A Novel Approach of BRELU RESNET Based Cyber Attack Detection System with BAIT Based Approach for Mitigation

Sangeetha Prabhu [1] & Nethravathi P. S. [2]

[1] Research Scholar, College of Computer Science and Information Science, Srinivas University, Mangalore, India.

ORCID: 0000-0002-8026-1133; E-mail: sangeethaprabhu96@gmail.com

[2] Professor, College of Computer and Information Sciences, Srinivas University, Mangalore, India.

ORCID: 0000-0001-5447-8673; Email: nethrakumar590@gmail.com

Subject Area: Computer Science.
Type of the Paper: Research analysis.
Type of Review: Peer Reviewed as per |C|O|P|E| guidance.
Indexed In: OpenAIRE.
DOI: https://doi.org/10.5281/zenodo.6530129
Google Scholar Citation: IJAEML

How to Cite this Paper:
Prabhu, Sangeetha, & Nethravathi, P. S., (2022). A Novel Approach of BRELU RESNET Based Cyber Attack Detection System with BAIT Based Approach for Mitigation. *International Journal of Applied Engineering and Management Letters (IJAEML)*, *6*(1), 243-258. DOI: https://doi.org/10.5281/zenodo.6530129

International Journal of Applied Engineering and Management Letters (IJAEML)
A Refereed International Journal of Srinivas University, India.

Crossref DOI : https://doi.org/10.47992/IJAEML.2581.7000.0134

Received on: 10/04/2022
Published on: 10/05/2022

Disclaimer: The scholarly papers as reviewed and published by the Srinivas Publications (S.P.), India are the views and opinions of their respective authors and are not the views or opinions of the S.P. The S.P. disclaims of any harm or loss caused due to the published content to any party.

A Novel Approach of BRELU RESNET Based Cyber Attack Detection System with BAIT Based Approach for Mitigation

Sangeetha Prabhu [1] & Nethravathi P. S. [2]

[1] Research Scholar, College of Computer Science and Information Science, Srinivas University, Mangalore, India.

ORCID: 0000-0002-8026-1133; E-mail: sangeethaprabhu96@gmail.com

[2] Professor, College of Computer and Information Sciences, Srinivas University, Mangalore, India.

ORCID: 0000-0001-5447-8673; Email: nethrakumar590@gmail.com

ABSTRACT

Purpose: *Industrial Control Systems become more vulnerable to digital attacks by merging communication groups and the Internet of Things, which could have severe implications. An Intrusion Detection System is essential in IoT businesses for identifying and stopping assaults. To ensure data privacy and security in the face of digital attacks, legislation and large enterprises should develop network security policies today. As people-based full frameworks have become more vital in today's society, they've also become targets for hostile activities, compelling both industry and research to concentrate more on dealing with local area disruption recognition issues. Contraption reviewing techniques have shown to be effective tools for resolving in-network interruption location issues.*

Design/Methodology/Approach: *This investigation yielded a very unique strategy for tackling hub moderation utilizing a Classification and Encryption method. The UNSW-NB15 dataset is acquired and divided into Data for preparation and testing from the start. The information is pre-handled and included are eliminated right away within the preparation time frame. The TWM Algorithm is then used to determine the relevant highlights from that moment onward. The BRELU-RESNET classifier then sorts the input into went after and non-went after categories. The compromised information is then saved in the security log record, and the typical data is encrypted using the ESHP-ECC computation. The shortest path distance is then calculated using Euclidean distance. Finally, the data is available. Finally, using the DSHP-ECC computation, the information is decrypted. If the information is available in the log document during testing, it is regarded as the sought-after data and is prevented from the transmission. If it is not present, then the process of digital assault recognition begins.*

Findings/Result: *The research is based on the UNSW-NB 15 dataset, which shows that the proposed method achieves an unreasonable awareness level of 98.34 percent, particularity level of 77.54 percent, exactness level of 96.6 percent, Precision level of 97.96 percent, review level of 98.34 percent, F-proportion of 98.15 percent, False Positive Rate of 22.46 percent, False Negative Rate of 1.66 percent, and Matthew's connection coefficient of 77.38*

Originality/Value: *This experimental-based research article examines the malicious activities in the cyberspace using BRELU-RESNET approach and mitigated by using BAIT based approach mechanism.*

Paper Type: *Research Analysis.*

Keywords: Cyber-attack detection, BAIT approaches, Feature Extraction, BRELU-RESNET, Attack node mitigation

1. INTRODUCTION :

Electrical and mechanical equipment, computers, and human-supervised manual processes make up industrial control systems (ICSs) [1]. They are mostly employed in industrial facilities and vital infrastructures, such as manufacturing sectors, chemical plants, electricity production, distribution networks, and others [2, 3]. Their activities have direct consequences on the environment, human safety and health, the economy, and national security [4]. In ICS, because data is often delivered across a wired or wireless network, it is extremely vulnerable to being tampered with by malevolent attackers. [5]. Furthermore, the incorporation of the IoT in ICSs allows hackers to exploit system weaknesses to conduct cyber-attacks [6, 7]. Typical cyber-attacks against ICSs include DoS, Man in the middle attack, SQL injections, Password attacks, Phishing, and so on [8]. By influencing and disturbing the physical process, cyber-attacks against ICSs are primarily aimed at causing damage or catastrophe (such as a hazardous accident or output loss) [9]. Attack detection systems are meant to avoid such assaults by effectively monitoring events in an information system and identifying signals of security vulnerabilities [10]. The most widely used strategy for attack detection is anomaly detection, which is the act of discovering abnormal occurrences or unexpected system behavior [11-12]. However, the majority of these algorithms have only been trained on particular sorts of assaults and are unable to identify unknown or novel attacks [13]**.**

Hence, to overcome these challenges and risks faced during attack detection, various anomaly detection algorithms are introduced. These approaches are integrated and implemented by using a variety of Machine Learning algorithms [14]. However, the majority of available algorithms ignore the unbalanced structure of ICS datasets, In real-world contexts, this leads to low detection rates or high false-positive rates. [15]. If the entire physical system was attacked at the same time, several of the present approaches would be ineffective [16]. Much research on fault-tolerant control has been conducted, and these studies can also give tools for attack-resilient control [17]. When it comes to evading surveillance and isolation, there are a few things to keep in mind, however, there are significant variations between fault-tolerant controls and attack-resistant which necessitates the use of particular approaches to solve security challenges in ICSs [18]. Thus, to offset the aforementioned problem, the work has proposed a framework called a novel approach of BRELU-ResNet based Cyber-Attack Detection System with BAIT based approach for mitigation, which guarantees the accurate detection of cyber-attacks and retains more authenticity of the network.

2. LITERATURE REVIEW :

According to the defense-in-depth strategy, Fan Zhang et al. [19] created a cyber-attack detection system that utilized data from the network, data from the host system, and data from the measured process characteristics. A multi-layer attack detection system was available saving the defenders valuable time before the physical system's unrecoverable repercussions occurred. The classic intrusion prevention layer, which included firewalls, data diodes, and gateways, was the initial protection layer. The second tier of securities ty comprised of data-driven algorithms for detecting cyber-attacks using system data and network traffic. M1 represented the classification model, while M2 denoted the big data analytics models. When attacks induced behavior divergence from normal functioning, early identification of intruders was possible with M1 and M2. If the second security measure fails to match suspicious attacks, the final security checkpoint examines the processed data and uses M3's modeling techniques to discover anomalous processes that could be triggered by a cyber-attack. The technique discovered physically damaging cyber-attacks before they had a substantial impact, according to the findings. The approach, however, was unsuccessful against modern cyber-attacks.

Abdulrahman Al-Abassiet al. [20] proposed a deep learning approach for creating balanced representations of unbalanced datasets. The representations were input into a deep learning ensemble model for detecting attacks that were created particularly for an ICS setting. Multiple unsupervised SAEs were used in the novel representations using a deep learning model from unbalanced datasets. Multiple Auto-Encoders (AE) were used in the SAE attack detection model to extract new representations from unlabelled input, resulting in various patterns. New representation from every SAE was then fed to a DNN and fused with the use of a fusion activation vector utilizing a super vector. Finally, a DT was used to distinguish models that have recently been combined and are launching attacks as a binary classifier. The strategy beat previous models that have been published in the literature, according to the findings. However, the approach was limited by a lack of backup capabilities.

Moshe Kravchiket al. [21] used 1 Dimensional Convolution Neural Networks (1D CNNs), which are classified as complete autoencoders), variation autoencoders (VAEs), and PCA, to create a technique to identify anomalies and cyber-attacks at the physical level Industrial control systems (ICS) data. In addition, the Kolmogorov-Smirnov test was used to identify features, and the frequency components of the time-domain signals were converted and represented using a short-time Fourier transform and energy binning, and the system was modeled both in terms of duration and frequency. The approach was tested on three widely used public datasets, and the findings demonstrated that it was resistant to such evasion assaults. The attacker was obliged to forego the desired physical to prevent influencing the system discovery. However, the concept was limited by its high energy use.

Nevius Kaja et al. [22] developed a two-stage intelligent Detection Mechanism that could detect and prevent such attacks. The solution included a two-stage design based on ML methods. The work's originality was the employment of a two-stage machine learning approach in the building of IDS after four-step efficient information pre-processing. The IDS employed K-Means to identify the attacks in the initial stages, then supervised learning during the second phase to classify the assaults and reduce the number of false positives. The approach's implementation resulted in computationally efficient IDS capable of detecting and classifying assaults with increased precision while lowering the number of erroneous positives. The IDS outperformed the present in terms of technology and state-of-the-art performance. The method, however, exhibited poor detection rates and a significant percentage of false positives.

To detect intrusions, Kaiyuan Jiang et al. [23] offered a hybrid sampling and deep hierarchical network technique for network intrusion detection. To decrease noise trends in the general category, the One-Side Selection (OSS) method was used, followed by the Synthetic Minority Oversampling Technique to improve the minority samples (SMOTE). As a consequence, a balanced dataset was constructed, allowing the system to completely comprehend the characteristics of extracted features while reducing model time training. Second, to obtain spatial characteristics, a Convolution Neural Network (CNN) was employed, and a Bi-directional long short-term memory (Bi-LSTM) was used to retrieve temporal features, leading to a deep hierarchical network model. The NSL-KDD and UNSW-NB15 datasets were used to validate the findings. The method, on the other hand, had a high packet loss limit and was poor at controlling network overhead. However, the concept was limited by its high energy use.

3. OBJECTIVES :

(1) To provide the deep ensemble technique for detecting the presence of a network assault.
(2) To create a model of the SHP-ECC procedure for removing the attacker from the network.
(3) To compare the proposed system's practicality with other state-of-the-art frameworks in terms of specific performance criteria.

4. METHODOLOGY :

4.1 Proposed Model for Cyber-Attack Detection and Mitigation System :

The increasing incidence of cyber-physical system (CPS) assaults in recent years has raised concerns about industrial control system cybersecurity (ICS). ICS cybersecurity attempts now depend very much on firewalls, information valves, and other intrusion detection and prevention systems, which may be insufficient to withstand rising cyber threats from persistent attackers. With the use of a deep learning method, earlier research has built a framework for identifying assaults. Although the attack node in the network was discovered, it was not disabled. As a solution to this challenge, an upgraded and formidable adversary model will be presented. As a result, employing a unique Classification and Encryption approach, a novel architecture for attack node mitigation is provided in this study. The input information is usually separated into two groups: training data and testing data. The total training data is initially pre-processed. The second stage is to extract features from the input training dataset. In the third step, the feature is optimized for selecting the important Features using TWMA. Then, the feature is trained using the proposed BRELU-RESNET Classifier. Here, the classifier classifies the data into the attack and normal data. If the data is attack data, save the Source IP Address into a secure log file using the BAIT approach. Next, if the data is normal, the data is ready for transmission. In Data Transmission, first, the data is encrypted using the ESHP-ECC algorithm. Next, the shortest path distance is calculated using Euclidean distance. In Destination, the data is decrypted using the DSHP-ECC algorithm. In testing, first, the testing data is checked in the Security Log File (SLF). If the data's originating IP address is already known, the data is stopped or an attack is detected. Figure 1 depicts a block schematic of the suggested framework.

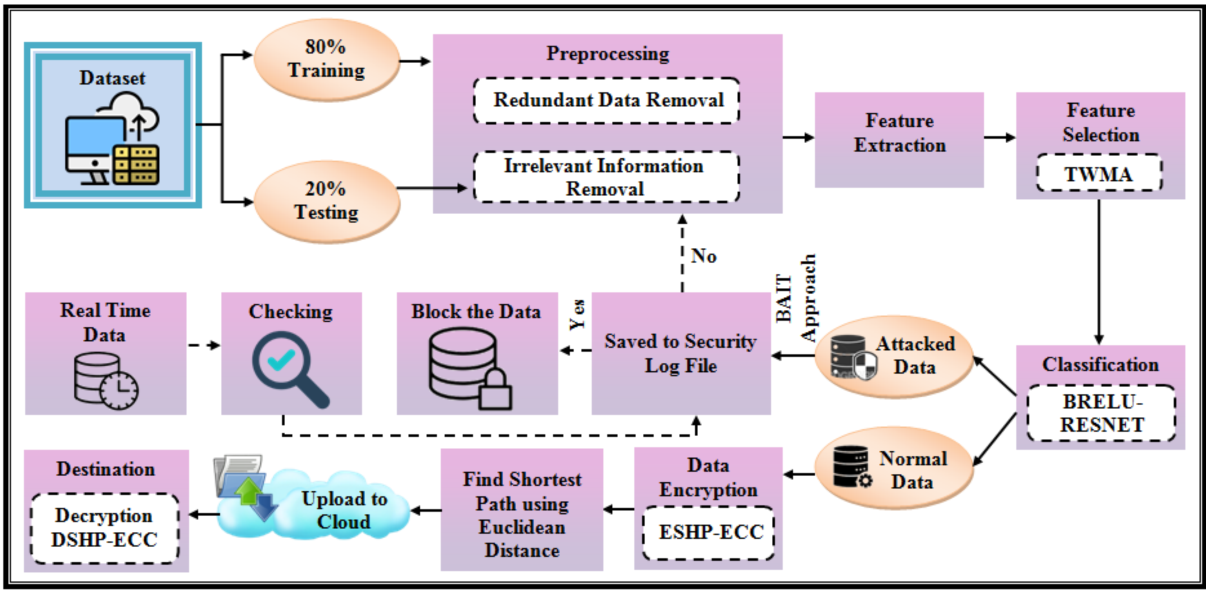

Fig. 1: The suggested cyber-attack detection and mitigation system's formwork [24]

4.2 Pre-processing :

The input dataset is separated to begin incorporating testing data and training data in the process of the cyber-attack detection technique. The data is initially in the training stage to turn the original data into data cleansing. Pre-processing is done to improve the classification's accuracy and shorten the cyberattack detection system's training period. The proposed work uses redundant data removal and irrelevant information removal as pre-processing steps.

- Redundant data means storing of same data in multiple locations. Redundant data removal is a technique to remove duplicate data from the dataset. This reduces the computational complexity and results in better generalization for the classifier [25].
- Irrelevant information removal is the process of removing the data that are not required for the detection of cyberattacks. The presence of such unrelated information may increase the processing time of the system and may result in an inaccurate attack detection rate. As a result, the system's efficiency is enhanced by removing the different data from the input dataset. The pre-processed data is then used to extract features.

4.3 Feature Extraction :

The method of extracting the number of features in a dataset by producing new features from existing ones is known as feature extraction [26]. The majority of the information from the initial set of features is contained in these additional features. Source IP address, destination port number, port number, a destination address, source bits per second, destination bits per second, transaction protocol, and so on are all extracted from the pre-processed dataset. The set of extracted features $x_{(i)}$ is mathematically articulated below,

$$x_{(i)} = x_{(1)}, x_{(2)}, \ldots\ldots x_{(n)} \qquad ---------- \qquad (1)$$

Where, n exemplifies the number of extracted features.

4.4 Feature selection by TWMA :

After feature extraction, the important features are selected using the novel Taxicab Woodpecker Mating Optimization (TWMA) technique. Woodpecker Mating Optimization (WMA) is a nature-inspired optimization algorithm developed by the mating behaviour of red-bellied woodpeckers. At the beginning of the mating period, male woodpeckers communicate with females by making a sound by pecking the trunks of the trees called drumming [27]. Depending on the quality of sound produced by the male woodpeckers, the female woodpeckers are attracted to them. Hence, the sound intensity of the male woodpeckers' drum mentions its capability to attract more female birds. By hearing the sound, the female woodpeckers move towards the male birds and the process of communication and flow of information between them occurs. On the other hand, if the sound intensity of some other male is closer to the female, then the female bird will attract this male and move towards this male. However, the problem with WMA is slow or premature convergence due to the loss of diversity within the population. To overcome this issue, Taxicab geometry is used to update the female woodpecker's position during movement. The usage of Taxicab geometry can effectively avoid the woodpecker population falling into local optimum and also eliminates the slow or premature convergence. This enhancement made in general WMA is called TWMA. The process of TWMA is detailed as follows.

Step 1: To begin the process of WMA, first, the woodpecker population (extracted features) is initialized as,

$$x_{(i)} = x_{(1)}, x_{(2)}, \ldots\ldots x_{(n)} \quad \text{--------} \quad (1)$$

Where shows the woodpecker population and n is the number of woodpeckers in the population.

Step 2: After population initialization, the fitness of each woodpecker is calculated to identify the best woodpecker. Afterward, the woodpecker population is separated into male and female groups. The male birds become the search agents and the one with the highest fitness is considered as x^* (the global best solution). Here, the fitness is computed in terms of classification accuracy. The fitness evaluation $f\left(x_{(i)}\right)$ is,

$$f\left(x_{(i)}\right) = f\left(x_{(1)}, x_{(2)}, \ldots\ldots, x_{(n)}\right) \quad \text{-----} \ (2)$$

Step 3: Next, the sound intensity of the woodpecker is calculated using the below equation,

$$\delta = \frac{2\pi^2 \gamma^2 \mathrm{A}^2 DS}{\Psi} \quad \text{------} \quad (3)$$

In equation (3), δ represents the sound intensity, γ indicates the sound frequency, A specifies the sound amplitude, D signifies the density of the medium in which sound is traveling, S illustrates the sound speed and ψ mentions the area of sound.

Step 4: The sound intensity of the woodpecker may change based on the source of the sound. Some sources may emit the sound in one direction. Consider a sphere in the region of a source with a radius ι, and then all the sound waves will pass through the surface of the sphere. Thus, the sound intensity (δ) in equation (3) becomes,

$$\delta = \frac{2\pi^2 \gamma^2 \chi DS}{\psi \cdot 4\pi\iota^2} \quad \text{--------} \quad (4)$$

Here, χ defines the propagation rate of sound waves and $4\pi\iota^2$ determines the area of the sphere. In equation (4), sound intensity depends on the distance between source and object.

Step 5: The shortest distance between source and object mentions the better sound quality received by the female woodpecker. The Taxicab distance is used to calculate the distance between the source and object, which overcomes the problem of premature convergence and obtains the global best solution. It is given below,

$$\iota = \sqrt{\left(x_m - y_f\right)} \quad \text{----------} \quad (5)$$

The above equation, x_m shows the sound source position (male woodpecker) and y_f points to the listener position (female woodpecker).

Step 6: Thereafter, the female woodpecker updates its position concerning the sound intensity of the male bird. The position updation process $\left(y_{f,j}^{\tau+1}\right)$ is mathematically modeled below,

$$y_{f,j}^{\tau+1} = y_{f,j}^{\tau} + r * \frac{\alpha_{f,j}^{\tau} \left\langle \beta^{x^*}\left(y_{x^*}^{\tau} - y_{f,j}^{\tau}\right) + \beta_{m,i}\left(x_{m,i}^{\tau} - y_{f,j}^{\tau}\right)\right\rangle}{2} \quad \text{-------------} \quad (6)$$

Where, $j = 1,2,\ldots,m$ shows the female woodpecker population, $y_{f,j}^{\tau}$ indicates the current position of $j-th$ woodpecker in τ^{th} iteration, $y_{x^*}^{\tau}$ represents the position of best woodpecker, $x_{m,i}^{\tau}$ is the position of $i-th$ male woodpecker, r is a random number with a uniform distribution in the range

[0, 1], $\alpha^{\tau}_{f,j}$ is a self-tuned random factor of $j-th$ woodpecker, β^{x^*} and $\beta_{m,i}$ are the attractiveness of the female bird to the male bird.

Step 7: The self-tuning random factor $\alpha^{\tau}_{f,j}$ is estimated using equation (7),

$$\alpha^{\tau}_{f,j} = r * \eta \qquad \text{----------} \quad (7)$$

$$\eta = ts\left(1 - \frac{\tau}{\tau^{\max}}\right) \qquad \text{----------} \quad (8)$$

This equation, ts defines the tangent sigmoid function, $\tau, \tau^{\max}$ models the current and maximum number of iterations respectively, α has a random value in the interval -2η to 2η. If $|\alpha| > 1$, the search agent deviates from the target, which leads to exploration, and if $|\alpha| < 1$, then the female bird joins with the male bird, which leads to exploitation.

Step 8: The attractiveness (β) of male and female woodpeckers is then calculated as,

$$\beta = (1 + \delta(i, j))^{-1} \qquad \text{-----------} \quad (9)$$

In (9), $\delta(i, j)$ depicts the sound intensity of $i-th$ male woodpecker heard by the female woodpecker. It is also called the step size of a female woodpecker because it specifies the closeness of the female woodpecker towards the male, β lies in the interval 0 and 1, the lower value defines the accurate movement of the female toward the male woodpecker.

Step 9: At each cycle, the male woodpecker population decreases, and finally, only one woodpecker will remain. A large male population increases the exploration in the initial phase. Hence, the decreasing population increases the exploitation and accuracy of the solution. The population size $(x_{m,i})$ in each iteration is computed as follows,

$$x_{m,i} = \left[round\left(\frac{n}{2} * \left(1 - \frac{\tau}{\tau^{\max}}\right)\right) + 1\right] \qquad \text{-------------} \quad (10)$$

Where, n mentions the total woodpecker population, $\tau, \tau^{\max}$ models the current and maximum number of iterations respectively.

Step 10: Finally, the decreased population of woodpeckers contains one woodpecker and the global best woodpecker x^*. Thus, equation (6) can be modified as,

$$y^{\tau+1}_{f,j} = y^{\tau}_{f,j} + r * \left\langle \alpha^{\tau}_{f,j} \cdot \left(y^{\tau}_{x^*} - y^{\tau}_{f,j}\right) \cdot \beta_{m,i} \right\rangle \qquad \text{-------------} \quad (11)$$

Step 11: During the movement of the female woodpecker towards the male, there is a possibility of deviation in direction; as well female birds might be attacked by other woodpeckers or hunting birds on the way. Thus, the female bird may change their path randomly to protect itself from danger. This random change in the pathway is called Run Away. This random escaping movement of the woodpecker consists of two types of movements, which are based on the sound intensity of x^* the male bird. The two types of movements (μ) are,

$$\mu = \begin{cases} R & \beta \geq \xi \\ \mathrm{P} & else \end{cases} \qquad \text{---------------} \quad (12)$$

Where, R, P specifies the runaway movement and x^* runaway movement respectively.

$$\xi = 0.8 \cdot \frac{\sum_{j=1}^{m-1} \beta^{j}_{x^*}}{m-1} \qquad \text{--------------} \quad (13)$$

Here, ξ mentions the threshold for the sound intensity.

Step 12: The position of the female woodpecker obtained from the runway is updated below,

$$\tilde{y}_{f,j} = L - (L - U) * r \qquad \text{---------} \quad (14)$$

In equation (14), $\tilde{y}_{f,j}$ is the position of $j-th$ the woodpecker after the runaway, r is a random number in the uniform distribution [0, 1] and L,U illustrates the upper and lower bounds of variables in that order.

Step 13: The x^* runaway movement is denoted further,

$$\mathrm{P} = \phi * \left(1 - \frac{\tau}{\tau^{\max}}\right) \qquad \text{------------} \quad (15)$$

In equation (15), ϕ is the runaway coefficient. The position of a female woodpecker from x^* runaway movement $\left(y^{x^*}_{f,j}\right)$ is modelled by,

$$y^{x^*}_{f,j} = y^{\tau}_{f,j} + \mathrm{P}^{bit}\left\langle y^{\tau}_{x^*} - y_r \right\rangle \cdot \mathrm{B} \qquad \text{-------------} \quad (16)$$

$$\mathrm{P}^{bit} = \begin{cases} 1 & r \le \mathrm{P} \\ 0 & else \end{cases} \qquad \text{-------------} \quad (17)$$

The above equation, r exemplifies a random number in the uniform distribution [0, 1] and B is a random number [-1, 1]. The process continues until the stopping criterion is met by comparing the position of $i-th$ the woodpecker with the former position and the position of the best woodpecker. Then, the better position is replaced with the other position. Finally, the optimal solution is obtained, that is, the selected best features $\left(X^{(k)}\right)$ articulated further,

$$X^{(k)} = X^{(1)}, X^{(2)}, \ldots\ldots, X^{(K)} \qquad \text{-------------} \quad (18)$$

Where, K shows the percentage of traits that have been chosen for further study classification. The pseudocode of the proposed TWMA technique is publicized below.

Pseudocode for Proposed TWMA Technique:

```
Input: Extracted Features $x_{(i)}$
Output: Selected features $\left(X^{(k)}\right)$
Begin
Create the initial population of woodpeckers
Compute $f\left(x_{(i)}\right) = f\left(x_{(1)}, x_{(2)}, \ldots\ldots, x_{(n)}\right)$
Obtain $x^*$ based on $f\left(x_{(i)}\right)$
While (stopping condition is not satisfied) do
Partition $x_{(i)}$
For $1 \le i \le n$
Determine the sound intensity $\delta$
Compute Taxicab distance
Choose $x_{m,i}$ ($i-th$ male woodpecker)
Evaluate $\beta^{x^*}$ and $\beta_{m,i}$
Analyze $\alpha^{\tau}_{f,j} = r * \eta$
Update woodpeckers' position $\left(y^{\tau+1}_{f,j}\right)$
Calculate sound intensity threshold $\xi$
If $\beta^{x^*,i} > \xi$
Estimate $\tilde{y}_{f,j} = L - (L - U) * r$
Else
Find out $x^*$ runaway movement $\left(y^{x^*}_{f,j}\right)$
End if
Appraise the new position of $y_{f,j}$
```

```
Renew $x^*$
End for
$\tau = \tau + 1$
End while
Obtain global best solution $\left(X^{(k)}\right)$
End
```

4.5 Classification by means of BRELU-RESNET classifier :

Next, the selected features $\left(X^{(k)}\right)$ are fed into the BRELU-RESNET classifier to classify the attacked data from non-attacked data. ResNet is used for classification because it overcomes the problem of degradation caused due to rise in network depth [28]. Convolutional layers, batch normalization layers, max pooling, flattening layer, and activation layers are all included in ResNet's design. The selected features are inputted to the ResNet, firstly, the input is convoluted with the 2*2 filter in the convolutional layer, and it produces the output with reduced feature dimension. The outcome from the convolutional layer is then sent into the batch normalization layer, which stabilizes the network's training time while decreasing the number of timestamps. Convolutional and batch normalization are separated into three levels. The data is then sent to the max-pooling layer, which downsamples the data [29]. Finally, the fully connected layer in the network comprises of average pooling layer and softmax layer to classify the outputs. However, RESNET has an overfitting problem due to the randomized nature of the activation function. Therefore, in the proposed work, instead of random value, Bernoulli's value is used in the Leaky Rectified Linear Unit activation function in the RESNET classifier. This modification in baseline RESNET is termed as BRELU-RESNET. The general structure of the ReSNet network is cited in figure 2.

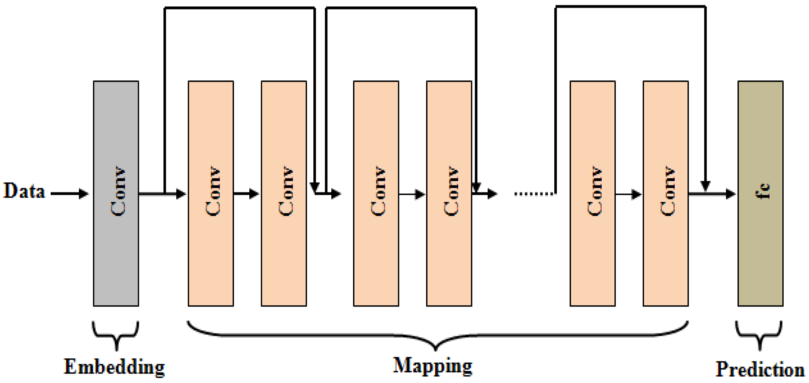

Fig. 2: RESNET architecture [24]

- Let $\left(X^{(k)}\right)$ be the input features and a filter (Γ) of size (a,b) is used in the convolution layer. The convolution formula can be as seen in the equation below,

$$conv\left(X^{(k)} * \Gamma\right) = \sum_{k=1}^{K} \left(X^{(k)} - a, X^{(k)} - b\right) \cdot \Gamma(a,b) \quad \text{--------------} \quad (19)$$

- The activation function is an important part of neural networks. The BRELU activation function $\left(f\left(X^{(k)}\right)\right)$ used in the proposed system is expressed as,

$$f\left(X^{(k)}\right) = \max\left(0, b\left(X^{(k)}\right)\right) \quad \text{------------} \quad (20)$$

Where, $b\left(X^{(k)}\right)$ is known as Bernoulli's distribution function, defined by

$$b\left(X^{(k)}(p,o)\right) = p \cdot o + (1-p)(1-o) \quad \text{----------} \quad (21)$$

The above equation, p, o differentiates the probability and possible outcome of $\left(X^{(k)}\right)$.

- The network layers in the BERLU-RESNET are capable of approximating any function asymptotically. The approximation of residual function $\partial X^{(k)}$ is,

$$\partial X^{(k)} = f\left(X^{(k)}\right) * X^{(k)} \quad \text{-------------} \quad (22)$$

Here, $f\left(X^{(k)}\right)$ is the target function which is formulated further,

$$f\left(X^{(k)}\right)=\partial X^{(k)}+X^{(k)} \quad \text{------------} \quad (23)$$

Hence, the output of the classifier separates the attacked data from the normal data and then the attacked data is stored in the security log file using the BAIT approach. While the normal data is encrypted using the Encrypted Secure Hash Probability-based Elliptic-curve cryptography (ESHP-ECC) algorithm [30]. Next, the shortest path between each node is calculated for transferring the data in the cloud efficiently.

4.6 Attack Mitigation System Based on BAIT Approach :

The malicious node is duped into sending the erroneous route request to the decoy route request using the BAIT mechanism. The goal of this method is to both identify and mitigate attack nodes. Malicious nodes, in general, market themselves as the most efficient and quickest path to their targets. In addition, the rogue node sends the source a route request packet, which is illegal. As a result, the source node in the proposed task sends the fake request to the destination address via the nearest node. When a malicious node receives a request, it answers even if it is not the target node. This request is fed into the flow table search. After then, the attack is identified and mitigated. When a malicious node sends a response, the source node compares it to the destination address. The source node deems the node malicious and denies the response request if the addresses do not match. The data flow will continue to flow normally if no attacker nodes are found. When a route request packet arrives at a non-attack intermediate node, it is transmitted to the destination node along with the address. You may feel assured that you will transfer the data once the shortest route between the source and the destination has been determined. The test data is compared to the security log file first during the test. The data will be prohibited for further processing if the source IP address of the information is already in the security log file. The cyber-attack detection and prevention process is carried out if it does not exist in the security log file. Figure 3 depicts the BAIT approach's overall structure.

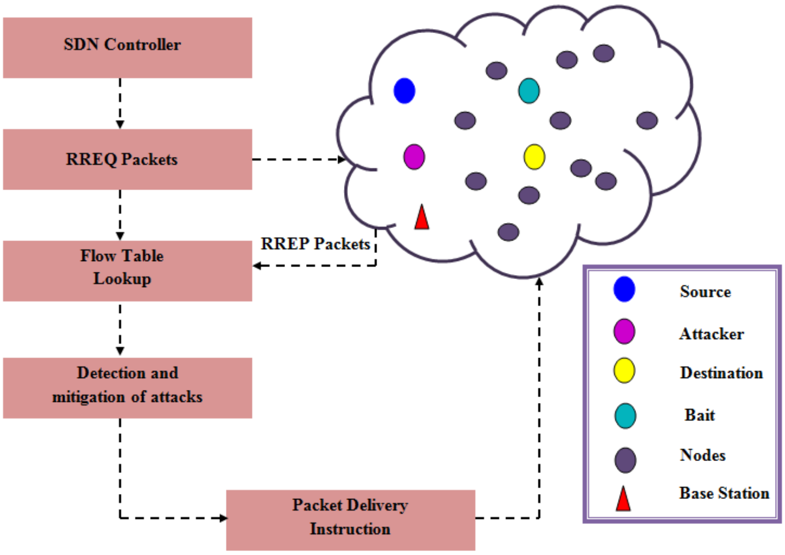

Fig. 3: General structure of BAIT [24].

4.7 Data Encryption using ESHP-ECC mechanism :

Elliptic-curve cryptography (ECC) is a public-key cryptosystem based on the elliptic curve hypothesis, which is a secure asymmetric encryption scheme used for data security. It generates public and private keys for each user through the elliptic curve properties. These keys are then used to encrypt and decrypt the data. In a conventional ECC technique, the keys are generated randomly. So, the attackers may easily hack the key information. In order to overcome this issue, the probability of ones and zeros are generated based on the randomly generated key value. Also, the key values are converted into a hash value using the secure hash method. Due to the alterations in the general ECC, the proposed technique is called as Encrypted Secure Hash Probability-based Elliptic-curve cryptography (ESHP-ECC) algorithm. The encryption process of ESHP-ECC is detailed below,

- At first, the elliptic curve equation used for key generation is given by,

$$Y^2 = X^3 + aX + b \qquad \text{--------} \quad (24)$$

In (24), a, b denotes the integers.

- Then, a random number (η) is generated from $[1, \quad n-1]$ and the probability of ones and zeros of this random number is calculated, defined as the private key. After that, the public key (ρ) is calculated as,

  $$\rho = \eta * \mathrm{B} \qquad \text{-----------} \quad (25)$$

  Here, B describes the point on the elliptic curve.
- Thereafter, these public and private keys are converted into a hash value using a secure hashing method. Secure Hashing Algorithm (SHA) is a cryptographic hash function that takes the keys as the input and produces a 160-bit (20-byte) hash value. The private and public keys after hashing are represented as η'' and ρ'' accordingly.
- Consider, M be the message to be transmitted and it has the point Q on the elliptic curve. Randomly select σ from $[1, \quad n-1]$. Two ciphertexts $\left(C^{(1)}, C^{(2)}\right)$ are calculated using equations (26), (27),

  $$C^{(1)} = \sigma * \mathrm{B} \qquad \text{------------} \quad (26)$$

  $$C^{(2)} = Q + \sigma * \rho \qquad \text{------------} \quad (27)$$

Where, $\left(C^{(1)}, C^{(2)}\right)$ defines the encrypted message that is transmitted to the cloud server through the shortest path.

4.8 Shortest Pathway Calculation :

Let, $\left(x_i = x_1, x_2, \ldots\ldots, x_N\right)$ be the number of sensor nodes available to transmit the encrypted message. The shortest path between each sensor node is identified for efficient data transmission as well as to reduce the computational time. Therefore, Euclidean distance $E^{(d)}$ is used to calculate the distance. It is formulated as follows,

$$E^{(d)} = \left\|\left(x_i - x_j\right)\right\|^2 \qquad \text{------------} \quad (28)$$

In this equation, x_j mentions the $j-th$ node. After computing the distance, the shortest pathway obtained is used to transmit the encrypted message. This encrypted message is further decrypted at the receiver side using Decrypted Secure Hash Probability-based Elliptic-curve cryptography (DSHP-ECC) algorithm.

4.9 Decryption through DSHP-ECC :

The encrypted message in equation (27) is decrypted using the below equation,

$$Q = C^{(2)} - \eta * C^{(1)} \qquad \text{-------------} \quad (29)$$

Where, Q specifies the original message.

5. RESULTS :

The detailed analysis of the outcome of the suggested structure is explained in this section. To demonstrate the work's efficacy, the performance analysis, as well as the comparative analysis, is carried out. The implementation of the proposed methodology is done by using MATLAB, and the data are obtained from the UNSW-NB15dataset, which is publically available on the internet.

5.1 Dataset description :

The Australian Cybersecurity Center's (ACCS) Network Scope Lab used the IXIA Perfect Storm engine to build a mix of operations using unprocessed network packets from the UNSWNB 15 dataset. Modern synthetic assaults and genuine modern normals UNSWNB154.csv, UNSWNB154.csv, UNSWNB153.csv, and UNSWNB152.csv are four CSV files containing a total of two million and 540,044 records: UNSWNB154.csv, UNSWNB151.csv, UNSWNB153.csv, and UNSWNB152.csv.

The training set has 175,341 records, whereas the test set contains 82,332 records, comprising normal and attack records.

5.2 Performance Analysis of the proposed BRELU-ResNet mechanisms :

The suggested BRELU-ResNet is compared with existing methodologies such as Artificial Neural Network (ANN), Convolution Neural Network (CNN), Adaptive Network-based Fuzzy Inference System in terms of sensitivity, False positive rate (FPR), accuracy, False negative rate (FNR), precision, recall, specificity, F-Measure, and Matthews correlation coefficient (MCC) (ANFIS). The comparative analysis is also done with the existing techniques to state the effectiveness of the model.

Table 1: Performance analysis of proposed BRELU-ResNet based on sensitivity, specificity, and accuracy.

Techniques	Performance metrics (%)		
	Sensitivity	Specificity	Accuracy
Proposed BRELU-ResNet	98.34	77.54	96.6
CNN	97.81	63.62	94.58
ANN	95.78	58.84	93.23
ANFIS	91.17	44.42	90.61

Table 1 demonstrates the performance analysis of the proposed BRELU-ResNet with various existing techniques, such as CNN, ANN, and ANFIS in terms of sensitivity, specificity, and accuracy.

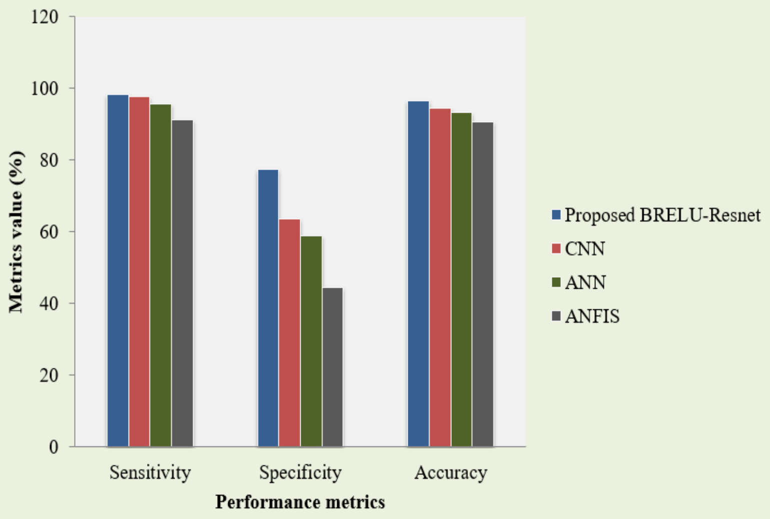

Fig. 4: Comparative analysis of proposed BRELU-ResNet based on Sensitivity, specificity and accuracy.

A clear view of tabulation 1 is given in figure 4. Figure 4 shows the comparative analysis of the proposed work. This comparative analysis clearly states that the proposed framework tends to attain higher Sensitivity, Specificity and accuracy values that range between 77.54%-98.345 whereas the existing techniques CNN, ANN, and ANFIS, attain the metrics values that range between 44.42%-97.81%, which is comparatively lower than the proposed BRELU-ResNet. As a consequence, the suggested strategy outperforms previous state-of-the-art methods and produces more notable outcomes in a variety of challenging situations.

Table 2: Performance analysis of proposed BRELU-ResNet based on precision, F-measure, and recall.

Techniques	Performance metrics (%)		
	Precision	Recall	F-measure
Proposed BRELU-ResNet	97.96	98.34	98.15
CNN	96.26	97.81	97.03
ANN	96.9	95.78	96.34
ANFIS	92.49	97.17	94.77

Tabulation 2 comprises the value of the performance metrics, like precision, F-Measure, and recall of the proposed BRELU-ResNet and the other existing works, like CNN, ANN, and ANFIS.

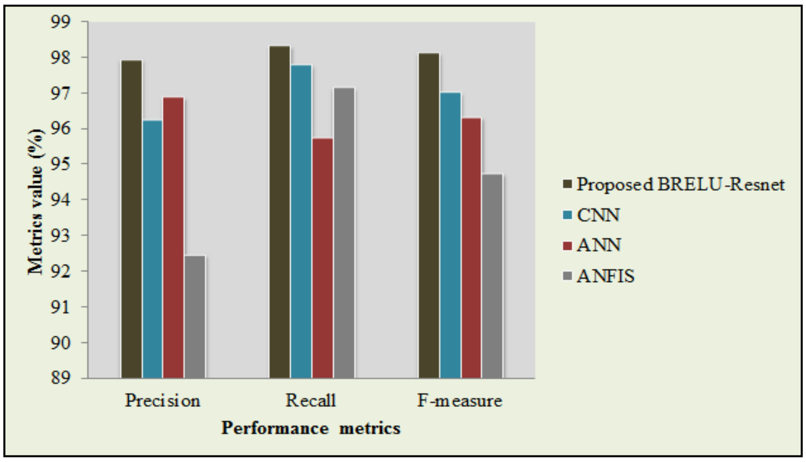

Fig. 5: Comparative analysis of proposed BRELU-ResNet based on precision, recall, and F-Measure.

A clear view of tabulation 2 is given in figure 5. Figure 5 shows the comparative analysis of the proposed work. This comparative analysis clearly states that the proposed framework tends to attain higher precision, recall, and F-Measure values that range between97.96%-98.34% whereas the existing techniques CNN, ANN, and ANFIS, attain the metrics values that range between 92.49%-97.81%, which is comparatively lower than the proposed BRELU-ResNet. As a consequence, the suggested strategy outperforms previous state-of-the-art methods and produces more notable outcomes in a variety of challenging situations.

Table 3: Performance analysis of proposed BRELU-ResNet with respect to FPR, FNR, and MCC.

Techniques	**Performance metrics (%)**		
	False Positive Rate	**False Negative Rate**	**Matthews Correlation Coefficient**
Proposed BRELU-ResNet	22.46	1.66	77.38
CNN	36.38	2.19	66.24
ANN	41.16	4.22	51.12
ANFIS	55.58	2.83	50.6

Table 3 depicts the performance evaluation of the proposed BRELU-ResNet and other existing techniques with respect to FPR, FNR, and MCC.

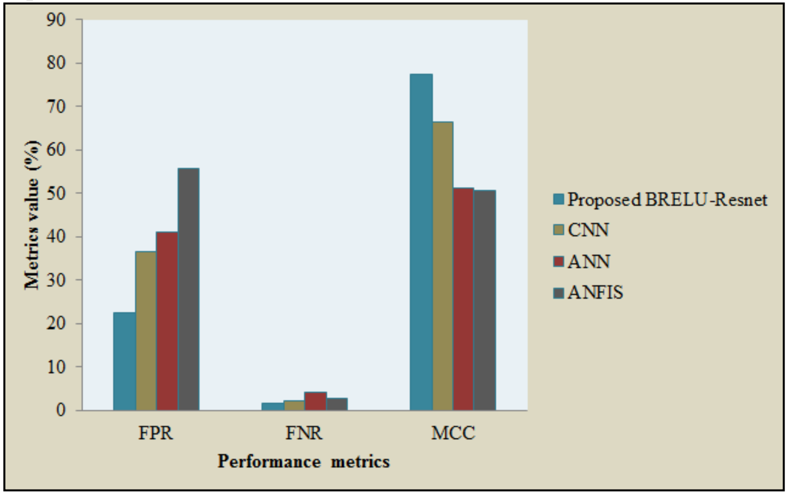

Fig. 6: Comparative analysis of proposed BRELU-ResNet in terms of FPR, FNR, and MCC.

Figure 6 compares the evaluation metrics such as FPR, FNR, and MCC of the proposed work with the existing works. The significance of the model is determined by the low value of FPR and FNR rates and the high value of MCC. In accordance with the above-mentioned statement, the FPR and FNR rates of the proposed work are low and the MCC rate achieved by the proposed work is higher than the existing approaches. Hence, the proposed method outperforms the other state-of-art methods and delivers better outcomes in the cyber-attack detection process.

6. DISCUSSION :

The proposed strategy achieves an over the top responsiveness level of 98.34 percent, particularity level of 77.54 percent, exactness of 96.6 percent, Precision level of 97.96 percent, review level of 98.34 percent, F-proportion of 98.15 percent, False Positive Rate of 22.46 percent, False Negative Rate of 1.66 percent, Matthew's relationship coefficient of 77.38 percent, according to the results obtained in the previous area, classified information. As a result, it can be deduced that the proposed methodology actually identifies digital assault; as a result, the organization's privacy is improved as well as more solidified, and it outperforms present methodologies. As a result, the suggested BRELU-ResNet-based Cyber-assault Detection architecture, which includes a BAIT-based far-reaching mitigation mechanism, protects the cloud server against gatecrashers. According to the claim, the proposed strategy achieves a level of awareness of 98.34 percent, particularity of 77.54 percent, and exactness of 96.6 percent, whereas current procedures such as CNN, ANN, and ANFIS achieve a level of responsiveness of 94.92 percent, explicitness of 55.62 percent, and precision of 92.80 percent, respectively. As a result, when compared to current works, the proposed BRELU-ResNet achieved better measurement rates. The meaning of the increased rate of accuracy, F-measure, and review is not totally clear. The methodology proposed, according to the claim, achieves 97.96 percent accuracy, 98.34 percent review, and 98.15 percent F-measure. Regardless, the current work achieves the average accuracy, review, and F-measure rate of 95.21 percent, 96.92 percent, and 96.04 percent, respectively. When compared to the intended task, this is a modest number. As a result, the proposed BRELU-ResNet reduces complexities and enhances the consistency of the digital assault recognition process. FPR and FNR's lower value effectively eliminates the misclassification or mis-expectation error. According to this claim, the proposed technique achieves 22.46 percent FPR and 1.66 percent FNR values. In any event, the existing methods' typical FPR and FNR upsides are 44.37 percent and 3.08 percent, respectively. The greater the MCC value, the more powerful the model; in this case, the proposed strategy obtains 77.38 percent of MCC, while the present methods obtain a typical MCC value of 55.98 percent. As a result, it is discovered that the proposed work is more reliable and outperforms existing approaches.

7. CONCLUSION :

The work has proposed a novel approach of BRELU-ResNet based Cyber-Attack Detection System with a BAIT-based approach for mitigation. This approach involved several operations designed to check cyber-attacks quickly. The work is pre-processed, feature extracted, feature selected, and classified for intrusion detection. The classification phase efficiently detects whether the data is normal or attack. If the data is normal, then the data transmission process begins. To ensure security, the encryption and decryption process is performed by the means of the SHP-ECC algorithm. The experimental analysis is then carried out, which includes performance analysis and a comparison study of the offered methodologies in terms of various performance measures to test the proposed algorithm's efficacy. The developed approach can handle various uncertainties and render more promising results. The publically available datasets called UNSW-NB 15 dataset is used for the analysis, in which the proposed method achieves 98.34% of sensitivity, 77.54% of specificity, 96.6% of accuracy, 97.96 % of Precision, 98.34% of recall, 98.15 % of F-measure, 22.46% of False Positive Rate, 1.66 % of False Negative Rate, 77.38 % of Matthew's correlation coefficient.

REFERENCES :

[1] Noorizadeh, M., Shakerpour, M., Meskin, N., Unal, D., & Khorasani, K. (2021). A cyber-security methodology for a cyber-physical industrial control system testbed. *IEEE Access*, *9(1)*, 16239-16253. Google Scholar↗

[2] Elnour, M., Meskin, N., Khan, K., & Jain, R. (2020). A dual-isolation-forests-based attack detection framework for industrial control systems. *IEEE Access*, *8(1)*, 36639-36651. Google Scholar↗

[3] Paridari, K., O'Mahony, N., Mady, A. E. D., Chabukswar, R., Boubekeur, M., & Sandberg, H. (2017). A framework for attack-resilient industrial control systems: Attack detection and controller reconfiguration. *Proceedings of the IEEE*, *106*(1), 113-128. Google Scholar↗

[4] Barrère, M., Hankin, C., Nicolaou, N., Eliades, D. G., & Parisini, T. (2020). Measuring cyber-physical security in industrial control systems via minimum-effort attack strategies. *Journal of information security and applications*, *52(1)*, 102471. Google Scholar↗

[5] Yang, J., Zhou, C., Yang, S., Xu, H., & Hu, B. (2017). Anomaly detection based on zone partition for security protection of industrial cyber-physical systems. *IEEE Transactions on Industrial Electronics*, *65*(5), 4257-4267. Google Scholar↗

[6] Adepu, S., & Mathur, A. (2018). Assessing the effectiveness of attack detection at a hackfest on industrial control systems. *IEEE Transactions on Sustainable Computing*, *6*(2), 231-244. Google Scholar↗

[7] Abana, M. A., Peng, M., Zhao, Z., & Olawoyin, L. A. (2016). Coverage and rate analysis in heterogeneous cloud radio access networks with device-to-device communication. *IEEE Access*, *4(2)*, 2357-2370. Google Scholar↗

[8] Sargolzaei, A., Yazdani, K., Abbaspour, A., Crane III, C. D., & Dixon, W. E. (2019). Detection and mitigation of false data injection attacks in networked control systems. *IEEE Transactions on Industrial Informatics*, *16*(6), 4281-4292. Google Scholar↗

[9] Ponomarev, S., & Atkison, T. (2015). Industrial control system network intrusion detection by telemetry analysis. *IEEE Transactions on Dependable and Secure Computing*, *13*(2), 252-260. Google Scholar↗

[10] Guo, H., Pang, Z. H., Sun, J., & Li, J. (2021). An output-coding-based detection scheme against replay attacks in cyber-physical systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *68*(10), 3306-3310. Google Scholar↗

[11] Han, S., Xie, M., Chen, H. H., & Ling, Y. (2014). Intrusion detection in cyber-physical systems: Techniques and challenges. *IEEE systems journal*, *8*(4), 1052-1062. Google Scholar↗

[12] Lu, K. D., Zeng, G. Q., Luo, X., Weng, J., Luo, W., & Wu, Y. (2021). Evolutionary deep belief network for cyber-attack detection in industrial automation and control system. *IEEE Transactions on Industrial Informatics*, *17*(11), 7618-7627. Google Scholar↗

[13] Genge, B., Siaterlis, C., Fovino, I. N., & Masera, M. (2012). A cyber-physical experimentation environment for the security analysis of networked industrial control systems. *Computers & Electrical Engineering*, *38*(5), 1146-1161. Google Scholar↗

[14] Baldoni, S., Battisti, F., Carli, M., & Pascucci, F. (2021). On the Use of Fibonacci Sequences for Detecting Injection Attacks in Cyber Physical Systems. *IEEE Access*, *9*(1), 41787-41798. Google Scholar↗

[15] Sui, T., Mo, Y., Marelli, D., Sun, X., & Fu, M. (2020). The vulnerability of cyber-physical system under stealthy attacks. *IEEE Transactions on Automatic Control*, *66*(2), 637-650. Google Scholar↗

[16] Jahromi, A. N., Karimipour, H., Dehghantanha, A., & 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), 13712-13722. Google Scholar↗

[17] Lv, Z., Han, Y., Singh, A. K., Manogaran, G., & Lv, H. (2020). Trustworthiness in industrial IoT systems based on artificial intelligence. *IEEE Transactions on Industrial Informatics*, *17*(2), 1496-1504. Google Scholar↗

[18] Haller, P., & Genge, B. (2017). Using sensitivity analysis and cross-association for the design of intrusion detection systems in industrial cyber-physical systems. *IEEE Access*, *5*(1), 9336-9347. Google Scholar↗

[19] Zhang, F., Kodituwakku, H. A. D. E., Hines, J. W., & Coble, J. (2019). Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data. *IEEE Transactions on Industrial Informatics*, *15*(7), 4362-4369. Google Scholar↗

[20] Al-Abassi, A., Karimipour, H., Dehghantanha, A., & Parizi, R. M. (2020). An ensemble deep learning-based cyber-attack detection in industrial control system. *IEEE Access*, *8*(1), 83965-83973. Google Scholar↗

[21] Kravchik, M., & Shabtai, A. (2021). Efficient cyber attack detection in industrial control systems using lightweight neural networks and PCA. *IEEE Transactions on Dependable and Secure Computing*, *10*(1),1-18. Google Scholar↗

[22] Kajaet, N., Shaout, A., & Ma, D. (2019). An intelligent intrusion detection system. *Applied Intelligence*, *49*(9), 3235-3247. Google Scholar↗

[23] Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network intrusion detection combined hybrid sampling with deep hierarchical network. *IEEE Access*, *8*(1), 32464-32476. Google Scholar↗

[24] Prabhu, S., & Nethravathi, P. S. (2022). Novel SHP-ECC Mechanism Architecture for Attack Node Mitigation and to Predict Future Community Intrusions. *International Journal of Applied Engineering and Management Letters (IJAEML)*, *6*(1), 176-183. Google Scholar↗

[25] Gopalakrishnan, T., Ruby, D., Al-Turjman, F., Gupta, D., Pustokhina, I. V., Pustokhin, D. A., & Shankar, K. (2020). Deep learning enabled data offloading with cyber attack detection model in mobile edge computing systems. *IEEE Access*, *8*(1), 185938-185949. Google Scholar↗

[26] Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications surveys & tutorials*, *18*(2), 1153-1176. Google Scholar↗

[27] Hussain, B., Du, Q., Sun, B., & Han, Z. (2020). Deep learning-based DDoS-attack detection for cyber–physical system over 5G network. *IEEE Transactions on Industrial Informatics*, *17*(2), 860-870. Google Scholar↗

[28] Tertytchny, G., Nicolaou, N., & Michael, M. K. (2020). Classifying network abnormalities into faults and attacks in IoT-based cyber physical systems using machine learning. *Microprocessors and Microsystems*, *77*(1), 103121. Google Scholar↗

[29] Ibor, A. E., & Epiphaniou, G. (2015). A hybrid mitigation technique for malicious network traffic based on active response. *International Journal of Security and Its Applications*, *9*(4), 63-80. Google Scholar↗

[30] Akyazi, U., & Force, T. A. (2014). Possible scenarios and maneuvers for cyber operational area. *European Conference on Cyber Warfare and Security*, *1*(10), 1-7. Google Scholar↗

****