Notice of Violation of IEEE Publication Principles

"Clustering Schemes in MANETs: Performance Evaluation, Open Challenges, and Proposed Solutions,"
by T. Rahman, I. Ullah, A. U. Rehman and R. A. Naqvi,
in *IEEE Access*, vol. 8, January 2020, pp. 25135-25158

After careful and considered review of the content and authorship of this paper by a duly constituted expert committee, this paper has been found to be in violation of IEEE's Publication Principles.

This paper contains a figure (Figure 18) that was copied from the paper cited below. The figure was copied without attribution (including appropriate references to the original author(s) and/or paper title) and without permission.

"BICSF: Bio-Inspired Clustering Scheme for FANETs,"
by A. Khan, F. Aftab and Z. Zhang,
in *IEEE Access*, vol. 7, March 2019, pp. 31446-31456

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 1, 2020, accepted January 25, 2020, date of publication January 30, 2020, date of current version February 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2970481

Clustering Schemes in Manets: Performance Evaluation, Open Challenges, and Proposed Solutions

TAJ RAHMAN[1,*], INAM ULLAH[2], ATEEQ UR REHMAN[2,3], AND RIZWAN ALI NAQVI[4,*]
[1]Department of Computer Science and IT, Qurtuba University of Science and Technology, Peshawar 25000, Pakistan
[2]College of Internet of Things (IoT) Engineering, Hohai University (HHU), Changzhou Campus, Changzhou 213022, China
[3]Department of Electrical Engineering, Government College University, Lahore 54000, Pakistan
[4]Department of Unmanned Vehicle Engineering, Sejong University, Seoul 05006, South Korea

Corresponding author: Rizwan Ali Naqvi (rizwanali@sejong.ac.kr)

This work was supported in part by the National Research Foundation of Korea under Grant NRF2019R1A2B2012337, and in part by Sejong University Faculty Research Fund.

Taj Rahman and Rizwan Ali Naqvi are co-first authors.

ABSTRACT Mobile ad hoc networks (MANETs) have recently received a lot of attention from researchers due to their ease of deployment and versatility. Clustering is one of the most efficient ways to coordinate routing in MANETs to improve network performance using different attributes and metrics. In this paper, the authors have extensively studied clustering schemes and divided the schemes into multiple types based on the Cluster Head (CH) selection criteria, which provides a good understanding of how each type of clustering algorithm differs from each other. The authors analyzed the performance of existing schemes based on the quality of service (QoS) metrics. Based on findings, the authors clarified some important tradeoffs between QoS metrics and also established some important factors influencing the efficiency of clustering schemes. These studies also contain open research challenges and proposed solutions to improve the performance of clustering schemes.

INDEX TERMS Clustering, cluster head, MANETs, protocol and schemes.

I. INTRODUCTION

Mobile ad hoc networks (MANETs) are evolving for real-time implementation in numerous fields of military and civilian. MANETs consist of nodes that on request, freely and arbitrarily organize a complex and temporary network topology without any infrastructure support [1], [2]. The ability of MANETs to cover a large transmission area to forward information from one node to another using multiple hops; make it suitable for deployment in the event of an emergency and natural disasters.

Clustering can overcome the problems of routing protocols and assists in improving scalability [3]. Clustering in MANETs provides a stable way to allocate resources efficiently and it also ensures stability in network structure by providing a hierarchical environment [4]. The main characteristics of MANETs can be achieved by using a clustering-based network. A large network is usually divided into smaller subgroups. These sub-groups are called clusters as shown in Fig. 1. In a cluster, a CH is selected among all the available nodes. A CH is used to manage the activities related to the cluster and can access all the member nodes [5], [6]. The CH has to control all the nodes in a network; therefore, the selection of CH becomes important for network survivability. The clustering algorithm performs two main functions namely cluster formation and cluster maintenance. Cluster formation involves the selection of CHs among nodes to form the hierarchical network. CH manages cluster activities like, managing cluster process, updating route information and new route discovery. In the cluster maintenance stage, multiple cluster related activities take place. These activities include CH re-election, cluster, inter and intra based clustering communication and efficient selection of the shortest path. A node known as gateway node provides a connection between clusters and is a member of more than one cluster as shown in Fig. 1. Clustering in MANETs divides the overall topology of the network into a random number of clusters without overlapping them with each other [7], [8]. This topological structure will reduce the overall number of control messages exchanged between the nodes and will

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

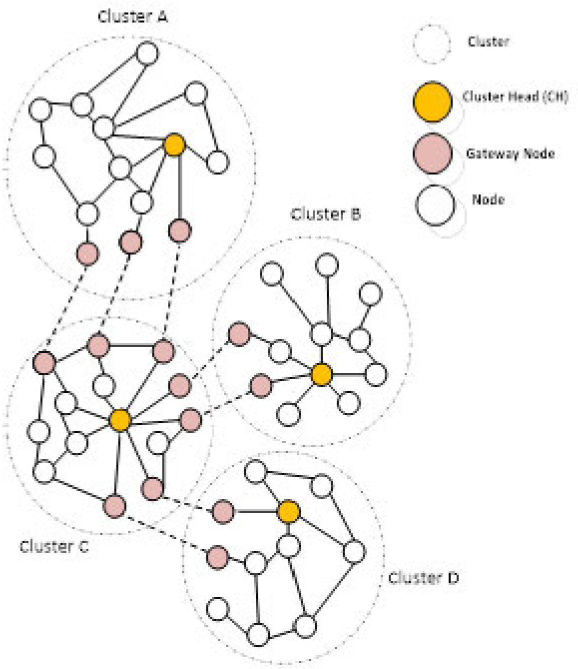

FIGURE 1. **Clustering in MANETs.**

enhance the overall communication capacity in large scale MANETs.

The cluster design in the clustering algorithm depends upon a few performance parameters and metrics. Some techniques and mechanisms can be used to fulfill design objectives by using these parameters. The designing stage of a clustering scheme depends upon network topologies, overall energy consumption, node mobility, and CH selection. These are a few of the factors which can directly affect the Quality of Service (QoS) of a network. In the past, the researcher's main emphasis was to design such clustering mechanisms to minimize the challenges related to the cluster formation and maintenance. In the existing literature, researchers have considered a few areas of clustering schemes without extensive performance comparison and evaluation [3]–[5], [7], [9]–[18].

The main contributions of our paper on clustering based MANETs are as follows:

1) In our paper, not only the comprehensive study of clustering schemes are considered but also schemes are divided into multiple types based on CH selection criteria which gives a proper understanding of how each type of clustering algorithm differs from one another. For a better understanding of every type of clustering scheme, an analysis table is provided for every category which consists of CH selection methodology, advantages, disadvantages, and drawbacks of every scheme.
2) Within every category, a direct comparison is made between different algorithms to illustrate their performance to each other.
3) Performance of each scheme is summarized in an evaluation table by taking QoS metrics.
4) Performance is evaluated based on QoS metrics to clarify which type of clustering scheme outperforms another and in what aspect.
5) Based on our study and evaluation, we have discussed some important performance tradeoffs in QoS metrics for MANET and also discuss some factors which can cause limitations and affect the performance. Our study also contains open research challenges of cluster-based MANET.
6) We have also proposed a cross-layer clustering framework and hybrid self-organization clustering model as a solution to improve QoS and minimize the research challenges of cluster-based MANET.

This paper outlines a comprehensive survey of clustering protocols proposed MANETs. The authors attempt to review and discuss the different aspects of clustering by considering multiple metrics and attributes. The goal of this survey is to provide a sound knowledge along with details of clustering algorithms to give readers a proper understanding of the clustering-based environment in MANETs. We also highlight the strengths and weaknesses of the existing algorithms according to their performance in MANETs that can help to identify alternative solutions and provides an opportunity for the application developers to adopt suitable strategies by using the comparison of multiple approaches based on the clustering mechanism. This paper is organized as follows: Section II explains about motivation and objective of this survey. Section III presents the taxonomy of the clustering algorithm in MANETs and also contains comparison along with performance analysis of each scheme. In Section IV, we provide a comparative performance evaluation of all schemes based on QoS metrics and discuss some important tradeoffs between QoS metrics. Section V summarizes the open challenges of MANET and in Section VI we have proposed two solutions to minimize the challenges of MANET. Section VII presents the comparison of the recent approaches regarding MANETs and finally, Section VIII summarizes and concludes this paper.

II. OBJECTIVE AND MOTIVATION

For cluster-based MANET, several survey literature already existed. In literature [10]–[17], [19], researchers have considered few numbers of clustering schemes without extensive performance comparison and evaluation. In [14] a comparative analysis is shown in a table which consists of advantages and disadvantages of every scheme but there is no direct comparison between different algorithms to illustrate their performance concerning each other. There is also no recommendations and solution to minimize the disadvantage. Similarly, in [15] comparison between different types of clustering schemes does not provide any knowledge about how to evaluate the performance and which type of clustering scheme is better as compare to another type. The study [16], [20], [21] provides good knowledge about power-efficient approaches in MANETs but the comparison is only performed within the schemes of each approach. This survey study lacks a comparison between different approaches to illustrate the performance of one approach concerning others. This survey suggests future research directions for researchers but does not provide any solution about how to minimize the limitations and open issues mentioned in the survey. A study [17] provides an overview of hybrid mechanisms for MANETs with future recommendations for researchers but does not provide any solution to minimize the limitations.

The above-mentioned existing literature provides us a motivation that we must consider the limitations of existing survey studies as our objectives. The first objective is to provide better performance comparison and analysis of schemes as compared to existing literature. The second objective is to consider proper criteria of performance evaluation as in a few of the existing literature evaluation criteria is not mentioned. In our study, we consider the QoS criterion to evaluate the performance of existing schemes. Most of the existing literature only evaluates the performance of existing schemes based on a few performance metrics but there is a limited explanation about the dependencies of one metric to another and how it can affect the overall performance of MANET. Therefore, our third objective is to explain the tradeoffs between QoS metrics for cluster-based MANET and its effect on overall performance. Most of the existing literature provides some open research challenges but very limited information is provided about how to minimize those challenges. Therefore, in this study, we have proposed a Cross-layer clustering framework and hybrid self-organization clustering model as a solution to improve QoS and minimize the research challenges of cluster-based MANET.

III. TAXONOMY OF CLUSTERING ALGORITHM

The main goal of the clustering algorithm is the selection of optimal CH based on the election rule-keeping in mind the performance metrics. Therefore, in this survey, we categorize clustering algorithms in MANETs into different types based on the criteria of CH selection and how the nodes are grouped to form a cluster. Fig. 2 shows the different types of clustering algorithms. All these types have their way of CH selection and in this section, we will discuss them in detail.

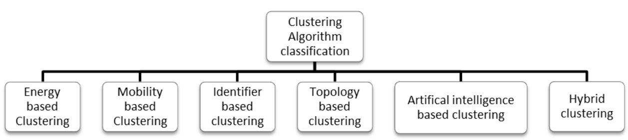

FIGURE 2. Classification of the clustering algorithm.

A. ENERGY-BASED CLUSTERING

In energy-based clustering, the energy level of a node is the key parameter for the selection of CH. As CH manages the cluster and performs cluster management so it causes excessive energy consumption which affects the network performance. Therefore, the power constraint of a node can be a useful parameter to select a CH because it directly influences the overall lifetime of the network.

1) FLEXIBLE WEIGHTED CLUSTERING ALGORITHM BASED ON BATTERY POWER (FWCABP)

FWCABP [22] utilizes multiple factors for the CH election. These factors include mobility (M_i), the degree of nodes, level of battery power (BP) and link connectivity coefficient (LCC). The basic aim of this algorithm is to prevent those nodes from being elected as a CH who have low battery power. FWCABP is adopted based on the size of the cluster and the battery power of the node is used for the maintenance. It minimizes overhead and decreases the overall number of clusters.

The cluster formation phase comprises of neighbor identification procedure and after identification, every node broadcasts a HELLO message to its neighboring nodes. This message informs the neighboring node about the status of a particular node. After receiving this message each node builds its neighbors list and updates it after a particular time interval. The combined weight (W_i) is calculated using equation (1).

$$W_i = w_1 * \Delta sp + w_2 * LCC + w_3 * BP + w_4 * M_i \quad (1)$$

Here $w_1 + w_2 + w_3 + w_4 = 1$. Node with the lowest weight is selected as CH. There are some cases for the initiation of the Cluster maintenance phase. These cases include new node arrival in a cluster coverage area or movement of a node outside its cluster coverage area. The cluster maintenance phase is also needed if a power level of a CH declines lower than a predefined threshold value. A large exchange of message packages can degrade the performance of the network because it causes an increase in network traffic during the CH selection process. FWCABP is a demand-based algorithm which is executed only when a node is moving and changing its relative distance between other nodes and CH.

2) ENHANCE CLUSTER-BASED ENERGY CONSERVATION (ECEC)

ECEC [23] is an enhanced version of Cluster-based Energy Conservation (CEC) algorithm. It ensures minimum connectivity of nodes which results in increased network lifetime. ECEC provides the ability to identify redundant nodes as a result of conserve energy. To select CH and gateway nodes, every node broadcasts a discovery message containing its ID, cluster-ID, position, velocity, estimated lifetime and transmission rate. In the cluster formation phase, CH is selected based on the value of estimated energy. The node having the highest estimated energy value is selected as CH. Gateway nodes are elected after the election of CHs to have a connection between clusters. ECEC uses a redundant removing technique to remove multiple gateway nodes between adjacent clusters. This technique selects the shortest and more efficient path to connect the adjacent clusters and remove some of the gateway nodes, as a result, this method provides an energy-efficient solution by conserving energy. This scheme provides a better network lifetime because of its feature of low power consumption at every node. However, ECEC exchanges more overhead to select the CHs and gateway nodes.

3) STRENGTH-BASED ENERGY EFFICIENT ALGORITHM (SEEA)

In SEEA [24] a node with the highest energy level is selected as a CH. The selected CH deals with data and processing. To propose an efficient algorithm for transmitting information without any loss, several constraints like energy consumption during transmission of packets, transmission power control, link weight, and maximum battery are considered. By using these constraints, network efficiency increases. To handle the case when the power of CH keeps on decreasing, a table for energy and signal level is maintained so select CH based on this. The distance from the source to destination with the node having the highest energy is calculated for packet transfer. An alternate path is chosen as a backup in case of the node running out of energy. SEEA saves the maximum energy of the node which will cause an increase in the lifetime of the network.

4) ENERGY EFFICIENT ALGORITHM USING MAX-HEAP TREE (EEAMHT)

In EEAMHT, the structured cluster is formed by using a max-heap tree [25]. Every node is assigned with an index number as per its energy level. The node which has the highest energy level will become a CH and max-heap tree root. All the other nodes will come under the CH forming a tree. Communication between CHs will occur directly without any gateway node. Multi-point relay (MPR) is used for communication between CHs. Network congestion is reduced using MPR, thus saving the energy of remaining CH. Connectivity between cluster members and CH is single hop.

During the cluster formation phase, the node having the highest energy will be at the root position of max-heap and will serve as a CH. The tree balancing algorithm will select the new CH if the energy of CH decreases as compared to other nodes. Using tree balancing, the node at the highest energy will become a new CH. CH can communicate directly with other member nodes but this communication is not direct as it requires CH assistance for intra-clustering-based communication. In EEAMHT CH always acts as an active node while the member node can switch its mode from active to sleep depending upon situation and workload. In a case when a node wants to communicate with a sleeping node, CH can control and instruct the sleeping nodes by sending an indicator message. If a sleeping node receives an indicator message from CH it means CH wants the node to switch it from sleep to active mode. A sleeping node in reply sends an acknowledgment message back to CH after switching its state which indicates it's ready for communication again. It can also switch its state back to sleep mode if it wants after communication with member nodes. The battery of CH will never be at a lower level so the cluster working does not affect as a result network is stable and has a better life.

5) ENERGY EFFICIENT CLUSTERING APPROACH FOR MANETS (EECA)

EECA [26] is proposed to minimize energy consumption. EECA consists of three phases. The basic objective of these phases is to initiate and maintain the clustering procedure. These three phases are neighbor discovery, cluster formation, and cluster maintenance respectively. In the neighbor discovery phase, every node broadcasts a query message and waits for a reply message until the timer expires. If no reply message receives within the time limit, it will retransmit the query message. In the second phase which is cluster formation, every node transmits a Hello message with its energy level to its neighboring nodes. Each node compares it's signal status, energy and mobility level with the value of neighboring nodes. The node with the lowest weight value among all nodes declares itself as a CH. During the cluster maintenance phase, the CH election takes place when a weight value of a CH becomes less than the pre-defined threshold value. Each node compares its weight value with its neighboring nodes and CH. If a certain node has the lowest value among all of its neighboring nodes, then it declares itself as a CH. EECA reduces the overhead and improves the procedure of the re-clustering process. It also extends the lifetime of a network but involves several message transmissions in the cluster formation and maintenance stages, therefore, it can cause an increase in the consumption of bandwidth and network traffic.

a) PERFORMANCE COMPARISON

To compare the schemes, we consider different metrics which are; Efficiency level that represents how effectively a scheme works by utilizing the available resources. The load and congestion metric represents the level of load on CH while performing cluster formation and management phase. CH change metric represents how frequently CH changes during the clustering procedure. The mobility metric represents the level of a scheme to adopt itself with variable speed of node. The stability metric represents how stable the performance of the scheme is during topological changes in MANETs. The analysis of energy-based algorithms represented in Table 1 shows the CH selection criteria of each scheme with advantages and disadvantages. From the original simulation results of the schemes, we evaluate the performance comparison as shown in Fig. 3. SEEA and ECEC are more stable as compared to EECA, FWCABP, and EEAMHT as depicted in Fig. 3. In FWCABP CH change is moderate but in MPGC, ECEC, and EEAMHT CH change are low. The possibilities of CH overlap in EECA and FWCABP is high as compare to other schemes. Load and congestion in FWCABP are high as compare to other schemes. The efficiency level of ECEC and SEEA is high as compare to EECA and FWCABP.

B. MOBILITY-BASED CLUSTERING

In mobility-based clustering, the mobility level of a node is the parameter used for the selection of a CH as a node

TABLE 1. Analysis of energy-based clustering algorithm.

Energy based clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
FWCABP [14]	Lowest energy value	Low clustering overhead	Increases network traffic during the CH selection.
ECEC [15]	Highest energy value	This algorithm reduces power consumption.	Minimum connectivity between nodes make it difficult to find an alternative link.
SEEA [16]	Highest energy value	Improve scalability by increasing the lifetime of a network and reduce the time delay.	Lead to an increase in network traffic.
EEAMHT [17]	Highest energy value	Minimizes power consumption while maximizes network lifetime	Delay at Inter-cluster communication affect the efficiency of the network.
EECA [26]	Highest energy value	Reduce energy consumption and extends the network lifetime.	Takes more exchange of control messages during clustering procedure.

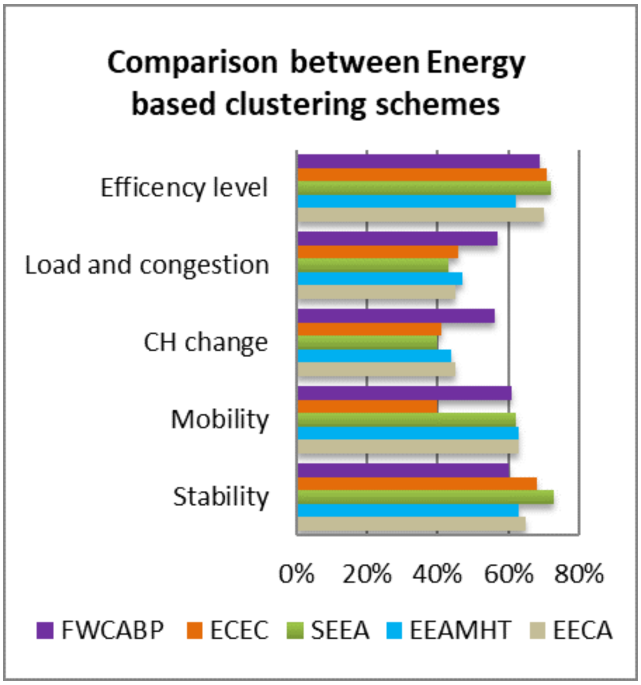

FIGURE 3. Performance analysis of the energy-based scheme.

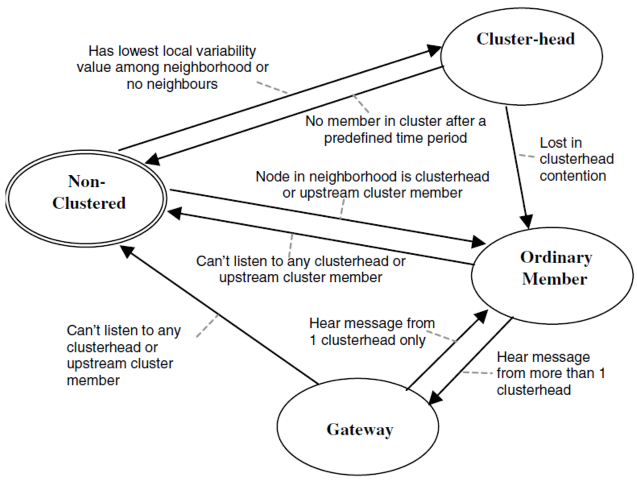

FIGURE 4. MobDHop node transition stage [27].

in MANETs are capable of moving randomly so in the clustering-based network it is very important to keep an eye on mobility level of every node inside a cluster. Therefore, the mobility level of a node is a vital constraint that can directly affect the efficiency and stability of a network so it can be a useful parameter to select a CH.

1) MOBILITY BASED D-HOP CLUSTERING ALGORITHM (MOBDHOP)

Clustering algorithms are mostly designed to form two-hop clusters that are not useful in the case of very large-scale networks. To make clustering more useful for large scale area MobDHop is proposed [27]. The MobDHop node transition stage is shown in Fig. 4. In the MobDHop algorithm, the network is divided using relative mobility metric into d-hop clusters. The number of CH is reduced by one hop radius to that CH creates d-hop clusters. The relative mobility metric is measured by varying the distance between two nodes over time.

The MobDHop algorithm needs to consider the local stability, the variation of estimated distance over time, the relative mobility between nodes, the estimated distance between nodes and the estimated mean distance. The difference in distance between two nodes at two successive intervals of time represents the relative mobility between these nodes. The successful packet received among nodes is calculated by using the probable distance between neighborhood nodes. The stability of each node is calculated which is a helpful metric in the selection procedure of a CH. The stability represents the relative mobility among each node. The slight difference in probable distance between nodes is due to the variation in relative mobility between two nodes. The estimated mean distance is considered which is a metric used to decrease the variation in probable distance.

The MobDHop cluster formation process consists of two stages. These stages are the discovery stage of the node and the merging stage in the cluster. In the discovery stage, the clusters that have the same direction and speed are grouped in the same cluster. In merging stage clusters join individual nodes to a cluster or they merge. The merging process will be successful only if the newly formed cluster achieves a required level of stability. The maintenance of the cluster is initiated when a node joins or leaves the cluster. MobDHop provides an objective of multi-hop-based communication to support a larger area. Based on group mobility pattern, the number of clusters is minimized therefore MobDHop keeps the number of overheads as low as possible to achieve a stable network by keeping the minimum information exchange during cluster formation, CH changes, and handovers.

2) CLUSTER-BASED MOBILITY CONSIDERED ALGORITHM (CMCA)

CMCA is proposed based on the load balancing approach in CHs [28]. Node mobility is taken one of the important characteristics of this clustering algorithm. Node mobility is used to design the cluster and appropriately placing a node. The node having the least mobility is selected as CH. The CH of each cluster is a local manager for that cluster which is responsible for data forwarding, inter-cluster routing and another complex task essential in a network. CMCA consists of three parts. Cluster formation being the first part. The clusters are categorized based on their mobility speed. Node mobility of each node is calculated and then that node is added in a cluster according to their categories. The cluster gateway during the cluster formation stage reduces the transmission overhead. The second part is used for CH selection. The node that has the least mobility value inside a cluster is selected as a CH. The CH is an identity label of a cluster. The last part consists of finding a path for the nodes to the desired destination for both inter and intra clustering communication.

3) STABLE CLUSTERING SCHEME FOR HIGH MOBILITY (SCHM)

In the SCHM algorithm [29] a node has the following three roles at any moment namely cluster head, a cluster member, and undecided. All the nodes in a network are initially in the undecided state. In the undecided stage, the node is still in a search of joining a cluster. The Regional Average Relative Speed (RARS) reveals the relative stability of a node to its neighboring nodes and can be calculated as:

$$v_i\text{−}RARS = \frac{\sum_{j=1}^{k} v_{ij}}{K} \tag{2}$$

where v represents the relative speed of a node. Every node calculates its RARS and sends it to its 1-hop neighboring nodes. Every node then compares the RARS with its own and if its own RARS has the smaller value then it will become CH by sending a message of cluster claim. Then the neighboring nodes join the new cluster and periodically send a message of cluster member to inform CH about their presence. SCHM also solves the issue of connection loss caused by the mobility of node by making the node to pick up a new CH having maximum predicted stay time and avoids the possible connection break.

4) MOBILITY PREDICTION BASED CLUSTERING ALGORITHM (MPBC)

MPBC [30] proposed for independent and random moving nodes by using the estimation of the speed of nodes relative to each other. Every node has information about the relative speed. MPBC consists of two stages; clustering and maintenance stage. During the clustering stage, each node compiles it's neighbor's list by broadcasting the Hello packets. The node's average relative speed is measured by exchanging of Hello packets between neighboring nodes which can be represented by using the equation giving below:

$$\overline{v_i} = \frac{1}{|N_{e,i}|} \sum_{j \in N_{e,i}} v_{i \to j} \tag{3}$$

Here $|N_{e,i}|$ represents the number of nodes in $N_{e,i}$ and the relative speed $v_{i \to j}$ is updated whenever receive a Hello packet from j nodes. Based on the relative mobility level, CH is selected and a node that has the lowest relative mobility is considered as the best candidate for CH. The cluster maintenance stage of MPBC solves the problem of the relative movement of a node by using a prediction-based method. The cluster maintenance stage takes care of cases such as a node that comes within the reach of two CHs and it also provides a good solution to these situations by selecting the connection with a proper cluster-based relative mobility. MPBC provides a long average lifetime of CH which results in cluster stability.

5) GROUP MOBILITY BASED CLUSTERING ALGORITHM (GMCA)

In [31], the movement of nodes creates group mobility because all nodes follow such a pattern of movement which is analogous as a result they form a group. The nodes are grouped according to their pattern of relative mobility (M). The relative mobility of node j concerning node i can be shown in Fig. 5. The relative mobility in GMCA represents by using equation (4).

$$M_{j,i}(t) = 10lg\frac{Df_{j,i}(t)}{DP_{j,i}(t)} \tag{4}$$

Here $Df_{j,i}(t)$ is the relative distance between node i and j while $DP_{j,i}(t)$ represents predicated relative distance between node i and j. The stability of the structure is achieved by performing group emergence and group partitioning according to the change in topology and mobility pattern of a specific node. A node in a group is a powerful candidate of CH if its

TABLE 2. Analysis of mobility based clustering algorithm.

Mobility based clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
MobDHop [27]	Mobility of nodes relative to each other (varying distance between nodes over time).	Achieve relative mobility and improve the stability taking into account group mobility pattern by minimizing the number of clusters.	Due to high dynamics more overhead occur in relative mobility.
CMCA [28]	Least mobile node selected as CH.	Achieve scalability, robustness to in large scale networks.	More hop count inside the cluster can affect CH efficiency.
SCHM [29]	Average relative speed of a node.	Provides more stable clustering in high mobility along with the solution to handle connection break up.	Require more energy to exchange control messages.
MPBC [30]	Lowest relative mobility.	Due to mobility prediction, it is possible to avoid some unnecessary cluster merges.	The less average lifetime of CH.
GMCA [31]	Relative lower mobility with great energy reserve and higher connectivity.	Enhance network stability, scalability and provides better performance during CH change.	Change in CH will affect during a mobility at high speed.

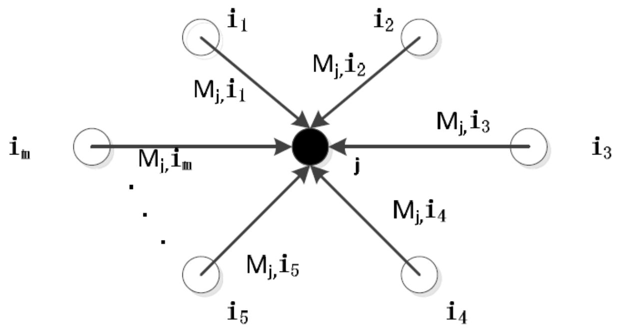

FIGURE 5. Relative mobility between nodes in GMCA [31].

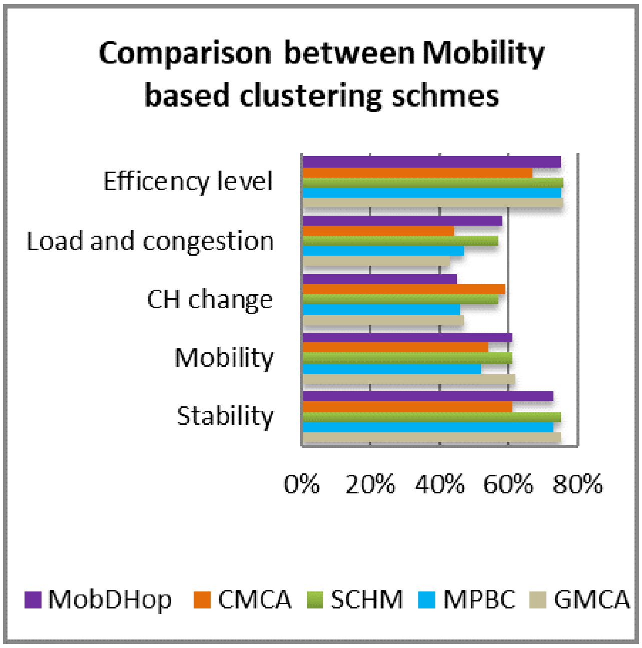

FIGURE 6. Performance analysis of mobility-based schemes.

mobility pattern is relatively slow as compared to other nodes and possesses good resources of energy as well as it contains stable connectivity with group nodes.

A node with the lowest change in relative mobility as compared to its neighboring nodes is selected as a CH. A CH invites every neighborhood node by sending a message to join a cluster. In reply, a node will respond by sending an acceptance message to CH. Nodes can transmit and receive packets to their neighborhood nodes that are in connection with CH. The mobility of each node is a metric used for the establishment and maintenance of a cluster. The mobility pattern of each node is estimated by using the consecutive packets received from neighborhood nodes.

a) PERFORMANCE COMPARISON

The analysis of mobility-based algorithms illustrated in Table 2 shows the CH selection criteria of each scheme along with the advantages and disadvantages of every scheme. Fig. 6 illustrates the performance comparison between these schemes. The stability level of CMCA is average as compared to MobDHop SCHM, MPBC and GMCA while GMCA and SCHM are more stable as compared to the disadvantages of every scheme. Fig. 6 illustrates the performance comparison between these schemes. The stability level of CMCA is average as compared to MobDHop, SCHM, MPBC, and GMCA while GMCA and SCHM are more stable as compared to MobDHop and MPBC. Overall CH change in all these schemes is low but in CMCA and SCHM, CH change is relatively high as compared to other schemes and the possibility of cluster overlapping is possible. All these schemes show better performance in low mobility. Load and congestion are balanced in GMCA and SCHM. The efficiency level of EWCA, EMSWCA and NWBCA are high as compared to DWCA and FWCA.

C. IDENTIFIER BASED ALGORITHM

The identifier (ID) based clustering algorithm is the algorithm in which every node inside the network has a unique ID. The node with the lowest or highest ID is chosen to be CH.

1) CLUSTER HEAD SELECTION AND UPDATE ALGORITHM (CHSUA)

CHSUA [32] is an ID-based clustering scheme consists of two parts namely cluster formation and cluster update. In the cluster formation stage, every node is assigned a unique ID and then it broadcasts a HELLO message regarding its status with neighboring nodes. This information on status is then stored in the neighborhood info list. The node that has the lowest ID is selected as a CH. An ordinary node that is already a part of the cluster will not participate again in the cluster formation stage and it is allowed to belong to multiple clusters. If a node belongs to multiple clusters at a time then it may act as a gateway node which then manages the relaying of data among clusters. In the cluster update, the performance of CH is monitored by updating the stability factor (S). This stability factor consists of a unique node identifier, node status information, node location information, and IP.

All nodes transmit their S value to their CH. The cluster members are arranged in increasing order of stability factor by CH and reassign the node IDs according to the ascending order of S. Smaller the stability factor, it is more likely a node will be selected as a CH. A node with the lowest S value is given the lowest ID because such a node indicates that it has consistent connectivity along with high channel quality, low mobility and has sufficient battery power. If nodes possess the same S value then the IDs are assigned according to the original ID sequence. A node having the lowest ID is assigned a new gateway ID by its CH. The CH broadcasts new IDs to all nodes in the cluster and then nodes update their IDs and initiate the cluster formation process.

2) LOWEST ID WITH ADAPTIVE ID REASSIGNMENT ALGORITHM (LIDAR)

LIDAR [33] is a clustering algorithm consists of the clustering and maintenance phase. At clustering stage node IDs are assigned to every node. A node having the lowest ID is selected as a CH that controls the topology information and communication between nodes. A gateway node is a node that connects two clusters and exists in the range of two clusters. All other nodes in a cluster act as ordinary nodes. Every node broadcasts its IDs by transmitting the HELLO control message called HELLO period (HP). In the cluster maintenance phase, the LID algorithm is initiated at the end of every HP. It adjusts formations of the cluster as per current topology status. Mobility and battery power of node is the metrics considered in the maintenance phase. Each node calculates its weight function value which is consists of battery life and mobility rate. This weight value is broadcast by every node to their local CH. The CH arranges the weight value of every node in descending order and re-assign an

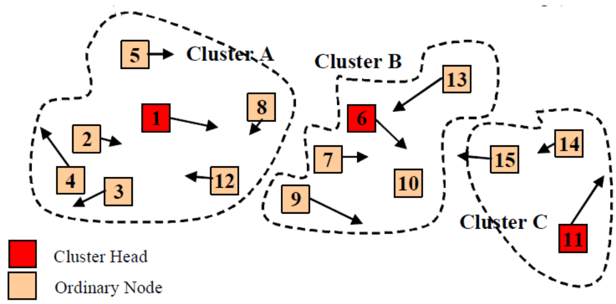

FIGURE 7. Status of nodes using LIDAR [33].

ID of every node according to their weight value as shown in Fig. 7. A node having the highest weight value has a lower ID assigned because these nodes have a low mobility rate and possess a sufficient amount of power. A node with the lowest ID has the highest possibility of being elected as CH in the next CH election procedure. After that CHs send the new IDs to their respective members and nodes update their ID values. LIDAR balanced the load in a clustering environment and provides stable clusters along with acceptable power consumption among nodes.

3) IMPROVED LOCATION AIDED CLUSTER-BASED ROUTING PROTOCOL (ILCRP)

ILCRP [34] algorithm has the utility of GPS. A CH selects its members that are m-hops away. Initially, all nodes start in undecided stage and they become a part of a cluster if their distance from a CH is m-hops. ILCRP consists of two main stages. In the first stage, the CH election takes place. Every node initially broadcasts a HELLO message. This message consists of the ID of a node along with its information about the location using a global position system. All the nodes broadcast their node value to m-hop neighbors using the INFO message. CH is selected based on node value and a node that has the highest node value is a strong candidate to be selected as CH. All nodes arrange a table which contains neighbor information. This table contains the ID of neighboring nodes along with their node values and information about its location. Nodes use a cluster table for inter-cluster routing. CH stores IDs of adjacent CHs and uses IDs of gateway node to reach them. In the second stage, clustering maintenance takes place. The mobile nature of nodes demands a cluster be configured and organized according to the changing environment. The mobility of CH, node mobility and node value reduction of the CH are the three main scenarios in cluster reconfiguration. By using these scenarios ILCRP increases the stability of clustering and performance of CH.

4) CLUSTER BASED ROUTE DISCOVERY TECHNIQUE FOR ROUTING (CBRD)

In CBRD [35], nodes are grouped to form clusters and CH manages these clusters. The Hello packet is transferred between nodes that contain node weight and ID. Based on the lowest ID, a CH is selected. Before the communication

TABLE 3. Analysis of ID-based clustering algorithm.

ID-based clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
CHSUA [32]	Lower ID	Increase QoS by reducing both CH update and change events.	Sensitive to high-speed node mobility.
LIDAR [33]	Lower ID	Maximum scalability and extended lifespan of the network.	Speed of CH to perform its operation is slow it will cause delay.
ILCRP [34]	Highest node value	Increases the delivery ratio and makes the route loop-free.	End-to-end delay increases due to an increased number of nodes.
CBRD [35]	Lower ID	Ensure better QoS by finding the shortest path between source node and destination.	A lot of burden on CH, can cause power drainage and delay.
OSCA [36]	Lower ID	Improve network lifetime and decrease the time delay.	Suitable for small scale networks. More nodes affect the network's performance.

procedure, the status of every node is checked to ensure member connectivity with CH. Each CH maintains a CH member table and a routing table. In the CH member table, all information of the member node is present and the routing table consists of information regarding routes of all member nodes. The route table contains node ID, next-hop sequence number and hops count. When a source node needs to transmit information to the destination node belongs to another cluster first it will transmit a REQ message to CH. When a CH receives a REQ message it will find the shortest path from the member node, it indicates member node needs information about the verified path. CH suggests the shortest path to ensure inter-cluster communication. CH will reorganize the routing table when a member node leaves the cluster or a new node joins the cluster and CH will inform all the nodes about the changes to manage route selection for inter-cluster communication.

5) AN OPTIMIZED STABLE CLUSTERING ALGORITHM FOR MANETS (OSCA)

OSCA [36] is an improved and stable ID-based clustering scheme in which a CH is selected based on the lowest value of ID assign inside a cluster. In the clustering establishment phase, every node sends a HELLO package, including two variables namely node ID and energy value, to its neighboring nodes. Before the initial clustering, every node is given an ID. This ID is called a node code. Each CH will have a table that consists of information about the IDs allocated to the nodes within a cluster. The node that has the lowest ID value is considered as CH. In the clustering maintenance stage, each node broadcasts a packet to its neighboring node to check each other's status. If a neighbor node is alive and its ID is smaller than the other neighboring node then all nodes will accept that node as a CH. The new selected CH will organize a table and sort the node IDs in descending order and then reassign the node ID to any new join node inside the cluster. Intra-cluster communication is handled by the node itself and inter-cluster communication is handled by CH. The node who has the same IDs in two CHs will act as a gateway node and this ID is called a group code because this node has a common ID for both clusters.

a) PERFORMANCE COMPARISON

The analysis between ID-based algorithms presented in Table 3 shows the CH selection criteria of each scheme along with the advantages and disadvantages of every scheme. Fig. 8 illustrates the performance comparison between these schemes. The stability level of ILCRP and CBRD is better as compared to CHSUA, LIDAR and OSCA and while LIDAR and OSCA are more stable as compared to CHSUA. CH change in LIDAR and CBRD is relatively high as compared to other schemes. Node mobility level in ILCRP is average while other schemes perform better only on low mobility levels. CH change in all these schemes is low but in CHSUA, ILCRP and CBRD possibility of cluster overlapping is possible. OSCA and LIDAR have high load and congestion while all other schemes show better performance in low mobility with a balanced load and congestion control. The efficiency level of ILCRP and CBRD is better as compared to LIDAR, CHSUA, and OSCA.

D. TOPOLOGY BASED ALGORITHM

The network topology depends on the node's transmitting range, communication links between network nodes and hops between nodes and CH. The goal of topology-based clustering is to control the topology of the network.

1) MULTI-HOP CLUSTERING BASED ON NEIGHBORHOOD BENCHMARK IN MOBILE AD-HOC NETWORKS (MHNB)

In MHNB [37] CHs are selected based on neighborhood benchmark (NB) which is a method to construct multi-hop clusters with a balanced size. The NB of CH is always higher than the NB of cluster nodes which avoids frequent CH change and ensures stability in clustering. The CH selection process for every node consists of a cluster radius which is the number of hops. Iteration in the CH selection process depends on the beckoning period of a HELLO packet from 1-hop neighboring nodes. After CH selection, a node handshake

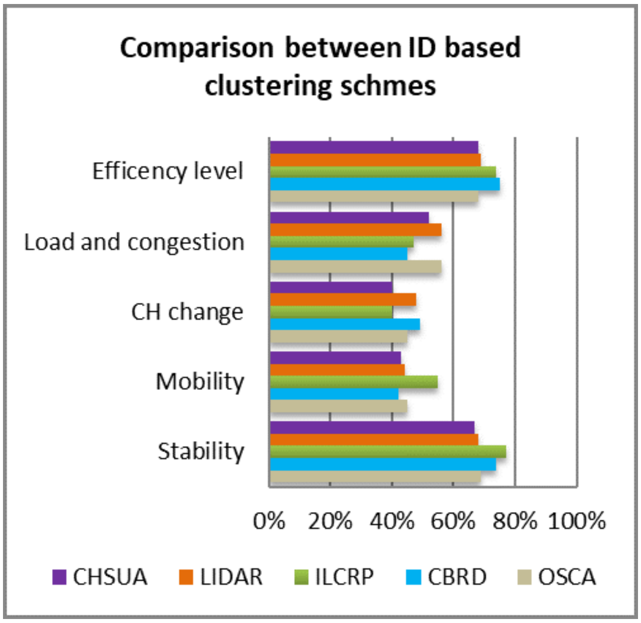

FIGURE 8. **Performance analysis of ID-based schemes.**

is executed to build the multi-hop cluster that ensures all clusters are connected properly and detects inconsistency between them.

In cluster maintenance phase CH multi-casts beaconing messages to all the cluster members. Each cluster member sends the acknowledgment message to its CH on receiving the beacon message. If the CH did not receive an acknowledgment from a member longer than the threshold time interval, then CH will consider the member is unreachable and deletes it from the list. If a non-CH node has not received the beacon message from its current CH longer than threshold time interval then it will consider its current CH to be unreachable and reselects its CH. The size balancing technique is also used in the cluster construction process that expands the clusters evenly in all directions. Hence cluster has a higher possibility to cover more nodes in the network.

2) CONNECTIVITY-BASED CLUSTERING WITH STRETCHING TECHNIQUE (CCST)

In the cluster establishment phase of CCST, every node transmits a HELLO message to its neighboring nodes [38]. After the broadcast of this HELLO message, each node will know its node degree as well as all the IDs of the neighboring nodes. A node that has the smallest ID as compared to its neighborhood nodes will set a timer for a predefined interval of time and waits for more messages from the other nodes. When the timer expires, that node becomes CH and broadcasts a cluster configuration message to every node. The nodes that are k-hops away become members and nodes that are to become CH candidates in the next selection procedure are also identified.

In the cluster maintenance phase, the stretching distance is used which provides freedom of movement to nodes because now nodes can stay in their cluster without joining other clusters as long as the intra-connectivity for a node is acceptable and the distances between nodes and CH do not exceed multiple hops. This algorithm improves the inter-connectivity of nodes and proposed that the node has to leave the current cluster if communication with other nodes is not well. A modified condition is proposed for a node to leave the cluster. This condition depends upon the degree of intra connection. If the value of the degree is less or equal to the critical value then the node will leave the cluster. To prevent a node frequent handover, every node that joins a new cluster follows a time restriction. Based on this restriction a node will become a member of another cluster when the time of restriction is completed.

3) ENHANCED SECTORIZED CLUSTERING SCHEME BASED ON TRANSMISSION RANGE (ESCS)

ESCS [39] ensures the stability of a cluster by dividing it into three sectors based on the nodes' residual energy. These sectors control the topology of clustering. In the initial clustering stage, every node stores its position information as well as of the neighboring CHs by HELLO messaging and periodically broadcasts it. This information consists of the node ID, the initial energy of a node and its information of position in Cartesian coordinates (x, y). Every node maintains a table of the single-hop neighboring node that has the information neighbor nodes received from HELLO messages.

Node energy plays a vital role and based on energy a decision will take place about the association of a node with a certain energy sector. When the energy level of a node is higher than the threshold maximum energy, it belongs to the high energy sector similarly a node belongs to the low energy sector if the value of node energy is less than the threshold minimum energy. If a node possesses energy value in between the value of maximum and minimum threshold limit then such a node is in the medium energy sector. The time for which a node will stay in a certain energy sector is called staying time and this metric will be helpful in the CH selection procedure as well as for the effective establishment of a clustering environment. CH for every cluster is selected by taking into account the residual energy of nodes and staying time. The node with maximum Combined Factor of residual energy and predicted staying time will be selected as a CH which results in the long life of the CH so that it can manage its cluster nodes effectively. The staying time is calculated by using the node relative velocity vectors from the past position information of neighbor CHs and itself. Every node judge whether it is moving towards or away from the CH with the help of measuring its relative velocity vector with the CH which controls the transmission range of cluster.

4) A TOPOLOGY MANAGEMENT SCHEME FOR LARGE SCALE MOBILE AD HOC NETWORKS (TMS)

TMS [40] is a topology-based scheme consisting of two phases where the first phase is cluster formation state while the other is topology maintenance. During the first phase,

TABLE 4. Analysis of topology based clustering algorithm.

Topology based clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
MHNB [37]	Neighborhood benchmark (NB)	Provides stability of the network with balanced sizes	The formation of cluster causes more communication overhead among nodes
CCST [38]	Node degree and smallest id	Reduces the number of CH with better intra-connectivity.	Maximizing the distance between two adjacent CHs can affect the speed of inter-clustering communication.
ESCS [39]	Residual energy of nodes and staying time.	Increase cluster stability by identifying the optimal transmission range.	Calculation of staying time and optimal transmission range on each node causes the delay.
TMS [40]	Clustering and QoS metrics	Improves scalability and QoS in large scale network, minimize the formation of extra clusters and reduces the number of CH re-selection.	A large number of clusters with multiple nodes may cause congestion.
KHSM [41]	Highest degree	Reduces the number of CHs and improves QoS for large scale MANETs.	Requires each node to exchange more control messages increases traffic load.

every node keeps the information of its local topology. A set of HELLO packets are broadcasted to detect the k-hop neighbor nodes within the range. The CH is elected by using clustering and QoS metrics. Clustering metrics include node mobility (M), node trust (T), battery remaining energy (E) and node density (D) while QoS metrics use Link Quality (S) and Local Available Bandwidth (LBW). Every node calculates its value of weight based on a combination of its QoS and clustering metrics. The node with the highest weight value is selected as an optimal CH.

After CHs selection, other nodes join suitable clusters by selecting neighbor CH that has the largest weight value. During the gateway selection process, every CH chooses gateways to communicate with the neighboring clusters. The gateway selection procedure depends on QoS metrics. A node that has attributes to be a gateway is selected if that node can connect to one or more neighboring clusters and has the highest QoS metrics. A topology can change due to the mobility of nodes and dynamic property of MANETs. Therefore, in topology maintenance parameters like CH weight value change, trust value change and phase node mobility are re-calculated again to maintain the efficiency of the algorithm during a topology change. A dominating set recovery mechanism is also proposed which is capable of selecting alternate dominating nodes from a set of nodes to handle link failures due to node movement. This solves the issues of topology imbalance and network instability.

5) K HOP SCALABLE MANAGEMENT SCHEME FOR LARGE SCALE MANETS (KHSM)

In KHSM [41] the main focus is on the construction and maintenance of network topology. The cluster topology construction consists of multiple phases which include cluster formation phase, node joining phase and gateway selection process. In the cluster formation stage, neighborhood discovery and CH selection take place for the formation of the cluster. D-Hello message is used for neighbor discovery. Based on the exchange of messages between nodes, neighborhood discovery takes place to construct a topology. In the CH election phase, each node generates a weight value based on remained battery energy, node mobility, node density, and trust value and broadcast it to k-hops neighborhood. Every node compares its weight value with the one received and declares itself as a CH if has the highest weight value among all neighbor nodes. After CH selection each CH selects its gateway nodes based on the value of the highest QoS and its location, A new node can join a cluster based on the approval message from a CH. The cluster maintenance stage is performed for CH reelection or to controls the situation of CH give up and to manage the changing cluster topology. KHSM effectively reduces the number of CH re-election for large scale topology and extend node membership time to keep the number of clusters comparatively low during the maintenance phase.

a) PERFORMANCE COMPARISON

The analysis of topology-based algorithms presented in Table 4 shows the CH selection criteria of each scheme along with the advantages and disadvantages of every scheme. Fig. 9 illustrates the performance comparison between these schemes. The stability level of KHSM and TMS is better than MHNB, ESCS, and CCST while ESCS is more stable as compared to MHNB and CCST. CH change in all these schemes is low but in KHSM the cluster overlapping is possible. All these schemes show better performance in low mobility while Load and congestion are balanced. The efficiency level of KHSM and TMS is better as compared to MHNB, ESCS, and CCST.

E. ARTIFICIAL INTELLIGENCE-BASED ALGORITHM

Artificial Intelligence (AI) based clustering, genetic algorithms, and fuzzy logic-based algorithms are used for every

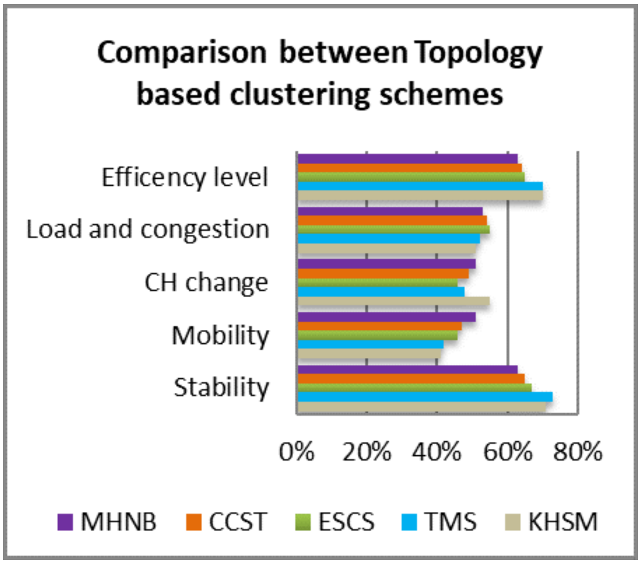

FIGURE 9. Performance analysis of topology based schemes.

decision-related event in clustering. Different techniques are used to identify, generate and select such a methodology that generates high-quality solutions for optimization and search problems regarding cluster-based MANET. A CH is also elected by searching the ultimate and most suitable node among all member nodes of a cluster.

1) GENETIC ALGORITHM SIMULATED ANNEALING CLUSTERING STRATEGY (GASA-CS)

In [42] GASA-CS is applied to dynamic weighted (DW). GASA is a computational intelligence method in which several nodes in a network is equal to randomly generated initial population. A CH is selected from a randomly generated initial population which produces an equal quantity of chromosomes as an integer string. Randomly generated L integer arrangements are used for N number of nodes. The range of these randomly generated arrangements is inside the array of $[1, N]$. These random arrangements derive L sets of CHs. Crossover operation is performed to select the better L sets of CHs and these newly elected CHs are replaced with the original ones.

The SA selection criteria used in this algorithm is to decide whether to select one from L sets of CHs or wait for the new L sets in the next generation. The above procedure is performed to achieve results until a certain number of iterations is reached otherwise it converges. The convergence gives a corresponding set of CHs which shows a global or sub-optimal solution is achieved and the corresponding set of CHs is obtained. The selection strategy is based on Boltzmann mechanism by creating L random arrangements to achieve copulation. The new selection of probabilistic strategy can be adaptive to optimize the selection procedure. The basic aim of this algorithm is to improve the lifetime of a network by reducing the load. The strategy proposed in this algorithm also useful to decrease the clustering cost.

2) GENETIC ALGORITHMS WITH HYPER-MUTATION FOR DYNAMIC LOAD BALANCED CLUSTERING (DLBCP)

DLBCP algorithm [43] provides a solution to produce a set of CHs that is selected considering all the nodes. A random set of CHs will result in a random permutation of node IDs. The random permutation of node IDs is used to represent a chromosome. Every node ID in the chromosome is called a gene and it guarantees the uniqueness of node ID in every chromosome. To randomly generate the corresponding permutation of node ID for every chromosome, genetic diversity is used. DLBCP aims to find the set of CHs that can handle the load balancing in cluster structure. To achieve the load balance, each CH has the same CH degree and the criterion to accomplish is the standard deviation of the CH degrees. Therefore, among solutions, there is a need to choose one with the least standard deviation and this procedure is achieved by using a fitness function. The selection of high quality is based on the fitness value using pair-wise selection. So, using the hyper-mutation method aim to maintain the diversity of the population via adjusting the mutation rate applied to the current population.

3) MULTI OBJECTIVE GENETIC ALGORITHM FOR OPTIMAL CLUSTER HEAD SELECTION (MOGA)

MOGA [44] is a genetic algorithm (GA) proposed to solve efficiently the problem of multi-objective optimization (MOO). The energy consumption of the CH depends on the level of several nodes in a cluster and overhead it has to handle for intra-cluster communication. This algorithm also discusses the clustering for energy-efficient operations and switching states of nodes. Nodes that are previously selected as CH are thus prevented from immediate re-election to conserve energy.

This algorithm consists of four steps. The first step consists of initialization of the population in which an upper limit of the population of nodes for a cluster has been adjusted to a threshold value so that the complexity of the algorithm doesn't become high. The second step is called fitness evaluation for the population in which the fitness score of every node of the population is calculated as per the functions of the objective. The third step is a sorting step in which members of a population are categorized based upon their fitness score. In fourth step genetic operators (GO) are used for obtaining the next generation of individuals including the selection of the best individuals, performing crossover between them, and the introduction of randomness by an optional mutation in the population. The best solutions from every step of evolution are placed in an archive. The archive is again sorted to find the best solution and is further optimized by passing from steps 1 to 4 again. The process ends when several successive attempts fail to improve the results in the archive. In the end, a weighted optimization among the archived results brings the solution for the current iteration.

4) FUZZY BASED PARAMETRIC CLUSTERING MODEL (FBPC)

In FBPC [45] nodes are defined randomly and parameters of each node including connectivity, stability, node degree,

TABLE 5. Analysis of AI-based clustering algorithm.

AI-based clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
GASA-CS [42]	Genetic algorithm	Optimize total number of CH, increase lifetime and decrease clustering cost.	The technique relies on iterative search can cause delay and the population of solutions are provided instead of single efficient solution.
DLBCP [43]	Genetic algorithm	Solves dynamic load balancing cluster problem on a large scale.	In rapidly changing environment performance degrade because more time is needed to search good solutions.
MOGA [44]	Genetic algorithm	Provide optimization of network by removing the problem of MOO. Improve network lifetime.	Optimization is lengthy which may cause delay and can affect network efficiency.
FBPC [45]	Fuzzy logic	Ensure stable network by reducing hop count.	Repeated process in each round incurs high overhead.
FRCA [46]	Fuzzy logic	Ensure efficient selection of the CH and improve efficiency.	An optimal number of overheads increase with high transmission distance.

and adaptive loss are calculated. These variables are analyzed using fuzzy operators. CH is identified after performing a fuzzy operation on these parameters. Clusters are generated to ensure optimized communication over the network. Analysis of degree is performed under the stability vector over the network to generate optimized clusters. Only those nodes are considered which have effective connectivity and stable for a certain threshold period. The fuzzy rule is used to pinpoint the connectivity aspects and achieves maximum stability on every node. The nodes are identified based on load limits and analysis of the communication parameter with the help of fuzzy operators. After identifying the nodes, the cluster adaptive communication is formed. The communication between a node to cluster and between clusters is controlled with the help of a fuzzy adaptive clustering model that utilizes the capabilities of node mobility effectively.

The analysis of architectural constraints and communication constraints are used for forming efficient and effective clustering by utilizing fuzzy-based rule formation. The analysis of distance and mobility is performed to generate a list of neighbors. Once neighbors are obtained, the distance and connectivity analysis is applied for CH identification. The cluster formation members are identified based on the communication parameters so that the more reliable cluster formation will be obtained.

5) FUZZY CLUSTER-BASED DYNAMIC ROUTING ALGORITHM (FRCA)

In CH selection, formation and cluster maintenance is performed using Sugeno fuzzy function [46]. This function covert the normalized input values into linguistic fuzzy variables. These variables categorized the value into three levels. These levels are low, medium and high. Fuzzy rules are used to find the optimal distance between CH and member node, optimal CH selection and energy monitoring of member nodes. The energy of each node is compared with its neighbor nodes along with the mobility level. Efficient CH is selected by using fuzzy and energy modeling. The other parameters of nodes like the distance between nodes and mobility are also considered before selecting a CH. These parameters provide a balance in the consumption of energy. Fuzzy state monitoring structure clusters adaptively and it is used for efficient clustering in the network. It is more effective in a situation when the mobility of nodes varies the size of networks. The fuzzy value of a node is used to prevent attacks from malicious nodes. Several overheads can be reduced due to the consideration of several parameters in the election of CH and improves routing performance. The higher value of available power (AP) for node means power is more stable. A node which has a higher value of AP is most likely to be elected as a CH and can ensure a longer network lifetime.

a) PERFORMANCE COMPARISON

The analysis of AI-based algorithms presented in Table 5 shows the CH selection criteria of each scheme along with the advantages and disadvantages of every scheme. Fig. 10 illustrates the performance comparison between these schemes. The stability level of FRCA and MOGA is high as compared to GASA-CS, FBPC, and DLBCP while GASA-CS is more stable as compare to FBPC and DLBCP. CH change in all these schemes is low but in FBPC and DLBCP possibility of the cluster, overlapping exists. These schemes show a better performance because of the balanced load and congestion control. GASA-CS, FBPC, and DLBCP moderate mobility levels of a node are considered while other schemes support low mobility levels of nodes. The efficiency level of FRCA, GASA-CS, and MOGA are high as compared to FBPC and DLBCP.

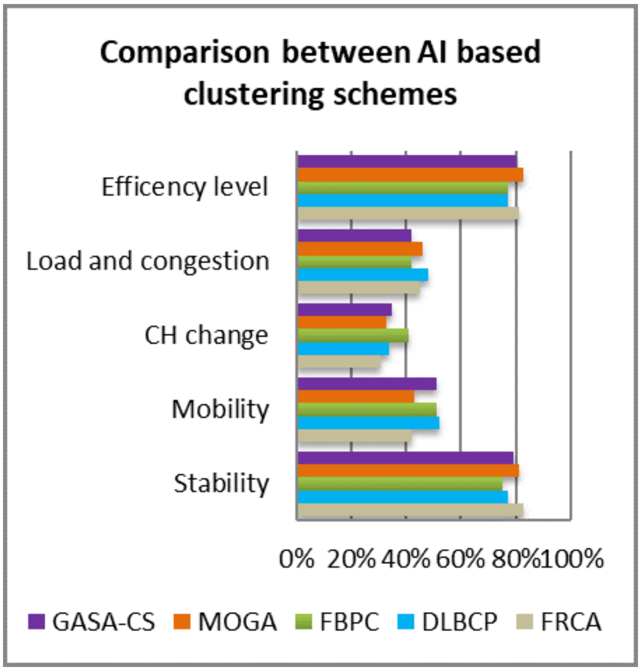

FIGURE 10. Performance analysis of AI-based schemes.

F. HYBRID ALGORITHM

Hybrid clustering algorithms are those algorithms that consist of a combination of two or more algorithms. These sorts of algorithms are adapted to provide more stable clustering inside the network.

1) CONNECTIVITY, ENERGY, AND MOBILITY DRIVEN CLUSTERING ALGORITHM (CEMCA)

CEMCA [47] elects a CH by ensuring network topology stability. Each node broadcasts its ID to the neighboring node within the transmission range. Every node can estimate the distance between them by the received signal strength. The CH is selected depending upon the value of a constraint named as the quality of the node. This constraint measures the level of suitability and it consists of the highest value of node degree and battery life. It also considers the lowest node mobility along with the stable transmission range of the node. The best candidate for CH is such a node that has the lowest mobility, sufficient amount of battery energy, a suitable number of neighboring nodes and an acceptable transmission range. For every node, SEMC calculates its transmission range. This range selection provides stability in node connectivity and energy preservation by the reduction in transmission range according to the distance between nodes. Every CH identifies several hops to reach its members and performs the construction of cluster members set. Each node recognizes the CH while every CH stores information of its member nodes which allow the clustering protocol to perform inter and intra-clustering communication.

2) A HYBRID APPROACH FOR NODE CO-OPERATION-BASED CLUSTERING (HANCC)

HANCC algorithm [48] consists of three basic procedures. The first process consists of evaluating the weighting scheme of cooperativeness in the system. The clustering is done based on nodes with high cooperativeness weight and that node becomes a CH. A CH ensures the communication between nodes based on clustered value by grouping them according to the cooperative weight. The behavior monitoring of these nodes is also based on the cooperativeness of the nodes. The spatial events are used to calculate the weight of the cooperativeness occurred at unique stages of communication. The cooperative weight of a node is based on removing the selfish nodes. These types of a node can cause path and message loss, therefore, these kinds of selfish nodes must be discarded.

After completing the computation of the cooperativeness of the nodes, the second step takes place which consists of a self-monitoring process where every node reorganizes itself. It does not allow any unauthorized node to get involved in the communication between nodes. The last step is to form a cluster by using a weighting-based scheme which provides efficient communication between the nodes involved in a clustering environment. It is necessary to assign an ID to every node and monitor its activities. The reorganization of a cluster takes place after calculating the weight of cooperativeness. The efficiency of the HANCC algorithm is defined by considering the performance of a single cluster without having selfish nodes.

3) HYBRID ALGORITHM FOR PRESERVING ENERGY AND ROUTING DELAY (HPED)

HPED [49] is based on the k-hop clustering structure that manages nodes according to their k-lowest ID. Every node maintains a table of neighbors containing the addresses of its neighboring node. If a source has to send some packets to the destination, it directly sends it using this table. A network contains normal and member nodes. Initially, a CH is selected among normal nodes having the k-lowest ID. In HPHC, there are two conditions by which a CH can change. One condition is a cluster merging scenario in which when two CHs move into the same cluster where one CH has to give up its position as per k-lowest ID and the other condition in which a CH leaves the cluster. Any node can act as a gateway node between two clusters that is useful to interconnect an adjacent cluster. Every node broadcasts the HELLO message to maintain network topology. The communication path error can be identified and recovered by using an Error recovery packet. If a neighborhood node link is unreachable, a node will inform the CH by sending a link recovery packet. This scheme decreases the end to end delay in a network and provides an acceptable packet delivery ratio. The topology of a cluster is adaptive that can be useful to handle different kinds of real-time situations.

TABLE 6. Analysis of the hybrid clustering algorithms.

Hybrid clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
CEMCA [47]	Quality of node	Reduces the number of CH changes and chooses an appropriate transmission range for every node.	Transmission range adapts itself according to the number of neighbors.
HANCC [48]	High cooperativeness range	SNode cooperativeness enhances the lifespan and packet data transmission, decreases communication failure.	High communication between nodes due to cooperativeness nature of the algorithm can increase number of overheads in large scale networks.
HPED [49]	k-lowest ID	Better load balancing along with a minimum average end-to-end delay.	Route re-discovery process can cause extra burden on member nodes and CH.
HTTCH [50]	Highest degree of node energy	Minimizes the number of clusters formation.	Nodes need to perform fuzzy operations it increases the delay.
FACO [51]	FNodes with higher stability factor and energy consider being CH.	Increases stability and CH is invoked only on demand, therefore, it also decreases the cost of communication.	Large end-to-end delay and the minimum lifespan of nodes cause management issues.

4) HYBRID THREE-TIER COMPETENT FUZZY CLUSTER ALGORITHM (HTTCH)

In [50], a hybrid scheme is proposed which is a combination of a modified fuzzy C-mean technique and cuckoo search (CS) algorithm. A modified fuzzy C-mean technique is used to find the optimal CH. The fitness value of each node is calculated and multiple solutions are generated based on the fitness value using fuzzy rules. These solutions are arranged in a merge table M(T)t according to the fitness value. CS algorithm represents a dynamic system with finite solutions. To find the optimal best solution from the merge table, CS is used for CH selection. As nodes change their state dynamically according to the state of their neighbors. A 3-Tier clustering mechanism based on request and acknowledgment is used for stable and reliable selection of nodes and links for communication within and outside the cluster. HTTCH provides the optimal selection of CH and ensures reliable communication within and between different clusters.

5) DYNAMIC CLUSTER FORMATION USING HYBRID FACO

In [51] fuzzy ant colony optimization (FACO) is used to find the optimal shortest path. The probability of each path between source and destination is calculated. The first step is to start the initialization of the population. In the next step evaluation of fitness takes place using ACO. Fuzzy rules are applied to select the shortest path with minimum delay. After the selection of the shortest path, data is transmitted by using the dynamic data forwarding mechanism of an ant colony. Neighbor and topology discovery consists of two levels. In the first level, a node first determines its neighbor nodes and organizes a list. By using the data of the list, identifies the possible paths for data transmission. In second level sender node will identify the best forwarder by considering a minimum number of hops. A node with the highest energy and with a higher level of stainability is elected as CH. This algorithm reduces the number of overheads and provides a robust solution.

a) PERFORMANCE COMPARISON

The analysis of hybrid algorithms presented in Table 6 shows the CH selection criteria of each scheme along with the advantages and disadvantages of every scheme. Fig. 11 illustrates a performance comparison between these schemes. The stability level of HPED and FACO is better than HANCC, HPED, and HTTCH. CH change in all these schemes is a low

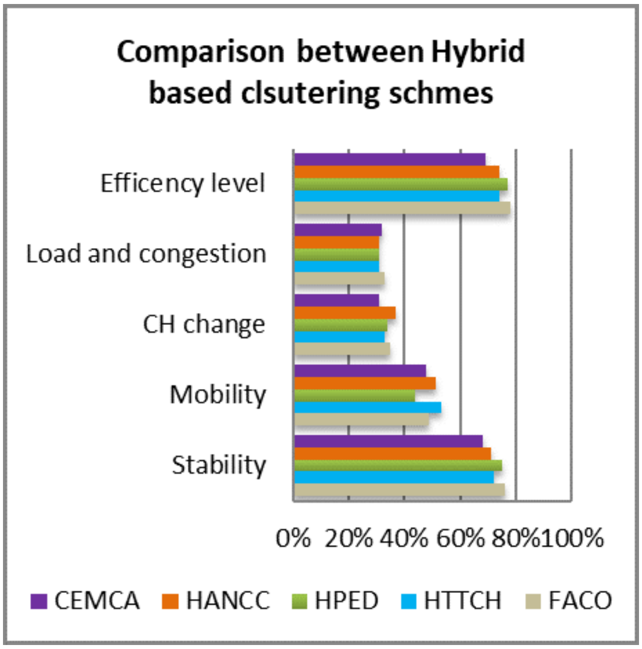

FIGURE 11. Performance analysis of hybrid schemes.

but small possibility of cluster overlapping also exists. All these schemes show better performance in low and moderate mobility. All schemes provide better load balancing and congestion control. The efficiency level of HPED and FACO is higher than HANCC, HPED, and CEMCA.

IV. PERFORMANCE EVALUATION OF CLUSTERING SCHEMES

In the previous section, we have comprehensively discussed the classification of clustering schemes along with analysis and comparison of schemes. This section contains the performance evaluation of the above-mentioned classification of schemes concerning QoS. The awareness of QoS in clustering-based MANETs is very important because it determines the overall performance of the network.

In MANETs QoS can be considered by taking into account multiple benchmark parameters. We categorize these parameters in MANET as follow:

1) Nodes and CH related parameters
Energy consumption and network lifetime are the metrics that depend on node and CH performance. CH always requires more energy to control the cluster.
2) Connectivity related parameters
Link quality and delay are connectivity related metrics that depend on the connection stability.
3) Topology related parameters
Hop count, number of overheads are related to the overall topology of the network.
4) Communication-related parameters
Packet delivery ratio, throughput, data rate, and bandwidth are related to the communication capabilities of the network.

In Fig. 12 we have summarized the constraints of QoS with their level of importance and focus of researchers. Table 7 presents the evaluation of clustering protocols according to the benchmark parameters of QoS. Researchers have mostly focused on energy consumption, network lifetime and throughput of the network. From the original simulation results of different types of schemes, we evaluate the performance using QoS metrics. Fig. 13 illustrates that AI and hybrid clustering schemes provide better results in terms of energy efficiency while energy-based clustering schemes show better energy efficiency as compare to ID-based clustering schemes. Fig. 14 illustrates the performance of clustering schemes concerning packet delivery ratio and throughput. AI, energy, and hybrid schemes show better performance as compared to other schemes. From Fig. 15, AI and hybrid schemes show better performance in terms of optimal selection of CH, as a result, there is a minimum average CH change in these schemes as compared to others. While topology and energy-based clustering schemes show better results as compared to others. Our evaluation determines that the overall performance of AI, Hybrid and topology-based schemes are better as compared to other types of clustering schemes.

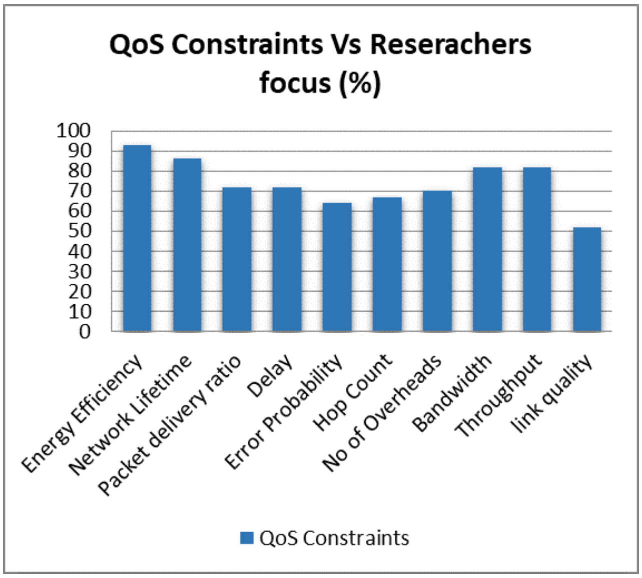

FIGURE 12. Researchers focus on clustering schemes.

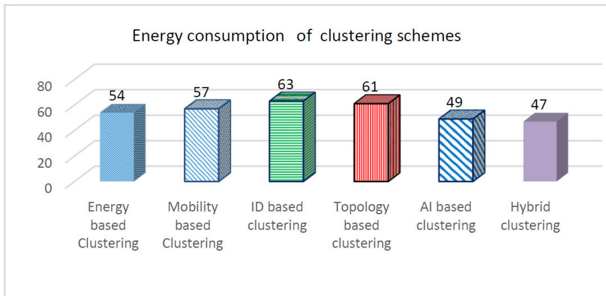

FIGURE 13. Energy consumption of clustering schemes.

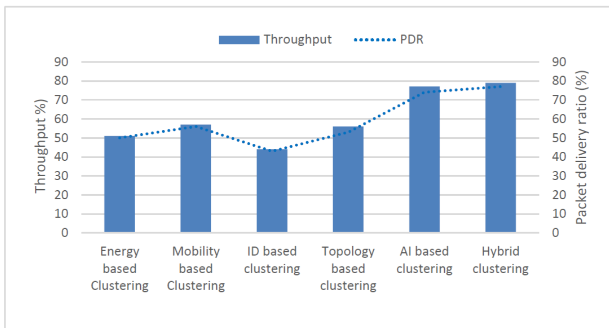

FIGURE 14. Throughput and PDR of clustering schemes.

Every type of scheme mentioned and evaluated above is designed to achieve some desirable objectives of the algorithm designer. These objectives are dependent on the QoS metric. From our study and evaluation, we have recognized that there are some important tradeoffs in QoS metrics. These trade-offs play a vital role in determining the overall performance of the network. Most important tradeoffs for MANETs are as given below:

TABLE 7. Performance evaluation of clustering algorithm.

	QoS parameters for MANETs											
Ref	Energy consumption	Network lifetime	Packet delivery ratio	End-to-end delay	No of nodes and clusters	Hop count	No of overheads	Throughput	link quality	Rate of CH change	Routing efficiency	Spatial reuse of resources
FWCABP [22]	×	✓	×	×	×	×	✓	✓	×	×	✓	✓
ECEC [23]	✓	✓	×	×	×	×	×	×	×	×	×	×
SEEA [24]	✓	×	✓	×	×	✓	✓	×	×	×	✓	×
EEAMHT [25]	✓	✓	×	×	×	✓	×	×	×	×	×	✓
EECA [26]	×	×	×	✓	×	×	×	✓	✓	×	✓	✓
MobDHop [27]	✓	✓	×	×	✓	×	×	×	×	✓	×	×
CMCA [28]	✓	✓	×	×	✓	✓	×	✓	×	✓	×	✓
SCHM [29]	×	✓	×	✓	×	✓	×	×	×	×	×	×
MPBC [30]	✓	✓	×	✓	×	✓	×	✓	✓	×	✓	×
GMCA [31]	✓	✓	×	×	×	✓	×	×	×	×	×	✓
CHSUA [32]	×	✓	×	×	✓	✓	×	×	×	✓	✓	×
LIDAR [33]	✓	✓	×	×	✓	×	×	×	×	✓	×	×
ILCRP [34]	×	×	×	✓	×	✓	✓	-	✓	×	✓	×
CBRD [35]	✓	×	✓	×	✓	×	✓	×	×	✓	✓	✓
OSCA [36]	×	✓	×	✓	✓	✓	×	×	×	✓	×	×
MHNB [37]	✓	×	×	✓	✓	×	✓	✓	×	✓	✓	×
CCST [38]	×	✓	×	×	✓	×	✓	✓	×	✓	×	×
ESCS [39]	✓	✓	✓	×	✓	×	✓	×	✓	✓	✓	✓
TMS [40]	×	✓	✓	×	✓	×	×	✓	×	✓	×	✓
KHSM [41]	×	✓	×	×	✓	×	×	✓	×	✓	×	×
GASA-CS [42]	✓	✓	×	×	✓	✓	×	×	×	✓	×	×
DLBCP [43]	×	×	×	×	×	×	×	×	×	×	✓	✓
MOGA [44]	✓	×	×	×	×	×	×	×	×	×	✓	×
FBPC [45]	×	×	✓	✓	×	✓	×	×	✓	×	✓	×
FRCA [46]	✓	✓	×	✓	×	×	×	×	✓	×	×	✓
CEMCA [47]	×	×	×	×	×	×	×	×	×	×	✓	✓
HANCC [48]	✓	✓	×	✓	×	✓	×	×	×	×	×	✓
HPED [49]	×	✓	×	×	✓	✓	✓	×	×	✓	×	×
HTTCH [50]	✓	×	✓	✓	×	×	×	✓	✓	×	✓	✓
FACO [51]	✓	×	✓	×	×	×	×	✓	✓	×	✓	×

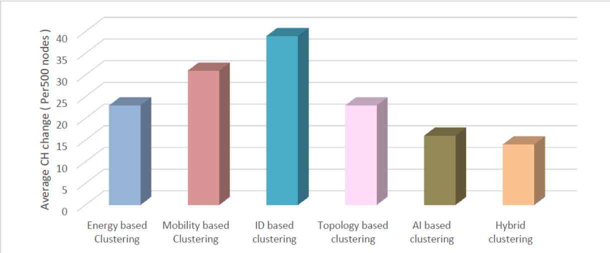

FIGURE 15. Average change in CH in MANETs schemes.

1) The tradeoff between energy consumption and overall performance
   Energy consumption is an important parameter that determines the overall performance of the network. A tradeoff between throughput and energy consumption can be achieved. At the acceptable cost of energy consumption, the algorithm designer can achieve desirable throughput and data rate. At the expense of affordable energy consumption robustness of the network can also be achieved by performing fault tolerance using multi-path clustering. Similarly, tradeoffs between Network lifetime and energy consumption can also be achieved. Network lifetime depends on the total number of active nodes available in the network. The lifetime of the network can be improved by decreasing the power consumption at every node. At the expense of minimizing the extra task and workload on the node, power consumption can be decreased and the network lifetime can be improved.
2) The tradeoff between throughput and delay
   In the case of link congestion, some packets can be dropped to avoid further congestion. Congestion of the network can be removed by reducing traffic on that link by forwarding the same dropped packet to another route as result robustness can be achieved and the overall throughput of the network can be improved at the acceptable cost of a small delay in transmission.
3) A tradeoff between the number of clusters and network lifetime
   Large cluster size with a greater number of member nodes can cause a frequent change of CH, more energy consumption at every node and monitoring topological changes during communication. More clusters in a network with small size are easy to manage the topology, therefore, several clusters can be increased at an acceptable rate to increase network life.
4) A tradeoff between packet delivery ratio and no of hops
   The packet delivery ratio represents the ratio of messages received at the receiver end. When a network becomes congested, the mechanism of route diversity can be used to select a link of less traffic load. As a result, the packet delivery ratio can be improved at the expense of a greater number of hops for data transmission.

The major challenge during algorithm design is to select the optimal tradeoff between the QoS metrics and the above-mentioned tradeoffs must be considered to achieve better performance. After reviewing multiple types of clustering schemes, from Table 7 we can point out some important factors which are difficult to handle and must be considered while designing a cluster-based scheme. These factors can cause limitations and affect the performance of clustering-based schemes in MANETs. These factors are as follow:

1) Variation in cluster size
   Clustering schemes require a better mechanism to control rapid variation in cluster size. The greater number of nodes and rapid topological changes in the network can cause variation in the size of the cluster. This can cause an increase in transmission overheads as a result communication complexity increases.
2) Coverage area
   Clustering schemes need a better mechanism to control the coverage area of a cluster. A better cluster emergence mechanism must be considered in a case when the number of nodes in a cluster increases the threshold value. It can improve overall network efficiency. The coverage area of a node is important in large scale high mobility scenarios because if a node cannot access its neighborhood node within its coverage area then a node can be isolated or it causes quick power drainage.
3) Fault tolerance
   Fault management between nodes is a special case when there is a temporary link failure between nodes inside clusters. This case needs more attention while designing a scheme.
4) Scalability
   On a large scale, MANETs more advance schemes are needed to check and improve the adaptability of clustering nodes.
5) Exchange of control packets during cluster formation and maintenance
   When a structure of the network is changing itself dynamically due to a large number of mobile nodes it can cause extra exchange of control packets. In the cluster formation and maintenance phase, a greater number of control messages can be exchanged between CH and nodes. These exchanges of messages sometimes require a significant amount of bandwidth and it also exhausts the energy of CH and node.
6) Optimal path selection
   The effective path selection is important for both inter and intra based clustering in MANETs. The optimal path can decrease the extra load by ensuring efficient communication between the source and destination node.
7) Optimal tradeoff selection
   For CH selection and cluster maintenance, there is always a tradeoff between different performance-based constraints. In a clustering scheme, there must be a

matrix to keep a balance between them to ensure the overall efficiency of a network.

8) Network dynamics
The dynamic nature of network topology is one of the difficult portions to handle while designing clustering-based schemes. This can affect the selection procedure of a CH. Dynamic topology can cause a delay in locating accurate information regarding the current state of a node while deciding on the optimal path and CH selection.

V. OPEN RESEARCH CHALLENGES

The open research challenges for cluster-based MANETs are as follow:

A. CLUSTER RELATED DESIGN CHALLENGES

1) Cluster maintenance according to the node positioning in a cluster
In a cluster total number of nodes, neighborhood nodes and distance among these nodes are important constraints to identify the quality of a link between nodes. If a node has a smaller number of neighboring nodes, then such a node will face problems in communication and as a result, its battery power will drain quickly. For optimal path selection between source and a destination node, cluster maintenance according to the position of a node is a very challenging open research topic for intra and inter based clustering.

2) Scalability of a cluster
The scalability of a network is a vital parameter and must be considered while designing a clustering-based scheme. A clustering protocol must be designed in such a way that it can work properly in the presence of a variable number of nodes. The scalability of a cluster must be organized in a way that a CH performs effective processing of all nodes data without any delay and load on the overall network. The scalability of a cluster for real-time application is still a difficult task and it's an open research domain with a lot of challenges.

B. COMMUNICATION RELATED CHALLENGES

1) Inter and intra-based data flow challenges
In clustering-based schemes, inter and intra-based clustering is still an open research direction for researchers. For effective management of traffic between clusters, a data flow model between nodes is essential that also needs to consider the unpredictability of node behavior and can identify the optimal link for traffic flow.

2) Mobility related challenges
Mobility is a constraint that must be considered while designing a clustering-based scheme. Fast-moving nodes can cause many issues related to node connectivity and cluster maintenance. Therefore, efficient mobility-based models are required which is a challenging task.

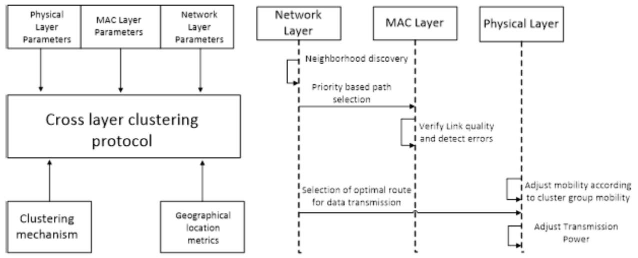

FIGURE 16. Architecture of cross-layer clustering protocol.

3) Challenges related to link selection and its quality
Efficient and reliable path selection is key for clustering-based communication. Due to unpredictable topological changes and mobility of nodes, it is essential to consider the path quality metrics that will not only identify a reliable link between nodes but also identify the optimal path selection. The designing of path quality metrics in the presence of mobility and topological changes is still an open challenge for researchers in clustering-based schemes.

C. PERFORMANCE RELATED CHALLENGES

1) Desired QoS
In clustering-based schemes, there are still some issues regarding packet forwarding and latency that affects the packet delivery ratio and also increase the network load. These issues need to be resolved to achieve a desirable QoS. A new study is needed to minimize the causes which affects the overall QoS of a network.

2) Fault tolerance
There is a need for multiple new constraints to be implemented in schemes to ensure reliable delivery of packets along with some fault tolerance mechanism to minimize the effects of damage caused due to path loss.

3) Reliability
The capability to transfer packets from source to destination node with a low percentage of damaged packets ensures the reliability of a network. To create a reliable end-to-end routing path for inter and intra-based clustering is a challenging task for algorithm designers.

VI. PROPOSED SOLUTIONS TO IMPROVE QOS FOR CLUSTER BASED MANET

To minimize the open research challenges of MANET and to improve overall QoS, we propose two solutions for MANET. Our first solution is based on the idea of cross-layer design and our second solution is based on an idea of self-organization [52], [53].

A. CROSS LAYER CLUSTERING FRAMEWORK

The cross-layer clustering framework (CLCF) can provide a better solution to resolve the open challenges related to cluster design, maintenance, and performance. The proposed architecture of the cross-layer clustering protocol for MANET

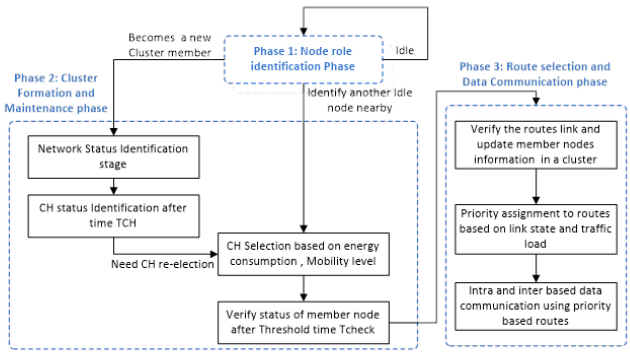

FIGURE 17. Working mechanism of proposed CLCF.

is shown in Fig. 16 which represents the cross-layer design between network, physical and MAC layers for cluster management and route selection for inter and intra-cluster communication. The cross-layer model can provide a better way to manage the cluster and select an optimal route with limited errors along with acceptable delays. The working mechanism of the proposed CLCF is illustrated in Fig. 17. CLCF consists of three phases. The node role identification phase is the first phase in which a node will identify its state. There are two states of a node named as idle and a member node. The idle condition represents a state in which a node is not a member of any cluster and identified as a free moving node. A member node means a node is a member of a cluster. If an idle node receives an invitation message (IM) from a CH, it can transmit an acknowledgment message (AM) to CH and as a result, CH will accept the request of the idle node to become a member node of a cluster by transmitting a confirmation message (CM). The second phase consists of cluster formation and maintenance. To effectively utilize the available resources and to minimize the number of idle nodes in MANET, we consider cluster formation procedure for idle nodes. If an idle node doesn't receive any IM from a CH but can identify another idle node inside its range. An idle node can transmit a cluster formation message (CFM) to an idle node with its node ID and energy level. If an idle node receives CFM, it will compare the energy level of its own with the energy level received from CFM. If the idle node has a more or equal energy level of another idle node, then it will declare itself as a CH and sends a CM message with the member id assign to another idle node, CH id and group mobility of a cluster. If the energy level of an idle node is less than in such a case that node will transmit acceptance message (ACM) with its node id to ensure that it accepts another idle node as a CH and waits for the CM message from newly elected CH. In the cluster re-election stage, a node with the highest energy level and medium mobility level will be elected as a CH. During the cluster maintenance stage, the CH will ensure the connectivity of its members after check and update the topological information in a MEMBER table. Phase 3 of CLCF consists of route selection and data communication. A priority is assigned to every available route in a cluster based on link-state and traffic load. The cross-layer model between clustering and MAC layer can offer a priority-based path selection to achieve limited delay and errors along with power-efficient intra and inter based clustering. The CLCF can improve the overall QoS and reliability of a network.

B. HYBRID SELF-ORGANIZATION CLUSTERING MODEL

To improve overall QoS in MANET, we proposed a hybrid self-organization clustering model (HSOCM) which is a combination of glowworm swarm optimization (GSO) and Krill Herd algorithm (KH). The Working mechanism of HSOCM is shown in Fig. 18. The selection of CH in HSOCM is done by applying the GSO approach [54] and cluster management is attained by applying a KH approach [55]. GSO algorithm is used in various fields of research for optimization. The foremost goal to use the GSO algorithm for CH selection in MANET is that we can assess the rank of an individual node by applying the fitness evaluation function of GSO based on the mobility and residual energy level of every node. Conferring to the GSO algorithm, the current rank of a glowworm can originate out by evaluating search space, luciferin level, and neighborhood range [54]. For MANET, luciferin level indicates the weight of a node grounded on the remaining energy level and mobility. Fitness function is evaluated and all nodes are graded conferring to their luciferin level for optimal CH selection. The node ranked highest among all nodes is elected as a CH. Cluster management is done using the KH process as presented in Fig. 18. Each krill's motion (Cluster member) is affected by CH's motion. Cluster participants should obey the speed of CH for efficient cluster management. By KH algorithm rules [55] of induced and foraging motion sideways with physical diffraction, each krill find motion solution grounded on the information of the current and previous position of CH and their motion. By applying crossover and mutation operation each krill updates its position conferring to the CH motion. Routes among the participant nodes are confirmed to guarantee effective inter and intra communication. HSOCM can provide better QoS with an ability of self-optimization and adaptivity conferring to the environmental variations and movement.

VII. COMPARISON OF THE RECENT APPROACHES REGARDING MANETS

In the above sections, the authors investigated state-of-the-art MANETS schemes as can be seen in Tables 1 to 7. These schemes include energy-based schemes, mobility-based schemes, ID-based schemes, topological schemes, AI-based schemes, hybrid schemes, etc. However, for a better understanding of the readers, in this section, the authors focused on the most recent approaches regarding clustering in MANETs.

A self-organized clustering algorithm for MANETs is presented in [53] which uses zone-based mobility to improve the scalability and stability of the network. The proposed algorithm uses the bio-inspired behavioral study of birds flocking for the creation and preservation of clusters. To this end, a dynamic cluster size management strategy is used

TABLE 8. Comparison of the most recent clustering algorithms.

Clustering algorithm	Cluster head selection methodology	Advantages	Drawback and issues
SOCZBM [53]	Self-organized	Reduce network congestion and improve the performance of the MANETs	Network lifetime is limited
FTBCA [56]	With trusted nodes	Eliminate misbehave nodes	Considerable amount of memory is used.
ECMS [57]	Revocation method	Prevent unauthorized access to the network	Only solve limited security issues.
SARS [58]	Dynamic repeated game model	Enhance energy efficiency and prevent the packets drop	Less resistive to DoS attacks.
BSAS [59]	Relay selection based method	Improve system capacity and energy consumption	Haven't consider security issues.
LWECM [60]	Light weight efficient cluster	Cluster-head life is prolonged by reducing the unnecessary load on cluster-head	higher maintenance cost
PCSM [61]	Multi-hop proximity aware clustering	Fits the MANET underlay and enhances the results of the search process	Traffic overhead is high
EBCH [62]	Dynamic hybrid topology (DHT)	Reducing the energy consumption during communication from source to destination	Lifetime of the network is limited

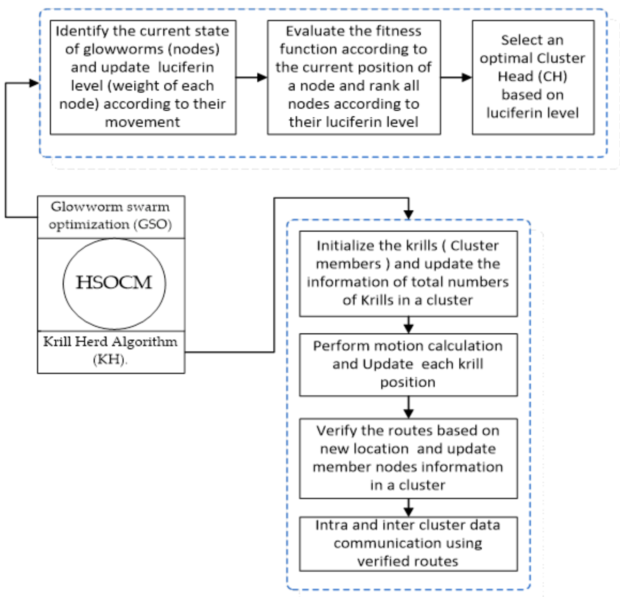

FIGURE 18. Working framework of the proposed mobility aware HSOCM.

to reduce network congestion and improve the performance of MANETs in group mobility. The proposed algorithm decreases energy consumption and increases the life of the network.

A recent approach is presented in [56], which investigates a secure multicast transmission. The authors presented Polynomial-based key management in Fuzzy Trust-based clustered networks (FTBCAs) for secure multicast transmission. The approach protects against both indoor and outdoor attackers and measures the presentation by interleaving the attack models. Similarly, in [57], [58] security-aware routing schemes (SARS) are presented. In the first scheme, the authors focused on the cluster-based certificate revocation scheme to isolate the malicious node and prevent the illegal admittance. The authors also proposed an energy connectivity mobility signal to noise ratio (ECMS) algorithm to identify the best cluster head in every cluster. The network sensor node is classified into three types such as warned node, revoked node, and normal node. In the certificate revocation, every node present in the network is monitored with the help of one-hop neighbors and used to gather malicious evidence of the nodes of the sensor. If any node is incorrectly recognized as a malicious node, its authentic nodes will send the justification packets to the correct gateways to correct the error. In the second scheme, an Infinitely-repeated game and cooperation method is considered to notice malicious sensor nodes and also to enhance energy-efficiency. This method aims to guarantee the malicious node in the system has a short/long run gain or loss.

A relay selection-based clustering scheme for high-density LTE Networks which is based on the Basic Sequential Algorithmic Scheme (BSAS) along with power control protocol is presented in [59]. The proposed scheme does not require any extra infrastructure such as other capacity improving schemes required small cells and relay stations. The BSAS is modified to make it more suitable for LTE networks and also to improve its performance. In the last two decades, MANETs have gained the attraction of the researchers due to its wide applications. Packets routing in these networks is a very difficult task due to the node's mobility and reliance on the small battery to remains active throughout the network. Cluster-based routing schemes for MANETs have advanced a vital substitute, enabling the network to be scalable and the capital to be well exploited. In clustering schemes, a cluster consists of a distinct node called a cluster head and acts as a controller in its cluster which accomplishes extra tasks as compared to other nodes in the network. Therefore, the cluster head becomes overloaded and drains its battery more rapidly. For this purpose, the authors in [60] presented a Light Weight Efficient Cluster (LWEC) based routing model for MANETs in which the life of cluster-head is prolonged by reducing the unnecessary load. Also, an efficient multi-hop proximity aware clustering scheme for mobile P2P systems (PCSM) is presented in [61] which based on the physical proximity of peers. This strategy aims at reducing the gap between the overlay of P2P and the network layer. The PCSM integrates three features to let the novel peer to professionally select the cluster to join such as cluster size, number of physical hops, and cluster head availability. The approach further aims to enhance the results of the search procedure in terms of the delay in a typical file-discovery and false-negative ratio.

MANETs are a collection of active nodes that facilitate communication from the transmitter to the receiver by either single or multi-hop processing. Within the network, nodes hold energy constraints that use an effective clustering mechanism to facilitate communication between nodes inside and outside clusters by scheduling dynamic hybrid topology (DHT). Therefore, in [62], the authors presented a clustering scheme named EBCH to reduce the consumption of energy during the process of communication from the transmitter to the receiver. Several parameters are investigated to determine the cluster head selection based on energy consumption because it is related directly to the network life-time. A comparative analysis of the clustering algorithms is shown in Table 8.

VIII. SUMMARY AND CONCLUSION

MANETs can be deployed as an emergency network in case of failure of the mains network due to any disaster situation because of its scalable and adoptive nature in the presence of environmental changes. Recently researchers have presented some new clustering schemes in MANETs to improve the overall performance of a network but these schemes still possess some limitations. To explain the classification of clustering-based schemes in MANETs and their limitations we have organized our paper in such a way that future researchers can understand the existed clustering schemes with their benefits and limitations. This survey also summarizes every category of clustering scheme concerning some important parameters that affect the overall performance as shown in Table 7. One of the design challenges for researchers in MANETs is to propose a scheme that is more scalable, robust and adaptive in case of high mobility. We have proposed a cross-layer clustering framework and HSOCM as a solution to improve QoS and minimize the research challenges of cluster-based MANETs.

More research should be conducted to consider high mobility scenarios for the procedure of CH selection and cluster maintenance. No specific clustering algorithm is preferable in all situations, therefore, there is a need for such a clustering algorithm that can be adaptive to handle different types of situations. More study is recommended to achieve the enhanced performance in term of QoS and connectivity between nodes while improving the overall lifespan of the network with minimum consumption of energy. Due to the latest advancement in technology researchers also need to design such a clustering-based scheme which can provide self-optimization, self-healing, and adaptivity according to the environmental changes to fulfill the needs of the future. This will motivate researchers to consider the tradeoff between different factors related to performance and will provide a suitable solution for network design. New clustering-based schemes with advanced features will open new doors for the deployment of MANETs in many daily life applications.

REFERENCES

[1] H. Yang and Z. Li, "A clustering optimization algorithm based on WCA in MANET," in *Proc. 6th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2015, pp. 1–5.

[2] B. Muneeswari and M. Manikandan, "Energy efficient clustering and secure routing using reinforcement learning for three-dimensional mobile ad hoc networks," *IET Commun.*, vol. 13, no. 12, pp. 1828–1839, Jul. 2019.

[3] J. Yu and P. Chong, "A survey of clustering schemes for mobile ad hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 7, no. 1, pp. 32–48, May 2005.

[4] H. Zhou and J. Zhang, "An efficient clustering algorithm for MANET based on weighted parameters," in *Proc. 8th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2015, pp. 145–148.

[5] M. Ashwin, S. Kamalraj, and M. Azath, "Weighted clustering trust model for mobile ad hoc networks," *Wireless Pers. Commun.*, vol. 94, no. 4, pp. 2203–2212, Jun. 2017.

[6] T. R. Siddiqi, H. Ning, H. Ping, and Z. Mahmood, "DPCA: Data prioritization and capacity assignment in wireless sensor networks," *IEEE Access*, vol. 5, pp. 14991–15000, 2017.

[7] S. Pathak and S. Jain, "A novel weight based clustering algorithm for routing in MANET," *Wireless Netw.*, vol. 22, no. 8, pp. 2695–2704, Nov. 2016.

[8] M. N. Anjum, C. Chowdhury, and S. Neogy, "Implementing mobile agent based secure load balancing scheme for MANET," in *Proc. Int. Conf. Opto-Electron. Appl. Opt. (Optronix)*, Mar. 2019, pp. 1–6.

[9] M. Aissa and A. Belghith, "An efficient scalable weighted clustering algorithm for mobile ad hoc networks," in *Proc. 3rd Int. Conf. Inf. Technol. e-Services (ICITeS)*, Mar. 2013, pp. 1–6.

[10] S. Dhamodharavadhani, "A survey on clustering based routing protocols in Mobile ad hoc networks," in *Proc. Int. Conf. Soft-Comput. Netw. Secur. (ICSNS)*, Feb. 2015, pp. 1–6.

[11] M. Alinci, E. Spaho, A. Lala, and V. Kolici, "Clustering algorithms in MANETs: A review," in *Proc. 9th Int. Conf. Complex, Intell., Softw. Intensive Syst.*, Jul. 2015, pp. 330–335.

[12] S. M. Almheiri and H. S. Alqamzi, "MANETs and VANETs clustering algorithms: A survey," in *Proc. IEEE 8th GCC Conf. Exhib.*, Feb. 2015, pp. 1–6.

[13] M. Gavhale and P. D. Saraf, "Survey on algorithms for efficient cluster formation and cluster head selection in MANET," *Procedia Comput. Sci.*, vol. 78, pp. 477–482, 2016.

[14] S. Mehta, P. Sharma, and K. Kotecha, "A survey on various cluster head election algorithms for MANET," in *Proc. Nirma Univ. Int. Conf. Eng.*, Dec. 2011, pp. 1–6.

[15] A. Bentaleb, A. Boubetra, and S. Harous, "Survey of clustering schemes in mobile ad hoc networks," *Commun. Netw.*, vol. 5, no. 2, pp. 8–14, 2013.

[16] W. A. Jabbar, M. Ismail, R. Nordin, and S. Arif, "Power-efficient routing schemes for MANETs: A survey and open issues," *Wireless Netw.*, vol. 23, no. 6, pp. 1917–1952, Aug. 2017.

[17] G. A. Walikar and R. C. Biradar, "A survey on hybrid routing mechanisms in mobile ad hoc networks," *J. Netw. Comput. Appl.*, vol. 77, pp. 48–63, Jan. 2017.

[18] M. Chatzidakis and S. Hadjiefthymiades, "Location aware clustering and epidemic trust management in mobile ad hoc network," in *Proc. 28th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2019, pp. 1–7.

[19] S. Sharma and S. Kumar, "Techniques for real-world implementation of a MANET," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COMITCon)*, Feb. 2019, pp. 519–524

[20] T. Rahman, X. Yao, and G. Tao, "Consistent data collection and assortment in the progression of continuous objects in IoT," *IEEE Access*, vol. 6, pp. 51875–51885, 2018.

[21] I. Ullah, J. Chen, X. Su, C. Esposito, and C. Choi, "Localization and detection of targets in underwater wireless sensor using distance and angle based algorithms," *IEEE Access*, vol. 7, pp. 45693–45704, 2019.

[22] A. H. Hussein, A. O. A. Salem, and S. Yousef, "A flexible weighted clustering algorithm based on battery power for mobile ad hoc networks," in *Proc. IEEE Int. Symp. Ind. Electron.*, Jun. 2008, pp. 2102–2107.

[23] A. Fathi and H. Taheri, "Enhance topology control protocol (ECEC) to conserve energy based clustering in wireless ad hoc networks," in *Proc. 3rd Int. Conf. Comput. Sci. Inf. Technol.*, vol. 9, 2010, pp. 356–360.

[24] S. Kaur, R. C. Gangwar, and R. Singh, "A strength based energy efficient algorithmic approach in MANET," in *Proc. Int. Conf. Soft Comput. Techn. Implement. (ICSCTI)*, 2015, pp. 141–145.

[25] M. Saxena, N. Phate, K. J. Mathai, and M. A. Rizvi, "Clustering based energy efficient algorithm using max-heap tree for MANET," in *Proc. 4th Int. Conf. Commun. Syst. Netw. Technol.*, 2014, pp. 123–127.

[26] A. R. Mohindra and C. Gandhi, "An energy-efficient clustering approach for collaborative data forwarding in heterogeneous MANET," *Int. J. Commun. Syst.*, vol. 30, no. 18, p. e3366, Dec. 2017.

[27] I. I. Er and W. K. G. Seah, "Mobility-based D-hop clustering algorithm for mobile ad hoc networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, vol. 4, Mar. 2004, pp. 2359–2364.

[28] S. Saha and R. Chaki, "Cluster based mobility considered routing protocol for mobile ad hoc network," in *Proc. Int. Conf. Comput. Sci. Inf. Technol.*, 2011, pp. 33–42.

[29] M. Ni, Z. Zhong, H. Wu, and D. Zhao, "A new stable clustering scheme for highly mobile ad hoc networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.

[30] M. Ni, Z. Zhong, and D. Zhao, "MPBC: A mobility prediction-based clustering scheme for ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4549–4559, Nov. 2011.

[31] M. Cai, L. Rui, D. Liu, H. Huang, and X. Qiu, "Group mobility based clustering algorithm for mobile ad hoc networks," in *Proc. 17th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Aug. 2015, pp. 340–343.

[32] H. Wu, Z. Zhong, and L. Hanzo, "A cluster-head selection and update algorithm for ad hoc networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–5.

[33] D. Gavalas, G. Pantziou, C. Konstantopoulos, and B. Mamalis, "Lowest-ID with adaptive ID reassignment: A novel mobile ad-hoc networks clustering algorithm," in *Proc. 1st Int. Symp. Wireless Pervas. Comput.*, Apr. 2006, p. 5.

[34] S. Mangai and A. Tamilarasi, "Improved location aided cluster based routing protocol for GPS enabled MANETs," in *Proc. Int. Conf. Future Gener. Inf. Technol.*, 2010, pp. 606–615.

[35] P. Thakur and A. Ganpati, "Cluster based route discovery technique for routing protocol in MANET," in *Proc. Int. Conf. Green Comput. Internet Things (ICGCIoT)*, Oct. 2015, pp. 622–626.

[36] S. Pathak and S. Jain, "An optimized stable clustering algorithm for mobile ad hoc networks," *EURASIP J. Wireless Commun. Netw.*, vol. 1, no. 1, p. 51, 2017.

[37] S. S. Yau and W. Gao, "Multi-hop clustering based on neighborhood benchmark in mobile ad-hoc networks," *Mobile Netw. Appl.*, vol. 12, nos. 5–6, pp. 381–391, Dec. 2007.

[38] K. T. Mai, D. Shin, and H. Choo, "Connectivity-based clustering with stretching technique in MANETs," in *Proc. 3rd Int. Conf. Ubiquitous Inf. Manage. Commun. (ICUIMC)*, 2009, pp. 200–206.

[39] S. Muthuramalingam, M. Sujatha, R. Surya, and R. Rajaram, "An enhanced sectorized clustering scheme based on transmission range for MANETS," in *Proc. Int. Conf. Recent Trends Inf. Technol. (ICRTIT)*, Jun. 2011, pp. 269–274.

[40] A. Bentaleb, S. Harous, and A. Boubetra, "A new topology management scheme for large scale mobile ad hoc networks," in *Proc. IEEE Int. Conf. Electro/Inf. Technol. (EIT)*, May 2015, pp. 31–37.

[41] A. Bentaleb, S. Harous, and A. Boubetra, "A k-hop scalable management scheme for large scale mobile ad hoc networks," *Wireless Pers. Commun.*, vol. 96, no. 4, pp. 6239–6271, Oct. 2017.

[42] X. Bli, "Genetic algorithm simulated annealing based clustering strategy in MANET," in *Proc. Int. Conf. Natural Comput.* Berlin, Germany: Springer, 2005, pp. 1121–1131.

[43] H. Cheng, "Genetic algorithms with hyper-mutation for dynamic load balanced clustering problem in mobile ad hoc networks," in *Proc. 8th Int. Conf. Natural Comput.*, May 2012, pp. 1171–1176.

[44] S. Sett and P. K. G. Thakurta, "Effect of optimal cluster head placement in MANET through multi objective GA," in *Proc. Int. Conf. Adv. Comput. Eng. Appl.*, Mar. 2015, pp. 832–837.

[45] M. Siwach and A. Singh, "Fuzzy based parametric clustering model for mobile network," in *Proc. Int. Conf. Soft Comput. Techn. Implement. (ICSCTI)*, Oct. 2015, pp. 49–53.

[46] K. P. R. Prabha and N. Jeyanthi, "A trust and fuzzy cluster based dynamic secure routing algorithm for mobile ad hoc networks," *Wireless Pers. Commun.*, vol. 98, no. 3, pp. 2959–2974, Feb. 2018.

[47] F. D. Tolba, D. Magoni, and P. Lorenz, "Connectivity, energy and mobility driven clustering algorithm for mobile ad hoc networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2007, pp. 2786–2790.

[48] T.-C. Huang, K.-P. Kho, and L. Tang, "Hybrid routing protocol based on the k-hop clustering structure for MANETs," in *Proc. 2nd Int. Conf. Intell. Netw. Intell. Syst.*, Nov. 2009, pp. 197–200.

[49] C. Sathiyakumar and K. Duraiswamy, "A hybrid approach for node co-operation based clustering in mobile ad hoc networks," *J. Comput. Sci.*, vol. 9, no. 2, pp. 147–154, Feb. 2013.

[50] J. Sathiamoorthy and B. Ramakrishnan, "A reliable data transmission in EAACK MANETs using hybrid three-tier competent fuzzy cluster algorithm," *Wireless Pers. Commun.*, vol. 97, no. 4, pp. 5897–5916, Dec. 2017.

[51] J. Sathiamoorthy and B. Ramakrishnan, "Energy and delay efficient dynamic cluster formation using hybrid AGA with FACO in EAACK MANETs," *Wireless Netw.*, vol. 23, no. 2, pp. 371–385, Feb. 2017.

[52] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.

[53] F. Aftab, Z. Zhang, and A. Ahmad, "Self-organization based clustering in MANETs using zone based group mobility," *IEEE Access*, vol. 5, pp. 27464–27476, 2017.

[54] J. Wang, Y.-Q. Cao, B. Li, S.-Y. Lee, and J.-U. Kim, "A glowworm swarm optimization based clustering algorithm with mobile sink support for wireless sensor networks," *Internet Technol. J.*, vol. 16, no. 5, pp. 825–832, 2015.

[55] A. L. Bolaji, M. A. Al-Betar, M. A. Awadallah, A. T. Khader, and L. M. Abualigah, "A comprehensive review: Krill Herd algorithm (KH) and its applications," *Appl. Soft Comput.*, vol. 49, pp. 437–446, Dec. 2016.

[56] G. Krishnasamy, "An energy aware fuzzy trust based clustering with group key management in MANET multicasting," in *Proc. 2nd Int. Conf. New Trends Comput. Sci. (ICTCS)*, 2019, pp. 1–5.

[57] G. V. Swaroop and G. Murugaboopathi, "Secure and reliable communication scheme for MANET using ECMS cluster head-based certificate revocation," *Cluster Comput.*, vol. 22, no. S5, pp. 11513–11525, Sep. 2019.

[58] S. Balaji, E. G. Julie, Y. H. Robinson, R. Kumar, P. H. Thong, and L. H. Son, "Design of a security-aware routing scheme in mobile ad-hoc network using repeated game model," *Comput. Standards Inter.*, vol. 66, Oct. 2019, Art. no. 103358.

[59] M. Hajjar, G. Aldabbagh, N. Dimitriou, and M. Z. Win, "Relay selection based clustering techniques for high density LTE networks," *Wireless Netw.*, vol. 25, no. 5, pp. 2305–2314, Jul. 2019.

[60] P. Dixit, A. Pillai, and R. Rishi, "A light weight efficient cluster based routing model for mobile ad hoc networks (LWECM)," *Int. J. Inf. Technol.*, pp. 1–7, Jan. 2019, doi: 10.1007/s41870-019-00291-0.

[61] M. Rahmani and M. Benchaïba, "PCSM: An efficient multihop proximity aware clustering scheme for mobile peer-to-peer systems," *J. Ambient Intell. Hum. Comput.*, vol. 10, no. 11, pp. 4243–4260, Nov. 2019.

[62] K. Sindhanaiselvan, J. M. Mannan, and S. K. Aruna, "Designing a dynamic topology (DHT) for cluster head selection in mobile adhoc network," *Mobile Netw. Appl.*, pp. 1–9, Jan. 2019, doi: 10.1007/s11036-019-01283-x.

TAJ RAHMAN received the B.S. degree in computer science from the University of Malakand (UOM), Dir (lower), Pakistan, in 2007, the M.S. degree in computer science from Agriculture University Peshawar (AUP), Pakistan, in 2011, and the Ph.D. degree in computer science from the School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), China. He is currently an Assistant Professor with the Department of Computer Science and IT, Qurtuba University of Science and Technology, Peshawar, Pakistan. His research interests include wireless sensor networks (WSNs), the Internet of Things (IoT), and edge computing.

INAM ULLAH received the B.Sc. degree in electrical engineering (telecommunication) from the Department of Electrical Engineering, University of Science and Technology Bannu (USTB), KPK, Pakistan, in 2016, and the master's degree in information and communication engineering from the College of Internet of Things (IoT) Engineering, Hohai University (HHU), Changzhou Campus, China, in 2018, where he is currently pursuing the Ph.D. degree with the College of IoT Engineering. He was an Assistant Engineer with ZTE Islamabad, Pakistan, in 2016. His research interests include wireless sensor networks (WSNs), underwater sensor networks (USNs), edge/fog computing, and artificial intelligence.

Dr. Ullah's awards and honors include the Best Student Award from USTB, in 2015, Top-10 students award of the College of Internet of Things (IoT) Engineering, in 2019, Top-100 students award of Hohai University (HHU), in 2019, and Jiangsu Province Distinguish International students award, in 2020.

ATEEQ UR REHMAN was born in Rawalpindi, Pakistan, in 1987. He received the B.S. degree in electrical (telecommunication) engineering from COMSATS University, Lahore, Pakistan, in 2009, and the M.S. degree in electrical engineering with a specialization in telecommunications from the Blekinge Institute of Technology (BTH), Karlskrona, Sweden, in 2011. He is currently pursuing the Ph.D. degree with the College of Internet of Things (IoT) Engineering, Hohai University (HHU), Changzhou Campus, China. From 2011 to 2012, he was a Lecturer with the Department of Electrical Engineering, University of South Asia, Lahore. From 2012 to 2016, he was with the Faculty of Engineering and Technology, The Superior College (University Campus) Lahore, Pakistan, as a Lecturer. Since 2016, he has been working as a Lecturer with the Government College University (GCU), Lahore. His research interests include biomedical signal processing, the Internet of Things (IoTs), Social Internet of Things (SIoTs), and big data.

RIZWAN ALI NAQVI received the B.S. degree in computer engineering from COMSATS University, Pakistan, in 2008, the M.S. degree in electrical engineering from Karlstad University, Sweden, in 2011, and the Ph.D. degree in electronics and electrical engineering from Dongguk University, South Korea, in 2018. From 2011 to 2012, he was a Lecturer with the Computer Science Department, Sharif College of Engineering and Technology, Pakistan. He joined the Faculty of Engineering and Technology, The Superior College, Pakistan, as a Senior Lecturer, in 2012. After his Ph.D. degree, he worked as a Postdoctoral Researcher with Gachon University, South Korea, from 2018 to 2019. He is currently working as an Assistant Professor with Sejong University, South Korea. His research interests include gaze tracking, biometrics, computer vision, artificial intelligence, machine learning, deep learning, and medical imaging analysis.