Enhancenig OLSR routing protocol using K-means clustering in MANETs

Y. Hamzaoui[1], M. Amnai[2], A. Choukri[3], Y. Fakhri[4]
[1, 2, 4] LARIT, Networks and Telecommunications Team, Faculty of Sciences Kenitra, IbnTofail University, Morocco
[3] Caddy Ayyad University, S.A.R.S Group, ENSA Safi, Morocco

Article Info	ABSTRACT
	The design of robust routing protocol schemes for MANETs is quite complex, due to the characteristics and structural constraints of this network. A numerous variety of protocol schemes have been proposed in literature. Most of them are based on traditional method of routing, which doesn't guarantee basic levels of Qos, when the network becomes larger, denser and dynamic. To solve this problem we use one of the most popular methods named clustering. In this work we try to improve the Qos in MANETs. We propose an algorithm of clustering based in the new mobility metric and K-Means method to distribute the nodes into several clusters; it is implemented to standard OLSR protocol giving birth a new protocol named OLSR Kmeans-SDE. The simulations showed that the results obtained by OLSR Kmeans-SDE exceed those obtained by standard OLSR Kmeans and OLSR Kmed+ in terms of, traffic Control, delay and packet delivery ratio.

Corresponding Author:

Y. Hamzaoui,
Faculty of Sciences Kenitra,
IbnTofail University,
Avenue de L'Université, Kénitra, Morocco.
Email: Younes_ieee@hotmail.com

1. INTRODUCTION

The ad hoc network is a new wireless communication system without the use of a centralized management infrastructure. It consists of a set of nodes (mobile device) that have a communication wireless device for communicating with entities located in their neighborhood. Each node can therefore directly reach its neighbors direct using its radio interface, or communicate with other nodes inside the network using intermediate nodes (located between the source and the recipient). These are responsible for commuting the messages and playing the role of the router, so it offers an autonomous network and gives the terminal user access to information at any time and from any place. These mobiles cooperate with each other to overcome the Ad hoc network constraints such as dynamic topology, lack of centralized monitoring points, limited bandwidth, etc. the latter requires designing a robust and non-traditional routing system to better management of the flow of information and ensure the quality of service in a dynamic and decentralized network.

Many routing protocols have been proposed [1] on different types of application. The research has not ceased to have efficient protocols that adapt to all mobility models [2]. The routing of the information in the Manets can be classified into two types: Flat and Hierarchical. In the first type all network nodes play the same role, this can overload the network, as well as other problems cause such as scalability and complexity, when the network becomes wider and denser. The second type of routing is used to support networks that are wide and dense. The clustering is a hierarchical structure that makes it possible to group geographically the nodes that are neighbors. This allows to each node to store all information about it group and only some information of other groups (clusters). This approach can reduce the cost of routing of

information in large and dense networks. Several researches have been proposed to improve the efficiency of the hierarchical protocols and to support this type of network and structure.

In order to better organize the network into several clusters with optimal structures, the researchers proposed various criteria for building and organizing the nodes in groups. This hierarchic structure leads to optimization and improvement of the routing protocols. Depending on the type of performance that can be ensured in the cluster structure, we use the suitable metric. Several metrics have been proposed to measure the physical and logical properties of the nodes (mobility, density, energy, etc.). A good metric makes it possible to differentiate easily between the nodes [2], to reflect the real behavior of the nodes. and facilitates the performance improvement of protocols [3].

In this article we refine our metric of mobility, defined in [4], taking into account any type of motion in a coverage area of a mobile node, this refined metric and other metrics (density, energy) will be the basis of a cluster building algorithm by using k-means method. To create the groups of mobile and to elect the cluster head, this new approach can generate a more stable cluster and cluster head. In this paper we start with a presentation of some related work. In the third section we present the problem formulation, and in the fourth section we define the clustering. Then in the fifth section we present the description of approach clustering algorithm, in the sixth section we explain our mobility metric, after in the seventh section we present the algorithm before we give the results in eighth section. In the ninth section we finish by a conclusion.

2. RELATED WORK

In A Several studies provide different clustering techniques to improve the network scalability and simplifying the complexity of routing problem to smaller groups of nodes. They usually differ on the criteria of selection of cluster head. In this part we propose some work done and that focuses on different criteria for the generation of the cluster head, some of this research are implemented in the OLSR [5] (Optimized Link State Routing Protocol) to improve it Qos.

Mobility Based Metric for Clustering (MOBIC): is the original proposal of Basu et al [6] suggesting that clustering, especially cluster head election, must consider mobility as a relevant criterion in order to ensure a certain stability of generated clusters. This algorithm is based on a local mobility metric called "relative mobility of nodes", it is revealed that the node with the low value is the least mobile, ie the most stable. Therefore, it is the node that is elected as a cluster head.

New approach named SALSA presented in [7], it is a distributed and self-organising clustering scheme assigning equal cluster management tasks to all nodes. In addition, a cluster balancing mechanism is introduced allowing nodes to be evenly distributed among clusters. Before the maximum capacity of a cluster is reached, it progressively starts assigning nodes to neighbor clusters. This contribution also proposes a cluster quality metric in order to assign nodes to the most suitable clusters, according to connectivity and free positions within clusters. Results confirmed the performance efficiency of the new scheme, providing stability and low maintenance overhead, even in the largest networks.

An optimized stable clustering algorithm for Ad Hoc (OSCA) proposed by [8], that will provide more stability to the network by minimizing the cluster head changes and reducing clustering overhead. In this algorithm, a new node is introduced which acts as a backup node in the cluster. Such backup node acts as cluster head, when actual cluster head moves out (or died) from the cluster. This practice keeps network availability without disturbance. Further, the priority of cluster head and backup node is calculated based on the node degree and the remaining battery life for mobile nodes. According to the experimental results that proposed an optimized stable clustering algorithm for mobile ad hoc networks (OSCA) algorithm, it will not only be able to make a network more stable by reducing number of cluster head changes but also reduce the clustering over- head.

The authors in [9] present a new approach to build clusters for Wireless Sensor Networks (WSN). The algorithm is based on the k-means method which is well known as a clustering technique. K-means clustering tends to find clusters of comparable spatial extent (density clustering). They try to enhance the clustering process by selecting nodes as clusters that are centric and have a high level of energy. This will give the same QoS results as given by the K-means approach with a reduction of energy consumption and a prolongation of the lifetime of the sensor network. For the simulation purposes, the authors have implemented our approach on the OLSR routing protocol. The approach proposed seems to give better results than the MaxMin approach.

In [10] an improvement of protocol OLSRMaxMin2c was proposed by the introduction of the cost of energy. The main objective of this new contribution, called OLSRMaxMin2C / Energy, is the optimization of energy consumption OLSRMaxMin2C/Energy. It consists of dividing the network into clusters. Cluster heads are elected according to their IDs and their residual energies (battery levels); this algorithm determines

the best path, in terms of energy cost. This involves calculating the energy required for each available path. It is an algorithm that optimizes the number of live nodes by always choosing the appropriate nodes for each task in the network.

3.   PROBLEM FORMULATION

The development of technology and the revolution of wireless technology have led to the existence of mobile devices in all areas of human activity. The connection of these mobiles in ad hoc mode becomes a necessity, by their simplicity of deployment and in the absence of a preexisting and expensive infrastructure. When the number of mobile increases and the topology changes rapidly by the high mobility of nodes; such a proactive routing protocol cannot support the network evolution, due to the generation of more message control and routing table, etc. Consequently the network becomes more sensitive, and it doesn't ensure a minimum quality of service. Developing or refining methods that will overcome these obstacles becomes a necessity. Many studies on routing optimization adopt the clustering method to reduce the costs product by density and mobility. It is based on the distribution of mobiles nodes into groups. This approach is proposed to reduce the storage and processing information within a cluster. In the literature we can have several clustering techniques; the K-means clustering technique is among the best methods known [11, 12] used in MANETs. This one will allow clusters to be more stable and the center will be better chosen. Choosing a less mobile and dense cluster head without taking into account their residual energy level; can result in the isolation of all cluster members from the rest of the network, in the case of the exhaustion of the battery [13]. The implication of the concept of energy in the process of clusters construction avoids the election of a node, well placed but with weak energy level.

In this paper we propose the use of K-means to produce optimal clusters. This method is based on the use of Euclidean distance to generate cluster centers. In our proposition we will introduce in the K means algorithm the three stabilization parameters (mobility, density, energy) as a distance vector. The cluster head will be elected according to these three parameters. The center of the cluster will be the one that respects both the three stability parameters in it group, it will be less mobile, denser and has an energy level that increases the life of the cluster. This approach allows to build the clusters that are highly resistant to the constraint structure of the ad hoc network. It is implemented to standard OLSR protocol giving birth a new protocol named OLSR-Kmeans-SDE.

4.   CLUSTERING

In Spatial clustering algorithms can be classified into four categories. They are the partition based, the hierarchical based, the density based and the grid based [14, 15]. According to [16] clustering in ad hoc networks can be defined as a theoretical arrangement of dynamic nodes corresponding to one or more specific properties in different subsets called "Cluster". An element of a cluster is characterized by a strong similarity of the components of its group, and a strong dissimilarity with respect to the members of other groups [17]. Each cluster is identified by a particular node called "Cluster head". Clustering allows a node to store only part of rather than all the information of the network topology. This simplifies the processing of the global topology [18]. This reduces the size of routing tables and thereafter the reduction of control messages generated by the routing system [19].

The use of clustering in MANETs has several advantages [20, 21], usually a cluster structure allows the node to play one of three roles:

– Cluster head: A cluster head is elected in the cluster formation process for each cluster. Each cluster should have one and only one cluster head.
– Gateway: A node is called a gateway node of a cluster if it knows that it has a bidirectional or uni-directional link to a node from another cluster.
– Members: All nodes within a cluster except the cluster head are called members of this cluster.

5.   DESCRIPTION OF APPROCH CLUSTERING ALGORITHM

In the absence of any assumption about the distribution of nodes in a mobile ad hoc network, an unsupervised classification of nodes into classes is required. We propose a model based on geometric considerations (grouping geographically close nodes). This proposal requests define a measure of the proximity between nodes. We begin by introducing the concepts of similarity, dissimilarity and distance between nodes, and then we will use the inertia concepts intra-class and inter-class to demonstrate mathematically how a classification method built homogeneous and distinct groups by translating it into a simple minimization problem.

C is a partition of X if and only if C satisfies the following properties:

- $C_a \subset X$ for all $C_a \in X$
- $\bigcup_{a=1}^{t} C_a = C$
- $C_a \cap C_b = \emptyset$ for (a, b) such that $a \neq b$

The property (3) expresses that the clusters formed are disjoint; each object of X belongs only to a single cluster of C. To define the homogeneity of observations set, it is necessary to measure a similarity between two observations. Then we introduce the concepts of similarity and dissimilarity. Dissimilarity is a function d which associates a value in IR + to each pair $(x_i x_j)$ such that:

$$d(x_i, x_j) = d(x_j, x_i)$$

$$d(x_i, x_j) = 0 \rightarrow x_i = x_j$$

Conversely, another possibility is to measure the resemblance between observations using a similarity. Similarity is a function s to which associates a value in IR* to each pair $(x_i x_j)$ such that:

$$s(x_i, x_j) = s(x_j, x_i) \geq 0$$

$$s(x_i, x_j) \geq s(x_j, x_i)$$

Thechoice of the distance is a key issue for classification methods. To offer a relevant measure of similarity between elements, it is necessary to well use the available information at the nodes.

The Minkowski distance is the most used to determine the similarity between elements:

$$d(x_i, x_j) = \left(\sum_{k=1}^{p} \left| v_k(x_i) - v_k(x_j) \right|^l \right)^{1/l}$$

Where vk (xi) is the value of the object xi on the variable vk.
Depending on the values taken by the parameter l, we talk about:

- Euclidean distance (l = 2);
- Manhattan distance (l = 1);
- Chebychev distance (l = ∞).

We note that the metrics commonly used to analyze the ad hoc performances such as density, mobility and energy can be used to express distance. Our objective is to divide the nodes into homogeneous and distinct clusters. To do this, we start from the definition of network inertia that can be modeled by a cloud of points.

Consider a network of n nodes ($x_1, \dots x_n$) and UG designates the centroid of nodes cloud.

$$U_G = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1}$$

The cluster inertia is by definition the sum of squares of distances from the center:

$$I = \sum_{i=1}^{n} d^2(x_i, u_G) \tag{2}$$

$$I = \sum_{i=1}^{n} \|x_i - u_G\|^2 \tag{3}$$

Assume that the network consists of P distinct clusters $C_1, \dots, C_P$. Each of these clusters having as centroid U_{ck}. We can then decompose the total inertia of the cloud of nodes as follows:

$$I = \sum_{k=1}^{p} \sum_{i=1}^{n} \|x_i - u_G\|^2$$

$$I = \sum_{k=1}^{p} \sum_{i=1}^{n} \|x_i - u_{c_k} + u_{c_k} - u_G\|^2$$

According to the theorem of Huygens,

$$I = \sum_{k=1}^{p} \sum_{i \in c_k} \left(\|x_i - u_{c_k}\|^2 + \|u_{c_k} - u_G\|^2 \right)$$
$$I = \left(\sum_{k=1}^{p} \sum_{i \in C_k} d^2(x_i, u_{C_K}) \right) + \left(\sum_{k=1}^{p} d^2(x_{C_k}, u_G) \right) \tag{4}$$

The first term called intra-cluster inertia calculates the sum of distances between nodes and their centroid. Low intra-cluster inertia indicates that the nodes in the same cluster are more nearer (Clusters are homogeneous).The second term called the inter-cluster inertia calculates the sum of distances between centroids of clusters and global centroid, that is to say the separation degree of clusters.

$$I = I_{intra} + I_{inter} \tag{5}$$

From a formal point of view, the optimal partition is:
- That minimizes the intra class inertia
- Or that maximizes the inter class inertia,

Thus the optimal partition would be defined as follows:

$$C_k^{optimal} = Argimin_{C \in C_k} \sum_{k=1}^{p} \sum_{i \in C_k} d^2(x_i, u_G) \quad (6) \tag{6}$$

The objective of classification algorithms based on this principle is the search of optimal partition.

In practice it is impossible to generate all patterns of clustering for evident reasons of complexity. We then seek a scheme sufficiently close to the optimum. This optimum is obtained in an iterative manner by improving an initial scheme randomly selected by reallocating objects around mobile centers. In order to partition the nodes into clusters, we used this technique (iterative reallocation) based on k-means algorithm.

5.1. K-means method

The k-means algorithm [15] is a partitioning method widely used in various application areas. From P separate clusters, the k-means algorithm assigns iteratively objects $(x_1, \dots x_n)$ at P centers of clusters $(u_1, \dots u_P)$, followed by calculating the positions of the new centers. The stopping algorithm is a criterion fixed by the user and can be:
- Achieve a limited number of iterations;
- The algorithm converges: clusters formed remain the same between two successive iterations;
- The inertia intra-cluster is not improving between two iterations (the algorithm is sufficiently close to convergence). To justify the K-means algorithm in view of our goal aimed to minimize the intra-cluster inertia, we demonstrate that the redefinition of centers of clusters and the reallocation of nodes (Sequence of two sub-stages of K-means algorithm) results in a decrease in intra-class inertia.We start from following considerations at the end of determined step j;
- Centers of clusters $\left(u_1^j, \dots u_p^j\right)$ have been calculated;
- The classes $\left(C_1^j, \dots C_p^j\right)$ were obtained by assigning at the center $\left(C_k^j\right)$ the n_j^i nearest nodes.We define the following quality at the end of step j;

$$Q_j = \sum_{k=1} \sum_{i \in C_k^j} \left\| x_i - u_k^j \right\|^2$$

The redefinition sub-step of new centers and reassigning of nodes (next iteration) requires:

a. Recalculate centers of clusters $\left(u_1^{j+1}, \dots u_p^{j+1}\right)$ based on points belong to each of clusters $\left(C_1^{j+1}, \dots C_p^{j+1}\right)$ possessing respectively n_k^j element. We have:

$$u_k^{j+1} = \frac{1}{n_k^j} \sum_{i \in C_k^j} x_i$$

and:

$$W_{j+1} = \sum_{k=1}^{p} \sum_{i \in C_k^j} \left\| x_i - u_k^{j+1} \right\|^2$$

where u_k^{j+1} the center of the cluster is C_k^j et W_{j+1} is the inetria intra-clusters clusters assosiated to clusters $\left(C_1^j, \dots C_p^j\right)$. We will have:

$$Q_i = \sum_{k=1}^{p} \sum_{i \in C_k^j} \left\| x_i - u_k^j \right\|^2$$
$$Q_i = \sum_{k=1}^{p} \sum_{i \in C_k^j} \left\| x_i - u_k^{j+1} + u_k^{j+1} - u_k^j \right\|^2$$

$$I = \sum_{k=1}^{p} \sum_{i \in C_k^j} \left(\left\| x_i - u_k^{j+1} \right\|^2 + \left\| u_k^{j+1} - u_k^j \right\|^2 + 2\langle x_i - u_k^{j+1}, u_k^{j+1} - u_k^j \rangle \right)$$

$$I = W_{j+1} + \sum_{k=1}^{p} \sum_{i \in C_k^j} \left\| u_k^{j+1} - u_k^j \right\|^2 + 2\sum_{k=1}^{p} \langle \sum_{i \in C_k^j} (x_i - u_k^{j+1}), u_k^{j+1} - u_k^i \rangle \geq W_{j+1}$$

The vector $\sum_{i \in C_k}(x_i - u_k^{j+1})$ being the null vector by definition of u_k^{j+1}

b.   Reassign nodes to the nearest centers. Then we obtain new clusters $\left(C_1^{j+1}, \dots C_p^{j+1} \right)$ and we define:

$$Q_{j+1} = \sum_{k=1}^{p} \sum_{i \in C_k^j} \left\| x_i - u_k^{j+1} \right\|^2$$

After reassignment all distances decrease because each node xi is assigned to the cluster center minimizing thereby the gap: $\left\| x_i - u_k^{j+1} \right\|^2$ We therefore have:

$$Q_{j+1} \leq W_{j+1}$$

Thus, for each j we have proved the following inequality:

$$Q_{j+1} \leq W_{j+1} \leq Q_j$$

So we have in particular:

$$W_{j+2} \leq Q_{j+1} \leq W_{j+1}$$

This shows that after each algorithm iteration the improvement of the nodes classification is sure within the meaning of I_{intra}criterion. Because the intra-cluster inertia of the optimal partition is the smaller, the margin of improvement is finite. This implies that the algorithm converges inevitably. The disadvantage of the k-means method is that the number of clusters is a parameter of the algorithm. This is not obvious for a mobile ad hoc network where nodes can join or leave the network randomly. Hence the need to make improvements to the method to be applicable to our case:

– Initial number of Clusters: At the beginning each node represents its own cluster. Thereafter we make a series of ascending partitions by combining the nodes belonging to the same neighborhood into same cluster until reaching a stable number of clusters.
– Distance: We can use the density, mobility or energy of nodes to express the distance metric.

6.   METRIC MOBILITY

The metric [22] proposed allows to calculate the stability of a node based on four parameters, the node that leaves the coverage area of node study, the nodes that joins the area zone, the node approaches to the studied node and the node that finally moves away from the examined node and that stays in their coverage area. The first two parameters will be retained by the collected control messages. The last two will be calculated by the calculating power for two successive received messages (eg Hello message in OLSR protocol). We define the following parameters that characterize our metric:

– N_{con} defines the number of nodes that converge on the studied node.
– N_{div} defines the number of nodes which diverge towards the outside of the node studied.
– N_{in} defines the number of nodes within the area of the studied node.
– N_{out} defines the number of nodes out of the coverage area of the studied node

Following these four types of movement we have created a metric that will rank the nodes between any of these four metrics of stability. We define each as as shown in Table 1.

<div align="center">Table 1. Classifying of metrics</div>

Cases	Mouvements parameters	Interval of γ
Cas 1 : Better stability	$\begin{cases} Ncon > Ndiv \\ Nin > Nout \end{cases}$	$\gamma \in [0\ \ 0.25[$
Cas 2: Stable	$\begin{cases} Ncon < Ndiv \\ Nin > Nout \end{cases}$	$\gamma \in [0.25\ 0.5]$
Cas 3 : less stable	$\begin{cases} Ncon > Ndiv \\ Nin < Nout \end{cases}$	$\gamma \in\ ]0.5\ 0.75\ [$
Cas 4 : poor stability	$\begin{cases} Ncon < Ndiv \\ Nin < Nout \end{cases}$	$\gamma \in [0.75\ 1]$

In the four classifications, we have the first case that reflects a better stability for the node in question and the latter which represents a poor stability of the studied node, intuitively the second is better than the third, because N_{in} is greater than N_{out} and secondly, even with $N_{div} > N_{con}$ the diverging node stay in the coverage area of the studied node. We determine subsequently metric degree of stability that will calculate for each category the best stability node, we use in the formula (7) the coefficient γ of flow defined in [23], we divide the coefficient of 4 intervals as shown in Table 1, and metric of stability degree of node i will be as follows:

$$M_i^\gamma(t) = \gamma\left(\frac{N_{div}}{N_{con}+N_{div}+N_{out}} + \frac{N_{out}}{N_{con}+N_{div}+N_{out}}\right) + (1-\gamma)\left(\frac{N_{con}}{N_{con}+N_{div}+N_{in}} + \frac{N_{in}}{N_{con}+N_{div}+N_{in}}\right) \quad (7)$$

The stability metric is defined in our proposition formula (7)[23]. The disadvantage of this approach is that the parameter γ must be fixed between three values (0.25, 0.5, 0.75), according to the flow of motion around node (dominant inflow $\gamma = 0.75$, dominant outflow $\gamma = 0.25$). To better choose the parameter γ, we have automated the determination of this parameter by refined the proposed equation. So the formula becomes (8):

$$M_i^\gamma(t) = \alpha\left(\frac{N_{div}}{N_{con}+N_{div}+N_{out}} + \frac{N_{out}}{N_{con}+N_{div}+N_{out}}\right) + \beta\left(\frac{N_{con}}{N_{con}+N_{div}+N_{in}} + \frac{N_{in}}{N_{con}+N_{div}+N_{in}}\right) \quad (8)$$

$$\alpha = \left(\frac{N_{con}+N_{in}}{N_{con}+N_{div}+N_{out}+N_{in}}\right) \text{ and } \beta = \left(\frac{N_{div}+N_{out}}{N_{con}+N_{div}+N_{out}+N_{in}}\right) \quad (9)$$

7.   DESCRIPTION OF ALGORITHM

We present below our algorithm in detail (k-means improved to make it applicable to the partition of nodes into clusters): The introduction of changes to the standard algorithm is a need, to use the k-means method for grouping nodes of a Manets into clusters; this change aims to determine the K parameter of the algorithm. In first it is assumed that each node in Manets is an own cluster; then followed by a sequence ascending partitions, in the end we reunite the nodes from the same neighborhood in the same cluster, until reaching a final number of clusters. Thereafter, the K-Means algorithm will be used to generate more stable clusters with their cluster head. The parameters used in the K-means calculation algorithm are the stability, density and energy residual of a mobile, so each node is identified by a vector (stability, density, energy). This vector will be the basis for expressing the distance between nodes.

We present our algorithm based on K-means to create the partition of the cluster nodes in Manets. Proposed algorithm
− Input: Mobile ad hoc network of n nodes.
− Output: Network virtually partitioned into P clusters
   Step 0
− Initialization with n centers $(u_1^0 \ldots u_n^0)$ each node is a cluster
− Creation of an initial partition P0 = $\{(C_1^0 \ldots C_k^0)\}$
− Initialize l to 1 ($list iteration index$). Assign to u_l^0 it is two-hop neigbours;
− $C_l^0 = \{x_i \in$ network | d $(x_i, u_l^0) >= 2$hops$\}$;
− Remove from list of centers the $n_{C_l^0}$ nodes assigned to centroid u_l^0;
− Move to the center $l + n_{C_l^0}$;
− Repeating steps b to do until all node are affected;
− Calculation of new centers of k cluster obtained $(u_1^1 \ldots u_k^1)$: nodes having the optimal values of (stability, density, energy);
   Step t
− Creation of new partition Pt = $\{(C_1^t \ldots C_k^t)\}$ by assigning to each centers its two-hop neighbors;
− The centers affected to other centers are removed from the list of centers;
− The isoled nodes are assigned to the list of centers;
− Calculate the centers of k clusters obtained $(u_1^0 \ldots u_k^0)$;
− Repeat steps 3 to 6 until that a stable partition is achieved (structure of partition P_{t+i} equals that the P_{t+i+1}) or reach n iterations.

We used the NS2.34 simulator to analyze the quality of service between OLSR Keans-SDE and the two other protocols (The classical Kmeans implemented in OLSR and OLSR Med+ [24]). We validate this analysis by studying the behavior of the protocol in terms of PDR, EED and Overhead [25]. The following Table 2 shows the simulation parameters.

Table 2. Simulation parameters

Parameter	Value
Protocol	OLSR, OLSR Kmeans-SDE, OLSR Med+
Simulation time	1200s
Source and type of traffic	CBR/UDP
Network size	1000mx1000m
Size paquet	512 bytes
Speed nodes	10m/s
Mobility model	RWP
Number of nodes	10 20 30 40 50 100

8. RESULTS AND DESCUSSION
8.1. Packet delivery ratio
We notice in the Figure 1, that when the network becomes dense and large the number of lost packet increases, and this is because of different phenomenon (topology change, battery exhaustion ...); the retained structure and refined algorithm in the our protocol proposed OLSR Kmeans-SDE allows to reduces the number of packets lost compared to our previous version.

8.2. End to end delay
Today, the real-time applications are invading the field of human activity. For this to minimize the delivery time of the packets between the transmitter and the receiver becomes a necessity. The method of clustering adopted and refined led us to have a small gain of transfer time (EED). As we see in the Figure 2, the End to End Delay (EED) of all routing protocol increase proportionally with the number of nodes; and the OLSR Kmeans-SDE responded the best.

8.3. Overhead
The overhead of three protocols are according to the network density in Figure 3, The OLSR Kmeans-SDE generates more overhead followed by OLSR Kmeans and OLSR Med+. It is noted that the division of nodes into groups allow a significant reduction in the controls messages broadcasted when the network becomes denser; however our protocol proposed OLSR Kmeans-SDE have more information to exchange (mobility, density, energy). For this we notice the small overhead compared with other protocol proposed.
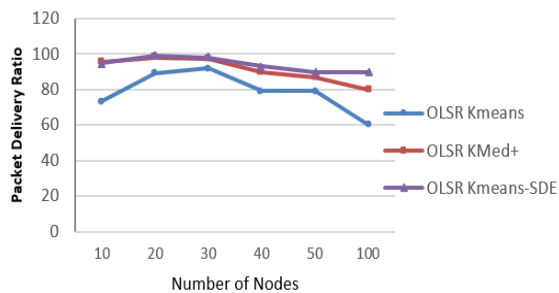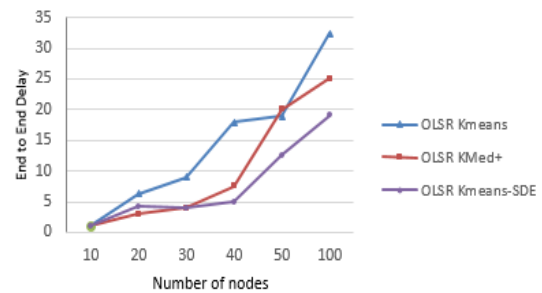

Figure 1. Packet delivery ratio
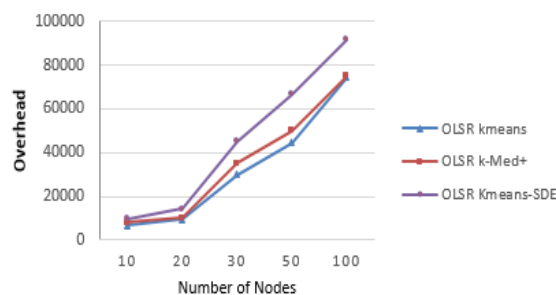

Figure 2. End to end delay

Figure 3. Control traffic overhead

9.   CONCLUSION

The clustering is one of the most important techniques used to organize the network into different groups, this can reduce the complexity in management of nodes and therefore simplify the processes routing information and increase the Qos in MANETs, we have used the strong method (K-means) to group the nodes into several clusters. We have confirmed that k-means improves the classification of nodes by demonstrating theoretically that the method reduces the intra-cluster inertia (therefore it increases the inter-cluster inertia) between two successive iterations. In this paper we have continued to refine our proposed protocol by adding more stability parameters. The routing protocol (OLSR Kmeans-SDE) brings some improvement to the level of EED and PDR, having a product that meets all the needs in terms of quality of service is not obvious. We win one type of performance and we lose the other. We started our study by implementing both clustering techniques (k-means, k-medoid), and we added metrics in the algorithms to have a robust protocol. That's what we achieved during our research. In future work, we will have a plan to minimize the overhead in OLSR Kmeans-SDE.

REFERENCES

[1]   D. T. Larose, "Discovering Knowledge in Data: An Introduction to Data Mining," *John Wiley & Sons, Inc.*, Hoboken, New Jersey. 2005

[2]   E. Forgy. W., "Cluster Analysis of Multivariate Data : Eficiency versus Interpretability Models," *Biometrics*, vol. 61, no. 3, pp. 768-769, 1965.

[3]   J. P. Benzecri, "L'analyse des données : la taxinomie," *Tome 1. Dunod*, Paris, 1973.

[4]   G. H. Ball, D. J. Hall, ISODATA, "An Iterative Method of Multivariate Analysis and Pattern Recognition," *Behavior Science*, vol. 153, 1967.

[5]   P. Jacquet, P. Muhletaler, A. Qayyum, A. Laoutiti, T. Clausen and L. Viennot, "Optimized Link State Routing Protocol," *in IEEE INMIC, (Pakistan),* Dec. 2001.

[6]   S. Dornbush, A. Joshi, "StreetSmart Traffic: Discovering and Disseminating Automobile Congestion Using VANET's," *Vehicular Technology ConferenceVTC2007 Spring. IEEE 65th*, pp 11-15, 2007.

[7]   E. Yoneki and J. Bacon, "Distributed Multicast Grouping for Publish/Subscribe over Mobile Ad Hoc Networks," *Wireless Communications and Networking Conference, 2005 IEEE*, vol. 4, pp. 2293-2299, 2005.

[8]   R. Khanna and H. Liu, "System Approach to Intrusion Detection Using Hidden Markov Model," *IWCMC '06 Proceedings of the 2006 international conference on Wireless communications and mobile computing*, pp. 349-354, 2006.

[9]   O.F. Hamad, M.Y. Kang, J. Jeon and J. Nam, "Neural Network's k-means Distance-Based Nodes-Clustering for Enhanced RDMAR Protocol in a MANET," *Signal Processing and Information Technology, IEEE 2008*, pp 192-197, 2008.

[10]   R. Badonnel, R. State, and O Festor, "Self-configurable fault monitoring in ad-hoc networks," *Ad Hoc Networks*, vol. 6, no. 3, pp 458-473, 2008.

[11]   A. Choukri, A. Habbani, M. ElKoutbi, "A hierarchical version of OLSR for MANET," *World Applied Sciences Journal,* vol. 21, no. 12, pp. 1739, 2013.

[12]   A. Choukri, A. Habbani, M. El.Koutbi, "Efficient heuristic based on clustering approach for OLSR," *Journal of Computer Networks and Communications*, 2013.

[13]   A. Sameer, Z. Zahriladha, and L. Herwansyah, "A new energy consumption technique for mobile Ad-Hoc networks," *International Journal of Electrical and Computer Engineering(IJECE),* vol. 9, no. 5, pp. 4147-4153, 2019.

[14]   A. Choukri, A. Habbani, M. ElKoutbi, "An Energy Efficient Clustering Algorithm for MANETs," *ICMCS,* 2014.

[15]   J. MacQueen, "Some methods for classification and analysis of multivariate observations," *In Proceedings of the Fifth Berkeley Symposium on Mathematical statistics and probability*, pp. 281-297, 1967.

[16]   R. Whitaker, S. Hurley, "Evolution of planning for wireless communication systems," *Proceedings of the 36th Annual Hawaii International Conference (HICSS'03)*, pp. 296, 2003.

[17]   Angione, G., Bellavista, P., Corradi, A. and Magistretti, E., "A k-hop Clustering Protocol for Dense Mobile Ad-Hoc Networks," *26th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW'06)*, pp. 10, 2006.

[18]   Daniel T. Larose,"Discovering Knowledge in Data: An Introduction to Data Mining," *John Wiley & Sons*, Inc., Hoboken, New Jersey, 2005.

[19]   R. Anant and M. Manas, "Mobility adaptive density connected clustering approach in vehicular ad hoc networks," *International Journal of Communication Networks and Information Security,"*. vol. 9, pp. 222-229, 2017.

[20]   Inn, E. and Winston, S., "Mobility-Based D-Hop Clustering Algorithm for Mobile Ad Hoc Networks," *IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 2359-2364, 2004.

[21]   M. Dyabi, M. Saadoune, A. Hajami, H. Allali, "OLSR Clustering Algorithm Based on Nodes Mobility," *International Review on Computer and Softaware, (IRECOS)*, vol. 10, no. 1, pp. 36-43, 2015.

[22]   Sanlin Xu, Kim L. Blackmore, Haley M. Jones, "An Analysis Framework for Mobility Metrics in Mobile Ad Hoc Networks," *Published in EURASIP J. Wireless Comm. and Networking,* 2007.

[23] Y. Hamzaoui, M. Amnai, A. Choukri, Y. Fakhri, "Novel Clustering Method Based on K-Medoids and Mobility Metric," *International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI)*, vol. 5, no. 1, pp. 29-33, 2017.

[24] Y. Hamzaoui, M. Amnai, A. Choukri, Y. Fakhri, "Enhancing OLSR Routing Protocol Using K-medoids Clustering in MANETs," *International Journal of Applied Engineering Research*, vol. 12, no. 2, pp. 200-206, 2017.

[25] A. Al-Maashri and M. Ould-Khaoua, "Performance Analysis of MANET Routing Protocols in the Presence of Self-Similar Traffic," *Proceedings 2006 31st IEEE Conference on Local Computer Networks*, Tampa, FL, pp. 801-807, 2006.

BIOGRAPHY OF AUTHORS

Ali Choukri is an Assistant professor at National School of Applied Sciences. He was awarded a Master of Computer Science and Telecom from the University of IbnTofail, Kénitra-Morocco in 2008. He got ENSET (higher Normal School of technical teaching) degree in 1992. He has a PhD from the School of Computer Science and Systems Analysis (ENSIAS). He is working within the MIS team of the Laboratory SIME, for studying ad hoc mobile intelligent communication systems, and wireless sensor networks. His research interests are in the areas of ubiquitous computing, Internet of Things, Delay/Disruption Tolerant Networks, Wireless Networks, QoS routing, Mathematical modeling and performance analysis of networks, Control and decision theory, Game theory, trust and reputation management, Distributed algorithms, Metaheurtistics and optimization, Genetic algorithms

Younes Hamzaoui obtained his DEUG degree on phsysics chemical in 2000 from Abdmaliksaadi University, Tanger city, morocco. Then, he received his degree on IEEA (Computers, Electronics, Electrical and Automation) in 2003. After in 2007, he obtained his DESA degree on information system and telecommunication. The authers is a professor of computer science in high school. He is a PhD student of management of Qos in routing protocol, also, he doing his research in the Laboratory in Computer Science and Telecomunications (LaRIT).

Mohamed Amnai received his bachelor's degree in 2000, in IEEA (Computers, Electronics, Electrical and Automation) from Molay Ismail's University, Errachidia city. Then, he obtained his master's degree in 2007, from Ibn Tofail University, Kenitra city. In 2011, he received his PhD on Telecommunication and computer science, from Ibn Tofail University in Kenitra city, Morocco. Since March 2014 he is a Professor, he has been an Assistant at National School of Applied Sciences Khouribga, Settat University, Morocco. The author is also an associate member of Research Laboratory in Computer Science and Telecommunications (LARIT), Team Networks and Telecommunications Faculty of Science, Kenitra, Morocco. Also, the author is an associate member of laboratory IPOSI National School of Applied Sciences Khouribga, Sultan Moulay Slimane University, Morocco.

Youssef Fakhri received his Bachelor's Degree (B.S) in Electronic Physics in 2001 and his Master's Degree (DESA) in Computer and Telecommunication from the Faculty of Sciences, University Mohammed V, Rabat, Morocco, in 2003 where he developed his Master's Project at the ICI Company, Morocco. He received a PhD in 2007 from the University Mohammed V - Agdal, Rabat, Morocco in collaboration with the Polytechnic University of Catalonia (UPC), Spain. He joined the Faculty of Sciences of Kénitra, Department of Computer Science and Mathematics, IbnTofail University, Morocco, as an Associate Professor in March 2009, he is the Laboratory head at LaRIT (Laboratory for Research in Computing and Telecommunications) in the Faculty of Kénitra, and Member of Pole of Competences STIC Morocco.