International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V70I2P211

Volume 70 Issue 2, 75-85, February, 2022
© 2022 Seventh Sense Research Group®

Original Article

A Trust Management Scheme for Intrusion Detection System in MANET using Weighted Naïve Bayes Classifier

Fouziah Hamza[1], S. Maria Celestin Vigila[2]

[1]*Research Scholar, Noorul Islam Centre for Higher Education, Kumaracoil.*

[2]*Associate Professor, Department of Information Technology, Noorul Islam Centre for Higher Education, Kumaracoil, Tamilnadu, India.*

[1]fouzi_ajmal@yahoo.com, [2] celesleon@yahoo.com

Abstract - *The primary research considerations are the development of intrusion detection and preventing mobile ad hoc networks (MANET) techniques with an exact detection rate and energy consumption with low packet loss. Node energy and node mobility are two major optimization challenges in MANETs, in which nodes move insecurely in all directions, and the topology is constantly changing. A significant clustering is carried out by the Emperor Penguin Optimization (EPO) algorithm. The cluster head selection is processed using a fuzzy strategy with a genetic algorithm (GA). The motive of this work is to use a trust management method based on DempsterShafer (D-S) evidence theory to identify intrusion behaviour. In addition, the attack pattern classification using the weighted Naive-Bayes method reduces the complexity of the classification. The features are extracted from the recognized pattern and finally transferred to the classifier for classification. Learning complexities can be overcome during the classification process using the Social Spider Optimization (SSO) technique. The proposed mathematical model detects the intrusion based on the final trust score. The robustness of the proposed model is executed based on attack detection rate, energy usage, and throughput for detecting and isolating the intruder. According to simulation results, the proposed solution effectively reduces IDS traffic and overall energy usage while maintaining a high attack detection rate and throughput.*

Keywords - *Intrusion detection system (IDS), clustering, cluster head election, classification, weight optimization, security.*

I. INTRODUCTION

A mobile ad hoc network (MANET) is a set of self-organizing mobile nodes that may connect with one another without the assistance of a solid infrastructure or a central coordinator. A node is any mobile device that can work together with the neighbouring devices, and here the MANET node works like both a host and a router. A node not in its communication range with the neighbouring node requests the intermediate node to transmit its message [1]. MANETs have diverse benefits over other networks in that they can be set up effectively, apart from the fact that they offer flexibility due to the unbound nature of the nodes. Due to MANET's very short battery life, it is impossible to keep IDS processing constant on every node of MANET [2] [41-43]. In addition, the existing MANET IDS systems do not consider the operating environments, so they result in each individual node being monitored with a uniform probability, regardless of whether the monitored node contains a summary of the description as malicious.

Several kinds of network anomalies are observed and are decreasing due to their dynamic and infrastructural nature [3]. The reasons for these anomalies are network congestion, faulty network equipment, active attacks and intruders. Intrusion is one of the critical anomalies affecting the integrity and availability of network services [4, 5]. The most well-known type of network attack is the Denial-of-Service (DoS) attack, which compromises the service for new authentic users [6-9]. There are many DoS attacks, namely black holes, wormholes, grey holes, and flooding [10-12]. Everyone on the network causes different types of security breaches. Some of the factors affecting wireless systems are disconnections, traffic floods, barriers to entry, or system problems [13]. MANET is an open communication medium; therefore, they are more prone to attack. The two major groups of attacks are active attacks and passive attacks.

Several methods have been proposed for detecting active attacks, but passive attacks are more complex in ad hoc network environments. Therefore, much research has been done to protect ad hoc networks from attackers. MANET [38] is developing a direct and easy implementation technique for IDS to stimulate the detection process on each network node. However, in such energy-constrained environments, this solution is ineffective. To deal with this, a common approach is if divide the MANET into clusters and select the cluster head for each cluster. Many optimization algorithms are efficiently used in cluster head (CH) selection [39, 40]. The CH's job is to defend the entire cluster by running its IDS. Intrusion detection theoretical models describe network security with enhanced decision-making. Usually, in these approaches, the decision-makers can take over the role, i.e., an attacking role or a defending role; the two goals are incompatible. In a network, an attacker aims to break into network activities. For dealing with intrusion detection problems, the game theory models [14] are most often used. These models perfectly illustrate the opposing goals of a defender and an attacker [15-18]. The interactive decision states are described and examined using game theory. It also ensures a robust group of devices for analyzing and forecasting the outcomes of complicated relationships between rational entities [19, 20].

A. Motivation & Major Contribution

In this study, the major attempt is to create a cluster-based model and a weighted Bayesian classifier for attack categorization. There are three models used for identifying the intrusion node in MANET. The CH is chosen initially based on energy consumption and mobility, encouraging compact cluster formation. Second, a trust-based management system and weighted Naive Bayes with social spider optimization are used to increase security. This hybrid combination minimizes the error rate and improves the accuracy of the classification. The social spider optimization has a higher convergence rate since this combination also reduces the time required. So, the type of attack is identified easily. Because of this combination, network energy is consumed, and the packet loss ratio is reduced. This work's key contribution is summarized below.

- Proposes a unique strategy for improving the intrusion detection rate of intrusion detection systems (IDSs) in a MANET based on clustering and the Naive Bayes Classifier. The technique increases the rate of IDS attack detection as much as feasible without sacrificing efficiency.
- Swarm optimization is proposed for cluster head selection and clustering with reduced network activity to validate the proposed strategy.
- Finally, the attack detection is carried out using the weighted Naive Bayes Classifier for attack detection.
- Simulation results proved that the proposed strategy for optimizing the IDS attack detection rate (DR) could result in significant energy savings and improved throughput.

B. Work Organization

The remaining part of the research work is constructed as: Recent literature works are provided in Section 2. A brief explanation of the proposed method using flow charts and algorithms are provided in Section 3. The graphical representations used to describe the simulation are given in Section 4. Finally, Section 5 comes to conclude the work.

II. LITERATURE REVIEW

R. Santhana Krishnan et al. [21] proposed a Modified Zone-Based IDS (MZBIDS), a novel intrusion detection system for MANETs. MZBIDS, when tested using modern methods, has a higher success rate in detecting malicious behaviour in persistent conditions while having a minor impact on overall network performance.

Neenavath Veeraiah et al. [22] proposed an efficient multipath routing protocol in MANET based on an optimization approach. In MANET, CH selection and IDS techniques such as fuzzy clustering and fuzzy naive Bayes (Fuzzy NB) were useful in addressing the energy and security crisis. The Bird swarm-whale optimization algorithm (BSWOA) was used to perform multipath routing using the safe nodes by routing. The best routes were chosen based on fitness criteria, including connectivity, energy, trust, and throughput.

Using machine learning approaches, Islabudeen, M. et al. [23] suggested an intelligent methodology for the Intrusion Detection and Prevention System (SA-IDPS) to alleviate attacks in MANET. The One Way Hash Chain feature was used to register mobile users in Trusted Authority. Packet Analyzer was used to check if an attack pattern had been discovered. It is implemented with a Type 2 fuzzy controller that considers packet header information. Logarithmic normalization and coding methods, which were time series and suitable for all applications, are considered in the pre-processing unit. Mutual information is used in the feature extraction unit to obtain the best set of characteristics for packet classification.

The authors used a clustering algorithm to group the MANET nodes and locate malicious nodes, and the clustered groups were then studied further. With the routing methodology of Destination Sequenced Distance Vector Routing (DSDV) [24], the watchdog protocol identified and classified the malicious nodes. Link failures were detected and classified by the proposed work based on the presence of malicious nodes. For detecting the malicious nodes, the cooperative bait detection approach in a MANET environment was used by the authors of [25]. The authors categorized and detected grey hole or collaborative blackhole attacks in the MANET system. In [26], the authors presented an approach to the DDoS attack. Firecol's score was updated to identify the potential attack in this approach. DGSOT plays an important role in clustering, relying on the routing level score.

The authors in [27] proposed a work based on incorrect data filtering in the routing nodes. When the mobile ad hoc network was first deployed, keys were provided to the nodes of the network using the Advanced Encryption Standard (AES). The use of the message authentication code validates the data. The forwarding node location and timestamp are examined for detecting the anomalies by the forwarding nodes. The location and timestamp are used to monitor node activity on the network. The attack simulation is identified in the network using the message authentication code. The trust score was designed based on the belief system, and a classification approach was used for attack detection. These have been very effective schemes against network attacks [28]. A hierarchical design was chosen to design effective IDS [29]. The major aim of these approaches was to consolidate the data that is deprived of full information about the common environment. The authors presented a novel clustering algorithm in [30] to overcome some of these disadvantages. The authors proposed a new self-organizing network with a tree structure called the Dynamically Growing Self-Organizing Tree (DGSOT) algorithm for hierarchical clustering. The proposed algorithm has improved the number of clusters at each hierarchical level.

They also presented a new cluster validation criterion with respect to the symmetric property of the data set's Voronoi division. A minimal spanning tree concept was used in this approach. A K-Level-Up Distribution (KLD) approach was used for improving the accuracy of clustering, which enhances the distribution of the data in the hierarchy construction. The result of clustering the self-organizing trees is visualized as a dendrogram. Cluster self-organization was well suited for this architecture [31]. For planning, self-organization has no external or central authority. The most important theories of self-organization were presented in [32, 33]. In [34], an algorithm called DGSOT was presented with some of the classified features that are consistent with the security enhancement of IDS. When designing new IDSs, the trust assessment parameters should be considered in particular [35, 36]. Some of the prime numbers for designing the trust matrix are cluster domain, trust hypothesis, and frequency. The trust matrix determines the performance of a node.

III. PROPOSED METHODOLOGY

In this work, a new IDS approach for MANETs is provided. The MANET CH election approach and the weighted MANET IDS are two separate aspects of this intrusion detection technology. By dividing the intrusion detection cycle among several cluster nodes, the previous part minimizes the consumption of memory and time of computation required to control the IDS. The cluster head node is chosen depending on the consumption of energy and mobile node mobility. The selected CH node is given the authority to run intrusion detection services for the amount of time. The suggested IDS scheme's second component is a weighted MANET IDS based on a Naive Bayes classifier, which executes the actual IDS operation. The election model elects the cluster head node, which executes the weighted MANET IDS.

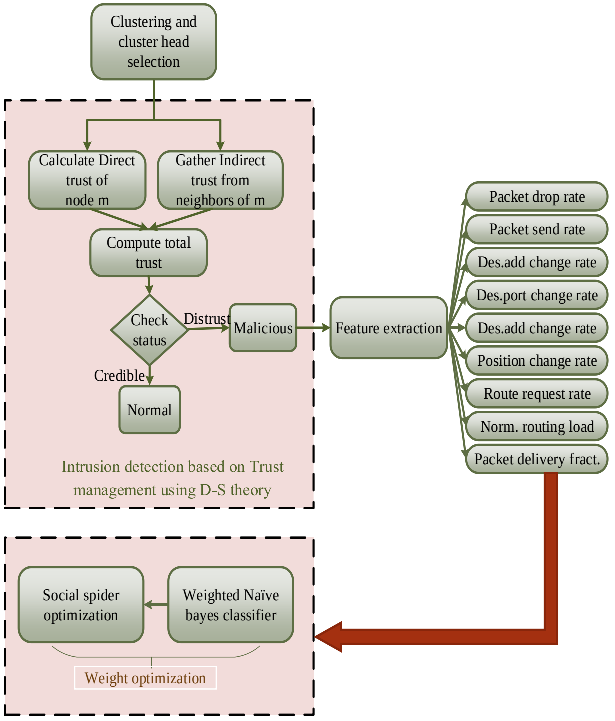

Fig. 1 Proposed Process Flow Diagram

A weighted IDS is constructed utilizing a tripartite model with MANET as the core network topology, as shown in Figure 1. CH election, cluster construction, and attack intrusion detection are three. Here, the D-S theory is adopted in a trust-based method for detecting the suspicious events of the mobile nodes in advance. Further, for normal and malicious nodes classification, the machine learning technique relied on Naive Bayes Classifier (NBC) to construct the Intrusion Detection System (IDS) is introduced. However, still, it has the limitation that different conditional attributes are independent of each other under the condition that the class decision attribute is known. At the same time, NBC considers that the conditional probability of each condition attribute has the same influence on results, which simplifies the algorithm logic and minimizes the computational complexity. Therefore, the proposed IDS model gives a certain weight to network intrusion features based on the impact of different features. The modified Naive Bayes algorithm follows the social spider optimization algorithm.

A. Clustering and cluster head selection

Effective cluster formation is carried out by using Emperor Penguin Optimization (EPO) [36]. After that, the selection of CH is processed by a fuzzy strategy with a Genetic Algorithm (GA) [37].

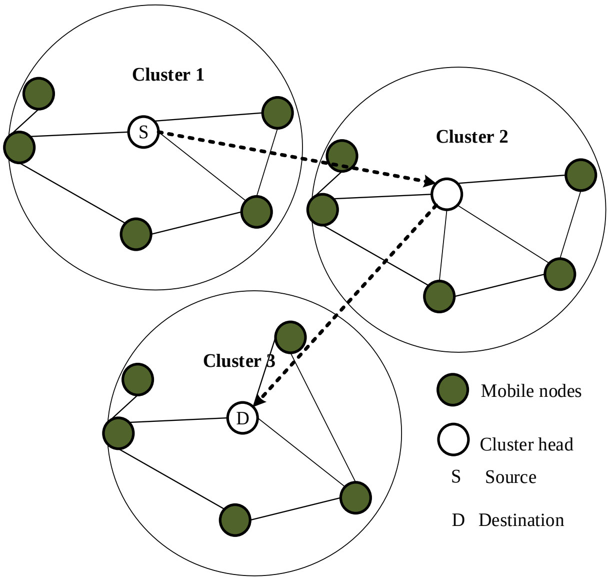

Fig. 2 Clustering and Cluster Head Selection

Both the random waypoint and the random direction models are similar. But only the difference is to travel and reach the destination. A direction is selected uniformly and randomly in the random direction model, whereas, in the random waypoint model, a destination is selected randomly. The node waits for a particular time upon reaching the boundary. After that, an alternative direction is selected uniformly and randomly. Hence, for the evaluations, a homogeneous distribution mobility model is provided.

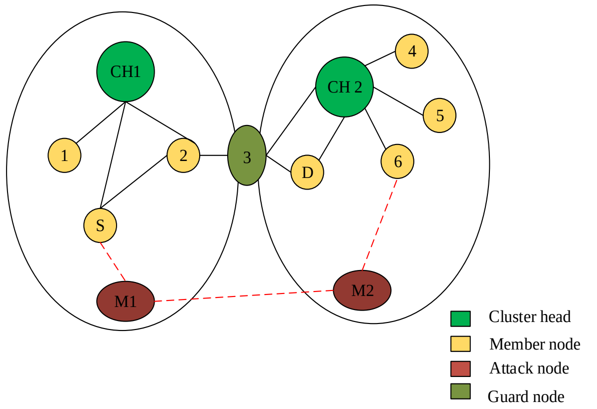

Fig. 3 Proposed Cluster-Based Intrusion Detection Technique

For intrusion detection, a trust model-based algorithm is presented. When an intrusion is suspected by a i^{th} cluster node in the first layer, the information is sent to the cluster head $CH(1,i)$. Then this information about the malicious node is sent to the second layer's cluster head $CH2$. All the cluster heads in the first layer are then informed $CH2$. Then the malicious node information is passed to all the cluster members by their cluster heads.

B. Trust management scheme for cluster-based intrusion detection

The major aim of the research is to make an IDS to the MANETs using cluster head and for classifying normal and malicious nodes, which degrade the performance of a network. The communication characteristics of neighbouring nodes are monitored by every member node in the clustered MANET. This collected data is then spontaneously transmitted to the CH. Based on the observed statistics of the evaluation node and the suggestions obtained from the neighbouring nodes, the cluster head calculates the trust score. Once the reputation of every node in the cluster is calculated, then it is saved in the cluster head.

a) Trust model

One node's belief about another, based on prior experiences, knowledge of entity behaviour, and/or recommendations from trustworthy entities, is referred to as trust. Based on D–S theory, the observation of the subject and the third party's suggestions can be utilized to determine the trust value. When the behaviour of an object is changed, then this trust value will be varied. Hence, the trust model comprises the following processes: (i) The Direct Trust Value (DT) among the object and service subject is calculated, and this value can be updated based on the perspectives of services (ii) The indirect trust values (IT) can be obtained by the service provider from member nodes using behaviour estimation (iii) The value of total trust can be obtained by synthesizing the DT and IT value.

The historical behaviour of direct contact is used to obtain the DT value in an entity. When an entity can't obtain the DT value, then the IT value will be calculated using the suggestions from third parties. At last, the weighted average algorithm is utilized for obtaining comprehensive trust value on the basis of the DT and IT value.

1) Direct trust value computation

Let DT be the direct trust value, and the judgement of the trust is represented as C_z. The DT between node m and node n can be calculated based on D–S theory:

$$DT_{m,n,z} = \left(pal_{m,n,z}\{T\}, pal_{m,n,z}\{F\}, pal_{m,n,z}\{T,F\}\right) \quad (1)$$

Where $\{T\}$, $\{F\}$ and $\{T,F\}$ denote entity trust, entity distrust and entity uncertain event, respectively. $pal\{\ \}$ Represents the basic reliability which indicates the occurrence degree of the event. It $DT_{m,n,z}$ can be calculated in two phases: initialization and update. Based on D-S theory, the computation of $DT_{m,n,z}$ being subject to the number of trust α_{z-1}, distrust β_{z-1} and uncertain γ_{z-1} behaviour. These behaviours will be updated continuously. If $C_z = (\alpha_{z-1}, \beta_{z-1}, \gamma_{z-1})$, then the following expression is used to update the behaviours.

$$\begin{cases} C_z = C_{z-1} + U \\ U = (u_1 + u_2 + u_3) \end{cases} \quad (2)$$

Where u1, u2, u3 are 0 or 1, and u1 + u2 + u3 = 1. The following two conditions must be considered while calculating the DT value between node m and node n:

- The evidence of trust is vacant when the node m isn't monitored the node n in the earlier context, i.e. $C_1 = (0,0,0)$ Hence, the maximum value is achieved by DT value due to the level of uncertainty between node m and n, i.e. $DT_{m,n,1} = (0,0,1)$.
- When n^th^ node's trust evidence is acquired earlier, i. e.

  $C_1 = (\alpha_0, \beta_0, \gamma_0)$, then the DT value can be computed by:

$$DT_{m,n,1} = \left(\frac{\alpha_0}{\alpha_0+\beta_0+\gamma_0}, \frac{\beta_0}{\alpha_0+\beta_0+\gamma_0}, \frac{\gamma_0}{\alpha_0+\beta_0+\gamma_0}\right) \quad (3)$$

2) Indirect trust value calculation

The third parties will recommend indirect trust value. Consider, $RT_{m,n}^{k}$ to represent the recommended trust (RT) value of m on n defined by the recommended node k, and $DT_{k,n}$ denotes the DT value of k on n. Hence, the neighbouring nodes of node n are used to determine its IT value. Let $RT_{m,n}^{k} = DT_{m,k} \otimes DT_{k,n}$, where, the neighbouring node is denoted as k, and the transfer operator is denoted as $\otimes$. Then $RT_{m,n}^{k} = pal_{m,n}\{T\}, pal_{i,j}\{F\}, pal_{i,j}\{T,F\}$, where $pal_{m,n}\{T\}$ represents the probability of the node m suggests node n as real. Because of $pal_{m,n}\{T\} = pal_{m,k}\{T\} \times pal_{k,n}\{T\}$ Then:

$$pal_{i,j}(\{F\}) = pal_{m.k}(\{F\}) \times pal_{k,n}(\{F\}) \quad (4)$$

$$pal_{i,j}(\{T,F\}) = 1 - pal_{m.k}(\{T\}) \times pal_{k,n}(\{T\}) - pal_{m,k}(\{F\} \times pal_{k,n}(\{F\})) \quad (5)$$

The trust value transfer may include enormous neighbouring nodes, which will be computed synthetically. If the set of neighbourhood nodes is $Neighbor(m \cap n)$, then the following expression is utilized to compute the IT value:

$$RT_{m,n} = \frac{1}{\left|Neighbor(m \cap n)\right|} \sum_{k=1,k\in Neighbor_{m,k}} ID_{k,n} \tag{6}$$

3) Final Trust score Calculation

The DT and IT values need to be combined to determine the exact trust value. The final trust value based on Equations (3) and (6) is given as:

$$Final\ Trust = DT_{m,n} \oplus RT_{m,n} \tag{7}$$

The eventual basic trust value transfer can be obtained by adopting an adaptive synthesis process. Because the complexities present in the computation will increase the cost of energy. The node n is recommended as trust node by the subject node m when the decision model fulfils the following condition:

$$\begin{cases} pal_{m,n}(\{T\}) - pal_{m,n}(\{-T\}) > \varepsilon \\ pal_{m,n}(\{T,-T\}) < \theta \\ pal_{m,n}(\{T\}) > pal_{m,n}(\{T,-T\}) \end{cases} \tag{8}$$

And node n is added into the mth node's trustworthiness list. A similar procedure is used to mark node n as distrust or uncertain.

Here, the data are aggregated that forms a data processing cycle. In every iteration, the security data aggregation procedure consists of 2 phases: setup and data transmission. In the setup, the clusters are created, and then CHs are chosen from every cluster. An authentication is given to the cluster head to manage their member nodes and coordinate the transmission time moderately. During the data transmission stage, the collected data are transferred to the cluster head by the member nodes. The data are aggregated in the CH, and then it will be transferred to the CH of the neighbouring cluster.

b) Steps for malicious node detection

The detailed steps for the malicious node detection procedure are provided below:

1: A request is sent to the neighbouring node m by the CH to demand the monitored data.

2: The cluster head identification is checked by node m. If node m belongs to this cluster head, then the data (collected statistics) are condensed and transferred to the cluster head. Or else it will discard the request message.

3: The DT value is combined with the IT value obtained from the neighbouring nodes of node m, that is $Neighbor(m) \in memSet(cluster\ head)$. Where cluster head's membership nodes are denoted as $memSet(cluster\ head)$ this value is the final trust value of node m.

4: The node m will be judged by the CH using the final trust value. The data obtained from node m will be aggregated if it is credible. It preserves the data in memory when its status is uncertain. The identity of the node m will be broadcasted by the cluster head when its status is distrust. Then, the data collected from this node will be discarded.

5: The data broadcasted from neighbouring node n is received by node m. Based on the decision interval, it will check the credible status of node n.

6: The transmission behaviour of the neighbouring node is monitored by the node m constantly. The neighbouring node's behaviour is considered normal when it properly transmits and receives data. If not, such a node is considered an abnormal node. The node m updates its neighbourhood behaviour table based on the monitoring results. Based on this, the value of α_{z-1} and β_{z-1} will be increased.

7: The direct trust value will be estimated based on these updated values of α_{z-1} , β_{z-1} and γ_{z-1} . The neighbourhood node trust list will reserve this value, and it will be utilized for evaluating the trust value in the subsequent round.

After finding the malicious node, it is essential for finding the type of malicious node. For identifying the type of malicious node, there is a need for feature extraction from the detected pattern.

C. Features of Interest

In this context, the features of interest are found out for a particular node which is useful for finding the type of attacks presented in the node. Choosing an optimum feature set is a complex process while doing the processes related to classification. Various factors are considered for the IDS in advance of choosing the features to be examined [30]. It must be capable of detecting and monitoring itself if an attacker compromises it. However, there should be a compromise between efficiency and effectiveness in selecting a feature set. The accessible traffic features are analyzed to monitor the radio range cluster. Every traffic feature is composed of subsequent concerns. From various sources and several layers, the packets are transmitted. The transmitted packets can be TCP data packets or route control message packets forwarded at the observed node. The route error message, request, and reply are included in the route control message packet.

The flow direction and the packet type are considered the first two attributes of the traffic feature. The flow direction considers the count of packets dropped, forwarded, sent or received. The packet type can be route-specific or data specific. The interesting information is routing control packet flow direction, MAC packet type, size of each packet, destination port, source port, a destination address, and source address. The packets are monitored by the local IDS in each node for every particular time to gather this information. For producing the desired features, once the raw data are gathered, they are processed by a simple program. Following feature vectors are used for the processing.

(a) Packet dropping rate with respect to the j-th node

$$R_{drop_j} = \left(drop_j / \mathrm{Re}c_{vj}\right) \Big/ \sum_{i=1}^{n} drop_i / \mathrm{Re}c_{vi} \qquad (9)$$

(b) Packet sending rate with respect to count

$$R_{send(count)_j} = send_j \Big/ \sum_{i=1} send_i \qquad (10)$$

(c) Packet sending rate concerning the size

$$R_{send(size)_j} = R_{send(count)_j} \Big/ \sum_{i=1}^{n} R_{send(count)_j} \qquad (11)$$

(d) Rate of changing the destination address

$$A_j = A_{send_j} \Big/ \sum_{i=1}^{n} A_{send_i} \qquad (12)$$

(e) Rate of changing port destination

$$P_j = p_{send_j} \Big/ \sum_{i=1}^{n} p_{send_i} \qquad (13)$$

(f) Rate of changing position

$$R_{\Delta position_j} = T_{\Delta position_j} \Big/ \sum_{i=1}^{n} T_{\Delta position_i} \qquad (14)$$

(g) The rate at which j-th node generates route request

$$R_{RREQ_j} = RREQ_j \Big/ \sum_{i=1}^{n} RREQ_i \qquad (15)$$

(h) Normalized Routing - Load

$$L_{(normalized)_j} = \left(route_j / recv_j\right) \Big/ \sum_{i=1}^{n} route_i / recv_i \qquad (16)$$

(i) Packet delivery ratio (PDR)

$$F_{packetdelivery} = \left(send_j / recv_j\right) \Big/ \sum_{i=1}^{n} send_i / recv_i \qquad (17)$$

The following section explains the parameters used in the feature vectors.

R_{drop_j} is the count of dropped packets by the j-th cluster node, $\mathrm{Re}c_{vj}$ denotes the count of received packets by the j-th cluster node, $R_{send(count)_j}$ indicates the count of transmitted packets by j-th cluster node based on the count, $R_{send(size)_j}$ denotes the overall size of transmitted data by the j-th cluster node, A_j is the count of addresses of destination for forwarding the data packets by the j-th cluster node, P_j represents various ports of destination for transmitting data by the j-th cluster node, $T_{\Delta position_j}$ represents the position changing time by the j-th cluster node, $RREQ_j$ denotes the count of route requests by the j-th cluster node, $RREP_j$ denotes the count of route replies done by the j-th cluster node, $route_j$ indicates the count of routing packets transmission by the j-th cluster node and n denotes the total count of network nodes.

D. Classification of Intrusion with weighted Naive Bayes classifier (WNBC)

This proposed classifier is used for finding the type of attack patterns based on features of interest. By considering that the variables are not dependent, the WNBC based classification solves this problem and considers conditional independence. For simplifying the computation, this assumption is made, and hence it is assumed to be "Naïve", and here learning complexities are reduced using social spider optimization. Moreover, high speed and accuracy are achieved when this approach is employed in large datasets.

The following subsections provide a more technical description of the classification.

a) Weighted Naive Bayes Classifier

This method gives conditional attributes for classification in which the resultant attribute (weight is 1). To replace the prior probabilities in the basic Naïve Bayes approach, the researchers propose a method based on the above observations by combining the Naïve Bayes classifier with various feature weighting approaches. In attack pattern detection, weighted prior probability is used in the proposed method to enhance the classification performance. Different weights are provided for different attributes based on classification importance, and this method is named weighted naive Bayes [31]. The weighted Naïve Bayes model is given as in the below equation,

$$C(F) = \arg\max_{C_i \in C} p(c_i) \prod_{K=1}^{n} p(F_k \mid c_i)^{w_k} \qquad (18)$$

Where the weight of attribute k is represented as w_k, in the classification, the impact of the attribute is high if the attribute weight is high. Determining the weights for various attributes is important in the weighted Naïve Bayes method. Therefore, the SSO algorithm is used to solve this issue.

b) Social Spider Optimization (SSO) algorithm for weighted Naive Bayes

The SSO algorithm [27] provides the optimal weights based on the membership functions. Each data from the sample is known due to the need for training features in the Bayesian classification. An important factor is to construct the membership function. In its own category, let the probability of data X be large if the probability of record should be large. However, in the other categories, the sum of probabilities of the data X should be minimum. A simple membership function is constructed based on this concept.

$$f = p(F \mid c_i) - \sum p(F \mid\sim c_i) \qquad (19)$$

The sum of probabilities of data that doesn't belong to c_i the class is referred to as $\sum p(F \mid\sim c_i)$. Using the SSO algorithm, the optimal weights are estimated for each record. After getting the optimal weights, the average of the weights of all features is estimated.

The algorithmic steps of SSO are given in table 1.

Table 1. ALGORITHMIC STEPS FOR SSO FOR WNBC

```
Step1: Read into the input features to calculate the first record's weights. Assign memory and generate the spider's population
Step 2: Each spider's $v_s^{tar}$ is initialized.
If the ending conditions are not completed, then do
Step 3: Calculate equation (19) for measuring fitness.
Step 4: Estimate intensity vibrations and position vibration
End if
Do pop for each spider
Step 5: Estimate vibrations intensity and Select the strongest vibration
Step 6:While $v_s^{best}$ intensity is greater than $v_s^{tar}$ then, Store $v_s^{best}$ as $v_s^{tar}$
        End while
Step 7: Update$C_s$.
Evaluate r random number from [0,1]
If $R > P_C$ then
Step 8: Update ms dimension mask.
        End if
Step 9: Originate $P_S^{FO}$
Step 10: Achieve an arbitrary walk.
Step 11: Specify constraints violation.
Step 12: When the weight evaluation data is finished, then stop; or else, start to evaluate the weight of the next record.
Step 13: Compute an average value of all weights.
```

E. Intrusion classification mechanism

A global alert message is launched, and a node initiates the process of cooperative identification once a neighbour node is detected as malicious. Using the linear opinion pool method, each node shares the self-evaluated probability values, and a combined decision is made.

$$p_{combined} = \sum_{i=1}^{n} W_i P_i \qquad (20)$$

Where the weights of each probability P_i are represented as W_i , and it has equal value when $\sum W_i = 1$. After combining the probability values, the system goes into the intrusion reaction stage if the final value > threshold value. Table 2 shows the corresponding feature vectors for intrusion behaviour.

Table 2. CORRESPONDING FEATURE VECTORS FOR INTRUSION BEHAVIOUR

feature vectors name	*Present Attack*
Sending rate (based upon total size)	Malicious activities-based Flooding attack
Sending rate (packet-count)	Flooding attack based malicious activities
Drop rate	Attack related to packet dropping
Fraction related to packet delivery	Sleep Deprivation
Rate of changing destination ports in packet sending	Malicious Flooding attack
Position Changing rate	Data combination
Destination changing rate address in packet sending	Malicious Flooding attack
Route Reply Rate	Routing table overflow
Normalized Routing Load (NRL)	Routing table poisoning
Route request Rate	Rushing attack

a) Global Reaction

In ad hoc networks, the intrusion reaction depends on the applications of network protocols. Some of the reactions are given below.

1) Between the nodes, the communication channels are re-initialized
2) The compromised nodes are identified
3) Exclude the compromised nodes, and the network is re-organized

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Here a developed model is implemented based on the sensing of data. Simulation is carried out with node availability, sensing range, transmission range and variable node density. In this domain, the nodes are arranged to the identical transmission. Simulation is carried out in Network Simulator 2 (NS-2) platform. By using some of the QoS parameters, the performance values are evaluated. The nodes are re-organized identically for the consecutive simulations in the system area. The proposed method (Weighted Naïve Bayes based Social Spider Optimization (WNBSSO) is compared with the recent existing works [32-35]

An experimental simulation is organized with 100 nodes with the same frequency B = 1 Mbit/s, and it has initial energy of 1J/bit/m2. Table 3 shows the performance used in the simulation.

Table 3. PARAMETERS USED IN SIMULATION

Parameter	value
Packet sending rate	1 packet/sec
Sink node	1
Initial energy	1J
Source node	1
Total number of nodes	100
Sink node	1
Total number of a relay node	98
Packet size	1024 bits
Distance between adjacent nodes	100 m
Total number of nodes	100
Deployment area	50m ×2500m

A. Throughput (T)

It's the overall packets sent to the target node successfully over a given time period. Bits per second (bps) is a unit of measurement that ranges from 0 to 100. When T it is high, the evaluation of the developed method is high, and when T it is low, it is minimum. It is given as:

$$T = \frac{Number\ of\ packets\ transmitted}{Time\ duration} \qquad (21)$$

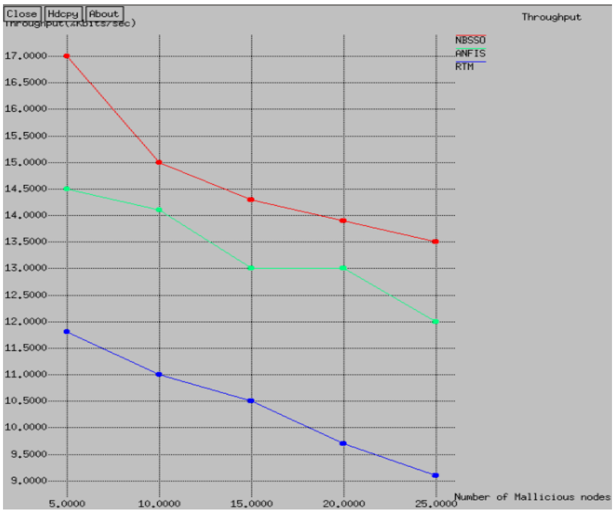

Fig. 4 Comparative analysis for throughput

Figure 4 demonstrates the comparative analysis of throughput (T). T is low when there is a higher amount of nodes and large when there are fewer number nodes. In this figure, the proposed NBSSO achieves 17 bits/sec for 5 nodes, 15 for 10 nodes, 14.3 for 15 nodes, 13.9 for 20 nodes and 13.5 for 25 nodes. Compared with two existing algorithms, the proposed approach yields better performance. In the case of RTM [32], it yields the worst performance (i.e., it produces minimum value for the throughput) than ANFIS [33] classifier.

B. PDR

It is considered the ratio of the total of data packets reached the destination to the sum of transmitted data packets.

$$PDR = \frac{Sum\ of\ packets\ received\ by\ receiver}{sum\ of\ data\ packets transmitted\ by\ transmitter} \qquad (22)$$

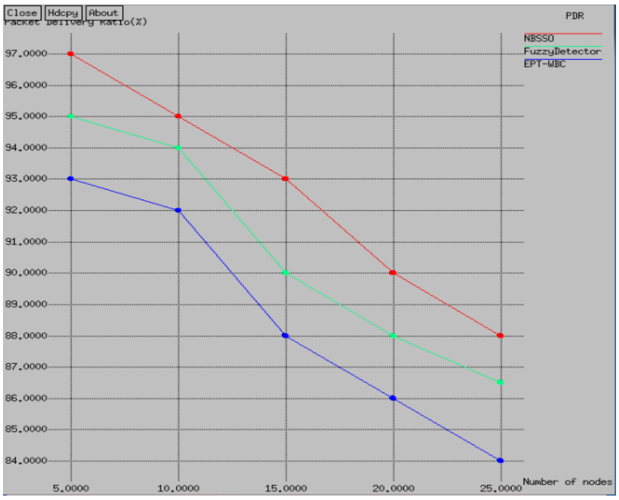

Fig. 5 Comparative analysis for packet delivery ratio

PDR must be maximum in case of effective transfer (Figure 5). When the PDR is high, the receiver can get all the data packets without losing any information. Here the PDR is maximum with a minimum number of nodes. In this figure, the proposed NBSSO achieves 97 % for 5 nodes, 95 for 10 nodes, 93 for 15 nodes, 90 for 20 nodes and 88 for 25 nodes. On comparing with the existing algorithms, the proposed work yields better performance. In the case of EPT-WBC [34], it yields the worst performance (i.e., it produces a minimum value for the packet delivery ratio) than the fuzzy detector [35].

C. Average packet loss ratio

It is described as the ratio between the overall amount of data loss within a particular period and the total amount produced at the particular node. This is given in percentage, and it ranges from 0 to 100.

$$Average\ packet\ loss\ ratio = \frac{Number\ of\ packets\ lost}{n} \times 100\% \qquad (23)$$

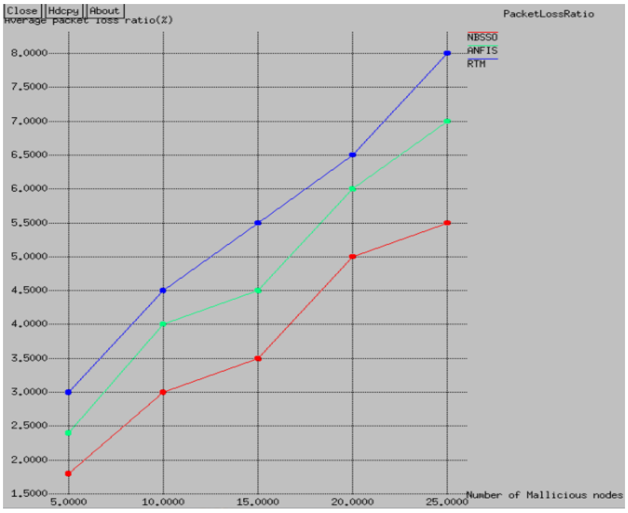

Fig. 6 Comparative Analysis for Average Packet Loss Ratio.

Figure 6 depicts the number of packets generated within the time. If the network has more nodes, then the average packet loss ratio is high and vice versa. The average packet loss ratio for the proposed system is 1.8, 3, 3.5, 5 and 5.5 for 5, 10, 15, 20 and 25 nodes. On comparing the three algorithms, RTM [32] yields less performance.

D. DR

It is expressed as the ratio among the count of exactly determined malicious nodes and the total count of malicious nodes in the network. In this paper, the malicious node is determined based on the type of attack. It is evaluated in percentage, and it varies from 0 to 100. It is formulated as,

$$DR = \frac{Number\ \ of\ \ malicious\ \ nodes\ detected}{sum\ \ of\ malicious\ \ nodes} \times 100\% \qquad (24)$$

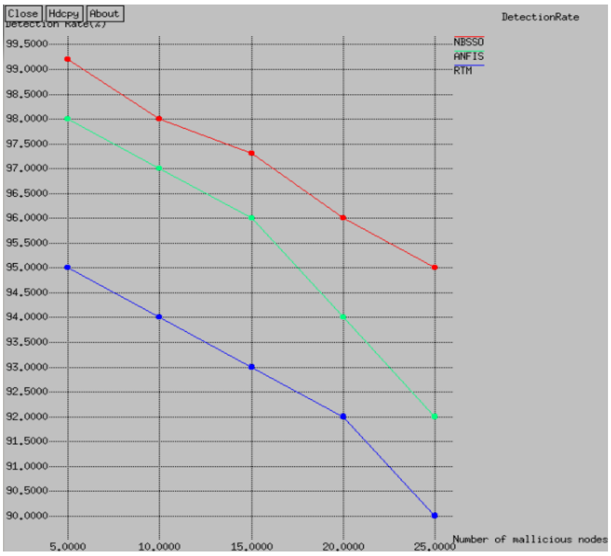

Fig. 7 Comparative analysis for detection ratio

Figure 7 depicts the evaluation study of the proposed IDS in MANET based on detection ratio. This ratio depends upon the overall malicious nodes. When it is high, the detection ratio is low, and the ratio is high, with a minimum number of malicious nodes. So the performance of the work is high with the maximum detection ratio. In NBSSO, the values of the detection rate are99.2, 98, 97.3, 96 and 95 for 5, 10, 15, 20 and 25 nodes. Compared with the two existing techniques, RTM [32] obtains a minimum detection rate than ANFIS [33], and since proposed, NBSSO has enhanced capacity for malicious detection.

E. Energy consumption

Energy consumption is huge when there are more malicious nodes.

$$E_{consumption} = (E_{RX} \times number\ \ of\ \ nodes) + E_{TX} \qquad (25)$$

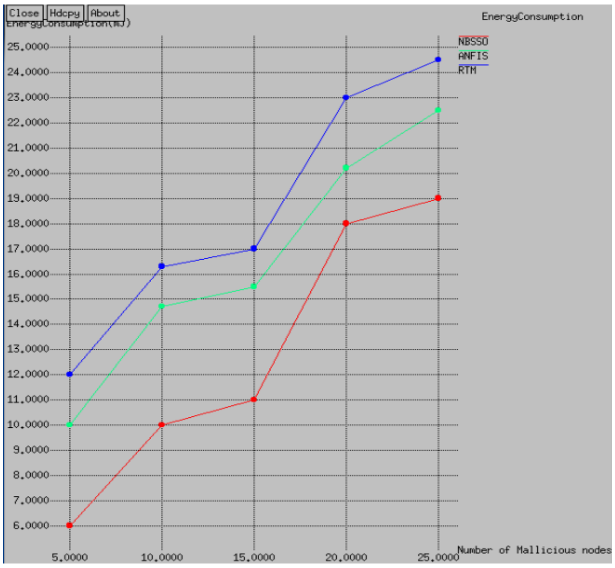

Fig. 8 Comparative analysis for energy consumption

Equation (19) is taken from [29], and it illustrates the evolution comparison of the current method with existing techniques. Figure (8) explains the performance of detection techniques for the proposed malicious node detection method. So energy consumption is more when more malicious nodes are presented in the network. In NBSSO, the values for the energy consumptionare6, 10, 11, 18 and 19 for 5, 10, 15, 20 and 25 numbers of nodes, respectively. Compared with the two existing techniques, RTM [32] obtains maximum energy consumption than ANFIS [33] and since the proposed NBSSO has enhanced lifetime for a large time.

F. Discussion

Providing security for MANET is a complicated process. It is necessary for securing the network over intrusions in MANET to assure the development of the service. For this purpose, the proposed NBSSO is introduced. The graphical analysis shows that the existing models like RTM and ANFIS, EPT-WBC and fuzzy detector obtains poor results compared to the proposed NBSSO on PDR, throughput, DR, and Average packet loss ratio. These models attained poor results since they require a huge amount of time to process, and there is no guarantee for authentication of the public key. The proposed model attained better results due to optimal CH selection.

Further, the proposed NBSSO performs the trust of nodes based on authentication. Features of interest are useful to find the kind of attacks that occurred in the node. The proposed classifier WNBC with SSO optimization is used to find the attack patterns based on features of interest. Hence, it is proved that this model can ensure security for the network over harmful intrusions.

V. CONCLUSION

This work motivates an efficient way to detect intrusion behaviour in MANET. Here malicious classification is done with a weighted hybrid naïve Bayes classifier prospered in identifying and separating malicious nodes using NS-2 simulation. Similarly, the proposed work concentrated on attack pattern classification with high efficiency. This work can detect malicious nodes from the mobile nodes before they start dropping data packets. So the network energy is consumed, and it can be processed for a long time. Simulation has been performed on the NS-2 simulation tool; the current detection scheme effectively notices the malicious nodes then avoids misbehaving acts. Thus the results enhance the system performance over throughput, energy consumption, average packet loss ratio, detection rate and packet delivery rate.

REFERENCE

[1] E.A. Shams, and A. Rizaner, A novel support vector machine-based intrusion detection system for mobile ad hoc networks, Wireless Networks, 24(5) (2018) 1821-1829.

[2] K. Murugan, and P. Suresh, Efficient Anomaly Intrusion Detection Using Probabilistic Hybrid Techniques in Wireless Ad Hoc Network, IJ Network Security, 20(4) (2018)730-737.

[3] K.P. Kumar, and B.R. Prasad Babu, A Simple and Cost-Effective Anomaly Detection Paradigm on the Basis of Computational Intelligence for Mobile Ad-Hoc Networks from a Security Viewpoint, In Computer Science On-line Conference, Springer, Cham, (2019) 78-86.

[4] H. Vegda, and N. Modi, Secure and Efficient Approach to Prevent Ad Hoc Network Attacks Using Intrusion Detection System, In Second International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, (2018) 129-133.

[5] R.K. Kovarasan, and M. Rajkumar, An Effective Intrusion Detection System Using Flawless Feature Selection, Outlier Detection and Classification, In Progress in Advanced Computing and Intelligent Engineering, Springer, Singapore, (2019) 203-213.

[6] P. Pandey, and A. Barve, An Energy-Efficient Intrusion Detection System for MANET, In Data, Engineering and Applications, Springer, Singapore, pp.103-117, 2019.

[7] A.M. Desai, and R.H. Jhaveri, Secure routing in mobile ad hoc networks, a predictive approach, International Journal of Information Technology, (2018) 1-12.

[8] G. Vaseer, G. Ghai, and D. Ghai, Novel Intrusion Detection and Prevention for Mobile Ad Hoc Networks, A Single-and Multiattack Case Study, IEEE Consumer Electronics Magazine, 8(3) (2019) 35-39.

[9] T. Poongodi, and M. Karthikeyan, Localized secure routing architecture against cooperative black hole attack in mobile ad hoc networks, Wireless Personal Communications, 90(2) (2016) 1039-1050.

[10] A. Lupia, and FD. Rango, Trust management using probabilistic energy-aware monitoring for intrusion detection in mobile ad-hoc networks, In 2016 Wireless Telecommunications Symposium (WTS) IEEE, (2016) 1-6.

[11] M. Bouhaddi, K. Adi, and M.S. Radjef, Evolutionary game-based defence mechanism in the manets, In Proceedings of the 9th International Conference on Security of Information and Networks, ACM, (2016) 88–95.

[12] T. Spyridopoulos, G. Karanikas, T. Tryfonas, and G. Oikonomou, A game-theoretic defence framework against dos/ddos cyber-attacks, Computers & Security, 38 (2013) 39–50.

[13] B. Subba, S. Biswas, and S. Karmakar, Intrusion detection in Mobile Ad-hoc Networks, Bayesian game formulation, Engineering Science and Technology, an International Journal, 19(2) (2016).782-799.

[14] D. Fudenberg, J. Tirole, Game Theory, Cambridge, Massachusetts. ISBN, 9780262061414. (1991).

[15] A. Nadeem, K. Ahsan, and M. Sarim, Illustration, Detection & Prevention of Sleep Deprivation Anomaly in Mobile Ad Hoc Networks, Mehran University Research Journal of Engineering and Technology, 36(2) (2017) 233-242.

[16] A.A. Korba, M. Nafaa, and Y. Ghamri-Doudane, Anomaly-based intrusion detection system for ad hoc networks, In 2016 7th International Conference on the Network of the Future (NOF), IEEE, (2016) 1-3.

[17] NJ. Patel, and R.H. Jhaveri, Detecting packet dropping nodes using machine learning techniques in Mobile ad-hoc network, A survey, In 2015 International Conference on Signal Processing and Communication Engineering Systems, IEEE, (2015) 468-472.

[18] S. Sankaranarayanan, and G. Murugaboopathi, Secure intrusion detection system in mobile ad hoc networks using RSA algorithm, In 2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM), IEEE, (2017) 354-357.

[19] A. Chaudhary, V.N. Tiwari, and A. Kumar, A cooperative intrusion detection system for sleep deprivation attack using neuro-fuzzy classifier in mobile ad hoc networks, In Computational Intelligence in Data Mining-Volume Springer, New Delhi, 2 (2015) 345-353.

[20] R.S. Krishnan, E.G. Julie, Y.H. Robinson, R. Kumar, T.A. Tuan, and H.V. Long, Modified zone-based intrusion detection system for security enhancement in mobile ad hoc networks, Wireless Networks, 26(2) (2020) 1275-1289, 2020.

[21] D. Singh, and S. Bedi, Trust aware intrusion detection system based on a cluster, International Journal of Computer Applications, 131(7) (2015) 7-13.

[22] N. Veeraiah, and BT. Krishna, An approach for optimal-secure multipath routing and intrusion detection in MANET, Evolutionary Intelligence, (2020) 1-15.

[23] M. Islabudeen, and K. Devi, A smart approach for intrusion detection and prevention system in mobile ad hoc networks against security attacks, Wireless Personal Communications, 112(1) (2020) 193-224.

[24] N. Lal, S. Kumar, A. Saxena, and V.K. Chaurasiya, Detection of malicious node behaviour via I-watchdog protocol in mobile ad hoc network with DSDV routing scheme, Procedia Computer Science, 49 (2015) .264–273.

[25] J.M. Chang, P.C. Tsou, I. Woungang, H.C. Chao, and C.F. Lai, Defending against collaborative attacks by malicious nodes in MANETs, A cooperative bait detection approach. IEEE Syst. J, 9(1) (2015) 65–75.

[26] M. Poongodi, S. Bose, Design of Intrusion Detection and Prevention System (IDPs) using DGSOTFC in collaborative protection networks, In Advanced Computing (ICoAC), Fifth International Conference on IEEE, (2013).

[27] M. Poongodi, S. Bose, N. Ganesh Kumar, The effective intrusion detection system using optimal feature selection algorithm, Int. J. Enterp. Netw. Manag. Forth Comingissue, http,//www.inderscience.com/info/ingeneral/forthcoming.php?jcode =ijenm. (2015).

[28] C. Manikopoulos, S. Papavassiliou, Network intrusion and fault detection, a statistical anomaly approach. Telecommunications network security, IEEE Commun. Mag, vol. 40 no. 10 (2002) 76–82.

[29] E. Aivaloglou, S. Gritzalis, Hybrid trust and reputation management for sensor networks, Wirel. Netw, 16(5) (2010) 1493–1510.

[30] F. Luo, L. Khan, F. Bastani, I.L. Yen, J. Zhou, A dynamically growing self-organizing tree (DGSOT) for hierarchical clustering gene expression profiles, Bioinformatics, 20 (2004) 2605–2617.

[31] J. Dopazo, J. Carazo, Phylogenetic reconstruction using an unsupervised growing neural network that adopts the topology of a phylogenetic tree, J. Mol. Evol, 44 (1997) 226–233.

[32] F. Heylighen, The science of self-organization and adaptivity. In, Kiel, L.D. (ed.) Knowledge Management, Organizational Intelligence and Learning, and Complexity, In The Encyclopedia of Life Support Systems (EOLSS). Eolss Publishers, Oxford. http,//www. eolss.net, (2001).

[33] F. Heylighen, Complexity and self-organization, In Bates, M.J., Maack, M.N. Encyclopedia of Library and Information Sciences, CRC Press, Boca Raton, (2009).

[34] MS. Oswaldo Aguirre, H. Taboada, A clustering method based on dynamic self-organizing trees for post-Pareto optimality analysis, Sciverse Science Direct, Procedia Computer Science, Conference Organized by Missouri University of Science and Technology 2011-Chicago, IL, 6 (2011) 195– 200.

[35] X. Li, F. Zhou, J. Du, LDTS, a lightweight and dependable trust system for clustered wireless sensor networks, In IEEE Transactions on Information Forensics and Security, (2013) 451–551.

[36] F. Bao, I.R. Chen, M.J. Chang, J.H. Cho, Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection, IEEE Trans. Netw. Serv. Manag, 9(2) 169–1832012.

[37] S. Sett, P.K. Thakurta, Effect of optimal cluster head placement in MANET through multi-objective GA, In2015 International Conference on Advances in Computer Engineering and Applications IEEE, (2015) 832-837.

[38] F. Hamza, S.M.C. Vigila, Review of Machine Learning-Based Intrusion Detection Techniques for MANETs, In Peng SL., Dey N., Bundele M. (eds) Computing and Network Sustainability. Lecture Notes in Networks and Systems, Springer, Singapore, 75(1) (2019) 367-374.

[39] F. Hamza & S.M.C. Vigila, Cluster Head Selection Algorithm for MANETs Using Hybrid Particle Swarm Optimization-Genetic Algorithm, International Journal of Computer Networks and Applications (IJCNA), 8(2) (2021) 119-129.

[40] F. Hamza and S.M.C. Vigila, Integrated Elephant Herd Optimization and Reputation Signaling Game for efficient Intrusion Detection in Mobile Ad hoc Networks. Design Engineering, (2021) 12946-12966.

[41] K. Spurthi, T.N.Shankar, An Improved Zone Routing Protocol For Secure And Efficient Energy Management International Journal of Engineering Trends and Technology, 69(1) 29-34.

[42] M.A. Hussain, B. Duraisamy, Review on Packet drop prevention in MANET by counter-based digester ACK International Journal of Engineering Trends and Technology, 68(8) 102-107.

[43] M. Tahboush, M. Agoyi, A. Esaid, Multistage Security Detection in Mobile Ad-Hoc Network (MANET) International Journal of Engineering Trends and Technology, 68(11) 97-104.