CHAPTER 3
PROPOSED METHODOLOGY

3.1 Introduction

Building on the foundation laid in Chapters 1 and 2, this chapter delves into the core of the research: a novel methodology to safeguard Vehicular Ad-hoc Networks from Denial-of-Service attack and types. Chapter 2 presumably focused on the vulnerabilities of VANET, particularly their susceptibility to DoS attacks. Moreover, the various types of DoS attacks, their impact on network performance, and the limitations of existing security solutions are provided. The performance degradation, manifested in reduced packet delivery rates, increased end-to-end delays, and impaired channel access, poses a significant risk to the safety of passengers and pedestrians. The attacks and its impact on application, transport and network layers are the modern day attacks to be identified as the attacks in the literature provided had limited scope.

This chapter unveils our proposed methodology for fortifying VANET security against DoS attack and types. The considerable deterioration in the network's performance must be handled due to its notable open and dynamic nature. The major objective is formulated to device a methodology to secure VANET against DoS attack and types by mitigation using the hybrid approach. The hybrid approach also includes handling of the undesirable consequences related to DoS attacks and types.

Despite advancements in DoS attack mitigation techniques for VANET, traditional security measures like authentication and digital signatures can increase processing time, potentially leading to communication delays. Moreover, these methods may hinder the establishment and maintenance of trust relationships between vehicles within the network. Existing researches based on machine learning for securing vehicular networks against DoS attacks and its variants secured the VANET but exposed to certain vulnerabilities due to its dependence on centralized entities. Artificial Immune System (AIS) based security mechanisms proposed in previous research had limited evaluation on the impact of AIS-based approach on network efficiency and delay. Mitigating DoS attacks using immune algorithms, a type of

artificial immune system (AIS), blocked malicious traffic but adequately not addressing false positives or negatives in detecting malicious behavior.

This chapter focuses on the methodology employed to secure VANET against DoS attacks by addressing the mentioned challenges in the existing research techniques. The hybrid approach proposed combines trust-based approach, machine learning and immune methods to detect and mitigate DoS attack and types effectively. Enhancing the resilience of VANET against malicious DoS attacks in ensuring the continuous and secure operation of vehicular communication networks is also considered for the safety and reliability of these critical networks.

To fulfill the major objective, a hybrid framework is proposed. The secondary objectives focus on achieving the following:

i. Minimization of Packet Loss
ii. Maximization of Throughput
iii. Maximize Accuracy of Detection and Classification
iv. Minimization of Processing Time
v. Maximization of Recall and Precision
vi. Minimization of Delay
vii. Maximization of Packet Delivery Ratio

The major objective is:

To device a three-phase methodology in securing VANET against DoS attack and types by mitigation using the hybrid approach.

The secondary objectives that aim to fulfill the major objective are:

i. To detect and classify DoS attacks with improved accuracy of detection and classification, recall, precision, minimum delay through optimizing the features.

ii. To detect, predict and isolate (Mitigate) DoS attack and types with increased accuracy of detection, recall, precision, minimum delay, improved packet delivery ratio.

iii. To enhance the security and reliability of VANET services with minimum packet loss, maximum throughput, minimum processing time and maximum packet delivery ratio.

Secondary objective (i) focuses on the detection and classification of DoS attacks. It aims to improve the accuracy, recall, precision, and speed of this process through feature optimization. Secondary objective (ii) expands on this by adding prediction and isolation (mitigation) of DoS attacks to the scope. It also introduces a new performance metric, packet delivery ratio. The objective (ii) is related to objective (i) with objective (ii) building upon the foundation of objective (i). It assumes that accurate detection and classification are prerequisites for effective prediction and mitigation. Both objectives share the common goals of improving accuracy, recall, precision, and minimum delay. Objective (ii) takes the process a step further by incorporating prediction to anticipate future attacks and isolation to proactively counter them.

3.2. Phases in the Proposed Methodology

A three-phase methodology is followed to fulfill the mentioned objectives and addressing the challenges in securing the VANET. The three significant steps are: Enhanced Feature Selection and Mitigation, Strengthening the Access Control and Mapping, and Immunization of Clusters and Routing.

The VANET scenario is simulated with nodes and traffic is generated using CIC - IDS 2018 dataset. The three-phase methodology is explained below. The initial step considered the dataset CIC - IDS 2018 for the feature selection and for the enhanced mitigation with further steps, the dataset considered is CIC – DDoS2019. The CIC – DDoS2019 (The 2019 Canadian Institute for Cybersecurity – Intrusion Detection Systems Evaluation Dataset) focusing on cyber security aspects. The dataset likely contains information on the attack patterns through network traffic characteristics.

The CIC – IDS2018 dataset included information of attacks relating to VANET communications like DoS, DDoS and Botnet. The network traffic characteristics are the features that differentiate malicious traffic from legitimate traffic. The features included are:

- Packet sizes
- Flow durations
- Source and destination IP addresses
- Protocol
- Flags and timestamps

As VANET rely on wireless communication, they are vulnerable to security attacks. Understanding DDoS attacks as in CIC-DDoS2019 are helpful for securing communications and designing intrusion detection systems for VANET. New attack vectors, such as reflection-based and exploitation-based attacks, inundate victims with overwhelming traffic. These attacks exploit vulnerabilities to target specific victim IP addresses while concealing the attacker's identity.

In this thesis, the DoS attack and types pertaining to VANET are taken and the VANET is simulated. The VANET simulation involves creating a network of virtual nodes representing vehicles and road side infrastructures. An open-source network simulator NS-3 is well-suited for simulating VANET communication scenarios. The built-in mobility models in NS-3 simulate the vehicular movements. The nodes, features and the network parameters with the values are explained in Chapter 4.

In this initial step, the dataset is considered with the VANET setup simulated representing the vehicles as nodes. The nodes programmed to move according to the traffic patterns and communicate with each other using designated communication protocols. The Ad hoc On-Demand Distance Vector (AODV) routing protocol for communication and is a popular choice for VANET due to its reactive nature and efficient resource management.

Following the dataset and the simulation is the three-phase methodology with the hybrid approach for securing VANET by mitigating DoS attack and types. The proposed methodology is shown in Figure 3.1.

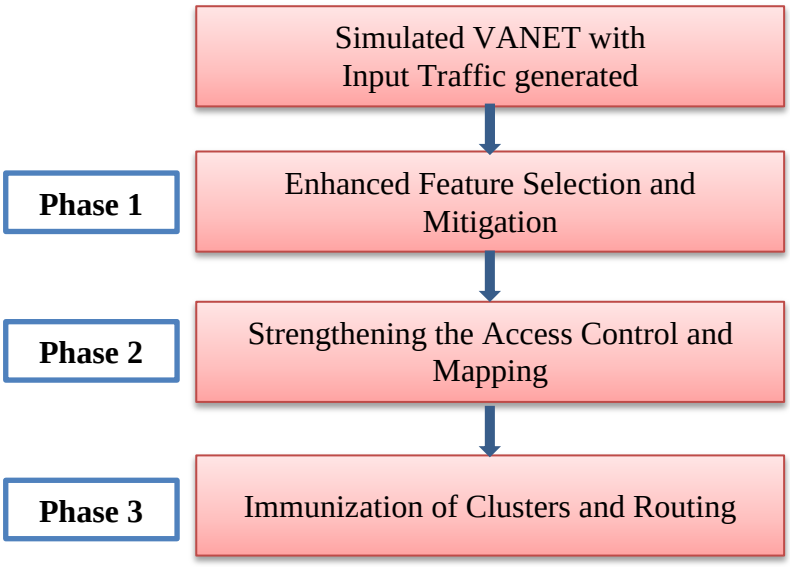

Figure 3.1.Proposed Three-phase Methodology

Phase 1: Enhanced Feature Selection and Mitigation

In the phase 1, a feature selection approach is implemented to detect attack patterns from the simulated VANET. Based on the fitness function, the attacks are detected and classified. The features from the traffic generated are calculated. The list of the features included in the CIC - IDS2018 dataset are:

i. **Flow Duration:** This feature measures the duration (in seconds) of a network flow.

ii. **Total Fwd Packets:** The total number of packets sent from the source to the destination.

iii. **Total Backward Packets:** The total number of packets sent from the destination to the source.

iv. **Total Length of Fwd Packets:** The total length (in bytes) of all packets sent from the source to the destination.

v. **Total Length of Bwd Packets:** The total length (in bytes) of all packets sent from the destination to the source.

vi. **Fwd Packet Length Max:** The maximum length (in bytes) of a packet sent from the source to the destination.

vii. **Fwd Packet Length Min:** The minimum length (in bytes) of a packet sent from the source to the destination.

viii. **Bwd Packet Length Max:** The maximum length (in bytes) of a packet sent from the destination to the source.

ix. **Bwd Packet Length Min:** The minimum length (in bytes) of a packet sent from the destination to the source.

x. **Fwd Packet Length Std:** The standard deviation of the forward packet lengths.

xi. **Bwd Packet Length Std:** The standard deviation of the backward packet lengths.

xii. **Flow Bytes/s:** The average number of bytes transferred per second for the flow.

xiii. **Flow Packets/s:** The average number of packets transferred per second for the flow.

xiv. **Fwd IAT Mean:** The average time (in milliseconds) between transmissions of consecutive packets in the forward direction.

xv. **Fwd IAT Max:** The maximum time (in milliseconds) between transmissions of consecutive packets in the forward direction.

xvi. **Fwd IAT Min:** The minimum time (in milliseconds) between transmissions of consecutive packets in the forward direction.

xvii. **Fwd IAT Std:** The standard deviation of the forward inter-arrival times.

xviii. **Bwd IAT Mean:** The average time (in milliseconds) between transmissions of consecutive packets in the backward direction.

xix. **Bwd IAT Max:** The maximum time (in milliseconds) between transmissions of consecutive packets in the backward direction.

xx. **Bwd IAT Min:** The minimum time (in milliseconds) between transmissions of consecutive packets in the backward direction.

xxi. **Bwd IAT Std:** The standard deviation of the backward inter-arrival times.

xxii. **Fwd PSH Flags:** The number of packets with the PSH flag set in the forward direction.

xxiii. **Bwd PSH Flags:** The number of packets with the PSH flag set in the backward direction.

xxiv. **Fwd URG Flags:** The number of packets with the URG flag set in the forward direction.

xxv. **Bwd URG Flags:** The number of packets with the URG flag set in the backward direction.

xxvi. **Fwd FIN Flag:** The number of packets with the FIN flag set in the forward direction.

xxvii. **Bwd FIN Flag:** The number of packets with the FIN flag set in the backward direction.

xxviii. **Total Fwd Headers Length:** The total length (in bytes) of all forward packet headers.

xxix. **Total Bwd Headers Length:** The total length (in bytes) of all backward packet headers.

xxx. **Fwd Packets/s:** The number of packets transmitted from the source to the destination per second.

xxxi. **Bwd Packets/s:** The number of packets transmitted from the destination to the source per second.

xxxii. **Min Packet Length:** The minimum length (in bytes) of any packet in the flow.

xxxiii. **Max Packet Length:** The maximum length (in bytes) of any packet in the flow.

xxxiv. **Packet Length Std:** The standard deviation of the packet lengths.

xxxv. **Flow Byts/s:** The average rate of byte transfer over the duration of the flow

xxxvi. **Flow Pkts/s:** The average number of packets transferred per second for the flow

xxxvii. **Fwd IAT Std:** The standard deviation of the forward inter-arrival times

xxxviii. **Bwd IAT Std:** The standard deviation of the backward inter-arrival times

xxxix. **Fwd Header Length:** The average length (in bytes) of forward packet headers.

xl. **Bwd Header Length:** The average length (in bytes) of backward packet headers.

xli. **Fwd Packets/s:** The number of packets transmitted from the source to the destination per second.

xlii. **Bwd Packets/s:** The number of packets transmitted from the destination to the source per second.

The calculated features including the number of transmitted TCP packets, the number of received packets, signal strength, latency, vehicle speed, direction, throughput, timestamp, input message ID, and layer ID are listed in Table 3.1. The values observed during the execution are given in Chapter 4.

Table 3.1: Parameters calculated for the Vehicular Systems

S.No.	Parameters	Specifications
1	No of TCP packets Transmitted	8 bytes
2	No of TCP Packets Received	8 bytes
3	Speed of the vehicles	35-40km/hr
4	Signal Strength	-10 to 25 dbm
5	Latency	2 ms
6	Direction	Omni direction
7	Throughput	40-80
8	Time Stamp	4 ms
9	Input Message ID	ID1
10	Layer ID	ID2

The features calculated are selected and optimized with optimized feature selection thus paving the way to detect the malicious behaviour based on DoS attack.

An enhanced way in feature selection and mitigation is proposed that included the new attacks based on modern reflective and exploitation. The traffic features

extracted from the dataset include 80 features from the TCP and UDP traffic pertaining to DDoS attacks. In order to detect the new DoS and DDoS attacks, the message rate, maximum limit and the elapsed time are considered for the feature selection. In addition, the trustiness values with their relationships, the credibility score are considered for attack detection with classification and the self-healing effect mitigates the malicious vehicles from the VANET.

Following methods are used in phase 1:

i. Glow-worm (GLW) swarm optimization technique with Single Layer Feed Forward Neural Network (SLFN) classifier

ii. Response Feedback Algorithm with Micro Cluster Outlier Detection Algorithm using Linear Regression (MCOD-LR)

iii. Kernel Density Estimation and Entropy-based SVM classifier

iv. Bayesian aggregate model with Self-healing effect of Artificial Immune System (AIS)

Phase 2: Strengthening the Access Control and Mapping

In the first phase, significant features relevant to DoS attacks are selected. This feature selection process optimizes the feature set extracted from the generated traffic, thereby improving the effectiveness of DoS attack classification in VANET. The detection and mitigation of attacks based on modern reflective and exploitation incorporates trustiness and self-healing effect in securing the VANET by isolating malicious vehicles.

In the phase 2, strengthening the traffic with trust nodes and encrypting the messages are implemented. The AODV protocol routes the packets and high chances of congestion in the VANET persists when the number of vehicles for communication is more in number. The routing process is eased for the protocol with proposed encryption and mapping among the rational vehicles. In the process of strengthening the access control and mapping, the irrational nodes are detected and classified leading to routing efficiency.

Strengthening the traffic and mapping the vehicles are incorporated using the proposed Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping and Deep Auto Spare Impasse Neural Network.

Phase 3: Immunization of Clusters and Routing

Immunization process in VANET improve network security and resilience by proactively detecting and mitigating threats, verifying data integrity, establishing trust-based relationships, and securing resources. The immunization countermeasures are similar to vaccines that train the immune system with data sanitization and verification techniques. These ensure the integrity and trustworthiness of information circulating within the network. After strengthening the access control as VANET's natural defense in adherence to the AODV protocol, the reliable and trustworthy nodes are determined. The trust scores are calculated, updated and shared for information exchange and collaboration. The low trust vehicles are isolated or penalized. Improving the VANET security and resilience through immunization with proactive detection, isolation and routing is enforced. The proposed Deep Trust Factorization Neural Network with Trust Score and Moth Flame Optimization with Cache Parallelized Circulation Link Routing are implemented in immunization.

The contributions are spread across the three steps highlighted in the Figure 3.2. The contributions fulfilling the major objective involve the objective function in maximizing the packet delivery ratio, minimizing the packet loss and maximizing the detection rate. The objective function for the optimization problem is considered in the proposed methodology in maintaining the quality of service and ensuring reliable communication between vehicles and promptly identifying and responding to DoS attacks to prevent network disruptions.

The phases are defined to detect and mitigate the DoS attacks and its types with the inception of optimizing the features for the detection and classification of DoS and Botnet attacks. An adaptive DoS attack detection and mitigation

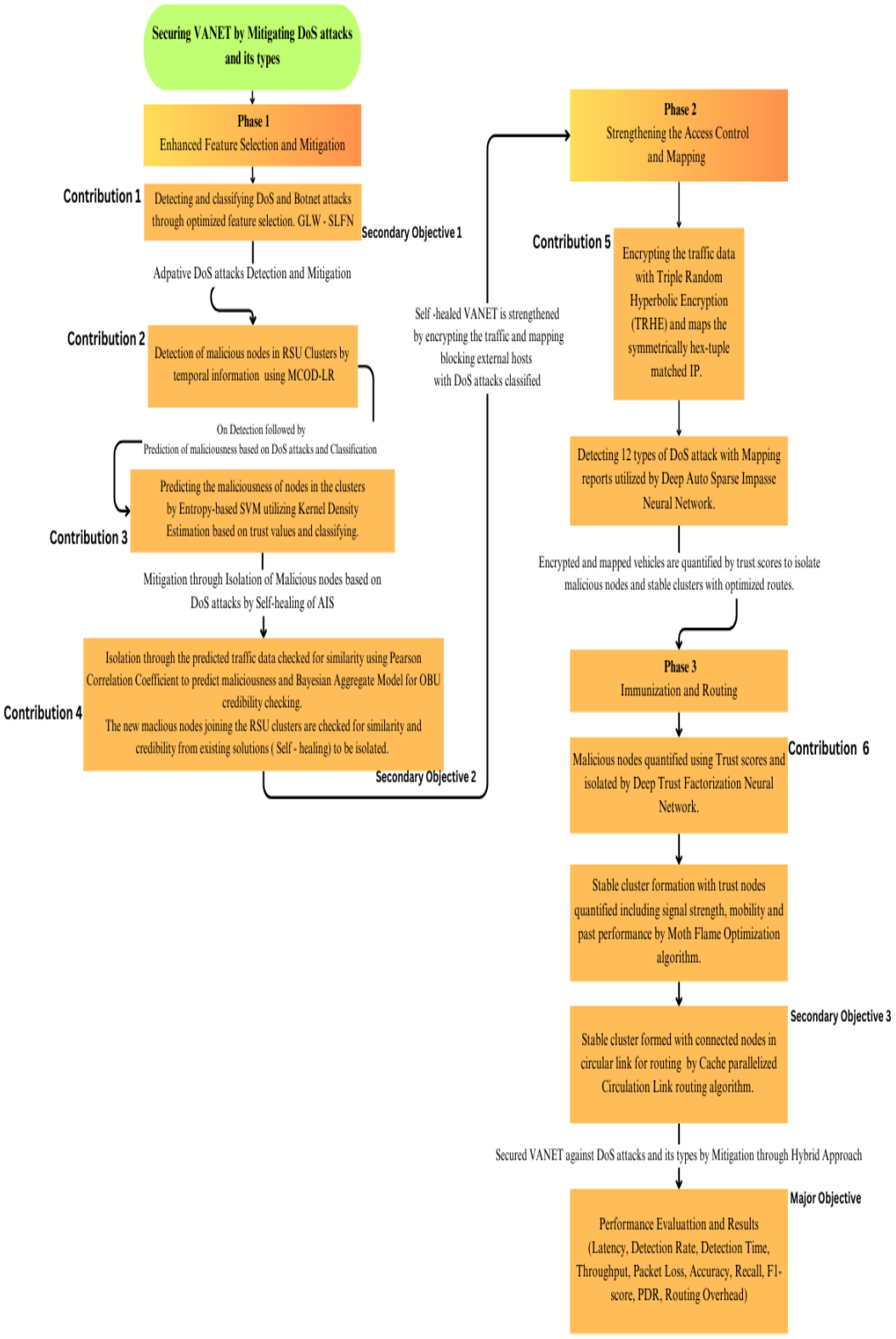

Figure 3.2 Three – phase Methodology with Contributions

3.3 Phase 1: Enhanced Feature Selection and Mitigation

3.3.1 Introduction

Glow-worm Swarm Optimization (GLW), a swarm intelligence technique, is utilized to identify the most crucial features for attack detection. This is achieved by analyzing variations in the values of these features. This optimization process improves the accuracy of attack detection. The considered attack types include Denial-of-Service (DoS) and Botnet attacks. The VANET due to its dynamic and complex nature becomes challenging in identifying the most relevant and effective features. Identifying the most relevant features from the dataset might be indicative of anomalies in VANET communication (e.g., unusual packet arrival patterns). Dataset considered is a high-dimensional containing 80 features and the features are elaborated in Chapter 4. Understanding of VANET communication and potential attacks with the dataset is essential. This allows identifying the most relevant features and tailoring the detection approach to the specific needs of VANET security.

Feature selection plays a crucial role in analyzing high-dimensional datasets. The selection of the most relevant features facilitates a reduction in data complexity, thereby enhancing the classifier's performance. Swarm intelligence-based feature selection algorithms are also particularly well-suited for high-dimensional datasets. Existing dimensionality reduction approaches based on features provided in Chapter 2 discussed on improving detection accuracy. Furthermore, accurate attack classification requires optimizing the feature set to identify the most relevant features that significantly contribute to improved detection accuracy and faster execution times. The GLW algorithm is a bio-inspired optimization technique based on the social foraging behavior observed in glowworms. This behaviour is similar to the clusters of the vehicles. Glowworms (agents) move towards brighter neighbors, eventually converging towards areas with higher glowing intensity.

Similarly, features with higher importance ("brighter glow") could attract other features. Through this metaphorical movement, features might converge into groups representing informative subsets. These subsets could then be evaluated for their effectiveness in intrusion detection. The use of the Glow-worm Swarm Optimization (GLW) technique in VANET allows for the selection of a subset of features that are most informative for detecting attacks. The glowworms represent the effective features pertaining to attacks and considered by the SLFN classifier based on the fitness measure. The attacks are detected

and classified using SLFN classifier with the optimized feature subset as the determined fitness for distinguishing between normal and attack traffic.

The existing machine learning techniques discussed in chapter 2 had limitations in dealing with the complex and evolving nature of cyber threats. VANET face constant changes due to vehicle movement, which can disrupt traditional routing protocols designed for static networks. Swarm intelligence algorithms as natural systems like ant colonies or bird flocks, are inherently adaptable to changing environments. These algorithms adjust their behavior based on real-time network conditions in VANET based on information sharing among nodes. This allows them to "learn" from past successes and failures of the collective swarm. The ability to learn from the past experiences with information sharing among nodes and its decentralized nature eliminates the single point of failure.

The contributions in the phase 1 are detailed in the following subsections.

3.3.2 Contribution 1: Optimized Feature Selection for Malicious Nodes Detection and Classification of DoS Attacks using Glow-worm (GLW) Single Layer Feed Forward Neural Network (SLFN)

In contribution 1, an optimized feature selection for malicious nodes detection based on DoS attacks using Glow-worm SLFN is applied. The objective is to optimize the features and to detect and classify DoS attacks. The contribution 1 consists of the Glowworm optimization technique and the most relevant features from GLW are fed as inputs to SLFN to detect and classify the attacks. Several features are extracted from the network environment to identify different types of attacks.

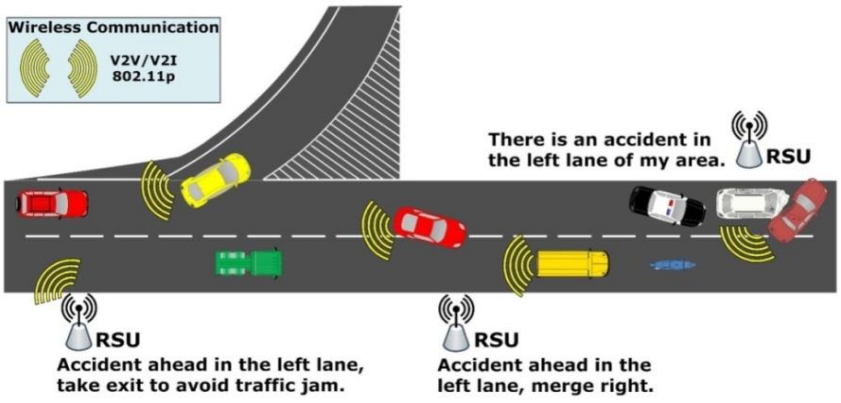

Figure 3.3: VANET Scenario with communications

The above Figure 3.3 shows the VANET scenario with the communications. Vehicles communicate using V2V communication, vehicles and roadside infrastructure communicate with one another using V2I communication and vice versa using I2V communication. The communications has the potential to revolutionize transportation safety, efficiency, and driver experience. Attacks known as DoS and DDoS, are a frequent occurrence in VANET and are designed to overwhelm or disrupt the network by saturating it with traffic or by interfering with the services it offers. DoS and DDoS attacks on VANET pose a significant threat to the transportation system by disrupting network operations. These attacks can severely compromise network security and effectiveness.

In order to successfully avoid assaults with reduced dimensionality and less processing time, the number of features must be handled based on the assaults. The presence of excessive features can introduce instability into the classifier, consequently compromising the accuracy of predictions. Therefore, the optimization technique GOF – SLFN is employed to select the optimal set of features, leading to improved performance in detecting and classifying attacks.

The novelty of this work begins with the establishment of a realistic VANET environment. This is achieved by integrating the SUMO traffic simulator with the OMNET++ network simulator. The microscopic traffic during simulation includes the following factors:

- Vehicle characteristics – maximum speed
- Road characteristics – lane, speed
- Lane change

By considering factors like microscopic traffic characteristics and the current state of each vehicle (with updates), the position of each vehicle on the road is calculated at a specific point in time. Simulation inputs encompass vehicle initial positions, their features (acceleration, maximum speed), road network topology, driving routes, terminus points, and speediness limits. Within the OMNET++ simulation environment, vehicles are modeled as nodes. From this representation, various features are extracted, including the number of transmitted and received TCP packets, signal strength, latency, vehicle speed

and direction, throughput, timestamps, input message IDs, and layer IDs. Detection of attacks relies on the features from the network. To enhance the accuracy and efficiency of traffic anomaly detection in VANET, it is crucial to identify and utilize the most relevant and significant features from the vast amount of traffic data.

VANET are characterized by their highly dynamic nature, where vehicles frequently enter and exit the network. This can make it difficult to track the behavior of individual vehicles and identify anomalous traffic patterns. In order to maintain accurate DDoS attack detection in the dynamic environment of VANET, characterized by the continuous entry and exit of vehicles, feature selection algorithms must exhibit adaptability. The identification of the most effective set of features for performance enhancement in VANET is achieved through the application of swarm optimization techniques.

The Glow-worm Swarm Optimization (GLW) technique has garnered significant attention in recent applications, demonstrating particular effectiveness in reducing the search time required to identify optimal solutions that enhance the objective function. The Glow-worm Optimization technique, modeled after the behavior of natural glowworms, incorporates a key concept called 'luciferin.' This factor represents the attractiveness of each solution, significantly influencing the interactions between 'glowworms' within the algorithm.

The proposed model began by developing a fitness function and then proceeded to design the GLW algorithm based on this function. The Glow-worm Algorithm involves a set of agents that are randomly distributed within the search space. In the Glow-worm Algorithm, each agent possesses a luminescence quantity known as luciferin. The agents are conceptualized as light emitting glowworms. Luciferin directly relates to the intensity with variable decision range ‘ri’, bounded by a circular sensor range ‘rs’. A glowworm within its current local-decision domain identifies a neighbor glowworm with brighter glow and gets attracted. Availability of information in the local-decision range facilitates the agents to make their decisions (Krishnanand, K. N et al., 2006).

The objective, defined by the fitness function, is the accurate detection of various attacks in vehicular systems. The GLW algorithm begins by initializing with random

values within the search space. The algorithm employs an iterative search process to identify the optimal value maximizing the defined fitness function. The GLW algorithm employs a three-stage process to locate the local optima within the current search space. Initially, all glowworms are assigned an equal amount of luciferin. Each iterative cycle of the Glow-worm Algorithm consists of two distinct phases.

Luciferin Update: The amount of luciferin associated with each glowworm is adjusted.

Movement Phase: Glowworms move within the search space based on a defined transition rule. The iteration has these stages as given below.

- Glow-worm attraction factor update stage
- Glow-worm movement stage
- Glow-worm neighborhood stage

3.3.2.1 (a) Glow-worm Attraction Factor Update Stage

The attraction factor is "luciferin" which produces the light enzyme on its own. Based on this attraction factor, each glow-worm interacts with neighbor glow-worms. The capability of luciferase enzyme is directly proportional to the present location of glowworm in the given search space. In this stage, it detects the local optimum value in the given search area rather than global value. Likewise, each glowworm attracted each other based on the highest luciferin values and updated with the current value. Similarly, the glowworm updates its attraction factor values nearer the optimum. The key rule of the attraction factor is given in the below equation (3.1) (Rama Mercy Sam Sigamani et al., 2020).

The glowworms hold the same luciferin value at the first cycle and the updation of luciferin has these values changed with the current position's function values. The luciferin is proportional to the sensed profile measured at that point. During the luciferin update phase, each glowworm adds to its previous luciferin level. In the case of a function optimization problem, this would be the value of the objective function at that point (Anurag Tiwari et al., 2023).

$$F_j(S+1) = (1-\delta)F_j(S) + \omega O_j(S+1) \quad (3.1)$$

where,

δ is the luciferin decay constant [0, 1]

ω denotes the luciferin enhancement constant

O_j (S+1) is the objective function at 'j' location at sample time 'S'

F_j is the attraction factor at 'j' location

The key rule of the attraction factor equation is the foundation for the movement and interaction of glow-worms in the GLW algorithm, contributing the search space exploration in finding better solutions preventing to get stuck in local optima.

The attraction factor equation is particularly significant for VANET for the following reasons:

- More efficient data transmission quickly
- Vehicles within a specific range influence each other's behavior
- Automatic adjustment ensuring efficient operation
- Self-organize and adapt to changing traffic conditions

In feature selection, the attraction factor prioritizes features from closer vehicles, implicitly favoring relevant and complementary information for each specific context. Vehicles collaborate and share information about suspected attacks through the attraction factor mechanism.

3.3.2.2 (b) Glow-worm Movement Stage

Probabilistic mechanisms are central to this stage of the Glowworm technique. Inspired by natural behavior, glowworms assess their own brightness relative to their neighbors. This comparison influences their attraction to one another. As shown in figure 3.4, the glowworms (a, b, c, and d) having high luciferin than glowworm 'e'. Since 'e' is in the sensor-overlap region of 'c' and'd', it can only in two possible directions (Krishnanand, K. N et al., 2006). Likewise, it moves entirely in the search space and detects the nearer optimum value.

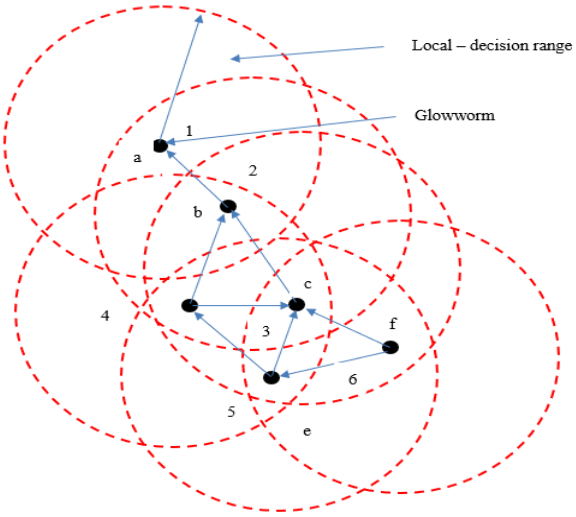

Figure 3.4: **Glowworm Movement Stage**

The probability movement of each glowworm 'i to j' is given below,

$$\varphi_j(S) = l_f(S) - l_i(S) \,/\, \sum_{k=M_i(S)} l_k(S) - l_i(S) \qquad (3.2)$$

where,

φ_j is the probability of movement to j

l_f is the attraction factor

l_i is the luciferin value

k is any neighbor of i within its neighborhood size

$$M_i(S) = \begin{Bmatrix} j: d_{ij} < r_d^i(S); \\ l_i(S) < l_j(S) \end{Bmatrix} \qquad (3.3)$$

where,

S denotes the step index and

d_{ij} represents the Euclidean distance (Rama Mercy Sam Sigamani et al., 2020).

The ratio of the combined attraction and luciferin "signal" received by 'i' from 'j' compared to the total "signal" received from all its neighbors is calculated. An elevated ratio is indicative of an increased probability of 'i' converging upon 'j'.\Discrete-time movements are given below,

$$G_i(S) = G_i(S) + \rho\,(G_{ij}(S) - G_i(S))/\left\|G_j(S) - G_i(S)\right\| \quad (3.4)$$

where,

G_i is the location of the glowworm at sample time S

|| || is the Euclidean norm operator

The GLW algorithm iteratively updates the positions of glow-worms based on their attraction and repulsion forces. This dynamic movement allows the algorithm to explore diverse regions of the feature space, potentially identifying features that are initially overlooked. The collaborative environment where vehicles share information about the features found relevant with the ability to learn from each other and collectively improve the feature selection process. The movement stage enables dynamic adaptation to constantly changing vehicle movement, weather conditions, and network events. The random displacement factor in this stage with the exploration of diverse areas of the feature space enables the glow-worms escape from local optima. This makes the algorithm less susceptible to noise and outliers in the data, leading to more robust and reliable feature selection.

Vehicles periodically share data about the features they have observed (e.g., average speed on their lane) and their locally calculated fitness scores based on their observations. This mimics how glowworms "sense" their surroundings. The movement mechanism is an iterative process in which vehicles within communication range exchange information about the features they possess and their calculated fitness scores based on the local data. Each vehicle updates the importance of its features (light intensity) based on the information received from its neighbors.

Features with low or inconsistent fitness scores from neighbors will see their importance decrease. This indicates these features might not be very helpful for detection of DoS attacks.

3.3.2.3 (c) Glowworm Neighborhood Stage

Normally, local optimum values are detected in the previous stages. The movement decision of the glowworms is governed by the local information. The radial sensor range determines the number of peaks through its function. Each agent possesses a

sensor range encompassing the entire workspace. All agents converge upon the global optimum, with local optima being disregarded.

In this phase, the local optimum values are ignored and the global optimum values are selected from the given workspace. It detects the multiple sensor locations and continues the attraction process in the neighborhood range. Multiple peaks are identified by varying the sensor range. For this purpose, glowworm agent 'i' is associated with a local-decision domain whose radial range is dynamic in nature bounded by a circular sensor range r_s and variable neighborhood ranger_d, $0 < r_d^i \leq r_s^i$. The enhancement movement is obtained using equation (3.5) (Rama Mercy Sam Sigamani et al., 2020).

$$r_d^i(S+1) = \min\{r_s, \max\{0, r_d^i(S) + \beta(S_t - |M_i(S)|)\}\} \qquad (3.5)$$

where,

β is a constant parameter

S_t is the parameter to control the number of neighbors

The Glowworm Neighborhood Stage is shown in Figure 3.5. The neighborhood stage plays a crucial role in feature selection for VANET by defining the scope of interaction between glow-worms. The neighborhood limits the interaction between glow-worms to those within a specific distance and identifies features relevant to the local traffic. The collaborative approach leverages the knowledge of multiple vehicles and quickly identifies emerging attack patterns. Limiting interaction to a local neighborhood reduces the computational complexity of the algorithm avoiding unnecessary communication overhead. The size and shape of the neighborhood can be dynamically adjusted to diverse DoS attack scenarios, including targeted attacks, flooding attacks, and denial-of-service attacks.

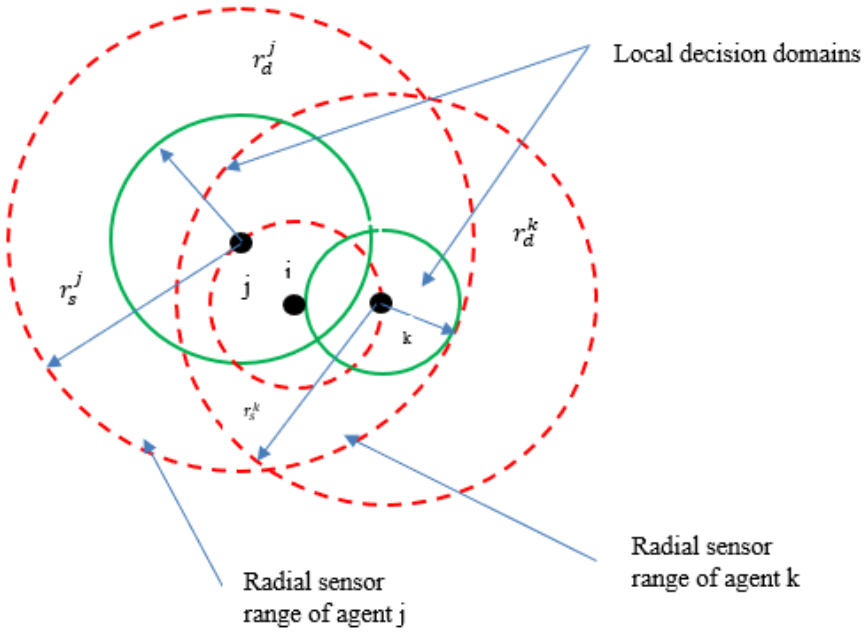

Figure 3.5: Glowworm Neighborhood Stage

The neighborhood stage encourages exploration of diverse features utilizing the information already acquired. The algorithm enabled localized attack detection, collaborative filtering, contextual awareness, scalability, and adaptability. The algorithm effectively utilized to improve DoS attack detection accuracy and response time, leading to a more secure and reliable VANET environment.

3.3.2.4 (d) Single Layer Feed Forward Network (SLFN) Classifier

The single layer feed forward network uses the principle of Extreme learning machines. This network leverages a single hidden layer to achieve high speed and accuracy in preparing velocity, leading to exceptional prediction/precision and the ability to approximate any function. SLFNs have a simpler architecture with only one hidden layer. This significantly reduces the training time and computational complexity, making them suitable for resource-constrained environments like VANET. Tuning the number of neurons in the hidden layer isn't strictly mandatory in SLFNs. The hidden layer neurons in SLFNs typically require a highly differentiable activation function, such as the sigmoid function. Highly differentiable functions allow for smoother and more efficient gradient updates, leading to better training results.

For a single-hidden layer SLFN, the system yield is given by equation (3.6).

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i(x) = h(x)\beta \quad (3.6)$$

where,

x→ input

β →output weight vector and it is given as follows as

$$\beta = [\beta_1, \beta_2, \ldots\ldots\ldots\ldots \beta_L]^T \quad (3.7)$$

H(x)→output hidden layer which is given by the following equation

$$h(x) = [h_1(x), h_2(x), \ldots\ldots\ldots\ldots h_L(x) \quad (3.8)$$

To determine output vector O the target vector, the hidden layers are represented by equation (3.9).

$$H = \begin{bmatrix} h(x_1) \\ h(x_2) \\ \vdots \\ h(x_N) \end{bmatrix} \quad (3.9)$$

The basic implementation of the SLFN uses the minimal non -linear least square methods which are represented in equation (3.10)

$$\beta' = H^* O = H^T (HH^T)^{-1} O \quad (3.10)$$

where,

H^*→ inverse of H known as Moore−Penrose generalized inverse.

Equation (3.10) can also be given as follows

$$\beta' = H^T (\frac{1}{C} HH^T)^{-1} O \quad (3.11)$$

Hence the output function can be found by using the equation (3.11)

$$f_L(x) = h(x)\beta = h(x)H^T (\frac{1}{C} HH^T)^{-1} O \quad (3.12)$$

SLFN provides enhanced functionality with good accuracy through the kernel function. Reduced errors in training and improved approximation are the merits of the SLFN. Classification and prediction are the functionalities of SLFN with its auto-tuned weights and activation function being non-zero.

The model has been executed with the input data for the 100 trials with 1500 collected traffic data. To evaluate the performance of the proposed SFLN classifier in detecting various attacks in transportation vehicular networks, the average accuracy across multiple trials was calculated. The sigmoid function was utilized as the activation function within the model's architecture. A detailed description of the SFLN classifier's algorithm can be found in Table 3.1. The fitness function for calculating the best features is given as follows

$$\text{Fitness\_Function} = W_B + \{i=1 \sum^{n} G_i \text{ X } F_i\}^{-1} \quad (3.13)$$

where,

W_B is the actual number of features used

$G_{i=}$ Cost function

F_i= No of feature to be selected**.**

The features pertaining to the malicious behavior of the vehicles are considered as the inputs to the Glowworm algorithm. The features selected as optimized have the values that provide the signatures of the malicious behavior due to their variations. The five acquired features were employed as input for the SLFN to perform attack classification and prediction. Labeled input data, consisting of seven features, is fed into the SLFN classifier. This classifier has a three-layer architecture designed specifically to differentiate between DoS and Botnet attacks. The proposed classifier classifies the attacks based on the above equation (3.13). This fitness function represents the potential feature subset with higher fitness values. The objective function formulated focus on feature count, and feature importance.

The GOF-SLFN algorithm, predicated upon the Glowworm Optimizer and integrated with a Single Layer Feedforward Network comprising 75 neurons, effectively detected DoS and Botnet attacks through the utilization of sigmoid learning. The proposed model exhibited superior performance compared to existing machine learning algorithms, achieving notable enhancements in detection accuracy, sensitivity, and selectivity.

The SLFN classifier with the optimized features selected by GOF detected two attacks DoS and Botnet. A comparative analysis of the classifier's performance was

conducted against existing classifiers, such as Single Layer Feedforward Network, Multi-Layer Perceptron, Random Forest, and ANN. The performance metrics accuracy, sensitivity and specificity with training and testing accuracies are considered for the comparison. The classifier exhibits enhanced intelligence in the detection of attacks.

The pseudo code of the proposed classifier is shown in Table 3.2.

Table 3.2: Pseudocode of SLFN Classifier

%Initialization

Start the process

Let 'N' be the training Glow worm Features (Input data) with an Activation Function and n Hidden neurons

Assigning the random values for Input weights are assigned and biases

Calculate the hidden Matrix H

Calculate the Output weight Matrix $'\beta'$

If $(\beta \leq= \beta 1)$

Then /Attack_1 is Classified (DoS Attacks)

Else if $(\beta \leq= \beta 2)$

Then /Attack_2 is Classified (Botnet Attacks)

Else if $(\beta \leq= \beta n)$

Then Attack_n is classified, End

The improvement in training shows the difference of 0.065 seconds difference obtained using the proposed when compared with the existing classifiers and the accuracy shows 8% increase for the proposed from the existing optimization algorithms and presented in Chapter 4. The minimal training time and impressive accuracy due to their single-hidden layer, SLFNs leverage auto-tuned weights and non-zero activation functions, making them a powerful tool for classification and prediction tasks in various domains.

3.3.3 Contribution 1 Merits

The contribution 1 applied as the optimized feature selection for malicious nodes detection and mitigation against DoS attacks using Glow-worm SLFN in the phase 1 has been illustrated with the steps involved in the algorithm. The optimized feature selection with classification with the consideration of two types of DoS attacks and launch of DoS attacks dealt with the VANET architecture with high density roads. Two scenarios were evaluated for the detection of DoS and Botnet attacks. The movement of vehicles, considered as nodes, was modeled using the Glowworm Optimization technique, which incorporates the locations of multiple sensors on vehicles and sustains the attraction process. The computational time for the identification of the optimal threshold by the search method remains as a challenge.

3.3.4 Contribution 1 Limitations

The approach in the contribution 1, Swarm Intelligence Optimization Technique GOF - SLFN, detected two class attacks namely DoS attacks and Botnet. The DoS attacks are possible with its variants in the VANET having the impact on its operations with unstable behavior. In the consideration of DoS and DDoS attacks, the selection of features based on the diverse characteristics of these attacks is essential for their effective detection and subsequent mitigation.

Given the dynamic nature of VANET, adaptive feature selection algorithms can be implemented to continuously adjust the selected feature subset in response to the evolving network conditions. Enhanced feature selection algorithms that use adaptive feature selection may be able to maintain accurate DDoS attack detection even in highly dynamic VANET. The replication and exploitation based DDoS attack include 12 types namely: MSSQL, SSDP, SYN flood and UDP based attacks.

Therefore, the contribution 2 is proposed to detect and mitigate the DoS attack types.

3.3.5 Contribution 2: Abnormal Behaviour Detection using Response Feedback Algorithm with Micro Cluster Outlier Detection Algorithm using Linear Regression (MCOD-LR)

The enhanced architecture to deal with multiclass attacks of DoS identified as the limitations from the contribution 1, the VANET architecture with NS3 addressing the

simulation area with increased vehicles and speed considerations. The detection and mitigation of multiclass attacks of DoS are considered with the proposed contributions 2, 3 and 4. The enhanced feature selection algorithm with adaptive selection is able to maintain accurate DoS and DDoS attack detection even in highly dynamic VANET [8].

Clustering optimizes overall network performance with internode distance and communication link capacity. Nodes in the cluster network can function as either cluster members (CMs) or cluster heads (CHs). The grouping helps to reduce communication overhead and enhances communication efficiency in the network. Clustering is done based on various factors, such as the position of the vehicles, speed, and direction. The dynamic movement of nodes within the network significantly increases communication latency between the server and the cluster head. This increased latency not only makes the system more vulnerable to security attacks but also contributes to a higher rate of packet loss within the Roadside Unit (RSU) cluster.

In this subsection, the enhanced feature selection proposed with the adaptive nature for the learning of maliciousness towards DoS attacks and the mitigation steps is explained.

The enhanced feature selection phase, specifically designed for detecting and isolating DoS attacks and their variations, is derived from the analysis of a VANET environment simulated using the NS3 platform. The Canadian Institute for Cybersecurity (CICIDS2019) dataset has been considered to include a variety of current reflective DDoS assaults namely, PortMap, NetBIOS, LDAP, MSSQL, UDP, UDP-Lag, SYN, NTP, DNS, and SNMP. In this setup, the parameters are analyzed when the VANET has been established.

In the simulated VANET of the proposed framework, vehicles as nodes form clusters and communicate in the network. Cluster communication is considered and the features triggering the malicious behavior in the cluster provide a better detection of maliciousness in the clusters. On detection of malicious clusters and the malicious nodes, mitigation strategies are applied in this phase. Isolation and prevention methods are applied to secure the clusters in the initial phase with the identification of features

revealing the maliciousness thus promoting the learning of maliciousness in the network with detection of attacks.

The objective of the contribution 2 is to detect and predict the DoS attack by analyzing the temporal information with variable speed range. The network overloading is also detected through the launch of the DoS attacks. In a VANET environment, RSUs interact with cluster members through cluster heads, which exhibit frequent changes due to the dynamic nature of vehicle movement. Attackers can exploit this dynamic by launching attacks that overload the network, resulting in packet failures within RSU-cluster communication. The identification of such malicious behavior within the VANET is achieved through the analysis of transmission response times obtained from temporal-based data monitoring. The proposed approach introduced a novel Response Feedback Algorithm with Micro Cluster Outlier Detection Algorithm using Linear Regression as the contribution 2.

The Response Feedback Algorithm leverages temporal information, considering variable vehicle speeds and accounting for factors such as data transmission delays, response times between the Roadside Unit (RSU), deviations from expected packet transmissions, packet losses, and the relative speed and position of vehicles. This approach focuses on variations in data transmission, response times, network traffic, and entropy values. The micro-cluster outlier detection method is employed to identify and detect potential attacks based on these variations.

The Micro Cluster Outlier Detection Algorithm with Linear Regression in VANET using temporal information for DoS Attack Detection (MCOD-LR) is a novel algorithm that uses a combination of micro cluster outlier detection and linear regression to detect the abnormal behavior based on DoS attacks in VANET. The MCOD-LR algorithm works based on the temporal information to identify anomalous traffic patterns that may be indicative of DoS attacks. Attacks are identified by comparing the observed temporal information with the derived forecasted transmission response time during data communication, and analyzing any deviations from the expected behavior.

The message rate determines the data transmission happening during the communications in VANET. When the redundant stream of packets is sent, changes in

the network traffic prove the successful DoS attack launch in VANET. The message rate is the threshold measure used to identify the anomalies in the new incoming messages. The network traffic traces with both normal and malicious traffic are used and DoS attacks are analyzed based on message rates. The current time for the incoming message with the previous time of messages is considered and the time elapsed is calculated. With the threshold based on the message rate, maximum limit based on the threshold and the elapsed time, the micro cluster outlier detection identifies the deviation in such message arrivals and detects the attack with linear regression.

The algorithm has the input as the new message and the output is the detection of the abnormal behavior. In the algorithm, 't' is the threshold and the counter is 'a'. The value of 't' is considered 0.5s as the new message's rate, where the current is 0.0025ms and the previous time value is 0.5ms.

The micro cluster outlier detection algorithm is given in Table 3.3.

Table 3.3: Micro Cluster Outlier Detection Algorithm using Linear Regression

Input: *New message (R)*

Output: *Detect abnormal behavior*

Start

Increase the counter by 1: a++ // based on parameters

If a% t = 0 then // t = threshold

E = current time – the previous time

s = t / E

else send the message to the RSU

end if

s input to MCOD

if s is normal, then notify the node

otherwise, find out whether the abnormality is because of attacks (by

using linear regression)

Steps in Micro Cluster Outlier Detection Algorithm using Linear Regression

The steps in the detection of DoS attacks using micro cluster outlier detection algorithm using linear regression in VANET are discussed below:

The steps at first consider the message arrival and the time elapsed leading to the message rates are calculated. This calculated measure is used for the detection of attacks with linear regression.

Step 1: Upon the reception of a new message (R), increment the message counter (a) by one.

Step 2: Calculate the maximum message limit using the modulo operation on the counter (a) and the threshold (t).

Step 3: Deliver the new message to the RSU if the result of the modular division is non-zero. Otherwise, record the current timestamp.

Step 4: Calculate the elapsed time (e) by subtracting the previously recorded timestamp from the current timestamp.

Step 5: Determine the new message rate (s) by dividing the threshold (t) by the elapsed time (e).

The RSU cluster network utilizes a hybrid approach of linear regression and micro-cluster outlier detection for attack identification. The linear regression model leverages the number of new messages and the counter value as input features to detect attacks. The micro-cluster outlier detection algorithm identifies instances where the number of new messages deviates significantly from expected patterns, triggering an increment of the counter. Temporal information serves as the basis for incrementing the counter in a typical network scenario.

The linear regression model, utilizing the number of new messages and counter values, is employed for attack identification. The mean absolute error associated with this regression line is calculated, with higher values indicating the occurrence of DoS attacks

within the VANET. This metric utilizes the analytical measures from the VANET and improves the performance in detecting the attacks with its low value.

The formula used for the mean absolute error is given below:

$$M = \frac{1}{n}\sum_{i=1}^{n}(|Xi - \acute{X}\imath|) \quad (3.14)$$

where,

Xi – the number of counters

$\acute{X}\imath$ – estimation of this value

Chapter 2 examines existing techniques for network anomaly detection, which, despite demonstrating high accuracy, exhibit a significant rate of false positives. The frequent reception of such false alarms presents a major drawback for anomaly detection systems for the following reasons:

- Operators waste time investigating alarms that do not indicate actual problems.
- False alarms can lead to unnecessary service interruptions as operators attempt to resolve non-existent issues.
- Investigating false alarms consumes valuable time and resources.

A key innovation in contribution 2 involved the use of linear regression analysis on incoming messages within VANET clusters to investigate the underlying causes of false positives.

The proposed contribution 2, Micro-cluster outlier Detection algorithm using Linear Regression in phase 1 detects the maliciousness behavior of RSU cluster improves the attack detection and thereby improving the latency. Following the identification of abnormal behavior, phase 2 aims to identify the malicious nature of attacks originating from network nodes. The malicious behavior in VANET due to DoS attacks is detected and they are to be classified and handled based on the impact as the escalation of the maliciousness leading to the unavailability of the services. The next subsection details with the contribution 3 based on trust-based approach with the classifier for detection and classification of DoS attacks.

3.3.6 Contribution 3: Novel Adaptive Nodal Detection Algorithm for the Prediction of Malicious DoS Attacks using Kernel Density Estimation and Entropy-based Support Vector Machine (SVM) Classifier

3.3.6.1 Introduction

The objective of the proposed approach is to predict the maliciousness of the nodes based on DoS attacks. The novel adaptive nodal detection algorithm utilizes the trust-based approach of the vehicle nodes and the recommendations of the neighbor vehicle nodes in the VANET. The vehicles while acting as senders in the data transmission considers their neighbor vehicles based on distance, speed and direction. During transmissions, to avoid the malicious vehicles, trust values updated in the vehicles are considered. The vehicles in the network share their observations to their neighbors by the push-based notification. The proposed approach in phase 1 utilizes the findings of contribution 3, specifically employing an entropy-based Support Vector Machine (SVM) classifier enhanced by Kernel Density Estimation (KDE).

3.3.6.2 Kernel Density Estimation with Entropy based SVM

By utilizing kernel density estimation based on the probability density function and incorporating trustiness values for the parameters, the network is assessed for malicious attacks.

The trust value for the following parameters is considered.

i. Vehicle density
ii. Average latency
iii. Packet delivery ratio
iv. Detection rate
v. Energy consumption

The proposed approach evaluates the trustworthiness of the RSU cluster network through the comparison of specific parameter values against established thresholds. The vehicle nodes affected are identified by the calculated trustiness values exceeding the threshold values.

The trustiness factors of the parameters in the VANET are given below:

i. Vehicle Density (VD)

The proposed approach employed a vehicle density metric to assess the RSU or cluster network's trustworthiness value. The nodes are designated as assault nodes if the vehicle density value exceeds the threshold value for vehicle density. The density of vehicles computed using the formula in equation (3.15).

$$Vehicle\ Density\ (VD) = n\ \mathrm{X} \frac{1000}{l} \qquad (3.15)$$

In the above equation, VD represents density of the vehicle (per km), 'n' – vehicle count on the road and l – road length covered by vehicles (m).

ii. Packet Delivery Ratio (PDR)

The trustworthiness of an RSU or the cluster network is determined by its packet delivery ratio. Nodes with a packet delivery ratio that surpasses a predefined threshold are flagged as potential attack nodes. The calculation of this packet delivery ratio is detailed in equation (3.16).

$$Packet\ Delivery\ Ratio\ (PDR) = \frac{Total\ no.of\ received\ packets}{Total\ no.of\ send\ packets} \qquad (3.16)$$

iii. Attack Detection Rate (ADR)

The proposed approach assesses the attack detection rate of the RSU or the cluster network. The formula for calculating this attack detection rate is presented in equation (3.17).

$$Attack\ Detection\ Rate\ (ADR) = 100\%\ X\ \left(\frac{Total\ no.of\ packets}{Total\ no.of\ detected\ attacks}\right) \qquad (3.17)$$

From the above formula, if the calculated value of the detection rate exceeds the threshold, the trust value is subject to a unitary increment. Conversely, if the token value is invalidated, the trust value is decremented by one.

iv. Average Latency (AL)

The average latency of the packets transaction between node 'n' to node 'm' in the RSU or the cluster network is considered to detect the attacks. If the average latency

exceeds the threshold, the proposed approach determines the node is an affected node based on the trustiness value. The average latency is calculated using the formula given in equation (3.18).

$$Average\ Latency\ (AL) = \frac{N}{T} \tag{3.18}$$

where,

N is the average number of packets in the network

T is the total amount of traffic

Energy Consumption (EC)

Energy consumed during communication between vehicle nodes, the RSU, and the cluster network is considered, with traffic data serving as a key input parameter. If the energy consumption value exceeds the threshold, the attacks are detected.

The classifier Entropy-based SVM classifies the maliciousness type as it predicts the changes in the trust factor value and the kernel density estimation. The flowchart exhibiting the working of the Algorithm 3.2 is shown in Figure 3.6.

The algorithm for the adaptive nodal attack detection is shown in the Table 3.4.

The given input is the parameters(x) and the output is maliciousness of attacks. The node's trust value is calculated by evaluating its past interactions and the recommendations it receives from its neighbors. This trust value is typically represented as a score between 0 and 1.

The range of trust values involves the trust values often normalized to a specific range, such as [0, 1] or [-1, 1], where 0 or -1 represents complete distrust and 1 represents complete trust.

Higher values indicate higher trust. Starting with the conservative value of the threshold as 0.5 and minimizing the false positives. Hence, the threshold value in the cluster network is 0.5 and the kernel density estimation value is 0.3.

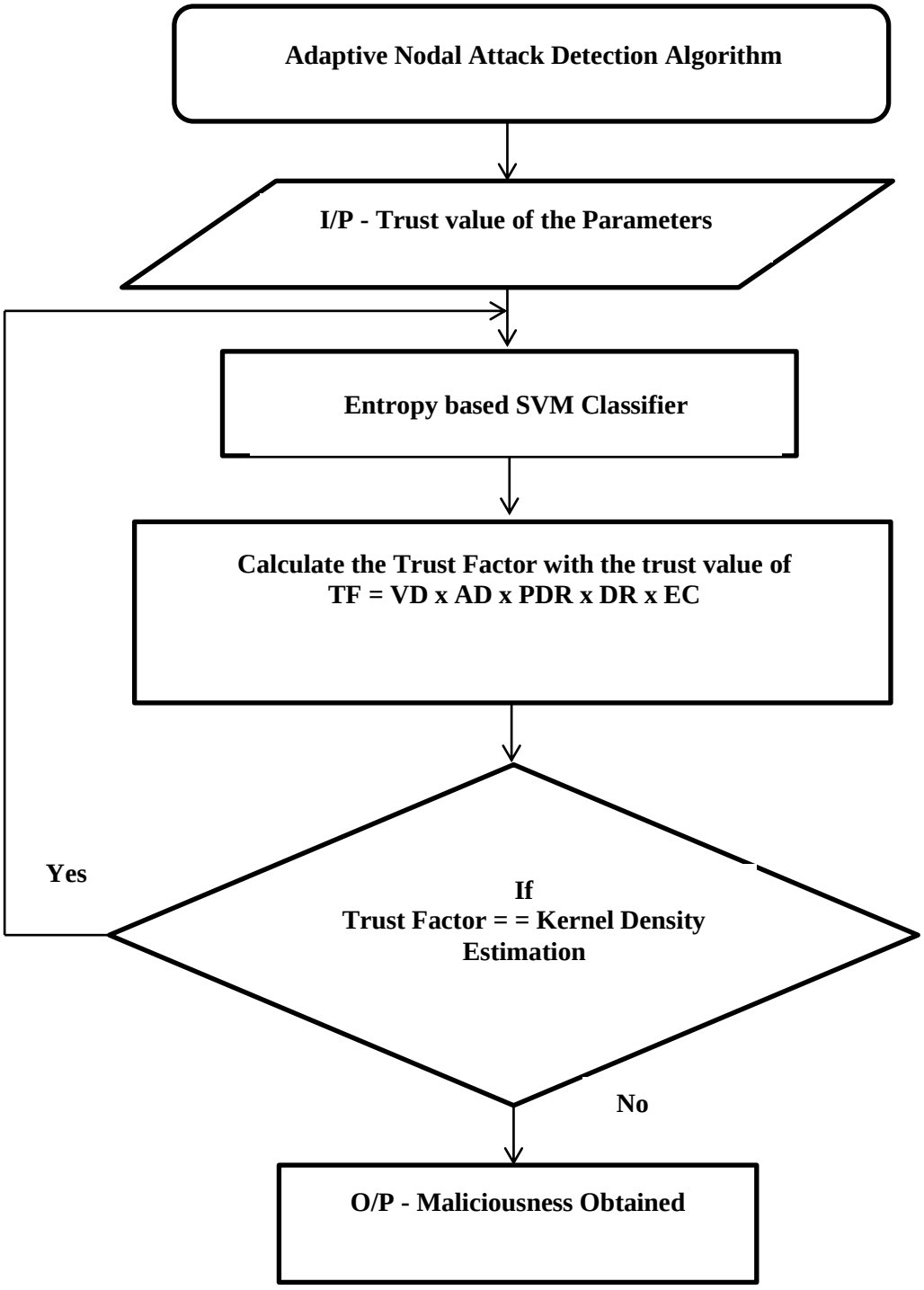

Figure 3.6: Flowchart of the Adaptive Nodal Attack Detection using Kernel Density Estimation with Entropy based SVM

Table 3.4: Adaptive Nodal Attack Detection Algorithm using Kernel Density Estimation with Entropy based SVM

Input: *parameters (x)*

Output: *maliciousness of attacks*

Start

If (the parameters (x) is equal to the threshold value in the cluster network)

// create a token for incrementing the trust value by one

Token j = j + 1

Else (the parameter (x) is not equal to the threshold value in the cluster network)

// create a token for decrementing the trust value by one

Token j = j – 1

Find trust factor $TF = VD \times AL \times PDR \times DR \times EC$

If (trust factor less than Kernel density estimation)

// find out how much the node is affected by the attack

SVM classify and predict the maliciousness of the attacks

End

The calculated trust values based on the network size are shown in the Table 3.5.

Table 3.5: Trust Value of the Nodes

Network Size	Trust Value
20	0.4
40	0.5
60	0.5
80	0.533333333
100	0.533333333
120	0.6
140	0.653333333
160	0.673333333
180	0.693333333
200	0.713333333

A description of the key steps involved in an adaptive nodal attack detection approach that incorporates Kernel Density Estimation (KDE) for feature extraction and Entropy-based Support Vector Machines (SVM) for classification is presented below.

Step 1: *Initialize the parameters: VD (Vehicle Density), AL (Average Latency), PDR (Packet Delivery Ratio), DR (Detection Rate), and EC (Energy Consumption).*

Step 2: *The Entropy-based SVM classifier assesses the trustiness of each node by considering the trust values associated with various parameters. Entropy, a key component of this classifier, is employed to quantify the randomness or unpredictability in the distribution of certain attributes within the headers of network packets.*

Then trust factor is found using the formula below

$$TF = VD \times AL \times PDR \times DR \times EC \qquad (3.19)$$

Step 3: *When a discrepancy is observed between the actual trust factor and the value predicted by Kernel Density Estimation (KDE), an entropy-based SVM classifier is employed to assess the node's maliciousness within the cluster network.*

The detection of malicious behavior within the communication framework between the RSU cluster network and vehicle nodes necessitates continuous evaluation of cluster behavior. Entropy is used to measure the randomness in the network packet header fields. The deviation in some packet header fields represents randomness compared to the other packet header fields. This detects the maliciousness in the nodes due to randomness. During DoS attacks within a VANET, the entropy associated with source and destination IP addresses tends to decrease, while the entropy of source and destination ports tends to increase. The detection of malicious nodes incorporates the entropy values associated with both packet type and packet size.

The entropy converges to a small value which is worth considering for the maliciousness node. The packet sizes of the similar sized packets also have the entropy value exceeding the value of the number of the packets. The packet distribution with the entropy decreased during the attack. The proposed approach integrates these fields into its Entropy-based SVM framework. By considering the deviation in randomness observed

within packet header fields and incorporating the trust factor derived from Kernel Density Estimation (KDE), the classifier effectively predicts the maliciousness of nodes.

The malicious node must be handled immediately as the escalation of the maliciousness in VANET causes catastrophic effects in the roads as the services becomes unavailable due to the attacks. Escalation of such maliciousness behavior due to DoS attack types are handled by prevention and isolation. Accordingly, phase 1 employs a novel reliance node estimation approach, as detailed in contribution 4, to facilitate the isolation of malicious nodes.

3.3.7 Contribution 4: Isolation using Reliance Node Estimation Approach using Pearson correlation coefficient and Bayesian aggregate model with Self-healing effect of Artificial Immune System (AIS)

3.3.7.1 Introduction

The objective of this approach is to isolate the predicted DoS attacks in the VANET. In addition to the detection of malicious behavior within the cluster network and at the nodal level, this approach incorporates an analysis of the factors that contribute to the vulnerability of different DoS attack types. Securing the VANET services with its availability while such assaults occur is the main aim at this stage. The diverse functions in the quality of the RSU cluster network with variations in the similarity between the nodes, varied features and functionalities of various manufacturers of the vehicles are considered in this approach.

3.3.7.2 Pearson correlation coefficient Method

The statistical method applied is the Pearson correlation coefficient for the linear association between two variables. The coefficient exhibits values between -1 and 1, 0 with negative correlation, positive correlation and no correlation respectively (Alshahrani, B. S et al., 2024).

The Pearson coefficient correlation was calculated by using the equation 3.20.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (3.20)$$

The degree of similarity between predicted data and actual VANET performance, considering factors like vehicle density, packet delivery ratio (PDR), detection rate, and energy consumption, was assessed using the Pearson correlation coefficient. This analysis compared predicted values to actual values for each of these variables. A high correlation coefficient (close to 1) for a particular variable indicates that the predicted values for that variable are very similar to the actual values. The Pearson correlation coefficient offers an effective means of quantifying both the strength and direction of the linear relationship that exists between any two variables within a VANET, including:

i. Vehicle density
ii. Energy consumption
iii. Packet delivery ratio (PDR)
iv. Detection rate

By analyzing these variables for closely positioned nodes, researchers can gain valuable insights into the interdependencies and correlations that exist within the dynamic VANET environment. The Pearson correlation coefficient for close nodes in VANET based on vehicle density, energy consumption, PDR, and detection rate can be used to understand the relationships between these variables.

A high Pearson correlation coefficient between vehicle density and energy consumption would indicate that vehicles in high-density areas consume more energy. A high Pearson correlation coefficient between PDR and detection rate would indicate that vehicles with a higher PDR are more likely to detect other vehicles.

The two steps considered in this method are discussed below.

Step 1: *Calculate the Pearson correlation coefficient between the predicted and actual values for each of the following variables:*

i. Vehicle density
ii. PDR
iii. Detection rate
iv. Energy consumption

Step 2: *Interpret the correlation coefficients:*

i. *A correlation coefficient with the value nearing to 1 indicates a strong positive correlation, suggesting a high degree of agreement between the predicted and actual values for that variable.*
ii. *A correlation coefficient near -1 signifies a strong inverse relationship between two variables. This implies that as one variable increases, the other consistently decreases. In the context of predictions, a strong negative correlation suggests that the model's predicted values for that variable deviate significantly from the actual observed values.*
iii. *A correlation coefficient nearing to 0 indicates no linear relationship between the variables, suggesting that the predicted values for that variable are not significantly associated with the actual values.*

The Pearson correlation coefficient measures the linear relationship between predicted data, specifically focusing on the association between low trustiness values observed across various node functions.

This coefficient ranges from -1 to 1, where:

-1: Indicates a strong inverse relationship in trust levels between nodes.
1: Indicates a strong direct relationship in trust levels between nodes.
0: Indicates no significant linear relationship between the trust levels of nodes.

The Pearson correlation coefficient quantifies the strength of the linear relationship between nodes. The quantifying value ranges -1 to 1 and tabulated in Table 3.6.

Table 3.6: The scale of Pearson's Correlation Coefficient

The scale of the correlation coefficient	Value
$0 < r \leq 0.19$	Very Low Correlation
$0.2 \leq r \leq 0.39$	Low Correlation
$0.4 \leq r \leq 0.59$	Moderate Correlation
$0.6 \leq r \leq 0.79$	High Correlation
$0.8 \leq r \leq 1.0$	Very High Correlation

In the VANET, the vehicular communications happen as the vehicle node moves from one access point to another and the network rearranging its connecting pattern. The quality of the RSU cluster communication with nodes availability and stability are focused due to the nodes density and the mobility using Pearson correlation method.

3.3.7.3 Bayesian Aggregate Model with the Self-healing Effect

In addition, onboard unit attached to the vehicle providing mobility with vehicular communication in appropriate direction is heterogeneous in nature based on the manufacturers. The wide range of features and functionalities found in vehicle nodes from different manufacturers makes them susceptible to attacks. This vulnerability arises from the continuous transmission of information, such as Basic Safety Messages (BSMs), between On-Board Units (OBUs), other vehicles, Roadside Units (RSUs), and other devices.

In phase 1, the proposed approach addresses the impact of diverse node behavior on detecting misbehavior. It assesses node credibility by employing a Bayesian aggregation model that leverages Pearson correlation coefficients. Figure 3.7 provides a visual representation of the RSU and OBU interactions.

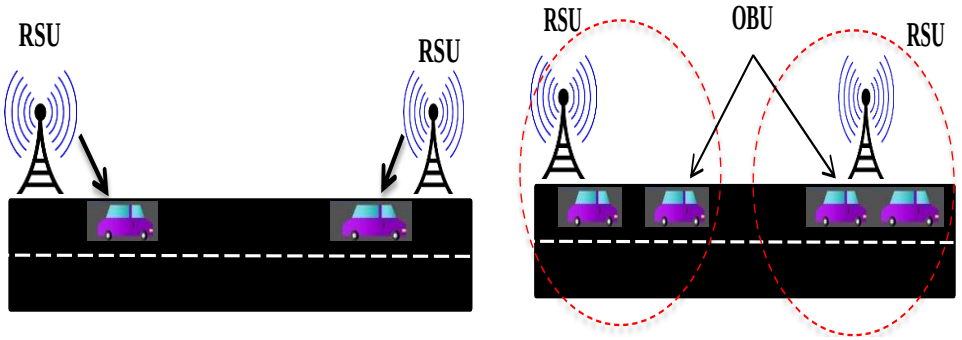

Figure 3.7: RSU and OBU

An onboard unit is a hardware device mounted on the vehicle; therefore, it also communicates with other OBUs and RSUs. The trustworthiness of other vehicles within the network is evaluated using a Bayesian aggregate model. This is crucial because even a reliable node can be compromised and inadvertently used to transmit malicious messages. Upon receiving event reports, onboard units in receiving vehicles evaluate the credibility

of both the data content and the source node. This evaluation process results in the generation of ratings for source vehicles (message senders). These ratings are then used to update the trust values of the source vehicles, contributing to an overall trust score. The proposed system evaluates vehicle credibility by considering both the source node's trust level and its security status. If the vehicle's ratings (credibility score) are deemed correct, a trust value of one is assigned. Otherwise, a trust value of zero is assigned.

The Bayesian aggregate model facilitates the continuous monitoring of the RSU cluster communication network for vulnerabilities. If the model detects any vulnerability in a node, its credibility score is immediately set to zero. This triggers the self-healing mechanism within the AIS, resulting in the isolation of the compromised node from the network. The steps of the Self-healing effect of the AIS algorithm are listed below.

Step 1: *Initialization: The algorithm begins by establishing initial predicted values for the communication within the cluster network.*

Step 2: *Sensing Activity: Vehicles actively engage in sensing activities by broadcasting Basic Safety Messages (BSMs). BSMs transmitted by OBUs incorporate a set of parameters including vehicle density, energy consumption, average latency, detection rate, and packet delivery ratio. A weighted approach is employed, where each parameter is assigned a weight determined by its corresponding credit value.*

Step 3: *Performance Evaluation: The performance of each parameter within the network is assessed.*

Step 4: *Decision Making: Based on the performance evaluation, a decision is made. If the performance of any parameter falls below the established threshold, the corresponding node is isolated from the cluster network.*

Figure 3.8 illustrates the step-by-step process of phase 1, which incorporates the Artificial Immune System (AIS). This approach effectively identifies malicious nodes by accurately assessing their behavior. By isolating intrusions and predicting attacks, it ensures the reliable operation of the VANET while accommodating the diverse features and functionalities of different nodes.

Furthermore, the Bayesian aggregate model plays a crucial role in evaluating OBU credibility. By accurately identifying and isolating malicious nodes, the model contributes to a significant reduction in energy consumption within the framework of phase 1.

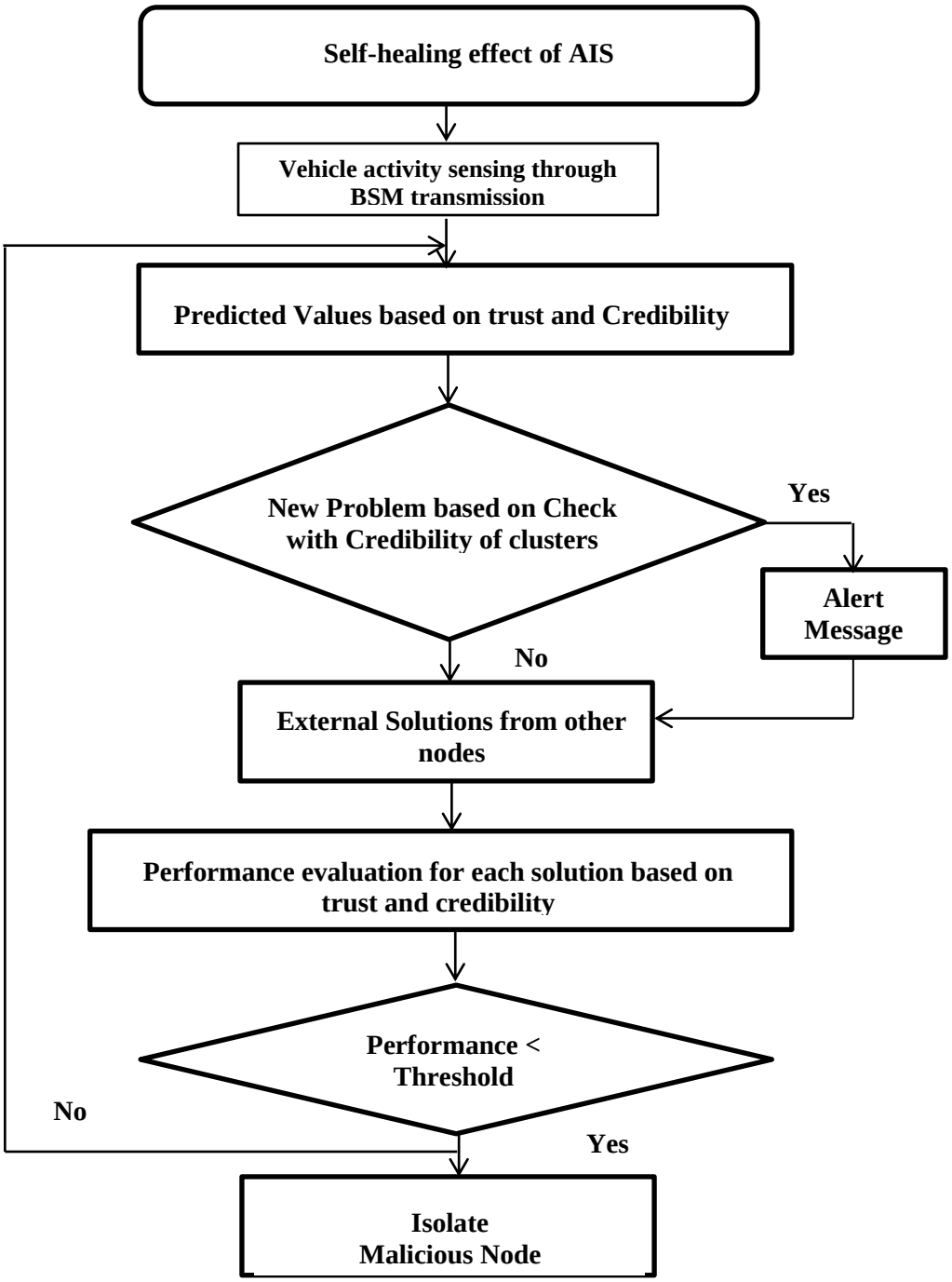

Figure 3.8 : Flow of the steps adapting Artificial Immune System

3.3.8 Phase 1 Merits

The proposed Response Feedback Algorithm, Adaptive Nodal Attack Detection, and Reliance Node Estimation methods were employed to evaluate VANET parameters and performance metrics. The results demonstrated that the first phase of the proposed system effectively addressed Denial-of-Service (DoS) attacks by successfully detecting and isolating malicious nodes. The RSU cluster communication network analyzes characteristics like vehicle density, energy consumption, average delay, packet delivery ratio, and detection rate to identify malicious nodes involved in DoS attacks. Subsequently, these nodes are isolated through the activation of the self-healing mechanism inherent to the AIS in the RSU cluster communication network.

3.3.9 Phase 1 Limitations

The proposed phase 1 handled the DoS attacks types based on detection and isolation steps. The self- healing effect of AIS predicted the DoS attacks in relation with the existing communication among the VANET. The dynamic topology inherent to VANET renders them susceptible to novel incoming attacks. These attacks are effectively isolated through the utilization of the self-healing mechanism inherent to the AIS. The AODV protocol in VANET provides the routes as and when needed and the links between vehicles connect and disconnect very often which makes routing process challenging due to the high mobility of nodes. Further enhancement is proposed in phase 2 focusing on the VANET services inspite of DoS attacks through routing in a secured way.

3.3.10 Phase 1 Summary

During the initial phase, suitable features were chosen for the identification and categorization of Denial-of-Service attacks. Phase 1 is considered as four contributions, in which, contribution 1, a GLW – SLFN technique selected the optimized features and detected the attacks. The two-class attack classification is performed based on DoS attack and Bot Net attack. GLW – SLFN provides better detection of DoS attack with classified results. VANET are vulnerable to DoS attack types disrupting the communications leading to unavailability of services. Contribution 2 proposed mitigated DoS attack types

on detection. The Response Feedback Algorithm analyzed network parameters relevant to DoS attacks. By monitoring transmission response times between the RSU and the network, the algorithm identified changes in transmission behavior.

The second contribution involved implementing micro-cluster outlier detection in conjunction with linear regression to identify anomalies and detect attacks within the VANET environment as new messages were received. To assess the maliciousness of classified attacks, the third contribution employed a novel adaptive nodal attack detection approach. This approach leveraged trustiness values and integrated entropy-based Support Vector Machines (SVM) with Kernel Density Estimation.

Contribution 4 focuses on isolating malicious nodes by introducing a novel reliance node estimation approach. This approach integrates self-healing mechanisms from the Artificial Immune System (AIS) with Pearson correlation analysis to assess the similarity between predicted and observed network behavior. Furthermore, a Bayesian aggregate model is employed to assess the trustworthiness of On-Board Units (OBUs), allowing the proposed system to effectively isolate malicious nodes.

The proposed system effectively detected, classified, and isolated attacks. Consequently, it demonstrated superior performance compared to conventional methods, achieving a 97% detection rate, a 39% reduction in energy consumption, and a 25% decrease in latency. The experimental setup and the measures are given in Chapter 4. The RSU cluster has the changing vehicle nodes as vehicles move in and out of range of each other for which the network have to be secured from unauthorized access and attacks. To secure the VANET in the changing topology and also to provide the operations inspite of attacks, phase 2 with contribution is proposed and explained in the following section.

3.4 Phase 2: Strengthening the Access Control and Mapping

3.4.1 Introduction

This section details the contribution as the combined approach with Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping using Deep Auto Sparse Impasse Neural Network. The combined approach is applied to enhance the routing process by providing the secured traffic through mapped vehicles as rational with

mapping report and detected the DoS attacks and DDoS attacks with the 12 types. The CIC-DDoS2019 dataset amounts to 22.6% records as normal traffic and 77.3% records as attack traffic with the total of 431,371 records. DDoS attacks frequently target TCP and UDP protocols at the application layer. Figure 3.9 illustrates two primary categories of DDoS attacks: reflection-based and exploitation-based.

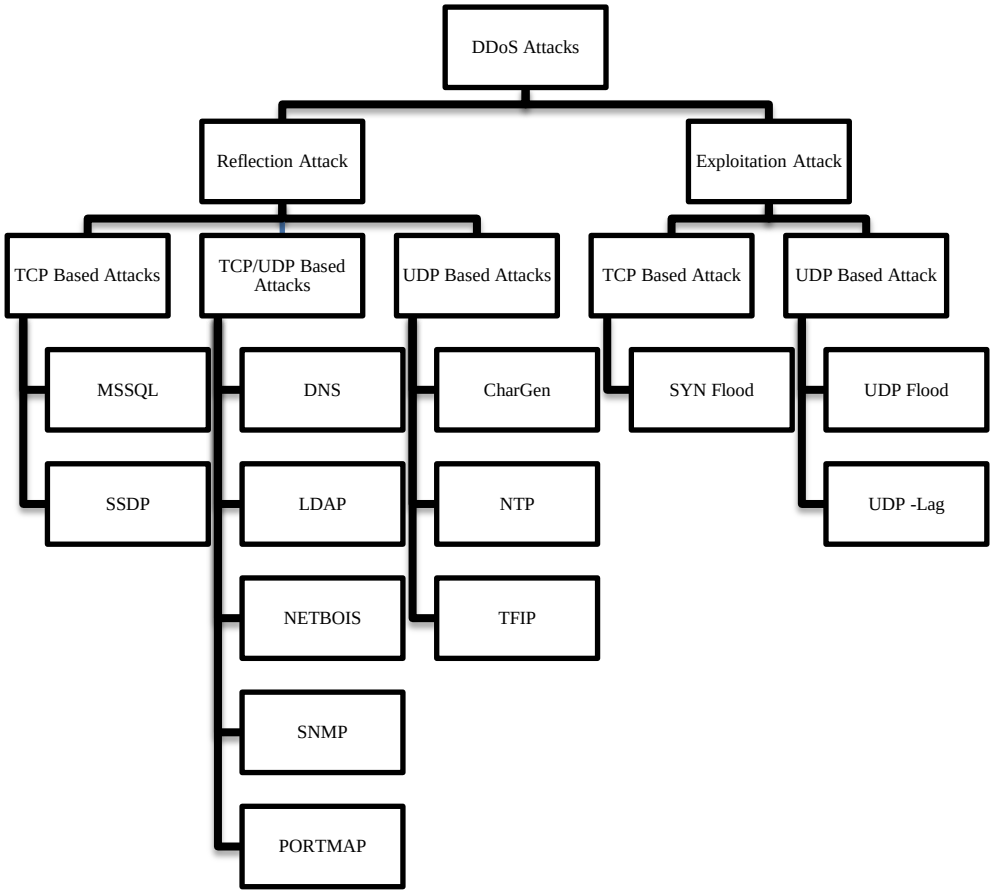

Figure 3.9: DDoS Attacks

TCP and UDP based attacks are included in both reflection-based and exploitation-based DDoS category. Handling of DDoS attacks is ensured in phase 2 with strengthening and mapping.

TRHE encrypts the traffic signal followed by mapping through hex-tuple matched mapping providing the mapping report and detects the twelve attacks by Deep auto sparse impasse Neural Network.

The multi-class attack detection problem is addressed through the Deep auto sparse impasse NN. This is aided by Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping to provide a mapping report for routing in the next phase. AODV protocol in VANET is generally not considered the best choice for networks with a high number of nodes based on the following factors:

- Broadcast Storms: AODV relies on route discovery broadcasts (RREQ messages) to find paths. In large networks, these broadcasts can create significant overhead, leading to further congestion and reduced network efficiency.
- Scalability Limitations: AODV's route discovery process may not scale well to very large networks with frequent node mobility and link changes. The number of broadcast messages can become overwhelming, consuming bandwidth and degrading performance.

In this step, the routing process is eased for the protocol by the Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping thus strengthening the access control.

The phase 2 with the proposed contribution and detailed implementation are discussed in the following subsection.

3.4.2 Contribution 5: Strengthening the Access Control and Detecting DDoS Attacks using Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping

Effective data transmission is crucial for the success of VANETs, facilitating seamless communication between vehicles and the surrounding infrastructure. Unfortunately, this vital data exchange is susceptible to a range of security threats. The vehicles must also be authorized in receiving the data transmitted. Authentication of vehicle nodes and access privileges tend to be vulnerable due to various attacks. However the need to encrypt data that is being transmitted between vehicles and authenticate the vehicles that are communicating with each other arises to route encrypted packets to the correct vehicles in VANET. Encryption and mapping are combined in this phase 2 to propose a Triple random hyperbolic encryption (TRHE) for encrypting the traffic with

data packets and hex-tuple mapping allowing vehicles to communicate securely and efficiently.

The TRHE is utilized to encrypt the data three times using a series of random hyperbolic transformations. The random hyperbolic matrices generated three times are used for encrypting the data traffic in the vehicular communication. In In Vehicular Ad-hoc Networks (VANETs), traffic encompasses the communication and data exchange that occurs both between vehicles themselves and between vehicles and roadside infrastructure. This data includes critical information such as vehicle location, speed, direction of travel, and prevailing traffic conditions, including accidents and other potential hazards.

The vehicle nodes encrypt the traffic using the recipient's public key. The encrypted traffic received by the recipient is decrypted using the private key. This process would ensure that all the traffic signals in the VANET network are encrypted, which would help to protect the data from being intercepted or decrypted by unauthorized users.

Triple random hyperbolic encryption is incorporated as the encryption method with the hyperbolic geometry to encrypt data and encrypting data is performed using a series of random hyperbolic transformations. In VANET, TRHE can be used to encrypt any type of data that is transmitted in VANET, including traffic data, vehicle location data, and sensor data. Each vehicle generates a key pair based on triple random hyperbolic encryption and the public key encrypts the traffic signals and private key decrypts the traffic signals.

The triple random hyperbolic encryption performs random encoding three times to encrypt traffic signals and determine the public key by plotting random values as coefficients in hyperbola to strengthen the access control in the middlebox. The random hyperbolic base point is chosen with the high order. The public key is computed with an integer within the higher order. The integer used to calculate the public key now serves as the private key, known only to the owner of the data transmission. The base point is meticulously chosen with a large order, rendering its computation infeasible without possession of the private key.

The implementation of TRHE algorithm is detailed as follows.

In hyperbolic encryption, to choose the coefficient the equation 3.21 as follows:

$$x2 - Dy2 = 1 \quad (3.21)$$

Equation 3.21 is the hyperbolic equation, and a base point $G = (x0, y0)$ with a large order $'r'$ is chosen. With $G^r = E$ given and an integer $'m'$ is selected with $m < r$ and $B = G^m \, mod \, q$ is computed. The private keys are formed by (G, B) . In the encryption process, the secret integer $'k'$ is chosen and 'H' and 'T' are calculated in the following equation 3.22 and 3.23:

$$H = G^k mod \, n \quad (3.22)$$

$$T = B^k w \, mod \, n \quad (3.23)$$

Equations 3.22 and 3.23 produce the cipher text (H, T), and in the decryption process, the 'R' is computed as the following equation 3.24:

$$R = H^m mod \, n \quad (3.24)$$

The following equation 3.25 recovers the data:

$$w = T/R \, mod \, n \quad (3.25)$$

The public key encrypting the traffic signals and private key decrypting the traffic signals are obtained. Encrypted traffic and vehicle mapping can be used together to improve the security by blocking external hosts and provide end-to-end network transparency. The proposed approach in phase 2 uses a process called hex tuple matched mapping to map all the network resources such as source IP, destination IP, source port number, destination port number, public key, and IP address range in an N to N mapping. The mapping structure is shown in Figure 3.10, and mapping the IP address in a tuple means the value is immutable.

Node:1 IP:192.168.0.1 Source IP: 192.168.0.3 Destination Source port:4356 Destination port:8975 IP:192.168.0.10 MAC:44:02:EF:9 C:09:4E Public key:C*f- JaNdRgUkXp2s5 v8y/ A?D(G+KbPeS	Node:2 IP:192.168.0.2 Source IP: 192.168.0.7 Destination IP:192.168.0.10 Source port:4595 Destination port:7885 MAC:A9:4E:47:AC :DA:B9 Public key:C*f- JaNdRgUkXp2s5v8 y/ A?D(G+KbPeS	Node:3 IP:192.168.0.3 Source IP: 192.168.0.1 Destination IP:192.168.0.8 Source port:4686 Destination port:8549 MAC:8C:9D:67:4C: 59 Public key:C*f- JaNdRgUkXp2s5v8 y/ A?D(G+KbPeS	Node:N IP:N Source IP: N Destination IP:N Source port:N Destination port:N MAC:N Public key:C*f- JaNdRgUkXp2s5 v8y/ A?D(G+KbPeS

Figure 3.10: The value stored in Hex Tuple Mapping

Once the scanning is initiated, hex-tuple matched mapping is used to map all the same IP addresses in a symmetrically matching hex-tuple value.

The misconfiguration discovery in the middle box will lead to performance degradation, making it vulnerable to attacks. In the proposed model, the middlebox controls the nodes and does N to 1 mapping IP address. N-1 mapping is a technique for mapping nodes in VANET that uses a combination of GPS receivers and V2V communication. In N-1 mapping, each node maintains a map of its neighbors, which is updated through V2V communication. The node also uses its GPS receiver to determine its own location. Each node combines its own location with the location of its neighbors and the network map created. This information is then shared with other nodes through V2V communication. Over time, each node will develop a complete map of the network. N-1 mapping is a versatile and effective technique for mapping nodes in VANET. It can be used to improve routing, safety.

3.4.3 Deep Auto Spare Impasse NN

To detect twelve variants of hybrid DDoS attacks by blocking external hosts and provide end-to-end network transparency, the approach uses Deep Auto Sparse Impasse NN. The system aggregates data from sensing and mapping reports to facilitate the detection of the attacker node within the network. The diagram of is given in Figure 3.11.

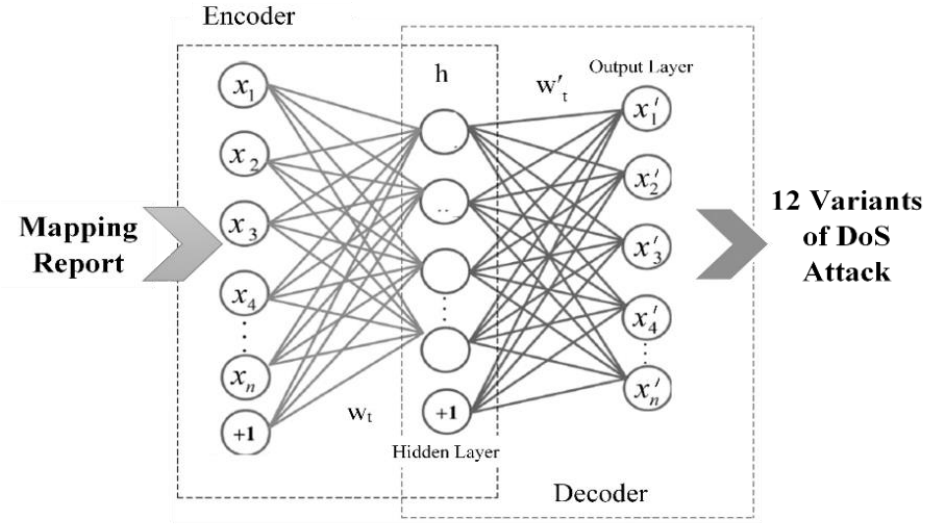

Figure 3.11: Deep Auto Sparse Impasse Neural Network

The Deep Auto Sparse Impasse Neural Network is a type of structure made up of numerous basic neurons acting as nodes or elements. These components are constantly working together in parallel. The connections between the neurons have a significant role in how the Deep Auto Sparse Impasse Neural Network functions. These neurons are linked together via links, and each link has weights that are adjustable values. The neural network architecture employs an input layer that receives the mapping report. To enhance efficiency, the number of connections between neurons is minimized. This reduction in connections not only conserves memory but also has a negligible impact on the network's accuracy. Notably, this optimized network demonstrates a prediction speed that is twenty-five times faster than a traditional neural network.

When dealing with Deep Auto Sparse Impasse neural networks, the number of layers and nodes are chosen using similar concepts as in standard neural networks, but there are several concerns that are unique to these sparse networks. Neurons are grouped logically in three layers that make up the Deep Auto Sparse Impasse NN. The Deep Auto Sparse Impasse NN is a three-layer neural network with a single neuron in the output layer and a variable number of neurons in the hidden layer. The values of each output vector member fall between [-1, 1]. Neurons on both layers have "tan-sigmoid" transfer functions. This function maps any input value, regardless of its magnitude (which can

range from negative infinity to positive infinity), to an output within the range of -1 to 1. The Deep Auto Sparse Impasse NN contains fewer active connections or parameters, resulting in more efficient and interpretable models. The connections that are present and that are pruned or set to zero are determined by the sparsity pattern. The pruning approach used in this sparse neural network eliminates the superfluous connections and reduces the number of parameters.

The unformatted training set is used in the Auto Sparse Impasse NN to provide the mapping report as the auto-encoder input data and it is shown in equation 3.26 below:

$$x = (x_1, x_2, \ldots\ldots, x_n) \tag{3.26}$$

The hidden and output layer neurons are activated by sigmoidal activation function which is shown in the below equation (3.27).

$$g(S_k) = \frac{1}{1+e^{-S_k}} \tag{3.27}$$

where, $'S_k'$ represents the cumulative input signal of the kth neuron in the hidden or output layer of the NN and it is given in equation (3.28) below.

$$S_k = \sum_{i=1}^{n}\left(w_{i,k}x_{i,k} + x_0 f_k\right) \tag{3.28}$$

where, $w_{i,k}$ is the the link weight from the previous layer's ith neuron to the hidden or output layer's kth neuron, $'x_0'$ is the input link weight of neuron and the offset of kth neuron is represented by $'f_k'$.

The below equation (3.29) shows the output data signal from the neural network having L number of neural layers,

$$h_{w,f}(x) = A^{(L)} \tag{3.29}$$

where, $A^{(L)}$is the value array of output layer neuron.

The twelve variants of DDoS attack are identified and predicted by this Sparse Impasse output layer. The twelve variants are tabulated in Chapter 4. Only a subset of connections or weights is active in the Deep Auto Sparse Impasse NN, with the remainder set to zero. Because zero-valued connections do not need to be processed, this sparsity

minimizes computing costs during training and inference. Consequently, fewer procedures are needed; resulting in quicker training time and this NN has high prediction accuracy for DDoS attacks. This mode effectively detects a broad spectrum of DDoS attacks, encompassing those targeting protocols such as NTP, DNS, LDAP, MSSQL, NetBIOS, SNMP, SSDP, UDP, UDP-Lag, WebDDoS, SYN, and TFTP.

3.4.4 Phase 2 Merits

The proposed encrypted mapping approach in phase 2 has been considered based on the attack detection rate and the deviation in detection accuracy. The proposed approach demonstrated a reliable detection process, maintaining accuracy regardless of the increasing number of vehicle nodes, leading to improved attack detection. The performance of the proposed approach was evaluated by comparing it with existing methods, including Trilateral Trust, H-IDS, Multi Filter, and SPPA. Results demonstrated that the proposed approach, utilizing encrypted traffic and hex-tuple mapping, achieved significantly higher attack detection accuracy.

3.4.5 Phase 2 Limitations

The traffic encrypted using TRHE reduces the data breaches and unauthorized access of data. The mapping through hex-tuple matched mapping provided N-1 mapping and end-end network transparency. The mapping report detected the twelve types of DDoS attacks. However, the continuous availability of the services in the VANET must be ensured through the rapidly changing topology and vehicles involve in mapping with external hosts and fluctuating changes in the network. To address the continuous availability of the services in the VANET, further enhancement is focused in phase 3 proposing the immunization and routing traffic away from overloaded areas of the network.

3.4.6 Phase 2 Summary

A novel attack detection approach is proposed in which Triple Random Hyperbolic Encryption with Hex-Tuple Matched Mapping using Deep Auto Sparse Impasse NN encoded the traffic three times in a random manner with the public key determination. Data breaches and unauthorized access are prevented. The end-to-end

transparency in the VANET through the mapping ensured the network secured form the external hosts and the approach enabled the detection of twelve types of DDoS attacks. The performance in detecting twelve types of DDoS attacks with encryption and mapping is assessed and found that the detection time is reduced providing higher accuracy. The continuous availability of the services with the encrypted traffic routed away from the overloaded traffic and also to detect and isolate the attacks due to the dynamic topology are addressed through the proposed phase 3. The contributions are discussed in the following section.

3.5 Phase 3: Immunization of Clusters and Routing

3.5.1 Introduction

The network transmission stability has been ensured through the encryption of the network traffic with mapping and the DoS attack types detected with the host blocked in the phase 2. The rapidly changing topology with the fluctuating changes in the network must consider interactions between the two vehicles and the history of the two vehicles for the continued services. Due to the dynamic nature of VANET, nodes may require adaptive routing strategies to accommodate the constantly evolving network topology. Disruptions in data transmission and increased congestion are often caused by varying connection status and network latency. To address the variations and provide continuous services without disruptions in network performance, the deep neural network based on the trust scores is considered in this phase as the changing conditions in the VANET due to the dynamic topology include varying nodes behavior.

Trust scores are calculated for each vehicle based on the frequency and success of their interactions with other vehicles and their historical behavior. Vehicles with high trust scores are deemed more reliable, while those with low scores are considered less trustworthy. Vehicles with extremely low trust scores may be temporarily isolated from the network to mitigate potential harm. Furthermore, VANET resilient feature to DDoS attacks and providing a more reliable service to their users has been proposed by optimization and routing focusing on dealing with the goal of DDoS attacks in impacting VANET with packet loss and overwhelmed traffic. Phase 3 contributions are based on identifying and isolating the source when number of packets is being dropped from a

specific IP address and also directing traffic away from congested areas and towards more reliable paths. The contributions proposed are based on Deep Trust Factorization Neural Network (DT-NN) with Moth Flame Optimization algorithm and Cache parallelized circulation link routing.

DDoS attack detection through trust score calculated and balanced cluster ensuring connectivity using optimization algorithm with routing to effectively manage the dynamic fluctuation of each node are discussed in the following subsection with steps.

3.5.2 Contribution 6: Deep Trust Factorization Neural Network with Trust Score

Phase 3 aims in obtaining the objective to enhance the security and reliability of VANET by developing effective mechanisms to identify and isolate malicious vehicles that launch DoS attacks, thereby disrupting network communication. Its operation relies on a combination of trust score calculation, deep learning, and adaptive trust updates to accurately identify and isolate malicious vehicles, safeguarding the integrity and reliability of VANET communication.

The contribution towards enhancing the security and reliability is provided by the Deep Trust Factorization Neural Network (DT-NN) based on trust scores. The operations in the DT-NN are listed below.

- Identifying and isolating malicious vehicles: DT-NN with trust scores identifies and isolates malicious vehicles in real time. This can help to prevent malicious vehicles from disrupting the operation of the VANET.
- Protecting critical infrastructure: DT-NN uses trust scores to protect critical infrastructure, such as roadside units (RSUs) and traffic lights from DoS attacks.
- Ensuring the reliability of VANET-based services: DT-NN leverages trust scores to enhance the reliability of VANET-based services, including traffic information and emergency response systems.

In this phase 3, DoS and DDoS attacks are mitigated by isolating the attacked node using DT-NN. The architecture of DT-NN incorporated in the phase 3 is responsible for detecting the trust values, thereby predicting the rational nodes, consists of an input layer,

a hidden layer, and one output layer. The working mechanism of DT-NN is based on the three functionalities

i. Trust score calculation
ii. Deep model training
iii. Trust score updates and malicious nodes detection

The computation of trust scores for each vehicle in the network. This involves gathering various network metrics that reflect the vehicle's behavior and adherence to protocols. These metrics include packet forwarding efficiency, participation in cooperative services, compliance with network protocols, response time, resource consumption, consistency of behavior, and feedback from other vehicles.

DT-NN utilizes a deep learning model to analyze trust score data, uncovering intricate patterns and relationships within this information. DT-NN continuously monitors network behavior and updates trust scores for each vehicle based on real-time observations. This ensures that trust scores remain relevant and reflect the latest changes in vehicle behavior.

The architecture of Deep Trust Factorization NN is shown as in Figure 3.12.

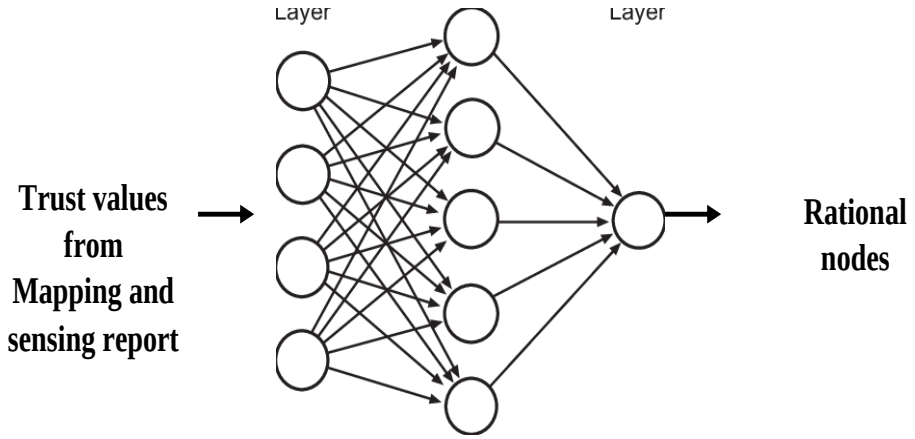

Figure 3.12: Deep Trust Factorization NN

The data gathered from the DDoS attack detection is given as the input for the trust factorization. The output layer comprises a single neuron, responsible for determining the trust score assigned to each node by the hidden layer. The number of input nodes in a

neural network model for network traffic analysis is determined by the specific features used to represent the traffic data. These features may include source IP address, destination IP address, port numbers, protocol type, and packet size. The initial layers of the hidden network extract low-level features from the raw input data. These features might include various characteristics of network packets, such as source and destination IP addresses, port numbers, packet sizes, and protocols.

By using multiple hidden layers, the network can learn to detect and represent increasingly abstract and complex features. As data flows through the hidden layers, the network learns to recognize patterns and relationships in the input data. These patterns might include traffic spikes, unusual traffic patterns, and abnormal behaviors that could indicate a DDoS attack. In the output layer, nodes are classified as rational and irrational nodes.

The random range of the VANET still remains as a threat in providing the stable operations in the VANET. The stability can be enforced for the transmission of data in vehicular networks through the optimization algorithms. For the transmission of data, the routing mechanism is considered for the clusters with the vehicle nodes having high PDR. The high PDR nodes form the balanced cluster group enhances the resilience of VANET against DoS attack types. This immunity is ensured to provide stability in the VANET with transmission of data through Moth Flame Optimization and Cache Parallelized Circulation Link Routing proposed in the phase 3. The immunization with the routing is detailed in the below subsection.

3.5.3 Moth Flame Optimization with Cache Parallelized Circulation Link Routing

Drawing inspiration from biological immune systems, immunization techniques have emerged as a promising approach for enhancing the resilience of VANET against malicious attacks and concurrently improving Packet Delivery Ratios (PDRs). The Moth Flame Optimization (MFO) algorithm, a nature-inspired metaheuristic inspired by the behavior of moths attracted to flames, effectively balances cluster formation and reliable node selection within VANETs. This is applied in phase 3 with the objective to detect the best path in the network by updating each position in the node in enhancing the resilience of VANET.

MFO-based immunization applied in this phase significantly improves PDR and network performance. The mechanism in the MFO-based immunization relies on the following steps.

Step 1: *With the cluster formed, MFO identifies the trusted nodes based on signal strength, mobility of nodes and the past nodes performance.*

Step 2: *MFO-based immunization techniques isolate malicious nodes from the network, preventing them from disrupting data transmission and compromising network security.*

The initial considerations are the vehicle nodes random location, its motion and followed by the search process. The fitness function is calculated based on the vehicle density. The search process continues in the transverse orientation and the spiral motion of the node with its next node in the search space. The updation of the vehicle node position and the fitness function with the sorting are performed during the process. The initial function considered is shown as in equation (3.30).

$$M = (I,P,T) \tag{3.30}$$

In equation (3.30),

I signifies vehicle location at random ($I:\phi \rightarrow \{M\}$),

P indicates the vehicle motion in the search space

($P:M \rightarrow M$),

T concludes the search process ($T:M \rightarrow true,false$) and the I function is represented below:

$$M(i,j) = \left(ub(i)\right) - lb(j)\ X\ rand(j) + lb(i) \tag{3.31}$$

In equation (3.31), ‘lb’ and ‘ub’ indicate the lower and upper bounds of the variable.

The search follows a transverse path. The initial spiral point originates from the current vehicle node, while the final spiral point aligns with the position of the next

vehicle node. The search range should remain within the defined search space. The spiral equation follows as,

$$S(M_i, F_j) = D_i \cdot e^{bt} \cdot cos(2\pi t) + F_j \quad (3.32)$$

In equation (3.32),

D_i refers to the space between the i^{th} node and j_{th} node, 'b' defines a fix to logarithmic spiral shape, and 't' specifies a number randomly between [-1 1].

The Moth Flame Optimization Algorithm (MFO), detailed in Table 3.7, maintains a balance between consideration and utilization through the spiral movement of each node around the next node within the search space. The algorithm initials with a vehicle and M_iA random vehicle. The number of vehicle is 30, the vehicle speed ranges from 20-60 km/hr, its radio frequency value is 2.47 GHz, the bandwidth value is 4.3 Mbps and the range value is 250m. The iteration took place to detect paths with a high packet delivery ratio (PDR). By optimizing cluster formation with the selection of reliable relay nodes and isolating malicious nodes, MFO can significantly improve PDR, reduce packet loss, and enhance VANET security. The malicious nodes are reduced in the network by MFO-based immunization algorithm and VANET is highly scalable due to its dynamic topology, further enhancement in efficiently handling a large number of nodes.

The fitness function moves the vehicles towards high packet delivery ratio (PDR). Efficient routing of packets around congestion and malicious nodes is still a challenge with the random range of vehicle nodes. This is also addressed by considering multiple paths simultaneously to avoid congested or unreliable links. The Cache parallelized Circulation Link routing (CCL) is applied to make time and frequency synchronization base channel hopping on a network. This effectively manages the dynamic fluctuation of each node and removes the transmission of dangerous malware files on the network.

Table 3.7: Moth Flame Optimization Algorithm

Initialize the parameters for the vehicle

Initialize the vehicle at a positionM_i randomly

For i =1 to ndo

calculate the fitness function f_i

end for

while iteration ≤Max_iteration do

Update the position of M_i

Calculate the number of vehicles

Evaluate the fitness function f_i

if iteration == 1 then

F=sort(M_(t-1,) M_t) and OF=sort(M_(t-1),M_t)

end if

for i=1 to ndo

for j=1to ddo

update the values of r and t

calculate the value of D concerning its corresponding vehicle

update M(i,j) respect to its corresponding moth using

end for

end for

end while

The Cache parallelized Circulation Link routing is applied to make time and frequency synchronization base channel hopping on a network to effectively manage the dynamic fluctuation of each node and remove the transmission of dangerous malware files on the network. The method eradications of the twelve variants of hybrid DDoS attacks without reducing the high packet delivery. The circular routing process packet and hoping process on circular link state is a process of the nodes; instead of changing the channel randomly, each node knows the sequence where they should be and always able to communicate. The optimization algorithm connects the node in the circular link to

make the nodes in regular circular updating with effective hopping between one node and another node.

On implementation of the phase 3, the number of bits transferred in the time and the deviation loss or packet loss on the network, the vehicle density and the total packets sent and received in the network, the duration to send and receive the packets has been recorded, the number of requests send to cause DDoS attacks such as DNS flood, HTTP flood, fragmentation attack, NTP amplification, Ping flood, and SNMP reflection, and the proposed model is capable to identify twelve types of DDoS attacks and protect the network from packet loss and other hybrid attacks.

By using CCL to make time and frequency synchronization base channel hopping on a network, it is possible to improve the performance of the network in a number of ways. CCL helps to reduce the amount of interference between nodes thus improving the throughput of the network. CCL can also help to reduce the amount of latency in the network, which can improve the responsiveness of the VANET.

CCL improves the security of the VANET by making it more difficult for malicious nodes to disrupt the network. CCL isolates the malicious nodes from the network, preventing them from sending packets or disrupting the routing of packets. Thus improving the performance of the VANET is achieved as CCL makes the time and frequency synchronization base channel hopping to achieve improved throughput, reduced latency, and enhanced security.

The parameters life and fall based on the objective function in securing VANET are considered on the following factors.

- Dynamic network conditions: Changes in traffic density, vehicle speed, and environmental factors can impact packet life and loss rates.
- Evolving attack patterns: New types of DoS attacks with different characteristics can affect the life and fall of detection and mitigation mechanisms.

The contributions are devised to secure the VANET based on the mentioned factors to sustain the reliability of the VANET services in minimizing the rate at which packets are dropped and maximizing the ability to detect the DoS attacks and mitigate.

The proposed methods are evaluated using the various parameters contributing to the resilience of the VANET in Chapter 4.

3.5.4 Phase 3 Summary

Overall, phase 3 has been proposed to provide immunity to the VANET network against the twelve types of hybrid DoS attack by Encrypted Access Hex-tuple Mapping Attack detection, a process in which all the traffic in the network is encrypted using triple random hyperbolic encryption and the middlebox does the N to 1 mapping of the IP address. The Stable Automatic Optimized Cache Routing used deep trust factorization NN to detect the irrational node by adding the trust score and Moth Flame Optimization algorithm to obtain the high packet delivery ratio. The Cache parallelized circulation link routing is applied to synchronize the time and frequency of each node, eliminating the network's malicious traffic.

The proposed phase 3 thus provided immunity to the DoS attack types with data transmission on the VANET architecture. The cache routing is introduced in which Deep trust factorization NN adds trust value for each node and the moth flame optimization is used to form a balance between the cluster to produce the high packet deliver ratio of 98% thereby detecting the malicious nodes and ensuring the linkage of each node in the cluster. Cache parallelized circulation link routing is used to provide multiple parallelized paths to each node and time and frequency synchronization to packets thereby removing unwanted traffic from the network so the response time of each node is reduced to 0.1 ms.

3.6 Chapter Summary

In VANET, the overwhelming of the network due to the increased messages among vehicles remains as the threats leading to unavailability of VANET services. The DoS attacks and its types create disruption of communications, unavailability of services with packet loss and packet drops affecting the stability of VANET. The method proposed in three phases contributed to the DoS and DDoS attacks mitigation addressing the issues relating to the unavailability of services and disruptions in communication. The stability of the VANET is ensured with immunization and routing methods. The proposed contributions in the phase 1 detected and isolated the malicious DoS attacked vehicles

through the features selected relevant to DoS attacks. The optimized feature selection GLW – SLFN and the enhanced feature selection using MCOD-LR, Kernel Density Estimation and Entropy-based SVM, Bayesian aggregate model with Self-healing AIS in the phase 1 are used to detect the malicious DoS attacks with classification. In order to mitigate on detection with isolation and prevention, phase 1 used Kernel Density Estimation and Entropy-based SVM, Bayesian aggregate model with Self-healing AIS. The self-healing effect inherent within the AIS significantly enhanced the robustness of the VANET against DoS attacks. This was accomplished through the successful identification, classification, and subsequent isolation of attacks, culminating in an increased detection rate of 97%, a 39% reduction in energy consumption, and a latency reduction of 25%.

Triple Random Hyperbolic Encryption (TRHE) with Hex-Tuple Matched Mapping in the phase 2 is used to encrypt the traffic signals and classified the twelve DoS attacks using the Deep trust factorization NN with the mapping reports. The stability and immunized behavior of the VANET are contributed in phase 3 by Deep Trust Factorization Neural Network with trust score, Moth Flame Optimization Algorithm and Cache parallelized circulation link routing. The immunized behavior of the clusters provided stability among changing topology and routing provided data transmission with reduced packet loss and increased PDR.