Advanced Threat Detection in Cyber-Physical Systems using Lemurs Optimization Algorithm with Deep Learning

Omar Ahmed Abdulkader[1,*] , Muhammad Jawad Ikram[1]

[1]Faculty of Computer Studies, Arab Open University, Riyadh, Saudi Arabia
Emails: o.abdulkader@arabou.edu.sa; m.ikram@arabou.edu.sa

Abstract

Cyber-physical systems (CPS) are significant to main organizations like Smart Grids and water conduct and are gradually helpless to an extensive range of developing threats. Identifying threats to CPS is of greatest significance, owing to their progressive frequent usage in numerous critical assets. Traditional safety devices like firewalls and encryption are frequently insufficient for CPS designs; the execution of Intrusion Detection Systems (IDSs) personalized for CPS is a crucial plan for safeguarding them. Artificial intelligence (AI) techniques have shown abundant probability in numerous areas of network security, mainly in network traffic observation and in the recognition of unauthorized access, misuse, or denial of network resources. IDS in CPSs and other fields namely the Internet of Things, is regularly considered through deep learning (DL) and machine learning (ML). This manuscript offers the design of an Advanced Threat Detection utilizing the Lemurs Optimization Algorithm with Deep Learning (ATD-LOADL) methodology in the CPS platform. The primary of the ATD-LOADL methodology is to focus on the recognition and classification of cyber threats in CPS. In the preliminary phase, the pre-processing of the CPS data takes place using a min-max scaler. To select an optimum set of features, the ATD-LOADL technique uses LOA as a feature selection approach. For threat detection, the ATD-LOADL algorithm uses a multi-head attention-based long short-term memory (MHA-LSTM) classifier. At last, the detection results of the MHA-LSTM method are boosted by the use of the shuffled frog leap algorithm (SFLA). The experimentation outcomes of the ATD-LOADL approach can be widely investigated on a benchmark CPS dataset. An experimentation outcome stated the enhanced threat detection results of the ATD-LOADL technique over other existing approaches

Keywords: Cyber-Physical System; Threat Detection; Lemurs Optimization Algorithm; Deep Learning; Hyperparameter Tuning

1. Introduction

A cyber-physical system (CPS) is the linkage of physical and cyber methods, while the interchange of information and data occurs in recent times. CPS plays a vital part in Internet of Things (IoT) based business and provides an extensive financial possibility [1]. CPS estimates the contact of network, physical, and computing methods and is dependent upon the IoT. It was developed as the cyber internet of physical things that provides a huge sort of services like e-health, e-commerce, smart cities, smart homes, etc. A huge amount of industrial tools can be organized wirelessly by accepting CPS which aids in handling difficult and huge industrial systems [2]. Interrelated modules of CPS have the capability to wisdom environments and procedure the IoT-based substances slightly. It has the flexibility to modify the developments in computing [3]. Besides, CPS is enclosed in numerous methods and employed in different areas such as transport, military, health care, communication, and several autonomous systems. It permits the remote commands and control of devices, machines, and systems that are vital in several business environments [4]. However, the wide execution of CPS has many safety threats that cause critical harm to organized physical objects and damage the consumers who entirely depend on them. The main safety devices to defend CPS gadgets from exterior attacks are trusted in firewalls, encryption, and anti-virus methods [5]. However, these devices cannot able to ensure all assaults, particularly by examining that attackers are always

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

developing their tactics. In this situation, utilizing Intrusion Detection Systems (IDSs) is important for identifying malicious conduct and shielding the CPSs from attacks [6]. IDSs use Machine Learning (ML) approaches to identify malicious actions by trusting training datasets [7]. But, many researchers in the study employ datasets gathered from common internet rules. These datasets are not appropriate for IDSs in CPSs, because they have a connection with the real present tool and lack traffic from classic protocols of CPS [8].

Therefore, IDSs must be executed on such methods so that protective activities can be taken earlier when there is permanent harm owing to these attacks. To identify attacks as well as accidental faults in CPSs, anomaly detection techniques have been developed to diminish these threats [9]. Deep learning (DL) offers better performance than traditional ML solutions. In addition, it is also detected that many DL methods have been presented in current journals for detecting CPS cyberattacks [10]. An extensively accepted sight to clarify the trouble of discovering cyberattacks on CPSs was certified to the complexity level when affixing cybersecurity over CPS.

This manuscript offers the design of an Advanced Threat Detection using the Lemurs Optimization Algorithm with DL (ATD-LOADL) method in the CPS platform. In the preliminary phase, the pre-processing of the CPS data takes place using a min-max scaler. To select an optimum set of features, the ATD-LOADL technique uses LOA as a feature selection approach. For threat detection, the ATD-LOADL technique uses a multi-head attention-based long short-term memory (MHA-LSTM) model. Lastly, detection results of the MHA-LSTM method are boosted through the usage of the shuffled frog leap algorithm (SFLA). The experimental outcomes of the ATD-LOADL methodology are widely investigated on benchmark CPS data.

2. Related works

Sharma et al. [11] developed a fundamental factor of IDS depending on the significant parameter safety. An effective and lightweight DL-based Convolutional Neural Networks (CNNs)-Bidirectional LSTM technique has been developed for the DDoS recognition that executes the features CNNs method to categorize traffic movements as benign and malicious in this research. Althobaiti et al. [12] project an innovative perceptive computing-based IDS method to attain safety in CPS. This technique includes a pre-processed step to eliminate the sound that occurs in the data. The binary bacterial foraging optimizer (BBFO) has been used for feature selection (FS). Moreover, the gated recurrent unit (GRU) technique has been utilized to classify the intrusions. Lastly, Nesterov-accelerated Adaptive Moment Estimation (NADAM) optimization has been utilized as a hyperparameter optimizer of the GRU method.

Ashraf et al. [13] developed an IDS technique for the present network atmosphere by considering the information from terrestrial and satellite systems. Integrating ML techniques, the research develops an ensemble method RFMLP that incorporates multilayer perceptron (MLP) and random forest (RF) to develop the solution of intrusion detection. For examining the efficacy of the projected structure, three dissimilar databases have been utilized. Li [14] presents a DL structure for the recognition and analysis of attacks on Photovoltaic (PV) Systems. Wide quantitative tests are led by using both micro-PMU and waveform data. Then, an adaptive hierarchical structure for the recognition and position of attacks in distribution methods has been offered. Furthermore, models like transfer learning (TL) and few-shot learning have been examined for improving the usage of labelled data models.

Deng et al. [15] proposed an effective attack recognition model by employing a difficult network-based FS and DL model, denoted as VFD-AE. Specifically, the technique mines the possible links of features and achieves actual FS by assuming the significance estimation model in difficult network theory. Besides, owing to the threat model being short in supply, an unsupervised recognition system by Auto encoder (AE) has been planned. Jahromi et al. [16] project a dual-level ensemble attack recognition and attribution infrastructure planned for CPS, and more exactly in an industrial control system (ICS). In the primary stage, a DT joined by a new ensemble deep representation learning approach was proposed to identify threats to imbalanced ICS atmospheres. In the next stage, an ensemble of DNNs was intended to simplify threat attribution.

In [17], the anti-honeypot enabled attack recognition method for ICPS has been proposed by employing the Reinforcement learning (RL) and Stakerlberg dynamic game (SDG) techniques. The connections between the attackers and ICPS protectors have been arrested via the BSDG method. RL condition and rewards work display numerous probable ICPS defence and aggressive attackers. It will seize a threat series in ICPS and classify the attackers in an efficient method. Duhayyim et al. [18] proposes an innovative SFSA-DLIDS technique. This method chiefly executes a min-max data standardization model for converting input data to a well-suited setup. The SFSA approach has been employed to pick a featured sub-set. In addition, a chicken swarm optimizer (CSO) with a deep stacked auto-encoder (DSAE) model has been applied to identify and classify the intrusion.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

3. The Proposed Method

In this manuscript, we offer the design of an ATD-LOADL methodology in the CPS platform. The purpose of the ATD-LOADL model concentrated on the detection and recognition of cyber threats in the CPS environment. Fig. 1 shows the workflow of ATD-LOADL methodology.

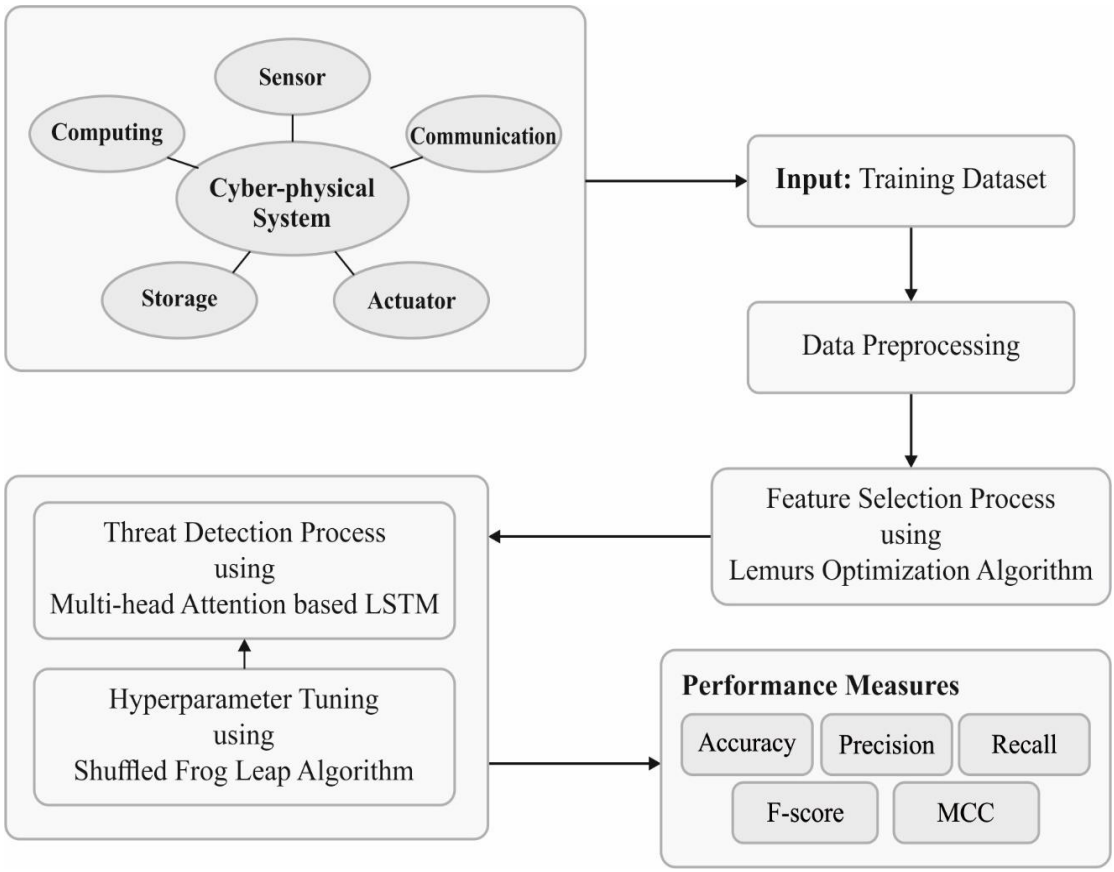

Figure 1. Workflow of ATD-LOADL methodology

A. Data normalization

In the preliminary phase, the pre-processing of the CPS data takes place using a min-max scaler. Min-max scaling, also called min-max normalization or feature scaling, is a data pre-processed method normally employed in ML and statistics [19]. The main intention of min-max scaling is to convert the numeric values of a dataset into an exact range, naturally [0, 1], while conserving the relative changes among the original values. This is attained by deducting the minimal value of the feature from every data point and then separating the outcome by the range (the difference between the maximal and minimal values). The formulation for min-max scaling is:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Here, X represents the original value, Xmin indicates the minimal value in the dataset, and Xmaxrefers to the maximal value. Min-max scaling is chiefly beneficial when features in a dataset have dissimilar scales, averting certain features from extremely influencing the learning procedure. It is an easy and effective method to regularize data, certifying that all features donate similarly to the model's training procedure and enhancing the values and convergence of numerous ML techniques.

B. LOA-based feature selection

At this stage, the ATD-LOADL technique uses LOA as a feature selection approach. LO is a great population-based technique, thus the lemur set is given in the matrix method [20]. Eq. (2) determines the input population matrix for LO model.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

$$X = \begin{bmatrix} l_1^1 & l_1^2 & \cdot & l_1^d \\ l_2^1 & l_2^2 & \cdot & l_2^d \\ \vdots & \vdots & \cdot & \vdots \\ l_n^1 & l_n^2 & \cdot & l_n^d \end{bmatrix} \quad (2)$$

Where the matrix of size $n \times d$ indicates X. The candidate solutions point to n, and the dimension represents d. The LO is used for resolving optimization problems such as Feature selection (FS), the steps of the LO technique are given below:

Step1: Determine the succeeding Lemur parameters: N Population, Max_{iter} represents the maximal iteration number. The dimension of the search range over the dataset size is represented as d. In addition, UB and LB are the up-and-low limitations of the problems.

Step2: Generate X decision variables at an i^{th} solution using the following expression:

$$X_i^j = \left(LB + \left(UB_j - LB_j\right)\right) \times r \quad (3)$$

Here r denotes the uniform distribution random value $\in [0, 1]$.

Step3: Evaluate the Free Risk Rate (FRR) which is the co-efficient of LO within the loop for every iteration:

$$FRR = HRR - t \times ((HRR - LRR)/\text{Max}_{iter} \quad (4)$$

In Eq. (4), the existing number of iterations is denoted as f. Max_{iter} indicates the iteration size. Low-Risk Rate (LRR) and High-Risk Rate (HRR) are the binary predetermined and constant values.

Step4: for each χj_i, evaluate the fitness values as follows:

$$Fit\left(x_i^j\right) = \alpha \times (1 - Acc) + \beta \times (s/S) \quad (5)$$

In Eq. (5), $Fit(x_i^j)$ denotes the fitness value, small s, and S is the overall and maximum amount of features nominated, and Acc refers to the accurateness of the subset extracted by $the\ KNN$ classifier to calculate the subset selected in all the iterations.

Step 5: it is classified into two processes to enhance the fitness value of lemurs. At first, the best near lemurs (bnl) have been defined which suggests choosing the solution with the lowest fitness value. bnl provides the superior feature for the current iteration based on the FS objective. Next, the global best lemur (gbl) is nominated in the overall population, which signifies the overall best solution.

Step6: Fix the random number value $r_1 \in [0,1]$, and equate it by FRR. Next, the position is upgraded for all the lemurs away from the risk-based model.

$$X_i^j = \begin{cases} x(i,j) + |(x(i,j) - x(bnl,j)| \times (r_3 - 0.5) \times 2;\ r_1 < FRR \\ x(i,j) + |(x(i,j) - x(gbl,j)| \times (r_3 - 0.5) \times 2;\ r_1 > FRR \end{cases} \quad (6)$$

In Eq. (6), r_1 denotes the arbitrary value $\in [0,1]$. The existing i^{th} lemurs of N^{th} the population is (i, j) which has been the candidate solution in j^{th} dimension. gbl refers to the global optimal lemur for the entire population at all iterations.

The fitness function (FF) replicates the classification accuracy and the amount of nominated features. It enlarges the exactness of classification and decreases the dimension of the nominated features. So, FF is applied to assess individual solutions, as shown in below Eq. (7).

$$Fitness = \alpha * ErrorRate + (1 - \alpha) * \frac{\#SF}{\#All\_F} \quad (7)$$

Where ErrorRate means the rate of error of classification applying the designated features. $ErrorRate$ Can be proposed as the ratio of improper categorized to the classification count arranged, transferred as a value between (0, 1). (Error Rate was the supplement of the accuracy of classification), $\#SF$ is the nominated features amount and $\#All\_F$ refers to the complete volume of feature in a unique database. α is employed to handle the importance of classifier superiority and sub-set length. α are agreed to 0.9 in our examinations.

C. Detection using MHA-LSTM model

For threat recognition, the ATD-LOADL technique uses the MHA-LSTM methodology. LSTM is a kind of Recurrent Neural Network (RNN) that is proposed for tackling the gradient vanishing problems that occur while training classical RNN [21]. Sepp Hochreiter and Jürgen Schmidhuber in 1997 introduced LSTM which has gained

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

popularity for many tasks including sequence data, such as recognition of speech, time series study, and Natural Language Processing (NLP).

At the core of LSTM are memory cells that could retain data over longer sequences, enabling them to capture relationships and dependencies in the input dataset. Every memory cell has three major elements:

Cell State (C_t): This determines the long-term memory of LSTM and also carries data from the previous time step-in.

Input Gate (i): This represents what data from the existing time step is to be kept in cell state C_t.

Forget Gate (f): This controls what data from the prior cell state to be discarded or forgotten.

Output Gate (0): This defines what amount of C_t is to be disclosed as the output.

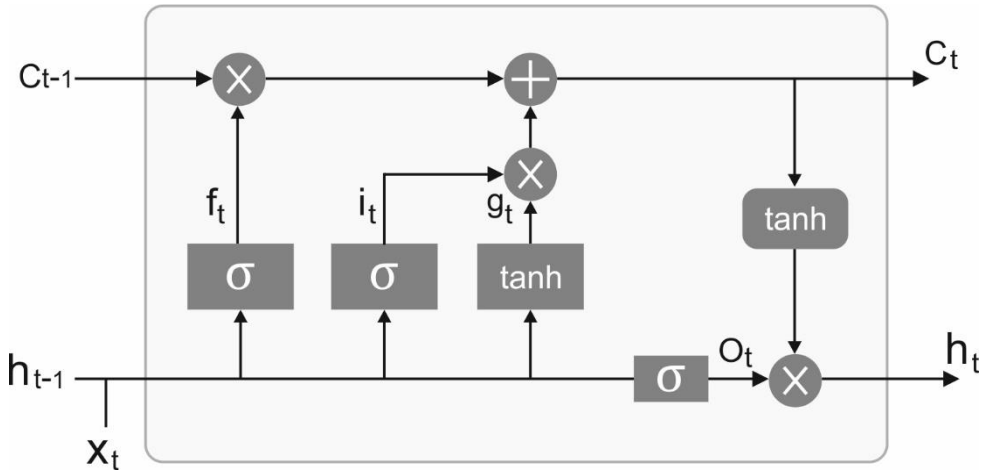

Figure 2. LSTM structure

Fig. 2 depicts the architecture of LSTM. The LSTM processes the input sequence step-wise, which updates the cell state and produces an output at the time step. The main concept behind LSTM is the usage of the above gates to regulate the data flow and employed as a sigmoid neural network layer that generates value within the range [0, 1].

In the forward pass of LSTM, the gate was calculated according to the present input, prior output, and prior C_t:

Input Gate (i):

$$i[t] = sigmoid(W_{ih}[t] + W_{ih}[h[t-1]] + b_i \quad (8)$$

Forget Gate (f):

$$f[t] = sigmoid(W_{fx}[t] + W_{fh}[h[t-1]] + b_f \quad (9)$$

Output Gate (0):

$$o[t] = sigmoid(W_{ox}[t] + W_{oh}[h[t-1]] + b_o \quad (10)$$

The input, forget, and output gates calculate the updated C_t and output at t time step:

Cell state update:

$$C[t] = f[t] * C[t-1] + i[t] * tanh(W_{cx}[t] + W_{ch}[h[t-1]] + b_c \quad (11)$$

Output:

$$h[t] = 0[t] * \tanh(C[t]) \quad (12)$$

Where the input at $the\ t$ time step is $x[t]$, the output of the prior time step is $h[t-1]$, and the weight matrix and bias vectors of the algorithm are W and b. LSTM could learn and recall long-term dependency in data sequence by selectively forgetting and updating data through the gate. Especially, this makes them well-suited for activities with temporal and context requirements. Backpropagation through Time (BPTT) is a backpropagation development that considers the successive properties of data while training LSTMs.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

LSTM has shown to be effective in processing and modelling data sequences, and is often used in a variety of applications where capturing and understanding long-term dependencies are crucial. MHA-LSTM integrates the successive modelling abilities of Long Short-Term Memory (LSTM) systems with the attention device resultant from transformers. This hybrid technique influences the capability of LSTMs to capture sequential dependencies in sequential data and improves it by presenting multi-head attention, permitting the system to instantaneously attend to dissimilar portions of the input sequence. By using this fusion, MHA-LSTM's main goal is to capture both short-term and long-term dependences more efficiently, providing enhanced performance in tasks that need nuanced accepting of consecutive designs.

D. Hyperparameter tuning using SFLA model

Eventually, the detection performances of the MHA-LSTM model can be boosted by the use of SFLA. SFLA is a novel population-based metaheuristic optimizer model that reproduces the memetic development of frog groups when observing a residence with a maximal quantity of obtainable food [22]. SFLA consists of definite as well as random plans for discovering an optimum response. The definite plan permits the method to employ surface-level data proficiently to guide the heuristic hunt. Accidental elements manage flexibility and control search design in this developed model. In this technique, every frog is measured as the best answer to the issue and a group of frogs creates a populace that transfers to grasp an exact goal. At the time of reaching the process, the population is separated into many sub-sets. The properties of frogs in every sub-group adjust result variables. After a definite quantity of evolutions, data is transferred among frogs during the procedure of merging sub-sets and generating novel populations and then a directed hunt is performed to define an optimum solution. This style endures until definite convergence situations develop.

In SFLA, an original populace of $sfla\_p$ frogs are arbitrarily created from probable solutions. The place or condition of the frog is a probable answer to the issue. This frog is then applied by paths and assemblies to specify values or problematic answers. In this technique, a whole early population is 1st separated into $sfla\_m$ clusters termed memplex. Various memplexes have $sfla\_n$ of a group of frogs that separately penetrates for an answer in hunt space. In every memplex, a sub-memplex is formed to prevent dropping in a local target. Each sub-memplex contains $sfla\_q$ frogs as well as nominated arbitrarily that depend on probability function which mentioned below:

$$P_j = \frac{2(sfla\_n + 1 - j)}{sfla\_n(sfla\_n + 1)}, j = 1,2, \dots, sfla\_n \tag{13}$$

Whereas P_j denotes the probability of selecting $the\ j$th frog for range and $sfla\_n$ signifies the number of frogs in memplex. In every memplex, frogs are organized as per a descendant sequence of fitness, where the possibility of choosing frogs is dropped. So, better-placed frogs in search space have a better opportunity of selecting an associate of sub-memplex. In every sub-memplex, the worst frog(P_w), executes diving depending on its individual skills and the location of the best frog (P_b). Then, the worst frog was nominated from the sub-memplex. The diving stage dimension for frog P_w is given below:

$$S_B = \begin{cases} \min\{int(rand.[P_b - P_w]).S_{\max}\} for a positive step \\ \max\{int(rand.[P_b - P_w]).-S_{\max}\} for a nagative step \end{cases} \tag{14}$$

Whereas, $rand$ denotes an arbitrary amount from the interval of zero and one; $S\_$ max is said to be the maximal dive length. In a subsequent stage, the worst frog place is revised as follows:

$$P'_w = P_w + S_B \tag{15}$$

If the novel Frog (P'_w) is enhanced than a new frog, then it is substituted by the original or else P_w is modified affording an optimum frog of total population (p) which is mentioned below:

$$S_G = \begin{cases} \min\{int(rand.[P_G - P_w]).S_{\max}\} for\ a\ positive step \\ \max\{int(rand.[P_G - P_w]).-S_{\max}\} for\ a\ nagative step \end{cases} \tag{16}$$

$$P''_w = P_w + S_G \tag{17}$$

Related to the preceding one, if P''_w frog is superior to the original frog (P_w), it is switched with P''_w frog and if either of those is confident, a novel frog at random is swapped with the poorest of sub-memplex. When the IT_{mem} stages of separating memplex into sub-memplexes, again all frogs united and re-separated into $sfla\_m$ memeplexes. This process continues to happen in program situations. By adopting this operation, the steadily usual fitness of the frog populace rises at the time of evolutionary phases and joins to a definite degree. With esteem to this procedure, P_G and P_w transformed to every iteration and the fitness variable was enhanced to unite the preferred answer. The SFLA model originates an FF to reach an upgraded classification solution. It determines an optimistic

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

number to suggest the higher values of the candidate results. In this study, the classification rate of error diminishing is reflected as FF, as provided in Eq. (18).

$$fitness(x_i) = ClassifierErrorRate(x_i)$$
$$= \frac{No.\ of\ misclassified\ instances}{Total\ no.of\ instances} * 100 \quad (18)$$

4. Experimental validation

This section validates the experimental analysis of the ATD-LOADL system under two datasets: NSLKDD2015 and CICIDS2017 datasets. Table 1 exemplifies the detailed description of 2 databases.

Table 1: Details on datasets

Classes	Datasets	
	NSLKDD2015	CICIDS2017
Normal	67343	25000
Anomaly	58630	25000
Total Instances	125973	50000

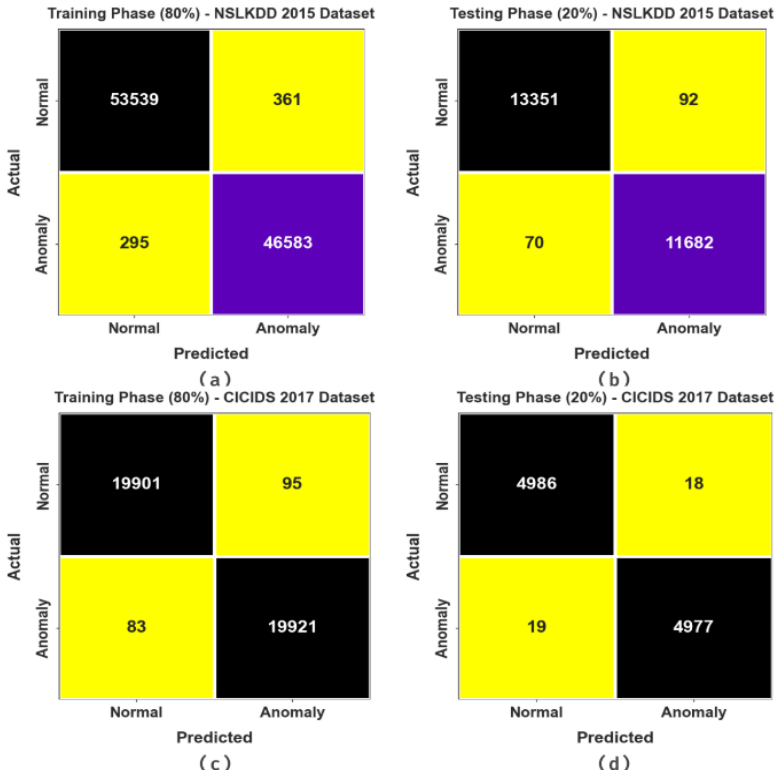

Figure 3. Confusion matrices of ATD-LOADL under 80%:20% TRAPS/TESPS (a-b) NSLKDD2015 and (c-d) CICIDS2017 datasets

Fig. 3 shows the confusion matrices acquired by the ATD-LOADL methodology under NSLKDD2015 and CICIDS2017 datasets. The accomplished result displays the adept recognition of normal and anomaly samples with all class labels.

The detection outcomes of the ATD-LOADL method under the NSLKDD2015 database are reported in Table 2 and Fig. 4. These obtained outcome demonstrate the ATD-LOADL method reaches effective identification of the normal and anomaly classes. Based on 80% of TRAPS, the ATD-LOADL methodology obtained an average $accu_y$ of 99.35%, $prec_n$ of 99.34%, $reca_l$ of 99.35%, F_{score} of 99.35%, and MCC of 98.69%. Additionally, based on 20% of TESPS, the ATD-LOADL model provides an average $accu_y$ of 99.36%, $prec_n$ of 99.35%, $reca_l$ of 99.36%, F_{score} of 99.35%, and MCC of 98.71%, respectively.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

Table 2: Detection outcomes of the ATD-LOADL system under the NSLKDD2015 dataset

NSLKDD2015 Database					
Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
80% of TRAPS					
Normal	99.33	99.45	99.33	99.39	98.69
Anomaly	99.37	99.23	99.37	99.30	98.69
Average	99.35	99.34	99.35	99.35	98.69
20% of TESPS					
Normal	99.32	99.48	99.32	99.40	98.71
Anomaly	99.40	99.22	99.40	99.31	98.71
Average	99.36	99.35	99.36	99.35	98.71

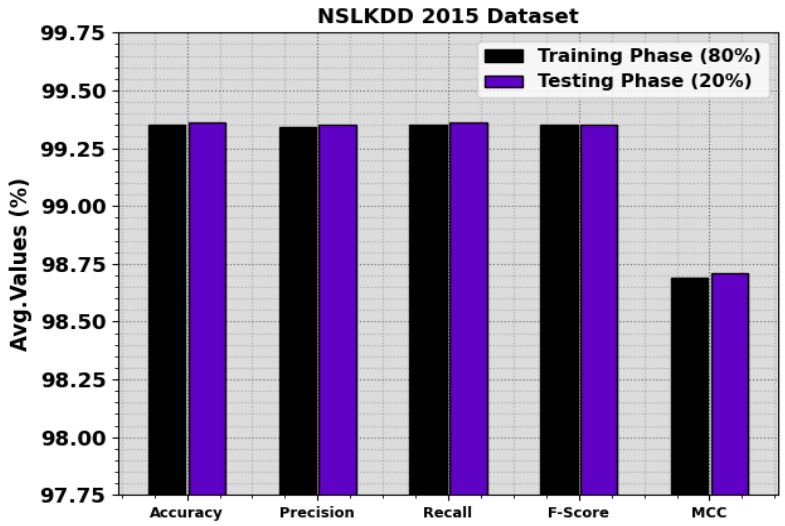

Figure 4. Average outcome of ATD-LOADL algorithm under NSLKDD2015 dataset

The $accu_y$ the curve for training (TR) and validation (VL) illustrated in Fig. 5 for the ATD-LOADL method at the NSLKDD2015 database provides an esteemed understanding of its efficiency in assorted epoch counts. Generally, it is a constant advancement at either TR or TS $accu_y$ with rising epoch counts, representing the efficiency of the technique in recognizing and learning patterns at both TS and TR data. The growing tendency in TS $accu_y$ highlights the model's flexibility for the TR database and its capacity for generating correct forecasts on unnoticed data, emphasizing proficiencies of strong generalizability.

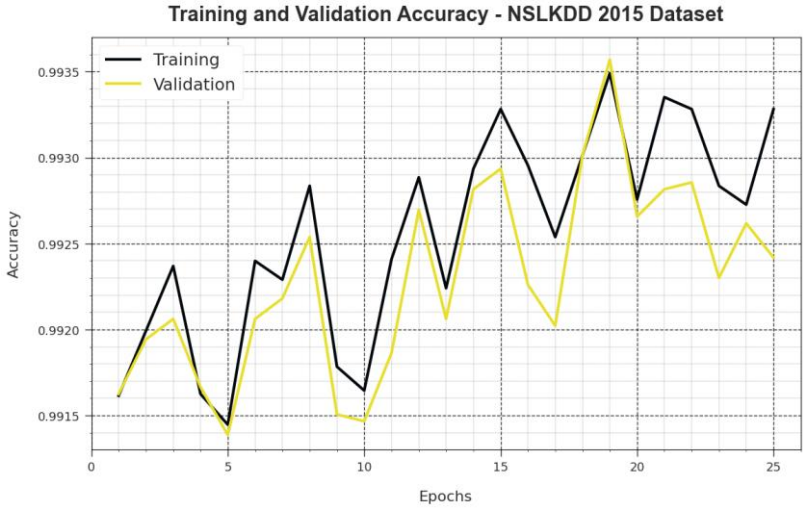

Figure 5. $Accu_y$ curve of the ATD-LOADL algorithm on the NSLKDD2015 dataset

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

Fig. 6 demonstrates a wide-ranging summary of the TS and TR loss for the ATD-LOADL methodology through the NSLKDD2015 data in varied epoch counts. The TR loss reliably diminutions as the method improves weights for decreasing classifier errors under 2 data. The loss investigation establishes the model's arrangement with the TR data, emphasizing its capability for excellently capturing patterns. Important can be a constant refinement of parameters in the ATD-LOADL approach, targeted at minimizing differences amongst predictions and real TR classes.

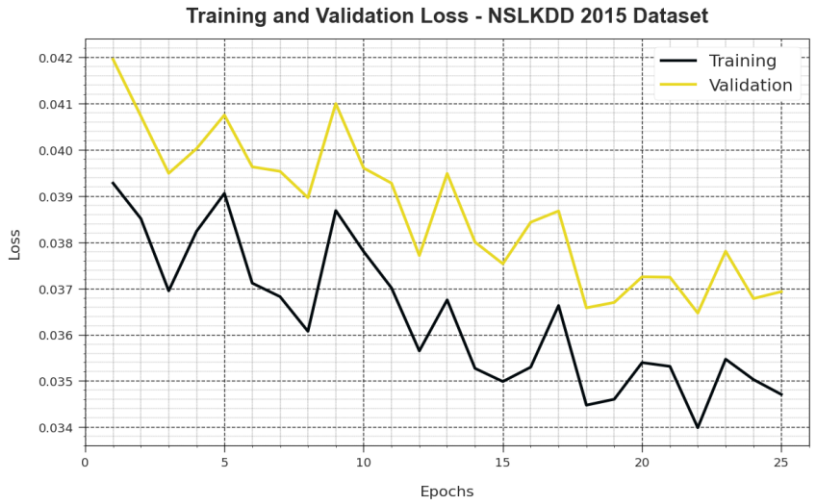

Figure 6. Loss curve of the ATD-LOADL technique under NSLKDD2015 dataset

The recognition study of the ATD-LOADL technique at the CICIDS2017 database is illustrated in Table 3 and Fig. 7. These experimental values display the ATD-LOADL methodology acquires effective identification with two classes. On 80% of TRAPS, the ATD-LOADL method offers an average $accu_y$ of 99.55%, $prec_n$ of 99.56%, $reca_l$ of 99.55%, F_{score} of 99.55%, and MCC of 99.11%. Also, with 20% of TESPS, the ATD-LOADL method gained an average $accu_y$ of 99.63%, $prec_n$ of 99.63%, $reca_l$ of 99.63%, F_{score} of 99.63%, and MCC of 99.26%, correspondingly.

Table 3: Detection of the ATD-LOADL methodology under the CICIDS2017 dataset

CICIDS2017 Database					
Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
805 of TRAPS					
Normal	99.52	99.58	99.52	99.55	99.11
Anomaly	99.59	99.53	99.59	99.56	99.11
Average	99.55	99.56	99.55	99.55	99.11
20% of TESPS					
Normal	99.64	99.62	99.64	99.63	99.26
Anomaly	99.62	99.64	99.62	99.63	99.26
Average	99.63	99.63	99.63	99.63	99.26

The $accu_y$ curves for VL and TR revealed in Fig. 8 for the ATD-LOADL methodology on the CICIDS2017 database provide a valued understanding of its efficacy in numerous epoch counts. Principally, it is a steady advancement at both TS and TR $accu_y$ with amplified epoch counts, suggesting the abilities of the method for identifying and learning patterns at these both data. The rising tendency in TS $accu_y$ emphasizes the model's adaptability for the database of TR and abilities to generate exact forecasts on hidden data, underscoring abilities of strong generalizability.

Fig. 9 illustrates a wide-ranging outline of the TS and TR loss of the ATD-LOADL algorithm on the CICIDS2017 database over several epoch counts. The TR loss steadily decreases as a model upgrade weight for diminishing classifier errors. These loss curves show the model's position through the data of TR, emphasizing its abilities for successfully capturing patterns. Noteworthy can be a nonstop development of parameters in the ATD-LOADL model, pointed at lessening discrepancies among actual and predicted TR classes.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

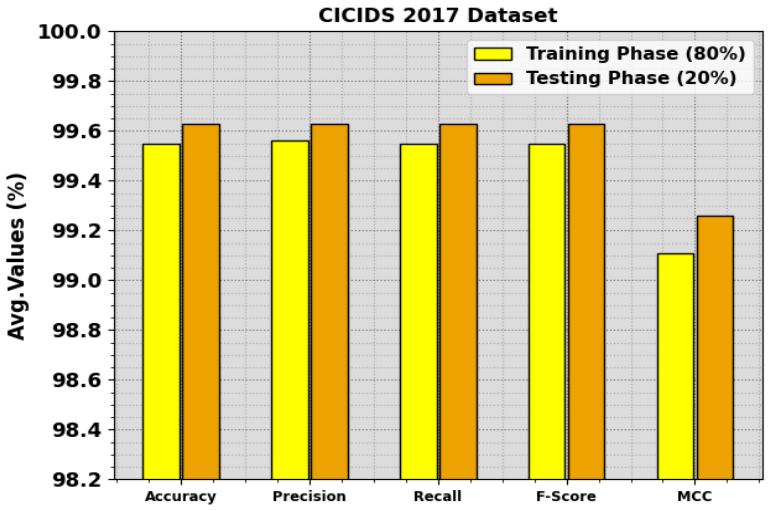

Figure 7. Average outcome of ATD-LOADL model under the CICIDS2017 dataset

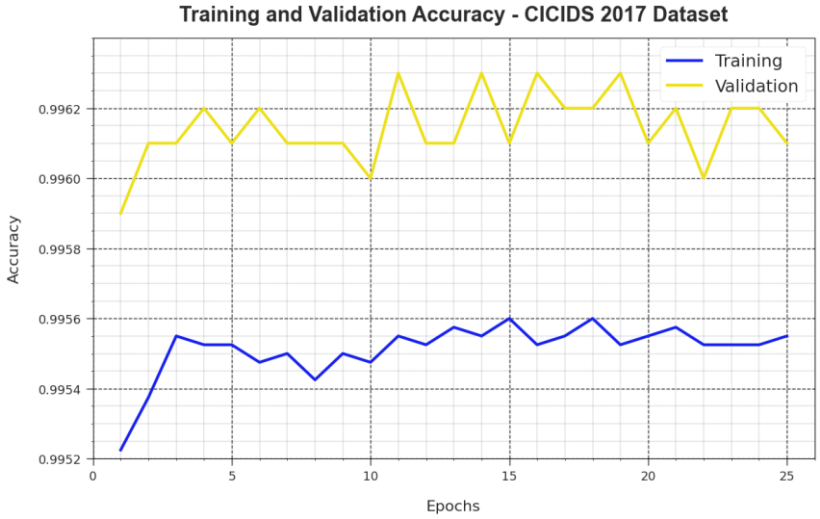

Figure 8. $Accu_y$ Curve of the ATD-LOADL model under the CICIDS2017 dataset

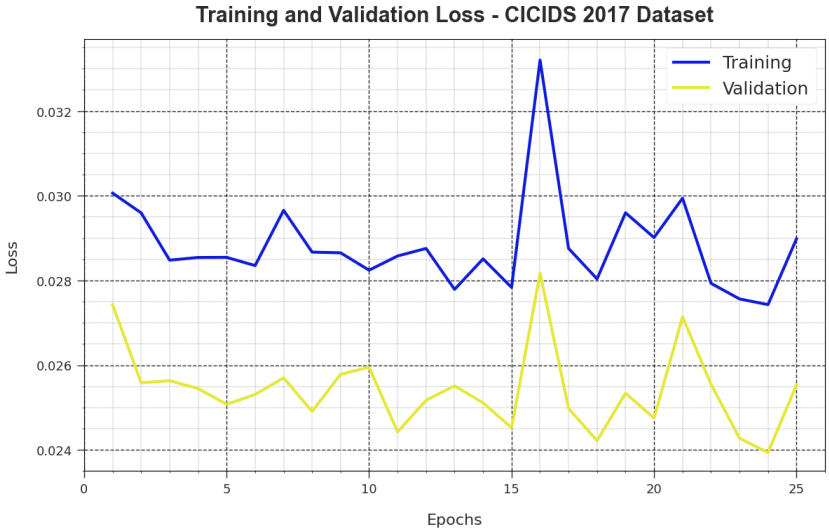

Figure 9. Loss curve of the ATD-LOADL technique with the CICIDS2017 dataset

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

The comparative results of the ATD-LOADL algorithm are defined in Table 4 and Fig. 10 [23]. The gained values stated that the OT method reaches worse results whereas the RF approach has resulted in somewhat higher performance. Moreover, the ATM-MFTDS, DBN, LSTM, and RNN systems gain reasonable results. Although the QDMOED-LTD model illustrates reasonable performance, the ATD-LOADL technique exhibits superior results with increased $accu_y$ of 99.63%, $prec_n$ of 99.63%, $reca_l$ of 99.63%, and F_{score} of 99.63%.

Table 4: Comparison outcome of the ATD-LOADL methodology with other algorithms

Methods	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
ATD-LOADL	99.63	99.63	99.63	99.63
QDMOED-LTD	99.52	99.52	99.52	99.52
ATM-MFTDS	98.72	99.27	99.09	99.08
DBN	98.59	98.77	98.33	98.25
LSTM	98.40	98.55	98.20	98.16
RNN	98.74	97.78	98.44	98.09
OT	93.84	95.93	92.54	95.48
RF	96.03	97.65	93.49	96.01

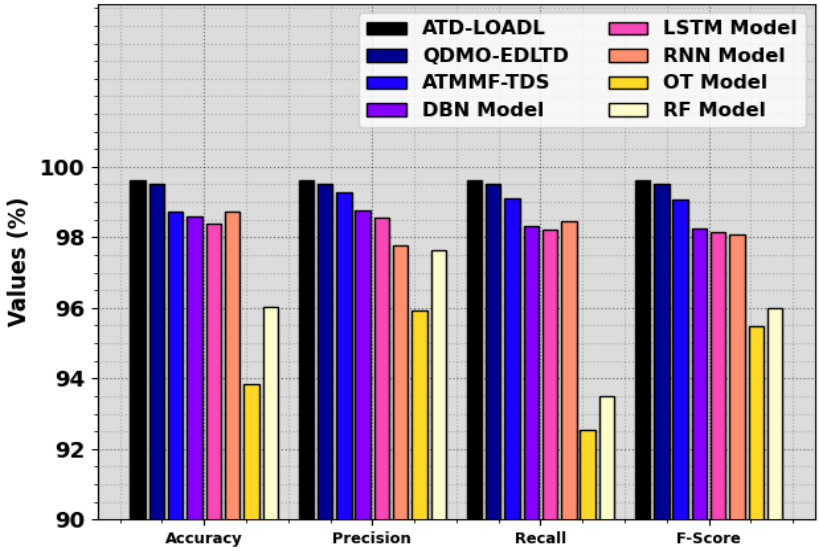

Figure 10. Comparison analysis of the ATD-LOADL model with other techniques

The computation complexity of the ATD-LOADL algorithm with recent models is measured concerning training time (TRT) and testing time (TST) in Table 5 and Fig. 11. The accomplished findings suggest that the AIMMF-IDS approach, DBN system, LSTM algorithm, RNN model, DT approach, and RF methodology demonstrate boosted and closer TRT and TST values. Meanwhile, the QDMO-EDLID model exhibits reasonable TRT and TST values of 0.80s and 0.54s. However, the ATD-LOADL technique demonstrates maximum performance with decreased TRT and TST of 0.31s and 0.19s, correspondingly. Thus, the simulation analysis ensured that the ATD-LOADL technique provides enhanced threat detection results.

Table 5: Computation complexity of the ATD-LOADL system with other methods

Methods	TRT (sec)	TST (sec)
ATD-LOADL	0.31	0.19
QDMO-EDLID	0.80	0.54
AIMMF-IDS	1.10	0.57
DBN	1.21	0.74

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

LSTM	1.31	0.72
RNN	1.31	0.75
DT	1.59	0.95
RF	1.60	1.00

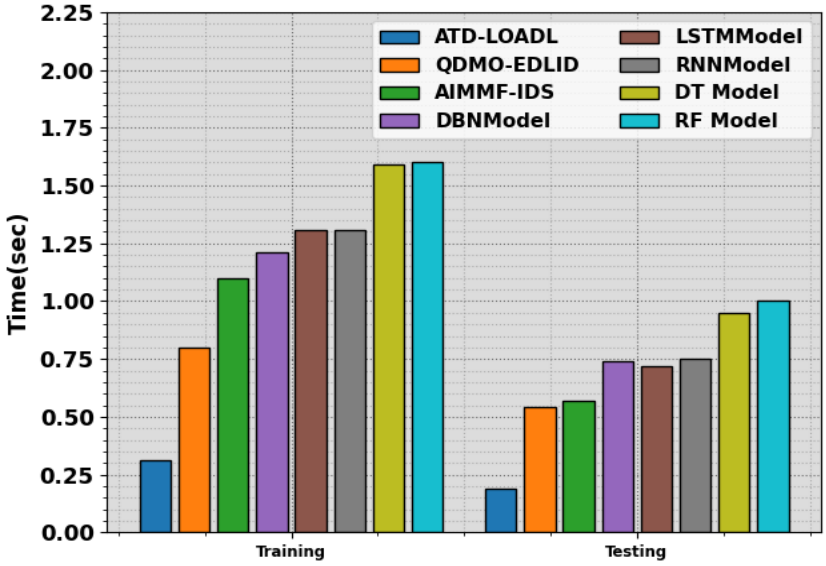

Figure 11. Computation complexity of the ATD-LOADL model with other methods

5. Conclusion

In this manuscript, we offer the design of an ATD-LOADL model in CPS. The main purpose of the ATD-LOADL methodology is to concentrate on the recognition and classification of cyber threats in the CPS. Initially, the pre-processing of the CPS data takes place using a min-max scaler. To select an optimum set of features, the ATD-LOADL technique uses LOA as a feature selection approach. For threat detection, the ATD-LOADL technique uses the MHA-LSTM classifier. At last, the detection outcomes of the MHA-LSTM methodology can be boosted by the use of SFLA. The experimentation outcomes of the ATD-LOADL method can be widely investigated on a benchmark CPS dataset. The experimentation outcomes stated the enhanced threat detection results of the ATD-LOADL technique over other existing approaches.

Funding: "The authors extend their appreciation to the Arab Open University for funding this work through AOU research fund No. (AOUKSA524008)"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] H. Liu, S. Wang, and Y. Li, ''Event-triggered control and proactive defense for cyber–physical systems,'' IEEE Trans. Syst., Man, Cybern. Syst., vol. 52, no. 10, pp. 6305–6313, Oct. 2022.

[2] S. Yadav and R. Kalpana, ''A survey on network intrusion detection using deep generative networks for cyber-physical systems,'' in Artificial Intelligence Paradigms for Smart Cyber-Physical Systems. Hershey, PA, USA: IGI Global, 2021, pp. 137–159.

[3] M. A. Alohali, F. N. Al-Wesabi, A. M. Hilal, S. Goel, D. Gupta, and A. Khanna, ''Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment,'' Cognit. Neurodynamics, vol. 16, pp. 1–13, Oct. 2022.

[4] T. Zoppi, M. Gharib, M. Atif, and A. Bondavalli, ''Meta-learning to improve unsupervised intrusion detection in cyber-physical systems,'' ACM Trans. Cyber-Phys. Syst., vol. 5, no. 4, pp. 1–27, 2021.

[5] D. Levonevskiy and A. Motienko, ''Modeling tasks of patient assistance and emergency management in medical cyber-physical systems,'' in Proc. Comput. Methods Syst. Softw. Cham, Switzerland: Springer, 2023, pp. 299–308.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024

[6] P. F. de Araujo-Filho, G. Kaddoum, D. R. Campelo, A. G. Santos, D. Macêdo, and C. Zanchettin, ‘‘Intrusion detection for cyber–physical systems using generative adversarial networks in fog environment,’’ IEEE Internet Things J., vol. 8, no. 8, pp. 6247–6256, Apr. 2021.

[7] R. F. Mansour, ‘‘Artificial intelligence based optimization with deep learning model for blockchain enabled intrusion detection in CPS environment,’’ Sci. Rep., vol. 12, no. 1, p. 12937, Jul. 2022.

[8] A. E. Ibor, O. B. Okunoye, F. A. Oladeji, and K. A. Abdulsalam, ‘‘Novel hybrid model for intrusion prediction on cyber physical systems’ communication networks based on bio-inspired deep neural network structure,’’ J. Inf. Secur. Appl., vol. 65, Mar. 2022, Art. no. 103107.

[9] M. Sharma, H. Elmiligi, and F. Gebali, ‘‘A novel intrusion detection system for RPL-based cyber–physical systems,’’ IEEE Can. J. Electr. Comput. Eng., vol. 44, no. 2, pp. 246–252, Spring 2021.

[10] J. Ali, ‘‘Intrusion detection systems trends to counteract growing cyberattacks on cyber-physical systems,’’ in Proc. 22nd Int. Arab Conf. Inf. Technol. (ACIT), Dec. 2021, pp. 1–6.

[11] Sharma, A., Rani, S., Shah, S.H., Sharma, R., Yu, F. and Hassan, M.M., 2023. An Efficient Hybrid Deep Learning Model for Denial of Service Detection in Cyber Physical Systems. IEEE Transactions on Network Science and Engineering.

[12] Althobaiti, M.M., Kumar, K.P.M., Gupta, D., Kumar, S. and Mansour, R.F., 2021. An intelligent cognitive computing based intrusion detection for industrial cyber-physical systems. Measurement, 186, p.110145.

[13] Ashraf, I., Narra, M., Umer, M., Majeed, R., Sadiq, S., Javaid, F. and Rasool, N., 2022. A deep learning-based smart framework for cyber-physical and satellite system security threats detection. Electronics, 11(4), p.667.

[14] Li, Q., 2023. Threat Detection and Diagnosis in Cyber-Physical Systems with Artificial Intelligence (Doctoral dissertation, University of Georgia).

[15] Deng, W., Yang, C. and Huang, K., 2022, July. VFD-AE: Efficient Attack Detection in Industrial Cyber-Physical Systems using Vital Feature Discovery and Deep Learning Technique. In 2022 41st Chinese Control Conference (CCC) (pp. 7479-7484). IEEE.

[16] Jahromi, A.N., Karimipour, H., Dehghantanha, A. and Choo, K.K.R., 2021. Toward detection and attribution of cyber-attacks in IoT-enabled cyber–physical systems. IEEE Internet of Things Journal, 8(17), pp.13712-13722.

[17] Zhou, X., Almutairi, L., Alsenani, T.R. and Ahmad, M.N., 2023. Honeypot Based Industrial Threat Detection Using Game Theory in Cyber-Physical System. Journal of Grid Computing, 21(4), p.59.

[18] Duhayyim, M.A., Alissa, K.A., Alrayes, F.S., Alotaibi, S.S., Tag El Din, E.M., Abdelmageed, A.A., Yaseen, I. and Motwakel, A., 2022. Evolutionary-based deep stacked autoencoder for intrusion detection in a cloud-based cyber-physical system. Applied Sciences, 12(14), p.6875.

[19] Ozsahin, D.U., Mustapha, M.T., Mubarak, A.S., Ameen, Z.S. and Uzun, B., 2022, August. Impact of feature scaling on machine learning models for the diagnosis of diabetes. In 2022 International Conference on Artificial Intelligence in Everything (AIE) (pp. 87-94). IEEE.

[20] Ra’ed, M., Al-qudah, N.E.A., Jawarneh, M.S. and Al-Khateeb, A., 2023. A novel improved lemurs optimization algorithm for feature selection problems. Journal of King Saud University-Computer and Information Sciences, 35(8), p.101704.

[21] Saleh, H., Amer, E., Abuhmed, T., Ali, A., Al-Fuqaha, A. and El-Sappagh, S., 2023. Computer aided progression detection model based on optimized deep LSTM ensemble model and the fusion of multivariate time series data. Scientific Reports, 13(1), p.16336.

[22] Pirgazi, J., Alimoradi, M., Esmaeili Abharian, T. and Olyaee, M.H., 2019. An Efficient hybrid filter-wrapper metaheuristic-based gene selection method for high dimensional datasets. Scientific reports, 9(1), p.18580.

[23] Almutairi, L., Daniel, R., Khasimbee, S., Lydia, E.L., Acharya, S. and Kim, H., 2023. Quantum Dwarf Mongoose Optimization with Ensemble Deep Learning Based Intrusion Detection in Cyber-Physical Systems. IEEE Access.

DOI: https://doi.org/10.54216/JCIM.150208
Received: May 04, 2024 Revised: July 08, 2024 Accepted: October 21, 2024