences*, vol. 569, pp. 706–727, 2021.
[20] F. Schroff, D. Kalenichenko and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 815–821, 2015.
[21] A. Hermans, L. Beyer and B. Leibe, "In defense of the triplet loss for person re-identification," arXiv Preprint arXiv: 1703.07737, 2017.
[22] D. Cheng, Y. Gong, S. Zhou, J. Wang and N. Zheng, "Person re-identification by multi-channel parts-based cnn with improved triplet loss function," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 1335–1344, 2016.
[23] X. Zhou, Y. Hu, W. Liang, J. Ma and Q. Jin, "Variational lstm enhanced anomaly detection for industrial big data," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3469–3477, 2020.
[24] J. Man and G. Sun, "A residual learning-based network intrusion detection system," *Security and Communication Networks*, vol. 2021, pp. 1–9, 2021.
[25] H. Zenati, C. S. Foo, B. Lecouat, G. Manek and V. R. Chan-drasekhar, "Efficient gan-based anomaly detection," arXiv Preprint arXiv: 1802.06222, 2018.
[26] G. Andresini, A. Appice, L. D. Rose and D. Malerba, "Gan augment-tation to deal with imbalance in imaging-based intrusion detection," *Future Generation Computer Systems*, vol. 123, pp. 108–127, 2021.
[27] H. Zenati, M. Romain, C. -S. Foo, B. Lecouat and V. Chandrasekhar, "Adversarially learned anomaly detection," in *2018 IEEE Int. Conf. on Data Mining (ICDM)*, IEEE, Singapore, pp. 727–736, 2018.
[28] D. Li, D. Chen, B. Jin, L. Shi, J. Goh *et al.,* "Mad-gan: Multivariate anomaly detection for time series data with generative adversarial networks," in *Artificial Neural Networks and Machine Learning–ICANN 2019: Text and Time Series: 28th Int. Conf. on Artificial Neural Networks, Part IV*, Springer International Publishing, Cham, Munich, Germany, pp. 703–716, 2019.
[29] G. Andresini, A. Appice, N. Di Mauro, C. Loglisci and D. Malerba, "Exploiting the auto-encoder residual error for intrusion detection," in *2019 IEEE European Symp. on Security and Privacy Workshops (EuroS&PW)*, IEEE, Stockholm, Sweden, pp. 281–290, 2019.
[30] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu *et al.,* "Deep autoencoding Gaussian mixture model for unsu-pervised anomaly detection," in *Int. Conf. on Learning Representations (ICLR 2018)*, Vancouver, Canada, 2018.
[31] G. Andresini, A. Appice, F. Paolo Caforio and D. Malerba, "Improving cyber-threat detection by moving the boundary around the normal samples," *Machine Intelligence and Big Data Analytics for Cybersecurity Applications. Springer*, vol. 919, pp. 105–127, 2021.
[32] G. Andresini, A. Appice, N. D. Mauro, C. Loglisci and D. Malerba, "Multi-channel deep feature learning for intrusion detection," *IEEE Access*, vol. 8, pp. 53346–53359, 2020.
[33] S. Moustakidis and P. Karlsson, "A novel feature extraction methodology using siamese convolutional neural networks for intrusion detection," *Cybersecurity*, vol. 3, no. 1, pp. 1–13, 2020.
[34] P. Illy, G. Kaddoum, C. M. Moreira, K. Kaur and S. Garg, "Securing fog-to-things environment using intrusion detection system based on ensemble learning," in *2019 IEEE Wireless Communications and Networking Conf. (WCNC)*, IEEE, Marrakesh, Morocco, pp. 1–7, 2019.
[35] G. Dlamini and M. Fahim, "Dgm: A data generative model to improve minority class presence in anomaly detection domain," *Neural Computing and Applications*, vol. 33, no. 20, pp. 13635–13646, 2021.

[36] X. Ma and W. Shi, "Aesmote: Adversarial reinforcement learning with smote for anomaly detection," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 943–956, 2020.
[37] N. Gupta, V. Jindal and P. Bedi, "Lio-ids: Handling class imbalance using lstm and improved one-vs-one technique in intrusion detection system," *Computer Networks*, vol. 192, pp. 108076, 2021.