

Adaptive k-means clustering for Flying Ad-hoc Networks

Ali Raza¹, Muhammad Fahad Khan¹, Muazzam Maqsood¹, Bilal Haider¹ and Farhan Aadil^{1*}

¹Department of Computer Science, COMSATS University Islamabad,
Attock Campus, Pakistan

*Corresponding author: Farhan Aadil

[E-mail : farhan.aadil@cuiatk.edu.pk]

*Received January 8, 2019; revised March 1, 2020; accepted March 20, 2020;
published June 30, 2020*

Abstract

Flying ad-hoc networks (FANETs) is a vibrant research area nowadays. This type of network ranges from various military and civilian applications. FANET is formed by micro and macro UAVs. Among many other problems, there are two main issues in FANET. Limited energy and high mobility of FANET nodes effect the flight time and routing directly. Clustering is a remedy to handle these types of problems. In this paper, an efficient clustering technique is proposed to handle routing and energy problems. Transmission range of FANET nodes is dynamically tuned accordingly as per their operational requirement. By optimizing the transmission range packet loss ratio (PLR) is minimized and link quality is improved which leads towards reduced energy consumption. To elect optimal cluster heads (CHs) based on their fitness we use k-means. Selection of optimal CHs reduce the routing overhead and improves energy consumption. Our proposed scheme outclasses the existing state-of-the-art techniques, ACO based CACONET and PSO based CLPSO, in terms of energy consumption and cluster building time.

Keywords: Energy optimization, Clustering, FANET, k-means, Routing, Transmission range optimization

1. Introduction

Flying ad-hoc network is a new type of vehicular ad-hoc networks (VANETs) and mobile ad-hoc networks (MANETs) [1]. It inherits many of its characteristics from VANET and MANET. It follows the same peer to peer communication mechanism. Nodes depend on each other for successful communication in the network [2]. Due to the fast movement of nodes, topology of network changes very frequently [3]. FANET also has some unique characteristics which differentiate it from other ad-hoc networks [4]. In FANET, UAVs are mobile nodes. Micro and macro UAVs fly in the free air space with the speed of almost 30-60 km/h [5]. They have relatively large inter-node distances and a clear line of sight for direct communication. Mobility patterns of nodes are also different in FANET [6]. FANET has a large number of applications from surveillance, monitoring to disaster management, environment sensing, ecological damage detection and communication relying [4, 7, 8].

UAVs are furnished with a small lithium battery which can stand flight time of about 30-35 minutes [1, 4]. UAV replacement strategy is defined to complete a long operation. When a UAV is near to the end of its battery energy, it returns towards the base station and an energized UAV is inducted into the network to complete the operation. This short lifetime and frequent replacement of UAVs puts a barrier in the wide applications of FANET [9].

Additionally, designing a reliable communication protocol in such high dynamic environments is also a big challenge [10]. Inefficient communication does not only degrade the performance of the network and inhibits to achieve the desired objective (i.e. collecting data and sending to the base station), but also waste the energy and shorten the lifetime of UAVs [11-13]. Broadly speaking, routing protocols of ad-hoc networks can be categorized in three paradigms: reactive protocols, proactive protocols, and cluster (or hierarchical) based routing protocols [4]. Routing tables are maintained in proactive routing protocols. Routes to other nodes are maintained in these tables. Routing tables are periodically updated through the exchange of route status messages. On the other hand, reactive routing protocols do not require to maintain routing tables. In an event to send data, a node finds a new route and continues to send until the current route disrupts. On the route disruption, the new route is to be found and this procedure is repeated until the node has sent the required data [6]. In FANET where high-speed movement of nodes causes to change network very frequently. In such high dynamic environment, maintaining an up to date routing table or finding a new route, again and again, puts a massive communication overhead [1].

Clustering aims to minimize the route finding and maintaining overheads [14, 15]. It divides the whole network into multiple logical groups. Nodes located geographically nearby to each other are grouped into a cluster. Each cluster maintains a cluster head (CH) which is liable for maintaining and controlling the communication in the cluster. It manages inter-cluster and intra-cluster communication. CH always acts as a first hop node for all of their cluster members (CMs) and it is in charge for delivering the message towards eventual destination. CHs exchange status messages with their CMs to keep updated information about the cluster [16]. The process of clustering and election of CH is in lieu of route request in proactive and reactive routing. Cluster overhead is mainly determined by cluster formation and maintenance phases [15, 17]. During cluster formation, a number of messages are exchanged between network nodes. For example, they exchange their position and other parameters to calculate node fitness. Using these parameters, they perform computations to elect CHs. The exchange of messages and performing computations to form clusters and elect CHs wastes the network resources such as bandwidth, computational power and battery energy etc. Cluster lifetime is

the time period between two cluster formation phases. The short cluster life period will call the cluster formation phases again and again, which puts huge computational and communicational overheads [16]. In the given flight time or simulation time of UAVs, longer the cluster lifetime, better will be the performance of the routing algorithm.

The key objective of this paper is to design a cluster based communication protocol for FANET that minimizes routing overhead and enhance the lifetime of nodes. Selection of transmission power is very much crucial for node energy consumption. The proposed framework select the optimal transmission rule of nodes according to the network requirement such as average distance among nodes. k-means [18] (with uniform distributed initial centroid) is used to perform clustering in the network. The value of 'k' i.e. a number of clusters is learned for each network topology. After clustering, a CH is elected for each cluster and nodes start communicating with each other and base station through their respective CHs. Nodes follow reference point group mobility model (RPGM) to coordinate their moments with CH.

The paper is further divided into following sections. Section 2 contains the related work, Section 3 is used to depict the material and methods used in the paper. Results are added in Section 4. Furthermore the conclusion is added in Section 5.

2. RELATED WORK

The different algorithms have been implemented for solving the clustering problem in ad-hoc Networks. The algorithms are based on grey wolf, ant colony and particle swarm optimization [19-21]. In the mentioned research the different algorithms produce the set of solutions. In the set of solution different nodes can be designated as a CH. The process is repeated for the particular number of times and the aforementioned method repeats to select the CH. The main advantage of these evolutionary techniques is that they don't stick on local optimum due to the involvement of the random factor.

Farhan et al. proposed CACONET [22], based on the ACO. The living behavior of ants is used to extract the set of rules. These rules are further molded according to the scenario of vehicular network. Afterward, these set of rules are implemented to find the set of population for selecting the CH. Consequently, the ACO based CACONET is used to provide the optimal number of clusters.

Shahzad et al. proposed the Comprehensive Learning PSO (CLPSO) [23] for clustering in MANET. CLPSO used the PSO to give optimal numbers of clusters and CHs. It considers the different parameters for CH selection like mobility, ideal degree, residual energy, and transmission power of the nodes. CLPSO assigns some weight for mentioned parameters and nodes are elected as CH basis on their weight. Hamid at al. [24] proposed another variant of PSO i.e. multi-objective PSO (MOPSO) to give multiple solutions for any problem. MOPSO claims that only one solution cannot be considered as enough in continuous nature problems.

GWO [21] used the living nature of wolves to exploit the different problems. The complete formation is divided into four steps such as, alpha, beta, delta, and omega. The strength level is also distributed from higher level (alpha) to lower level (beta) and so on. Afterwards, Fahad et al. [25] implement the GWO based algorithm called as GWOCNET for solve the above mentioned problem.

The usage of artificial intelligence is increasing day by day. Whereas, there is a drawback of using these method is high complexity. These methods consume the high resources for the simulation. This is due to the randomness nature of these algorithms [26]. Therefore, the

randomness converges these algorithm very early. However, in FANETs the UAVs frequently changes the topology due to the mobile nature of UAVs.

Apart from artificial intelligence based clustering algorithms, k-hop clustering has also been used in ad-hoc networks such as KHOPCA and GWO. K-hop clusters are those where CMs are at most k-hops away from CH. CM are connected with CH either directly or through intermediate nodes. The main advantage of k-hop clustering is that it has relatively less number of cluster heads. In the literature, a network topology having less number of clusters and their CHs is considered to be a better approach. It minimize the routing overhead and save nodes energy. But k-hop clustering is suitable for networks where nodes mobility is relatively static or slow. In FANET, very nodes mobility is high, large cluster area and high number of cluster members causes frequent changes in cluster structure which results in high clustering overhead. Moreover, in k-hop clustering nodes keeps the information about their CHs and intermediate nodes which connects them to CH. In FANET like networks where node mobility is high, intermediate nodes changes very frequently and causes the nodes to update their information accordingly.

The graph based clustering [27] has been used for the learning in kernel space to shown impressive performance on a number of benchmark data sets. Another model [28] has been proposed for boosting the performance of data clustering, semisupervised classification, and data recovery significantly, primarily due to two key factors: 1) enhanced low-rank recovery by exploiting the graph smoothness assumption and 2) improved graph construction by exploiting clean data recovered by robust principle component analysis and clustering has been used for kernel-driven similarity learning [29].

Clustering [30] has been proposed for wireless sensor network for network management based on the policy for independently organize the network.

In 1-hop clustering [31], nodes only contain ID of their CH. They send messages directly to CHs. If the destination node located within the same cluster, CH send directly to destination node. If destination node is situated in other cluster, CH send the message to CH of destination node, which ultimately delivers the message to destination node.

3. PROPOSED MODEL

This section provides the detailed description of the proposed model. The proposed model has two objectives:

- 1) Minimize computational complexity
- 2) Minimize the communicational overhead.

Both objectives will result in an enhanced lifetime of nodes. The first objective is achieved by using k-means with uniformly distributed initial centroids for clustering. This keeps the clustering mechanism much simple and converges towards optimal solution very quickly. The second objective is achieved by electing optimal CHs and enhancing the lifetime of the cluster. **Fig. 1** shows the pictorial view of proposed model, and algorithm 1 outlines the complete process of the proposed model. Each step of the algorithm is described below:

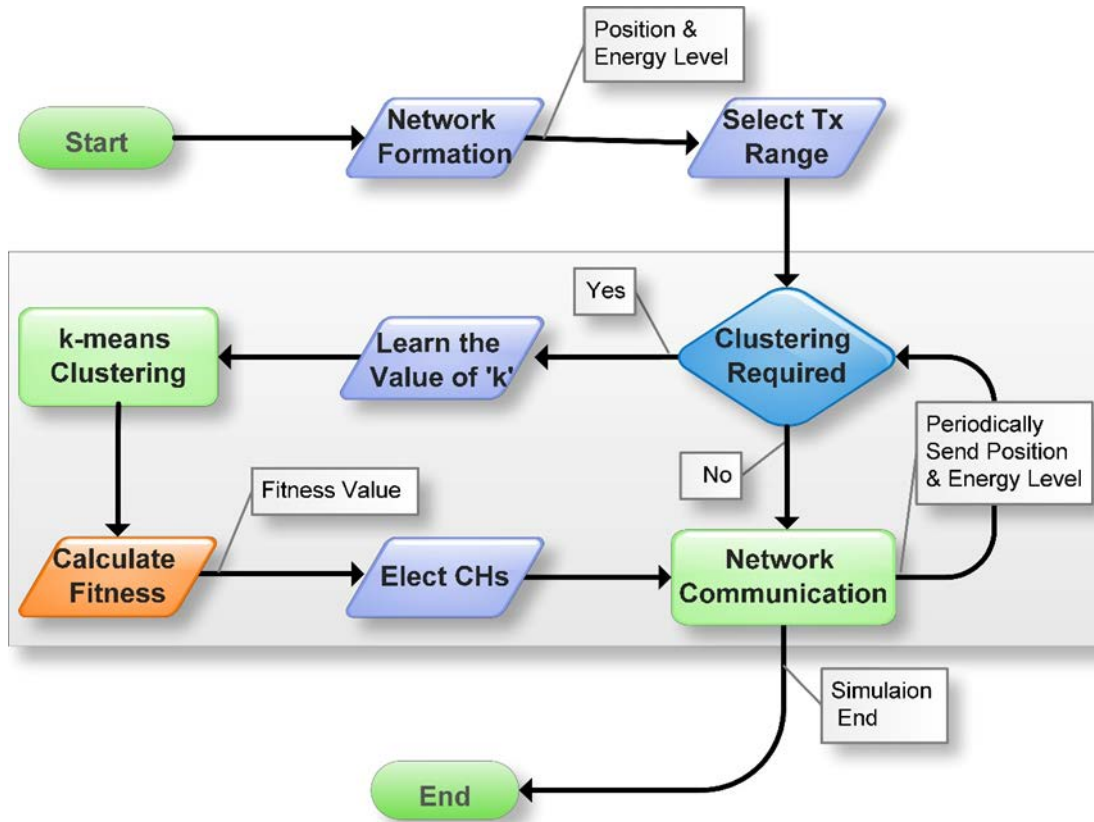


Fig. 1. Adoptive k-means clustering model for communication in FANET

3.1. NETWORK FORMATION

The network is initialized after the beginning of flight of UAVs. UAV's are equipped with multiple sensors. Some of them are task-oriented such as cameras, temperature or humidity sensors. Whereas, others provide assistance to achieve the task, such as GPS and altitude sensors deliver the three dimensional position of node. To communicate with other nodes, we assume that UAVs are also equipped with wireless communication chip that has different four strength levels that relates to power range of 400, 600, 800 and 1000 meters correspondingly. Initially, the higher level of power is selected by the node and later on signal power is adjusted accordingly based on the three dimensional location. This will provide the saving of energy for the nodes.

Selection of transmission power is directly related to transmission range. Equation 1 determines the relation between transmission power and range [8]. Transmission range is the power of signal at the time of emitting from the antenna at transmitting node.

$$P_T = P_R + 20 \log\left(\frac{4\pi R}{\lambda}\right) - G_T - G_R \quad (1)$$

P_T is transmit power required to achieve R range. P_R is the receiver sensitivity (i.e. minimum ability to detect an RF signal), G_T and G_R are antenna gains of transmitting and receiving node. Equation 1 dictates that a node with higher transmission power will have long transmission range and better connectivity with neighbour nodes. It also has a higher degree of the

neighbourhood to which it is directly connected. On the other hand, high transmission range will consume more energy that shortens the lifetime of the node.

Algorithm 1: Cluster based Routing protocol for FANET

```

1.  START
2.  Position the node in the 3D grid
3.  Node  $n_i \in N$  broadcasts the position and residual energy information
    message
4.  Calculate average distance D for each  $n$  in the same sub-grid
5.  Adopt optimal  $P_T$  based on  $D$ 
6.  WHILE nodes are alive
7.      IF un-clustered nodes >20% of all nodes
8.          Learn the values of  $k$ 
9.          Select  $k$  number of uniformly distributed
            Centroids  $C_k$ 
10.          $n_{i,c} \in N \leftarrow k\text{-means (Node positions, } k, C_k)$ 
11.          $f\_val_i \leftarrow \text{Fitness}(\text{distance}, d\_diff, \text{residual energy})$ 
12.         FOR each Cluster  $C$ 
13.              $CH_C \leftarrow \max(f\_val)$ 
14.         END FOR
15.     ELSE
16.         Nodes follow the RPGM for their
            movement
17.         CMs start communication through CH
18.     END IF
19.     Periodically exchange distance, no. of
        neighbors, residual energy
20. END WHILE
21. END

```

Although the increase in transmission power increases the range, however, it has been evident through extensive experiments that increasing the transmission power two times does not double the range.

3.2. K-MEANS CLUSTERING

The process of grouping the nodes into multiple clusters is one of the most crucial tasks of the proposed model. It determines the performance of the algorithm. An efficient clustering process groups the most related nodes that lead to enhance the lifetime of cluster and reduce routing overhead. It can also accommodate subsequent topological changes in the network. In the proposed model, k-means is used to perform clustering of nodes in the FANET. k-means is one of the classical and most widely used clustering algorithm in the field of data science. To the best of our knowledge, k-means had not been used in the field of ad-hoc networks to perform clustering of network nodes. k-means requires three parameters (i.e. nodes position, the value of 'k' and position of initial centroids) as an input, perform clustering and returns association of each node to a particular cluster at the output. 'k' defines the number of clusters whereas centroid is the centre (mean) point of the cluster. k-means starts with choosing 'k'

number of random initial centroids. After that, positions of all nodes are scanned and Euclidean distance of each node from every centroid is computed. Assignment of nodes to a particular cluster is made based on the shortest distance from the centroid. Meanwhile, the mean value of each cluster is computed and centroids are updated to the mean value. Subsequently, the process of assignment and updating the mean is repeated until there is no change in the mean value and CM of all clusters.

3.3. 'k' LEARNING

'k' parameter determines the number of clusters that the k-means will produce. Given a fixed grid area the number of clusters is inversely proportional to the size of the cluster. Larger the size of the cluster, less the number of clusters will be generated. In ad-hoc networks, size of the cluster is dependent on average transmission range of nodes. Because all cluster member must be directly connected with their CHs, transmission range limits the size of the cluster. Long transmission range will correspond to larger cluster size. In other words, transmission range has an inverse relationship with a number of clusters and will be considered as an important parameter for setting the value of 'k'. If transmission range of nodes is known, the value of 'k' can be anticipated and set statically.

As mentioned earlier, transmission range is not known in advance. It is set dynamically during runtime based on the requirement of current network topology. If the nodes are sparsely located, they require high transmission range; if they are nearby located to each other, short transmission range can fulfil their needs to remain in contact with other nodes. Because of runtime setting of transmission range, we cannot assign the 'k' value at start time. As the setting of transmission range is at runtime, the setting of 'k' value will also be at runtime.

Transmission range is set according to the network requirement, 'k' value will also be set according to the network requirement.

Algorithm 2: Learning the value of k

```

1.   START
2.   FOR  $i=1$  to number of Solutions
3.     Remaining nodes  $\leftarrow$  All nodes
4.     WHILE (Remaining nodes  $\neq 0$ )
5.       Randomly select Leader node  $N_L$  from
         Remaining nodes
6.       Member  $\leftarrow$  All nodes in the transmission
         range
         of  $N_L$ 
7.       Remaining nodes  $\leftarrow$  Remaining nodes
         Member
8.     END WHILE
9.     TempK(i)  $\leftarrow$  Total number of  $N_L$ 
10.  END FOR
11.  k  $\leftarrow$  TempK with the highest Frequency
12.  END

```

In the proposed framework, some sort of brute force strategy is followed to learn the value of 'k'. We exhaustively tested different values of 'k' and choose the optimal one. Algorithm 2 describes the procedure for learning the value of 'k'. A search space for multiple solutions is created. Number of elements in a solution corresponds to a 'k' required to fulfil network

requirement (i.e. all nodes must lay in the direct transmission coverage of at least one CH). While constructing a single solution, elements are added to the solution subject to following conditions:

- Element is not already present in the solution
- Element does not come under the direct transmission range of any element in the solution.

A node (named Leader node) is elected randomly and added to the solution if it satisfies above mentioned two conditions. All nodes in the transmission coverage of Leader node are removed from remaining nodes. This procedure (while loop) is repeated until all nodes are evaluated in terms of Leader or member nodes. Number of elements in the solution (i.e. 'k') are stored in the temporary array. Multiple solutions (up to the size of search space) are created in the same way and the most frequent value from the temporary array is elected for the final value of 'k'.

The size of the cluster depends on the transmission range of CH. In the particular network scenario, the most frequent value of 'k' will provide enough amount of clusters that each node will be associated with at least one CH. It is also to be noted that more optimal the value of 'k', longer will be the cluster lifetime and less energy will be consumed. If the 'k' value is more than an optimal number of cluster for a specific network, more number of CHs will have to send data at long haul or to the other CHs in the network. This reduces the scalability of the network and increases inter-cluster communication. If 'k' value is less than optimal value, size of the clusters might be too large to satisfy network requirement (i.e. few of nodes does not have direct access to their CH). Even if it satisfies the network requirements, a large number of nodes in a cluster unbalance the network traffic. It introduces additional packet delays and congestion (both network and node congestion).

3.4. SELECTION OF INITIAL CENTROIDS

The performance of k-means is also much sensitive to the selection of initial centroids. Standard k-means use the random selection which unstable its performance. The randomness does not ensure the unique results. Subsequently, the selection of centroid also not make the effect on performance, whereas the number of iteration does make the effect on the performance. To enhance the efficiency and accuracy of k-means, Zahra S. et al [13] highlights 20 algorithms for initial centroids selection. Each algorithm is suitable for different applications based on the distribution of data in data set.

3.5. SELECTION OF CHs

Once the k-means performs the clustering of the network and groups the nearby located nodes into a single cluster, each node is associated with at least and at most one cluster. In order to start communication within the cluster, a CH is to be elected for each cluster. CH owns the responsibility to regularize and control the communication in the cluster. Selection of optimal CH enhances the cluster lifetime and balance the network traffic as well.

In the proposed model, selection of CHs is based on the fitness of nodes. Fitness determines the potential of a node to become CH. A node with higher potential has the ability to balance the network traffic, saves the energy of itself and other nodes, and prolongs the cluster life. Fitness value of a node is computed using equation 2.

$$Fitness = \frac{w1 * Energy_{Residual}}{(w2 * dist)(w3 * D_diff)} \quad (2)$$

$Energy_{Residual}$ is the remaining energy of the node; $dist$ is the sum of distances from CM; and D_diff is the delta difference. $w1$, $w2$ and $w3$ is the, weight of respective parameters. User

defines the value of these weights according to the importance of the parameter in a particular application scenario. D_diff is the variable used for balancing the load factor on clusters. It is used to equally distribute the load on CHs. The equally distribution means having the same number of cluster members for each cluster. It is the difference between ideal degree and node's number of neighbors. Larger the difference between ideal degree and node's neighbourhood, more unbalance the network will be. For example, let a network is formed of 30 nodes, which are divided into 3 clusters. If one cluster has 20 members whereas other two cluster possesses only 5 members each. Surely, this will unbalance the network traffic where one cluster is overwhelmed and other two are underutilized Delta Difference is also called as the degree of neighbors with respect to distance and can be computed by using equation 3.

$$D_diff = (Ideal_Degree - Node_Degree) \quad (3)$$

If situation is arrived where more than one nodes have same value for all three parameters i.e. energy, distance and d_diff , then a node with the higher degree (i.e., higher neighbour's association) is elected as CH. It increases the cluster stability and prolong the network lifetime.

3.6. NETWORK COMMUNICATION AND NODE MOVEMENT

After the completion of clustering process, when nodes are grouped into clusters and CHs are elected, nodes start communicating with other nodes or base station. Inter cluster and intra-cluster communication is performed through CHs. CMs do not maintain a full routing table. They only keep and update the latest information about their first hop node i.e. CH. It is the responsibility of CHs to maintain routes to other nodes. In order to provide network-wide access to nodes, CHs exchange route information of CM with other CHs. It is to be noted that rather than all nodes, only the CHs exchange route information; thereby reducing the huge amount of routing overhead. All communication took place through CHs. Whenever a node is required to send data, it will send to CH. If the destination node is also its CM, it will send directly to destination. If the destination node is from other clusters, CH will acquire the route to the destination from other CHs. Then CH of source node send data to the CH of destination node, which will eventually deliver to the destination.

In the proposed model, nodes follow Reference Point Group Mobility Model (RPGM). The proposed model ruminates the CHs as orientation points and all cluster member keep informed their place according to association of their CHs. Base station controls the position of CHs whereas CMs follow the CHs and adjust their direction, speed and position such that they always remain in the transmission coverage of CH.

4. EXPERIMENTAL RESULTS AND ANALYSIS

This section describes the experimental setup and evaluates the performance of the recommended model. Results are matched against Machine learning based clustering procedures i.e. ACO based CACONET and PSO based CLPSO. Performance of these clustering algorithms is compared in term of cluster lifetime, cluster building time and energy consumption.

The experimentations are conceded in MATLAB by considering the different factors such as; Grid Sizes and number of nodes. Parameters setting for simulation are shown in Table 1. We directed 10 simulations for each scenario and afterwards compute the average value.

In the simulated scenario, the initial network is formed in three-dimensional free space where position and direction of nodes are set randomly. The transmission power of nodes is selected according to its neighbourhood. When the CHs are elected, CM coordinate their movement with CHs and communicate through CHs. CHs consume more energy as compared to CM because they have to relay messages of their members. We use the term pending (un-clustered) nodes in order to define the clustering (re-clustering) criteria. Pending nodes are those nodes which require clustering. In the first iteration of the while loop, all nodes are pending nodes; whereas in subsequent iteration, two conditions are evaluated to declare a node as pending: 1) When energy of CH falls below 20% of energy of a cluster member that possess the highest energy in the cluster, all nodes in that cluster will be termed as pending nodes, 2) and while moving in the operational area, when two clusters come close enough such that they can merge into a single cluster, node of these two clusters will also be termed as pending nodes. When the total number of pending nodes in the network increases 20% of all nodes, clustering will be recalled. If the pending nodes are less than 20% of nodes in the network, then CHs will keep maintaining the cluster structure and nodes communicate through the elected CHs.

Table 1. Simulation Parameters

<i>Parameter</i>	<i>Value</i>
Grid Size	2 x 2 km^2 and 3 x 3 km^2
Number of Nodes	20, 30, 40, 50, and 60
Node Velocity	15 to 20 m/s
Minimum Distance between Nodes	2 m
Mobility Model	RPGM
Simulation Runs	10
Simulation Time	120 seconds
Position Exchange Interval	2 seconds
Node energy level at start time	80-Watt/hour
Transmission Range	400, 600, 800 and 1000 meters
Transmission Frequency	2.45 GHz
Constant Bit Rate	100 kbps
Receiver Sensitivity	-90 dBm
Stall Iterations (for CACONET and CLPSO)	10
Number of Solutions (for 'k' learning)	1000
Ideal Degree	10

4.1. CLUSTERS BUILDING TIME

The time required to make the clustering is called as cluster building time. The proposed algorithm takes the different input values. On the basis of following input values the CH is selected and their cluster members are also formed. Therefore, the time consumed from the taking the input and producing the output is referred as cluster formation or building time. This is also known as computational complexity of the algorithm.

As, UAVs are limited to memory and power so high cluster formation time lead towards the negative impact on the performance. It will also reduce UAVs life time by consuming more energy. The proposed method is compared with the existing methods; CLPSO and CACONET. The k-mean based algorithm consumes the less building time as compare to others. This is because of the reason that CLPSO and CACONET is based on the randomness and converges

iteratively. As shown in **Figs. 2 and 3**, increasing in the number of nodes will also increase the formation time of clusters. However, proposed model uses k-means to perform clustering, which generates only one solution and sequentially update it in order to move towards a global optimal solution. It is famous for its low time complexity. The low complexity support and provisioning the time for the route finding. It also helps in saving the energy of nodes.

The computational complexity of k-means is denoted by $ktn(\Theta)$, where ' k ', ' t ' and ' n ' are three variables,

k ; total clusters.

t ; total iterations

n ; total nodes

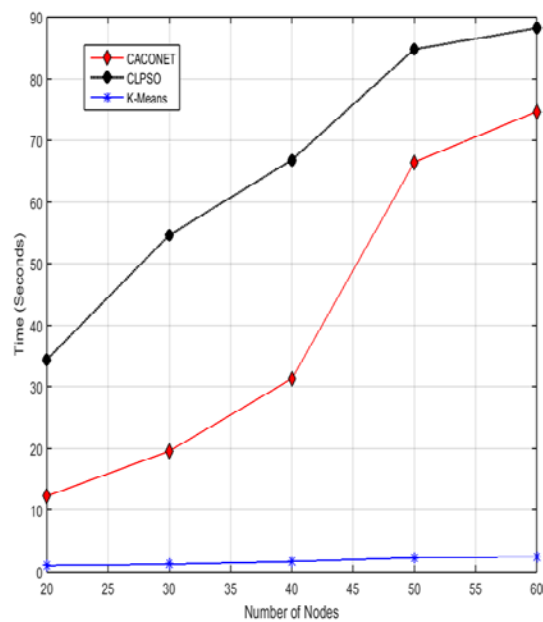


Fig. 2. Cluster Building Time with Grid size 2000x2000m

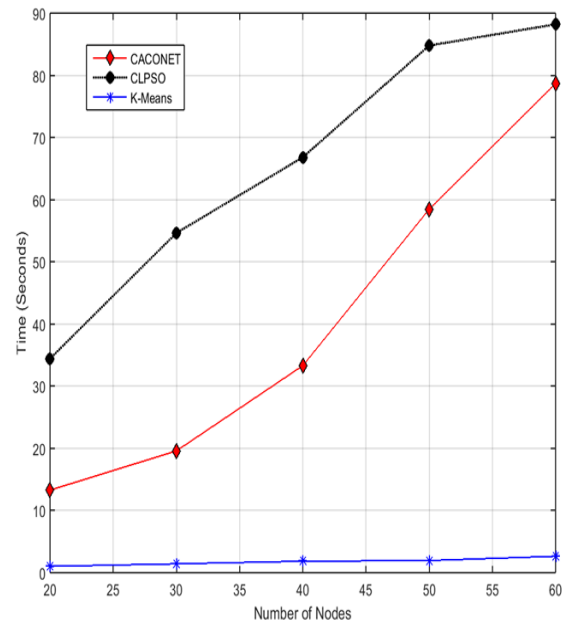


Fig. 3. Cluster Building Time with Grid size 3000x3000m

4.2. CLUSTER LIFETIME

The time elapsed from the formation of cluster to the moment at which cluster formation end is termed as cluster lifetime. It is dependent on multiple factors such as relative mobility, energy consumption rate of nodes and total number of clusters in the network. When the mobility is high, nodes may change their cluster assignments more rapidly. In the same way, when energy consumption rate is high, an elected CH may soon become invalid and re-clustering will be required. With the shorter lifetime of cluster, it enforces to recall the clustering factor again and again. Due to which the demand of computational power increases. **Figs. 4 and 5** show the contrast of cluster lifetime among proposed model, CACONET, and CLPSO. The proposed model performs much poor in this metric. Lifetime of cluster with k-means is very low as compared to the other two. The reason behind this poor performance is that k-means generate ' k ' number of clusters where the value of ' k ' is chosen as a most frequent value from multiple values of ' k ' for a given network environment. The value of ' k ' is insufficient to optimize the cluster size and enhance the cluster lifetime. Size of the cluster is too big to accommodate all cluster in the transmission coverage of CHs. The results show that the increment in the number

of nodes causes the decrement in the lifetime of cluster. This can be due to the mobile nature of UAVs which produces frequent change in topology.

4.3. ENERGY CONSUMPTION

The resources are scarce in FANETs and energy is one of the primary factor to maintain the network. The small dry batteries are normally used to provide the energy for micro UAVs.

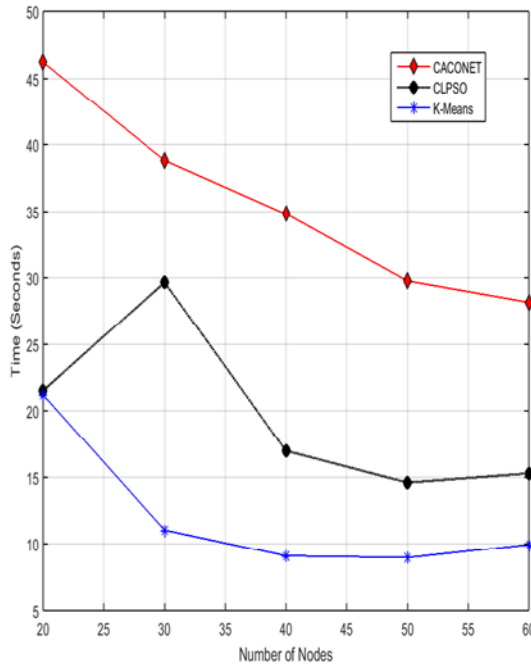


Fig. 4. Cluster Life Time with Grid size 2000x2000

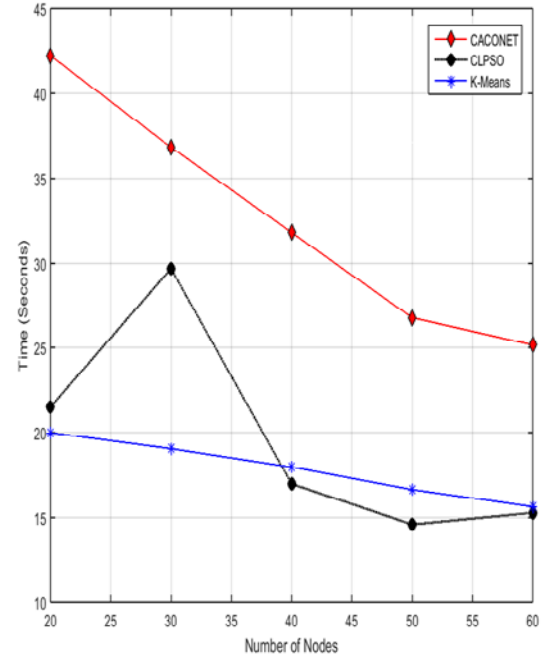


Fig. 5. Cluster Life Time with Grid size 3000x3000m

Due to the limited supply of energy to UAVs it is the obstacle in the way to application of UAVs. Therefore, the energy optimization is considerable topic now a days by the researchers. Basically there are three main mechanism for usage of energy in UAVs;

- i) Requirement of energy for flying in air.
- ii) Consumption of energy by sensors.
- iii) Demand of energy for communication purpose.

Most part of the energy consumed by UAVs is due to the power required for communication. Equation 4 shows the total energy consumed from transmission to reception [6, 24].

$$E_{com} = E_{Tx} + E_{Rx} \quad (4)$$

Where

$$E_{Tx}, L = P_T \times \frac{L}{\text{Bitrate}}$$

P_T represents maximum transmits power and 'L' shows number of bits transmitted. The proposed method is used for energy of whole network to calculate the total consumption of energy. CLPSO and CACONET for the time span of 120 seconds. **Figs. 6 and 7** show that energy consumption of the network increase with an increase in a nodes number. It is clearly shown in figures that, proposed model outclass the existing methods. The proposed model

demands the lower energy consumption. The reason of the low demand is due to the management of transmission power and optimize method used for the selection of CHs.

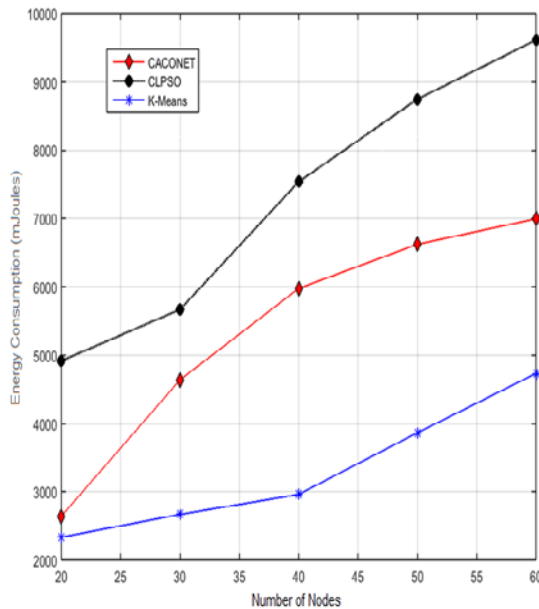


Fig. 6. Network Energy Consumption with Grid size 2000x2000m

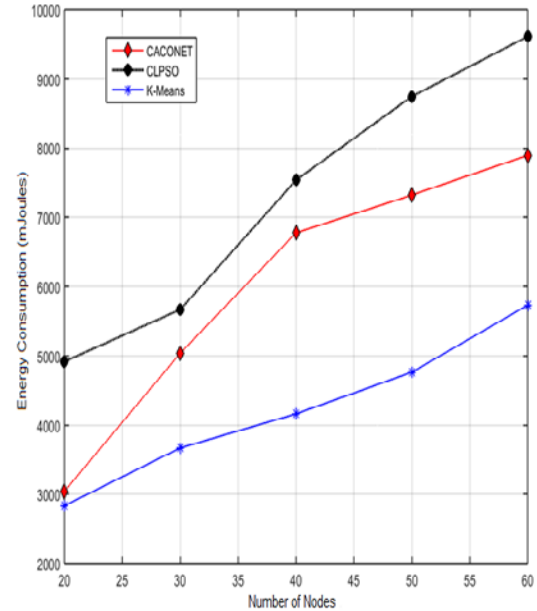


Fig. 7. Network Energy Consumption with Grid size 3000x3000m

5. CONCLUSION

The novel method for the FANET is proposed, which is based on the intelligence of k-means. There are two main limitation of mobile nodes such as, limited energy and optimized routing. In the proposed technique the routing method is optimized by using the clustering method. Energy consumption of UAVs are also reduced by managing the transmission power and effective clustering method. Afterwards, the outcomes are compared with the existing algorithms which are also based on machine learning techniques; CACONET and CLPSO. Results show that intelligent k-means outperforms the other two approaches. The results are compared by keeping focusing the factors of clusters building time and cluster lifetime and energy consumption. k-means has poor performance in respect of cluster lifetime. Cause of this poor performance is inefficient learning of 'k'. The future work can be the optimizing the 'k' value and select the optimal number of clusters that is best suitable for current network scenario and will enhance the lifetime of clusters. The different mobility models can also be considered for efficient routing.

6. References

- [1] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, "Flying ad-hoc networks (FANETs): A survey," *Ad Hoc Networks*, vol. 11, pp. 1254-1270, 2013. [Article \(CrossRef Link\)](#)
- [2] W. Shi, H. Zhou, J. Li, W. Xu, N. Zhang, and X. Shen, "Drone Assisted Vehicular Networks: Architecture, Challenges and Opportunities," *IEEE Network*, vol. 32, pp. 130-137, 2018. [Article \(CrossRef Link\)](#)

- [3] F. Aadil, S. Rizwan, and A. Akram, *Vehicular Ad Hoc Networks (VANETs), Past Present and Future: A survey*, Springer-Verlag Berlin Heidelberg, 2011.
- [4] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Communications Surveys & Tutorials*, vol. 18, pp. 1123-1152, 2016.
- [5] L. A. Tony, D. Ghose, and A. Chakravarthy, "Avoidance maps: A new concept in UAV collision avoidance," in *Proc. of Unmanned Aircraft Systems (ICUAS), 2017 International Conference on*, pp. 1483-1492, 2017. [Article \(CrossRef Link\)](#)
- [6] J. Xie, Y. Wan, J. H. Kim, S. Fu, and K. Namuduri, "A survey and analysis of mobility models for airborne networks," *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 1221-1238, 2014. [Article \(CrossRef Link\)](#)
- [7] A. Bujari, C. E. Palazzi, and D. Ronzani, "FANET application scenarios and mobility models," in *Proc. of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, pp. 43-46, 2017. [Article \(CrossRef Link\)](#)
- [8] Z. Kaleem and M. H. Rehmani, "Amateur Drone Monitoring: State-of-the-Art Architectures, Key Enabling Technologies, and Future Research Directions," *IEEE Wireless Communications*, vol. 25, pp. 150-159, 2018. [Article \(CrossRef Link\)](#)
- [9] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on Unmanned Aerial Vehicle Networks for Civil Applications: A Communications Viewpoint," *IEEE Communications Surveys and Tutorials*, vol. 18, pp. 2624-2661, 2016. [Article \(CrossRef Link\)](#)
- [10] Z. Zheng, A. K. Sangaiah, and T. Wang, "Adaptive communication protocols in flying ad hoc network," *IEEE Communications Magazine*, vol. 56, pp. 136-142, 2018. [Article \(CrossRef Link\)](#)
- [11] I. Jawhar, N. Mohamed, J. Al-Jaroodi, D. P. Agrawal, and S. Zhang, "Communication and networking of UAV-based systems: Classification and associated architectures," *Journal of Network and Computer Applications*, vol. 84, pp. 93-108, 2017. [Article \(CrossRef Link\)](#)
- [12] V. Sharma and R. Kumar, "Cooperative frameworks and network models for flying ad hoc networks: a survey," *Concurrency and Computation: Practice and Experience*, vol. 29, p. e3931, 2016. [Article \(CrossRef Link\)](#)
- [13] S. Zahra, M. A. Ghazanfar, A. Khalid, M. A. Azam, U. Naeem, and A. Prugel-Bennett, "Novel centroid selection approaches for KMeans-clustering based recommender systems," *Information sciences*, vol. 320, pp. 156-189, 2015. [Article \(CrossRef Link\)](#)
- [14] F. Aadil, S. Khan, K. B. Bajwa, M. F. Khan, and A. Ali, "Intelligent Clustering in Vehicular ad hoc Networks," *TIIS*, vol. 10, pp. 3512-3528, 2016.
- [15] M. Gerla and J. T.-C. Tsai, "Multiclustet, mobile, multimedia radio network," *Wireless networks*, vol. 1, pp. 255-265, 1995. [Article \(CrossRef Link\)](#)
- [16] F. Aadil, W. Ahsan, Z. U. Rehman, P. A. Shah, S. Rho, and I. Mehmood, "Clustering algorithm for internet of vehicles (IoV) based on dragonfly optimizer (CAVDO)," *The Journal of Supercomputing*, vol. 74, pp. 4542-4567, 2018. [Article \(CrossRef Link\)](#)
- [17] J. Joy, V. Rabsatt, and M. Gerla, "Internet of Vehicles: Enabling safe, secure, and private vehicular crowdsourcing," *Internet Technology Letters*, vol. 1, p. e16, 2018. [Article \(CrossRef Link\)](#)
- [18] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 24, pp. 881-892, 2002. [Article \(CrossRef Link\)](#)
- [19] M. Dorigo, V. Maniezzo, and A. Colnari, "Ant system: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, pp. 29-41, 1996. [Article \(CrossRef Link\)](#)
- [20] J. Kennedy, R. Eberhart, "Particle swarm optimization," in *Proc. of ICNN'95 - International Conference on Neural Networks*, vol.4, pp. 1942 - 1948, 1995. [Article \(CrossRef Link\)](#)
- [21] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014. [Article \(CrossRef Link\)](#)
- [22] F. Aadil, K. B. Bajwa, S. Khan, N. M. Chaudary, and A. Akram, "CACONET: ant colony optimization (ACO) based clustering algorithm for VANET," *PloS one*, vol. 11, p. e0154080, 2016. [Article \(CrossRef Link\)](#)

- [23] W. Shahzad, F. A. Khan, and A. B. Siddiqui, "Weighted clustering using comprehensive learning particle swarm optimization for mobile ad hoc networks," *International Journal of Future Generation Communication and Networking*, vol. 3, pp. 61-70, 2010.
- [24] H. Ali, W. Shahzad, and F. A. Khan, "Energy-efficient clustering in mobile ad-hoc networks using multi-objective particle swarm optimization," *Applied Soft Computing*, vol. 12, pp. 1913-1928, 2012. [Article \(CrossRef Link\)](#)
- [25] M. Fahad, F. Aadil, S. Khan, P. A. Shah, K. Muhammad, J. Lloret, *et al.*, "Grey wolf optimization based clustering algorithm for vehicular ad-hoc networks," *Computers & Electrical Engineering*, vol. 70, pp. 853-870, 2018. [Article \(CrossRef Link\)](#)
- [26] P. Kaur and A. Singh, "Nature-Inspired Optimization Techniques in VANETs and FANETs: A Survey," *Advanced Computational and Communication Paradigms*, pp. 651-663, 2018. [Article \(CrossRef Link\)](#)
- [27] Z. Kang, L. Wen, W. Chen, and Z. Xu, "Low-rank kernel learning for graph-based clustering," *Knowledge-Based Systems*, vol. 163, pp. 510-517, 2019. [Article \(CrossRef Link\)](#)
- [28] Z. Kang, H. Pan, S. C. Hoi, and Z. Xu, "Robust Graph Learning from Noisy Data," *IEEE transactions on cybernetics*, vol. 50, pp. 1833-1843, 2020. [Article \(CrossRef Link\)](#)
- [29] Z. Kang, C. Peng, and Q. Cheng, "Kernel-driven similarity learning," *Neurocomputing*, vol. 267, pp. 210-219, 2017. [Article \(CrossRef Link\)](#)
- [30] S.-H. Cha, J.-E. Lee, M. Jo, H. Y. Youn, S. Kang, and K.-H. Cho, "Policy-based management for self-managing wireless sensor networks," *IEICE transactions on communications*, vol. 90, pp. 3024-3033, 2007. [Article \(CrossRef Link\)](#)
- [31] S. Chinara and S. K. Rath, "A survey on one-hop clustering algorithms in mobile ad hoc networks," *Journal of Network and Systems Management*, vol. 17, pp. 183-207, 2009. [Article \(CrossRef Link\)](#)



Ali Raza received the MCS degree in Computer Science from Virtual University, Lahore, Pakistan, in 2013. HE completed his MS in Computer Science from COMSATS University Islamabad, Attock Campus, Pakistan. He is currently pursuing his PhD from University of Engineering and Technology, Taxila, Pakistan. His current research interests include Flying ad hoc Networks, bioinformatics, and bio-inspired algorithms.



Muhammad Fahad Khan received his MS degree in Computer Science the Networks from COMSATS Institute of Information Technology, Attock, Pakistan, in 2018. He also did the BS in telecommunication and networking from the same institute in 2016. Currently he is working as Lecturer in the Computer Science Department, COMSATS University Islamabad, Attock Campus, Pakistan and pursuing his PhD from Sunway University Malaysia. His current research interests include data mining, bio-informatics, and bio-inspired algorithms.



Muazzam Maqsood has done his Ph.D. in Software Engineering from University of Engineering and Technology, Taxila. He has completed his MS degree in 2013 from University of Engineering and Technology, Taxila. Currently, he is serving as an assistant professor at Computer Science Department, COMSATS University Islamabad, Attock Campus. His research interests include Machine Learning, Speech Processing, and Recommender System.



Bilal Haider is serving as assistant professor since 2013 in Computer Science Department, COMSATS University Islamabad, Attock Campus. He is a Ph.D. candidate in Computer Science at Bahria University Islamabad Campus. Bilal Haider received his degree of Masters of Science in Computer Science from Griffith College Dublin. He also served as Manager Information System in CIIT Attock Campus till 2017. His research interests include Vehicular ad hoc Networks, Intelligent Transportation Systems, and bio- inspired algorithms.



Farhan Aadil received his B.S. degree in Computer Science from Allama Iqbal Open University, Pakistan in 2005. He pursued a career in the computer science for 4 years (2005 to 2009). He received his M.S. degree in Software Engineering and PhD degree in Computer Engineering from University of Engineering and Technology, Taxila, Pakistan in 2011 and 2016 respectively. He is currently working as assistant professor in COMSATS University Islamabad, Attock Campus, Pakistan. His research interests include Vehicular ad hoc Networks, Intelligent Transportation Systems, and bio-inspired algorithms.