International Journal of Data Science and Artificial Intelligence (IJDSAI)
Volume 3, Issue 2, March – April (2025)

RESEARCH ARTICLE

SYBIL ATTACK DETECTION IN VANET USING CNN BASED LSTM FRAMEWORK MODEL

Mageshkumar N.[1,*] and Joshua Bapu J. [2]

[1] Assistant Professor, Department of Computer Science and Technology, Madanapalle Institute of Technology & Science, Andhra Pradesh, India

[2] Associate Professor, Department of Electronics and Communication Engineering, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Tamil Nadu, India

[*]*Corresponding e-mail: mageshkumarns19@gmail.com*

Abstract – Vehicular Ad Hoc Networks (VANETs) is a promising technology for autonomous driving that provides many benefits to the user's conveniences by improving road safety and driving comfort. However, Sybil attack is a major threat in VANETs where an attacker uses multiple fake identities to send false messages, disrupting safety-related applications. To overcome this, a novel jelly Fish Optimized deep Learning for Sybil Attack DETECTion (FOSA-DETECT) has been proposed to detect sybil attack in VANETs and raise the traffic efficiency. The proposed FOSA-DETECT method utilizing the Jelly Fish Optimization for feature extraction to enhance classification accuracy. Convolutional Neural Network (CNN) based Long Short-Term Memory (LSTM) classification technique classifies the extracted feature into two classes such as Normal and Attacked. Measures including specialty, F1score (F1S), Accuracy, Precision (PR), Recall (RC) are used to assess the suggested approach. Compared to current models, experimental results using VeReMi datasets show higher accuracy. The accuracy of the FOSA-DETECT in the VeReMi dataset is 1.6%, 2.8%, and 2.1% higher than that of the current MDFD, MA-DQN and I-LeeNet techniques respectively.

Keywords *– Vehicular Ad Hoc Networks, Sybil attack, Jelly Fish Optimization, Convolutional Neural Network (CNN) based Long Short-Term Memory (LSTM).*

1. INTRODUCTION

T VANETs is a promising technology to address the challenging issues in the intelligent transportation system (ITS) such as accident avoidance, traffic monitoring and transport efficiency. VANETs enable a vehicle to directly communicate with neighboring vehicles (vehicle-to-vehicle, V2V) as well as roadside infrastructures (vehicle-to-infrastructure, V2I) [1,2]. Here, the group members exchange information about their own statuses as well as other information about the current environment and traffic patterns. This makes VANETs particularly vulnerable to a type of attack known as a Sybil attack [3].

According to a report published by National Highway Traffic Safety Administration, VANETs can provide a wide range of communication-based vehicle safety and non-safety applications such as intersection collision avoidance, cooperative collision warning, emergency electronic brake lights, traffic flow control and enhanced route guidance and navigation. Nevertheless, VANETs inherit all security vulnerabilities from the wireless networks, which becomes the major issue to apply this technology into practice [4,5]. Many types of attacks can be launched in VANETs, but one of the most harmful is Sybil attack [6]. As mentioned, many safety or non-safety applications in VANETs such as cooperative collision warning and enhanced route guidance and navigation need cooperation of other vehicles. This requires one vehicle gets enough credible information from legitimate vehicles [7,8].

However, in Sybil attack, malicious node generates multiple fake identities to create many untrusted virtual nodes (Sybil nodes) in VANETs. This violates the fundamental assumption in implementing those applications. Fake information reported by a single malicious vehicle may not be highly convincing because most of the VANET applications require several vehicles to reinforce a particular information before accepting as a truth [9-11]. A Sybil attacker pretends multiple vehicles in order to reinforce false messages. To overcome these issues, a novel APPROACH has been proposed to detect sybil attack in VANETs accurately [12,13]. The following are the main contributions of the suggested work:

- The main objective of the study is to provide an efficient method to detect sybil attack in VANETs to enhance the traffic efficiency.
- Improve data quality, the collected data like speed, location, and road conditions from Vehicular networks through V2V communication are first pre-processed using data cleaning, data logging and data reduction.
- The FOSA-DETECT technique leverages feature extraction by using Jelly Fish Optimization to capture

features efficiently and raise the classification model's accuracy.

- The extracted features are then used to train the CNN-LSTM model for classification, allowing it to correctly classify data into two classes such as Normal and Attacked.
- The effectiveness of the suggested technique is evaluated utilizing parameters like precision, accuracy, recall (RC), f1-score.

The remainder of the document is arranged as follows. The literature review is thoroughly discussed in part II. A description of the Sybil attack detection in VANETs is provided in part III. The results and observations of the experiment are presented in part IV. The conclusion and next steps are in part V.

2. LITERATURE REVIEW

In 2023, Chulerttiyawong, D. and Jamalipour, A., [14] suggested an intelligent Sybil attack detection approach for FANETs-based Iot using physical layer to identified a threat on flying ad hoc network (FANET) paradigm. This proposed scheme can achieve a high correct classification accuracy of above 91%.

In 2023, Chen, Y., et al [15] suggested an multi-source data fusion detection (MDFD) framework for Sybil attacks to identify Sybil attacks often hide the real identity of the attacker with the help of a legitimate pseudonym. The average detection accuracy of the MDFD framework for four types of compound attacks is as high as 97.69%.

In 2024, Sultana, R., et al [16] suggested a deep learning-based Sybil Attack Detection mechanism to identifies the Sybil nodes by detecting the associativity between senders at a time instant using common vehicle. It achieves improved detection performance through convolutional neural network (CNN) and a combination of CNN with long short-term memory (LSTM) models in varying network scenarios.

In 2024, Rajendra, Y., et al [17] suggested a novel Sybil attack detection scheme that leverages Verifiable Delay Functions (VDFs) and location data to solve the issue that leading to fabricated congestion reports and corrupting traffic management data.

In 2024, Bhanja, U et al [18] suggested a novel is proposed using fuzzy logic controllers (FLCs) to detect both the Sybil and the DDoS attacks in VANET. The proposed model yields better accuracy, sensitivity, and recall value compared to the existing techniques. Margin of error for the attack detection is also estimated for 95% of the confidence interval.

In 2024, Suman, P., et al [19] suggested an Improved LeeNET (I-LeeNet) architecture to identify and mitigate unidentified attack types to provide real-time solutions to unidentified attacks. The average accuracy achieved by the proposed method is 97.21% on i-VANET, 97.75% on ToN-IoT, and 96.66% on the CIC-IDS 2017 dataset.

In 2025, Ajin, M. and Shaji, R.S., [20] suggested an effective Sybil attack detection model namely Adaptive Bald Eagle Search Optimization (ABESO) based Multi-Agent-Deep Q Neural network (MA-DQN). The detection results affirm that the efficiency of the proposed ABESO based DQN approach is superior and outperformed previous methods.

3. PROPOSED METHOD

In this portion, a novel FOSA-DETECT has been proposed to detect sybil attack in vehicular network accurately. Initially the datas are gathered from vehicle sensors and Road Side Units (RSUs) includes speed, location and IDs. The pre-processed utilizing methods such as Data cleaning, Data reduction and Data logging to enhance the data quality. After pre-processing the essential features are extracted using Jelly Fish Optimization technique to better Detection of Security Threat. Finaly the extracted features are fed into CNN-LSTM model which classifies the data into two classes such as Normal and Attacked. In Figure 1, the FOSA-DETECT approach's overall workflow is displayed.

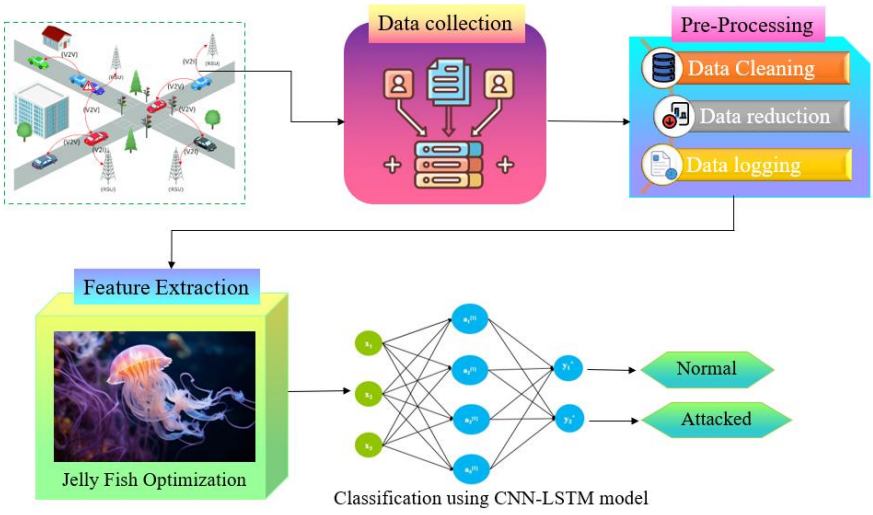

Figure.1 Proposed FOSA-DETECT Methodology

3.1 Data Collection

VANET are networks of moving automobiles that serve as either nodes or routers for message exchanges between other vehicles. Here sensors like GPS camera, RADAR are used to capture data from vehicles. These data's have information about speed, location, vehicle ID etc.

3.2. Data Pre-processing

This process removes duplicates, fills missing values that may cause incorrect or misleading analyses. Pre-Processing stage enhances the quality of data for accurate classification.

Data Cleaning

Data cleaning entails fixing errors, eliminating outliers, and adding missing values. Sensor datasets that contain broken sensor data and settling time zones must be cleaned. Errors in the data collection process or problems with the sensor itself are recorded as broken sensor data.

Data reduction

Data reduction technique is used to simplify large datasets by reducing their size while keeping important information. It helps improve processing speed and model performance. In VANETs, it helps to remove unnecessary data, making attack detection faster and more efficient.

Data logging

Data logging refers to the process of recording and storing information about the data during the pre-processing stage to ensure transparency, monitoring and perform debugging. In VANET records vehicle and communication data for analysis and security to supports attack detection.

3.3. Feature Extraction using Jelly fish Optimization

In feature extraction, vehicular network data is extracted from unprocessed data. The FOSA-DETECT technique leverages feature extraction by using Jelly fish Optimization to capture features efficiently and raise the classification model's accuracy. The search-feeding behaviour and drive designs of jellyfish in the water served as an inspiration for the development of the algorithm. Jellyfish Optimization Algorithm is based on both exploration and exploitation Phase.

Exploration Phase

In JSO, the variety of food at various locations are different, hence comparing the food proportions through the jellyfish, the optimal position is traced efficiently. This randomness helps in capturing diverse patterns, similar to collecting varied features in VANETs for effective attack detection.

$$Z_i(t+1) = Z_i(t) + rand(0,1) * \left(Z'' - \beta * (0,1)\right) \times \mu \quad (1)$$

Where, $Z_i(t+1)$ is relevant updated position, β is coefficient of the distribution ($\beta > 0$), μ is mean location; In the swarm motion, the jellyfish are passive (type A, i) as well as active (type B, i).

$$Z_i(t+1) = Z_i(t) + \gamma * rand(0,1) * (U_n - L_n) \quad (2)$$

Where, $\gamma > 0$ is motion coefficient, $U_n - L_n$ is upper and lower bound values. Type B is a jellyfish motion (j) opposite to type A motion.

Exploitation Phase

In spiral foraging of Jellyfish Optimization, a few fish lead the movement in a spiral path, influencing nearby fish to follow and form a unified hunting group. Similarly, in VANET feature extraction, key data patterns guide the selection of related features, forming a consistent set for accurate attack detection. This coordination improves optimization and classification.

$$Z_i = Z_a + rand(0,1) \times (Z_b - Z_c) \quad (3)$$

Where, Z_i is new location of the jellyfish, $Z_a, Z_b \; and \; Z_c$ arbitrarily selected positions which are described through the population size. The selected features are then provided to the classification process.

3.4. Classification using CNN-LSTM model

The CNN-LSTM method is used for classification. This technique is an efficient choice for classification due to its capability to utilize the merits of both convolutional and recurrent networks, integrated with the power of attention mechanisms. The CNN-LSTM model gives superior accuracy, robustness, and the ability to handle diverse data types, making it highly effective for classification tasks. Algorithms such as Convolutional Neural Network (CNN) based Long Short-Term Memory (LSTM) have demonstrated their efficacy in analyzing malware by examining their behavioural patterns. These methodologies are revolutionizing the approach to studying malicious software and augmenting our capacity to mitigate emerging security risks.

The collaboration method amongst gates makes LSTM more adaptable in memory and learning. LSTM frequently exceeds the conventional RNN. The LSTM unit structure, with equation is described as follows,

$$j_t = \sigma\left(Z_j y_t + V_j g_{t-1} + d_j\right) \quad (4)$$

$$f_t = \sigma\left(Z_f y_t + V_f g_{t-1} + d_f\right) \quad (5)$$

$$p_t = \sigma\left(Z_p y_t + V_j g_{t-1} + d_p\right) \quad (6)$$

whereas j_t, f_t, and p_t correspondingly relate toward switching states of the input, output, and forget gates,$p_t$exists the present unseen state; y_t exists input sequence value at the present time-step, Z and V signify the three gates weight matrices, d denotes the bias vector; and σ characterizes the activation function of sigmoid.

$$s_j = F\left(P, M_j\right) \quad (7)$$

$$\alpha_i = softmax(s_j) = \frac{\exp(s_j)}{\sum_{i=1}^{N} \exp(s_j)} \quad (8)$$

The first attention contains P, M, and V from input features. V signifies the input feature vectors, and P and M are feature vectors applied to estimate the attention weights. If an attention network is not presented, then a single set of P's is required to input it for training the network.

CNN extracts spatial patterns, while LSTM captures time-based behavior. The model learns to distinguish between normal and attacked data. Finally, it classifies each input as either Normal or Attacked.

4. RESULT AND DISCUSSION

This section explores the effectiveness of the suggested strategy using a range of evaluation measures and gives the experimental findings. Here, the suggested categorization model is implemented using a Python programming tool. Accuracy, precision, F1-score and recall are the evaluation metrics used in the suggested approach.

4.1. Dataset Description

The proposed FOSA-DETECT model is evaluated for vehicular network attack detection using a publicly available Vehicular Reference Misbehavior dataset (VeReMi) dataset. This VeReMi dataset is a benchmark dataset used for evaluating misbehavior detection in VANETs. It contains simulated vehicle communication data with various types of malicious and normal.

4.2. Evaluation Metric

This section explains the measures that were used to evaluate the suggested approach. The effectiveness of the recommended strategy has been evaluated using the F1-Score, Recall, Precision and Accuracy measures.

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (9)$$

$$\mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{FP+Tp}} \quad (10)$$

$$\mathrm{F1-score} = \frac{2*(\mathrm{PR}*\mathrm{Recall})}{(\mathrm{PR+Recall})} \quad (11)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{(\mathrm{TP+FN})} \quad (12)$$

4.3 Performance Analysis

According to the experimental results, the suggested FOSA-DETECT technique has been compared with current techniques for sybil attack detection in vehicular networks, including MA-DQN, MDFD and I-LeeNet.

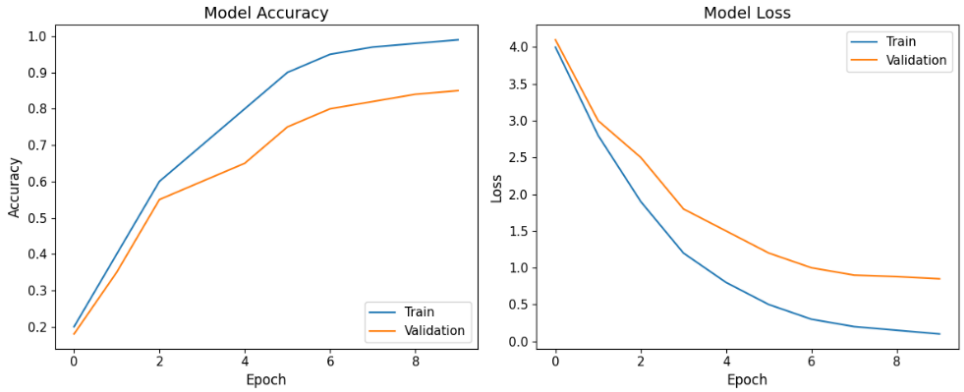

Figure 2. Accuracy and Loss curve for FOSA-DETECT method

Figure 2 shows the Model accuracy and loss over 10 epochs, training accuracy nears 100%, validation accuracy reaches 85%, both losses decrease steadily. The validation loss stays slightly higher, indicating effective learning with minimal overfitting. The consistent gap between training and validation metrics suggests stable model generalization.

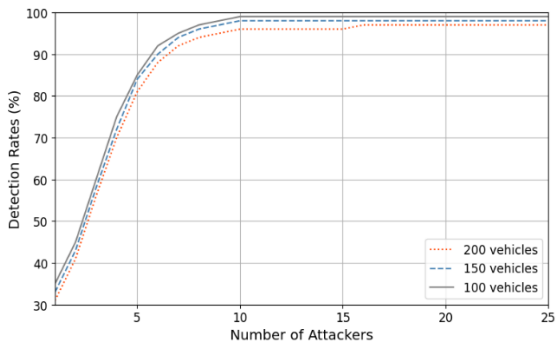

Figure 3. Detection rate vs. attackers for different vehicle counts.

Figure.13 shows that the attack detection rate and the number of attackers for three scenarios with 100, 150, and 200 vehicles, the detection is higher when the number of attackers increases independently from the number of vehicles. The number of vehicles has little effect on the detection rate because the number of attackers is very small compared to the total vehicles.

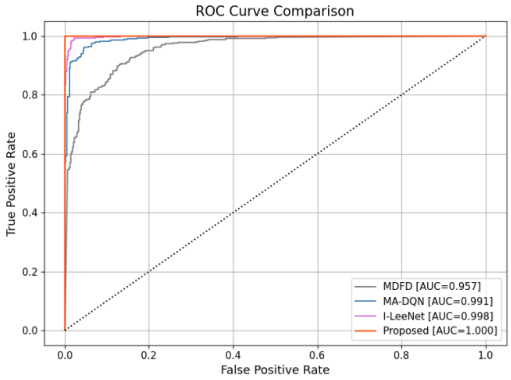

Figure 4. ROC curve

Figure.4 shows the ROC curve comparison graph for MDFD, MA-DQN, I-LeeNet, and proposed, it achives the lowest FPR. It indicates improved reliability and reduced

misclassification. The proposed FOSA-DETECT model achieves the highest AUC indicating near-perfect detection performance with minimal false positives.

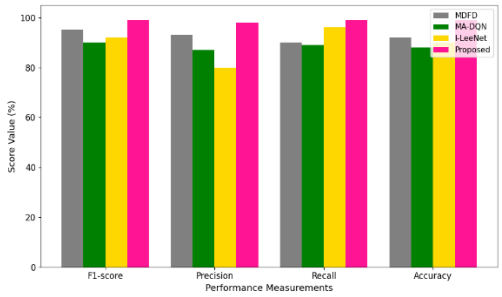

Figure 5. Performance Comparison

Figure 5 shows the Performance comparison of four models as MDFD, MA-DQN, I-LeeNet and Proposed across F1-score, Precision, Recall, and Accuracy. The Proposed FOSA-DETECT model consistently outperforms others with the highest scores in all metrics, indicating superior classification performance.

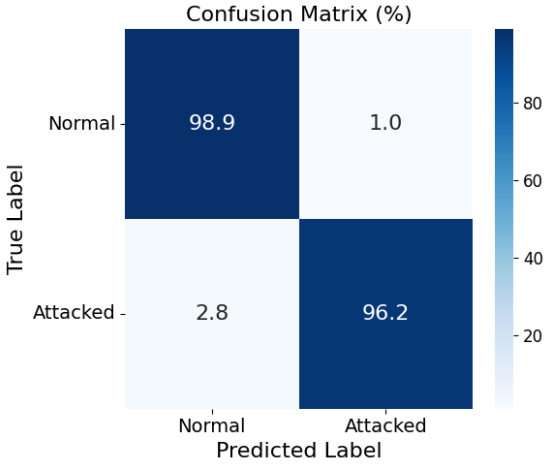

Figure 6. confusion matrix

Figure6 illustrates the Confusion matrix showing performance between Normal and Attacked classes. The model achieves 98.9% accuracy for Normal and 96.2% for Attacked, with minimal misclassification, indicating strong detection capability and highlighting its effectiveness in distinguishing between normal and malicious behavior.

5. CONCLUSION

In this research, a novel FOSA-DETECT has been proposed to detect Sybil attack in vehicular networks to enhancing security of the system. The Jelly fish Optimization is utilized for feature extraction in order to efficiently capture features and improve the classification model's accuracy. By fed these extracted features into the CNN-LSTM method, the network is able to classify data reliably into two classes: Normal and attacked. In addition to increasing attack detection, this FOSA-DETECT method raises Improved Traffic Management and reduced Network Overhead. For the VeReMi dataset, proposed FOSA-DETECT method F1-Score, accuracy, recall, precision are 98.61%, 96.21%, 96.85%, and 97.34%, respectively. The accuracy of FOSA-DETECT methodology is 2.8%, 1.6%, and 2.1% higher than the current MDFD, MA-DQN and I-LeeNet methods. In future, this research can be extended to detect various attacks, such as tunneling attacks, impersonation attacks, etc. This can be done by simulating the environment of the security attacks and then detecting the malicious information broadcasted using rogue nodes detection techniques.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING STATEMENT

Not applicable.

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES

[1] A. Balaram, S.A. Nabi, K.S. Rao, and N. Koppula, "Highly accurate sybil attack detection in vanets using extreme learning machine with preserved location", *Wireless Networks*, vol. 29, no. 8, pp. 3435-3443, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] N.C. Velayudhan, A. Anitha, and M. Madanan, "Sybil attack detection and secure data transmission in VANET using CMEHA-DNN and MD5-ECC", *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 2, pp. 1297-1309, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] P.K. Pareek, P. Siddhanti, S. Anupkant, S.O. Husain, and I. Bhuvaneshwarri, "Ensemble Based Machine Learning Classifier for Sybil Attack Detection in Vehicular AD-HOC Networks (VANETS)", *In 2024 Second International Conference on Data Science and Information System (ICDSIS),* pp. 1-4, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] S.M. Mathur, N. Jadhav, and R. Gupta, "Sybil Assault Detection in VANET with Efficient Network Analysis*", In 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon)*, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] H.B. Tulay, and C.E. Koksal, "Sybil attack detection based on signal clustering in vehicular networks", *IEEE Transactions on Machine Learning in Communications and Networking*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Y. Zhong, H. Yang, Y. Li, B. Yang, X. Li, Q. Yue, J. Hu, and Y. Zhang, "Sybil Attack Detection in VANETs: An LSTM-Based BiGAN Approach", *In 2023 International Conference on Data Security and Privacy Protection (DSPP),* pp. 113-120, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] P. Remya krishnan, and R. Koushik, "Decentralized Distance-based Strategy for Detection of Sybil Attackers and Sybil Nodes in VANET", *Journal of Network and Systems Management*, vol. 32, no. 4, pp. 91, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] A. Barve, and P.S. Patheja. A hybrid deep learning based enhanced and reliable approach for VANET intrusion detection system. Cluster Computing, vol. 27, no. 9, pp. 11839-11850, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] S. Rakhi, and K.R. Shobha, "LCSS based Sybil attack detection and avoidance in clustered vehicular networks", *IEEE Access*, 11, pp. 75179-75190, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] H.S. Patel, M.F. Khan, A. Rahbar, and N.P. Yashwanth, January. XGSybil: A Framework for Modeling and Identifying Sybil Attacks in Intelligent Transportation Systems. In 2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-7, 2025[CrossRef] [Google Scholar] [Publisher Link]

[11] Y. Wei, W. Qi, Z. Li, Y. Han, and Y. Lai, "MVDetector: Malicious Vehicles Detection Under Sybil Attacks in VANETs", *In International Conference on Information Security (pp. 307-322). Cham: Springer Nature Switzerland*, (2024). [CrossRef] [Google Scholar] [Publisher Link]

[12] R. Sultana, J. Grover, and M. Tripathi, "Intelligent defense strategies: Comprehensive attack detection in VANET with deep reinforcement learning", *Pervasive and Mobile Computing*, 103, p.101962, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] A. Borah, and A. Paranjothi, "Sybil Attack Detection in VANETs using Fog Computing and Beamforming", *In 2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON),* pp. 0626-0631, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] D. Chulerttiyawong, and A. Jamalipour, "Sybil attack detection in internet of flying things-iot: A machine learning approach", *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 12854-12866, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Y. Chen, Y. Lai, Z. Zhang, H. Li, and Y. Wang, "MDFD: A multi-source data fusion detection framework for Sybil attack detection in VANETs", *Computer Networks*, vol. 224, pp. 109608, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] R. Sultana, J. Grover, M. Tripathi, M.S. Sachdev, and S. Taneja, "Detecting Sybil attacks in VANET: exploring feature diversity and deep learning algorithms with insights into sybil node associations", *Journal of Network and Systems Management*, vol. 32, no. 3, pp. 51, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Y. Rajendra, V. Subramanian, and S.K. Shukla, "Sybil attack detection in ultra-dense VANETs using verifiable delay functions", *Peer-to-Peer Networking and Applications*, vol. 17, no. 3, pp. 1645-1666, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] U. Bhanja, A. Majhi, S. Sahu, and D. Parida, "Detection of Sybil & DDoS attacks in VANET using intelligent technique", *International Journal of Computers and Applications*, vol. 46, no. 10, pp. 811-829, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] P. Suman, S. Padhy, N. Kumar, A. Suman, A. Singh, K.K. Singh, Á.K. Castilla, and T.S.S. AL-Zahrani, "An improved deep learning-based intrusion detection for reliable communication in VANET", *IEEE Transactions on Consumer Electronics*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] M. Ajin, and R.S. Shaji, "Enhancing Security in Vanets: Adaptive Bald Eagle Search Optimization Based Multi-Agent Deep Q Neural Network for Sybil Attack Detection", *Vehicular Communications*, pp. 100928, 2025. [CrossRef] [Google Scholar] [Publisher Link]

AUTHORS

Mageshkumar N, M.Tech, Ph.D., currently serves as an Assistant Professor in the Department of Computer Science & Technology at Madanapalle Institute of Technology & Science (an Autonomous Institution) in Madanapalle, Andhra Pradesh – 517325. He obtained his Bachelor's Degree in Computer Science and Engineering from the University of Madras in 2004. Later, he completed his Master's in Information Technology at Sathyabama Institute of Science and Technology in 2009, followed by a Ph.D. from the same institute. His areas of specialization include Network Communication, Cryptography and Network Security, Cloud Computing, Big Data, and Information Retrieval. With over 15 years of teaching experience across various Engineering Colleges, he has been honored with the Best Teacher Award twice. Dr. Mageshkumar is a member of IAENG and CSTA and has authored more than 10 papers in reputed international journals and presented over 10 papers in IEEE International Conferences.

Joshua Bapu J received his B.E degree in Electronics and Communication Engineering in 2006 from Anna University, Chennai, India , M.E degree in Applied Electronics in 2009 from Anna University, Tirunelveli, India and Ph.D in Information and Communication Engineering in 2022 from Anna University, Chennai, India . Presently he is working as an Associate Professor in Department of Electronics and Communication Engineering at Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, India. His research interests includes Image Processing and Embedded systems. He has published several Research papers in International Journals. He is a life time member of Life member in Indian Society for Technical Education (ISTE) and Life member in International Association of Engineers (IAENG).

Arrived: 20.03.2025
Accepted: 22.04.2025