Article

Mobility, Residual Energy, and Link Quality Aware Multipath Routing in MANETs with Q-learning Algorithm

Valmik Tilwari [1], Kaharudin Dimyati [1,*], MHD Nour Hindia [1], Anas Fattouh [2] and Iraj Sadegh Amiri [3,4,*]

[1] Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia; valmiktilwari@gmail.com (V.T.); nourhindia@hotmail.com (M.N.H.)
[2] Academy of Innovation, Design, and Technology (IDT), Division of Computer Science and Software Engineering, Mälardalen University, 72123 Västerås, Sweden; anas.fattouh@mdh.se
[3] Computational Optics Research Group, Advanced Institute of Materials Science, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
[4] Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
* Correspondence: kaharudin@um.edu.my (K.D.); irajsadeghamiri@tdtu.edu.vn (I.S.A.)

check for updates

Abstract: To facilitate connectivity to the internet, the easiest way to establish communication infrastructure in areas affected by natural disaster and in remote locations with intermittent cellular services and/or lack of Wi-Fi coverage is to deploy an end-to-end connection over Mobile Ad-hoc Networks (MANETs). However, the potentials of MANETs are yet to be fully realized as existing MANETs routing protocols still suffer some major technical drawback in the areas of mobility, link quality, and battery constraint of mobile nodes between the overlay connections. To address these problems, a routing scheme named Mobility, Residual energy and Link quality Aware Multipath (MRLAM) is proposed for routing in MANETs. The proposed scheme makes routing decisions by determining the optimal route with energy efficient nodes to maintain the stability, reliability, and lifetime of the network over a sustained period of time. The MRLAM scheme uses a Q-Learning algorithm for the selection of optimal intermediate nodes based on the available status of energy level, mobility, and link quality parameters, and then provides positive and negative reward values accordingly. The proposed routing scheme reduces energy cost by 33% and 23%, end to end delay by 15% and 10%, packet loss ratio by 30.76% and 24.59%, and convergence time by 16.49% and 11.34% approximately, compared with other well-known routing schemes such as Multipath Optimized Link State Routing protocol (MP-OLSR) and MP-OLSRv2, respectively. Overall, the acquired results indicate that the proposed MRLAM routing scheme significantly improves the overall performance of the network.

Keywords: MANETs; MRLAM; Q-Learning algorithm; MP-OLSR; RWP

1. Introduction

In recent years, advances in wireless communication and internet service technologies have tremendously heightened the prospect of wireless mobile computing applications [1]. Nowadays, several wireless communication technologies such as MANETs [2], Vehicular Ad-hoc Networks (VANETs) [3], Cognitive Radio (CR) [4], Wireless Sensor Networks (WSN) [5], Device to Device Communication (D2D) [6], Coordinated Multipoint (CoMP) [7], Internet of things (IoT) [8], Carrier Aggregation (CA) [9], Passive Optical Network (PON) [10], Massive Multiple Input Multiple Output (MIMO) [11], Non-Orthogonal Multiple Access (NOMA) [12,13], and Fog Computing [14] are now

becoming increasingly popular due to their compelling performances and application in the vast area of wireless communications. MANETs and VANETs are prominent among other wireless applications related to D2D networked systems. Smart devices in fifth generation (5G) wireless systems can be loaded not only with a variety of sensors, but also by different means of communication that enable access to the Internet. Technologies such as Wi-Fi, Bluetooth, NFC, and more recently Wi-Fi P2P (Wi-Fi Direct), can be used to enrich the smart side of these devices by enabling D2D communication and providing more functionality to enhance the user experience [15,16] and facilitate interactivity with the surrounding environment [17–22]. However, such level of user experience and interactivity may become seriously reduced in remote areas with intermittent cellular services and/or lack of Wi-Fi coverage. In times of calamitous events, such as an earthquake, communication infrastructures may become dysfunctional or degraded due to failure or damage, in such a situation, the MANET paradigm becomes the most convenient way to provide rapid deployment and self-manage wireless link with the end users. These functionalities and features of MANET have offered a broader possibility for extensive research in building the future 5G wireless network community [23–25]. MANET is a self-generated and self-structured wireless network which does not rely on a centralized infrastructure for its operation. In other words, all mobile nodes/devices in a MANET network collaborate with each other and act as a router for one another, thereby providing a robust and effective operation throughout the whole network. Mobile nodes are incorporated with routing functionality, and each node can join and/or leave the network at will depending on the capacity of its energy resources and nature of its network topology [26]. MANET is characterized by event such as constant change in network topology, which often leads to frequent link failure, degraded transmission quality, and reduced network throughput [27]. To overcome these issues, it is imperative to design, develop and implement a new generation of routing protocols that support robust and efficient routing in MANETs.

Routing protocols in MANETs are classified into three categories, namely; proactive, reactive, and hybrid routing protocols, depending on the nature of underlying routing information and update mechanism they employ. The proactive routing protocols are also known as table-driven routing protocols because of their tendency to store updated routing information in the form of tables. Notable examples of proactive routing protocols include destination-sequenced distance-vector (DSDV) [28] and optimized link state routing protocol (OLSR) [29]. Each node maintains and stores the network topology information in a tabler form in order to maintain a consistent network view as the network topology constantly updates periodically due to the very nature of ad-hoc wireless network. From time to time, nodes that run the proactive routing protocol need to exchange up-to-date routing information table among themselves. Whenever a node needs to transmit packets, it firstly extracts routing information from the maintained table with which it routs the packets. Thus, less time is required for a route discovery process which leads to reduced end-to-end delay for data transmission between the source and destination nodes. In contrast, this periodic exchange of routing information and route request packets in the route discovery process leads to high control packet overhead throughout the network. On the other hand, the reactive routing protocols are also known as on-demand routing protocols due to their on-demand manner for the route selection process. Some good examples of reactive routing protocols are dynamic source routing (DSR) [30] and ad hoc on-demand distance vector (AODV) [31]. Nodes running the reactive routing protocols do not need to maintain any prior routing information but exchange routing information of the network only when there is a need for communication. Thus, source node based on reactive routing protocols simply discovers routes in an on-demand manner to establish a connection to a destination node; hence, few control packets overhead are generated for the maintenance of network topology information. Conversely, during the route discovery initiation process in reactive protocols, all nodes exchange topology information with one other, which typically involves flooding of packets in the network. Then, the reactive protocol takes time to gather and analyze the network topology information. This process tends to prolong the end-to-end delay for data transmission between source and destination nodes [32]. Hybrid routing protocols such as zone routing protocol (ZRP) [33] and Secure Link State routing (SLSP) [34] handle

routing activity by dividing the network into zones while combining the best features of reactive and proactive routing protocols to make a timely and informed routing decision. For instance, in the event that a node wants to transmit packets to a destination node within a certain geographical zone of its immediate neighboring network, it will choose the proactive approach to handle this task. On the other hand, if the destination node is outside the geographical zone of its immediate neighbor, it will rather use the reactive routing protocols. However, this approach increases the complexity, computational cost, and energy consumption of mobile nodes [35] during the route selection process in the hybrid routing protocol, which degrades the network performance due to repeatedly changing network zones and constant switching between routing protocols.

2. Contribution

The MANETs reactive routing protocols are more practical due to their ability to maintain a reasonable network convergence time and minimize flooding of packets in the route establishment process, especially in real-time communication scenarios. In practical terms, it is evident that MANETs do not rely on a fixed infrastructure or any centralized ancillary for coordination and control between the mobile nodes. Thus, mobile nodes are simultaneously considered as hosts as well as routers when establishing a connection between end users in the network. Therefore, whenever a route needs to be established between a pair of mobile nodes, the status of intermediate nodes must be considered before determining the routing in MANETs. Routing in MANETs involves establishing a stable and reliable route between a source and destination nodes. However, factors such as channel link quality, mobility, and energy resource constraints of mobile nodes, remain major challenges. These challenges are briefly discussed as follows:

Link Quality: The dynamic time-varying characteristics of the MANETs pose link failure. To overcome this, it is of paramount importance to consider the reliability of link quality as it interacts with the Media Access Control (MAC) layer for the selection of the optimal route.

Mobility: The network topology in MANETs is highly dynamic and changes in an unpredictable manner due to high displacement and random addition/subtraction of intermediate nodes in the network, which is capable of degrading the overall performance and can disrupt the and Quality of Service (QoS) of the network.

Energy constraints: Limited energy resources and processing power of mobile nodes are major source of constraints in the design of an optimal route, which incur severe impacts on route stability and lifetime of the network.

Although, number of multiobjective routing metrics have been suggested by researchers in the literature as discussed in Section 3; however, they did not combine multiple parameter metrics with decision-making techniques such as Q-learning algorithm to simplify information exchange throughout the network and reduce routing overhead. To the best of our knowledge, none of these studies addressed mobility, link quality, and energy efficiency of nodes simultaneously with the multipath concept in the selection of the optimal route. Therefore, in this paper, we have considered link quality, mobility, and energy resource constraints of mobile nodes factor in the proposed scheme.

The rest of this paper is organized as follows. Section 3 reviews the existing literature pertaining to OLSR reactive routing scheme. Section 4 describes the proposed MRLAM routing scheme in detail and presents simulation parameters and the scenario used to validate the proposed scheme. Section 5 presents the results and discussion. Finally, the conclusion and future work are presented in Sections 6 and 7, respectively.

3. Related Works

Multipath routing protocols have been extensively researched in the last few decades and several MANETs paradigms which are said to possess the ability to provide reliable communications, ensure load balancing, and improve QoS have been designed by researchers [36]. Multipath routing is an additional heuristic dimension introduced into routing paradigms with the intention to improve

end-to-end delay, reduce packet overhead, and maximize network lifetime. In the first version of OLSR reactive routing protocol introduced in [37], some specific nodes named as Multi-Point Relay (MPR) are selected during routing to minimize traffic overhead and copes with message flooding problem in the network. These MPR nodes are selected two hops away from neighbor node to be used for forwarding packets to the destination. Thus, excessive load induces on MPR nodes contributes to the rapid depletion of their battery and increases the chance of link failure. Kots and Kumar et al. [38] proposed a Fuzzy logic based novel routing scheme, Neuro-fuzzy, a fuzzy-genetic techniques for the quality attributes of MPR nodes selection. Meanwhile, they applied the Soft Computing (SC) technique for the selection of qualified MPR node to improve energy efficiency and network lifetime for reliable data transmission in MANETs. Novel strategies for selection of MPR nodes using a simplified OLSR routing protocol were proposed in [39] without the addition of a signaling message to forwarding packets. In addition to that, the main aim is to optimize the network performance by reducing the number of signaling messages, collision & traffic congestion, and energy consumption in the network. In [40], the authors proposed a proactive Multipath Optimized Link State Routing Protocol (MOLSR), in which new cross-layer metric and node discovery algorithm are employed to build two disjoint paths between source and destination node. The first disjoint path is used for data communication, whereas the second path acts as an alternate path in case link failure occurs in the first path. Moreover, the MOLSR routing protocol broadcasts a smaller number of topological messages as compared to OLSR for the discovery of multiple paths between source to destination node. The results prove that it provides a better route selection mechanism and reduces packet overhead for the network. In [41], a modified OLSR routing scheme called OLSRM for reliable routing and energy conservation in the network is proposed, which selects mobile nodes based on the energy awareness and drain rate. The OLSRM aims to maximize MPR lifetime and minimizes routing overhead in the network.

Due to the scalability problem of flat routing schemes in ad-hoc networks, in [42] the authors presented the Heterogenous OLSR routing (HOLSR) scheme, which is more suitable for large-scale heterogeneous networks. The HOLSR scheme utilizes multiple interfaces and reduces control message overhead in the network to improve the performance of the routing mechanism. The authors in [43] introduced a multipath routing scheme called Quality OLSR routing (QOLSR), in which path selection criteria is based on the multiple QoS routing metric such as bandwidth and delay. The paths are loop-free and multiple node disjoint, computed by using shortest–widest path algorithm. Moreover, a correlation factor is presented which defines the number of links between each two disjoint paths to minimize the interferences between multiple paths to achieve better QoS with improved network resource utilization. Sarkar et al. [44] proposed a point-based mobility factor computing technique for route selection which provides better routing performance for frequent topology changing scenarios in MANETs. The point-based mobility factor value is estimated based on the pause time, speed and moving direction of the mobile nodes for the static and dynamic scenario. Moreover, the Mobility Aware Routing (MAR) technique based on the mobility factor is used for the route selection process. It provides better QoS and performance for the static and highly dynamic network. An energy-aware routing model named enhanced intellects masses optimizer energy-efficient and secure optimized link state routing (EIMO-ESOLSR) is proposed in [45] for energy efficient and secure MPR selection in OLSR routing scheme. The MPR selection in the route discovery process is based on the willingness and Composite Eligibility Index (CEI) for each node in the network. The willingness value is estimated based on the available bandwidth, lifetime, and queue occupancy metric of the network nodes, whereas the CEI value is calculated based on the power factor, misbehaving probability, and forwarding behavior of the network nodes. The results show that the proposed EIMO-ESOLSR routing scheme reduces the energy consumption, provides a secured selection of MPR nodes, and minimizes flooding of packets in the network.

One of the limitations in MANETs is the traffic congestion, that occurs when the excessive flow of packets injected on a single mobile node and carry most of the network traffic that affect the quick depletion of the node's battery. Due to this, routing approaches suffer from balancing

the load among the intermediate nodes of the network, which can further significantly degrade the network performance. In addition to that, the authors in [46] addressed a network utility maximization (NUM)-based congestion control paradigm for cross-layer designs for wireless networks. It considered that the problem of cross-layer congestion control arises from the several dimensions, specifically, elastic or inelastic traffic, same or multitimescale, and single or multipath transmissions. Moreover, it gives comprehensive discussions on various mathematical techniques and solutions used for cross-layer congestion control problems. The Successive Convex Approximation (SCA) methods is also discussed, which strongly depend on initial conditions and network scenarios. Overall, it concluded that the cross-layer congestion control in wireless networks can be solved by utilizing network generation network such as SDN and Cloud-RAN with efficient central controller and network statistics. In [47], the authors introduced Proactive Source Routing protocol (PSR) to facilitate opportunistic data forwarding in MANETs. Each node has full information about the network topology and periodically exchanges messages with neighbors' nodes. The exchanged messages contain link cost information based on the number of nodes along the path from source to destination. The PSR routing maintains more network topology information than distance vector routing, making it ideal for source routing since it incurs less routing overhead with improve data transmission capability. In [48], a routing protocol called Least Common Multiple based Routing (LCMR) is proposed for MANETs which load balances among the different possible paths. The LCMR routing distributes packets and properly uses them for computation of route along multiple paths to minimize the overall routing time.

For link instabilities, energy resource constraints and node mobility of frequently topology changing network, [49] introduced Multipath-ChaMeLeon (M-CML) routing protocols to stabilize and enhance the QoS and performance of the network. The introduced routing protocol helps to minimize routing overhead and energy consumption of mobile nodes by cutting down the amount of generated of duplicate packets in the network. In a dynamically changing network topology, link quality is badly impacted, resulting in frequent link failure in the network. To solve this, a Smooth Mobility and Link Reliability-based OLSR (SMLR-OLSR) routing scheme is proposed in [50]. Moreover, Semi-Markov Smooth and Complexity Restricted mobility model (SMS-CR) is used to provide enough smoothness and low complexity for reliability enhanced MPR selection that facilities longer MPR lifetime and less routing overhead in the network. In [51], authors proposed the swarm intelligence based Ant-based Energy Aware Disjoint Multipath Routing Algorithm (AEADMRA), which consider "node's energy consumption" as one of node selection criterion. The proposed algorithm is a subset of swarm intelligence, which utilizes the ability of ants to solve complex problems by cooperating among themselves. The results show that it delivers better performance in terms of reducing routing complexity and improving route discovery process. A scalable routing protocol named Link-stability and energy Aware Routing protocol (LAER) has been proposed in [52], which is based on the joint metric of link stability and energy drain rate. A Biobjective Integer Programming (BIP) model is adopted to achieve optimal solution in the route selection process. The BIP model selects next hop node among the neighbor nodes, which has better link stability and less energy consumption rate. The LAER scheme prolongs the network lifetime and maintains the robustness of the network. A parallel Disjointed Multipoint routing algorithm-based OLSR called DMP-EOLSR is proposed in [53], which is a hybrid routing protocol that involves both reactive and proactive routing protocols. The DMP-EOLSR considers living time of node as well as living time of links between nodes based on energy consumption and moving state of mobile nodes during route computation. The living time of nodes depends on the residual energy of the nodes, whereas the living time of links between nodes depends on speed and direction of the movement. Moreover, in order to find multiple parallel disjoint paths or link disjoint paths, assessment factors of node energy and an iterative is used based on the modified Dijkstra algorithm. The Dijkstra algorithms are well-known schemes for exploring multiple disjointed paths between source and destination nodes based on topology information of the network. The proposed algorithm not only improves the stability of the selected paths with better

parallel transmission capacity, but also reduces the computation cost of intermediate nodes in route selection and route recovery mechanism.

Another approach proposes a MP-OLSR routing protocol in [54], where the Dijkstra algorithm has been modified in order to find multiple routes for sparse and dense network. It utilized two cost function to generate node disjoint and link disjoint paths. Results proved that the proposed algorithm is able to achieve great flexibility by using different link cost metrics and cost function. In addition, two new concepts were introduced, namely; route recovery and loop detection in order to improve the packet delivery ratio and reduce the chance of optimal link failure for network topology that changes frequently. The MP-OLSR routing improves overall network performance, especially in high mobility and heavy loaded network scenarios. Yi et al. [55] proposed multipath extension version of MP-OLSR stated as MP-OLSRv2, which belongs to the category of hybrid multipath routing protocol that involves proactive and reactive routing concepts for data transmission. The MP-OLSRv2 uses two incremental cost functions for the link cost between the nodes to generate multiple node-disjoint and link disjoint paths. It discovers multiple disjointed paths instead of single disjointed path for data transmission with the multipath Dijkstra algorithm in a frequently changing network topology. The MP-OLSRv2 mechanism is divided into two phases (topology sensing and route computation) to maintain the multiple routes from source to destination pairs. It delivers reliable communication and facilitates traffic load distribution into multiple paths, thereby ensuring load balancing among the nodes of the network. Multipath data transmission is a way of increasing network throughput while maintaining robust link with reliable transmission in the network. The selection of highly qualified intermediate nodes in the route mapping process still remains one of the most important and critical issues for efficient data transmission in MANET networks. Thus, this is one of the key factors mind in our design and implementation of the MRLAM routing protocol proposed in this paper. However, the MP-OLSR and MP-OLSRv2 routing scheme do not consider residual energy, mobility, and link quality status of the intermediate nodes, when selecting the route from source to destination nodes. In the MRLAM scheme, a route is established between source and destination nodes with the active participation of intermediate nodes that have sufficient amount of energy resource, less mobility, and better link quality.

4. System Model

A communication network in Mobile Ad-hoc network environment can be represented as a graph $G = (V, L)$ wherein V is the set of nodes and L is the link between the nodes. The link between the source node s and destination node d is denoted by (s, d) i.e., $s, d \in V$. Whenever node s is the neighbor of node d, both of which can directly communicate with each other. In a centralized network, the network topology information is maintained by a network controller. In a decentralized network, network topology information is maintained by the information exchange between the node and its neighbors. The proposed routing scheme is for a decentralized network, in which all mobile nodes work as hosts as well as a router for other nodes to build network infrastructure. In the proposed scheme, the route selection and data transmission is based on the collaboration among network nodes for topological information exchange among each other. Due to the dynamic movement of nodes, network topology continuously changes in an unpredictable manner. In this scenario, if two nodes do not share a direct link with each other, they will utilize routing protocols to establish a connection, using intermediate nodes between them to transfer data. The proposed MRLAM routing scheme utilizes some key factors, such as the energy consumption, link quality of channel, and mobility of nodes in the route computation, to enhance the energy efficiency and QoS of communicating nodes as discussed below:

4.1. Energy Consumption Estimation of Mobile Nodes

The energy consumption of nodes plays a significant role in the selection of robust and qualified intermediate nodes, which collaborate to improve the overall network performance. The mobile nodes

in the MANETs are battery operated and their energy is limited. Therefore, battery life of the nodes should be taken into account while selecting a set of intermediate nodes for route establishment to a destination in the network. Moreover, the mobile nodes operate in four states, which are; transmission, receiving, idle, and sleep states. The energy consumption in idle and sleep states is low; hence, the transmission and receiving states of nodes are considered in the design of the routing algorithm of our proposed scheme. In the proposed scheme, the energy of each node is linearly discharged as a function of current load based on the Coulomb counting technique [56], which accumulates the dissipated Coulombs from the beginning of the discharge cycle. It estimates the residual energy based on the difference between the accumulated value and prerecorded full charge value of the battery capacity. The proposed scheme utilizes a generic radio energy model [57] to estimate the consumed energy for each state of node such as ($E^{\mathrm{i}}_{Initial}(t)$, $E^{i}_{Residual}$, $E^{i}_{Consumption}$ and $E^{i}_{operation}(t)$). Energy consumption in different states of nodes and duration of time in each state of nodes, affect the performance for the proposed routing scheme. The initial energy level of the intermediate node i at the time t is denoted by $E^{\mathrm{i}}_{Initial}(t)$ and the residual energy $E^{i}_{Residual}(t + \tau)$ of node i at the period of time τ is calculated as follows:

$$E^{i}_{Residual}(t + \tau) = E^{\mathrm{i}}_{Initial}(t) - E^{i}_{Consumption}(t + \tau) \qquad (1)$$

where $E^{i}_{Consumption}(t + \tau)$ refers to the energy consumption of node i for a period of τ, which is the energy consumed during transmission, receiving, exchanging routing information control packets, and during internal operation time of nodes. In the proposed scheme, energy consumption is estimated based on the circuitry power consumption, number of packets exchanged, and time spent in each state. Thus, the energy consumption in a transmission state of the node i for transmitting n number packets can be calculated as follows:

$$E^{i}_{Transmission}(t + \tau) = n \times P^{i}_{Transmission}(t + \tau) \qquad (2)$$

whereas $P^{i}_{Transmission}(t + \tau)$ denotes the amount of power consumed during transmission of n number of data packets and exchanging routing information control packets with neighbor's nodes at the period of τ time, the unit is in Watts per second. Similarly, the energy consumed by node i during the receiving state for transmitting m numbers of packets and including control packets at the period τ is calculated as follows:

$$E^{i}_{Recieve}(t + \tau) = m \times P^{i}_{Recieve}(t + \tau) \qquad (3)$$

where $P^{i}_{Recieve}(t + \tau)$ represents power consumed by node i during exchange control messages and receiving m number of packets at a time τ. Moreover, all the nodes consume energy when they perform internal operation such as connecting, managing, catching, and updating database at the time of τ period, which is denoted by $E^{i}_{operation}(t + \tau)$. Therefore, the total energy consumption of node i at the time period τ in all transmission, reception and operation states can be calculated as follows:

$$E^{i}_{Consumption}(t + \tau) = E^{i}_{Transmission}(t + \tau) + E^{i}_{Recieve}(t + \tau) + E^{i}_{operation}(t + \tau) \qquad (4)$$

Finally, the residual energy of node i updated at time τ is denoted as follows:

$$E^{i}_{Residual}(t + \tau) = E^{\mathrm{i}}_{Initial}(t) - \left\{ E^{i}_{Transmission}(t + \tau) + E^{i}_{Recieve}(t + \tau) + E^{i}_{operation}(t + \tau) \right\} \qquad (5)$$

The above expression provides information on the amount of remaining energy and the amount of energy that will be consumed by node i at time $t + \tau$. Based on this, MRLAM routing scheme gives higher priority to nodes which have higher available residual energy and less energy consumption rate while establishing a connection between end users. Overall, this approach maximizes the network lifetime with decreased chance of link failure in the network.

4.2. Node Mobility Estimation

Node mobility estimation governs the unpredictable movement of a node in a network, which helps to estimate the probability of a link failure and link instability in advance. The random waypoint model (RWP) is extensively used for estimation of node's mobility and for simulation analysis in the mobile wireless network due to its simplicity and better QoS [58]. Several parameters, such as speed, move time, pause time, and separation distance of the mobile nodes, are configured for the estimation of node's mobility. The RWP model periodically estimates the mobility of nodes and selects the random destination point ("waypoint") in the network based on the above parameters. Once the nodes reach their selected waypoints, nodes are pause for defined duration of time (pause time), and this process is repeated. The waypoints are uniformly and randomly chosen in the network area. In the proposed scheme, nodes can move randomly in an area of operation and pause for an arbitrary period of time. The maximum and minimum velocity ($v_{\max}$ and $v_{\min}$) of each node is configured based on the RWP model along with its calculation of pause time duration, so as to simulate real deployment scenarios. Network nodes are uniformly distributed in the RWP model between the velocities range $[v_{\max} - v_{\min}]$, where $v_{\max}$ and $v_{\min}$ are the minimum and maximum velocities of the nodes, respectively. The node speed distribution $f_{(s,d)}(v)$ in the dynamic network is described as below:

$$f_{(s,d)}(v) = p_{mov} \cdot \frac{1}{v \ln(v_{\max}/v_{\min})} + p_{pause} \cdot \delta(v) \tag{6}$$

where the velocity function v is defined in the range of $v \in [v_{\max} - v_{\min}]$, p_{pause} and $p_{mov} = (1 - p_{pause})$ are the probabilities of nodes in the pause and moving states in the network respectively. $\delta(v)$ defines the Dirac delta function whose value updates according to the velocity of the node. The value of $\delta(v)$ varies from 0 to 1, when the velocity of a node is maximum, $\delta(v)$ is updated to 0, and when the velocity of the node is minimum, $\delta(v)$ become 1. The pause time of nodes is considered to be $t_p \geq 0$, and the value of probability of pause time p_{pause} can be calculated as follows:

$$p_{pause} = \frac{t_p}{t_p + E[D] \frac{\ln(v_{\max}/v_{\min})}{(v_{\max} - v_{\min})}} \tag{7}$$

where $E[D]$ denotes expected value of distance cover due to node movement on the whole trip. If a node starts from the pause state p_{pause}, it pause time is set to be t_p. On the other hand, if the node starts from the moving state p_{mov}, it will select the node speed range of $v \in [v_{\max} - v_{\min}]$. Recall that, p_{pause} is the probability of a mobile node in the pause state, while t_p is the time spent by the node during pause state. The mobility of nodes in a distributed dynamic network is calculated as follows:

$$\text{Mobility} = \min_{v \in [v_{\max} - v_{\min}]} \sum_{(s,d) \in N} f_{(s,d)}(v) \tag{8}$$

The stability of an established path between the source and destination nodes depends on the mobility of the intermediate nodes which form the path. Therefore, a path will be unstable if the mobility of the constituent nodes of the path is high. Based on the above mobility value of nodes in the network, the MRLAM routing scheme prioritizes nodes with less mobility, meaning that the route's stability and network's lifetime are increased, resulting in reduced probability of link failure.

4.3. Link Quality Estimation of Mobile Nodes

The link quality of the channel is one of the key factors in the selection of the reliable route, which is estimated based on the Expected Number of Transmission (ETX) and Expected Any-path Count (EAX) parameters [59]. The ETX is the expected number of transmissions required before a packet is successfully delivered to the next-hop which also determines link quality and packet loss on both directions of a link. Node *i* is the intermediate node whose probabilities of receiving and

forwarding data packets are λ_i and μ_i, respectively. These values are measured through the probe message that is transmitted using dedicated link before actual data transmission. Each node broadcasts a probe message to all of its neighbor nodes and maintains records of transmitted probe message for a designated period of τ seconds. Accordingly, the packet delivery ratio probability from a source node to neighbor nodes is calculated as follows:

$$\lambda_i = \frac{n_w}{w/\tau} \tag{9}$$

whereas n_w denotes the number of probe messages delivered in a period of w seconds satisfying the condition $w > \tau$. The probability that a neighbor's node i receives a packet from at least one forward node is $p_{for} = 1 - \prod_{i>s}(1 - \mu_s\lambda_i)$, and the probability that node i successfully delivers a packet to at least one backward node is $p_{back} = 1 - \prod_{d>i}(1 - \mu_i\lambda_d)$. Consequently, the expected number of transmissions that node i needs to take can be calculated as:

$$ETX = \frac{1}{p_{for} \times p_{back}} = \frac{1}{(1 - \prod_{s>i}(1 - \mu_s\lambda_i))(1 - \prod_{d>i}(1 - \mu_i\lambda_d))} \tag{10}$$

The ETX calculates the expected number of transmissions needed to successfully deliver a packet from source to destination through the forward node set i. The ETX metric selects all forwarded candidates and prioritizes them based on the Equation (10). However, the ETX metric also includes forwarded sets that increase packet overhead in the network. To address this issue, the EAX metric is employed to select only the optimal forwarded sets and prioritize them based on opportunistic routing. This decreases interference with neighbor's node and minimized packets overhead, and number of transmissions to some extent. The hop-count parameter is useful for optimized routing performance with minimum hop-count value. In addition to that, EAX minimizes the number of intermediate nodes and number of transmissions required for reliable packet's delivery on optimal route without contrarily affecting the performance of network. $C^{s,\,d}$ defines the forwarded set of candidates from node s to d and $C_i^{s,\,d}$ defines the node i forwarded set prioritizing range from 0 to 1. Therefore, the delivery probability between source and $C_i^{s,\,d}$ is p_{for} and vice versa (considering both the forward data and backward ACK transmissions). Then, the EAX value from source to destination through forwarded set can be calculated as follows:

$$EAX(s,\,d) = \frac{1 + \sum_i EAX(C_i^{s,\,d},d)(1 - \prod_{s>i}(1 - \mu_s\lambda_i))\prod_{j-1}^{i-1}\left(1 - \prod_{d>i}(1 - \mu_s\lambda_j)\right)}{(1 - \prod_{s>i}(1 - \mu_s\lambda_i))} \tag{11}$$

In the above expression, the EAX metric selects and enables the potential candidate pairs of intermediate nodes based on the packet's delivery probability. The selection of potential candidates in the optimal route from source to destination nodes is performed as follows. Firstly, the ETX metric potential candidate $C_{potential}^{s,\,d}$ is determined based on the best path, such that if $ETX(s,d) > ETX(j,d)$, node j will be added to the potential candidates in the route. Secondly, the subset of potential nodes $C_{potential}^{s,\,d}$ is selected as the actual candidates set $C^{s,\,d}$ having the smallest ETX value to the destination. Also, a potential node is added to the forwarded set $C^{s,\,d}$ when it decreases the value of $EAX(s,\,d)$ by a factor δ, which is defined as a configurable parameter. This step iterates until the new potential node is added in the network.

The selection of intermediate potential candidate nodes with ETX and EAX metrics is shown in Figure 1. It can be observed that source node s, destination node d, and next-hop neighbor nodes b, e, and f of node c possess a different ratio of delivery probabilities. The ETX metric selects all three neighbor nodes such as b, e, and f because paths from these nodes to d have smaller ETX values than that from s to d. These three neighbor's nodes are sorted in ascending order based on the ETX value obtained by using Equation (10). On the other hand, the EAX select only two nodes, such as b and e, because with these two candidates, the EAX from s to d is less than the EAX from f to d, adding f to

the candidate set does not decrease EAX between *s* and *d*. Furthermore, EAX prioritizes node *b* over node *e* as it possesses a smaller value of EAX based on Equation (11). Next hop selection comparison between ETX and EAX potential intermediate candidate's selection and prioritization from source to destination nodes by ETX and EAX is illustrated in Table 1.
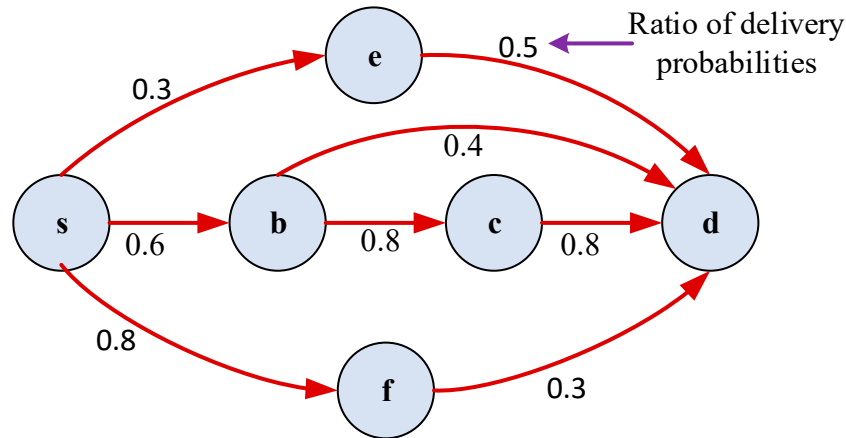

Figure 1. The general scenario for potential node selection.

Table 1. Selection of potential intermediate nodes in optimal route based on the Expected Number of Transmission (ETX) and Expected Any-path Count (EAX) metrics.

Methods	(Source, Destination)	Hop Count	Nodes in Route	Priority
ETX	*(s, d)*	3	*b, e, f*	*e > b > f*
EAX	*(s, d)*	2	*b, e*	*b > e*

4.4. Node Selection in the Optimal Route Based on Q-Learning Algorithms:

Q-Learning is a prevalent paradigm of reinforcement learning algorithms applied in wireless multipath networks for decision making policy to improve network performance. Some salient attributes of Q-Learning that make it the most widely adopted algorithm are its simplicity and model-free nature. Moreover, Q-Learning generates an optimal reward value within environments' state–action pairs that are described by Markov decision processes [60]. Q-learning observes agent behavior through trial-and-error interactions with a dynamic environment through a tuple (S, A, R) such as states, actions, and rewards as shown in Figure 2.
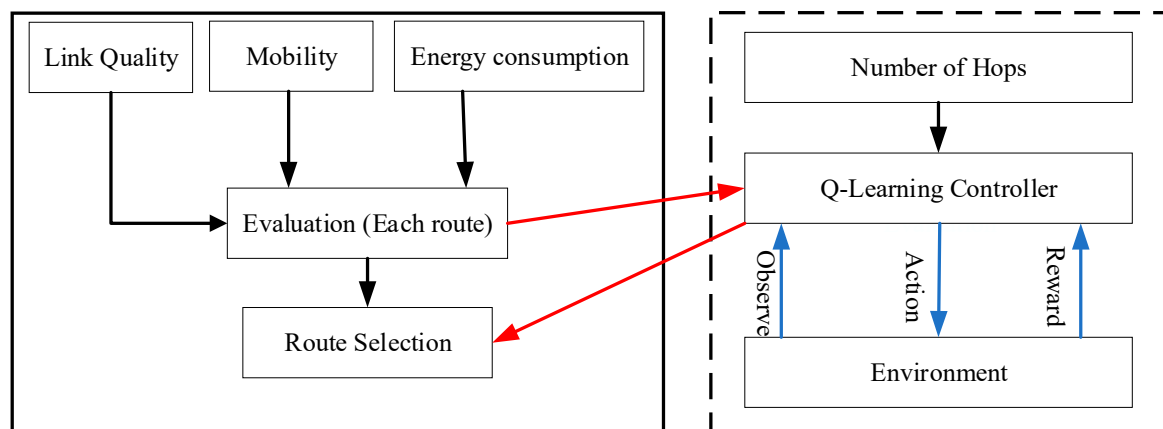

Figure 2. Q-learning approach for route selection.

　　　The objective of the agent is to maximize its long term rewards of the selection procedure which are obtained from immediate and discounted rewards for the estimation of the Q-value $Q_t^i(s_t^i, a_t^i)$, ($s \in S$, $a \in A$) by taking a given action a_t^i at the given state s_t^i. The Q-Learning algorithm learns in a nondynamic environment and adopts the best policy to solve routing problems in a distributed manner by utilizing the Q-value equation as shown below.

$$Q_{t+1}^i(s_t^i, a_t^i) \leftarrow (1-\alpha) \times Q_t^i(s_t^i, a_t^i) + \alpha\left[r^i(s_{t+1}^i) + \gamma \times \max_{a \in A} Q_t^i(s_{t+1}^i, a_{t+1}^i)\right] \tag{12}$$

where α is defined as the learning rate and ranges between $0 \leq \alpha \leq 1$, it depends upon variation of Q-value that changes with dynamic topology of the network. If Q-value changes very fast, learning rate value goes to 1 and the agent takes only new reward value. The parameter $\gamma \in [0, 1]$ is a discount factor, which determines the importance of future rewards. Low value, i.e., $\gamma = 0$, indicates the system is myopic and merely takes results of the current action into account. By contrast, as γ close to 1, future rewards play an important role in determining optimal actions. In addition, $\max_{a \in A} Q_t^i(s_{t+1}^i, a_{t+1}^i)$ is the model of maximum expected future reward, which selects possible action a^i in the next state s_{t+1}^i. Based on different decision factors, i.e., link quality, mobility, and energy consumption of nodes, as described above, a new reward function for node i updated at each instant $t + 1$ is formulated and is as follows:

$$R_{t+1}^i = w_1 \times EXT(s, d) + w_2 \times \text{Mobility} + w_3 \times \text{Residual Energy} \tag{13}$$

where $w_1, w_2,$ and, w_3 are the corresponding weight assigned to each criterion based on the available status of nodes (i.e., mobility, link quality, and residual energy) which range from 0 to 1 and their sum equals to 1. The above criteria's value and weights facilities the sensitivity of decision factors for future reward value such as R_{t+1}^i. Therefore, Q-value of node i is updated with a previously stored value and the new reward R_{t+1}^i which is calculated using the following equation as below:

$$Q_{t+1}^i(s_t^i, a_t^i) \leftarrow (1-\alpha) \times Q_t^i(s_t^i, a_t^i) + \alpha\left[R_{t+1}^i(s_{t+1}^i) + \gamma \times \max_{a_{t+1}^i \in A} Q_t^i(s_{t+1}^i, a_{t+1}^i)\right] \tag{14}$$

　　　Lastly, based on the expression, next-hop node which has higher Q-value is selected for the optimal route. The flowchart of MRLAM routing scheme shown in Figure 3 describes the process of selection of optimal route from source to destination based on Q-learning algorithms. The MRLAM routing scheme selects the optimal route applying decision policy-based Q-learning algorithms, which can effectively reduce the redundant information in the network topology and enhance overall network performance. The route selection based on mobility, residual energy, and link quality with the Q-Learning algorithm is described below:

| Q-Learning Algorithms for Route Selection	
1.	At every edge and node between source and destination node
2.	**Initialize** $Q_i(s_t^i, a_t^i) \; \forall (s \in S, a \in A)$
3.	Choose action a_t^i from the state s_t^i using best policy derived Q_{t+1}^i
4.	**Q-value** at time as:
	$Q_{t+1}^i(s_t^i, a_t^i) \leftarrow (1-\alpha) \times Q_t^i(s_t^i, a_t^i) + \alpha\left[r_{t+1}^i(s_{t+1}^i) + \gamma \times \max_{a \in A} Q_t^i(s_{t+1}^i, a_{t+1}^i)\right]$
5.	Based on mobility, residual energy and link quality new reward function
	$R_{t+1}^i = w_1 \times Mobility + w_2 \times RB + w_3 \times Link\ Quality$
6.	**Update Q-value** with new reward function as:
	$Q_{t+1}^i(s_t^i, a_t^i) \leftarrow (1-\alpha) \times Q_t^i(s_t^i, a_t^i) + \alpha\left[R_{t+1}^i(s_{t+1}^i) + \gamma \times \max_{a \in A} Q_t^i(s_{t+1}^i, a_{t+1}^i)\right]$
7.	Select the optimal route with update Q-value
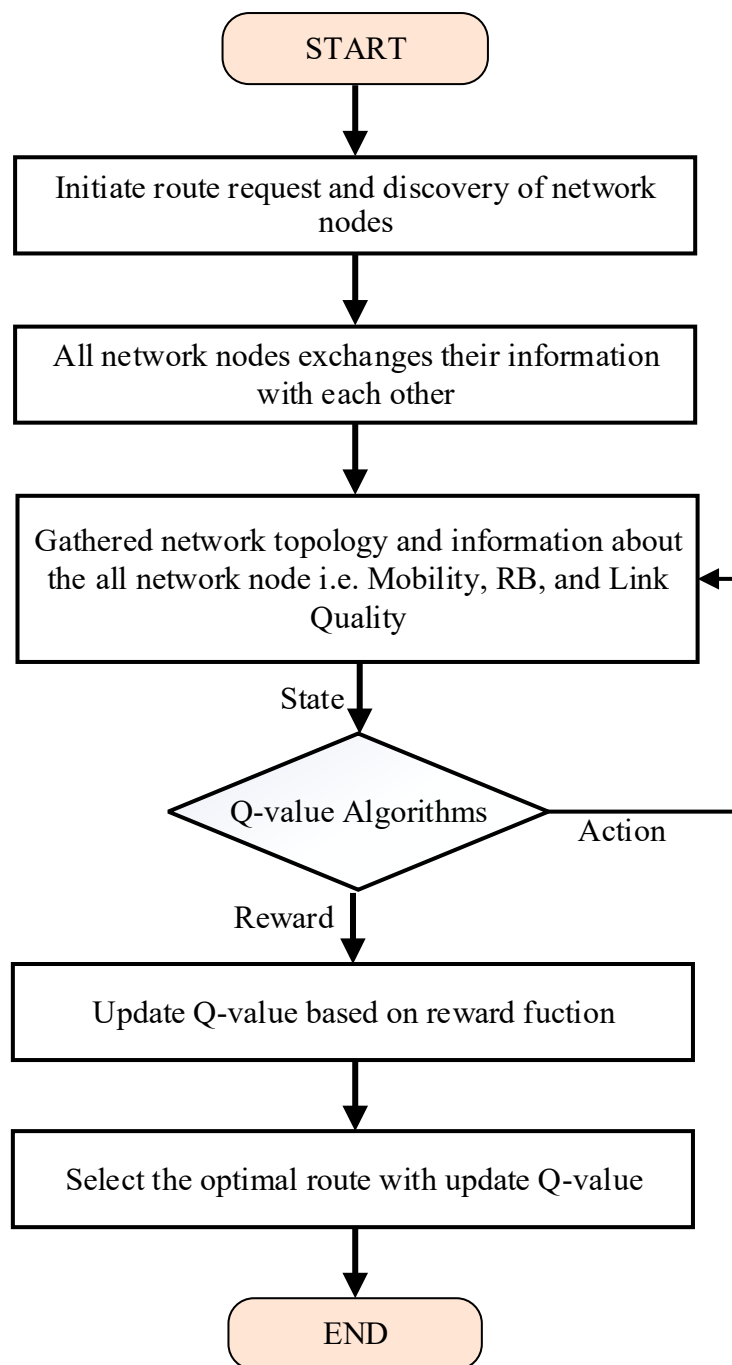
Figure 3. Route selection procedure of the Mobility, Residual energy and Link quality Aware Multipath (MRLAM) routing scheme.

4.5. Simulation Setup

Extensive simulations have been conducted with MATLAB 2018a simulator to evaluate the performance of the MRLAM routing scheme, which is compared with MP-OLSR and its extended version MP-OLSRv2 routing schemes, using different speed of node scenario. Random topologies with a maximum of 49 nodes are generated over a rectangular field in the area of 1000 m × 1000 m. The nodes are placed in the middle of the simulation area and seven data sources are randomly chosen for the scenarios and all sources are transmitting their data to the destination node. The Constant Bit Rate (CBR) generates traffic into networks having a size of the CBR 512 bps from sources and the simulation runs in total for 200 times. 802.11a standard and modules of the wireless physical

layer are utilized in the simulation in order to provide a high level of accuracy. The User Datagram Protocol (UDP) has been used as the transport layer protocol, which, in contrast to Transmission Control Protocol (TCP), provides a simple transmission model. Random Waypoint Mobility model is used as the mobility standard that enables node speed variation from 10 m/s to 60 m/s. The other performance evolution scenarios are presented in Table 2 as follows:

Table 2. Simulation parameters.

Simulation Parameters	Value
Routing protocols	MRLAM, MP-OLSR, MP-OLSRv2
Simulation time	200 s
Traffic type	CBR (UDP)
Battery capacity	3600 mAh
Propagation model	Two ray ground
Generic energy model	$P_{Transmission} = 1300$ mW and $P_{Recieve} = 900$ mW
Signal transmission power	31.623 mW
Random Waypoint Mobility model	Minimum velocity $v_{min} = 10$ m/s, Maximum velocity $v_{max} = 60$ m/s
Learning rate α	[0, 1] based on Q-value
Discount rate γ	0, current reward value 1, future reward value
Dirac delta function $\delta(v)$	0, when v_{max} or 1 when v_{min}
Pause time t_p	10 s
Battery model	Linear battery model
Channel frequency	2.4 GHz
Transmission range	270 m
MAC Layer protocol	IEEE 802.11
Physical layer model	PHY 802.11b

4.6. Evaluation Criteria

The objective of this extensive simulation is to evaluate the performance of the MRLAM routing scheme and the following metric is adopted for the performance evaluation:

(i) *Throughput:* A network throughput represents the average amount of data productivity that is delivered during a period of the network operation time, between the source and destination node pairs. It is expressed in Kbps that can be defined as:

$$\text{Throughput} = \frac{\text{Number of packet recieved} \times 8}{\text{Network operation time}} \tag{15}$$

(ii) *Average End-to-End Delay (Avg. EED):* It is the ratio of total time taken for data transmission between source and destination node to the number of packets received at the destination node, which can be calculated as follows:

$$\text{Avg. EED} = \frac{\text{Total time taken for packets transmission}}{\text{Number of packets recieved}} \tag{16}$$

(iii) *Packet Loss Ratio* : *(PLR):* It is represented as the percentage ratio of the number of packets dropped by the malicious nodes that are selected in the route to the total number of packets sent during the transmission, which is calculated as follows:

$$PLR = \frac{N_s - N_r}{N_s} \times 100 \tag{17}$$

where N_s and N_r denotes the number of bits transmitted from the source and received at destination nodes respectively.

(iv) *Energy Consumption* : It is defined as the total amount of energy consumed by all nodes for key transmission throughout the duration of the simulation. Energy consumption of each node is obtained at the end of each simulation, factoring in the initial energy of each node. The energy consumption formula for transmitting data is:

$$E_{consumption} = \frac{E_{Total}}{\text{Number of packets succesfully transmitted}} \tag{18}$$

where E_{Total} is the total energy of node which is 3600 mAh in the simulation model.

(v) *Energy Cost:* The energy cost metric represents operational model's energy efficiency for period of data transmission with end node pairs that can be evaluated as the energy consumption of nodes to the total number of successfully received packets in the network, as shown below:

$$E_{\cos t} = \frac{\text{Average Energy Consumption}}{\text{Total Packet Recieved}} \tag{19}$$

(vi) *Convergence Time:* Convergence occurs in the network due to frequent network topology changes; meanwhile, the intermediate nodes independently run routing algorithms and recalculate parameters values. The intermediate nodes update routing information and build new routing table based on the parameter's information. It is calculated based on the time required before all of the intermediate nodes can reach a consensus regarding the updated network topology.

5. Results and Discussion

In this section, results acquired from the simulation of MRLAM, MP-OLSR, and MP-OLSRv2 routing schemes under different node speed scenarios are presented. The results of mentioned routing schemes are compared on the following metric, i.e., throughput, average end-to-end delay, packet loss ratio, energy consumption, energy cost, and convergence time. Furthermore, critical analysis of acquired results is also conducted in this section.

5.1. Throughput

The throughput performance of the MRLAM, MP-OLSR, and MP-OLSRv2 routing schemes with increasing node speed shown in Figure 4 indicates that the MRLAM scheme constantly provides acceptable and better throughput performance than the MP-OLSR and MP-OLSRv2 routing schemes in all scenarios of node speed. Higher throughput achieved through the MRLAM scheme is indicative of the mobility awareness consideration factored in during the optimal route selection process, whereas the other two routing schemes do not consider mobility awareness, especially when a link failure occurs as a result of high node's mobility. The MP-OLSR routing employs route recovery and loop detection, while the MP-OLSRv2 employs the topology sensing and route computation mechanism for the selection of an intermediate node in the route. These mechanisms need to transmit extra packets in the network when the route is established, which lead to reduce number of packets received at the destination at the particular time period. The route selected by the MRLAM scheme usually has lower mobility or more energy level than other routes; therefore, the link is more stable and ultimately experiences very few packet drops, which in turn maximizes the throughput. The link failure of nodes

in MRLAM routing is estimated by using ETX and EAX parameters. Based on the values of these parameters, the proposed routing scheme diverts the packet flow toward the intermediate nodes which have better link quality. As the node speed increases, the number of link breaks with MRLAM remains lower than MP-OLSR and MP-OLSRv2 routing schemes and accordingly, the throughput decreases constantly. Throughput decreases from 47.39 kbps to 39.87 kbps for MP-OLSR, from 49.17 kbps to 40.87 kbps for MP-OLSRv2, and from 50.86 kbps to 42.63 kbps for MRLAM when the node speed varies from 10 m/s to 60 m/s respectively. The proposed routing scheme selects the route based on the status of link quality metric, which reduces the number of transmissions and increases channel utilization efficiency upon which reliable link is selected. This, in turn, increases the throughput of the whole network. Compared with the other schemes, critical analysis of the simulation results demonstrates that MRLAM scheme clearly performs better when it comes to selecting intermediate nodes that are more qualified in terms of low mobility, energy level, and link quality. As a result, the route recovery process invokes less frequent link failure; therefore, improved overall network throughput is achieved.
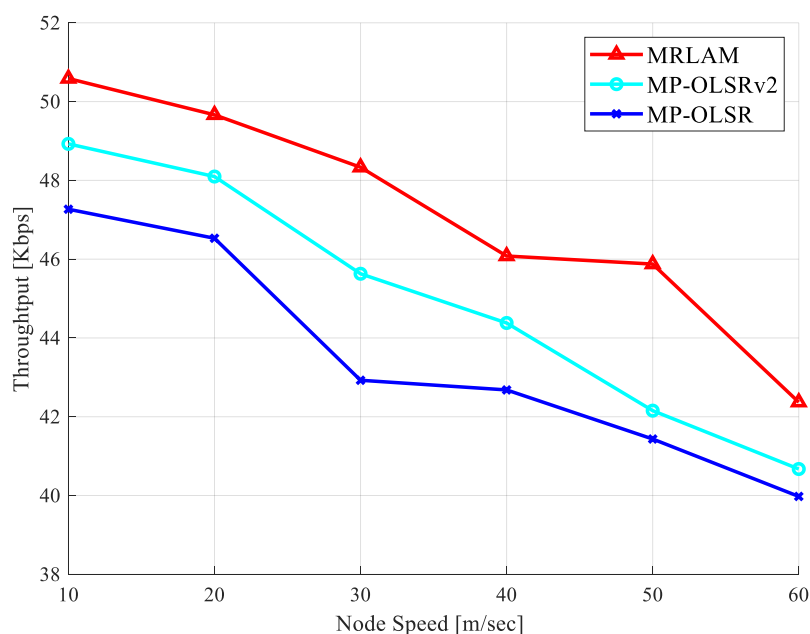
Figure 4. Network throughput with different speed of nodes.

5.2. Average End-to-End Delay (Avg. EED)

Figure 5 illustrates the end-to-end delay comparison of the MRLAM, MP-OLSRv2, and MP-OLSR routing schemes with a variation of node speed between 10 m/s and 60 m/s. It can be seen from the figure that all the schemes show almost similar results when the node speed is till 30 m/s. As the node speed increases from 40 m/s, the end-to-end delay of the conventional MP-OLSR and MP-OLSRv2 routing schemes substantially increases. It is due to the fact that both routing schemes do not consider mobility awareness factor of intermediate nodes in the selection of an optimal route, which leads to increased delay at high-speed node scenarios. The MP-OLSR and MP-OLSRv2 routing schemes select the intermediate nodes based on the multiple node-disjoint and link disjoint metric cost functions, respectively. Therefore, these routing schemes forward packets through a longer route from source to destination, that induces propagation and transmission delays in the network. Moreover, MRLAM scheme maintains and controls end-to-end delay by exploiting EAX metric that minimizes the number of intermediate nodes and number of retransmissions required for reliable packet's delivery on optimal route. The MRLAM utilizes pause time and moving time factor for mobility estimation of nodes from the RWP model. Therefore, it stabilizes the dynamic network to effectively deliver the packet with less time. In addition, using the Q-learning algorithm, the MRLAM selects the intermediate nodes which

have low mobility and better link quality when determining the best path from available alternate paths for forwarding data packets toward the destination that leads to minimized end-to-end delay. Overall, it can be observed that MRLAM reduces the end to end delay by approximately 15% and 10% compared to MP-OLSR and MP-OLSRv2 routing schemes, respectively, at 60 m/s node speed.
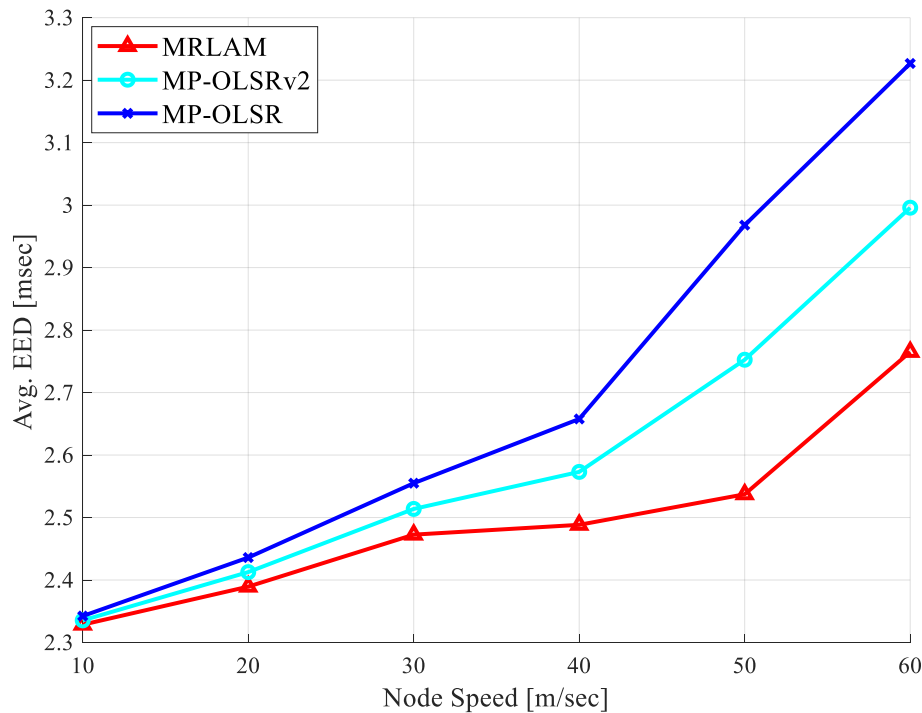

Figure 5. End-to-end delay with different speed of nodes.

5.3. Packets Loss Ratio (PLR)

The MRLAM routing scheme maintains a PLR less than 30% for all scenarios of node speed which is a significant performance improvement for efficient data transmitted in MANETs environment. These results attribute the effectiveness of the proposed scheme to selection of optimal route synchronized with Q-learning process for multiple parameters during the route establishment process. The Q-learning algorithm provides higher rewards to the node which has better link quality with a neighbor node, which reduces the chances of frequent link failure and increases the successful delivery probability of packets. The MRLAM scheme exploits EAX that decreases the number of intermediate nodes in the selection of optimal route, the impact of which reduces the number of dropped packets during data transmission. In addition, EAX decreases the number of trials for data retransmission and reduces link failure probability resulting from the dynamic nature of networks, thereby keeping a lower PLR than other multipath schemes. In establishing a reliable route, the MRLAM scheme avoids the nodes which change their positions frequently and quickly exhaust their battery power. For high-speed node scenarios, it can be seen from Figure 6 that MRLAM scheme achieves better performance compared to both routing schemes, and the improvement percentages are approximately 30.76% and 24.59% compared to MP-OLSR and MP-OLSRv2, respectively.
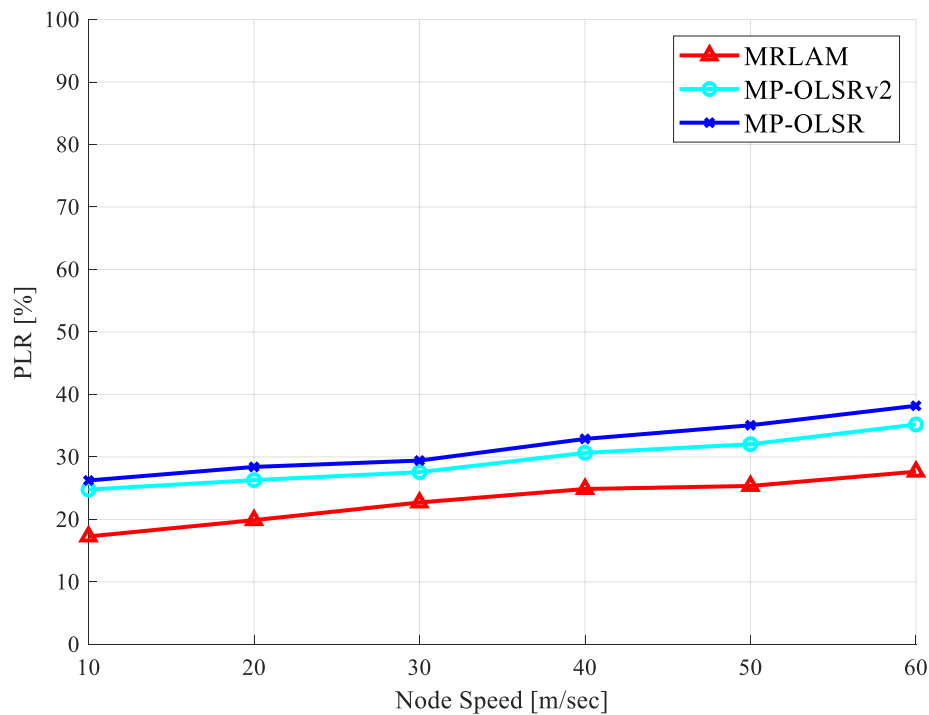
Figure 6. Packet loss ratio with different speed of nodes.

5.4. Energy Consumption

Comparison of node energy consumption during operation time of the network is shown in Figure 7. The result shows that MRLAM performs better than the other routing schemes in terms of energy consumption during path establishment and in exchanging routing topological information as it selects intermediate nodes based on low EAX value for probe messages. Moreover, it has already been established that MRLAM selects the path with less chances of link failure, resulting in lower energy consumption as it forwards data packets toward destination node. If effect, the Q-learning technique provides high rewards value to the nodes which have lower energy consumption, thereby enhancing energy utilization during data transmission. However, as the node speed increases, the energy consumption of nodes increases. Energy consumption increases from 56.67 mAh to 57.26 mAh for MP-OLSR, from 56.43 mAh to 56.926 mAh for MP-OLSRv2, and from 56.15 mAh to 56.72 mAh for MRLAM approximately, when the node speed increases from 10 m/s to 60 m/s. Overall, it can be observed that MRLAM has less energy consumption rating than MP-OLSR and MP-OLSRv2 scheme because the source node forwards traffic flow towards intermediate nodes which have the highest level of energy as opposed to its counterpart routing schemes.
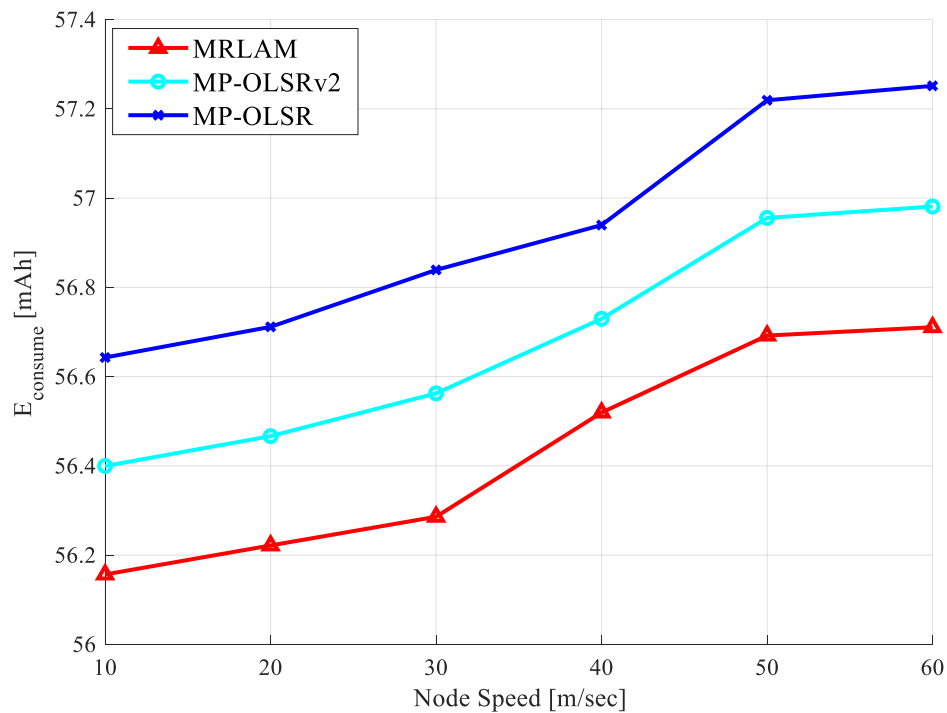
Figure 7. Node energy consumption with different speed of nodes.

5.5. Energy Cost

Simulation results for energy cost per packet are shown in Figure 8, the proposed routing scheme attains lower energy cost due to the node selection mechanism it employs which select only the nodes with highest energy level. Displacement of intermediate nodes in the path causes the variation of energy levels of nodes whilst increasing mobility. The MRLAM scheme also selects nodes with better link quality and lower speed to decrease energy consumption and packet loss ratio, this awareness is not used in the other routing protocols. The increase in a node's movement in the network increases the complexity of network topology for sensing and route computation process. However, the proposed MRLAM routing scheme takes advantage of utilizing Q-learning algorithm along with minimum energy consumption node for transmission packets that helps to reduce the energy cost as well as complexity of the network. The Q-learning algorithm periodically updates the state-action value with learning rate α based on energy consumption of intermediate nodes as well as the number of packets transmitted and received. Moreover, there is reduced network packet flooding and energy cost of packets forwarded to the destination in MRLAM scheme compare with the other schemes. From the perspective of the overall node speeds, critical analysis of the simulation results shows a decrease in energy cost by 33% and 23% as compared to the MP-OLSR and MP-OLSRv2 routing schemes, respectively.
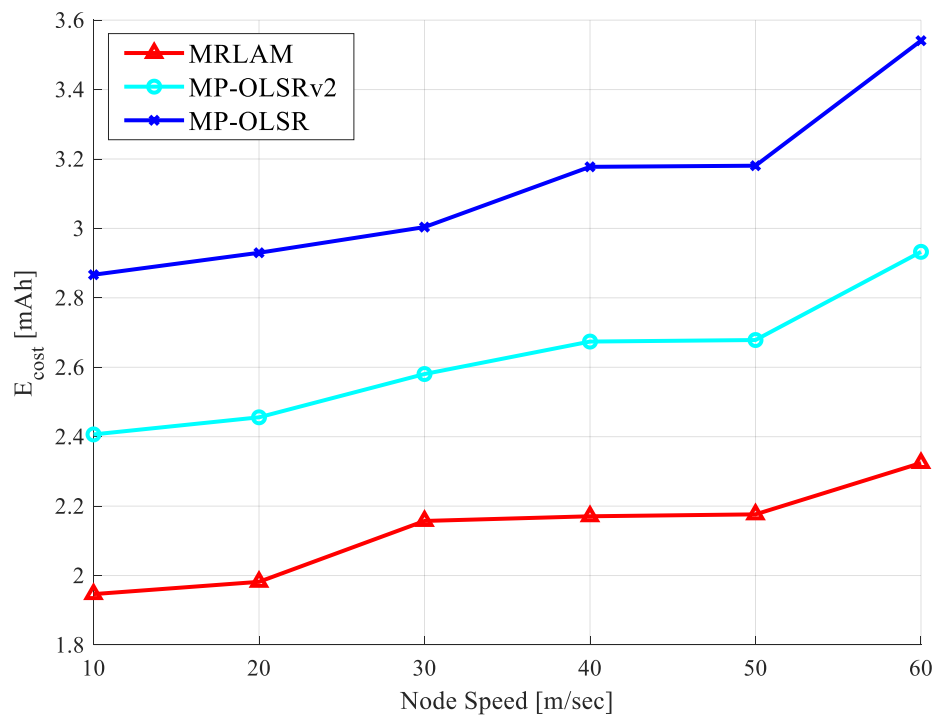
Figure 8. Node Energy Cost with different speed of nodes.

5.6. Convergence Time

The convergence time for all routing schemes is depicted in Figure 9. It can be seen from the figure that all the schemes show similar results when the node speed is up to 30 m/s. As the node speed increases to 40 m/s, the convergence time of the conventional MP-OLSR and MP-OLSRv2 routing schemes substantially increases. This because other routing schemes select the longer path from source to destination nodes based on the link cost function and multiple disjoint metrics, while the MRLAM chooses the shortest path with more stable nodes and better links which reduces data packet transmission delay. In addition, the MRLAM routing scheme utilizes the mobility and link quality status of the mobile nodes, which leads to minimizing the frequent change of network topology. Meanwhile, it uses the Q-learning algorithm which quickly updates the routing information of the network with any changes occurring in the network, and this reduces the convergence time of the network. In the MRLAM routing scheme, fewer intermediate nodes need to converge, thus the load on any given node or communication link is minimized. Therefore, it reduces the calculating costs of intermediate nodes of the paths and improves communication efficiency. Overall, at the node speed of 60 m/s, the proposed MRLAM routing scheme recorded the lowest convergence time of 16.49% and 11.34% in comparison to the MP-OLSR and the MP-OLSRv2 routing schemes, respectively.
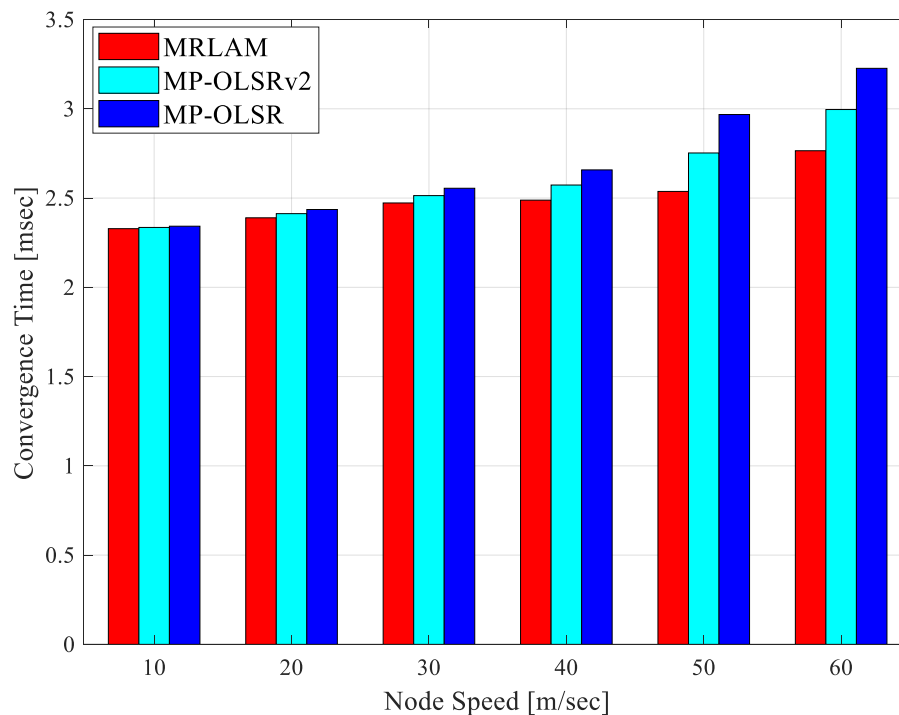
Figure 9. Convergence time of routing schemes with different speed of nodes.

6. Conclusions

This paper represents an evaluation and comparative study of three routing protocols in MANETs environment, conducted under a series of simulations with varying speed of node scenario. A multipath routing scheme called MRLAM is proposed in which the routing decision is made based on residual energy, link quality, and mobility status of the network nodes. In addition, the MRLAM scheme aggregates multiple parameters into a single metric using the mechanics of the Q-learning process to make an optimal routing decision. The MRLAM routing protocol evaluates the status of nodes during the route computation and topology sensing process to determine the optimal routes in a convergence network. Moreover, the proposed scheme has the ability to cope with link failure and sustains the network lifetime by avoiding nodes with lower residual energy and higher mobility in selecting optimal and stable paths between all pairs of source and destination nodes. The simulation results show that MRLAM has approximately 33% and 23% less energy cost per packet compared to MP-OLSR and MP-OLSRv2 routing schemes, respectively. Moreover, results also corroborate the premise that MRLAM attains better throughput in all scenarios of node speed with successfully delivered packets. In addition, the frequent change in the network topology increase the computational complexity of existing MP-OLSR and MP-OLSRv2 routing processes as the route computation process for discovery of new routes becomes more intricate. In the proposed routing scheme, the Q-learning aggregates multiple parameters related to energy, mobility, and link quality into a comprehensive metric to dramatically reduce the complexity and avoid the control overhead caused by separately broadcasting multiple parameters. Furthermore, the results show that the MRLAM scheme decreases the packet loss ratio up to 30.76% and 24.59%, approximately as compared to MP-OLSR and MP-OLSRv2 routing schemes respectively. Overall, the proposed MRLAM scheme evidently outperforms the conventional MP-OLSR and its extended version MP-OLSRv2 routing schemes in terms of throughput, average end-to-end delay, packet loss ratio, and energy consumption rate.

7. Future Works and Challenges

The MRLAM routing scheme with reinforcement by the Q-learning algorithm can be used to improve the QoS not only for MANETs, but also for scenarios of highly mobile nodes, such as drones

and remote-controlled vehicles. Due to the large computation cost of Q-learning algorithms, reducing the energy consumption for large numbers of mobile nodes should be considered for future research. The MRLAM routing scheme can be further extended in other large-scale network deployments and popular multihop wireless network scenarios, including typical WSN, VANETs, and MANET-IoT scenarios. Moreover, the queue length of the network nodes should be considered for the selection of routes to minimize the traffic congestion and packet overhead in the network. This will result in a new scheme that can suit more scenarios and meet the requirements of various applications.

References

1. Yazıcı, V.; Kozat, U.C.; Sunay, M.O. A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management. *IEEE Commun. Mag.* **2014**, *52*, 76–85. [CrossRef]
2. Wang, X.; Li, J. Improving the network lifetime of manets through cooperative mac protocol design. *IEEE Trans. Parallel Distrib. Syst.* **2015**, *26*, 1010–1020. [CrossRef]
3. Huang, C.-F.; Chan, Y.-F.; Hwang, R.-H. A comprehensive real-time traffic map for geographic routing in vanets. *Appl. Sci.* **2017**, *7*, 129. [CrossRef]
4. Khan, A.A.; Rehmani, M.H.; Reisslein, M. Cognitive radio for smart grids: Survey of architectures, spectrum sensing mechanisms, and networking protocols. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 860–898. [CrossRef]
5. Lu, X.; Wang, P.; Niyato, D.; Kim, D.I.; Han, Z. Wireless networks with rf energy harvesting: A contemporary survey. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 757–789. [CrossRef]
6. Tehrani, M.N.; Uysal, M.; Yanikomeroglu, H. Device-to-device communication in 5g cellular networks: Challenges, solutions, and future directions. *IEEE Commun. Mag.* **2014**, *52*, 86–92. [CrossRef]
7. Qamar, F.; Dimyati, K.B.; Hindia, M.N.; Noordin, K.A.B.; Al-Samman, A.M. A comprehensive review on coordinated multi-point operation for lte-a. *Comput. Netw.* **2017**, *123*, 19–37. [CrossRef]
8. Elijah, O.; Rahman, T.A.; Orikumhi, I.; Leow, C.Y.; Hindia, M.N. An overview of internet of things (iot) and data analytics in agriculture: Benefits and challenges. *IEEE Internet Things J.* **2018**, *5*, 3758–3773. [CrossRef]
9. Chayon, H.R.; Dimyati, K.; Ramiah, H.; Reza, A.W. An improved radio resource management with carrier aggregation in lte advanced. *Appl. Sci.* **2017**, *7*, 394. [CrossRef]
10. Udeshi, D.; Qamar, F. Quality analysis of epon network for uplink and downlink design. *Asian J. Eng. Sci. Technol.* **2014**, *4*, 78–83.
11. Zhang, J.; Dai, L.; Li, X.; Liu, Y.; Hanzo, L. On low-resolution adcs in practical 5g millimeter-wave massive mimo systems. *IEEE Commun. Mag.* **2018**, *56*, 205–211. [CrossRef]
12. Liaqat, M.; Noordin, K.A.; Latef, T.A.; Dimyati, K. Power-domain non orthogonal multiple access (pd-noma) in cooperative networks: An overview. *Wirel. Netw.* **2018**, 1–23. [CrossRef]
13. Pham, Q.-V.; Hwang, W.-J. α-fair resource allocation in non-orthogonal multiple access systems. *IET Commun.* **2018**, *12*, 179–183. [CrossRef]
14. Yue, J.; Hu, Z.; He, R.; Zhang, X.; Dulout, J.; Li, C.; Guerrero, J.M. Cloud-fog architecture based energy management and decision-making for next-generation distribution network with prosumers and internet of things devices. *Appl. Sci.* **2019**, *9*, 372. [CrossRef]
15. Gachhadar, A.; Hindia, M.N.; Qamar, F.; Siddiqui, M.H.S.; Noordin, K.A.; Amiri, I.S. Modified genetic algorithm based power allocation scheme for amplify-and-forward cooperative relay network. *Comput. Electr. Eng.* **2018**, *69*, 628–641. [CrossRef]

16. Noordin, K.A.B.; Hindia, M.N.; Qamar, F.; Dimyati, K. Power allocation scheme using pso for amplify and forward cooperative relaying network. In Proceedings of the Computing conference, 2018, London, UK, 10–12 July 2018; Springer: Cham, Switzerland, 2018; pp. 636–647.

17. Hindia, M.N.D.N.; Qamar, F.; Majed, M.B.; Rahman, T.A.; Amiri, I.S. Enabling remote-control for the power sub-stations over lte-a networks. *Telecommun. Syst.* **2019**, *70*, 37–53. [CrossRef]

18. Hindia, M.H.D.N.; Qamar, F.; Rahman, T.A.; Amiri, I.S. A stochastic geometrical approach for full-duplex mimo relaying model of high-density network. *Ad Hoc Netw.* **2018**, *74*, 34–46. [CrossRef]

19. Hindia, M.N.; Fadoul, M.M.; Abdul Rahman, T.; Amiri, I.S. A stochastic geometry approach to full-duplex mimo relay network. *Wirel. Commun. Mob. Comput.* **2018**, *2018*. [CrossRef]

20. Dai, L.; Wang, B.; Yuan, Y.; Han, S.; Chih-Lin, I.; Wang, Z. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Commun. Mag.* **2015**, *53*, 74–81. [CrossRef]

21. Feng, D.; Lu, L.; Yuan-Wu, Y.; Li, G.Y.; Feng, G.; Li, S. Device-to-device communications underlaying cellular networks. *IEEE Trans. Commun.* **2013**, *61*, 3541–3551. [CrossRef]

22. Qamar, F.; Abbas, T.; Hindia, M.N.; Dimyati, K.B.; Noordin, K.A.B.; Ahmed, I. Characterization of mimo propagation channel at 15 ghz for the 5g spectrum. In Proceedings of the 2017 IEEE 13th Malaysia International Conference on Communications (MICC), Johor Bahru, Malaysia, 28–30 November 2017; pp. 265–270.

23. Abbas, T.; Qamar, F.; Ahmed, I.; Dimyati, K.; Majed, M.B. Propagation channel characterization for 28 and 73 ghz millimeter-wave 5g frequency band. In Proceedings of the 2017 IEEE 15th Student Conference on Research and Development (SCOReD), Putrajaya, Malaysia, 13–14 December 2017; pp. 297–302.

24. Qamar, F.; Siddiqui, M.H.S.; Dimyati, K.; Noordin, K.A.B.; Majed, M.B. Channel characterization of 28 and 38 ghz mm-wave frequency band spectrum for the future 5g network. In Proceedings of the 2017 IEEE 15th Student Conference on Research and Development (SCOReD), Putrajaya, Malaysia, 13–14 December 2017; pp. 291–296.

25. Agiwal, M.; Roy, A.; Saxena, N. Next generation 5G wireless networks: A comprehensive survey. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 1617–1655. [CrossRef]

26. Wang, D.-L.; Sun, Q.-Y.; Li, Y.-Y.; Liu, X.-R. Optimal energy routing design in energy internet with multiple energy routing centers using artificial neural network-based reinforcement learning method. *Appl. Sci.* **2019**, *9*, 520. [CrossRef]

27. Cao, Y.; Sun, Z. Routing in delay/disruption tolerant networks: A taxonomy, survey and challenges. *IEEE Commun. Surv. Tutor.* **2013**, *15*, 654–677. [CrossRef]

28. Zhang, H.; Wang, X.; Memarmoshrefi, P.; Hogrefe, D. A survey of ant colony optimization based routing protocols for mobile ad hoc networks. *IEEE Access* **2017**, *5*, 24139–24161. [CrossRef]

29. Haque, I.T. On the overheads of ad hoc routing schemes. *IEEE Syst. J.* **2015**, *9*, 605–614. [CrossRef]

30. Taha, A.; Alsaqour, R.; Uddin, M.; Abdelhaq, M.; Saba, T. Energy efficient multipath routing protocol for mobile ad-hoc network using the fitness function. *IEEE Access* **2017**, *5*, 10369–10381. [CrossRef]

31. Kuo, W.; Chu, S. Energy efficiency optimization for mobile ad hoc networks. *IEEE Access* **2016**, *4*, 928–940. [CrossRef]

32. Bai, F.; Sadagopan, N.; Krishnamachari, B.; Helmy, A. Modeling path duration distributions in manets and their impact on reactive routing protocols. *IEEE J. Sel. Areas Commun.* **2004**, *22*, 1357–1373. [CrossRef]

33. Hurley-Smith, D.; Wetherall, J.; Adekunle, A. Superman: Security using pre-existing routing for mobile ad hoc networks. *IEEE Trans. Mob. Comput.* **2017**, *16*, 2927–2940. [CrossRef]

34. Rosati, S.; Krużelecki, K.; Heitz, G.; Floreano, D.; Rimoldi, B. Dynamic routing for flying ad hoc networks. *IEEE Trans. Veh. Technol.* **2016**, *65*, 1690–1700. [CrossRef]

35. Torrieri, D.; Talarico, S.; Valenti, M.C. Performance comparisons of geographic routing protocols in mobile ad hoc networks. *IEEE Trans. Commun.* **2015**, *63*, 4276–4286. [CrossRef]

36. Gupta, L.; Jain, R.; Vaszkun, G. Survey of important issues in uav communication networks. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 1123–1152. [CrossRef]

37. Clausen, T.; Jacquet, P. *Optimized Link State Routing Protocol (Olsr)*; IETF: Fremont, CA, USA, 2003.

38. Kots, A.; Kumar, M. The fuzzy based qmpr selection for olsr routing protocol. *Wirel. Netw.* **2014**, *20*, 1–10. [CrossRef]

39. Boushaba, A.; Benabbou, A.; Benabbou, R.; Zahi, A.; Oumsis, M. Multi-point relay selection strategies to reduce topology control traffic for olsr protocol in manets. *J. Netw. Comput. Appl.* **2015**, *53*, 91–102. [CrossRef]

40. Xuekang, S.; Wanyi, G.; Xingquan, X.; Baocheng, X.; Zhigang, G. Node discovery algorithm based multipath olsr routing protocol. In Proceedings of the 2009 WASE International Conference on Information Engineering, Taiyuan, China, 10–11 July 2009; pp. 139–142.

41. Joshi, R.D.; Rege, P.P. Implementation and analytical modelling of modified optimised link state routing protocol for network lifetime improvement. *IET Commun.* **2012**, *6*, 1270–1277. [CrossRef]

42. Villasenor-Gonzalez, L.; Ying, G.; Lament, L. Holsr: A hierarchical proactive routing mechanism for mobile ad hoc networks. *IEEE Commun. Mag.* **2005**, *43*, 118–125. [CrossRef]

43. Badis, H.; Al Agha, K. Qolsr multi-path routing for mobile ad hoc networks based on multiple metrics: Bandwidth and delay. In Proceedings of the 2004 IEEE 59th Vehicular Technology Conference, Milan, Italy, 17–19 May 2004; pp. 2181–2184.

44. Sarkar, S.; Datta, R. Mobility-aware route selection technique for mobile ad hoc networks. *IET Wirel. Sens. Syst.* **2017**, *7*, 55–64. [CrossRef]

45. Kanagasundaram, H.; Kathirvel, A. Eimo-esolsr: Energy efficient and security-based model for olsr routing protocol in mobile ad-hoc network. *IET Commun.* **2018**, *13*, 553–559. [CrossRef]

46. Pham, Q.; Hwang, W. Network utility maximization-based congestion control over wireless networks: A survey and potential directives. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 1173–1200. [CrossRef]

47. Wang, Z.; Chen, Y.; Li, C. Psr: A lightweight proactive source routing protocol for mobile ad hoc networks. *IEEE Trans. Veh. Technol.* **2014**, *63*, 859–868. [CrossRef]

48. Bhattacharya, A.; Sinha, K. An efficient protocol for load-balanced multipath routing in mobile ad hoc networks. *Ad Hoc Netw.* **2017**, *63*, 104–114. [CrossRef]

49. Ladas, A.; Deepak, G.C.; Pavlatos, N.; Politis, C. A selective multipath routing protocol for ubiquitous networks. *Ad Hoc Netw.* **2018**, *77*, 95–107. [CrossRef]

50. Li, Z.; Wu, Y. Smooth mobility and link reliability-based optimized link state routing scheme for manets. *IEEE Commun. Lett.* **2017**, *21*, 1529–1532. [CrossRef]

51. Wu, Z.-Y.; Song, H.-T. Ant-based energy-aware disjoint multipath routing algorithm for manets. *Comput. J.* **2008**, *53*, 166–176. [CrossRef]

52. De Rango, F.; Guerriero, F.; Fazio, P. Link-stability and energy aware routing protocol in distributed wireless networks. *IEEE Trans. Parallel Distrib. Syst.* **2012**, *23*, 713–726. [CrossRef]

53. Huang, M.; Liang, Q.; Xi, J. A parallel disjointed multi-path routing algorithm based on olsr and energy in ad hoc networks. *J. Netw.* **2012**, *7*, 613. [CrossRef]

54. Yi, J.; Adnane, A.; David, S.; Parrein, B. Multipath optimized link state routing for mobile ad hoc networks. *Ad Hoc Netw.* **2011**, *9*, 28–47. [CrossRef]

55. Jiazi, Y.; Benoît, P. Multipath Extension for the Optimized Link State Routing Protocol Version 2 (OLSRv2). RFC 8218, Category Experimental, WG MANET, IETF. 2017.

56. Rong, P.; Pedram, M. An analytical model for predicting the remaining battery capacity of lithium-ion batteries. *IEEE Trans. Very Large Scale Integr. (Vlsi) Syst.* **2006**, *14*, 441–451. [CrossRef]

57. Nguyen, T.D.; Khan, J.Y.; Ngo, D.T. A distributed energy-harvesting-aware routing algorithm for heterogeneous iot networks. *IEEE Trans. Green Commun. Netw.* **2018**, *2*, 1115–1127. [CrossRef]

58. Ge, X.; Ye, J.; Yang, Y.; Li, Q. User mobility evaluation for 5g small cell networks based on individual mobility model. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 528–541. [CrossRef]

59. Xu, W.; Jiang, M.; Tang, F.; Yang, Y. Network coding-based multi-path routing algorithm in two-layered satellite networks. *IET Commun.* **2017**, *12*, 2–8. [CrossRef]

60. Kiran, Y.V.; Venkatesh, T.; Murthy, C.S.R. A reinforcement learning framework for path selection and wavelength selection in optical burst switched networks. *IEEE J. Sel. Areas Commun.* **2007**, *25*, 18–26. [CrossRef]