ITC 2/54
Information Technology and Control
Vol. 54 / No. 2/ 2025
pp. 629-642
DOI 10.5755/j01.itc.54.2.39543

An Early Warning Model for Industrial Network Security Issues: A Crafted Strategy for High Accuracy Based on Machine Learning Approach

Received 2024/11/20 | Accepted after revision 2025/04/07

HOW TO CITE: Le, X., Zhao, Y. (2025). An Early Warning Model for Industrial Network Security Issues: A Crafted Strategy for High Accuracy Based on Machine Learning Approach. *Information Technology and Control*, 54(2), 629-642. https://doi.org/10.5755/j01.itc.54.2.39543

An Early Warning Model for Industrial Network Security Issues: A Crafted Strategy for High Accuracy Based on Machine Learning Approach

Xiang Le
School of Computer, Beijing University of Technology, Beijing 100124, China
Ningbo Holly Sys. Information Security Research Institute Co., Ltd., Ningbo 315100, Zhejiang, China

Yong Zhao
School of Computer, Beijing University of Technology, Beijing 100124, China

Corresponding author: zhaoyong08@bjut.edu.cn

An industrial network has become an important infrastructure. As industrial networks develop, their cybersecurity problems become more and more prominent. The attacks currently realized to networks turn out to be advancing quicker than ever, and their destructive force also continuously gets bigger. Thus, the available early warning technology for industrial network security issues requires more accuracy and timeliness since a serious amount of delays occurs in real cases. The article proposes a strategy with high accuracy based on a machine-learning algorithm. Nonlinear high-dimensional data with different feature characteristics in cyber-attacks and low training efficiency of conventional early warning models to predict attacks are underlined as a significant part of the problem to deal with. Thus, the manuscript suggests a feature selection method based on the Tuna Swarm Optimization (TSO) algorithm to filter out redundant features and reduce the data's dimensionality. Then, the Extreme Learning Machine (ELM) and Auto-Encoder (AE) are combined to construct the model called Extreme Learning Machine-Auto Encoder (ELM-AE) to be implemented as the basis of the early warning model for industrial network security. Afterward, the improved Whale Optimization Algorithm (I-WOA) is used to optimize the parameters of the ELM, to construct the obtained optimization model. Finally, the obtained optimization model is applied to detect attacks on industrial cy-

ber security systems as an early warning method. Eventually, the proposed model is tested by constructing an evaluation index system on how effective the early warning system functions. The experimental results show that the proposed warning model for industrial network security issues has high warning accuracy and efficiency concurrently, which provides an advanced early warning model for network attacks. The proposed model with 92.64% precision and 51.84 s average execution time excels over other methods.

KEYWORDS: Industrial Network; Network Attack; TSO; ELM; AE

1. Introduction

An industrial network is composed of a complex industrial system and Internet technology. It houses several technologies such as global industrial technology, advanced computer technology, analytical process technology, sensor technology, network technology, and other highly integrated new products [24]. An industrial network breaks through a relatively closed environment of conventional industrial systems, all emerging systems, production equipment, and external network connections. Its application is so important that the economic growth and social development of each country depend on these new industrial production and constructed infrastructure systems, bringing a new driving force [12].

The future development of industrial networks has vast opportunities. However, several defects are also accompanied by Internet-based production systems [29]. Since production equipment systems have been continuously in use and are connected to the Internet for online supervision to repair defects and regularly surveil and update systems, they are open to network attacks, thus posing an increasing threat to industrial systems' security. If a security breach occurs, it will not only cause huge economic losses but also jeopardize public safety. To mitigate the issues that exist in industrial network security, a growing demand for early warning methods for industrial network security is placed in the sector [7].

In the complex and changing industrial network environment, although early warning technologies for conventional network security such as firewalls, intrusion detection systems, and a series of security components are deployed for critical applications, these technologies' instant reactions are not enough to protect industries' applications and a delay in a fraction of a second to react occurs when an attack happens, and the accuracy level of an early warning system is not satisfactory at all. To resolve this problem, industrial networks must have accurate and efficient warning models [9].

Aiming at providing a solution to the above problems, the article researches the limitations of the conventional early warning technologies for industrial networks' security issues, points out the available security risks of industrial networks, designs a feature selection method used for detecting network attacks based on the Tuna Swarm Optimization (TSO) algorithm, and adopts the Improved Whale Optimization Algorithm (I-WOA) to optimize the Extreme Learning Machine (ELM)-Auto Encoder (AE) algorithm and proposes I-WOA-ELM-AE algorithm be used as an early warning model to detect attacks without a delay on industrial network security, which provides early protection for malicious intrusion behaviors in industrial networks and enhances the warning capability of industrial networks.

The research contributes to literature as follows: 1. By combining 2 efficient algorithms, which are the Extreme Learning Machine (ELM) [13] and Auto-Encoder (AE) [8] called Extreme Learning Machine-Auto Encoder (ELM-AE); 2. Choosing effective attributes used in the proposed algorithm based on the Tuna Swarm Optimization (TSO) algorithm [10]; 3. Optimizing the parameters of the Extreme Learning Machine (ELM) using the improved Whale Optimization Algorithm (I-WOA) [25]; 4. Proposing the early warning model called the IWOA-ELM-AE AE to detect attacks without a delay on industrial network security.

2. Related Work

An early warning technology for industrial network security has a very important guiding role in the construction of Internet-based industrial information systems. The current academic and industrial research on early

warning technology for industrial network security has continued for a long period and achieved certain solid research results, especially in the fields of electricity production and computer security. Although early warning technology for industrial network security effectively prevents hacking and thus, protects the security of industrial network information, a complete early warning system has not yet been formed so far.

Janabi et al. [14] proposed a network attack detection method based on Convolutional Neural Networks (CNNs) for Software-Defined Networks, which can successfully detect a variety of attacks with good real-time performance and provide some early warning effects. Neira et al. [19] designed a cooperative system for early warning signals to predict DDoS attacks in advance, which recognizes the attack signals before the attacker launches an attack, thus enabling sufficient preparation time before a DDoS attack occurs. However, other types of defenses to attacks are not examined, so the practical aspects remain to be further verified. Azzam et al. [3] proposed a preliminary metric to gauge the success likelihood of an attack in real-time for stealth attacks on industrial control systems and applied the metric to a linear time-invariant (LTI) system for testing, demonstrating that the metric can provide early warning of a potential stealth attack before it causes damages. However, the method did not provide an in-depth analysis and lacked a comparison of the accuracy of the warning. Abdalzaher et al. [1] focused on the application of artificial intelligence in IoT smart systems and proposed trust techniques based on machine learning models to cope with the cyber security of smart systems in IoT, but the universality of the technique was not explored. Chakkaravarthy et al. [5] designed a robust Intrusion Detection Honeypot (IDH) to address the limitations of the available security systems that are complex and time-consuming and unable to warn of ransomware. In IDH, the Honeyfolder module uses the Social Leopard Algorithm to model the decoy folder and is used for attack warnings. AuditWatch was designed to verify the entropy of files and folders. Cep engine was used to aggregate the security data from multiple sources to confirm the ransomware signature and respond instantly, the designed IDH significantly improved the ransomware detection time, detection rate, and response time. The IDH was designed to improve the effectiveness of ransomware alerts significantly, however, was not compared in depth with other methods and had some limitations. Bou-Harb et al. [4] proposed a novel system, CSC-Detector, for early warning of network scanning activities, in response to darknet activities. The system can be used for early warning of network attacks orchestrated in the Dark Web by using a fingerprinting engine to obtain activity data from Dark Web traffic, an inference engine to generate insights, and an analysis engine to infer activity. However, the practical effects of the method were not presented for the middle and late stages of network scanning activities, which has some limitations. Kumar et al. [15] developed a DL-TIF automation framework in IoT-enabled Maritime Transportation Systems (MTS), which consisted of a Deep Feature Extractor (DFE), CTI Driver Detection (CTIDD), and CTI Attack Type Identification (CTIATI). Although the cumulative research has achieved certain results in dealing with industrial network security threats, they usually target only a single type of network attack and is difficult to achieve a balance between accuracy and efficiency. Therefore, to address this difficulty, the article proposes an early warning model based on a new machine learning-based cybersecurity, which can realize high-accuracy strategies and can balance the efficiency of early warning to improve the level of industrial cybersecurity. More up-to-date research can be found in [21, 23, 2, 27].

3. Feature Extraction of Cyber Attacks

3.1 The Presentation of the Problem

Features of cybersecurity data in industrial networks contain a large number of features with different characteristics, and it is required to determine which ones fit into security warning tasks. In early warning security issues, not all features are highly relevant to the judgment of network attacks, i.e., many features are redundant in network attack cases. Directly utilizing raw high-dimensional data for early warning judgment tasks may encounter a "dimensionality curse" problem, and at the same time, processing a large amount of high-dimensional data will take up higher computational resources, storage space, and time.

Therefore, a dimensionality reduction method to detect better features used in the construction of an early warning model is conducted in this paper. Thus, this process filters out redundant features that do not contribute to or have less influence on detection tasks. By doing so, the feature set is simplified and the optimal set of the features is attained. Hence, the resource cost is largely reduced, and the difficulty of the model training is avoided, which prevents overfitting issues and improves the efficiency of the constructed early warning model.

3.2 The Design of the Methodology

This subsection abstracts the feature selection problem for industrial cybersecurity by assuming an initial cybersecurity feature dataset, $B = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$, and the sample matrix $X = (x_1, x_2, \ldots, x_m)^T$. Each row represents a set of industrial cybersecurity samples, each column represents a set of conjugate cybersecurity features, and the matrix of cyberattack types is denoted by $Y = (y_1, y_2, \ldots, y_m)^T$, so, each row corresponds to a network attack type. If there are m samples and n features, $x_i = (x_{i1}, x_{i2}, \ldots, x_{im})$ is a vector of the n-dimensional features space of the network.

$\sigma = span\{f_1, f_2, \ldots, f_j, \ldots f_n\}$, then x_{ij} denotes the value of the sample x_i on the jth feature f_i.

The TSO algorithm can effectively improve the convergence speed and optimization accuracy of the algorithm as well as have better stability when compared with other optimization algorithms such as SSA, DE, and GWO. Therefore, the manuscript picks the TSO algorithm to detect features that will be employed in the constructed early warning model based on the original cybersecurity feature dataset B [22]. To improve the efficiency of the constructed model, it is necessary to filter out features that contribute less to the proposed model, and the features are selected to determine a subset of the features Z for the evaluation function $f(z)$ when a smaller number of features are used for $f(z)$. The bigger the $f(z)$ is, the better the effect of the subset Z.

Furthermore, the TSO algorithm simulates 2 foraging behaviors of tuna populations: spiral-form foraging and parabolic-form foraging. When the original dataset of network security is used as input, it is processed according to the previously mentioned treatment that partitioned the data into the set of network security features and the corresponding set of network attack types, and the matrix of network security features as well as the matrix of network attack types generated respectively. Finally, the generated feature subsets are utilized as tuna individuals.

The TSO algorithm uses a random selection of initial positions in the search space for population initialization [20]:

$$X_i = Rand * (UB - LB) + LB, i = 1, 2, \cdots, N \quad (1)$$

where X_i denotes the ith individual, which represents the initial subset of features. UB and LB represent the upper and lower bounds of the search space.

This random initialization has a drawback in that the generated initial positions could be unstable, which may cause an unbalanced distribution of the initial population, thus affecting the capability of global exploitation in the early-stage warning problems. In the article, a good point set is used to initialize the population, which is an effective uniform distribution, when compared to a random method, it helps the initial population be more uniformly distributed in the search space, thus, obtaining a stronger global development capability.

The population based on the set of good points is initialized as follows.

1 Construct a good point set $X = \{x_1, x_2, \ldots, x_N\}$, which contains N points that represent the number of populations.

2 Compute the jth component of x_i, where $x_{ij} = 2 \cdot i \cdot \cos\left(\frac{2\pi j}{p}\right)$.

3 Map the good point set into the search space, $x_{ij} = LB + \mathrm{mod}(x_{ij}, 1) \cdot (UB\text{-}LB)$.

Spiral-form foraging and parabolic-form foraging are defined as follows: Spiral-form foraging: When the TSO algorithm spirals the search, the need for a subset of excellent features in a wide search space in the pre-iteration is on. Hence, the algorithm will randomly generate a subset of features as a reference target. With the increase in the number of iterations, the algorithm also gradually searches a small search scope, at this time no longer a randomly generated feature subset is needed, but in the current optimal subset around the search. In summary, the model of spiral form search is delineated as follows:

$$X_j^{t+1}=\begin{cases}\begin{cases}\alpha_1\cdot(X_{\text{rand}}^t)+\beta\cdot(|X_{\text{rand}}^t-X_i^t|)+\alpha_2\cdot X_i^t, i=1\\ \alpha_1\cdot(X_{\text{rand}}^t)+\beta\cdot(|X_{\text{rand}}^t-X_i^t|)+\alpha_2\cdot X_{i-1}^t,\\ i=1,2,\cdots,N\end{cases}\\ \begin{cases}\alpha_1\cdot(X_b^t)+\beta\cdot(|X_b^t-X_i^t|)+\alpha_2\cdot X_i^t, i=1\\ \alpha_1\cdot(X_b^t)+\beta\cdot(|X_b^t-X_i^t|)+\alpha_2\cdot X_{i-1}^t\\ i=1,2,\cdots,N\end{cases}\end{cases} \quad (2)$$

$$\alpha_1=a+(1-a)\cdot\frac{t}{T} \quad (3)$$

$$\alpha_2=(1-a)-(1-a)\cdot\frac{t}{T} \quad (4)$$

$$\beta=\mathrm{e}^{bl}\cdot\cos(2\pi b) \quad (5)$$

$$1=\mathrm{e}^{3\cos\left(\pi\left(\left(T+\frac{1}{t}\right)-1\right)\right)} \quad (6)$$

where X_i^t denotes the i-th subset of features at t iterations, X_b^t denotes the current optimal subset of features, α_1 and α_2 represent the weighting coefficients. T represents the maximum number of iterations, b denotes a random number in [0,1], and a represents a constant.

Parabolic form foraging: In addition to the spiral form search, a portion of the tuna population searches in a parabolic form using the optimal individual as a reference. The other part searches around the optimal individual, and the parabolic form search is modeled as:

$$X_i^{t+1}=\begin{cases}X_b^t+rand(X_b^t-X_i^t)+TF.p^2\cdot(X_b^t-X_i^t),\\ rand<0.5\\ TF.p^2, rand>0.5\end{cases} \quad (7)$$

$$\log p=\frac{t}{T}\log\left(1-\frac{t}{T}\right), \quad (8)$$

where TF represents a random number with a value of 1 or -1. 50% of the feature subsets will be searched parabolically concerning the current optimal subset, and the other 50% will be searched in the vicinity of the current optimal subset.

3.3 The Steps of the Method

Due to the large amount of cybersecurity data, to reduce the computational cost and improve the efficiency of the feature selection process, this paper adopts ELM as a cybersecurity feature detection tool to design, to search for the subset of cybersecurity features with higher accuracy.

In the optimization problem, it is crucial to choose a good fitness function that should consider the specific optimization objectives. In the feature selection, 2 objectives appear to be more important, the first one is to detect the correct rate as high as possible, and the second one is to select the number of features as small as possible. Therefore, to evaluate the quality of the feature subset, both objectives are encapsulated into the fitness function, and the fitness function f is defined by

$$f=\alpha*(1-acc)+\beta*\frac{n}{N}, \quad (9)$$

where α and β denote the weight coefficients, acc represents the accuracy of ELM for detecting a subset of network security features, n denotes the number of features in the current subset of network traffic features, and N denotes the number of all features.

The steps of the selection of the network security features based on the TSO algorithm are shown in Figure. 1.

Step 1: Preprocess the data and set the TSO parameters.

Step 2: Initialize the population individuals in the TSO algorithm, i.e., the subsets of network security features, calculate the evaluation scores of each feature subset using the fitness function in Equation (9), and record the optimal feature subset and the worst feature subset.

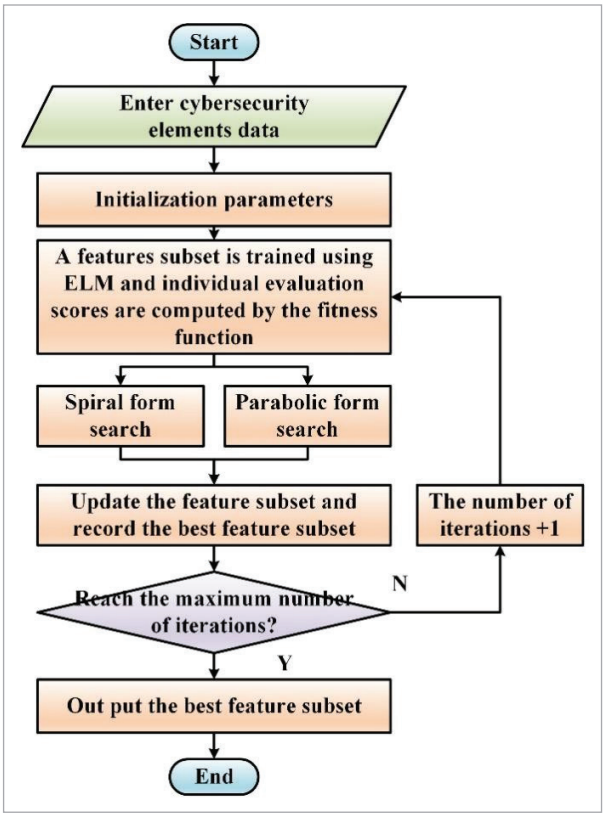
Step 3: Perform the update of the feature subset using Equations (2) and (7) and recalculate the evaluation scores of each feature subset, if the updated feature subset has higher scores, then replace it, otherwise keep it unchanged.

Step 4: Select the feature subset with the best evaluation score and the worst feature subset from the whole feature subset.

Step 5: Judge whether the algorithm is terminated or not, if satisfied then output the current optimal feature subset, otherwise return to Step 3.

Aiming at resolving the problem of the high dimensional dataset, which affects the efficiency and ac-

Figure 1
The steps of the feature selection of the network security based on the TSO algorithm.



curacy of the constructed early warning model, we abstract the problem of selecting the features into an optimization problem and make use of the TSO algorithm to find a subset of features that is conducive to the detection, to guarantee the detection accuracy and reduce the complexity of the data processing, and to provide the front-end support for the construction of the early warning model.

4. An Early Warning Model for Industrial Network Security

Compared to the deep neural network, the machine learning model's shallow network structure reduces the operation's required performance, saves computational resources, and has high classification accuracy for multi-source data. The data to be processed by the early warning model is text data, which requires a high degree of real-time accuracy. If it cannot be analyzed accurately within a reasonable period to defend against cyber-attacks in advance, there will be a lag and lead to the inability to respond to cyber-attacks promptly. Therefore, this paper's main concern is to focus on establishing an early warning model based on machine learning.

4.1 The Design of the Model

In this paper, when a warning model is designed for industrial network security, we mainly consider the combination of the ELM and AE to propose a model, since the gradient descent algorithm will be implemented, which tends to result in low training efficiency and being trapped in a local optimization when there are too many nodes within the layers in the network, and the classification accuracy of the ELM for large-capacity data samples tends to be difficult to meet the actual demand. Therefore, the output of the ELM can also be used as the input, which can simplify the process and thus make ELM have the function of AE so that the generalization performance, training efficiency, and classification accuracy of the proposed ELM-AE can be improved.

The input sample capacity of the ELM is set to n_e and the number of input and output nodes are assigned to m_e then, the input matrix formed by the input variables can be denoted as

$X=\{x_1,x_2,\ldots,x_n \mid x \in R^m\}$ and x can also act as the output variables. The ELM-AE implicit layer variables are delineated as follows [11]:

$$h = g(\omega_{ELM-AE} X + \alpha_{ELM-AE}), \tag{10}$$

where ω_{ELM-AE} and α_{ELM-AE} denote the input weights and the input bias, respectively, both of which are randomly generated. To ensure the generalization performance of the ELM-AE, orthogonal operations are performed on both, then [17]:

$$\begin{cases} \omega_{ELM-AE} \cdot \omega_{ELM-AE}^T = 1 \\ \alpha_{ELM-AE} \cdot \alpha_{ELM-AE}^T = 1 \end{cases} \tag{11}$$

Equation (11) is attained. The hidden layer variable $h_{ELM\text{-}AE}$ is related to the training output x':

$$h_{ELM-AE} \cdot \alpha_{ELM-AE} = x' \tag{12}$$

Then, the Loss function is set as follows:

$$\min Loss_{ELM\_AE} = \frac{\left\|\alpha_{ELM-AE}\right\|^2 + C\left\|x - x'\right\|}{2} \tag{13}$$

where C denotes the regularization parameter, which regulates the generalization performance of the ELM-AE. The optimization $\alpha_{ELM\text{-}AE}$ is resolved based on the least squares method. Let

$$\mathbf{K} = C^{-1}\mathbf{E} + \mathbf{H}^T_{ELM\_AE}\mathbf{H}_{ELM-AE} \quad \textbf{then}$$
$$\alpha_{ELM-AE} = \begin{cases} \mathbf{K}\mathbf{H}^T_{ELM-AE}X, n_e < n_m \\ \mathbf{H}^T_{ELM-AE}\mathbf{K}\mathbf{H}^T_{ELM-AE}X, n_e > n_m \end{cases}, \tag{14}$$

where **E** denotes the unit matrix.

In this paper, an early warning model for industrial network security based on ELM-AE is constructed whose steps are given as follows:

Step 1: Obtain data samples and divide the whole data set into the training set and test set of different anomaly-type data.

Step 2: Use the training set to train ELM-AE and obtain the trained ELM-AE detection model.

Step 3: Input the test set into the trained ELM-AE detection model to obtain the detection results.

Step 4: After all the test sets have been detected, the detection results are counted and the anomalies of the data are analyzed.

Although the ELM-AE can complete the effective classification of network security features to a certain extent to achieve early warnings, the ELM itself will be not reliable due to the randomness of the link weights leading to the final output results having strong volatility, so this paper designed the I-WOA algorithm used to achieve the parameter optimization of the ELM.

4.2 The Optimization of the Model

WOA algorithm has the advantages of a simple mechanism and few parameters, which can significantly improve the operation efficiency of the model when it is used to realize the parameter optimization of the machine learning model, which is more comprehensive than the TSO algorithm, so this paper chooses the WOA algorithm to be used in the realization of parameter searching of the ELM-AE.

4.2.1 Basic Principles of WOA

The WOA belongs to a new type of intelligent optimization algorithm, which in essence constructs a grid search scheme by imitating the behavior of humpback whales in hunting fish and shrimp. Humpback whales will first analyze and determine the location of the prey by moving the spit bubbles to make the bubbles surround the prey. Humpback whale hunting behavior is mainly divided into the following 3 stages [6, 18]:

1 Prey hunting stage

Since the location of the prey is unknown, the WOA should first assume that the existing optimal solution is the location of the target prey or a near point, and then gradually update its location:

$$S = \left|\lambda F^*(t) - F(t)\right| \tag{15}$$

$$F(t+1) = F^*(t) - \eta S, \tag{16}$$

where S denotes the rounding step, $F(t+1)$ denotes the position of the solution at the $t+1$-th iteration, $F^*(t)$ denotes the position of the optimal solution at the t-th iteration, and $F(t)$ denotes the position of the solution at the t-th iteration. η and λ represent random coefficient vectors. At each iteration, when a better solution than the current optimal solution is generated, $F^*(t)$ is dynamically updated, η and λ computed as follows:

$$\begin{cases} 2\delta\gamma - \delta \\ \delta = 2 - \dfrac{2t}{t_{\max}} \\ \lambda = 2\gamma \end{cases}, \tag{17}$$

where δ characterizes the control parameter, the iterative operation is gradually reduced from 2 to 0 linearly, $t_{\max}$ represents the upper limit of iteration. This paper is set to 500, $\gamma \in [0,1]$ representing the random value.

2 Bubble attack stage

In this stage there are 2 types of situations:

a Contraction of the prey range. At this time, the new position of an individual can take the value between the current individual position and the optimal individual position, and the process is consistent with Equation (16).

b Spiral update method. Spiral updating of the position is performed to resolve for the individual-prey spacing D:

$$F(t+1) = F^*(t) + De^{\rho t}\cos(2\pi l) \tag{18}$$

where $D = |F^*(t) - F(t)|$, ρ denotes the logarithmic helix shape constant, and τ represents a random value in [-1,1].

Here, 0.5 is used as a baseline, and a random probability $p \in [0,1]$ is generated to determine the hunting method. When $|\delta|<1$ and $p<0.5$, then the hunting range is contracted. When $|\delta|<1$ and $p>0.5$, then a spiral update is used.

3 Prey search phrase

The random search starts when $|\delta|>1$. Individuals are randomly screened to update the position until the iteration limit is reached, and its search expression is given as follows:

$$S = \left|\lambda F_{rand} - F(t)\right| \tag{19}$$

$$F(t+1) = F_{rand} - \eta S, \tag{20}$$

where F_{rand} denotes a randomized position vector and affects whether an individual carries out the current phase or not.

4.2.2 Improvements to WOA

1 Chaotic initialization of populations

Since the initial population in the WOA is often randomly generated, the imbalance of its distribution will weaken the optimization efficiency. Chaos theory, however, due to its nonlinearity and dynamic randomness, can promote the diversity of initial populations, so that the particle optimization region covers the whole domain and does not deprive the original population of its randomness. Therefore, this paper adopts Cubic chaotic mapping to perform the initialization of the population as follows:

$$x_{n+1} = \sigma x_n (1 - x_\eta^2) \tag{21}$$

where x_n denotes the chaotic sequence of the humpback whale population and $x_n \in (0,1)$. σ denotes the regulation covariate where $\sigma \in (1,5,3)$.

2 Nonlinear regulation of control parameters

Adjusting η can correct the optimization region of the WOA, and η is mainly affected by δ, which is proportional to the size of the optimization region. The convergence rate is relatively slow because of the δ decreasing linearity in the WOA. Therefore, nonlinear adjustment is introduced to improve the optimization performance of the WOA ensuring that the trend of δ remains unchanged, and the improvements are presented as follows:

$$\delta(t) = \delta_{mitial} - (\delta_{mitial} - \delta_{final}) \cdot \tan\left(\frac{\pi}{4}\left(\frac{t}{t_{\max}}\right)^2\right), \tag{22}$$

where $\delta_{initial}$ and δ_{final} denote the initial and final values of the control parameters, respectively. In the early stage of improvement, the decreasing speed is slow and mainly realizes the global optimization control, while in the later stage, the value decreases faster and the algorithm is terminated by fast convergence.

3 Adaptive weighting factor

Inertia weight is a key parameter in the WOA concerning the performance of global optimization and local optimization, so the adaptive weight factor is introduced in the WOA to complete the updated regulation of individual positions:

$$w(t) = w_{\max} - (w_{\max} - w_{\min}) \cdot \sin\left(\frac{\pi}{2}\left(\frac{t}{t_{\max}}\right)^2\right), \tag{23}$$

where $w(t)$ characterizes the current inertia weight, $w_{\max}$ and $w_{\min}$ characterize the upper and lower bounds of the inertia weight, respectively. At the beginning stage, the inertia weights converge to $w_{\max}$ and the focus is on realizing the global optimization, while at the later stage, the inertia weights converge to $w_{\min}$ and the focus is on realizing the local optimization to complete the operation.

Combining Equations (16) and (18), Equation (24) is attained:

$$F(t+1) = \begin{cases} w \cdot F^*(t) - \eta \cdot S, p < 0.5 \\ D \cdot e^{\rho t} \cdot \cos(2\pi l) + wF^*(t), p > 0.5 \end{cases} \tag{24}$$

4.3 The Steps of the Proposed Model

The process of applying the I-WOA to the parameter optimization of the ELM-AE to achieve the optimized early warning model for industrial cybersecurity system is presented as follows:

Step 1: Obtain data samples and divide a training set and a test set of different anomaly-type data.

Step 2: Input the training set into the optimized ELM-AE model to train the model

Step 3: Initialize the WOA population and population parameters by Chaos theory.

Step 4: Calculate the fitness value of the WOA population.

Step 5: Assume that the existing optimal solution is the location of the target prey or a near point, and determine whether the current p-value is less than 0.5, if so, contact the hunting range. At this time, the new position of the individual can take the value between the current individual position and the optimal individual position, and vice versa, the position is updated in a spiral.

Step 6: Start a random search immediately, and randomly screen individuals to update their positions until the iteration limit is reached.

Step 7: Determine whether the termination conditions are satisfied, if yes, terminate the optimization, otherwise output the optimal link weight matrix and bias matrix, and determine the optimal network structure of the ELM-AE based on the output results.

Step 8: Input the test set data into the trained optimal ELM-AE model to detect abnormal data.

5. Experimental Analysis

5.1 Data Sources and Experimental Settings

The dataset, UNSW-NB15 [26] is used, a cybersecurity dataset proposed by the UNSW Cybersecurity Laboratory in 2015 that is more in line with the characteristics of modern networks. It is composed of a mixture of real-life normal traffic and anomalous traffic from contemporary attacks, containing one type of normal traffic and nine types of anomalous traffic. All the experiments are conducted under Intel(R) Core (TM) i9-10850K CPU @ 5.20GHz, 16GB of RAM, 2TB of hard disk, and Windows 10 environment, and the simulation experiments are run using MatLab R2023A tool.

5.2 The Evaluation Indicators of The Model Performance

To evaluate the proposed early warning model more comprehensively, recall, accuracy, precision rates, and F-score are selected. These indicators can be calculated by the confusion matrix.

The recall rate (*Rec*) indicates the ratio of the number of samples in which the alert model determines an actual cyber attack as a cyber attack to the total number of samples of actual cyber attacks:

$$Rec = \frac{TP}{TP + Fm} . \tag{25}$$

Accuracy rate (*Acc*) indicates the percentage of the alerting model that judges normal data as normal and cyber-attack data as cyber-attack:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} . \tag{26}$$

Precision rate (*Pre*) indicates the ratio of the number of samples in which the alert model judged a cyber attack as a cyber attack to the number of samples in which all were judged as cyber attacks:

$$Pre = \frac{TP}{TP + FP} . \tag{27}$$

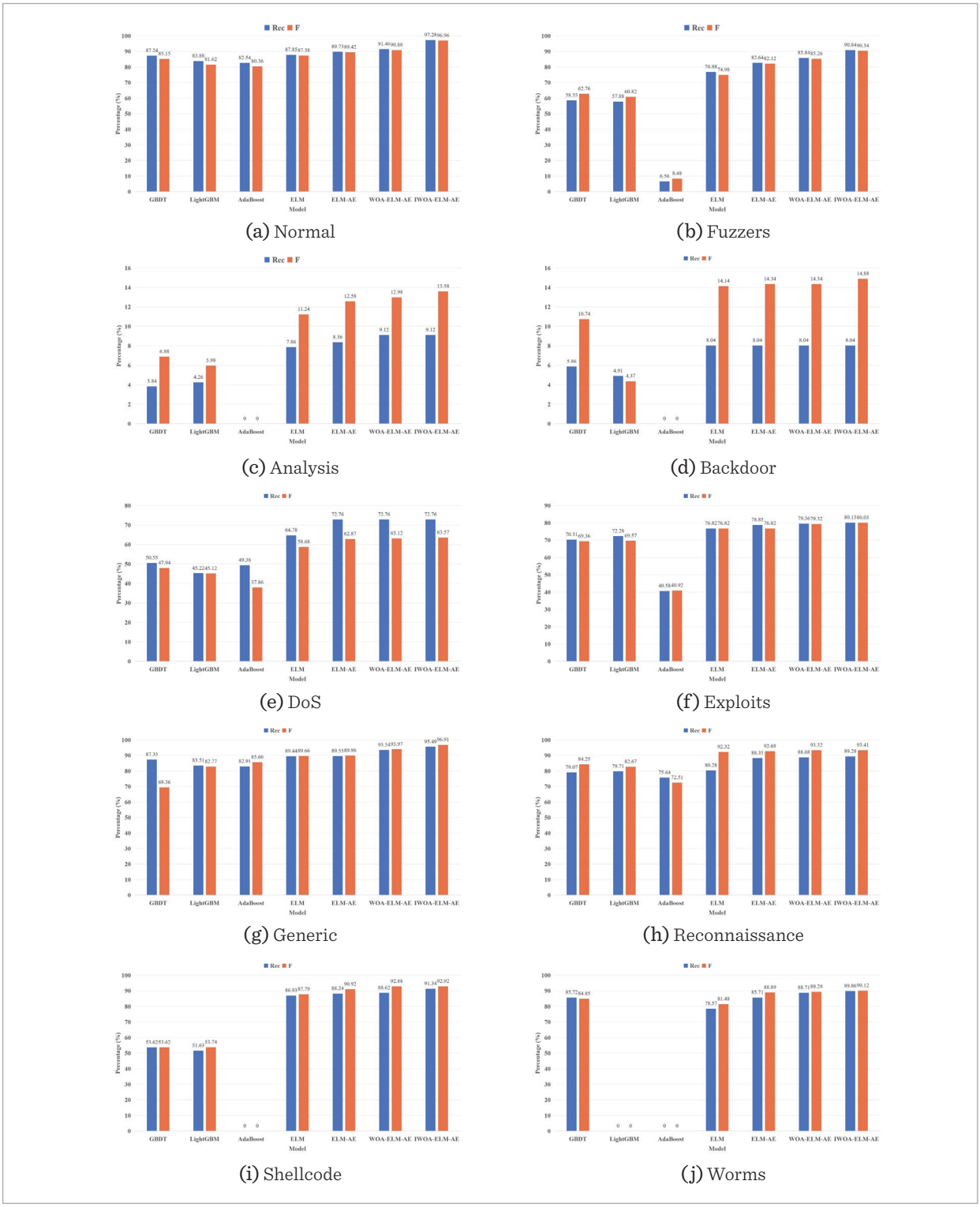
F-score (*F*) combines the accuracy and recall of the proposed early warning model and can evaluate the comprehensive performance of the algorithm by Equation (28)

$$F = \frac{2}{\frac{1}{Rec} + \frac{1}{Pre}} . \tag{28}$$

5.3 The Analysis of the Model Testing's Results

The UNSW-NB15 dataset was subjected to feature selection using the TSO algorithm. The machine learning models GBDT [30], Light-GBM [16], AdaBoost [28], ELM, ELM-AE, and WOA-ELM-AE are picked for side-by-side comparisons to validate the effectiveness of the proposed I-WOA-ELM-AE, as shown in Figure 2.

Figure 2
The classification results of the different models.



(a) Normal (b) Fuzzers

(c) Analysis (d) Backdoor

(e) DoS (f) Exploits

(g) Generic (h) Reconnaissance

(i) Shellcode (j) Worms

Seven distinct models are employed to execute multi-classification tasks on the UNSW-NB15 dataset, focusing on ten diverse categories. The outcomes of this classification are depicted in Figure 2. A thorough examination of the experimental data reveals that the proposed I-WOA-ELM-AE model demonstrates generally superior performance compared to conventional machine learning models in the context of the cybersecurity data utilized for testing. Notably, AdaBoost emerges as the least effective model in this comparison. Further analysis indicates that models leveraging the ELM-AE, particularly in the categories of Fuzzers, DoS, Exploits, Generic, and Reconnaissance, exhibit a distinct advantage in terms of Recall and F-score metrics. However, for categories such as Analysis and Backdoor, the performance of the models is adversely affected, likely owing to an imbalance in the distribution of data samples. This imbalance results in diminished efficacy, as exemplified by the AdaBoost algorithm, which registers a Recall and F-score of 0. The proposed I-WOA-ELM-AE model, while differing from other models, does not demonstrate exceptional performance in these categories.

In categories with smaller sample sizes, such as Shellcode and Worms, significant variability is observed among the machine learning models. This is particularly evident in the Worms category, where both GBDT and AdaBoost exhibit a Recall and F-score of 0. In contrast, the proposed I-WOA-ELM-AE maintains better performance, underscoring its stability even in scenarios of limited data availability.

Table 2 presents a comprehensive summary of the combined alert results generated by various algorithms for each type of network security data. The superior classification efficacy of the ELM-AE-based model over other models is highlighted. The proposed I-WOA-ELM-AE, following optimization, shows significant improvements across all indicators, markedly outperforming other models. Collectively, these findings suggest that the detection capabilities of the proposed I-WOA-ELM-AE are consistently high, thereby affirming its effectiveness in the field of cybersecurity data classification.

Table 1
Comprehensive test results for UNSW-NB 15.

Models	Rec	Pre	Acc	F-score
GBDT	62.84%	75.45	86.78	62.44
Light-GBM	50.76	52.24	84.25	52.06
AdaBoost	36.46	39.22	74.18	35.26
ELM	68.54	83.71	89.88	70.44
ELM-AE	69.77	87.65	90.67	70.93
WOA-ELM-AE	72.44	89.87	91.44	71.59
IWOA-ELM-AE	74.87	92.64	97.62	74.12

To further test the model performance, the average time used for 20 operations of the seven models was computed separately. The time statistics were achieved by using the program's automatic timing statements, and a comparative analysis of computational efficiency among various machine learning algorithms is presented. The results are shown in Table 1. The ELM algorithm demonstrates a notably expedited computation time of only 37.41 seconds. In stark contrast, algorithms such as GBDT, Light-GBM, and AdaBoost exhibit considerably prolonged computation durations, each surpassing the 300-second threshold. Further examination reveals that the ELM-AE, an enhanced version derived from the foundational ELM, records a computation time of 45.86 seconds. This duration, while marginally longer than that of the ELM, remains significantly lower than the aforementioned algorithms. Additionally, the integration of the WOA with the ELM-AE referred to as the WOA-ELM-AE, results in a computation time of 87.66 seconds. This outcome indicates a pronounced iterative delay when the WOA is employed for optimizing the ELM-AE. Subsequent improvements to the WOA, leading to the creation of the I-WOA, demonstrate a reduction in computation time when applied to the ELM-AE (I-WOA-ELM-AE), clocking in at 51.84 seconds. This reduced duration, in comparison to the WOA-ELM-AE, signifies a substantial enhancement in optimization efficiency attributed to the improvements in the WOA.

Even though the computation times of both I-WOA-ELM-AE and ELM-AE exceed that of the original ELM, it is imperative to consider their classification performance. The proposed I-WOA-ELM-AE, in particular, exhibits superior classification capabilities. When evaluating the comprehensive performance of these algorithms, it is evident that the enhanced versions, despite their increased computational time, offer considerable advantages, particularly in terms of classification accuracy and efficiency. This

suggests a favorable trade-off between computational speed and classification performance, positioning the enhanced algorithms as viable options in scenarios where accuracy is paramount.

Table 2
The Comparison of Computing Efficiency.

Models	Average time for 20 operations (s)
GBDT	318.6
Light-GBM	324.5
AdaBoost	365.2
ELM	37.41
ELM-AE	45.86
WOA-ELM-AE	87.66
I-WOA-ELM-AE	51.84

The proposed I-WOA-ELM-AE exhibits superior performance in terms of early warning for network security threats. It demonstrates a remarkable capability to achieve high accuracy and efficiency in its predictive capabilities. This enhanced accuracy and efficiency are crucial for timely and reliable early warning systems in network security issues, particularly in the context of rapidly evolving cyber threats. Moreover, the practicality of the proposed I-WOA-ELM-AE in industrial network security scenarios has been explored. In these environments, where the stakes of network security breaches can be particularly high, the proposed I-WOA-ELM-AE has shown commendable performance. Its capability to provide accurate and efficient early warnings makes it a valuable tool for maintaining the integrity and security of industrial networks. The effectiveness of the proposed I-WOA-ELM-AE in delivering precise early warnings, combined with its efficiency in processing and analyzing network security data, positions it as a significant advancement in the field of network security. Its application in industrial contexts further underscores its practicality and reinforces its relevance in addressing contemporary network security challenges. The findings suggest that the proposed I-WOA-ELM-AE represents a notable contribution to the development of a more robust and reliable early warning model of network security systems, particularly in an era where cyber threats become increasingly sophisticated and pervasive.

6. Conclusion

For real-time early warning requirements of industrial network security, a feature selection method for network security features based on the TSO algorithm is proposed, which reduces the redundant features of network security data and improves the functionality of machine learning models. To improve the accuracy and efficiency of the early warning system for network security issues, the ELM-AE is used to establish a basic warning model, and then I-WOA is used to realize the parameter optimization of the ELM-AE for the parameter selection problem of the ELM-AE, to design an early warning model for industrial network security issues, which result in a proposed model called I-WOA-ELM-AE. Finally, the effectiveness of the proposed algorithm is tested by experiments. The adopted strategy has both high early warning accuracy and efficiency, which has important practicality in the protection of industrial network security. In the future, we will further study the applicable security warning technology to complex interconnected industrial networks considering encryption, introduce blockchain technology to realize higher-level protection for network security issues, and further improve the level of industrial network security. The proposed model with 92.64% precision and 51.84 s average execution time excels over other methods.

Author Contributions

Conceptualization, methodology, software, resources, Le, X.; validation, formal analysis, investigation, data curation, Zhao, Y.; writing—original draft preparation, Le, X. Both authors have read the manuscript.

Funding

This work was supported by grants from the National Key R&D Program of China NO. 2021YFB3101700.

Data Availability Statement

Data sharing is not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Abdalzaher, M. S., Fouda, M. M., Elsayed, H. A. M., Salim, M. Toward Secured IoT-Based Smart Systems Using Machine Learning. IEEE Access, 2023, 11, 20827-20841. https://doi.org/10.1109/ACCESS.2023.3250235
2. Aydın, H., Aydın, G. Z. G., Sertbaş, A., Aydın, M. A. Internet of Things Security: A Multi-Agent-Based Defense-System Design. Computers and Electrical Engineering, 2023, 111, Part B, 108961. https://doi.org/10.1016/j.compeleceng.2023.108961
3. Azzam, M., Pasquale, L., Provan, G., Nuseibeh, B. Grounds for Suspicion: Physics-Based Early Warnings for Stealthy Attacks on Industrial Control Systems. IEEE Transactions on Dependable and Secure Computing, 2022, 19(6), 3955-3970. https://doi.org/10.1109/TDSC.2021.3113989
4. Bou-Harb, E., Assi, C., Debbabi, M. CSC-Detector: A System to Infer Large-Scale Probing Campaigns. IEEE Transactions on Dependable and Secure Computing, 2018, 15(3), 364-377. https://doi.org/10.1109/TDSC.2016.2593441
5. Chakkaravarthy, S. S., Sangeetha, D., Cruz, M. V., Vaidehi, V., Raman, B. Design of Intrusion Detection Honeypot Using Social Leopard Algorithm to Detect IoT Ransomware Attacks. IEEE Access, 2020, 8, 169944-169956. https://doi.org/10.1109/ACCESS.2020.3023764
6. Chen, X., Cheng, L., Liu, C., Liu, Q., Liu, J., Mao, Y. A WOA-Based Optimization Approach for Task Scheduling in Cloud Computing Systems. IEEE Systems Journal, 2020, 14(3), 3117-3128. https://doi.org/10.1109/JSYST.2019.2960088
7. Cui, D., Sun, G., Li, Y. Design and Implementation of Petrochemical Port Safety Monitoring and Early Warning System Based on Deep Learning Algorithm. 2023 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 2023, 1-5. https://doi.org/10.1109/ICDSNS58469.2023.10245276
8. Dong, G., Liao, G., Liu, H., Kuang, G. A Review of the Autoencoder and Its Variants: A Comparative Perspective from Target Recognition in Synthetic-Aperture Radar Images. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(3), 44-68. https://doi.org/10.1109/MGRS.2018.2853555
9. Dong, W., Wang, G., Yan, Q., Liu, Y. Design of Network Security Situation Awareness and Early Warning System Based on Big Data. 2023 International Conference on Networking, Informatics, and Computing (ICNETIC), Palermo, Italy, 2023, 749-753. https://doi.org/10.1109/ICNETIC59568.2023.00159
10. Fu, C., Zhang, L. A Novel Method Based on Tuna Swarm Algorithm Under Complex Partial Shading Conditions in PV System. Solar Energy, 2022, 248, 28-40. https://doi.org/10.1016/j.solener.2022.10.056
11. Hashmi, M. F., Bellare, T. B., Suresh, A., Naik, B. T. Football Event Classification Using Convolutional Autoencoder and Multilayer Extreme Learning Machine. IEEE Sensors Letters, 2022, 6(10), 1-4. https://doi.org/10.1109/LSENS.2022.3209366
12. He, Y., Wei, Y., Cao, J. Key Technologies of 5G Wireless Communication Network Physical Layer Based on Information Security Early Warning Model. 2023 XXXVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS), Sapporo, Japan, 2023, 1-4. https://doi.org/10.23919/URSIGASS57860.2023.10265619
13. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K. Extreme Learning Machine: Theory and Applications. Neurocomputing, 2006, 70, 489-501. https://doi.org/10.1016/j.neucom.2005.12.126
14. Janabi, A. H., Kanakis, T., Johnson, M. Convolutional Neural Network Based Algorithm for Early Warning Proactive System Security in Software Defined Networks. IEEE Access, 2022, 10, 14301-14310. https://doi.org/10.1109/ACCESS.2022.3148134
15. Kumar, P., Gupta, G. P., Tripathi, R., Garg, S., Hassan, M. M. DLTIF: Deep Learning-Driven Cyber Threat Intelligence Modeling and Identification Framework in IoT-Enabled Maritime Transportation Systems. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(2), 2472-2481.
16. Li, W., He, J., Lin, H., Huang, R., He, G., Chen, Z. A LightGBM-Based Multiscale Weighted Ensemble Model for Few-Shot Fault Diagnosis. IEEE Transactions on Instrumentation and Measurement, 2023, 72, 1-14. https://doi.org/10.1109/TIM.2023.3291742
17. Li, X., He, L., Zhu, Z., Wang, C., Shi, J., Du, G. Double-Scale Convolutional Autoencoder and Extreme Learning Machine for Parameter Identification of DC Bus Capacitor in Power Electronic Transformer. IEEE Transactions on Industrial Informatics, 2023, 19(8), 9102-9112. https://doi.org/10.1109/TII.2022.3224968
18. Liang, Z., Han, Q., Zhang, T., Tang, Y., Jiang, J., Cheng,

Z. Nonlinearity Compensation of Magneto-Optic Fiber Current Sensors Based on WOA-BP Neural Network. IEEE Sensors Journal, 2022, 22(20), 19378-19383. https://doi.org/10.1109/JSEN.2022.3205701

19. Neira, D., Araujo, A. M., Nogueira, M. An Intelligent System for DDoS Attack Prediction Based on Early Warning Signals. IEEE Transactions on Network and Service Management, 2023, 20(2), 1254-1266. https://doi.org/10.1109/TNSM.2022.3223881
20. Obayya, M., Arasi, M. A., Alruwais, N., Alsini, R., Mohamed, A., Yaseen, I. Biomedical Image Analysis for Colon and Lung Cancer Detection Using Tuna Swarm Algorithm with Deep Learning Model. IEEE Access, 2023, 11, 94705-94712. https://doi.org/10.1109/ACCESS.2023.3309711
21. Pour, M. S., Nader, C., Friday, K., Bou-Harb, E. A Comprehensive Survey of Recent Internet Measurement Techniques for Cyber Security. Computers & Security, 2023, 128, 103123. https://doi.org/10.1016/j.cose.2023.103123
22. Qais, M. H., Hasanien, H. M., Alghuwainem, S. Transient Search Optimization: A New Meta-Heuristic Optimization Algorithm. Applied Intelligence, 2020, 50, 3926-3941. https://doi.org/10.1007/s10489-020-01727-y
23. Reka, S. S., Dragicevic, T., Venugopal, P., Ravi, V., Rajagopal, M. K. Big Data Analytics and Artificial Intelligence Aspects for Privacy and Security Concerns for Demand Response Modeling in Smart Grid: A Futuristic Approach. Heliyon, 2024, 10(15). https://doi.org/10.1016/j.heliyon.2024.e35683
24. Wang, W., Jia, D., Xu, J., Huang, X. Challenges and Solutions for Network Security in the Information Age. 2023 International Seminar on Computer Science and Engineering Technology (SCSET), New York, NY, USA, 2023, 33-38. https://doi.org/10.1109/SCSET58950.2023.00017
25. Xu, J., Li, B. Uncertain Utility Portfolio Optimization Based on Two Different Criteria and Improved Whale Optimization Algorithm. Expert Systems with Applications, 2025, 268, 126281. https://doi.org/10.1016/j.eswa.2024.126281
26. Zeeshan, M., et al. Protocol-Based Deep Intrusion Detection for DoS and DDoS Attacks Using UNSW-NB15 and Bot-IoT Data-Sets. IEEE Access, 2022, 10, 2269-2283. https://doi.org/10.1109/ACCESS.2021.3137201
27. Zhang, J. Distributed Network Security Framework of Energy Internet Based on Internet of Things. Sustainable Energy Technologies and Assessments, 2021, 44, 101051. https://doi.org/10.1016/j.seta.2021.101051
28. Zhang, P.-B., Yang, Z.-X. A Novel AdaBoost Framework with Robust Threshold and Structural Optimization. IEEE Transactions on Cybernetics, 2018, 48(1), 64-76. https://doi.org/10.1109/TCYB.2016.2623900
29. Zhang, Y. Application Research of Computer Artificial Intelligence Technology in Network Security System. 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE), Jinzhou, China, 2023, 1058-1062. https://doi.org/10.1109/ICSECE58870.2023.10263436
30. Zhang, Z., Jung, C. GBDT-MO: Gradient-Boosted Decision Trees for Multiple Outputs. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(7), 3156-3167. https://doi.org/10.1109/TN-