ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20230613000723
Original scientific paper

A Combined Scheme Based on Artificial Immune System for Selective Forwarding Attack Detection in Event Driven Wireless Sensor Networks

T. YUVARAJA*, S. SUMITHRA, M. JEYALAKSHMI, M. PREMKUMAR

Abstract: In general, the wireless sensor networks (WSNs) can be easily targeted by attacks due to their open nature. Among these attacks, the selective forwarding attack is particularly challenging to detect within WSNs. In this type of attack, malicious nodes deliberately discard received data packets, making it difficult to identify such attacks. Existing methods for selective forwarding attacks often suffer from low accuracy or high algorithm complexity, when dealing with Distributed Denial of Service attacks (DDoS). We propose an artificial immune system based on the danger model to detect network attacks. Our approach combines variational mode decomposition (VMD) and LSTM to identify selective forwarding attacks among other DoS attacks. We determine an optimal danger threshold and compare the outcomes to confirm the presence of attacks. This approach improves detection accuracy while minimizing computation requirements. The simulation results demonstrate that our proposed method achieves a low missing detection rate (*MDR*) of 0.6% and maintains a false detection rate (*FDR*) below 3.3%. Furthermore, in comparison with previous works, our system produces low algorithm complexity, making it more efficient in practical implementations.

Keywords: clustering; LSTM; selective forwarding; variational mode decomposition (VMD); wireless sensor networks.

1 INTRODUCTION

Wireless Sensor Networks (WSNs) [1] are self governing networks that communicate wirelessly, gathering and transmitting environmental data [2]. Event driven Wireless Sensor Networks (EWSNs) [3] are widely used for data collection purposes. Unlike conventional networks that constantly gather information, EWSNs gather the information once predetermined events are identified by the SN within the arrangement. EWSNs prove especially valuable in applications where real-time response to events is critical [4]. By transmitting data only during event occurrences, EWSNs conserve energy, minimize network congestion, and operate for extended durations without frequent battery replacements. However, due to open nature, EWSNs often lack robust security measures, rendering them vulnerable to malicious attacks [5]. Among various attacks, the selective forwarding attack [6] is particularly challenging to detect. In a selective forwarding attack, attackers randomly drop packets, resulting in unusual packet loss within EWSNs. This attack becomes even more challenging to distinguish from genuine packet loss within harsh environmental conditions caused by factors, which adversely affect channel quality. Consequently, intrusion detection systems (IDS) face significant difficulties in discerning the unusual behavior associated with these attacks caused by adverse weather conditions. In contrast to conventional networks, the security of EWSNs requires careful consideration by both detection accuracy and energy consumption challenges [7]. Since the limited resources of nodes, the direct utilization of IDSs designed for networks is impractical due to their computational complexity. Consequently, an effective detection scheme for EWSNs must possess characteristics such as low complexity, a minimal *MDR*, and a low *FDR*, in order to fulfill the unique requirements of EWSNs.

2 RELATED WORKS

The objective of ACK based techniques is to detect attacks by utilizing additional data entrenched in the forwarding packets [8]. The approach involves inserting certification data into the forwarded data packets of nodes. The node generates multiple ACK to validate the forwarding of its packets. By searching and identifying ACK message behavior, malicious nodes are accurately detected. Another provenance-based method was introduced. Through the analysis, it was observed that the transmission loss of normal nodes differed significantly from that of malicious nodes [9]. A countermeasure is suggested to achieve high detection probabilities in small networks. Rather than examining the transmission processes, this approach integrated packet drop time to enhance the accuracy of detection. However, earlier methods did not effectively address the challenges posed in WSNs and struggled to classify different types of nodes under such conditions [10]. Typically, a deciding node known as the watchdog node is designated to gather node behavior and often renew the trust value. The illegitimate nodes are identified by comparing with forwarding rates and estimated thresholds [11]. Every node first receives a reputation value to identify abnormal nodes. However, as the channel quality changes, the detecting capacity declines. A reputation model-based solution to identifying attacks [12] builds a reputation framework and verifies incursions using reputation values. Additionally, in order to resist selective forwarding attacks, authors incorporated the reputation, greatly increasing the PDR. It is recently developed to explain the crisis of incorrect conditions, and it worked effectively in dynamic contexts. The rise of ML and AI has led to increased adoption of AI-based techniques in attack detections. However, these approaches often have long learning periods and inadequate response speeds. The utilization of CFRs as a crucial feature for identifying suspicious nodes has gained considerable attention. However, most DCAs are not parameter-free and struggle during clustering. On the other hand, DNNs, a notable subset of ANNs, excel at feature extraction for classification purposes. This study employs Long Short-Term Memory (LSTM) and explores techniques that employ DL method for detecting attacks. Various algorithms such as DNN and LSTM network have been used to secure WSNs. The PSO combined with the BPNN to reduce consumption, while others have proposed an authentication scheme, which is sensitive to signal imperfections. However, it limited suitability for massive WSN nodes and specific areas of application. Recently, RNN has been proposed, which has shown effectiveness in attack revealing. Therefore, LSTM model with additional advancements aims to overcome this challenge, conserve nodes' energy, and improve detection accuracy [13-15]. The proposed research introduces a multistage detection method for combating attacks in EWSNs. In the training phase, normal nodes' forwarding rates are processed using VMD. Then LSTM is trained with data to make predictions. In detection phase, threshold selection rule is employed to spawn anomaly points for each node, and a voting method is designed to classify nodes as legitimate, suspicious, or malicious. The contributions of the research are: utilizing VMD-LSTM for attack detection, selecting dynamic thresholds based on prediction errors, and designing an effective scheme for classification. The designed model aims to enhance detection accuracy and adaptability in challenging network conditions.

3 PROPOSED VMD + LSTM FOR EWSNS

In EWSNs, the attacks are launched by malicious nodes, which randomly discard received data, causing anomalous forwarding behavior characterized by unusual rates. Conversely, legitimate nodes also experience reduced forwarding rates due to the adverse effects of operating in ruthless conditions, resulting in degraded performance. As a consequence, the forwarding rates are overlapped, presenting significant difficulties in distinguishing between them. To tackle this challenge, we aim to introduce an innovative solution that harnesses the power of deep learning models to detect attacks in crucial conditions.

3.1 Overview

The work flow of our method is shown in Fig. 1. Our scheme consists of the training and the testing phase. Each phase comprises four steps to effectively detect selective forwarding attacks:

Training phase:

- Learn the normal node behavior using the VMD+LSTM model. It is trained to predict the future value of every node. With the help of predicted and actual values, prediction error is calculated.

Testing phase:

- Create a decision-making function that uses the prediction errors to dynamically choose a threshold. Overriding error points are regarded like AP.
- Implement the NVM to differentiate AP. If the NVM results fail to reach a consensus, the resultant node is labeled as suspicious.
- Constantly monitor the entire network. If a node is flagged as suspicious more than a specified threshold (K), it is confirmed as malicious one.

By dividing the scheme into two stages and following these four steps, our approach aims to effectively train the model, set dynamic thresholds, employ neighbor voting, and continuously screen the network to detect and identify malicious nodes involved in selective forwarding attacks.

3.2 Data Set

The objective of our system is to understand the typical forwarding behavior exhibited by regular one and leverage the understanding to estimate the value of individual nodes within future rounds. By analyzing the prediction errors, we can potentially identify malicious nodes exhibiting abnormal behavior. In order to train the model, a dataset displayed by regular nodes was generated using MATLAB 2022a. This function simulated random packet loss by producing in the range of [0.8, 1] which is represented as DS:

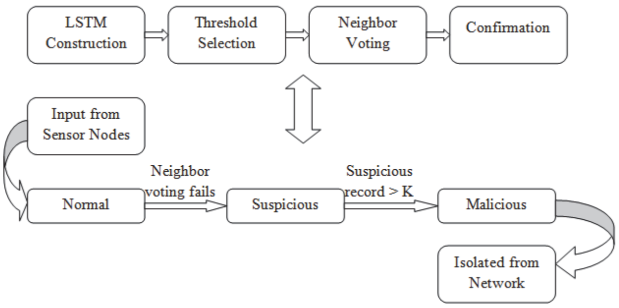


Figure 1 Proposed VMD + LSTM model workflow

$$DS=\{DS_1, DS_2, ..., DS_{2000}\} \quad (1)$$

It is important to note that in real scenarios the datasets can be readily available. However, for simplicity, we chose to simulate the making practice. The normal forwarding rates with varying rounds are depicted in Fig. 2, validating the accuracy of regular nodes and confirming that they align with our expectations.

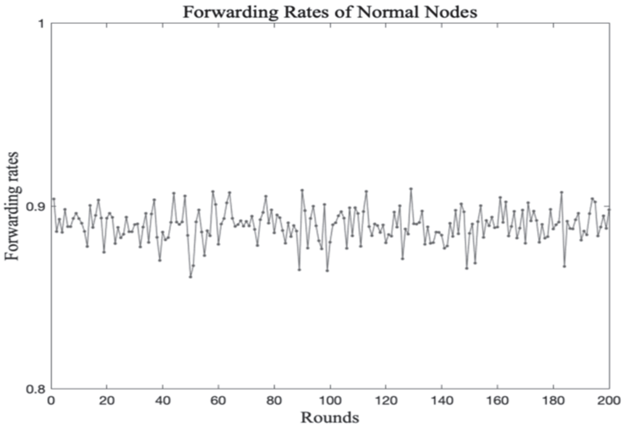


Figure 2 Generated normal forwarding rates

In the detecting stage, the forwarding rate (*FR*) of a specific node *i* in the *k*th round is calculated using the equation:

$$FR_{ik} = \frac{FP_{ik}}{Rp_{ik}} \quad (2)$$

Here, FP_{ik} represents forwarding packets count, and RP_{ik} represents received packets count in the *k*th round. This calculation allows us to determine the forwarding rate for each node. Consequently, we can represent the N_i as:

$$N_i = \{FR_{i1}, FR_{i2}, ..., FR_{in}\} \tag{3}$$

Here, n refers to the simulation round number. It is important to note that all the data points in the series exist in a 1D gap, making model to effectively process and analyze the data.

3.3 Training Model

By employing the VMD technique to decompose and smooth the data, we enhance the performance of the LSTM section. The complete training process is illustrated in Fig. 3. Initially, the processed forwarding rate (FR) series undergo decomposition using VMD. Subsequently, subsets of the series, such as IMF k (intrinsic mode function k), are utilized for training and fine-tuning the LSTM model. Finally, we save the trained LSTM model to be employed in the detecting stage. This integration of VMD and LSTM enables improved modeling and analysis of the data, contributing to the overall effectiveness of our approach.

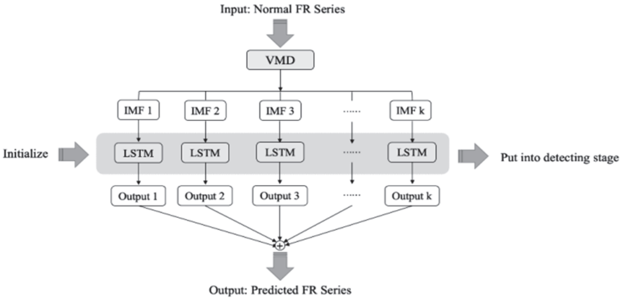


Figure 3 Training phase of proposed model

```
Algorithm 1 : VMD+LSTM Model
Input:
- Normal forwarding rate dataset $DS$
- Forwarding rate of every node $N$
Output:
- Error series of each node
Parameters:
- $K$: VMD component number
- $\alpha$: Initial learning rate
- $ite$: Training epoch
- $S$: Size of the sliding window
Begin
Decompose the input series $DS$ into $k$ subsets using VMD
For each iteration $ite$ do:
  Train the LSTM model with the $k$ subsets
End for
For each node $i$ in the length of $node\_id$ do:
  For each value $j$ in the length of $Nji$ do:
    Use the trained LSTM model to predict the forwarding rate $FRij$
    Calculate the error $eij$ for node $i$ and value $j$: $eij = |FRij - FRij|$
  End for
  Return the error series $e$ of node $i$
End for
Return the error series of each node
End
```

After the training, our next step is to compute the prediction errors for each node. Since LSTM exhibits short-term prediction characteristics, we employ a sliding window approach to enhance our prediction. For example, we set the size as S, and it varies gradually during the process. Subsequently, we use FR_2, FR_3, ..., FR_{s+1} to predict FR_{s+2} and so on. After the prediction, the error ei is examined between FR_i and $\widehat{FR_i}$:

$$e_i = \widehat{FR_i} - FR_i \tag{4}$$

3.4 Detection Phase

Once we have obtained the $\mathrm{e_i}$ the dynamic threshold selection approach is used to detect anomaly points. These points, also known as outliers, exhibit distinct characteristics that deviate significantly from the normal data. The presence of these points tends to enhance the σ and μ of the entire dataset. To address this, our method determines a threshold dynamically, where any data point surpassing the threshold is identified as abnormal and classified as an anomaly.

```
Algorithm 2: Detection Model
Input:
- Training dataset DS of normal forwarding rates
- Forwarding rate dataset N of node i
Output
- Detection result for node i
Begin
Perform VMD-LSTM predictions on node i's dataset N to obtain the error series e_i (see Algorithm 1).
Detect local anomaly points in e_i using the selected threshold ε.
Conduct the neighbor voting process for node i:
  If the voting ratio is higher than V:
    Classify it as a normal node.
  Else:
    Mark it as a suspicious node.
  If the number of suspicious records exceeds K:
    Confirm it as a malicious node.
  End if
End if
Return the detection result for node i.
End
```

Through the application of the Law of Averages, following the statistical 3σ rule, data points outside the 3σ range are typically classified as outliers. To optimize our anomaly detection capability while minimizing computational overhead, we restricted the search region to the interval [0, 3]. It is evident that it exhibits lower volatility compared to those of malicious nodes. As a result, the number of detected anomaly points for malicious nodes far exceeds that for normal nodes. This observation aligns with our intended outcome, which is to differentiate normal nodes from malicious nodes by analyzing the errors. The LSTM-NV scheme's detection complexity is primarily determined by the computational requirements during the detection stage. Overall, the complexity level is $O(m \times \log n)$, which is suitable for practical implementation. In summary, the complexity is based on structure of the LSTM network, while the detection complexity is influenced by optimal thresholds.

4 RESULTS AND DISCUSSION

4.1 Simulation Parameters

Tab. 1 shows that the simulation space is adjusted from 100 to 1000 nodes, in a 300 m circle shape. The harsh radius is set to 75 m, and the starting energy of the nodes was set to 1 J. For the purpose of evaluating the model's efficacy, the proportion of malicious nodes was increased from 5% to 20%. The table further displays that the VMD-LSTM network's parameters are split into 7:3 for training and testing, and SGD is chosen. The PSO technique optimized the initial learning rate within the range [0.001, 0.1] with a sliding window size of 20. Additionally, particular values were chosen for the inertia weight and learning factor throughout the PSO process. The Normal nodes' forwarding rates were established as a uniform distribution between [0.8, 1], whereas malevolent nodes' forwarding rates were modeled as a combination of Poisson distributions. In order to replicate dynamic scenarios, value is adjusted to [0.6, 0.8], based on how far away to the harsh region. The locations of network conditions also varied once in 10 rounds.

Table 1 Simulation parameters

Parameter	Value
Total node number	[100, 300, 500, 1000]
Ratio of malicious nodes	[5%, 10%, 15%, 20%]
Hidden neuron number 1	[20, 200] (optimized by PSO]
Hidden neuron number 2	[20, 200] (optimized by PSO]
Initial learning rate	[0.001, 0.1] (optimized by PSO]
VMD component number	4
Sliding window size	20
Iteration epoch	300
Inertia weight	ω max = 0.8, ω = 0.6
Learning factors	$C1$ = 1.5, $C2$ = 1.5

4.2 Evaluation parameters

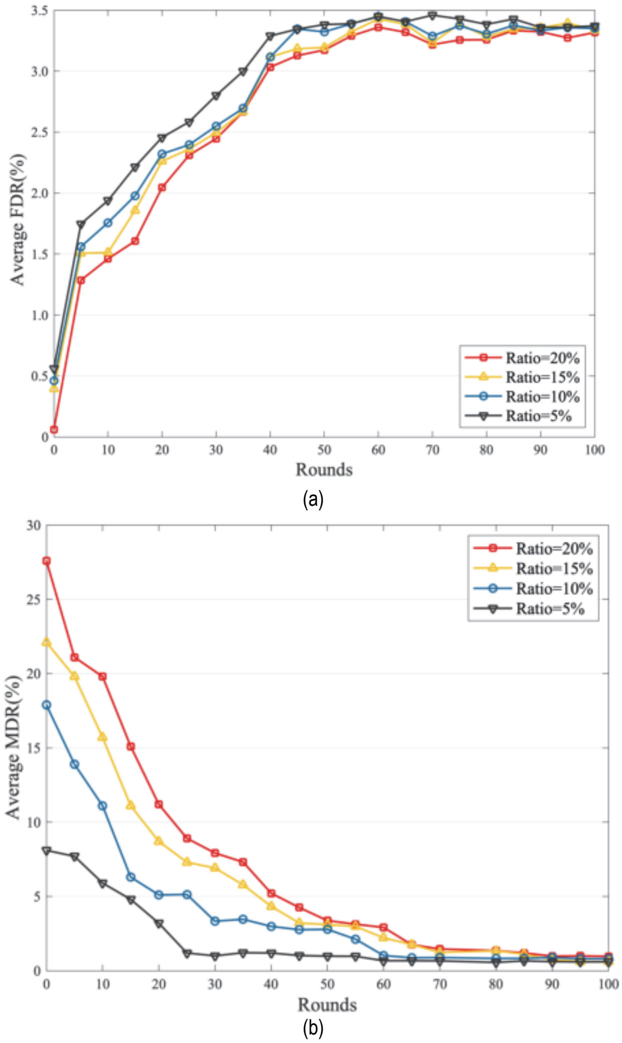
To comprehensively review the detection capability of our scheme, it is necessary to establish metrics for evaluation. The three metrics used in this context are referred to as *MDR*, *FDR* and *DAR*. These metrics provide quantitative measures to evaluate the performance and effectiveness of the detection process.

$$MDR = \frac{N_{nm}}{N_{nm} + N_{mn}}$$
$$FDR = \frac{N_{nm}}{N_{nm} + N_{nm}} \tag{5}$$
$$DAR = \frac{N - N_{mn} - N_{nm}}{N}$$

4.3 Discussion

Fig. 4 to Fig. 6 provide an analysis of *MDR*, *FDR* and *DAR* with varying malicious ratios. If the detecting system is malfunctioning and *MDR* is close to null, *FDR* will unavoidably increase sharply.

FDR and *MDR* values are lower and higher respectively when the criteria are less strict because more malicious nodes can elude detection. They exhibit distinct patterns, but both tend to stabilize after approximately 50 simulation rounds, as depicted in Fig. 4. After 50 rounds, stable *FDR* results across different malicious ratios hover around 3.3%, with the 20% ratio yielding the lowest *FDR*. This can be attributed to the presence of a larger number of malicious nodes, which poses significant challenges for detection. As a result, the detection condition becomes less stringent, reducing the chances of misjudging legitimate as suspicious and resulting in the low *FDR* along with every part of ratios. In the same way, the *MDR* stabilizes at around 0.6% after approximately 50 rounds. Fig. 5 illustrates the average stable results and showcases connection among two metrics. It confirms that suspicious node increases, *MDR* tends to rise while *FDR* decreases to some extent. However, both *MDR* and *FDR* remain in the region of 0.7% and 3.5% respectively, demonstrating the performance of the proposed method.



(a)

(b)

Figure 4 (a) Average *FDR* and (b) *MDR* with varying rounds

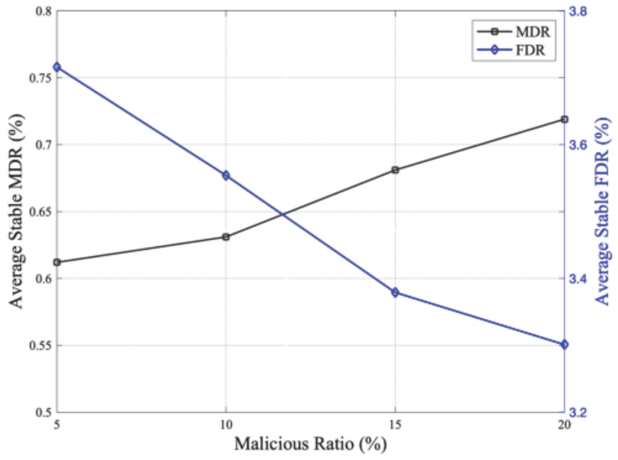


Figure 5 Malicious ratio vs *MDR* and *FDR*

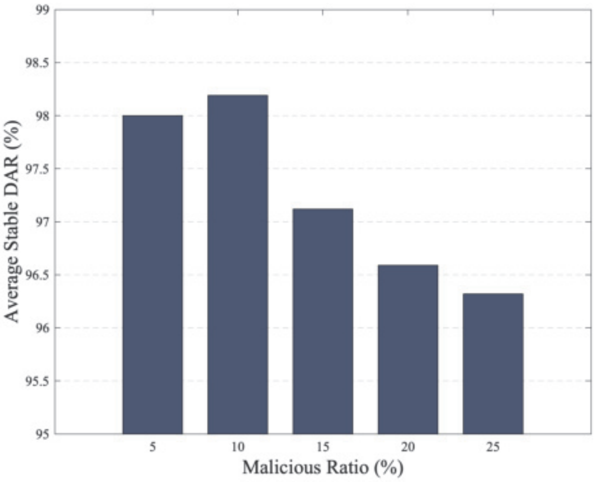


Figure 6 Average *DAR* vs malicious ratio / %

In Fig. 6, it is noteworthy that all *DAR* values exceed 95%, indicating excellence of the method. The trend observed is a slight increase followed by a decrease. In theory, a lower malicious ratio leads to better detection accuracy. The main reason behind this observation is as follows: once the ratio is 5%, suspicious count is minimal, and FPR have significant impact on outcome, thereby affecting accuracy negatively. On the other extreme, if malicious ratio increases, errors will occur in NVM. For instance, if all nearby nodes are suspicious, it impacts integrity of the voting process. We conducted a comprehensive comparison between our proposed scheme and other recently proposed methods for the attack detection. Assessment encompassed various categories and factors, including *MDR*, *FDR*, complexity and the consideration of network conditions. The detailed comparison results are presented in Tab. 2. All the methods including our proposed scheme achieved excellent *MDR* and *FDR* results, with values below 5%. Notably, the DCA-SF [8] method demonstrated the low *MDR* and *FDR* values, although it also exhibited the highest complexity among compared methods. It is worth mentioning that the simulation settings varied across the methods.

Table 2 Performance analysis of proposed method

Scheme	*MDR* / %	*FDR* / %	Harsh Area	Complexity
Proposed VMD + LSTM	0.68	3.31	Mobile	$O(m\times \log n)$
DPC [7]	0.92	0.95	-	$O(n\times \log n)$
DBSCAN [11]	0.67	4.54	Immobile	$O(n\times \log n)$
DCA-SF [8]	0.12	0.09	-	$O(n)$
Danger Model [10]	1.39	4.31	Immobile	$O(\mathrm{d}n)$

Among the compared schemes, the DPC-based scheme [7], and DCA-SF scheme [8] did not consider the network conditions while the reference [11] and [10] analyzed in mobile area to examine the accuracy. However, these schemes did not account for time-variant scenarios. Furthermore, our proposed scheme demonstrated the low complexity scheme, indicating its superiority by their speed in computation. In summary, our proposed scheme was compared with other effective methods for selective forwarding attack detection, covering various aspects. The results showed that all methods achieved ideal *MDR* and *FDR* results, with the DCA-SF method exhibiting the lowest values. However, our scheme stood out with its lower complexity and consideration of time-variant harsh environments.

5 CONCLUSIONS

This paper introduces a VMD and LSTM methodology for detecting selective forwarding attacks in EWSNs. The proposed scheme consists of four phases to identify the suspicious nodes. When the system reaches stability, simulation findings show *MDR* of around 0.6% and *FDR* of about 3.3%. The suggested approach yields high *DAR* values, surpassing 95% across various malicious node ratios. Our method is less sophisticated than those already in use. The neighbor voting technique should be improved in future studies to overcome approximation issues in distributed WSNs. Conducting experiments with hardware might also be helpful for more research and study.

6 REFERENCES

[1] Gharaei, N., Bakar, K. A., Hashim, S. Z. M., & Pourasl, A. H. (2019). Inter-and intra-cluster movement of mobile sink algorithms for cluster-based networks to enhance the network lifetime. *Ad Hoc Networks*, *85*, 60-70. https://doi.org/10.1016/j.adhoc.2018.10.020

[2] Phasinam, K., Kassanuk, T., Shinde, P. P., Thakar, C. M., Sharma, D. K., Mohiddin, M. K., & Rahmani, A. W. (2022). Application of IoT and cloud computing in automation of agriculture irrigation. *Journal of Food Quality*, 1-8. https://doi.org/10.1155/2022/8285969

[3] Othman, R. A., Darwish, S. M., & Abd El-Moghith, I. A. (2023). A Multi-Objective Crowding Optimization Solution for Efficient Sensing as a Service in Virtualized Wireless Sensor Networks. *Mathematics*, *11*(5), 1128. https://doi.org/10.3390/math11051128

[4] Subramani, S. & Selvi, M. (2023). Comprehensive review on distributed denial of service attacks in wireless sensor networks. *International Journal of Information and Computer Security*, *20*(3-4), 414-438. https://doi.org/10.1504/IJICS.2023.128828

[5] Luo, S., Lai, Y., & Liu, J. (2023). Selective forwarding attack detection and network recovery mechanism based on cloud-edge cooperation in software-defined wireless sensor network. *Computers & Security*, *126*, 103083. https://doi.org/10.1016/j.cose.2022.103083

[6] Premkumar, M., Ashokkumar, S. R., Jeevanantham, V., Mohanbabu, G., & Anu Pallavi, S. (2023). Scalable and energy efficient cluster based anomaly detection against denial of service attacks in wireless sensor networks. *Wireless Personal Communications*, *129*(4), 2669-2691. https://doi.org/10.1007/s11277-023-10252-3

[7] Ding, J., Zhang, H., Guo, Z., & Wu, Y. (2021). The DPC-based scheme for detecting selective forwarding in clustered wireless sensor networks. *IEEE Access*, *9*, 20954-20967. https://doi.org/10.1109/ACCESS.2021.3055026

[8] Vinoth Kumar, K., Jayasankar, T., Eswaramoorthy, V., & Nivedhitha, V. (2020). SDARP: Security based Data Aware Routing Protocol for Ad hoc Sensor Networks. *International Journal of Intelligent Networks (Elsevier-KeAi)*, *1*, 36-42. https://doi.org/10.1016/j.ijin.2020.05.005

[9] Fu, H., Liu, Y., Dong, Z., & Wu, Y. (2019). A data clustering algorithm for detecting selective forwarding attack in cluster-based wireless sensor networks. *Sensors*, *20*(1), 1-23. https://doi.org/10.3390/s20010023

[10] Premkumar, M., Sundararajan, T. V. P., & Mohanbabu, G. (2022). Dynamic Defense Mechanism for DoS Attacks in

Wireless Environments Using Hybrid Intrusion Detection System and Statistical Approaches. *Tehnički vjesnik*, *29*(3), 965-970. https://doi.org/10.17559/TV-20210604113859

[11] Huang, X. & Wu, Y. (2022). Identify selective forwarding attacks using danger model: Promote the detection accuracy in wireless sensor networks. *IEEE Sensors Journal*, *22*(10), 9997-10008. https://doi.org/10.1109/JSEN.2022.3166601

[12] Liu, Y. & Wu, Y. (2021). Employ DBSCAN and neighbor voting to screen selective forwarding attack under variable environment in event-driven wireless sensor networks. *IEEE Access*, *9*, 77090-77105. https://doi.org/10.1109/ACCESS.2021.3083105

[13] Premkumar, M., Ashokkumar, S. R., Mohanbabu, G., Jeevanantham, V., & Jayakumar, S. (2022). Security behavior analysis in web of things smart environments using deep belief networks. *International Journal of Intelligent Networks*, *3*, 181-187. https://doi.org/10.1016/j.ijin.2022.10.003

[14] Gopinath, S., Vinoth Kumar, K., Elayaraja, P., Parameswari, A., Balakrishnan, S., & Thiruppathi, M. (2021). SCEER: Secure cluster based efficient energy routing scheme for wireless sensor networks. *Materials Today: Proceedings (Elsevier)*, *45*(2), 3579-3584. https://doi.org/10.1016/j.matpr.2020.12.1096

[15] Premkumar, M., Sundararajan, T. V. P., & Vinoth Kumar, K. (2019). Various Defense Countermeasures against DoS attacks in Wireless Sensor Networks. *International Journal of Scientific & Technology Research*, *8*(10), 2926-2935.

Contact information:

T. YUVARAJA, Associate Professor
(Corresponding Author)
Department of ECE,
Kongunadu College of Engineering and Technology,
Thottiyam, India
E-mail: kstyuvaraja@gmail.com

S. SUMITHRA, Professor
Department of ECE,
J.J College of Engineering and Technology,
Trichirappalli, India
E-mail: drsumithra30@gmail.com

M. JEYALAKSHMI, Associate Professor
Department of ECE,
SSM Institute of Engineering and Technology,
Dindigul, India
E-mail: jeyame20@gmail.com

M. PREMKUMAR, Associate Professor
Department of ECE,
SSM Institute of Engineering and Technology,
Dindigul, India
E-mail: prem53kumar@gmail.com