IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 21, 2021, accepted January 23, 2021, date of publication January 28, 2021, date of current version February 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3055422

A Recurrent Reward Based Learning Technique for Secure Neighbor Selection in Mobile AD-HOC Networks

K. SAKTHIDASAN SANKARAN[1], (Senior Member, IEEE), N. VASUDEVAN[1], K. R. DEVABALAJI[2], THANIKANTI SUDHAKAR BABU[3], (Senior Member, IEEE), HASSAN HAES ALHELOU[4], (Senior Member, IEEE), AND T. YUVARAJ[5]

[1]Department of ECE, Hindustan Institute of Technology and Science, Chennai 603103, India
[2]Department of EEE, Hindustan Institute of Technology and Science, Chennai 603103, India
[3]Department of Electrical and Electronics Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad 500075, India
[4]Department of Electrical Power Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia 2230, Syria
[5]Department of EEE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 600077, India

Corresponding authors: Hassan Haes Alhelou (alhelou@tishreen.edu.sy) and Thanikanti Sudhakar Babu (sudhakarbabu66@gmail.com)

ABSTRACT Mobile ad-hoc network is an assortment of distinct attribute-based mobile devices that are autonomous and are cooperative in establishing communication. These nodes exploit wireless links for communication that causes injection of the adversaries in the network. Therefore, detection and mitigation of adversaries and anomalies in the network are mandatory to retain its performance. To strengthen this concept, in this article, a novel secure neighbor selection technique using recurrent reward-based learning is introduced. This proposed technique inherits the benefits of conventional routing and intelligent machine learning paradigm for classifying the states of the nodes based on their communication behavior. Thorough learning of the behavior of the nodes unanimously at all the hop-levels of communication enables establishing secure and consistent routing and transmission paths to the destination. The performance of the proposed technique is estimated using the metrics throughput, packet delivery ratio, and delay and detection ratio. Experimental analysis proves the consistency of the proposed technique by improving throughput, packet delivery ratio, and detection ratio under less delay.

INDEX TERMS Attack detection, behavior modeling, machine learning, MANET, reward function.

I. INTRODUCTION

Research focus over mobile ad-hoc networks (MANETs) has increased significantly in the present years due to its on-demand communication and infrastructure-less configuration abilities. MANETs consist of mobile self-disciplined mobile nodes (such as electronic devices with a communication unit, mobile phones, wireless sensors, etc.) that interconnect using intrinsic routing ability. For communication and packet data exchange, the nodes establish wireless links by discovering reliable neighbors. The nodes are capable of transmitting data packets to the destinations located away from its radio range. In such cases, to scale the physical transmission distance, the sender banks on intermediate nodes. The intermediary nodes are the same as that of the sender act as a hub/ relay to forwarding packets. The intermediary node receives the information from the sender and forwards it to the neighbor connecting the destination. The source forms transmission routes using the discovered neighbors. Neighbor discovery, path monitoring and preservation, and path expurgation process are decided using the network layer protocol employed by the sender. Due to infrastructure-less support, ease of deployment, and technological interoperability, MANET finds its application in disaster recovery networks, emergency scenarios, commercial and residential applications, defense, etc. [1], [2].

MANET nodes face different challenges due to their resource-restricted access and utilization, including security and privacy issues due to the wireless communication medium. Wireless medium provides open access to other networks and users for interoperability and data sharing. The open-access nature of the medium is exposed to security risks where an intruder breaches the communication of the nodes [3]. Network dynamicity with time, node mobility,

The associate editor coordinating the review of this manuscript and approving it for publication was Qiuye Sun.

and lack of centralized administrative support are some of the issues that permit intruder or adversary to breach the communications of the network. A malicious user or node that enters the communication between the nodes gains information about the network. It then launches its attack to restrict resource availability, route failure, packet drop, spoofing, etc. [4]. The type of attack depends on the nature and purpose of the adversary fissuring into the network. However, this results in network outcome vitiating, unnecessary resource exploitation, and privacy of the users. Therefore, cooperative routing and attack mitigation systems are designed for encountering such attacks in MANET [5].

Graph-theory-based network partitioning techniques were presented so far for realizing the detection of decentralized recognition in the fast response speed on using the characteristics highlights of power flow in the independency of various groups [6]. One such application aims at employing deployed sensors in the pipeline network, the data driven detection problem occurred by failure of device or the interruption of network which in turn hinders the implementation of pipeline monitoring status. This issue could be solved so as to evade data leakage [7].Mitigating the effects of adversaries and anomalies in MANET requires optimal detection and mitigation systems. With the development in decision making and artificial intelligence computing model, decision-based detection and performance optimization methods are commonly employed in MANETs. These systems rely on the input observed from the characteristics and behavior of the nodes for accurate decision-making. The events in route discovery and communication are modeled based on the decisions performed by the computing systems. Decision making and intelligent computing algorithms aid the optimization process by detecting the adversary based on the reputation and trusts it has assessed from the current and past behavior of the nodes. Some decision making systems incorporate location and sensing attributes, require the support of external networks such as cloud for providing reliable communication in MANET [8], [9]. The contributions in this article are listed as follows:

i) Designing a secure neighbor selection technique that exploits recurrent reward for discovering optimal path nodes to provide end-to-end communication reliability.
ii) Improving the security in MANET communication by classifying the nodes based on their states as modeled by the behavior and reward. The reward is estimated using a recurrent machine learning algorithm for determining the reliability of the neighbor.
iii) Performing a comparative analysis of the proposed method with the existing techniques to prove its consistency.

II. RELATED WORKS

Li *et al.* [10] introduced diagnosing anomalies with provenance and verification (DAPV method) for improving the routing efficiency of MANETs. DAPV is legitimate in detecting direct and indirect routing adversaries. For this purpose, DAPV utilizes peer logs and Merkle based hash trees for secure neighbor detection and authentication. This method achieves better detection and time cost under average overhead.

Evolutionary self-cooperative trust (ESCT) is introduced by Cai *et al.* [11] for mitigating routing distraction in MANETs. This secure routing scheme replicates the human intervention process and estimates the trust of the path nodes at different levels. The reliable characteristic of the scheme is its uncompromised ability towards known attacks. The decision making is novel at each level of attack detection. This scheme achieves reliable network outcome measured using energy utilization, delay,and packet delivery ratio.

Vaseer *et al.* [12] proposed a behavior-based analysis algorithm as a countermeasure for different kinds of attacks. This algorithm is scalable in defending against denial of service, vampire, and user-to-root attacks. The different attacks are identified based on the behavior and response of the nodes to its neighbor. This response ensures the capability of the nodes for handling data packets and forwarding them. With the help of conventional ad-hoc routing and layer 4 transport protocols, the authors have proved the consistency of the algorithm verified using delay and packet delivery ratio.

A trust-based multi-objective optimization (MOO) is introduced by Wang *et al.* [13] for improving the service outcome of MANETs. This optimization aids multi-dimensional trust in assessing node behavior at the time of network resource allocation and communication. Specifically, this trust optimization method achieves less service cost and improves the service quality in the communications between nodes and service providers.

Zhang *et al.* [14] exploited communication, energy, and recommendation based trust evaluation for improving its adaptability and distributed nature in the MANET scenario. The adaptive nature of the trust model is estimated using the packet transmission rate, reserving energy, distance, familiarity, and reliability of the nodes. The trust model achieves better one-to-one validation of the metrics, reducing the impact of attackers. This helps to verify the reliability of the node for the successive transmission based on mean trust. This helps to improve the rate of detection by precisely estimating the trust of the nodes.

Keerthika and Malarvizhi [15] introduced a hybrid weighted trust-based artificial bee colony 2-opt algorithm for evading black hole attacks in MANET. The artificial bee colony algorithm is exploited for discovering the available paths to the destination, and the 2-opt algorithm estimates the trust of the nodes by evaluating fitness function. These two algorithms are integrated to provide a hybrid trust model that reduces the distance to the sink and delay with better packet delivery.

MASID-R-SA is an optimized intrusion response system designed by Mechtri *et al.* [16] as a measure of evading attack-severity. The intensity and nature of the detected adversaries in the routing path are reported effectively to the predecessor and source node to provide better routing choices. This helps to reduce the false alarms in the network irrespective of single or multi-attack MANET scenarios. By confining the false detection rate, this method achieves better packet delivery and less delay.

Demidov *et al.* [17] discussed the issues in challenges in detection and mitigating cyber security threats in ad-hoc wireless networks. The authors have incorporated a neural network for analyzing the challenges in ad-hoc networks with better approximation. This neural network training relies on likelihood maximization and recurrent graph analysis models for improving the security solutions for ad-hoc networks.

Moudni *et al.* [18] assimilated the adaptive neuro-fuzzy interference system (ANFIS) and particle swarm optimization (PSO) for detecting and mitigating black hole threats in MANET routing. Fuzzy is used to frame rules based on packet forwarding ratio and destination concentric sequence number. The satisfied rules are operated using PSO for selecting conditional routes that satisfy the rules. Nodes violating the rules are discarded from the transmission process, and hence the combined method suppresses false alarm and improves detection.

Efficient stream region sink position analysis (ESRSPA) is modeled by Vignesh and Santhosh [19] for improving the attack detection in MANET routing process. In the stream region analysis, the presence of attack is detected by fetching sink location information and other legitimate nodes. Based on the position of the sink, the transmission path is determined. The nodes that fail to satisfy weight constraints in each stream are classified as attackers. The routes with such nodes are discarded. This method is reliable by improving network throughput and packet delivery under controlled overhead.

Zhang *et al.* [20] projected a reputation management system for improving the resilience of MANETs against attacks. Network efficiency is levitated by adapting reputed neighbors for route discovery and communication. The reputation of the nodes is individually collected, and then the cumulative path value is determined to select transmission paths. The reputation of the nodes is estimated using subjective and recommendation based trust.

Sakthidasan *et al.* [21] proposed an optimal route selection technique for wireless sensor networks (WSNs) using whale optimization algorithm. This route selection technique integrates optimization and affinity propagation for identifying transmission paths. In particular, this route selection technique is designed for clustered WSNs in a time-critical scenario for energy optimized transmissions.

Sun *et al.* [23] presented a novel controller that were distributed coordinated integrated with a multi-agent-dependent consent algorithm that is employed for the distribution generations in the internet of energy. After that, the tasks that were decomposed, flow of information, models of the presented approach were analyzed. Huang *et al.* [24] investigated the problem of economic dispatch in the distributed fashions of microgrid. For addressing this issues, the delay free dependent distributed algorithm was projected for assigning the entire energy demand optimally along with the intention of reducing the agminated operations cost.

Wang *et al.* [25] suggested line inductance that was stability based domain assessment technique for the weak grids along with CPLs. Initially, the matrix of source impedance and the matrix of load admittance are built separately. Li *et al.* [26] established a model of double-mode energy management intended for the system of multi-energy. This is formed by several bodies of energy. By such model, all participants were capable of responding adaptively to the switching mode change. Besides, a distributed dynamic novel event-triggered Newton-Raphson process was suggested for solving the issue of energy management at the fashion of fully distributed.

Yushuai *et al.* [27] investigated the issue of cooperative distributed energy management of the several multi-bodies along with the optimal energy consideration or generation of entire participant in the single bodies and the optimized distribution of energy at the lines that were interconnected among any pairs of energy bodies.The novelty of the proposed SNS-RR is that it classifies the state of the nodes on the basis of their reward and transition. This state is of the nodes is not estimated on the other existing methods such as [20], [15], & [16]. In another method, such as [13], [18], [21] and [22], optimization is used as an external measure. In this case, routing is considered a whole paradigm in discovering routes to the destination. Different from these approaches, the proposed method analyses the independent nodes for their transition state using multiple metrics associated with the node. The proposed technique estimates the reward using a recurrent machine learning algorithm for determining the reliability of the neighbor. In a machine learning process, the input fetched is recursively analyzed for all the possible combinations of output and communication features such that the best solution is achieved often. The smart computational and operational feature of the learning algorithm is exploited for complex input processing with ease. With a reward function, this is retained for a prolonged time to select a mere optimal node in the routing process.

III. SECURE NEIGHBOR SELECTION (SNS) USING RECURRENT REWARD (RR) FUNCTION

Administering security in autonomous MANET is a challenging task due to varying node characteristics. The cooperative communication of nodes and transmitting the nature of the neighbors conclude its reliability. The presence of attackers in the routing path degrades MANET performance, and therefore the source needs to select robust and reliable neighbors for packet handling. As the characteristics of the neighbor changes with time and communication, seamless analysis of their attributes is essential to determine its reliability. The following sections describe the proposed neighbor selection method with its working functions. The design goal of the proposed SNS is to improve the network throughput by mitigating the impacting attackers in the routing path. Different from the conventional routing procedures, multi-metric analysis is adapted in this method. With the help of multi-metric analysis the favorable transmission condition is identified for the available node. This helps to improve the availability of the nodes and transmission routes throughout the data handling process. Therefore, the change in neighbor characteristics does not impact the transmission resulting in convergence.

A. NETWORK REPRSENTATION

We employ a MANET consisting of N mobile nodes communicating using wireless edges called vertex (V). The nodes interconnect in a random-mesh topology forming a graph $G = (N, V)$. The communication/coverage range R and distance d between the nodes decide the availability of V. A typical MANET scenario is portrayed in Figure 1.

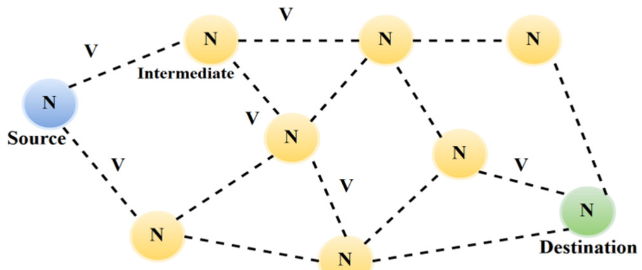


FIGURE 1. **Typical MANET scenario.**

The source transmits packets to the destination through the neighbors that form the routing path. The source and destination are located far away from each other, and therefore, multi-hop transmission is adopted. We assume that the nodes are mobile under random mobility, and this owns a less impact over network functions. As that article is concentrating on secure neighbor selecting, the attack mode is not specified in this context.

IV. PROPOSED METHODOLOGY

The proposed SNS with RR is featured by integrating machine learning (ML) in MANET operations. Ad-hoc routing protocols that are familiar in the network layer discover shortest distance nodes to form routes to the destination. In this SNS, the route formation and neighbor selection are assisted using multi-metric node attributes other than distance. The multi-attribute evaluation increases the reliability of the selected path by incorporating efficient neighbors. Therefore, the proposed method is divided into three phases, namely: State determination, multi-attribute analysis, and route selection. In the state dissemination, the transmitting and receiving functions of the nodes are relied. Based on this, the availability of the node is intimated to the other neighbors' in-range. This state determination helps to prevent contamination of nodes from the attackers in data handling. In the multi-attribute analysis, the favorable transmission condition associated with the nodes is identified for easing transmission. Finally, the available nodes that are classified under the attribute analysis are selected for forming routes to the destination and hence the transmission is performed. The machine learning for neighbor selection relies on the attributes analyzed.

A. STATE DETERMINATION

The state of a node is determined based on its network functions. The nodes perform transmitting/receiving and idle functions, and the state is thus given as {TR, I}. An idle state node has a chance of being malicious, and so, the states are expanded as {TR, I, M} where M represents the malicious node. The chance of a node to be malicious relies on TR and I states of the node (i.e.) ρ (M) = {ρ (TR) ∪ ρ (I)} ∩ρ(M). The probability of a node to be neighbor ρ_n is given by

$$\rho_n = \rho(TR) \cup \rho(I) < \rho(M) \tag{1}$$

The probability featuring in equation (1) relies on the previous TR and I states of a node. It is mandatory to estimate the node's performance in these states to verify its consistency. It is obvious that the routing protocol selects a node in TR state for transmitting packets. Therefore, we define a state consistency set independently for the nodes.

In equation (2), the state representative for the set of nodes is given. Here, the conditions in I (i.e.) i ≠ j represents a node cannot be present in both TR and I at the same transmitting time (t_{tx}).

$$\left.\begin{array}{ll} TR = \{n_i\}, & i \in N \\ I = \{n_j\}, & j \in N, i \neq j \\ M = \{n_k\}, & k \in N, k = i || k = j \end{array}\right\} \tag{2}$$

Contrarily, a node in state M belongs to the one present in TR or I. Therefore, the chances for selecting a neighbor is modified from equation (1) as

$$\rho_n = \rho(TR) \wedge \rho(M) || \rho(I) \wedge \rho(M) || [\rho(TR) \cup \rho(I)] \wedge \rho(M) \quad, \forall n \in N \tag{3}$$

The above equation is a logical representation of a node states such that the malicious states are least required with TR, and I states. On the other hand, the chances of I state other than TR is most required over the malicious state. This condition is estimated for all the n nodes in the set N.

The nodes satisfying equation (3) is therefore considered for transmitting packets. The evaluation of multiple attributes associated with the node communication confirms the state of the node. A general state representation with switch-over probability is illustrated in Figure 2.

A node from TR to I and vice-versa is a conventional process depending on their transmitting and idle time (t_{id}). Instead, a node if swapped to M state is fixed with t_{tx}=t_{id}= 0, from either TR or I state correspondingly. This means that the node in M cannot be moved to TR or I state as a security measure.

B. MULTI-ATTRIBUTE ANALYSIS

In conventional node attribute analysis, direct and indirect reputation is assessed. This reputation relies on the packet probing and receiving attribute of the sender and receiver

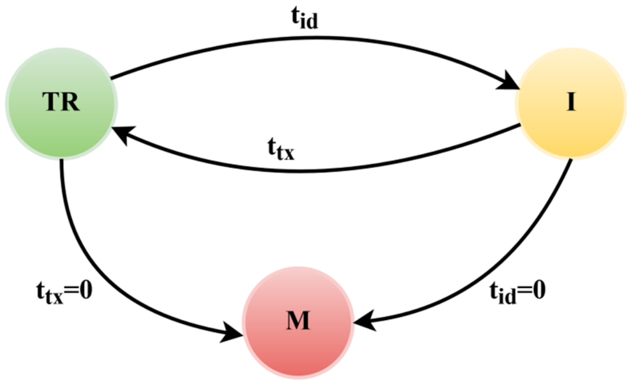


FIGURE 2. State switch-over.

nodes. For preventing indirect reputation/ attribute malfunction, direct reliability estimating methods are employed. Let $\nabla\tau$ be the reliability value of the node that is competition for neighbor selection. We consider the multi-attribute factors such as packet dispatching rate (p_{dr}), transmission quality (q_{tx}) and node availability (a_n). Here, the factors p_{dr} and a_n are estimated for a direct neighbor and q_{tx} is estimated for the path through the neighbor $n \in N$. The predecessor node estimates these attributes at the time of selection. The autonomous nature of the nodes is preserved, such that each node is responsible for selecting its own forwarding neighbor. A node is analyzed independently for its reward, and the decisions for selecting it relies on its predecessor. This helps to retain the distributive nature of the nodes. Equation (4) presents the estimation of the fore-mentione4d factors:

$$\left.\begin{aligned} p_{dr} &= \frac{p_{tr}}{p_{rx}} \\ q_{tx} &= hop \times C_{hx} + (1-d) \times C_h \\ a_n &= p_{dr} \times t_{tr} \end{aligned}\right\} \tag{4}$$

where, p_{tr} and p_{rx} are the packets transmitted and received by the neighbor n, C_{hx} is the number of HELLO control messages generated for 'x' transmission and C_h is the total HELLO messages generated. Unit d represents the distance between the nodes.The attribute a_n relies on the transmitting and receiving time of the neighbor. The availability of the node is accounted for until the successful transmitting time of the node. With the estimating of individual attributes, the reliability value of a node is computed using equation (5) as

$$\tau = w_1 \times a_n + w_2 \times p_{dr} + w_3 \times q_{tx} \tag{5}$$

where w_1, w_2 and w_3 are the adjustable weights for node availability, p_{dr} and q_{tx} respectively. The value of w_2 and $w_3 \in [0, 1]$ (i.e.)$w_2+w_3 = 1$ whereas, $w_1= 0$ or $w_1 = 1$ and therefore, if a_n is true then $w_1 = 1$ or 0 otherwise. The rate of change in node attribute relies on p_{dr} and q_{tx} for all $w_1 = 1$. If $w_1 = 1$, then the node is in I or M state. The attributes are analyzed for four state changeover of a node, namely:

- TR to I
- I to TR
- TR to M and
- I to M

These four changeover conditions are assessedbased onthe reward assigned. Here, $\Delta\tau$ is assessed as a reward for the above change-overs. The reward $Q(\Delta\tau)$ is estimated in $\{t_{tx}, t_{tx+i}, \ldots .t_{tx+x}\}$ for I to TR and in $\{t_{id}t_{id+1}, \ldots ,t_{id+x-1}$ for I to TR switch-over provided $t_{tx}\neq,t_{id}$. The reward is computed using equation (6)

$$\begin{aligned} Q(\Delta\tau, TR) &= Q(\Delta\tau, TR) + \beta\left[x_i(\Delta\tau, TR) - Q(\Delta\tau, TR)\right], \\ & \quad i = \{t_{tx}, t_{tx+i}, \ldots .t_{tx+x}\} \\ Q(\Delta\tau, TR) &= Q(\Delta\tau, TR) + \beta\left[x_i(\Delta\tau, TR) - Q(\Delta\tau, TR)\right], \\ & \quad i = \{t_{tx}, t_{tx+i}, \ldots .t_{tx+x}\} \end{aligned} \tag{6}$$

The unit β is the reward factor, and it lies between 0 and 1. In the estimation of reward for I, the reward that is achieved in the previous TR state is equated with verifying the difference in p_{dr} or q_{tx}. The reward is estimated for change-overs (i) and (ii). The cases (iii) and (IV) tend to a malicious state wherein we define a threshold for TR and I states to verify node reliability. The state threshold S_{th} is defined using equation (7)

$$S_{th} = \begin{cases} 1 - \dfrac{p_{dr}}{q_{tx}}, & \forall n \in TR \\ 1 - \dfrac{hop}{q_{tx}}, & \forall n \in I \end{cases} \tag{7}$$

If $Q(\Delta\tau, TR||I) > S_{th}$, the changeover in case (i) and (ii) is permitted else, the node is added to state M from TR and I respectively. The state representation for equation (6) in accordance with equation (7) is illustrated In Figure 3(a) The conditions for the case (i) and case (ii) are tabulated in Table 1.

TABLE 1. Conditions and Changeover for Case (i) and (ii)

Reward	Condition 1	Condition 2	State Change-Over
$Q(\Delta\tau, TR)$	$> S_{th}$	$w_1 = 1$	Remains in *TR*
	$> S_{th}$	$w_1 = 0$	*TR* to *I*
	$< S_{th}$	$w_1 = 0$ *or* 1	*TR* to *M*
$Q(\Delta\tau, I)$	$> S_{th}$	$w_1 = 1$	*I* to *TR*
	$< S_{th}$	$w_1 = 0$ *or* 1	*I* to *M*
	$> S_{th}$	$w_1 = 1$	Remain to *I*

The process of reward estimation is recurrent to ensure optimal neighbor selection in any t_{tx}. The recurrent learning of node attribute is segregated between TR and I. The nodes that are moved to M state from TR or I are not recurrently analyzed. Similarly, the case (i) and (ii) are analyzed only if $w_1 = 1$ for all n in either TR or I. Depending on the recurrent analysis process, the selection of the node in the routing path is determined. From the analysis of case (i) and (ii), neighbors satisfying the condition in Table 1 with $w_1 = 1$ are initially selected for handling packets. In Table 2, the number of nodes satisfying both conditions 1 and 2 with respect to the number of recurrent iterates is presented.

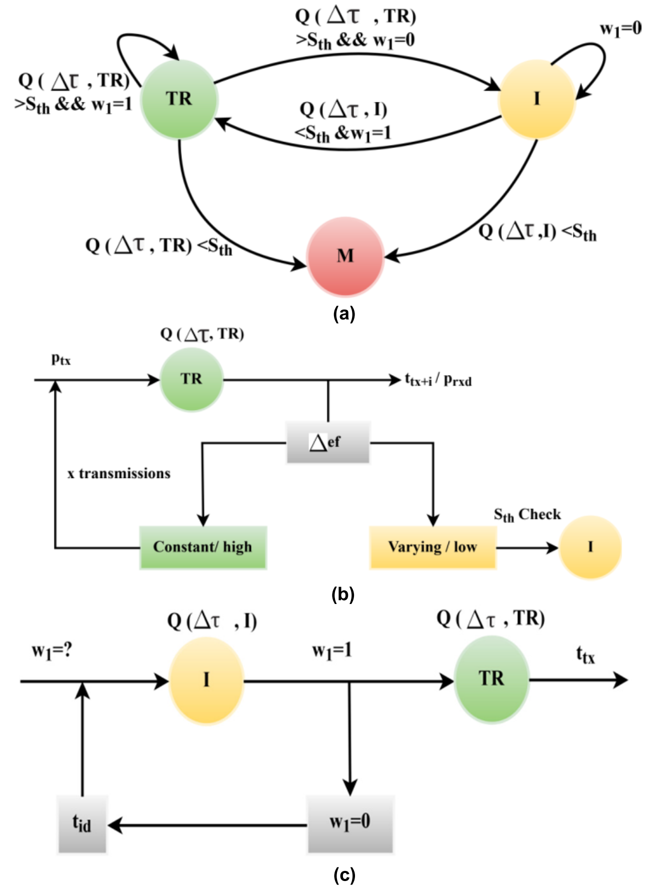


FIGURE 3. (a) State representation for case (i) and (ii). (b) Recurrent reward learning for TR state. (c) Recurrent reward learning for I state.

TABLE 2. Nodes Satisfying Condition 1 and 2 in Different Iterates

Iteration	Condition 1	State	Condition 2	State
10	47	TR	33	I to TR
20	29	TR	64	I
30	54	TR to I	40	I
40	54	TR to I	30	I to TR
50	60	TR	24	I
60	44	TR	39	I to TR
70	38	TR	51	I
80	62	TR to I	17	I
90	43	TR	6	I to M
100	4	TR to M	13	I to TR

C. ROUTE SELECTION

The reliability factor of a node does not remain unchanged as observed in different time intervals of $\{t_{tx}, t_{tx+1}, \ldots .t_{tx+x}\}$. From equation (4), q_{tx} and a_n relay on p_{dr} of a node such that the p_{tx} decides the reliability of the neighbor and the path. The attributes of the route nodes are analyzed to ensure the reliability of the neighbor and the transmission route. The communication attributes such as packet transmission, delay, etc. are analyzed by the predecessor node. The predecessor node estimates the transition state, reward of the one-hop direct nodes in the network. Based on the obtained attributes, it estimates the selection of the neighbors in the network for a new communication path. Variation in neighbor attributes is effectively monitored and assessed using a recurrent learning method. For ensuring better packets delivery from each of the intermediate, the packet received rate at the destination (p_{rxd}) is computed using equation (8)

$$p_{rxd} = \prod_{i=1}^{hop} 1\left[-\left(\frac{p_{txi}-p_{rxi}}{p_{txi}}\right)\right] \tag{8}$$

In the above equation, p_{txi} and p_{rxi} represents the packet transmitted and received at the ith mediate hop. The above equation is used to verify the rate of packets delivered at the destination at each t_{tx} period. Now, the sum of p_{tx} at each intermediate node is equated with equation (8) as

$$\nabla ef = p_{rxd} = \sum_{i=1}^{hop} p_{rxi} \tag{9}$$

where ∇ef is an equity factor. The factor determines the state of the neighboring node estimated for the currently employed path nodes. The estimation of ∇ef is continuous at $\{t_{tx}, t_{tx+i}, \ldots .t_{tx+x}\}$ intervals to verify if there is a change in p_{rxd} or p_{rx}. The process of recurrent learning determines $Q(\Delta\tau, TR)$ of the path node. This process is illustrated in Figure 3(b).

The path nodes with their mean reward value in different iterates, as determined in theTR state is tabulated in Table 3.

TABLE 3. Observed Path Nodes and Mean Reward in TR State

Iterates	Path Nodes	Mean Reward
20	29	0.84
40	30	0.76
60	83	0.712
80	0	0.147
100	13	0.893

The learning constraint is different for a I state node that is illustrated in Figure 3 (c).

The w_1 weight for a_n determines the selection and rejection of a node in the source transmission path. The node satisfying $w_1 = 1$ condition adapts routing and packet handling. Now, the reward of the node is estimated at t_{tx} intervals are satisfying the probabilities in equation (1) and (2). This process is considered for all the nodes along the routing path. The case (ii) and case (iv) nodes as classified in the previous section are discarded eventhoughtheir $w_1 = 1$. This helps to prevent unnecessary verification of routes that increase routing complexity. Besides, selecting nodes satisfying the opportunistic conditions in Table 1 helps to ensure better packet delivery with reduced delay and drop. The path nodes with their mean reward value in different iterates as determined in I state are tabulated in Table 4.

The starred value in Table 3 represents the last retained value by the nodes before entering the idle state. Figure 4 represents the ratio of TR and I path nodes with respect to the iterates 20, 40, 60, 80, and 100.

From the above route selection process, the following Table 5 presents the identified nodes in the TR state and the selected routes to the destination.

TABLE 4. Observed Path Nodes and Mean Reward in I State

Iterates	Path Nodes	Mean Reward
20	64	0.74
40	54	0.82
60	0	0.609*
80	79	0.844
100	0	0.748*

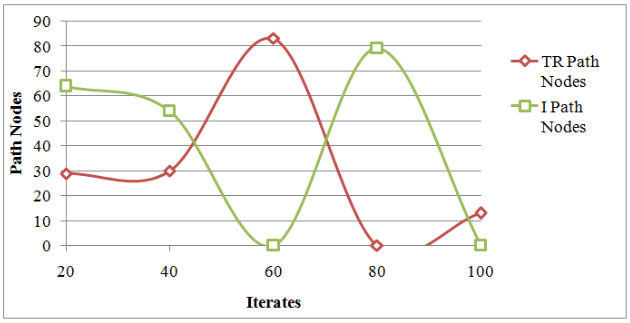


FIGURE 4. Comparison of path nodes in TR and I state as observed in the reward method.

TABLE 5. Nodes in TR State and Routes

Iterates	TR Nodes	Routes	Mean Reward
10	40	6	0.813
20	29	4	0.84
30	17	2	0.933
40	30	4	0.76
50	54	7	0.672
60	83	6	0.712
70	11	1	0.877
80	0	0	0.147
90	24	3	0.744
100	13	1	0.893

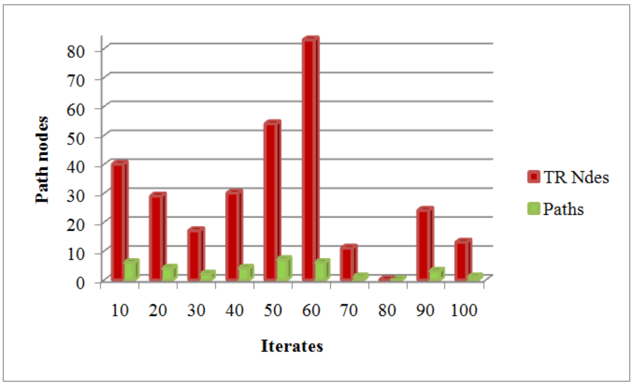


FIGURE 5. TR nodes and available routes.

Figure 5 illustrates the TR nodes and the number of available routes in iteration from 10 to 100.

V. PERFORMANCE ANALYSIS

The Performance of the proposed SNS-RR is evaluated using network simulator experiments (NS2).We define a MANET scenario with 100 nodes, distributed in a 1000 × 500m^2 network region.The network is packed with 100 mobile nodes that are dispersed randomly. The movement of the nodes is randomly manipulated to follow different vector and magnitude across the network region. Each node is equipped with a transceiver with a communication range of 50m. The nodes lying within this range are said to be direct neighbors, and the rest are assessed using multiple hops. The routing protocol operates from the network layer, where the data is transferred using a constant bit rate (CBR) application. Table 6 presents the detailed simulation parameters and their values.

TABLE 6. Simulation Parameters and Values

Simulation Parameter	Value
Network Area	1000 × 500m^2
MANET Nodes	100
Mobility	Random Way Point
Transmission Range	50m
Packet Length	512bytes
Traffic Class	Constant Bit Rate
Pause Time	10ms

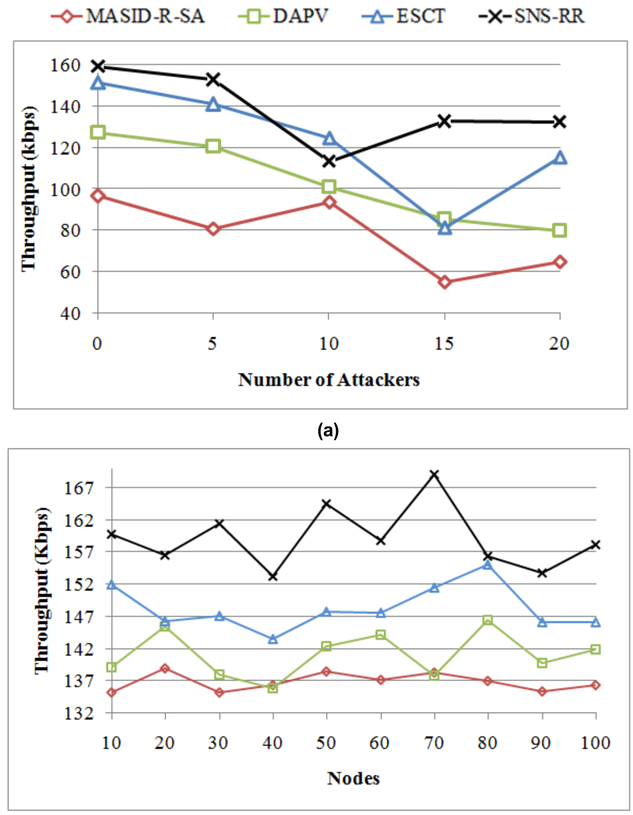
The proposed SNS-RR is compared with the existing MASID-I-RA [7], DAPV [1], and ECST [2] for comparative analysis. In comparison, we consider throughput, packet delivery ratio, and delay, and detection ratio for analysis. In this comparative analysis, the nodes, hop count, and attack % are varied to validate the consistency of the proposed SNS.

A. THROUGHPUT COMPARISON

Figure 6(a) and 6(b) illustrates the throughput comparison between the existing and proposed methods for attackers and nodes. Throughput fluctuates with the rate of change of attackers. The state of the nodes is classified based on its reliability value. The reliability determines the state of the nodes using a reward-based learning method to ensure secure neighbor selection. The nodes that are present in the TR state have opted for transmission. To streamline better transmission through the selected nodes, p_{rxd} the metric is estimated for improving packet delivery. The role of malicious nodes/attackers in the selected path is evaded by moving them to M state. The nodes in the M state are not opted for transmission, preventing malicious nodes from participating in packet forwarding. The factor p_{dr} estimated in equation (4) determines the packet delivery at each hop retaining a_n of the attackers.

B. PACKET DELIVERY RATIO COMPARISON

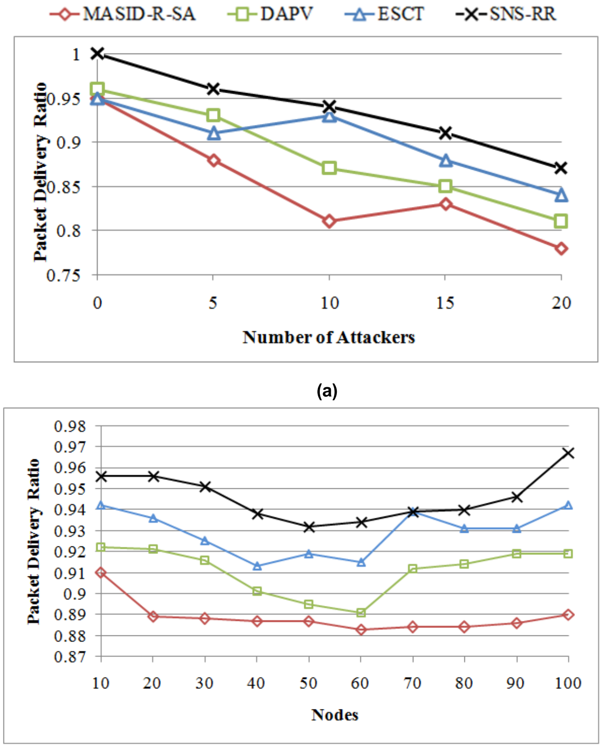
A Comparative study of packet delivery ratio between the proposed SNS-RR and the existing methods is presented in Figure 7(a) and 7(b) for attackers and nodes. There are two factors considered for improving the delivery ration more specifically. The p_{dr} and TR state of the node determines its position in the transmission route. As mentioned in the throughput enhancement, the nodes at each hop ensure maximum packet delivery. The equality factor computed using



(a)

(b)

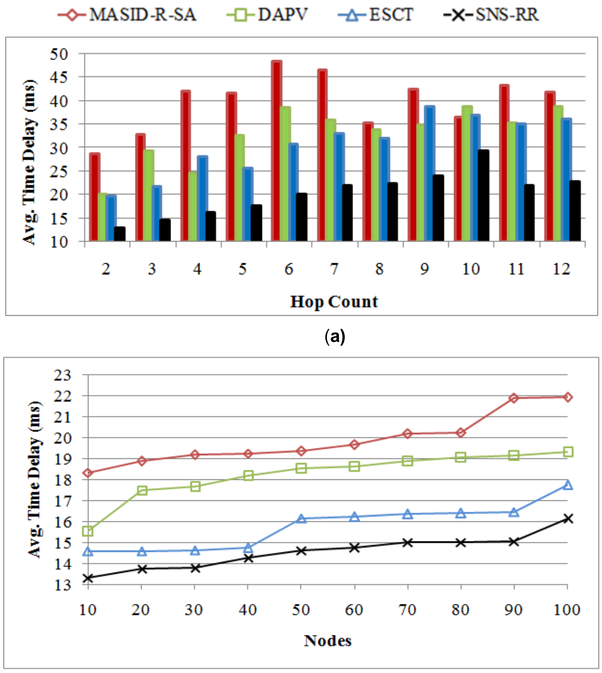
FIGURE 6. **(a) Throughput vs attackers. (b) Throughput vs nodes.**



(a)

(b)

FIGURE 7. **(a) Packet Delivery Ratio vs Attackers. (b) Packet delivery ratio vs nodes.**



(a)

(b)

FIGURE 8. **(a) Avg. delay vs hop count. (b) Avg. delay vs nodes.**

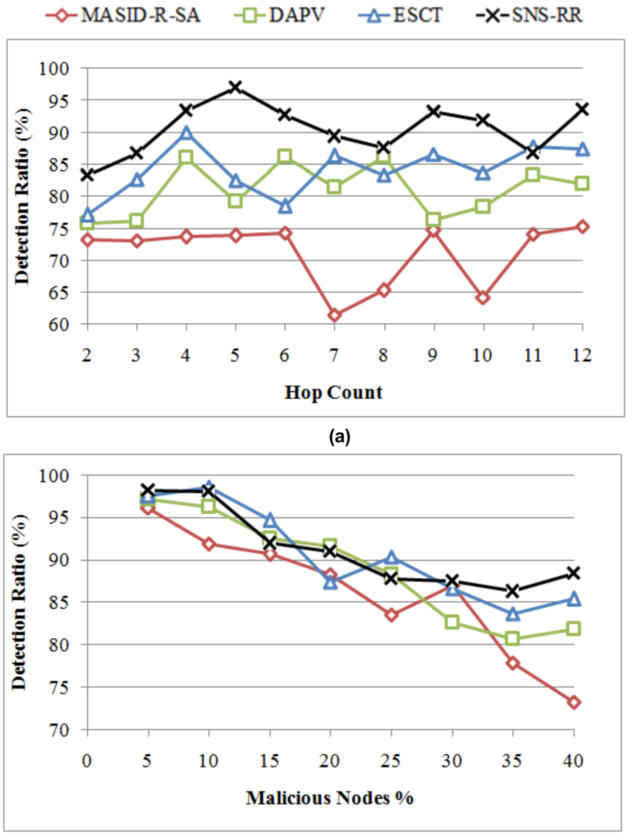
equation (9) verifies the balanced packet delivery at the destination.

C. AVERAGE DELAY COMPARISON

The delay experienced in the network varies with the hop count. An increase in hop count increases the injection rate of adversaries in the routing path. In the proposed SNR-RR, the routing path is established using nodes in TR and I states. Besides, the recurrent reward-based learning designed based on metrics defined in equation (4) and on $\nabla\tau$ improves the successful packet transmission rate. This means the rate of packet drop or paused transmission is avoided in the network. The time spent on packet re-transmission and frequent neighbor selection is restricted based on $Q(\nabla\tau, \mathrm{TR})$ and $Q(\nabla\tau, \mathrm{I})$ in all the hops for each t_{tx}. Therefore the average time difference between source and destination (i.e.) $(t_{tx} - t_{rx})/(\text{hop} \times \text{nodes})$ is less in the proposed SNR-RR. The learning process ensures reliable neighbors to participate in the routing and transmission process, reducing delay. [Refer to Figures 8(a) and 8(b)].

D. DETECTION RATIO COMPARISONS

Figure 9(a) and (b) portrays the detection ratio of the malicious nodes with respect to hop-count and malicious nodes ratio. To classify the nodes, SNR-RR makes use of three states. The nodes that fail the $Q(\nabla\tau, TR) < S_{th}$ or $Q(\nabla\tau, I) < S_{th}$ are moved to M state, and hence these nodes are restricted from participating in routing and transmission process. Moreover, the probability factors described in equation (1) and (3) are mandatory in determining the legitimacy and participation of the node in routing and



(a)

(b)

FIGURE 9. (a) Detection ratio with respect to hop count. (b) Detection ratio for malicious nodes %.

TABLE 7. (a) Comparative Analysis Summary for Different Attackers. (b) Comparative Analysis Summary for Hop Count. (c) Comparative Analysis Summary for Nodes

(a)

Metrics	MASID-R-SA	DAPV	ESCT	SNS-RR
Throughput (Kbps)	64.96	79.62	115.2	132.21
Packet Delivery Ratio	0.78	0.81	0.84	0.87

(b)

Metrics	MASID-R-SA	DAPV	ESCT	SNS-RR
Detection Ratio(%)	75.33	81.96	87.37	93.62
Avg. Time Delay (ms)	41.74	38.62	35.94	22.78

(c)

Metrics	MASID-R-SA	DAPV	ESCT	SNS-RR
Throughput (Kbps)	136.29	141.93	146.1	158.17
Packet Delivery Ratio	0.89	0.919	0.942	0.967
Avg. Time Delay (ms)	21.896	19.305	17.735	16.135

transmission. There are two more conditions to neglect a node (i.e.) minimum $\{p_{dr}\}$, minimum $\{q_{tr}\}$ and minimum $\{\nabla ef\}$. This process is repeated throughout the available hops in the source designed path. The reward-based learning process ensures the $\nabla\tau$ of the path nodes using the factors described in equation (4). Therefore the differentiation of ${}^{\mathrm{M}}/_{(\mathrm{TR}+\mathrm{I})}$ is less as categorized by the states and reward learning. Therefore the rate of detection in both malicious density and hop count is high in the proposed SNS-RR. In Tables7(a), 7(b), and 7(c), the comparative analysis of the above discussion is summarized.

The proposed method is found to maximize throughout and packet delivery ratio by 33.75% and 18%, respectively. From the summary in Table 7(b), the proposed SNS-RR is found to achieve 12.07% high detection ratio and 14.11% less delay.

The proposed SNS-RR is found to maximize throughput and packet delivery ratio by 31.73% and 15% respectively and reduces the time delay by 17.87%.

VI. CONCLUSION

In this article, we introduce a secure neighbor selection (SNS) based on the recurrent reward (RR) machine learning process. In this proposed neighbor selection, the states of the nodes are pre-classified, and then the reward of the neighbors is estimated based on their communication characteristics. The estimated reward is recursively analyzed for selecting optimal path nodes throughout the transmission. The classified nodes in all the hop-levels to the destination are verified for the unanimous criterion for improving the throughput and packet delivery ratio of the network by improving the anomaly detection ratio and suppressing delay. The proposed SNS-RR is found to achieve 33.75% and 31.73% high throughput and 18% and 15% high packet delivery ratio for different attackers and nodes. Similarly, the proposed method reduces average delay by 14.11% and 17.87% for the different hop count and nodes.

In the future, the proposed work is planned to incorporate decision-making methods for analyzing the attributes and information of the front-end application for better security establishments. The significance of mobile computing and swift neighbor detection for real-time application support is planned to be integrated along with the decision-making process. This would help to improve the scalability and application-specific node selection.

Further, in the future, the proposed work is planned to incorporate decision-making methods for analyzing the attributes and information of the front-end application for better security establishments. The significance of mobile computing and swift neighbor detection for real-time application support is planned to be integrated along with the decision-making process. This would help to improve the scalability and application-specific node selection.

REFERENCES

[1] P. Bellavista, G. Cardone, A. Corradi, and L. Foschini, "Convergence of MANET and WSN in IoT urban scenarios," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3558–3567, Oct. 2013.

[2] G. Liu, Z. Yan, and W. Pedrycz, "Data collection for attack detection and security measurement in mobile ad hoc networks: A survey," *J. Netw. Comput. Appl.*, vol. 105, pp. 105–122, Mar. 2018.

[3] Z. Ali Zardari, J. He, N. Zhu, K. Mohammadani, M. Pathan, M. Hussain, and M. Memon, "A dual attack detection technique to identify black and gray hole attacks using an intrusion detection system and a connected dominating set in MANETs," *Future Internet*, vol. 11, no. 3, p. 61, Mar. 2019.

[4] A. Yasin and M. Abu Zant, "Detecting and isolating black-hole attacks in MANET using timer based baited technique," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–10, Sep. 2018.

[5] M. Rmayti, R. Khatoun, Y. Begriche, L. Khoukhi, and D. Gaiti, "A stochastic approach for packet dropping attacks detection in mobile ad hoc networks," *Comput. Netw.*, vol. 121, pp. 53–64, Jul. 2017.

[6] D. Ma, X. Hu, H. Zhang, Q. Sun, and X. Xie, "A hierarchical event detection method based on spectral theory of multidimensional matrix for power system," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access, Aug. 9, 2019, doi: 10.1109/TSMC.2019.2931316.

[7] X. Hu, H. Zhang, D. Ma, and R. Wang, "A tnGAN-based leak detection method for pipeline network considering incomplete sensor data," *IEEE Trans. Instrum. Meas.*, early access, Dec. 18, 2020, doi: 10.1109/TIM.2020.3045843.

[8] Z. Movahedi, Z. Hosseini, F. Bayan, and G. Pujolle, "Trust-distortion resistant trust management frameworks on mobile ad hoc networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1287–1309, 2nd Quart., 2016.

[9] H. Xia, F. Xiao, S.-S. Zhang, X.-G. Cheng, and Z.-K. Pan, "A reputation-based model for trust evaluation in social cyber-physical systems," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 792–804, Apr. 2020.

[10] T. Li, J. Ma, Q. Pei, H. Song, Y. Shen, and C. Sun, "DAPV: Diagnosing anomalies in MANETs routing with provenance and verification," *IEEE Access*, vol. 7, pp. 35302–35316, 2019.

[11] R. J. Cai, X. J. Li, and P. H. J. Chong, "An evolutionary self-cooperative trust scheme against routing disruptions in MANETs," *IEEE Trans. Mobile Comput.*, vol. 18, no. 1, pp. 42–55, Jan. 2019.

[12] G. Vaseer, G. Ghai, and D. Ghai, "Novel intrusion detection and prevention for mobile ad hoc networks: A single- and multiattack case study," *IEEE Consum. Electron. Mag.*, vol. 8, no. 3, pp. 35–39, May 2019.

[13] Y. Wang, I.-R. Chen, J.-H. Cho, A. Swami, and K. S. Chan, "Trust-based service composition and binding with multiple objective optimization in service-oriented mobile ad hoc networks," *IEEE Trans. Services Comput.*, vol. 10, no. 4, pp. 660–672, Jul. 2017.

[14] D.-G. Zhang, J.-X. Gao, X.-H. Liu, T. Zhang, and D.-X. Zhao, "Novel approach of distributed & adaptive trust metrics for MANET," *Wireless Netw.*, vol. 25, pp. 3587–3603, Aug. 2019.

[15] V. Keerthika and N. Malarvizhi, "Mitigate black hole attack using hybrid bee optimized weighted trust with 2-Opt AODV in MANET," *Wireless Pers. Commun.*, vol. 106, no. 2, pp. 621–632, May 2019.

[16] L. Mechtri, F. D. Tolba, and S. Ghanemi, "An optimized intrusion response system for MANET," *Peer Peer Netw. Appl.*, vol. 11, no. 3, pp. 602–618, 2017.

[17] R. A. Demidov, P. D. Zegzhda, and M. O. Kalinin, "Threat analysis of cyber security in wireless adhoc networks using hybrid neural network model," *Autom. Control Comput. Sci.*, vol. 52, no. 8, pp. 971–976, Dec. 2018.

[18] H. Moudni, M. Er-rouidi, H. Mouncif, and B. E. Hadadi, "Black hole attack detection using fuzzy based intrusion detection systems in MANET," *Procedia Comput. Sci.*, vol. 151, pp. 1176–1181, Jan. 2019.

[19] M. Vigenesh and R. Santhosh, "An efficient stream region sink position analysis model for routing attack detection in mobile ad hoc networks," *Comput. Electr. Eng.*, vol. 74, pp. 273–280, Mar. 2019.

[20] S.-S. Zhang, S.-W. Wang, H. Xia, and X.-G. Cheng, "An attack-resistant reputation management system for mobile ad hoc networks," *Procedia Comput. Sci.*, vol. 147, pp. 473–479, Jan. 2019.

[21] K. Sakthidasan, N. Vasudevan, P. K. G. Diderot, and C. Kadhiravan, "WOAPR: An affinity propagation based clustering and optimal path selection for time-critical wireless sensor networks," *IET Netw.*, vol. 8, no. 2, pp. 100–106, Mar. 2019, doi: 10.1049/iet-net.2018.5081.

[22] G. Kyriazis and A. Rouskas, "Joint access and backhaul power Consumption optimization in heterogeneous mobile broadband networks," *J. Green Eng.*, vol. 6, no. 4, pp. 337–368, 2017.

[23] Q. Sun, R. Han, H. Zhang, J. Zhou, and J. M. Guerrero, "A multiagent-based consensus algorithm for distributed coordinated control of distributed generators in the energy Internet," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3006–3019, Nov. 2015.

[24] B. Huang, L. Liu, H. Zhang, Y. Li, and Q. Sun, "Distributed optimal economic dispatch for microgrids considering communication delays," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 8, pp. 1634–1642, Aug. 2019.

[25] R. Wang, Q. Sun, D. Ma, D. Qin, Y. Gui, and P. Wang, "Line inductance stability operation domain assessment for weak grids with multiple constant power loads," *IEEE Trans. Energy Convers.*, early access, Sep. 2, 2020, doi: 10.1109/TEC.2020.3021070.

[26] Y. Li, D. W. Gao, W. Gao, H. Zhang, and J. Zhou, "Double-mode energy management for multi-energy system via distributed dynamic event-triggered Newton-raphson algorithm," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5339–5356, Nov. 2020.

[27] L. Yushuai, W. Gao, W. Gao, H. Zhang, and J. Zhou, "A distributed double-Newton descent algorithm for cooperative energy management of multiple energy bodies in energy Internet," *IEEE Trans. Ind. Informat.*, early access, Oct. 12, 2020, doi: 10.1109/TII.2020.3029974.

K. SAKTHIDASAN SANKARAN (Senior Member, IEEE) received the B.E. degree from Anna University, in 2005, the M.Tech. degree from SRM University, in 2007, and the Ph.D. degree from Anna University, in 2016. He is currently an Associate Professor with the Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, India. He has published more than 50 articles in refereed journals and international conferences. He has also published three books to his credits. His research interests include image processing, wireless networks, cloud computing, and antenna design. He is member in various professional bodies. He is an active reviewer in Elsevier, Springer, and Taylor & Francis Journals, and an editorial board member in various international Journals.

N. VASUDEVAN received the bachelor's degree from Madurai Kamaraj University, the master's degree from REC Trichy, and the Ph.D. degree from Anna University, Chennai. He was the dean for about four years and a principal for four years. He has published many articles in both national and international journals. His research interests include VLSI design and communication processor design, image processing, wireless networks, cloud computing, and antenna design.

K. R. DEVABALAJI received the bachelor's degree in electrical and electronics engineering and the master's degree in power electronics and drives from Anna University, Chennai, and the Ph.D. degree from the Vellore Institute of Technology, Deemed University, Vellore, in recognizing his significant contribution in the area of distribution system. He is currently an Assistant Professor (SG) of the Department of Electrical and Electronics Engineering, Hindustan Institute of Technology and Science, Chennai. He is also involved in the funded project titled "Optimal Dispatch of Virtual Power Plant Using Cyber-Physical Controller For Real -Time EMS", funded by Royal Academy of Engineering, U.K., with wroth of £80,000. This project is expected to be complete by March 2021.

THANIKANTI SUDHAKAR BABU (Senior Member, IEEE) received the B.Tech. degree from Jawaharlal Nehru Technological University, Ananthapur, India, in 2009, the M.Tech. degree in power electronics and industrial drives from Anna University, Chennai, India, in 2011, and the Ph.D. degree from VIT University, Vellore, India, in 2017.

He was associated as a Postdoctoral Researcher with the Institute of Power Engineering, Universiti Tenaga Nasional (UNITEN), Malaysia, from 2019 to 2020. He is currently working as an Associate Professor with the Department of Electrical Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, India. He has published more than 60 research articles in various renowned international journals. His research interests include the design and implementation of solar PV systems, renewable energy resources, power management for hybrid energy systems, storage systems, fuel cell technologies, electric vehicles, and smart grids. He has been acting as an Editorial Board Member and a Reviewer for various reputed journals, such as the IEEE and IEEE Access, IET, Elsevier, and Taylor and Francis.

HASSAN HAES ALHELOU (Senior Member, IEEE) is currently a Faculty Member with Tishreen University, Lattakia, Syria. He has published more than 30 research articles in the high quality peer-reviewed journals and international conferences. He has also performed more than 160 reviews for high prestigious journals including the IEEE Transactions on Industrial Informatics, the IEEE Transactions on Industrial Electronics, *Energy Conversion and Management*, *Applied Energy*, the *International Journal of Electrical Power & Energy Systems*. He has participated in more than 15 industrial projects. His current research interests include power systems, power system dynamics, power system operation and control, dynamic state estimation, frequency control, smart grids, micro-grids, demand response, load shedding, and power system protection.

He is included in the 2018 and 2019 Publons list of the top 1% best reviewer and researchers in the field of engineering. He was a recipient of the Outstanding Reviewer Award from *Energy Conversion and Management* Journal, in 2016, *ISA Transactions* Journal, in 2018, *Applied Energy* Journal, in 2019, and many other Awards. He was also a recipient of the Best Young Researcher in the Arab Student Forum Creative among 61 researchers from 16 countries at Alexandria University, Egypt, 2011.

T. YUVARAJ received the B.E. degree in electrical and electronics engineering and the M.E. degree in power electronics from Anna University, Chennai, India, in 2011 and 2013, respectively, and the Ph.D. degree from VIT University, Vellore, India, in 2017. He is currently a Senior Assistant Professor of the Department of EEE, Saveetha Institute of Medical and Technical Sciences, Chennai, India. He has published more than 35 Web of science/Scopus indexed journals. His research interests include the optimal allocation of compensating devices in the distribution networks using optimization algorithms. He is a member of IET and IAENG. He has also served as reviewer to some reputed journals.