**Dynamic K-means Algorithm for Optimized Routing in Mobile Ad-hoc Networks**

Zahra Zandieh Shirazi and Seid Javad Mirabedini

Department of Computer Engineering Science, Damavand University, Damavand, Iran.

**ABSTRACT**

In this paper, a dynamic K-means algorithm to improve the routing process in Mobile Ad-Hoc networks (MANETs) is presented. Mobile ad-hoc networks are a collocation of mobile wireless nodes that can operate without using focal access points, pre-existing infrastructures, or a centralized management point. In MANETs, the quick motion of nodes modifies the topology of network. This feature of MANETs is lead to various problems in the routing process such as increase of the overhead massages and inefficient routing between nodes of network. A large variety of clustering methods have been developed for establishing an efficient routing process in MANETs. Routing is one of the crucial topics which are having significant impact on MANETs performance. The K-means algorithm is one of the effective clustering methods aimed to reduce routing difficulties related to bandwidth, throughput and power consumption. This paper proposed a new K-means clustering algorithm to find out optimal path from source node to destinations node in MANETs. The main goal of proposed approach which is called the dynamic K-means clustering methods is to solve the limitation of basic K-means method like permanent cluster head and fixed cluster members. The experimental results demonstrate that using dynamic K-means scheme enhance the performance of routing process in Mobile ad-hoc networks.

**KEYWORDS**

MANET, Route optimization, Clustering, Dynamic K-means algorithm.

**1. INTRODUCTION**

With the prompt development of wireless communication technologies, the number of Mobile ad-hoc networks studies has been increasing remarkably in recent years. A MANET includes categories of mobile nodes which communicate with each other without existing of any communication infrastructure or centralized management point. Due to swift movement of mobile nodes and change the network topology, routing is the largest challenge for this kind of networks.

Routing in a network is the act of finding paths to send network traffic from source to destination. In MANET, high mobility of nodes is cause of various routing problems regarding discovery time, delay, packet loss and power consumption. The routing process should be capable to keep optimized routing table after happening any changes in the network topology. The ground breaking attempts to establish optimized routing methods in wireless networks presented in [1-4]. Widmer and her team have been proposed an optimized routing scheme for wireless networks with flat structure [5]. Aiming to make a considerable contribution to the routing process of hierarchical wireless networks, Kartalcetin in [6] proposed to effectively employ an entropy based algorithm to find the best path in this kind of networks.
In the MANET literature, several approaches have been presented to fight routing challenges using optimization techniques. In [7], Gurpreet et al. proposed a MANET optimization method in which Ad-hoc on -demand Distance Vector (AODV) was used to rapid adaptation to dynamic link conditions, low processing and memory overhead, low energy consumption, and quick routes to destinations within the ad hoc network. This scheme uses destination sequence numbers to guaranty loop freedom in MANETs. Tiwari proposed another approach to effectively employ Located Based-Adaptive Routing to reduce overhead massages and delay in Mobile ad-hoc networks [8]. In this scheme, each mobile node has geographical location data of itself, its groups and the destination. When a message should be send, each node forwards the traffic in an eager fashion to the location of the final destination. The requirement this scheme is usually expensive. The reader should be aware that the use of Tree-Seed Algorithm (TSA) to achieve more optimized path in MANETs has been already addressed in [8, 9]. An obvious difference between TSA optimization and aforementioned methods is that multiple path from source to destination is selected and routing will be done proportional with the volume of network traffic.

Cluster-based routing is a solution to moderating the drawbacks of MANETs in terms of both the characteristics of the transmission medium and the routing protocols. The main idea behind the clustering algorithms is structuring of the network into groups of overlapping clusters creating a hierarchical structure. Each cluster includes a node called cluster head opted as cluster head according to a distinct criterion or a combination of metrics such as mobility, power, and density. Clustering methods is a way to recording paths between clusters instead of nodes.

The papers offering a complete clustering-based method for route optimization in MANETs are greatly abundant. The authors are aware of futurist research of [10, 11], where Mobility Based Metric for Clustering has been used to effectively find the optimized paths in MANETs with group mobility behavior, in which an colony of mobile nodes travels with similar speed and direction, as in highway traffic. In this method, by calculating the variance of a mobile node’s speed regarding each of its neighbors, the totality positional speed of mobile nodes is estimated. A low variance quantity demonstrates that this mobile node is relatively less mobile to its neighbors. Consequently, cluster heads are chosen from mobile nodes with low variance quantity in their neighborhoods.

Power-aware clustering is one of the most popular classifications of cluster-based route optimization methods in MANET. In [12], Kojic et al. proposed an impressive kind of power-aware clustering methods which is called load balancing clustering (LBC). LBC provides a node is selected a cluster head it is favorable for it to stay as a cluster head up to maximum determined slot of time, or budget. The budget is a manually defined limitation placed on the algorithm and can be changed to meet the unique features of the system, i.e., the battery life of mobile nodes. In this scheme each mobile node has a floating, virtual ID (VID), and the amount of VID is match as its ID number at first. Initially, mobile nodes with the superior IDs in their local area obtain the cluster head role. LBC restricts the maximum time slots that a node can render as a cluster head steadily, so when a cluster head depletes its time slots, it resets VID to 0 and becomes a non-cluster head node. When two cluster heads travel into the reach range, the one with superior VID achieves the cluster head role. When a cluster head abdicates, a non-cluster head with the bigger VID value in the neighborhood can renovate the cluster head operation [13].

In this paper, a devised dynamic K-means algorithm, which takes advantages of electing cluster head and cluster members in a floating fashion is explained. The metrics that we use to performance assessment of proposed approach can be found in [14].

The remained of this study is organized as follows. In Sect. 2, we proffer a concise summary related to basic K-means clustering approach and will introduce the several important criteria in
Mobile ad-hoc networks. The proposed dynamic K-means clustering algorithm for MANETs is described in Sect. 3. Section 4 shows experimental results obtained on simulation environment to demonstrate the efficiency of the proposed algorithm. Finally, we give our conclusions in Sect. 5.

2. METHODOLOGY

2.1 Basic K-means Clustering Algorithm

Clustering methods are used in many different applications, such as data mining and ad hoc networks. Several clustering schemes have been proposed for ad hoc networks. K-means clustering technique is one of the significant clustering algorithms that can solve many routing problems in MANETs [15]. Due to easy implementation and fast convergence, K-means clustering is an applicable clustering method specifically in mobile ad hoc networks. In contrast, there are some limitations like inadequate distribution of nodes in clusters, fixed cluster head and cluster members.

K-means is the most important flat clustering algorithm. Its objective is to minimize the average squared Euclidean distance from their cluster centres. The first step of K-means is to select as initial cluster centres K. The algorithm follows a simple way to sort out a specific data group through a distinct number of clusters (assume k clusters). The main idea is to determine k centroids, each centroid belongs to one cluster.

These centroids should be located to minimizing the objective function of algorithm. For this reason, a squared error function has been defined as objective function. A square error function is explained as follows [16]:

$$f = \sum_{j=1}^{k} \sum_{i=1}^{n} |x_{i,j}^{(l)} - c_{j}|^2$$ \hspace{1cm} (1)

Where, $|x_{i,j}^{(l)} - c_{j}|^2$ denote distance measure between a node location $x_{i,j}^{(l)}$ and centre of cluster $c_{j}$. In fact, this term calculates total distance of the n nodes location from their respective cluster centres. The example of clustering and electing cluster centres in K-means algorithm is shown in Fig.1.

The K-means method will implemented as follows:

1. Elect K nodes into the space demonstrated by the objects that become clustered. These nodes opted as initial group centroids.
2. Allocate each object to the set that has the adjacent centroid.
3. After all objects have been allocated, estimate the location of the K centroids.
4. Iterate Steps 2 and 3 up to the centroids no longer move. This process a segregation of the objects into clusters from which the criteria to be minimized can be repeated.

2.2 Mobile Ad hoc Networks Criteria

There are many criteria which displays the performance of the mobile ad hoc network. This section takes a quick look to some of these criteria.
1) **Throughput**: It is favourable to devise a wireless routing protocol with the maximum probability of data delivery, minimum probability of packet drop. Therefore, in wireless networks, the effort has always been concentrate to an end-to-end high capacity path.

2) **Route Discovery Time**: the aim of all route discovery mechanism is to construct, maintain and repair paths between nodes in a network. In a MANET, overhead packets can be propagated to regions where they have no sense but nonetheless consume inadequate network resources. Therefore, reducing the route discovery time means to protect ineffective use of network resources in MANETs.

3) **Energy consumption**: Energy consumption is a vital challenge in all kind of wireless networks where the battery lifetime limitations the autonomy of network nodes. Electing a non-optimized path with unreliable link probably would be cause of produce more delay in order to higher retransmission rates, that finally results is increase of energy consumption. For energy saving, most of the researches concentrate on the communication protocol design. For example, the routing protocol Parkins [17] uses a manipulated Ad Hoc on-demand Distance Vector (AODV) routing to be applied by low-power devices.

4) **Delay**: it is desirable to select the network path with minimum delay. The delay of a network path determines how long it takes for a packet to move across the network source from one node or endpoint to destination node. It is typically calculated in multiples of seconds. Delay may vary smoothly, depending on the position of the specific pair of corresponding node.

5) **Packet Drop**: Packet loss occurs when one or more packets of data travelling across a computer network fail to reach their destination. In MANETs, wireless link transmission errors, mobility and congestion are major cause of packet drop.

![Fig. 1 Example of basic K-means clustering method](image)

### 3. Proposed Optimization Scheme

The proposed K- means clustering scheme is compatible with IP-based Ad hoc networks. In this scheme, the disadvantageous of basic K-means algorithm have been eliminated and the strength point of previous methods is preserved. Accomplishment of Dynamic K-means clustering algorithm is established according to 3 steps. Each step of this proposed scheme is presented as follows.
3.1 Implementation Step

At the beginning of the dynamic K-means clustering scheme, information of node location and energy consumption are used to promote the clustering process of basic K-means method and create an adequate distribution for node into the clusters. For this purpose, the K numbers of cluster centers are designated in such a way that the limitation of basic algorithm like inappropriate distribution of nodes has been covered. The formulation of intended cluster center $C$ in the proposed scheme with $n$ nodes is presented as follows:

$$C = \frac{\sum_{i=1}^{n} x_i}{n}$$  \hspace{1cm} (2)

Where $X_i$ denotes location of nodes in the network. After specify centre nodes $C$, the average distance between $C$ and all the nodes of network $R$ is calculated as presented below:

$$R = \frac{\sum_{i=1}^{n} |X_i - C|}{n}$$  \hspace{1cm} (3)

Then, the location of means $m_i$ ($m_{i_x}$, $m_{i_y}$) in each cluster according to $R$ and $C$ values is computable.

$$m_{i_x} = R \times \cos\left(\frac{360}{k} \times (i-1) \times \frac{\pi}{180}\right) + C_x$$ \hspace{1cm} (4)

$$m_{i_y} = R \times \sin\left(\frac{360}{k} \times (i-1) \times \frac{\pi}{180}\right) + C_y$$ \hspace{1cm} (5)

In this scheme, the calculations of means are sorely important because the appropriate locations of selected means are cause of a fair nodes distribution in each cluster.

After determining the position of K-means, $n$ nodes of network should properly distribute into $K$ clusters. For this purpose, the presented equivocation is used:

$$\text{avg}_{\text{min}} \sum_{i=1}^{K} \sum_{x_{i,s_j}} |x_{i,j} - m_i|^2$$ \hspace{1cm} (6)

Where, $S = [S_1, S_2, \ldots, S_k]$ denotes Clusters and $\text{avg}_{\text{min}}$ is function of minimum average distance. In this method, each node is joined to closest cluster.

In order to creation clusters with a fair distribution, minimum average distance of nodes to each cluster has been estimated and each node exactly attach to one nearest cluster. Therefore, the final created clusters can be described as follows:

$$S_{i}^{(3)} - \left\{ x_{i,j} \left| x_{i,j} - m^{(3)}_{i} \right|^2 \leq \left| x_{i,j} - m^{(2)}_{i} \right|^2 \right\}$$ \hspace{1cm} (7)

When all the nodes would be attached to clusters, the formation of all the clusters is finalized. An example of dynamic K-means clustering method is illustrated in Fig. 2.
3.2 Head Clusters Mobility Step

At the end of clusters recruitment, the step of head cluster election will launch. Due to mobility as a natural trait of mobile ad hoc network, employing a node as constant cluster head is impairing because in the clustering scheme, one cluster head can be exhausted when it serves too many mobile hosts.

In the proposed approach, initial election of cluster head is performing based on distance to cluster center and total energy of each node. All nodes obviously have an equal energy after the first distribution of nodes in clusters. Thus, the only metric for initiatory opting of cluster heads would be the minimum distance to cluster center. In the proposed scheme, only one cluster head is allowed per cluster [18].

Since the nodes with highest metrics have been chosen as cluster heads, each cluster should be send a hello massage to neighbors and presented itself as head cluster. Cluster members announce their interest to join the groups of cluster head after receiving hello massage. The neighbors of a cluster head become members of that cluster and can no longer take part in the election process. In this way, the stage of initial head cluster selecting is completed.

In the dynamic K-means clustering algorithm, in order to keep mobility in cluster head election and protecting network from drawbacks of constant head cluster, each node can serve as a cluster head for a time slot T. The cluster head election process should be holding after passing each time slot T.

At the termination of each time slot T, each cluster head broadcast a cluster head candidate massage along the span area of its members that are within its cluster range.

Let P be the probability of electing node n as cluster head and Z is a group of nodes which do not select as cluster head in the r previous time slots T. New metric of each node presented as equation (8):

\[
T(n) = \begin{cases} 
\frac{P}{1-P \cdot (r \mod p)^r} \cdot \frac{E_{\text{curr}}(n)}{E_{\text{init}}(n)} & \text{if } n \in G \\
, & \text{otherwise}
\end{cases}
\]
Where, $E_{cur}$ denotes remaining energy of node and $E_{init}$ marks initial energy of each node.

Each members of cluster should be calculates its metric in accordance to below equation and send back to head clusters. Eventually, the node with maximum amount of metric is chosen as a cluster head.

### 3.3 Stability Step

When the process of head cluster election in the each time slot $T$ is complemented, cluster state is going to stable stage. Due to optimized routing in this state, each cluster head appoint a proportionate TDMA (Time Division Multiple Access) schedule for sending traffic from cluster nodes. Each cluster member receives this schedule and should send traffic only in the specific time which belonged to itself in TDMA. Cluster members after sending traffic in their own time should go to sleep in order to energy saving [19]. At the beginning of next slot time $T$, elected head cluster should refresh TDMA schedule and send this upgraded TDMA to all cluster nodes. Each node should send its own traffic according to new version of TDMA. All traffic of cluster nodes route by means of cluster head using proactive routing method.

### 4. EXPERIMENTAL RESULTS

Simulations have been directed to evaluate the performance of the dynamic K-means clustering algorithm. In this simulation, we applied proposed clustering scheme on experimental MANET environment with size 2500*2500 meters. Also, the two ray fashion has been chosen as propagation model of network traffic. Another specification of experimental environment described in the Table.1. Moreover, a schematic plane of simulated environment is shown in Fig. 3.

The proposed dynamic clustering algorithm for MANETs takes into account of several assessment criteria, route discovery time, delay, total route error sent, receiving packet rate, throughput and total packets dropped.
The absolute value of each of these criteria can take a large range at different points on the network. In this paper the average values for each criterion has been considered. Our simulations have been concentrated on comparing aforementioned criteria measurements between basic K-means clustering method and proposed clustering scheme in MANET.

Route discovery time criterion in MANTES is period of time that is necessary for finding the best path from source to destination [20]. In order to solve the problems of basic K-means method regarding to fair distribution of nodes in clusters and updating the route information, applying dynamic proposed algorithm leads to significant decrease in the route discovery time which observable in Fig. 4.

![Fig. 4 Rout Discovery Time for Simulated Algorithms](image)

<table>
<thead>
<tr>
<th>Simulation Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
</tr>
<tr>
<td>Size of Network</td>
</tr>
<tr>
<td>Packets destination</td>
</tr>
<tr>
<td>Size of Packets</td>
</tr>
<tr>
<td>Physical Layer Protocol</td>
</tr>
<tr>
<td>Number of Nodes</td>
</tr>
<tr>
<td>Transmission Layer Protocol</td>
</tr>
<tr>
<td>Nodes Speed</td>
</tr>
<tr>
<td>Power of Nodes</td>
</tr>
</tbody>
</table>

Table. 1 Simulation Specifications
Furthermore, to study the vital parameters of mobile ad hoc networks, the measure of delay in the basic K-means algorithm and proposed dynamic K-means method is estimated. From Fig.5 is clear that the proposed clustering algorithm is same behavior with the basic K-means algorithm in terms of delay criterion. It confirms that the proposed Dynamic clustering scheme keeps efficient characteristics of basic K-means method coincide with enhance in the other impressive properties of this popular clustering technique.

In Fig.6 and Fig.7, we depict the plots of packet sent rate and packet received rate in both experimented algorithms. The packet sent rate implicates the volume of traffic which sends via source nodes to specific destinations. Likewise, the packet received rate is volume of received traffic of corresponding destination nodes. Due to absence of any interference in sending process, the packet send rate of both compared methods is identical. On the other hand, in Fig. 7 it is empirically observable that the created expansions in the dynamic K-means clustering are cause of remarkable performance improvement in comparison with basic K-means clustering in terms of packet received rate.

![Fig. 5 Delay for Simulated Algorithms](image-url)
Figure 8 demonstrate an analogy between the number of dropped packets in the dynamic K-means scheme and basic K-means algorithm. Total packets dropped is an indicator to assess the performance of clustering schemes to combat the vital challenges of MANETS like channel noise, interference and lack of appropriate management of clusters. From Fig 7, it is clear that the total packet dropped in both methods is very close to each other at the beginning of simulations. However, the number of dropped packets decreases abruptly in the proposed scheme by passing of time. It confirms the applicable performance of the proposed clustering scheme for MANET in terms of network reliability.
One of the most important criterions for evaluates the performance of a routing method in the wireless networks is throughput. From a telecommunication point of view, throughput or network throughput is the rate of successful packet delivery over a communication channel [21]. Throughput is usually estimated in bits per second (bit/s or bps), and usually in data packets per second (p/s) or data packets per time slot.
By numerical experiments, we illustrate that our proposed dynamic K-means clustering scheme can achieve dramatic enhancement in the measure of throughput in comparison to basic K-means algorithm when passing of times leads to stability of network. This progression in the proposed method is shown in Fig. 9.

Eventually, Fig.10 indicates the estimated total route error sent criterion for the basic K-means algorithm and proposed clustering scheme in MANETs. The route error sent in the wireless ad hoc networks as result of nodes mobility and inefficient distribution of nodes into clusters leads to wrong path finding. This empiric outcome verify again the results achieved in the previous experiments: the proposed dynamic K-means clustering scheme offers considerably lower than the total route error sent for the other scheme compared.

5. CONCLUSIONS

In this survey, a new dynamic K-means clustering scheme for optimized routing in mobile ad hoc networks is proposed.

The presented algorithm is includes three steps for optimizing the routing process of MANETs. Due to high mobility of nodes and electing fixed cluster head, there are many routing problems in the basic K-means method. The main objective of this proposed algorithm is to covering the drawbacks of basic K-means algorithm by means of applying an effectual method for fair distributions of network nodes to cluster and electing the most appropriate cluster head. The significant remark in the presented algorithm is making efforts to preventing the challenges of fixed cluster head via substituting the cluster head after time slot T.

The empirical experiments illustrate that our dynamic K-means clustering algorithm for MANETs can obtain dramatic improvements in terms of vital parameters of a wireless network such as route discovery time, throughput and total dropped packets in comparison with basic K-means algorithm. From a practical point of view, we want to note that this dynamic K-means algorithm can be utilized in other usages of clustering methods.
Currently, we aim to combine the proposed dynamic K-means clustering scheme with other efficient clustering method to achieve better routing performance in the wireless ad hoc networks. Future work will investigate more efficient solutions under the clustering framework to optimize the routing process.

REFERENCES


