Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.98549384

RELATIVE SPECTRAL FEATURE ANALYSIS-BASED CLONE ATTACK DETECTION AND ENHANCE ROUTING IN WIRELESS SENSOR NETWORKS USING ARTIFICIAL NEURAL NETWORKS

S.Bhuvana
Research Scholar, Department of Computer Science and Engineering, Dr.MGR Educational and Research Institute, E-mail: bhupreethi@gmail.com

Dr.S.Kevin Andrews
Associate Professor, Department of Computer Applications, Dr.MGR Educational and Research Institute, E-mail:kevin.mca@drmgrdu.ac.in

Dr.M.S.Josephine
Professor, Department of Computer Applications, Dr.MGR Educational and Research Institute, E-mail:Josephine.mca @drmgrdu.ac.in

Dr.V.Jeyabalaraja
Professor, Department of Computer Science and Engineering, Velammal Engineering College,Chennai, E-mail: jeyabalaraja@gmail.com

Abstract

Wireless sensor network is recent trend development in remote technologies for ubiquitous computing in various applications, like monitoring, sensing, geo location-based applications. By the sense the transmission nodes are most probably affected by close attacks leads devastating communication defects. From the communication, ensuring security is the important aspect for communication nodes to protecting the data transmission without any malicious attacks to resolve this problem, we propose an advanced transfer learning model to identifying the clone attacks based on Neural Fuzzy intensive-Sub spectral scaling feature selection (NFI-SSFS) to secure using Cooperative Secure Optimal Link Stability Routing Allocation (CS-OLSR). The communication logs is collected to consume the variance feature level of packet difference rate under memory and transmission defect fact with sport of False injection impact rate (FIIR) and Time stamp communication behaviour rate (TSCBR). Then the intensive feature factor is obtained using NFI-SSFS to marginalize the clone attack rate. Then CS-OLSR is applied to ensure the secure routing based on the identified clone attack region. The proposed system effectively identifies the data replacement node effetely to find the clone attacks. This system makes effective clone attack route finding approach to ensure the security to transfer another route to make data safely.

Keywords: Attack detection, Wireless mobile network, Machine learning, ANN, node behaviour analysis, feature selection and classification

1. Introduction

Wireless Sensor Networks (WSNs) integrate sensor nodes with great potential. The main task of a sensor node is to monitor and sense the deployment area, use sensors from the environment to collect data, process data, and communicate with further nodes. This type of

harmful attack, where single or multiple nodes illegally declare their identities as duplicates, is called a node duplication attack [1-2].WSN works as an infrastructure-less network and mobile nodes can communicate with each other through radio links. WSNs are highly susceptible to node replication attacks (cloning attacks). It is possible to replicate multiple clones and deploy clones in different network locations to find the same Identity (ID) from a compromised node. Also, an attacker can compromise a sensor node. Clones can be used to validate all credentials for a member when the member appears [3]. Characteristics of cloning attacks: 1) There are defects in the hardware architecture, and the resources of physical capture and destruction are limited so that the attack can be launched quickly; 2) Various insiders and it is considered as a severe destructive threat leading to outside attacks [4].

However, existing solutions can develop and overcome wormhole attacks. Also, it requires more hardware, has higher delivery latency, fails to provide higher throughput and packet delivery rates, and consumes more power. Additional vulnerabilities are exposed if an attacker can inject attack data while accessing (eavesdropping) critical industrial data [5]. This will have disastrous consequences, and the latter situation will be severe in the industry [6].

Industrial wireless cyber-physical schemes are susceptible to malicious node attacks such as clone node attacks. Existing clone detection schemes are based on upper layer monitoring or physical layer channel state information. Schemes based on high-level observations are easily discredited, While schemes based on channel state information perform well against slander, they are heavily influenced by channel conditions. The physical layer uses reputation and backpropagation neural network for clone detection to improve detection accuracy. The proposed scheme accumulates physical layer reputation through channel-level information and feeds it to a neural network. The cloud server first performs attack detection through group detection. If a group is classified as attacked, the corresponding edge processor conducts attack monitoring to identify specific clone nodes. In the attack source tracking phase, multiple reputations of each node are used for detailed detection. Extensive testing was performed on the Universal Software Defined Radio Peripheral Platform. Numerical results show that this scheme significantly improves the detection accuracy.

Wireless sensor networks (WSNs) are often used in harsh environments, where attackers can physically capture some nodes, first reprogram them, and then replicate them in multiple clones, so that the network can be easily controlled. Several distributed solutions have recently been proposed to address this fundamental problem. However, these solutions are not satisfactory. First, they are energy and memory demanding: a serious drawback for any protocol used in the resource-constrained environment of WSNs. Additionally, they are susceptible to the specific enemy models presented in this paper. First, we will analyse the desirable properties of distributed algorithms for detecting node duplication attacks. Second, we demonstrate that known solutions to this problem do not fully meet our needs. Third, we propose a new self-correcting, Randomized, Efficient, and Distributed (RED) protocol to detect node duplication attacks, which we demonstrate satisfies the introduced requirements. Finally, detailed simulations show that our protocol is efficient in terms of communication, storage, and computation; are more efficient than competing solutions in the literature; and are resistant to the novel attacks presented in this paper, but other solutions are not.

But all the works presented above use the results of independent classifiers without emphasizing the dependence of combined classes. Indeed, under the assumption of gaps between classes, attack detection is easier to implement, faster to evaluate, and to reduce the amount of training data needed to evaluate attacks. Therefore, including feature selection and classifier techniques can achieve better performance. Following this trend, our work tends to combine a classifier with three feature selections (protocol, service, and flag). In our work, The EELAPs and artificial neural networks (ANNs) achieve higher classification accuracy than other classifier models, but their applicability is limited due to long training time for large datasets. Therefore, multiple feature selection techniques are integrated with EELAP and ANN classifiers to obtain accurate attack detection results.The contribution of this study is to propose detection of cloning attacks using channel-based machine learning. To identify malicious attacks, channel responses between sensor peers are explored as a form of spatially and temporally distinct fingerprints. In addition, machine learning-based methods are used to provide more accurate certification rates. Specifically, by connecting to devices at the edge.

We adopt a channel-difference based threshold detection method to provide machine learning algorithms with a labelled set of offline training samples, avoiding manual label generation. Therefore, our proposed scheme is lightweight for resource-constrained industrial wireless devices because only online results are required. Extensive simulations and tests were performed in real industrial environments. Both results show that a recognition accuracy of 84% can be achieved at a suitable threshold without human labelling.

2. Related work

The author proposed that adversaries can intercept legitimate nodes and extract stored credentials such as identities by deploying Wireless Sensor Networks (WSN) in remote and harsh environments. However, low-cost sensor nodes are vulnerable to node cloning and duplication attacks due to inherent characteristics, such as a lack of memory, batteries, and tamper-proof hardware [7]. The author proposed that a channel-based Machine Learning (ML) model can be used to detect cloning and Sybil attacks. The spatially and temporally differentiated fingerprints can be analyzed to classify replies to channels between sensor peers [8]. The author proposed that human cognitive processes' evolutionary self-cooperative trust (ESCT) scheme reflects belief-state information to prevent various derivative attacks. However, these appearances make designing routing protocols for MANETs challenging [9].

The author proposed that the cloned node has a different physical location, but the requested identity conflicts with the captured node. Spatial differences can detect cloned nodes and then be tracked using physical layer Channel State Information (CSI). However, despite the computational complexity, conventional cloning attacks in sensor systems are a challenging problem to detect with cryptographic methods [10]. The author proposed that the Bacterial Aging Optimization Algorithm (BFOA) can be used as a belief-based safety, energy-efficient navigation algorithm. The optimal hop number for routing optimization can be found and provide additional estimates for MANETs. However, mobile node power outages affect the node's ability to transmit packets and depend on the overall lifetime [11]. The author proposed a preliminary survey of program selection criteria based on device type, detection method, deployment strategy, and detection coverage by Clone. Also, the requirements of the introduced existing methods can be classified [12].

The author proposed an energy-efficient location-awareness protocol (EELAP) in densely deployed wireless sensor networks using Clone. It can guarantee successful detection of clone attacks and maintenance of moral network lifetime [13]. The author predicts that WSNs can often be used in adverse environments. Also, an adversary can physically capture a given node and then reprogram and clone it into multiple clones to quickly gain control over the network [14]. The author proposed that different scenes could be created by changing the movement (position) of the nodes. Performance metrics used in performance analysis include throughput, end-to-end latency, and packet transfer rate [15]. The author recommends using multi-path routing in wireless networks to progress the unique single-path routing and burden tolerance. Routing Protocol Genetic Algorithm with Hill Climbing (GAHC) is clearly defined as given that a hybrid GAHC algorithm that can select the best route among various routes [16].

The author overcomes the shortcomings of wireless sensor network clone attack detection methods. The MSCD method can be implemented as a multi-based wireless sensor network clone detection for low-resource consumption [17]. The author proposed implementing a lightweight one-shot hash using a Counterfeit Clones (CC) scheme to protect the location privacy of data link layer nodes by masking the MAC address. [18]. The author proposed a distributed Low-Storage Clone Detection (LSCD) protocol for WSNs. A detection path can be designed in the vertical direction of the observation path using observation nodes arranged along the loop path. WSN can counter the threat of cloning attacks and control the attacking network by conducting various attacks [20]. Most cases cluster ensemble approach are implemented with Fuzzy logic based on ANN to resolve the detection problem [21]. In addition MANET resources Clusters are implemented bases on feature selection and classification model [22, 23]. The author proposed that establishing IDS and creating clone detection routes can be implemented in a ring-structured network without hot spots to balance resource consumption.

3. Proposed methodology

Towards the development identifying the clone attacks based on Neural Fuzzy intensive-Sub spectral scaling feature selection (NFI-SSFS) to secure using Cooperative Secure Optimal Link Stability Routing Allocation (CS-OLSR). Then the intensive feature factor is obtained using NFI-SSFS to marginalize the clone attack rate. The feature analysis takes importance to analyse the Non relevance factor. This gives importance to spectral margins which covers the specific threshold margins actively compared with each other. The communication logs is collected to consume the variance feature level of packet difference rate under memory and transmission defect fact with sport of False injection impact rate (FIIR) and Time stamp communication behaviour rate (TSCBR)

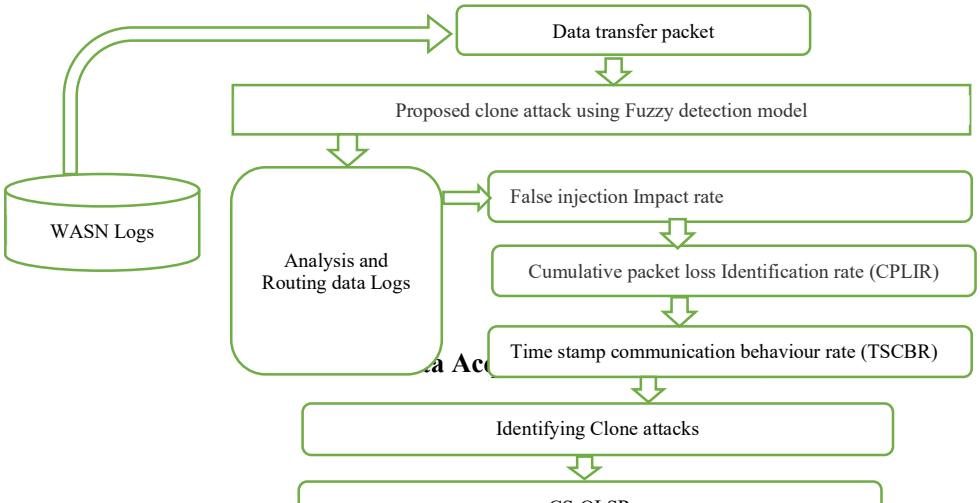

Figure 1: Proposed system architecture diagram NFI-SSFS- CS-OLSR

The training and testing result are carry on the fact of filtered margin features which is actively diffused to carry on the ideal margins. Figure 1 shows proposed system architecture diagram NFI-SSFS- CS-OLSR. Then CS-OLSR is applied to ensure the secure routing based on the identified clone attack region. The proposed system effectively identify the data replacement node effetely to find the clone attacks. These features are splinted into class by reference to identify the risk. Initial stage the route table and track the behaviour information are marginalized into active values. The ideal values are comparatively posed with ideal and active margin comparison to improve the defection dataset.

3.1 False injection Impact rate

By analysing the clone defect, we carries out packet variation defect rate carried out identify the packet replacement, packet injection, memory contamination level feature are monitored during communication, the absolute mean square error rate is identified through the variation feature limits. The estimation of transmission Feature importance denotes to a class for assigning weight to the input features to find the clone behaviour analysis model. This method identifies the False data injected during the data transition which pointes the memory and intrusion descriptive mode of feature variation weight. Therefore the most crucial feature weight is calculated by the prognosis of the pre-processed data set.

Step 1: Find the Expected readings by:

expected_sensor_values=f(Sensor Values,Network state)

Step 2:Find the False data injection

To inject the false data into the sensor node by flipping the values of randomly selected sensors.

- Generate a list of random numbers:
- $random_{numbers[i]} = r \; where \;\; r \sim U\,(1, len(sensor_{values})$ (1)

  r ~ U(a, b)-→ represents a random variable r following a uniform distribution between a and b, inclusive.
  The notation $random_{numbers[i]}$ = r denotes that the random number r is assigned to the i-th position of the random_numbers list.
- Iterate over the random numbers and flip the corresponding sensor values:
  - If the sensor value is 0, set it to 1:

    If false_sensor_values[$random_{numbers[i]}$] == 0:
    false_sensor_values[$random_{numbers[i]}$] = 1
  - If the sensor value is 1, set it to 0:

    If false_sensor_values[$random_{numbers[i]}$] == 1:

false_sensor_values[$random_{numbers[i]}$ **] = 0**

- Return the false sensor values: **false_sensor_values**

Step 3:Find the impact for false data:

- To find the impact of the false data injection by measuring the difference between the expected and false sensor values.
- Calculate the absolute difference between the expected and false sensor values:

diff[i]=abs(expectedsensor_values[i]-false_sensor_values[i] (2)

- Calculate the maximum difference:

max_diff = np.max(diff) (3)

- Return the maximum difference as the impact of the false data injection: **return max_diff**

Step 4: Find the impact_of_noise:

To find the impact of noise in the sensor readings by measuring the standard deviation of the sensor values.

$$m = \frac{\Sigma(sensor\_values)}{n} \quad (4)$$

$$std_{dev} = \frac{\sqrt{\Sigma(x-m)\hat{}2))}}{2n} \quad (5)$$

x represents each element in the sensor_values array,
m is the mean (average) of the sensor_values array, calculated as mean = sum(sensor_values) / n,
n is the total number of elements in the sensor_values array, and
sqrt is the square root function.

- Return the standard deviation as the impact of the noise: std_{dev}
- noise_impact= std_{dev}

Step 5: The False Injection Impact Ratio:

- This function calculates the FIIR (False Injection Impact Ratio) value by combining the above functions.
- Calculate the FIIR value as the ratio between the impact of the false data injection and the impact of the noise:

fiir_value = false_data_impact / noise_impact (6)

This section proficiently identifies the packet variation using False Injection Impact Ratio (FIIR) technique.

3.2 Cumulative packet loss Identification rate (CPLIR)

In this step, we evaluate the IDS weight according to the path's congestion distribution flag ratio and assess whether each communication packet's delay is designed for a negative loss rate. These characteristics depend on the traffic flow and the packet transmission during the communication rates. CPLIR values measure different factors related to packet flow rate

transmission. Packet flow dependency estimates are based on the overall features of the entire dataset.

Input: User Access Trace UaT, User Profile Usp.

Output: Security Taxonomy ST

Step1: Initialize the data logs

Step2: Read UaT and UsP

Step3: Compute process trust Set TSeS.

Step4: Compute the service level trust For each access trace Ti

Identify list of attributes accessed by the service Als.

Als = ∑_(i=1)^(size(UaT)) ⟦UaT(i).User==Usp(j).User && Attributes∈Usp(j) ⟧

Compute Attributes sanctioned As = ∑_(i=1)^(size(Ar)) ⟦Ar(i)∈Usp(U)⟧

Identify Trust level attributes Ta = Als ∩As

Add to trust set TSes = SeS ∪Ta -- (10) End

Step 5: Compute the Trust level from access rate For each attribute Ai

Compute total number of access Tna = ∑_(i=1)^(size(UaT)) ⟦ UaT(i).Service→Ai⟧ Compute access frequency Af = Tna⁄(size(UaT))

End

Step6: Sort attributes according to access frequency.

Step7: Compute minimum frequency Mf = Min (TSeS (Ai).Af)

Step 8: Compute maximum frequency Mxf = Max (TSeS (Ai).Af)

Step 9: Initialize Number of class Nc.

Step 10: Split minimum and maximum frequency

The outlier detection feature indicates that edges have a higher influence ratio measure. This may affect the interdependence of the service level integrity ratios.

3.3 Time stamp communication behaviour rate (TSCBR)

This degree examines exchange data pricing between supply and vacation destinations that rely on relative power to marginalize penetration pricing. As this technique keeps records belonging to specific time windows, the collection of individual information gains entry rights. Once the accumulation degree and achievement price are measured for all time windows, a guideline of thumb can be developed primarily based on various values.

The packet variation in transmission time evaluates the time dependency factor based on the user log activities to trance the domain factor in communication window. It carried the difference along the realization state of delay facts in packet exchange medium to consider as clone.

Algorithm

Input: Collective logs- pre-processed Ps dataset

Output: Time stamp communication behaviour rate

Step1: Initialize Class reference Nc

Step2: Find access variation user rate Ur and packet log l, Packet transmission service S ←ps.

Step3: Create class by variation of transmission from sensor node variation

Step4: Substitute packet call log Twl = $\sum_{i=1}^{size(data)} Ur(i).packet == Di \;\&\&\; Ur(i).Time = Tw_{\mathrm{I}}$

Step5: For all log access trace on clone behaviour on packet Level Twl

Step6: Verify the user behaviour variation in transition level

Step7: Trace call sensor node Dt = $\sum_{i=1}^{size(Twl)} Ur(i).packet == Di$

Step8: Process the packet attain factor on size variation

Sr = $Packetsize(Si)$

Step 9: Calculate occurrence quantity.

Step10: Variance packet Quantity OQ = $\frac{size(Twl).Sr}{size(Dt)}$

Step11: Estimate the support route rate

Step12: Support rate SRr = $\frac{\sum_{i=1}^{size(OQ)} Di..Status==Succe}{size(Twl)}$

Step13: retain Node reality Nr←SRr

Step 14: Modify the route support rate

Step 15: Stop

The above algorithm process the time stamp factor by accessing the time factor based clone attack through packet difference. During the frequency in packet variation level difference is carried out by packet replace level to retain the node reality to modify the node support level.

3.4 Neural Fuzzy Intensive-Sub spectral scaling feature selection (NFI-SSFS)

At this point, the behavioural characteristics of the network by the clone attackers are seen through the service-optimized access control. It estimates the average weight of relevant terms contained in the Memory contaminated features. Sensor node behaviour and transmission structures allow you to quickly discover relationships between features. $X = \{x1, x2, \cdots, xn\}$ And $Y = \{y1, y2, \cdots, ym\}$ the entropy be calculated concerning the user behavioral actions to be calculated by input interactive Ai and Aj features relation. The Experimental Feature (EF) shows to comparing below

$$Ef(x) = \sum_{i}^{n} \begin{matrix} ni\ (Ai) == 1\ which\ is \\ Equals\ to\ another\ input\ I(xi) \log 2i(x2), \end{matrix} \quad (7)$$

$$Ef(y) = \sum_{j}^{m} \begin{matrix} mi(Aj) == 1\ which\ is \\ Equals\ to\ realtional\ input\ seriesI(yi) \log 2I(Y2) \end{matrix} \quad (8)$$

$\begin{vmatrix} x_1y_1 & x_1y_2 & \dots & x_ny_m \\ I(x_1y) & I(x_1y_2) & \dots & I(x_ny_m) \end{vmatrix}$, The Attaching feature data connection among the ways of behaving is the entropy value of An and B.

The independent feature from A and B are the Relative Feature (RF) defined as

$$RF\ (x, y) = E(X) + E(Y) - E\ (xy) real\ Latticeof\ user\ behavioral\ (0, 1) joint\ relative\ features \quad (9)$$

Relates feature of independent relations of X and Y from sources of feature user behavior dependencies of joint relation. These composite relational properties are considered to have removed features that do not require further classification.

ANFIS, or Adaptive Neuro-Fuzzy Inference System, integrates the strengths of Artificial Neural Networks (ANNs) and Fuzzy Logic (FL) within a unified framework. This powerful approach combines the ability to rapidly learn from data with the adaptive interpretability needed to capture intricate patterns and understand nonlinear relationships.

Fuzzy rules can be used to capture the relationships between these factors and identify potential clone nodes.

Fuzzy rules of clone Rule 1: IF (Node Density is High) AND (Node Energy Level is Low) THEN (Node is a Potential Clone)

Rule 2: IF (Node Distance to Neighbors is Small) AND (Node Communication Overhead is High) THEN (Node is a Potential Clone)

Rule 3: IF (Node Communication Pattern is Similar to Other Nodes) AND (Node Location is Suspicious) THEN (Node is a Potential Clone)

Rule 4: IF (Node Behavior Deviates from Normal) AND (Node Received Signal Strength is Abnormal) THEN (Node is a Potential Clone)

Rule 5: IF (Node has Duplicate ID) AND (Node is Transmitting on Same Frequency) THEN (Node is a Potential Clone) .

3.5 Cooperative Secure Optimal Link Stability Routing Allocation (CS-OLSR)

After the clone fact considers the relative feature variation the cooperative is used to improve the security. By defining R is the route in Lookup transmission which contains all the energy constraints nodes in multicast network T(s, D) i.e., R→ source's' and destination 'D' create a cluster group set Is represented as $C_G \rightarrow \boldsymbol{D} \subseteq \{\boldsymbol{R} - \{\boldsymbol{s}\}\}$.The Relation lookup transmission R_T (s, D)→ $D \subseteq \{R - \{s\}\}$ by considering the minimum energy at delay constraints features level to make efficient transmission without clone point of attack region,

$$Energy\ \big(p(s,d)\big) = \min\{R_T(e), e\ \in p(s,d)\} \quad (10)$$

$$Traffic\ \big(p(s,d)\big) = \sum_{e\in p(s,d)} delay(e) + \sum_{e\in p(s,d)} traffic(n) \quad (11)$$

$$Route\big(T(s,D)\big) = \sum_{e\in p(s,d)} Dist(e) + \sum_{e\in p(s,d)} Tolerance(n) \quad (12)$$

The network be optimized with intension of route management by considers the energy enhancement in each transmission having the duty cycle. The presence of nodes from Source S at $.n \in R$ Belongs with intermediate clusters. The dynamic transmission be updated on RT having the energy consumption rate at K- number of clusters.

$$E_h = lE_e + l\varepsilon_s d^2 + \left(\frac{n}{k} - 1\right) l(E_e + E_{BF}) \quad (13)$$

The sensor nodes having the energy consumption with low level latency is considered and non-cluster formation active nodes are avoided during the transmission.

$$E_{nh} = lE_e(1+k) + l\varepsilon_s d^2 + (klE_{BF}) \quad (14)$$

The maximum support energy consumed by the cluster Head is depends on the transmission range and density of the medium. The Lookup finds the minimal congestion to route the transmission.

$$E_d = lE_e + l\varepsilon_l d^2 + \left(\frac{n}{k} - 1\right) l(E_e + E_{BF}) \quad (15)$$

The non-residual energy consumption nodes is discarded from the routing and relay constraints is considered to add-on nodes to balancing the routing in each duty cycle.

$$E_{nd} = lE_e + l\varepsilon_s d^2 \quad (16)$$

The initial value:

$$\varepsilon_s = \frac{\frac{10pJ}{bit}}{m^2}, E_e = \frac{50nJ}{bit}, \quad (17)$$

$$\varepsilon_l = \frac{\frac{0.001pJ}{bit}}{m^4}, E_{BF} = 5nJ/bit \quad (18)$$

The lookup transmission find the energy level density during the dynamic propagation, each QoS services are considered into absolute transmission mean rate. The mean rate considers the minimum energy at data bits rate handled 20 nodes averages taken mean transmission range 100*100 at RT transmission in single broadcast medium takes the relational medium.

The minimal energy consumption of an actual rate

$$E_{Tx}(k,d) = min \begin{cases} E_{elec} \times k + \in_{fs} \times k \times d^2 \ , if \ d < d_0 \\ E_{elec} \times k + \in_{mp} \times k \times d^4, if \ d \geq d_0 \end{cases} \quad (19)$$

$$d_0 = \sqrt{\left(\frac{\in_{fs}}{\in_{mp}}\right)} \quad (20)$$

$$E_{Rx}(k) = E_{elec} \times k \quad (21)$$

The sensor elective consumption is $E_{elec,}$ in look up transmission on each amplifying unit in sliding window as refer duty cycle talks c_{mp} in multipath cooperative transmission medium. This improves the optimized transmission at equalized emery path in maximum probability to improve the life time of the network. This CS-OLSR based energy optimization will further reduce the energy and execution time of network. Hence, our projected CS-OLSR based routing mechanism can be efficient and effective offloading scheme to improve the clone attack detection based secure routing in WSN

4. Result and discussion

The proposed implementation is tested on python language with an anaconda environment using publicly available clone darknet dataset. Clone attack can be effectively detected by comparative parameters such as classification accuracy, sensitivity, specificity, false ratio and time complexity with the help of a confusion matrix.This volume describes descriptively the results and discussions of the proposed methods. And in this, 30 services are used, according to which the parameters are shown in Table 1.

Table 1: Environment and values processed

Parameters	Values
Simulation Tool	Anaconda, Jupyter notebook
Simulation language	Python
Name of the dataset	Clone Darknet dataset
No of users/ records	500/ 2500
Number of classes	High / medium / low

The comparison algorithms are EELAP, GAHC, and LSCD carried out based on Multi-Factor clone attack Detection System. The following parameters are calculated by the confusion matrix.

Dataset Description: Darknet datasets contain many types of properties, including Non-Numerical data that needs to be converted to numerical form for processing. Perform data transformation to convert non-numeric data of categorical features to numeric format. The value of the Darknet dataset includes the protocol type. Protocol categories are converted into the numeric format by assigning numbers to individual protocol categories.

: IP	Src Port	Dst IP	Dst Port	Protocol	Timestamp	Flow Duration	Total Fwd Packet	Total Bwd packets	...	Active Std	Active Max	Active Min	Idle Mean	Idle Std	Idle Max	Idle Min	Type	packets	target
.11	57158	216.58.220.99	443	tcp	24/07/2021 04:09:48 PM	0	1	1	...	20	20	20	20	20	0	0	Non-Tor	AUDIO-STREAMING	normal
.11	57159	216.58.220.99	443	udp	24/07/2021 04:09:48 PM	0	1	1	...	20	20	20	20	20	0	0	Non-Tor	AUDIO-STREAMING	normal
.11	57160	216.58.220.99	443	tcp	24/07/2021 04:09:48 PM	0	1	1	...	20	20	20	20	20	0	0	Non-Tor	AUDIO-STREAMING	risk
.11	49134	74.125.136.120	443	tcp	24/07/2021 04:09:48 PM	0	1	1	...	20	20	20	20	20	0	0	Non-Tor	AUDIO-STREAMING	risk
.11	34697	173.194.65.127	19305	tcp	24/07/2021 04:09:48 PM	0	1	1	...	20	20	20	20	20	0	0	Non-Tor	AUDIO-STREAMING	risk
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
.11	46461	10.152.152.10	53	tcp	24/02/2018 02:24:58 PM	1	0	0	...	1	1	1	1	1	1	1	Non-Tor	Browsing	normal
.11	53984	54.169.125.186	80	udp	24/02/2018 02:24:58 PM	1	0	0	...	1	1	1	1	1	1	1	Non-Tor	Browsing	normal

Figure 2: Darknet dataset features

In the above figure 2 shows the Darknet dataset after analysis the preprocessing step remove null values or unwanted values from the dataset. Utilized the current dataset feature sequential element and $max\,relational\,feature(F_w)$ is utilized to maximum relational features find to threshold values and calculate the feature weight in the order.

Table 2: Dark net Dataset Parameters

Standard Parameters Taken	Active state parameters
Source IP: Source IP Address	Fwd PSH Flags
Source Port: Source Port	Bwd PSH Flags

Destination IP: Destination IP Address **Destination Port:** Destination Port	Fwd URG Flags Bwd URG Flags
Timestamp: Timestamp for when traffic was sent	Fwd Header Length Bwd Header Length
Protocol: Internet Protocol Version	Fwd Packets/s Bwd Packets/s
FrwdPacketLengthMax FrwdPacketLengthMin FrwdPacketLengthMean FrwdPacket Length Std	Subflow FrwdPackets Subflow FrwdBytes Subflow Bawd Packets Subflow Bawd Bytes
BwdPacketLengthMax BwdPacketLengthMin BwdPacketLengthMean BwdPacket Length Std	FWD Init Win Bytes Bwd Init Win Bytes Fwd Act Data Pkts Fwd Seg Size Min
Flow Bytes/s Flow Packets/s	Active Mean Active Std Active Max/ Active Min
Protocol: Internet Protocol Version	SYN Flag Total RST Flag Total PSH Flag Total ACK Flag Total URG Flag Total CWE Flag Total ECE Flag Total Down/Up Ratio

Table 2 described, Standard and active parameters are taken for the Dark net dataset. Dataset values are evaluated in the language of python using the framework Anaconda. It is a predicting tool for data. And the best feature value estimation is divided into training and testing using deep learning feature selection techniques. Darknet dataset features are efficient for estimating the IDS from traffic and improving the detection accuracy.

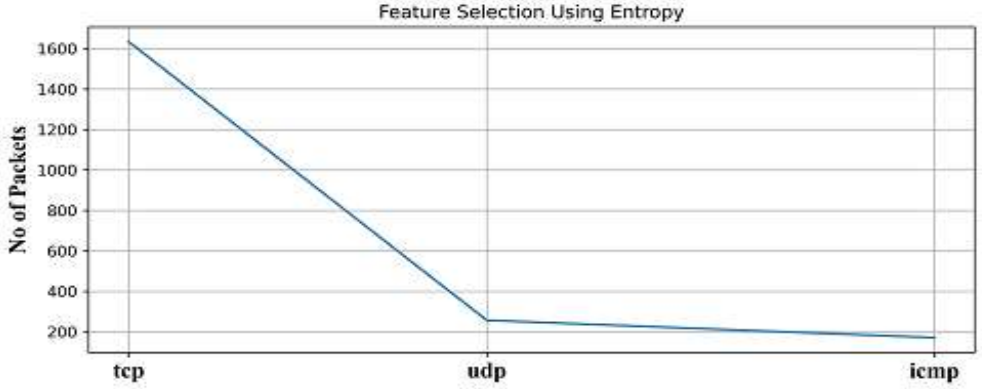

Figure 3: Screenshot feature Selection using Entropy

In the above figure 3, analysis the features selection using CS-OLSR optimization based on transmission protocols and using this algorithm for weightage (using no. of packets using Type, Timestamp, protocols etc.) calculation. In the features selection particularly selected values for packet streaming and features to take during transmitting the packets. This analysis the maximum range dataset values and how many percentage to considered take unique features

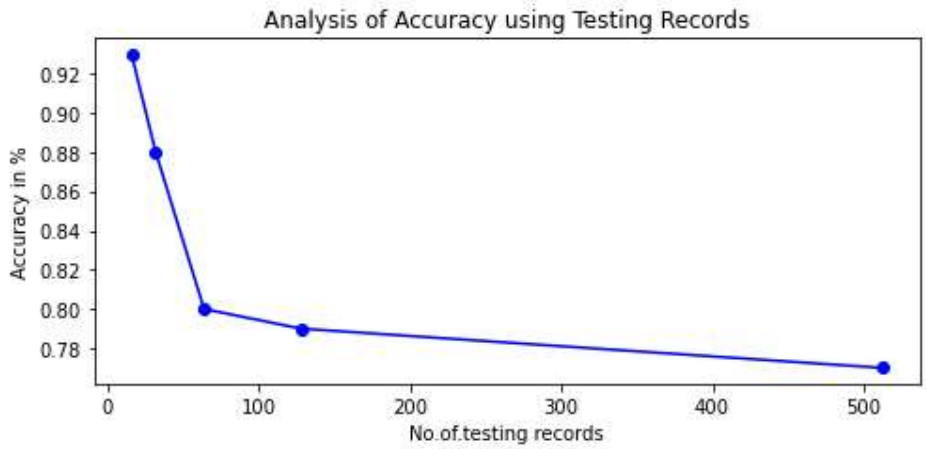

Figure 4: Analysis of Accuracy

Figure 4 described as, Accuracy of comparison method using dark net dataset and evaluating the results based on the TP, TN, FN and FP ranges. In the proposed method CS-OLSR shows the accuracy level is 0.93%, comparing the other methods are LSCD is 0.88%, GAHC is 0.80% and EELAP is 0.77%. Proposed method shows the better accuracy compared to the previous methods.

Table 3: Performance on clone node accuracy vs. no of services

Clone node Detection Accuracy in % vs No of Services			
Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
EELAP	70.9	73.6	78.3
GAHC	76.2	77.1	81.8
LSCD	78.7	82.4	87.5
CS-OLSR	91.9	96.6	95.8

Table 3 describes the IDS accuracy performance vs no of services with different techniques like EELAP, GAHC, LSCD and the proposed Ids based feature Analysis Model (CS-OLSR).

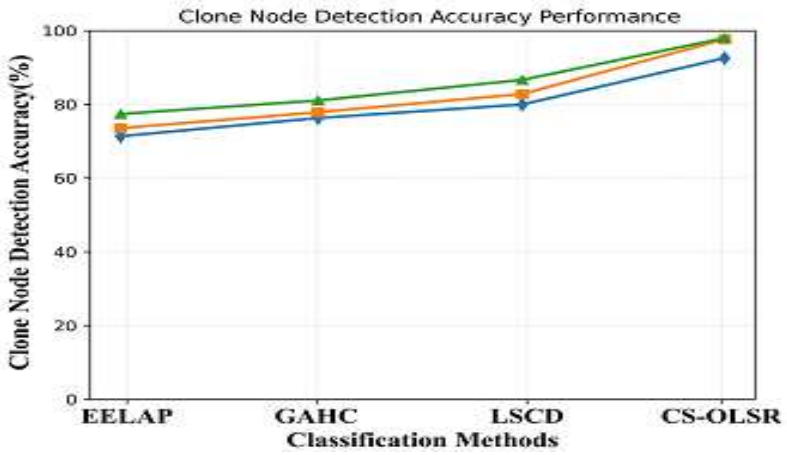

Figure 5: Impact of Clone node detection accuracy performance

Figure 5 denotes impact of intrusion detection classification accuracy performance with various services like 10, 20 and 30. The proposed technique CS-OLSR method attained 98.3% for 30 services also the previous EELAP attained 78.3%, GAHC 81.8%, and LSCD attained 87.8%. Nonetheless, the proposed method produces better performance than other techniques.

Table 4: Impact of Sensitivity performance

Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
EELAP (%)	68.6	71.8	77.1
GAHC (%)	73.2	75.6	80.7

LSCD (%)	76.7	81.5	86.2
CS-OLSR (%)	89.3	94.5	95.1

Table 4 describes the impact of sensitivity performance the proposed compared with previous techniques.

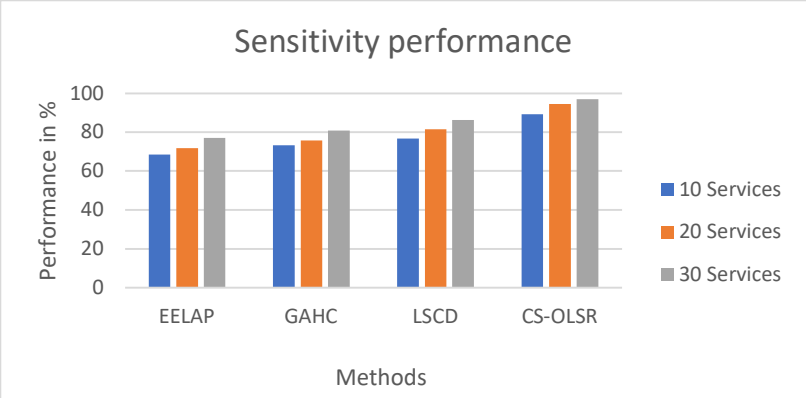

Figure 6: Analysis of sensitivity performance

Figure 6 shows the sensitivity performance for IDS detection using CS-OLSR algorithm. The proposed algorithm provide result is 96% of sensitivity performance for 30 services; similarly the exiting algorithm provide results are EELAP is 76% of Sensitivity performance, GAHC is 81% of Sensitivity performance and LSCD is 85% of Sensitivity performance for 30 services.

Table 5: Impact of Specificity performance

Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
EELAP (%)	71.2	72.6	78.4
GAHC (%)	78.7	76.8	83.6
LSCD (%)	79.4	82.8	87.4
CS-OLSR (%)	90.8	95.2	96.8

Table 5 describes the analysis of specificity performance measures in different number of services such as 10, 20, and 30 services. The proposed technique provide better result than previous approaches.

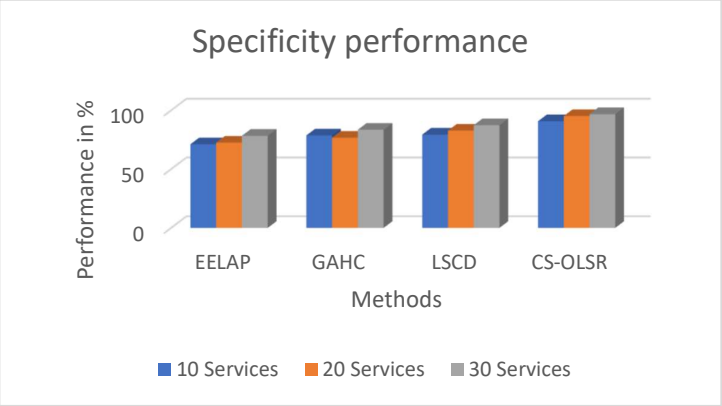

Figure 7: Analysis of Specificity performance

Figure 7 illustrate the analysis of specificity performance the proposed and previous approaches comparison result presented. The proposed CS-OLSR algorithm has 97% of Specificity performance for 30 services; similarly the existing algorithm results are EELAP is 77% of specificity performance, GAHC is 82% of specificity performance and LSCD is 86% of specificity performance for 30 services.

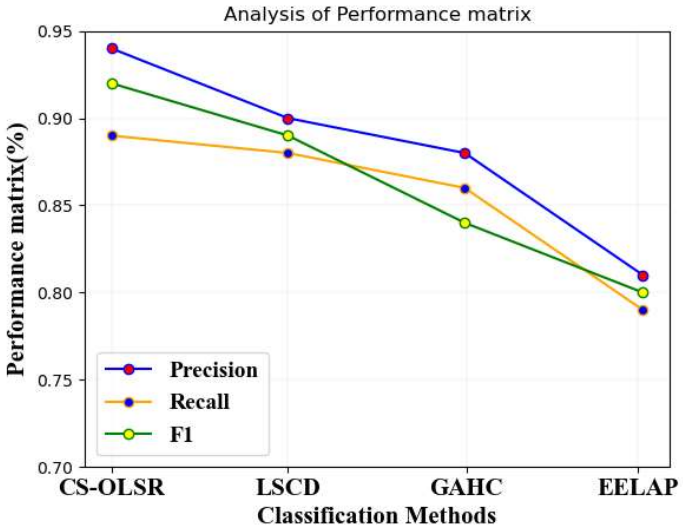

Figure 8: Analysis of overall performance accuracy

Figure 8 described based on the performance matrix calculated based on the confusion matrix (TP, FP, and FN) estimations depend on the dataset truth value rate. In the proposed method CS-OLSR analysis the precision score is 0.94%, recall score is 0.89% and F1 score id 0.92%.Comparing the previous methods, FLSP precision is 0.90% recall is 0.88% and f1 is 0.89%,DRL precision is 0.88% recall is 0.86% and f1-score is 0.84%,MDDLF-IoT precision score is 0.81%,recall is0.79% and f1 score is0.80 and BR-IoT precision score is 0.79%,recall is 0.74% and f1 score is 0.71%. Experimental results evaluating the testing data of 500 records based on these measurements, in the proposed method comparatively show higher performance better than previous methods.

Table 6: Analysis on false classification ratio

False Classification Ratio in % vs No of Services

Comparison methods/ services	10 Services	20 Services	30 Services
EELAP	28.4	25.1	20.7
GAHC	23.5	21.4	16.9
LSCD	20.7	16.9	12.6
CS-OLSR	6.7	2.4	1.3

Analysis of false classification ratio the proposed comparison with previous methods performance is listed in table 6.

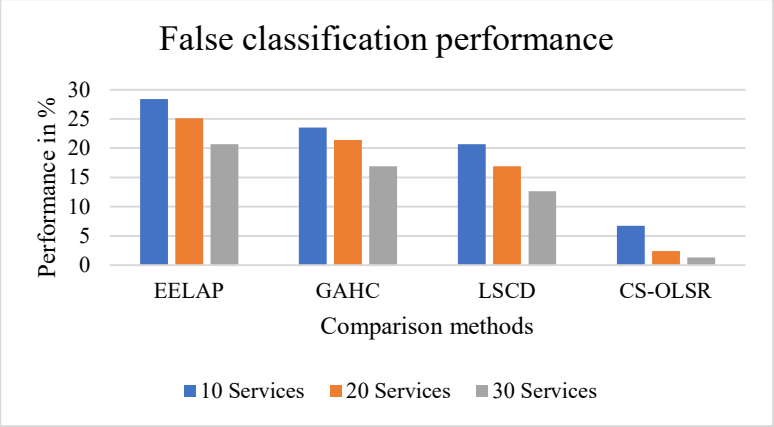

Figure 9: Impact of false classification ratio

Figure 9 illustrates impact of false classification ratio performance for IDS with various services like 10, 20 and 30 services. In this graph, X-axis is a comparison methods moreover Y-axis performance gradually decrease with each method. The proposed Service Specific Payload Inference Analysis Model (CS-OLSR-LSVNN) method achieves 1.3% false classification performance for 30 services besides EELAP achieves 20.7% of false classification performance, GAHC method achieves 16.9%, and LSCD method achieves 12.6%.

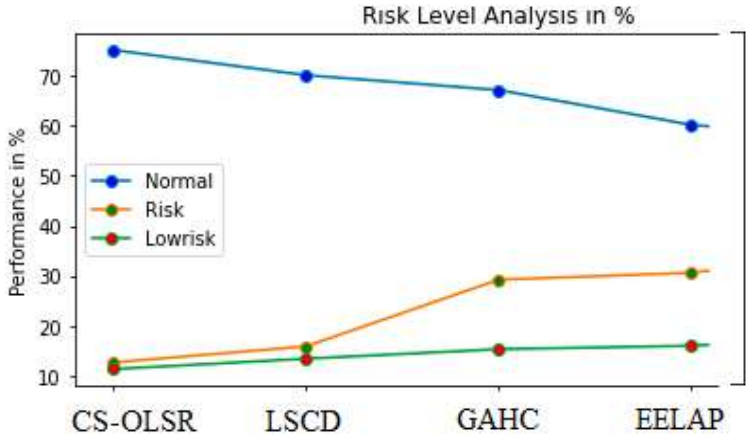

Figure 10: Analysis of clone feature dependency risk rate

Figure 10 described as Risk level identification for IDS detection in cyber security using darkenet dataset features. Proposed approach CS-OLSR calculated risk level 12.67%, low risk

11.34% and normal level 75.14% using maximum weighted feature to improving the normal features, and reducing the risk features. Previous methods of LSCD are evaluating the normal 70.12%, low-risk at 15.89% and risk is 13.42%, GAHC is evaluating the normal is 60.19%, the low-risk level is 30.61% and risk is 16.01, and EELAP analysis the normal level of 56.87%, Low-risk is 35.21% and the risk is 18.23%.

Table 7: Impact of time complexity performance

Time complexity in seconds vs No of Services			
Comparison methods/ services	**10 Services**	**20 Services**	**30 Services**
EELAP	62.4	66.1	67.2
GAHC	48.6	52.8	58.6
LSCD	31.6	35.5	41.2
CS-OLSR	17.5	21.6	28.3

Table 7 describes the impact of time complexity performance vs no of services. The proposed CS-OLSR has 28.3 sec for IDS classification besides EELAP has 67.2 sec, GAHC has 58.6 sec and LSCD has 41.2 sec for IDS classification.

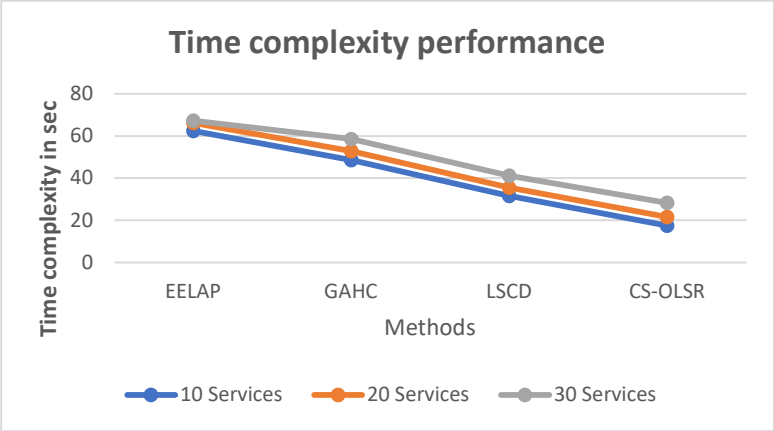

Figure 11: Result of time complexity performance

Figure 11 denotes result of time complexity performance the proposed CS-OLSR technique compared with other methods like EELAP, GAHC and LSCD. In figure X-axis presents comparison methods besides Y-axis presents time complexity performance in seconds with each methods. However the proposed method produced less time complexity result than previous techniques.

Table 8. Performance on various measures

Comparison methods/ services	Detection Rate %	False Ratio %	Time Complexity in sec
EELAP	78.6	18.6	62.5
GAHC	81.3	16.2	48.4
LSCD	87.2	10.6	31.5
CS-OLSR	98.2	1.3	17.9

Table 8 denotes the proposed CS-OLSR performance of various measures based on detection rate, false ratio and time complexity. The proposed FGWO- LSVNN techniques gives better performance than prevailing methods. Analysis of these feature values based on the Data size and bit rate, in the proposed approach better Response ratio comparing the previous methods.

Table 7 described the Response ratio and time complexity based on the source to-destination data transfer without traffic. Analysis of these feature values based on the Data size and bit rate, in the proposed approach better Response ratio comparing the previous methods.

$$Response\ ratio(time) = \frac{Data\ size}{bit\ rate} \quad (22)$$

$$Total\ Time\ complexity = [O\{\propto \{O(N) + O(N) + O(t) + O(t) + O(x) + O(xy) + O(x^2y)\}\}] \quad (23)$$

$$= O(\propto (N + T + x^2y)) \quad (24)$$

Where, T is time, x-arrival time, and y –response time, in the equation based on the best detection takes O (n) time which is represented in equation 45. The data Response ratio can be obtained from the data size and bitrate, time complexity is based on data arrival and response time. It estimates features of proposed and previous methods calculations.

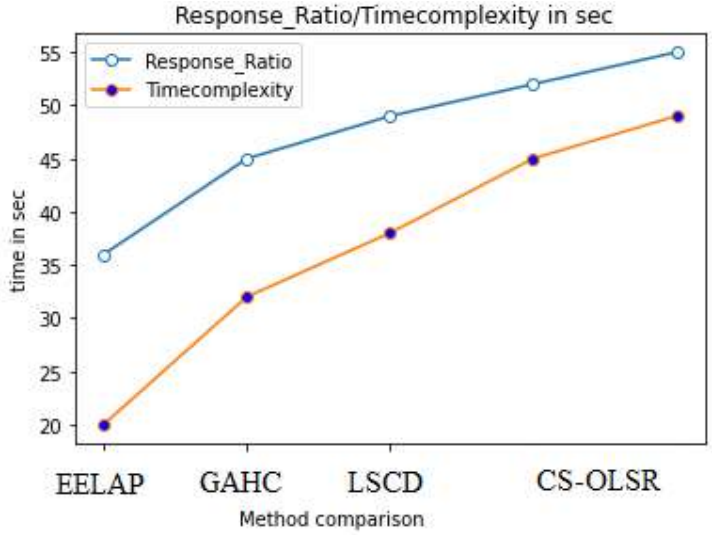

Figure 12: Analysis of Response Ratio

Figure 12 described the Response ratio for using the features and evaluating the response time. Comparing the proposed and previous methods based on data size and bit rate sending and receiving response time. In the proposed method, CS-OLSR is 36s immediate response for Central Processing Unit (CPU), and compared to previous methods takes a time MLPGCN is 45s, DRL is 49s, GAHC is 52s and EELAP is 55s. Time complexity is CS-OLSR

is the 20s, LSCD is 32s, GAHC is 38s, EELAP is 45s.The proposed time for CPU response ratio and reduced time complexity is better than previous methods.

5. Conclusion

An advanced transfer learning model detects cloning attacks based on neural fuzzy dense sub spectral scaling feature selection (NFI-SSFS) implemented effectively with cooperatively secure optimal link stability routing assignment (CS-OLSR) achieve high performance to detect clone. Communication logs are collected to consume characteristic levels of variation in under-memory packet discrepancy rate and transmission error facts, along with driving Error Injection Impact Rate (FIR) and Time stamped Communication Behaviour Rate (DSCPR). NFI-SSFS is used to derive dense Eigen factors to marginalize the clone attack rate. CS-OLSR is then used to ensure secure routing based on the identified cloned attack regions. The proposed system can effectively identify data replacement nodes and thus effectively detect cloning attacks. Further the future work concentrates deep Learning model based optimized neural Network. To concentrate the projection of dimensionality problem will highlight to resolve and improve the performance of clone attack detection to enrich the security.

References

[1] J. R. Dora and K. Nemoga, “Clone node detection attacks and mitigation mechanisms in static wireless sensor networks,” J. Cybersecur Priv., vol. 1, no. 4, pp. 553-579, 2021, doi:10.3390/jcp1040028.

[2] H. R. Shaukat et al., “Hybrid multi-level detection and mitigation of clone attacks in mobile wireless sensor network (MWSN),” Sensors (Basel), vol. 20, no. 8, p. 2283, 2020, doi:10.3390/s20082283[Green Version], Google Scholar.

[3] P. P. Devi and B. Jaison, “Protection on Wireless Sensor Network from Clone Attack using the SDN-Enabled Hybrid Clone Node Detection Mechanisms” Comput. Commun., vol. 152, pp. 316-322, 2020, doi:10.1016/j.comcom.2020.01.064.

[4] J. Xu et al., “Collective memory for detecting nonconcurrent clones: A localized approach for global topology and identity tracing in IoT networks” in IEEE Internet Things J., vol. 8, no. 7, pp. 5762-5777, 1 April1, 2021, doi:10.1109/JIOT.2020.3032127.

[5] M. Tahboush and M. Agoyi, “A hybrid wormhole attack detection in mobile ad-hoc network (MANET)” in IEEE Access, vol. 9, pp. 11872-11883, 2021, doi:10.1109/ACCESS.2021.3051491.

[6] F. Pan et al., “Clone detection based on BPNN and physical layer reputation for industrial wireless CPS” in IEEE Trans. Ind. Inform., vol. 17, no. 5, pp. 3693-3702, May 2021, doi:10.1109/TII.2020.3028120.

[7] M. Numan et al., “A systematic review on clone node detection in static wireless sensor networks” in IEEE Access, vol. 8, pp. 65450-65461, 2020, doi:10.1109/ACCESS.2020.2983091.

[8] S. Chen et al., “Automated labeling and learning for physical layer authentication against clone node and Sybil attacks in industrial wireless edge networks” in IEEE Trans. Ind. Inform., vol. 17, no. 3, pp. 2041-2051, Mar. 2021, doi:10.1109/TII.2020.2963962.

[9] R. J. Cai et al., "An evolutionary self-cooperative trust scheme against routing disruptions in MANETs" in IEEE Trans. Mob. Comput., vol. 18, no. 1, pp. 42-55, Jan. 1 2019, doi:10.1109/TMC.2018.2828814.
[10] F. Pan et al., "Clone detection based on physical layer reputation for proximity service" in IEEE Access, vol. 7, pp. 3948-3957, 2019, doi:10.1109/ACCESS.2018.2888693.
[11] U. Srilakshmi et al., "A secure optimization routing algorithm for mobile ad hoc networks" in IEEE Access, vol. 10, pp. 14260-14269, 2022, doi:10.1109/ACCESS.2022.3144679.
[12] K. Cho et al., "Classification and experimental analysis for clone detection approaches in wireless sensor networks" in IEEE Syst. J., vol. 7, no. 1, pp. 26-35, Mar. 2013, doi:10.1109/JSYST.2012.2188689.
[13] Z. Zheng et al., "Energy and memory efficient clone detection in wireless sensor networks" in IEEE Trans. Mob. Comput., vol. 15, no. 5, pp. 1130-1143, May 1 2016, doi:10.1109/TMC.2015.2449847.
[14] M. Conti et al., "Distributed detection of clone attacks in wireless sensor networks" in IEEE Trans. Depend. Sec. Comput., vol. 8, no. 5, pp. 685-698, Sept.-Oct. 2011, doi:10.1109/TDSC.2010.25.
[15] E. O. Ochola et al., "Manet reactive routing protocols node mobility variation effect in analysing the impact of Black Hole attack" in SAIEE Afr. Res. J., vol. 108, no. 2, pp. 80-92, Jun. 2017, doi:10.23919/SAIEE.2017.8531629.
[16] U. Srilakshmi et al., "An improved hybrid secure multipath routing protocol for MANET" in IEEE Access, vol. 9, pp. 163043-163053, 2021, doi:10.1109/ACCESS.2021.3133882.
[17] C. Tang and D. Han, "A low resource consumption clone detection method for multi-base station wireless sensor networks" in IEEE Access, vol. 8, pp. 128349-128361, 2020, doi:10.1109/ACCESS.2020.3007388.
[18] "Counterfeit Clones: A Novel Technique for Source and Sink Location Privacy in Wireless Sensor Networks" Al Ahmadi, "in S. A. Al-Ahmadi, Access, vol. 10, pp. 62693-62701, 2022, doi:10.1109/ACCESS.2022.3182660.
[19] K. O. Dong et al., "Senior member, IEEE 'LSCD: A low-storage clone detection protocol for CyberPhysical systems' Mianxiong," IEEE Trans. Comput. Aid. Des. Integr. Circuits Syst., vol. 35, no. 5, May, 2016.
[20] Z. Zhang et al., "A clone detection algorithm with low resource expenditure for wireless sensor networks," J. Sens., vol. 2018, no. Mar., 1-16, 2018, doi:10.1155/2018/4396381.
[21] Gopalakrishnan Subburayalu, Hemanand Duraivelu, Arun Prasath Raveendran, Rajesh Arunachalam, Deepika Kongara & Chitra Thangavel (2021) Cluster Based Malicious Node Detection System for Mobile Ad-Hoc Network Using ANFIS Classifier, Journal of Applied Security Research, DOI: 10.1080/19361610.2021.2002118.
[22] Gopalakrishnan, S. and Kumar, P. (2016) Performance Analysis of Malicious Node Detection and Elimination Using Clustering Approach on MANET. Circuits and Systems, 7, 748-758. doi: 10.4236/cs.2016.76064.
[23] D. Hemanand, G. . Reddy, S. S. . Babu, K. R. . Balmuri, T. Chitra, and S. Gopalakrishnan, "An Intelligent Intrusion Detection and Classification System using CSGO-LSVM Model for

Wireless Sensor Networks (WSNs)", Int J Intell Syst Appl Eng, vol. 10, no. 3, pp. 285–293, Oct. 2022.