Clustering of Wireless Ad Hoc Networks: An Efficient Probabilistic Technique

Md. Shafakhatullah Khan[#1], Mourad Mohamed Henchiri[*2]

*#*Information Systems Department, College of Economics, Management and Information Systems, University of Nizwa, P.O.Box:33, PC:616, Birkat al Mouz, Nizwa, Sultanate of Oman.*

***Abstract*— Now a day's the emergence of the wireless devices and the usage of wireless ad-hoc Networks, is increased. Almost all the wireless device have option to connect to wireless ad hoc Network to access data or information (resource sharing) from other devices, to share resources from the wireless ad-hoc Network they need a connectivity to the network. The smart/advanced devices contains a Wi-Fi connectivity and Wi-Fi hotspot tools built-in, which get connection from network services as well as provides network service to other devices by converting the device as router(using Wi-Fi Hotspot). Major problems identified in wireless ad hoc networks are organizing, scalability and routing, which affects the commercial success of MANET. Clustering of mobile nodes among separate domains has been proposed in many papers as the successful approach to solve those problems. In this work we introduce sophisticated and dynamic technique of organizing nodes into clusters in wireless ad hoc network. This technique is reliable and more efficient than the well-known algorithm WCA (Weighted Clustering Algorithm) [1]. Here we have developed new algorithm for the implementation of Probabilistic Weighted Clustering Algorithm (PWCA) with the help of Weighted Clustering Algorithm (WCA).**

***Keywords*— Wi-Fi, Wireless Ad hoc Network, MANET, Clusters, WCA.**

I. INTRODUCTION

Wireless communication and the lack of centralized administration pose numerous challenges in mobile wireless ad-hoc networks (MANETs) [3]. Node mobility results in frequent failure and activation of links, causing a routing algorithm reaction to topology changes and hence increasing network control traffic [2]. Ensuring effective routing and Quality of Service (QoS) support while considering the relevant bandwidth and power constraints remains a great challenge. Given that MANETs may comprise a large number of MNs, a hierarchical structure will scale better [4].

Hence, one promising approach to address routing problems in MANET environments is to build hierarchies among the nodes, such that the network topology can be abstracted. This process is commonly referred to as clustering and the substructures that are collapsed in higher levels are called clusters [5]. The concept of clustering in MANETs is not new; many algorithms that consider different metrics and focus on diverse objectives have been proposed [5]. However, most existing algorithms fail to guarantee stable cluster formations. More importantly, they are based on periodic broadcasting of control messages resulting in increased consumption of network traffic and mobile hosts (MH) energy.

In clustering procedure, a representative of each subdomain (cluster) is 'elected' as a cluster head (CH) and a node which serves as intermediate for inter-cluster communication is called gateway. Remaining members are called ordinary nodes. The boundaries of a cluster are defined by the transmission area of its CH [7].

Clustering in mobile ad hoc networks can be defined as the virtual partitioning of dynamic nodes into various groups. Groups of the nodes are made with respect to their nearness to other nodes. Two nodes are said to be neighbors of each other when both of them lie within their transmission range and set up a bidirectional link between them [6]. Clustering is an important approach to solving capacity and scalability problems in mobile ad hoc networks where no physical infrastructure is available. The connected dominating set (CDS) is a special cluster structure where the CH's form a connected network without using gateways. Certain nodes, known as CH's, are responsible for the formation of clusters each consisting of a number of nodes (analogous to cells in a cellular network) and maintenance of the topology of the network. The set of cluster heads is known as a dominant set. A CH does the resource allocation to all the nodes belonging to its cluster. Due to the dynamic nature of the mobile nodes, their association and dissociation to and from clusters disturb the stability of the network and thus the configuration of CH's is unavoidable. This is an important issue since frequent CH changes adversely affect the performance of other protocols such as scheduling, routing and resource allocation that rely on it. The choice of the CH's is here based on the weight associated to each node: the smaller the weight of a node, the better that node is for the role of CH [8].

II. LITERATURE REVIEW

A. Clustering

In clustering procedure, a representative of each subdomain (cluster) is 'elected' as a cluster head (CH) and a node which serves as intermediate for inter-cluster communication is called gateway. Remaining members are called ordinary nodes. The boundaries of a cluster are defined by the transmission area of its CH. With an underlying cluster structure, non-ordinary nodes play the role of dominant forwarding nodes, as shown in Figure 1.

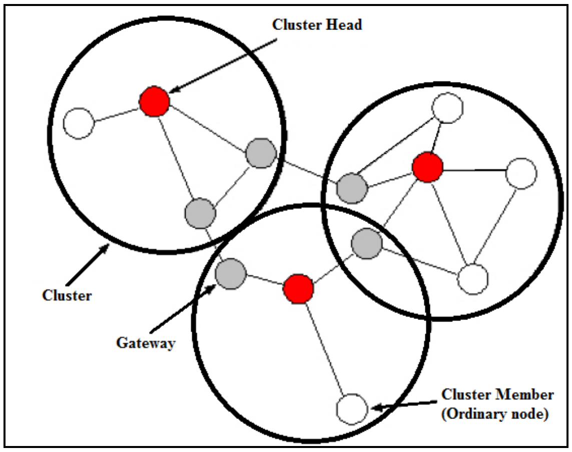

Figure 1: Cluster Heads, Gateways, and Ordinary noders in Wireless Ad Hoc Network.

Cluster architectures do not necessarily include a CH in every cluster. CHs hold routing and topology information, relaxing ordinary MHs from such requirement; however, they represent network bottleneck points. In clusters without CHs, every MH has to store and exchange more topology information, yet, that eliminates the bottleneck of CHs. Yi et al. identified two approaches for cluster formation, *active* clustering and *passive* clustering [9]. In *active* clustering, MHs cooperate to elect CH's by periodically exchanging information, regardless of data transmission. On the other hand, *passive* clustering suspends clustering procedure until data traffic commences [10]. It exploits on-going traffic to propagate "cluster-related information" (e.g., the state of a node in a cluster, the IP address of the node) and collects neighbor information through promiscuous packet receptions. Passive clustering eliminates major control overhead of *active* clustering, still, it implies larger setup latency which might be important for time critical applications; this latency is experienced whenever data traffic exchange commences. On the other hand, in active clustering scheme, the MANET is flooded by control messages, even while data traffic is not exchanged thereby consuming valuable bandwidth and battery power resources.

B. Lowest ID Cluster Algorithm

This is an identifier based clustering algorithm in which each node is assigned a distinct ID and the cluster formation is done based on these identifiers. In the lowest ID cluster algorithm, the node with a minimum ID is chosen as the CH [11, 12]. A node will always broadcast the list of nodes within its range (including itself). The CH is the node that will only hear nodes with ID higher than itself. Thus, the neighbor nodes of the CH will be having ID's higher than that of the CH. A node that lies within the transmission range of two or more CH is called the gateway node and they are used for routing between different clusters in a network.

The drawback of this scheme is that, the lowest ID scheme considers only the lowest node ID that is arbitrarily assigned numbers. It is not considering any other qualifications of a node for selecting the node as a CH. The ID of the node does not change with the time and for a long period; the node may have to be the CH. Therefore, there is a chance for certain nodes to have power drainage due to serving as CH's for longer period of time.

C. Highest Degree Algorithm

The highest degree algorithm is also known as connectivity-based algorithm. Here, the degree of a node is calculated based on its distance from the other node. There exists a link between those nodes, if the Euclidean distance between the two nodes is less than the range. In the highest degree algorithm, a node that has maximum degree is chosen as a cluster head [11, 12]. The neighbors of a cluster head become members of that cluster. The algorithm does not limit the number of nodes in a cluster. Therefore, when many nodes are there in the cluster, the throughput drops and the system performance is reduced.

D. Distributed Cluster Algorithm

The Distributed Clustering Algorithm is a modified version of the Lowest Identifier algorithm. For each cluster, it chooses a node with locally lowest ID among all the neighbor nodes as a CH. Every node can determine its cluster and transmits only one message during the algorithm [11]. Since it uses node ID for the selection of CH's, it inherits the drawbacks of the Lowest Identifier heuristic.

E. Weighted Cluster Algorithm

The weighted cluster algorithm elects the cluster head based on the factors like node mobility, number of nodes a CH can handle, transmission power etc…[11, 13, 14, 15]. The cluster head must not be over-loaded and therefore a pre-defined threshold value is used which indicates the number of nodes each CH can support. The weighted clustering algorithm selects a CH according to the weight value of each node. The weight associated to a node *vi* is defined as:

$$W_{vi} = w_1\Delta_{vi} + w_2 D_{vi} + w_3 M_{vi} + w_4 P_{vi} \quad (1)$$

The node with the minimum weight is selected as a CH. w_1, w_2, w_3, and w_4 are weighting factors. The weighting factors are chosen such that $w_1+w_2+w_3+w_4=1$. M_{vi} is the measure of mobility. It depends on the average speed of every node during a specified time T. Δ_{vi} is the degree difference. D_{vi} is defined as the sum of distances from a given node to all its neighbors. This factor is related to the energy consumption since more power is needed for the long distance communication. The parameter P_{vi} is the cumulative time for which a node is retained as the CH. This factor is related to measure the power consumption. The cluster head election continues until all the nodes in the network is covered. No two CH's can be immediate neighbors.

F. Distributed Weighted Clustering Algorithm

This algorithm is an enhanced version of WCA to achieve distributed clustering set up and to extend lifetime span of the system. This algorithm differs from WCA in which it localizes configuration and reconfiguration of clusters and poses restriction on the power requirement on the CH's. This algorithm provides better performance than WCA in terms of the number of reaffiliations, end-to-end throughput, overheads during the initial clustering set up phase, and the minimum lifespan of nodes [17].

III. Methodology

A. Problem Identification

In the above mentioned all approaches they focused on find the CH using different techniques and algorithms. The process of finding the CH's when the node are moving is a time taking process, this consumes high power, when the nodes are moving automatically CH's also moves, when CH's moves few nodes will be with connected CH's within a range few will get disconnected as usual, when CH's moves we need to check the condition that two CH's should not be immediate neighbors. So there will be a frequent change in the CH's. In mobility assigning CH's is a challenging task, this task should be done quickly to avoid inconsistency in the connectivity of Ad hoc Network, to avoid this issues so many algorithms has been proposed, some proposed how to reduce CH's, few others proposed Weight factor, battery energy, Lowest ID's, Highest Degree and many more. The process of creating clusters will be same, the issue is how to find the CH's as quickly as possible so that we can avoid inconsistency and save energy. Therefore the main objective of this work is to increase the efficiency of Wireless Ad hoc Network by finding the CH's when the node are moving in a very less time.

A cluster head is selected based on two aspects, the pheromone value associated with each node and its visibility. Visibility refers to the number of nodes that will be covered if the node is added into the cluster head set [18]. Visibility keeps changing as topology changes.

In this work we are working with the Visibility. The visibility value associated with a node is updated for each iteration of the algorithm. For each iteration, a node is selected as the cluster head and the next cluster head is selected based on the visibility of its neighbor nodes. This process continues until all the nodes in the network are covered. A node is said to be covered if it is a cluster head or falls in the range of an already selected cluster head.

Each time a node is selected as a cluster head, its Visibility value is updated. Thus, possibility of a node to be selected as cluster head depends on the weight, probability value derived which is more than 0.5 * standard deviation of distance of all the node's values which are in the range/cluster and visibility which changes as the algorithm proceeds through the various iterations. The probability of each node to be selected as cluster head is calculated based on Visibility.

$$Vis(iter) = \frac{Wt_{vi} * \alpha + [vis_{vi}(iter)] * \beta}{\sum_{vi=0}^{n} Wt_{vi} * \alpha + [vis_{vi}(iter)] * \beta} \quad \text{............(2)}$$

The equation (2) gives the probability of each node to be selected as a cluster head. It is represented by *Vis*. Wt_{vi} is the weight associated with each node *vi*. Wt_{vi} is dependent on the degree associated with each node.

$vis_{vi}(iter)$ is the visibility of each node for an iteration. α Controls the relative importance of visibility measure and β controls the Visibility value.

B. Algorithm

```
Initialize each node
        Weight = 0;
        isHead = false;
        visibility= 0;
        convector = empty;
Find neighbor
        If Euclidian distance<Range
        Edge exists;
        Else
        No edge exists;
Increment weight
Set α, β value
        Iterate i to n times
        Select i^th node as Cluster head
        If(distance(Cluster head)>0.5*standard deviation(distance[All nodes within the range]))
        Select as Cluster heads until all nodes are covered;
        Update visibility of selected cluster heads
        Count++;
Iterate i to count times
        If(cluster head is immediate neighbor)
        Prune cluster head with less weight;
Find final set of cluster heads with maximum probability(with weight).
```

Figure 2: Alogrithm to create cluster heads

The algorithm shows that the Vis_{vi}**,** the probability of Visibility to choose a node is proportional to the degree of each node and the pheromone concentration factor.

Note: The modification is done in the selection of cluster head based on maximum probability. Instead of that we have used standard deviation; it reduces cluster head finding time, Computation time, Consume very less energy, and it is fast.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 12, December - 2013

C. Simulation

The proposed technique is simulated in java with nodes being taken as random values at different time instances. Number of probabilistic objects and probability of change of the node locations are vary and accuracy of the neighbor cluster head selection is measured. The tool we have used to develop this application is NetBeans IDE, The base IDE includes an advanced multi-language editor, debugger and profiler integration, file versioning control, and unique developer collaboration features.

IV. RESULT ANALYSIS

A. Simulation Results

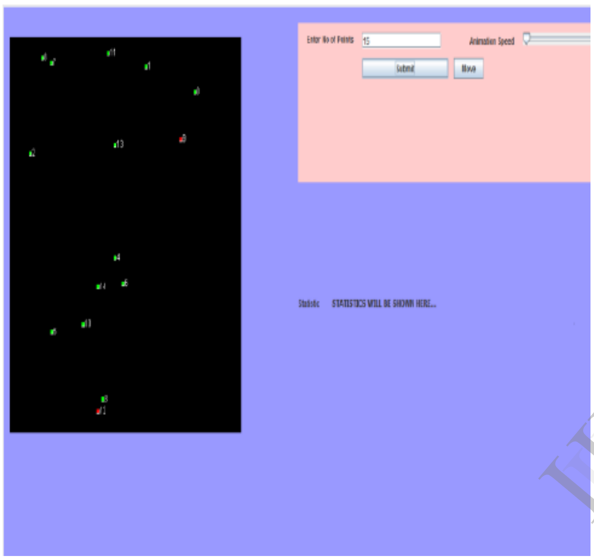

Figure 3: When nodes are initialized

Figure 4: When nodes are moving Selection of Cluster Heads

B. Graph Analysis

TABLE I
FUNDAMENTAL DIFFERENCE BETWEEN PRESENT AND PROPOSED SYSTEMS CLUSTER HEAD SELECTION TIME

S.No	No of Nodes	CHST of Present	CHST of Proposed
1	25	22	7
2	50	14	9
3	75	19	8
4	100	14	8
5	125	47	9
6	150	14	7
7	175	14	12
8	200	14	8
9	225	15	12
10	250	14	9
11	275	15	8
12	300	15	8
13	325	15	8
14	350	15	8
15	375	15	9
16	400	15	9
17	425	16	9
18	450	15	7
19	475	17	10
20	500	16	9

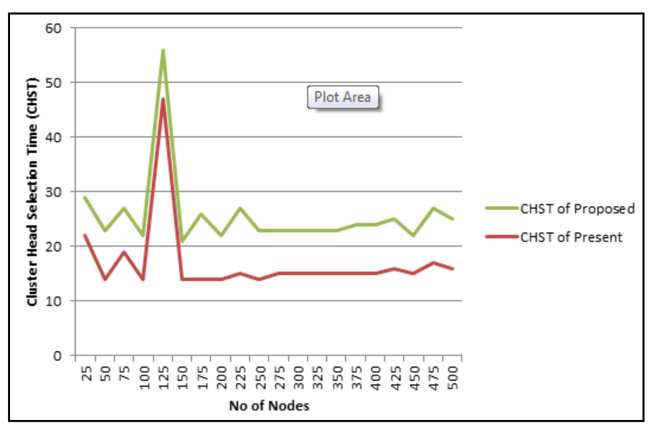

Figure 5: Graph analysis of No of Nodes Vs Cluster Head Selection Time between Present and Proposed Systems

In the above graph Fig 5 we have taken nodes up to 500. To make the nodes uncertain we moved the node objects; we are generating 10 instances for each node at 10 different positions when it is moved, as the probability of their occurrence at different positions from time to time. The node objects will be generated randomly.

V. CONCLUSION AND FUTURE WORK

In this paper we have proposed a Weighted Clustering Algorithm by making some modification and improvements on existing algorithms discussed in [14]. As demonstrated, our algorithm selects the cluster head when they are moving within a little time and saves energy.

In future, some new parameters can be added to weight computation of nodes, consideration of transmission power, this technique is tested on simulation environment; it can be implemented in a real ad hoc system to evaluate its performance in real world scenarios.

REFERENCES

[1] Christian Bettstetter. The Cluster Density of a Distributed Clustering Algorithm in Ad Hoc Networks. IEEE International Conference Communications; Page(s): 4336 - 4340 Vol.7; 2004.

[2] X. Hong, K. Xu, M. Gerla, **"**Scalable Routing Protocols for Mobile Ad Hoc Networks", IEEE Network, 16(4), pp. 11-21,July-Aug 2002.

[3] C. Perkins, "Ad Hoc Networking", Addison-Wesley, 2001.

[4] C. R. Li, M. Gerla, "Adaptive Clustering for Mobile Wireless Networks", IEEE Journal of Selected Areas in Communications, 15(7), pp. 1265-1275, 1997.

[5] J. Yu, P. Chong, "A Survey of Clustering Schemes for Mobile Ad Hoc Networks", IEEE Communications Surveys, 7(1), pp. 32-48, March 2005.

[6] Suchismita Chinara, Santanu Kumar Rath. A Survey on One-Hop Clustering Algorithms in Mobile Ad Hoc Networks. Journal Network System Management; 17:183 207; 2009.

[7] Damianos Gavalas, Grammati Pantziou, Charalampos Konstantopoulos, Basilis Mamalis, "Clustering of Mobile Ad Hoc Networks: An Adaptive Broadcast Period Approach". IEEE Communications, ICC-06, Vol. 9, pp. 4034 – 4039; 2006.

[8] Bhaskar Nandi, Subhabrata Barman, and Soumen Paul, "Genetic Algorithm Based Optimization of Clustering in Ad-Hoc Networks". IJCSIS, Vol. 7 No.1, 2010.

[9] Y. Yi, M. Gerla, T. Kwon, "Efficient Flooding in Ad hoc Networks: a Comparative Performance Study", Proceedings of the IEEE International Conference on Communications (ICC'2003), 2003.

[10] Y. Yi, M. Gerla, T. Kwon, "Passive Clustering (PC) in Ad Hoc Networks", Internet Draft, draft-ietf-yi-manet-pac-00.txt, 2001.

[11] M. Gerla and J. T.-C. Tsai, "Multicluster, mobile, multimedia radio network," Wirel. Netw., vol. 1, no. 3, pp. 255–265, 1995.

[12] R. Agarwal, M. Matwani, "Survey of Clustering Algorithms for MANET", International Journal on Computer Science and Engineering, Vol. 1(2), 2009, pp. 98-104.

[13] G. Chen, F. Nocetti, J. Gonzalez, and I. Stojmenovic, "Connectivitybased k-hop clustering in wireless networks," Proc. of the 35th Annual Hawaii International Conference on System Sciences (HICSS‟02)-Volume 7. Washington, DC, USA: IEEE Computer Society, 2002, p. 188.3.

[14] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: A weighted clustering algorithm for mobile ad hoc networks," Cluster Computing, vol. 5, no. 2, pp. 193–204, 2002.

[15] M. Chatterjee, S. K. Das and D.Tunjnt, "An On-Demand Weighted Clustering Algorithm (WCA) for Ad hoc Networks", Proc. of IEEE Globecom‟00, pp. 1697-1701.

[16] S. Basagni, "Distributed Clustering for Ad Hoc Networks" In Proceedings: International Symposium on Parallel Architectures, Algorithms, and Networks, 1999, pp. 310-315.

[17] W. Choi and M. Woo, "A Distributed Weighted Clustering Algorithm for Mobile Ad Hoc Networks," Proceedings of Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services, 2006. doi:10.1109/ AICT-ICIW.2006.11.

[18] C. K. Ho, Y .P .Singh, H. T. Ewe, "An Ant Colony Optimisation approach to Building Clusters in Ad hoc Networks", Proc. MMU International Symposium on Information and Communication Technologies, 7th-8thOct2004, Malaysia, pp.TS1B1.

A. Md.Shafakhatullah Khan has Masters in Technology, Masters in Computer Applications. Currently he is working as Lecturer in the Department of Information Systems, College of Economics, Management and Information Systms, University of Nizwa, Sultanate of Oman. He has 7+ years of work experience in both Industry and Academics. Md. Shafakhatullah Khan, is member in IAENG (International Association of Engineers).

B. Mourad Mohamed Henchiri has Masters in IS and Network's Management, is working as Instructor in the Department of Information Systems, College of Economics, Management and Information Systems, University of Nizwa, Sultanate of Oman. He has 7+ years of work experience in both Industry and Academics.